

VERITROOPER Scout — All Q&A Pairs



Every question in this audit (889 total): the question, the ground truth, what the baseline (vanilla rag) answered (with its verdict), and what VERITROOPER answered. The heading verdict is the baseline (vanilla rag)'s — so misses on the deployed system show in red.

1. PASS · Exception · `SEC\_10K\_multi\_0651\_00`

- **Question:** What is excluded from principal investments?
- **Ground truth:** those that are consolidated on the Firm's balance sheets
- **✓ Baseline (vanilla RAG) (PASS):** The provided evidence does not specify what is excluded from principal investments. It only states that the Firm's approximately 18.6 million Visa B-2 shares are included in its principal investment portfolio.
- **✓ VERITROOPER (PASS):** Principal investments exclude investments consolidated on the Firm's balance sheets.

2. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_04`

- **Question:** What is the total revenue generated by NVIDIA's data center business in 2025?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The provided evidence does not state NVIDIA's total Data Center revenue for 2025, so it cannot be determined from this information.
- **✓ VERITROOPER (PASS):** The evidence does not provide NVIDIA's data center revenue for 2025.

3. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1409\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the percentage of risk participations for other letters of credit compared to total commitments?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** For 2025, risk participations for other letters of credit were **\$514 million**, compared with **\$4.529 billion** of total commitments:

$$\frac{514}{4.529} \times 100 \approx 11.35\%$$

Thus, risk participations were approximately **11.4% of total commitments**.

- **✓ VERITROOPER (PASS):** The evidence does not provide total commitments, so the percentage of risk participations for other letters of credit cannot be calculated.

4. PASS · Precision · `SEC\_10K\_multi\_1548\_01`

- **Question:** What was the impairment charge recorded by the Company during the three months ended December 31, 2025?
- **Ground truth:** 960 million
- **✓ Baseline (vanilla RAG) (PASS):** The Company recorded an impairment charge of **\$960 million** related to the BodyArmor trademark.
- **✓ VERITROOPER (PASS):** The Company recorded an impairment charge of **\$960 million**.

5. PASS · Cause & Effect · `SEC\_10K\_multi\_0137\_01`

- **Question:** What could happen if JPMorganChase is unable to attract or retain qualified employees?
- **Ground truth:** materially and adversely affected performance and competitive position
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase's performance and competitive position could be materially and adversely affected. It could also face reduced operational efficiency or increased costs, particularly if restrictive immigration or travel policies limit recruitment and retention.
- ✓ **VERITROOPER (PASS):** JPMorganChase's performance and competitive position could be materially and adversely affected.

6. PASS · Precision · `SEC\_10K\_multi\_0494\_00`

- **Question:** What was the total average assets as of December 31, 2024?
- **Ground truth:** 4,125,167
- ✓ **Baseline (vanilla RAG) (PASS):** Total average assets for 2024 were **\$3,822,224 million** (approximately **\$3.82 trillion**).
- ✓ **VERITROOPER (PASS):** Total average assets were \$4,125,167 million as of December 31, 2024.

7. PASS · Precision · `SEC\_10K\_multi\_0254\_01`

- **Question:** What is the year for which the results are compared to the previous year?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** Fiscal year 2026, compared with fiscal year 2025.
- ✓ **VERITROOPER (PASS):** The year is 2025, compared with 2024.

8. FAIL · Precision · `SEC\_10K\_multi\_1345\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the net change for DVA on fair value option elected liabilities?
- **Ground truth:** 1,025
- ✗ **Baseline (vanilla RAG) (FAIL):** The net change in DVA on fair value option-elected liabilities was **\$(773) million after tax in 2025 (\$1.025) billion pre-tax**.
- ✓ **VERITROOPER (PASS):** The DVA net change was \$(1,025), with a \$252 tax effect, resulting in \$(773) net of tax.

9. PASS · Cause & Effect · `SEC\_10K\_multi\_0131\_00`

- **Question:** What happens as a consequence of higher interest rates?
- **Ground truth:** JPMorganChase could incur losses as well as increased operational costs related to servicing a higher volume of delinquent loans.
- ✓ **Baseline (vanilla RAG) (PASS):** Higher interest rates can lead to:
 - Fewer commercial and residential real-estate loan originations
 - Losses on underwriting exposures and more client downgrades
 - Higher client financing costs, increasing credit-loss allowances and net charge-offs
 - Deposit outflows as customers shift to higher-yielding products
 - Losses on available-for-sale securities
 - Reduced financial-market liquidity
 - Higher funding costs
 - Weaker payment performance and more delinquencies on variable-rate loans, causing losses and increased servicing costs

Collectively, these effects can adversely affect JPMorganChase's earnings, liquidity, and capital, especially if high rates persist.

- ✓ **VERITROOPER (PASS):** Higher interest rates can reduce loan originations, increase client financing and JPMorganChase funding costs, cause deposit losses and losses on available-for-sale securities, reduce market liquidity, and increase credit losses, delinquencies, and servicing costs. They may also lower client asset values and related fee revenue, adversely affecting JPMorganChase's earnings, liquidity, and capital.

10. PASS · Cause & Effect · `SEC\_10K\_multi\_1462\_00`

- **Question:** What is the consequence of the First Republic acquisition resulting in a bargain purchase gain?
- **Ground truth:** A preliminary estimated bargain purchase gain of \$2.7 billion, which was later adjusted to \$2.9 billion.
- ✓ **Baseline (vanilla RAG) (PASS):** Because the fair value of First Republic's net assets acquired exceeded the purchase price, JPMorganChase recognized an after-tax bargain purchase gain of approximately **\$2.9 billion**.
- ✓ **VERITROOPER (PASS):** The acquisition produced a final bargain purchase gain of \$2.9 billion, reflecting that the estimated fair value of net assets acquired exceeded the purchase price.

11. PASS · Precision · `SEC\_10K\_multi\_0588\_00`

- **Question:** What percentage of total wholesale credit exposure was criticized as of December 31, 2025?
- **Ground truth:** 3.4
- ✓ **Baseline (vanilla RAG) (PASS):** Approximately **3.4%** of total wholesale credit exposure was criticized as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** Approximately 3.4% of total wholesale credit exposure was criticized as of December 31, 2025.

12. PASS · Cause & Effect · `SEC\_10K\_multi\_0424\_00`

- **Question:** What is the consequence of the absence of the \$7.9 billion net gain related to Visa shares?
- **Ground truth:** Noninterest revenue was lower.
- ✓ **Baseline (vanilla RAG) (PASS):** The absence of the \$7.9 billion Visa-share gain recorded in 2024 significantly reduced 2025 noninterest revenue, particularly in Corporate, contributing to its decline from \$7.6 billion to \$911 million.
- ✓ **VERITROOPER (PASS):** The absence of the \$7.9 billion Visa shares net gain drove noninterest revenue down to \$911 million in 2025 from \$7.6 billion in 2024.

13. PASS · Cross-Reference · `SEC\_10K\_multi\_0715\_00`

- **Question:** What is the basis on which the guidance may be applied, and what is the planned date of adoption?
- **Ground truth:** The guidance may be applied on a prospective or modified retrospective basis, and the Firm's planned date of adoption is not specified.
- ✓ **Baseline (vanilla RAG) (PASS):** The guidance may be applied on a **prospective, modified, or retrospective transition basis**. The Firm is still evaluating its planned adoption date; the required effective date is **January 1, 2028**, with early adoption permitted.
- ✓ **VERITROOPER (PASS):** The guidance may be applied on a prospective or modified retrospective basis. The Firm is still evaluating its planned adoption date.

14. PASS · Exception · `SEC\_10K\_multi\_1434\_00`

- **Question:** What is excluded from the revenue figures?
- **Ground truth:** ratios
- ✓ **Baseline (vanilla RAG) (PASS):** Revenue figures exclude taxes collected from customers and remitted to governmental authorities.
- ✓ **VERITROOPER (PASS):** The revenue figures exclude the provision for credit losses.

15. FAIL · Cross-Reference · `SEC\_10K\_multi\_1568\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total amount of unused backup lines of credit for general corporate purposes as of December 31, 2025, and how does it relate to the Company's capital allocation priorities?
- **Ground truth:** \$6.2 billion
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, the Company had approximately **\$6.2 billion** (\$6.15 billion) in unused backup lines of credit for general corporate purposes.

These lines support liquidity and financial flexibility alongside the Company's capital allocation priorities:

1. Investing in business operations
 2. Growing dividend payments
 3. Expanding its beverage portfolio through acquisitions
 4. Repurchasing shares with excess cash over time
- ✓ **VERITROOPER (PASS):** As of December 31, 2025, the Company had \$6.2 billion in unused backup lines of credit. These support liquidity alongside other funding sources, while capital allocation priorities include business investment, dividend growth, acquisitions, and share repurchases.

16. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0516\_01`

BAD QUESTION

- **Question:** What is the total amount of deposits that are covered by deposit insurance protection at JPMorgan Chase as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, deposits covered by insurance protection totaled approximately **\$1,000.7 billion**:

\$2,559.3 billion total deposits – \$1,558.6 billion uninsured deposits = \$1,000.7 billion insured deposits.

- ✗ **VERITROOPER (FAIL):** Deposits covered by insurance protection totaled approximately **\$1,000.7 billion** as of December 31, 2025 (\$2,559.3 billion total deposits less \$1,558.6 billion uninsured).

17. PASS · Cause & Effect · `SEC\_10K\_multi\_1527\_01`

- **Question:** What is the consequence of the Firm's capital being removed from investment products?
- **Ground truth:** The intent is to ensure the fund is of sufficient size to represent a viable offering to clients, enabling pricing of its shares, and allowing the manager to develop a track record.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm removes its seed capital after the product is large enough to be viable, can price its shares, and has established a track record.
- ✓ **VERITROOPER (PASS):** The evidence does not state any consequence of removing the Firm's capital from investment products.

18. PASS · Cause & Effect · `SEC\_10K\_multi\_0293\_00`

- **Question:** What happens as a result of higher revolving balances in Card Services?
- **Ground truth:** Net interest income increases.
- ✓ **Baseline (vanilla RAG) (PASS):** Higher revolving balances in Card Services drove an increase in net interest income (NII).
- ✓ **VERITROOPER (PASS):** Higher revolving balances in Card Services predominantly drove higher net interest income.

19. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0714\_02`

BAD QUESTION

- **Question:** What is the name of the specific accounting standard that the new guidance is intended to clarify?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The guidance is intended to clarify **Revenue from Contracts with Customers (ASC 606)**, specifically the accounting for share-based noncash consideration from customers.
- ✗ **VERITROOPER (FAIL):** The standard is **Revenue from Contracts with Customers**.

20. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1537\_01`

- **Question:** What is the total cost of a widespread product recall?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not quantify a total cost. It states that a widespread product recall could cause significant losses from recall expenses, destroyed inventory, lost sales, liability claims, and negative publicity.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the total cost of a widespread product recall.

21. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0159\_01`

- **Question:** What is the average annual salary of JPMorganChase's executive officers?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The average annual salary cannot be determined from the provided evidence. Executive compensation details are incorporated by reference into JPMorganChase's 2026 Definitive Proxy Statement, but no salary amounts are included here.
- ✓ **VERITROOPER (PASS):** The evidence does not provide salary information for JPMorganChase's executive officers, so the average annual salary cannot be determined.

22. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0728\_04`

- **Question:** What is the typical recovery rate for a credit derivative?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a typical numerical recovery rate. It states that recovery rates are implied by CDS spreads and adjusted for differences between derivative creditors and senior unsecured creditors.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a typical recovery rate for a credit derivative.

23. FAIL · Cross-Reference · `SEC\_10K\_multi\_0064\_00`

BAD QUESTION

- **Question:** What is the total amount of cash and cash equivalents at the beginning of the period for the year ended January 25, 2026, and the year ended January 26, 2025?
- **Ground truth:** 10,605 and 8,589
- **✗ Baseline (vanilla RAG) (FAIL):** Cash and cash equivalents at the beginning of each period were:
 - Year ended January 25, 2026: **\$8.589 billion**
 - Year ended January 26, 2025: **\$7.280 billion**

Total: \$15.869 billion.

- **✗ VERITROOPER (FAIL):** Cash and cash equivalents at the beginning of the period were \$8,589 million for 2026 and \$7,280 million for 2025, totaling \$15,869 million.

24. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0678\_02`

- **Question:** What is the date by which the Firm's earnings-at-risk scenarios are next evaluated?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** No specific next evaluation date is provided. The Firm states that its earnings-at-risk scenarios are evaluated periodically.
- **✓ VERITROOPER (PASS):** The evidence does not specify a date; it only states that the scenarios are evaluated periodically.

25. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1585\_05`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average return on investment for the Company's investments in limited partnerships?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** For the alternative-energy limited partnerships:
 - **2025 ROI:** $\$241\text{M} \div \$306\text{M} = 78.8\%$
 - **2024 ROI:** $\$323\text{M} \div \$226\text{M} = 142.9\%$

The **average annual ROI was approximately 110.8%.**

On a combined basis, total benefits of \$564M divided by \$532M invested equals **106.0%.**

- **✓ VERITROOPER (PASS):** The evidence does not provide enough information to calculate the average return on investment for the Company's limited partnership investments.

26. PASS · Cross-Reference · `SEC\_10K\_multi\_0130\_01`

- **Question:** What are the potential risks to JPMorganChase's businesses due to political and geopolitical developments?
- **Ground truth:** JPMorganChase's businesses could suffer losses on its market-making positions or in its investment portfolio, reduce its liquidity and capital levels, or hamper its ability to deliver products and services to its clients and customers.
- **✓ Baseline (vanilla RAG) (PASS):** Political and geopolitical developments could create economic uncertainty and market volatility that may:
 - Weaken investor and consumer confidence, disrupt labor markets, trade, supply chains, and access to credit, or cause inflation, recession, currency fluctuations, and widening credit spreads.
 - Trigger tariffs, sanctions, retaliatory measures, military actions, government shutdowns, immigration changes, or shifts in fiscal, monetary, foreign, and trade policies.
 - Heighten geopolitical tensions-particularly involving Russia, the Middle East, and China-and increase targeted cyberattack risks.
 - Harm clients' and counterparties' financial condition or credit ratings and force JPMorganChase to forgo business opportunities.

- Cause investment and market-making losses, higher credit-loss allowances or charge-offs, and reduced liquidity and capital.
- Hamper product and service delivery, weaken operating results, financial condition, or credit ratings, and expose the firm to prolonged litigation.
- ✓ **VERITROOPER (PASS):** Political and geopolitical developments could create economic uncertainty, market volatility, trade and supply-chain disruptions, cyberattacks, sanctions, government shutdowns, inflation or recession, credit contraction and currency fluctuations. These effects could cause JPMorganChase to suffer investment or market-making losses, reduce liquidity and capital, increase credit losses, disrupt client services, weaken operations, financial condition or credit ratings, forgo opportunities, or face prolonged litigation.

27. PASS · Conditional · `SEC\_10K\_multi\_1594\_00`

- **Question:** If the Company determines it is no longer probable that the minimum performance threshold will be achieved, what action is taken regarding previously recognized compensation expense?
- **Ground truth:** All previously recognized compensation expense is reversed.
- ✓ **Baseline (vanilla RAG) (PASS):** The Company reverses all previously recognized compensation expense in the reporting period when that determination is made.
- ✓ **VERITROOPER (PASS):** The Company reverses all previously recognized compensation expense in the reporting period when that determination is made.

28. PASS · Cross-Reference · `SEC\_10K\_multi\_0153\_00`

- **Question:** What percentage of net issued shares were received by Ashley Bacon and Jeremy Barnum as a result of PSUs vesting, and who was the Chief Financial Officer at that time?
- **Ground truth:** 50%, Jeremy Barnum
- ✓ **Baseline (vanilla RAG) (PASS):** Ashley Bacon and Jeremy Barnum each received **50% of the net issued shares** resulting from PSUs vesting on March 25, 2026. **Jeremy Barnum** was the Chief Financial Officer.
- ✓ **VERITROOPER (PASS):** Ashley Bacon and Jeremy Barnum each received 50% of the net issued shares from PSUs vesting. Jeremy Barnum was the Chief Financial Officer.

29. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0134\_02`

- **Question:** What is the average cost of a cybersecurity breach to JPMorganChase?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide an average cost of a cybersecurity breach to JPMorganChase.
- ✓ **VERITROOPER (PASS):** The evidence does not state the average cost of a cybersecurity breach to JPMorganChase.

30. PASS · Precision · `SEC\_10K\_multi\_0263\_01`

- **Question:** What is the year for which the Net interest income - reported is shown?
- **Ground truth:** 2023
- ✓ **Baseline (vanilla RAG) (PASS):** Net interest income - reported is shown for **2025, 2024, and 2023**.
- ✓ **VERITROOPER (PASS):** Net interest income-reported is shown for 2025, 2024, and 2023.

31. PASS · Precision · `SEC\_10K\_multi\_1600\_01`

- **Question:** What is the amount of loss carryforwards that must be utilized within the next five years?
- **Ground truth:** 340 million
- ✓ **Baseline (vanilla RAG) (PASS):** \$340 million.

- ✓ **VERITROOPER (PASS):** \$340 million.

32. PASS · Cause & Effect · `SEC\_10K\_multi\_1185\_01`

- **Question:** What happens because of the Firm's right to call for additional margin when the fair value of collateral declines?
- **Ground truth:** Credit risk arising from lending activities subject to collateral maintenance requirements is generally mitigated.
- ✓ **Baseline (vanilla RAG) (PASS):** The right to call for additional margin mitigates credit risk, so these collateralized exposures generally do not require an allowance for credit losses.
- ✓ **VERITROOPER (PASS):** This right mitigates credit risk, so these exposures generally do not require an allowance for credit losses.

33. PASS · Exception · `SEC\_10K\_multi\_1316\_01`

- **Question:** What does the preferred stock take precedence over in the event of liquidation or dissolution?
- **Ground truth:** the Firm's common stock with respect to the payment of dividends and the distribution of assets
- ✓ **Baseline (vanilla RAG) (PASS):** It takes precedence over the Firm's common stock regarding the payment of dividends and distribution of assets.
- ✓ **VERITROOPER (PASS):** It takes precedence over the firm's common stock regarding the payment of dividends and distribution of assets.

34. PASS · Cause & Effect · `SEC\_10K\_multi\_1541\_01`

- **Question:** What may happen if our business continuity plans do not effectively resolve issues in a timely manner?
- **Ground truth:** We could experience delays in reporting our financial results, and we may lose revenue and profits.
- ✓ **Baseline (vanilla RAG) (PASS):** We could experience delays in reporting financial results and lose revenue and profits due to an inability to timely manufacture, distribute, invoice, and collect payments for concentrates or finished products.
- ✓ **VERITROOPER (PASS):** The company could experience delays in reporting financial results and lose revenue and profits due to an inability to timely manufacture, distribute, invoice, and collect payments for concentrates or finished products.

35. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1279\_00`

- **Question:** What is the average time it takes for reimbursement of these advances?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The advances are generally reimbursed "within a short period of time"; no specific average duration is provided.
- ✓ **VERITROOPER (PASS):** The evidence does not specify an average reimbursement time; it only states that reimbursement generally occurs within a short period.

36. PASS · Precision · `SEC\_10K\_multi\_0037\_01`

- **Question:** What is the number for Exhibit 104?
- **Ground truth:** 104
- ✓ **Baseline (vanilla RAG) (PASS):** Exhibit 104 is the **Cover Page Interactive Data File**.
- ✓ **VERITROOPER (PASS):** The number for Exhibit 104 is **104**.

37. PASS · Cause & Effect · `SEC\_10K\_multi\_0044\_01`

- **Question:** What happens as a consequence of failing to develop or monetize new products and technologies?
- **Ground truth:** Our financial results could be adversely affected.
- ✓ **Baseline (vanilla RAG) (PASS):** The company's financial results could be adversely affected, particularly if the products or technologies are not widely adopted.
- ✓ **VERITROOPER (PASS):** The company's financial results could be adversely affected.

38. FAIL · Precision · `SEC\_10K\_multi\_0304\_00`

MISINTERPRETED

- **Question:** What was the value of mortgage loans 90 or more days past due and insured by U.S. government agencies at December 31, 2023?
- **Ground truth:** 123 million
- ✗ **Baseline (vanilla RAG) (FAIL):** At December 31, 2023, mortgage loans 90 or more days past due and insured by U.S. government agencies totaled **\$182 million** (\$123 million in CCB and \$59 million in CIB).
- ✓ **VERITROOPER (PASS):** \$123 million.

39. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0135\_00`

- **Question:** What is the average annual cost of maintaining adequate data management processes for JPMorganChase?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose an average annual cost for maintaining adequate data management processes at JPMorganChase.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the annual cost of maintaining adequate data management processes, so an average cannot be calculated.

40. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1062\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of the defined benefit pension and OPEB plans' liabilities as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, the defined benefit pension and OPEB plans' total projected benefit obligations were **\$14.724 billion**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total value of the defined benefit pension and OPEB plans' liabilities as of December 31, 2025.

41. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1595\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average annual return on the U.S. pension plan's assets for the year 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The expected average annual return on the U.S. pension plan's assets for 2025 was **6.75%**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the average annual return on the U.S. pension plan's assets for 2025.

42. PASS · Exception · `SEC\_10K\_multi\_0067\_01`

- **Question:** Under what circumstance do RSUs not vest?
- **Ground truth:** unless continued service
- ✓ **Baseline (vanilla RAG) (PASS):** RSUs generally do not vest if the recipient fails to meet the continued service or employment requirement-for example, by leaving the company before the vesting date-subject to specified exceptions or eligibility provisions.
- ✓ **VERITROOPER (PASS):** RSUs generally do not vest if the recipient fails to continue providing service during the four-year vesting period.

43. FAIL · Precision · `SEC\_10K\_multi\_0934\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total fair value of trading assets and liabilities?
- **Ground truth:** 665,587
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, the total fair value was:
- **Trading assets:** \$606.382 billion gross; **\$57.777 billion net**
- **Trading liabilities:** \$596.738 billion gross; **\$46.329 billion net**
- ✓ **VERITROOPER (PASS):** The total fair value of trading assets and liabilities was \$665,587 for the first period and \$653,218 for the second period.

44. FAIL · Cross-Reference · `SEC\_10K\_multi\_1588\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total amount of assets held for sale as of December 31, 2025, and how does it compare to the total amount of liabilities held for sale?
- **Ground truth:** The total amount of assets held for sale as of December 31, 2025, is \$5,342 million, and the total amount of liabilities held for sale is \$2,570 million.
- ✗ **Baseline (vanilla RAG) (FAIL):** Assets held for sale totaled **\$5.342 billion** as of December 31, 2025. The evidence does not provide the complete total for liabilities held for sale, so an exact comparison cannot be determined; it only shows **\$816 million** of accounts payable and accrued expenses within those liabilities.
- ✗ **VERITROOPER (FAIL):** As of December 31, 2025, assets held for sale totaled \$5.342 billion, compared with liabilities held for sale of \$2.570 billion; assets exceeded liabilities by \$2.772 billion.

45. FAIL · Precision · `SEC\_10K\_multi\_1614\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the number of the certification executed by John Murphy?
- **Ground truth:** 32.1
- ✗ **Baseline (vanilla RAG) (FAIL):** The certification executed by John Murphy is **Exhibit 31.2**.
- ✓ **VERITROOPER (PASS):** 32.1

46. FAIL · Precision · `SEC\_10K\_multi\_1113\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total loan balance at December 31, 2025?
- **Ground truth:** 1,493,429
- ✗ **Baseline (vanilla RAG) (FAIL):** The total retained loan balance at December 31, 2025 was **\$1.409 trillion** (\$1,408,905 million).
- ✓ **VERITROOPER (PASS):** The total loan balance at December 31, 2025 was **\$1,493,429 million**.

47. PASS · Exception · `SEC\_10K\_multi\_0052\_01`

- **Question:** What could result in substantial costs to us and our ability to compete?
- **Ground truth:** actions to adequately protect our IP rights
- ✓ **Baseline (vanilla RAG) (PASS):** Actions to adequately protect our intellectual property (IP) rights could result in substantial costs, and failure to protect those rights-or being prohibited from making or selling products-could harm our ability to compete.
- ✓ **VERITROOPER (PASS):** Actions to adequately protect our intellectual property rights could result in substantial costs, while failure or prohibition from making or selling products could harm our ability to compete.

48. FAIL · Cross-Reference · `SEC\_10K\_multi\_0216\_01`

BAD QUESTION

- **Question:** What is the cumulative total return for JPMorgan Chase & Co. common stock over the five-year period ending December 31, 2025, and how does it compare to the S&P 500 Index?
- **Ground truth:** The cumulative total return for JPMorgan Chase & Co. common stock was \$289.18, which is higher than the S&P 500 Index return of \$196.12.
- ✗ **Baseline (vanilla RAG) (FAIL):** Over the five years ended December 31, 2025:
- **JPMorgan Chase common stock: 189.18% cumulative total return** (\$100 grew to \$289.18)
- **S&P 500 Index: 96.12% cumulative total return** (\$100 grew to \$196.12)

JPMorgan Chase outperformed the S&P 500 by **93.06 percentage points**, assuming dividends were reinvested.

- ✗ **VERITROOPER (FAIL):** JPMorganChase common stock returned 189.18% cumulatively, versus 96.12% for the S&P 500 Index, outperforming it by 93.06 percentage points.

49. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0576\_03`

- **Question:** What is the average FICO score of the credit card loans in the Firm's portfolio at December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The average FICO score is not disclosed. At December 31, 2025, 84.6% of retained credit card loans had refreshed FICO scores of 660 or higher, 15.2% were below 660, and 0.2% had no FICO score available.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the average FICO score of the Firm's credit card loans at December 31, 2025.

50. PASS · Cross-Reference · `SEC\_10K\_multi\_1173\_01`

- **Question:** What is the weighted-average extension of loans for single modifications, and how does it compare to the extension for interest rate reduction and term extension?
- **Ground truth:** 20 months for single modifications, and 36 months for interest rate reduction and term extension.
- ✓ **Baseline (vanilla RAG) (PASS):** For 2024 commercial and industrial loans, single term-extension modifications extended loans by a weighted average of **20 months**. Interest-rate-reduction and term-extension modifications extended them by **36 months, 16 months longer**.
- ✗ **VERITROOPER (FAIL):** Single-modification term extensions averaged 20 months, compared with 36 months for interest rate reduction and term extension-16 months longer.

51. FAIL · Cross-Reference · `SEC\_10K\_multi\_0583\_00`

WRONG CALC/QUOTE

- **Question:** What is the total amount of loans retained after 5 years, and what percentage of investment-grade loans is that?
- **Ground truth:** 175,215, 68%
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, loans retained after five years totaled **\$189.81 9 billion**, and **68%** of total retained loans were investment-grade.
- **✓ VERITROOPER (PASS):** Loans retained after 5 years totaled **\$175,215 million**, and **68%** of total loans retained were investment-grade.

52. FAIL · Precision · `SEC\_10K\_multi\_1501\_01`

UNSUPPORTED NUMBER

- **Question:** What is the decrease in deposits with banks due to change in rate?
- **Ground truth:** (2,792)
- **✗ Baseline (vanilla RAG) (FAIL):** The decrease due to changes in rates was **\$4.352 billion** in 2025 versus 2024:
 - U.S.: \$2.792 billion
 - Non-U.S.: \$1.560 billion
- **✗ VERITROOPER (FAIL):** The decrease in deposits with banks due to the change in rate was **\$2.792 billion**.

53. PASS · Exception · `SEC\_10K\_multi\_0063\_00`

- **Question:** What does the auditor's opinion NOT apply to?
- **Ground truth:** material misstatement, whether due to error or fraud
- **✓ Baseline (vanilla RAG) (PASS):** The auditor does **not provide a separate opinion on the critical audit matter or the related accounts or disclosures**.
- **✓ VERITROOPER (PASS):** The auditor's opinion does not apply separately to the critical audit matter or the related accounts or disclosures.

54. PASS · Exception · `SEC\_10K\_multi\_1217\_00`

- **Question:** Under what circumstance does the Firm NOT consolidate residential mortgage securitizations?
- **Ground truth:** when it is not the servicer or does not hold a beneficial interest in the trust that could potentially be significant to the trust
- **✓ Baseline (vanilla RAG) (PASS):** The Firm does not consolidate residential mortgage securitizations when it:
 - Is not the servicer and therefore lacks power to direct the trust's most significant activities; or
 - Does not hold a beneficial interest that could potentially be significant to the trust.
- **✓ VERITROOPER (PASS):** The Firm does not consolidate residential mortgage securitizations when it is not the servicer or does not hold a beneficial interest that could potentially be significant to the trust.

55. PASS · Precision · `SEC\_10K\_multi\_1358\_01`

- **Question:** What is the percentage limit for covered transactions between a banking subsidiary and its affiliates in relation to total capital?
- **Ground truth:** 10
- **✓ Baseline (vanilla RAG) (PASS):** Covered transactions are limited to:

- **10%** of the banking subsidiary's total capital for any single affiliate.
- **20%** of total capital in aggregate for all affiliates.
- ✓ **VERITROOPER (PASS)**: Covered transactions with any particular affiliate are generally limited to 10% of total capital, while aggregate transactions with all affiliates are limited to 20%.

56. FAIL · Precision · `SEC\_10K\_multi\_0747\_01`

BAD QUESTION

- **Question**: What is the total value of asset-backed securities?
- **Ground truth**: 11,316
- ✗ **Baseline (vanilla RAG) (FAIL)**: The total fair value of available-for-sale asset-backed securities was **\$23.906 billion** as of December 31, 2025.
- ✗ **VERITROOPER (FAIL)**: The total value of asset-backed securities is **2,823**.

57. PASS · Cross-Reference · `SEC\_10K\_multi\_1037\_00`

- **Question**: What is the typical duration of the revenue-sharing agreements with co-brand partners, and how does it relate to the recognition of card income?
- **Ground truth**: The terms of these agreements generally range from five to ten years.
- ✓ **Baseline (vanilla RAG) (PASS)**: Co-brand partner agreements typically last **five to ten years**. Related partner payments affect card income based on their nature:
 - **Loyalty rewards**: reduce card income when incurred.
 - **New account originations**: deferred and recognized as a reduction of card income over **12 months**.
 - **Sales-volume payments**: reduce card income when the related interchange income is earned.
 - **Marketing payments**: expensed as incurred as marketing expense, rather than reducing card income.
- ✓ **VERITROOPER (PASS)**: The agreements typically last five to ten years. Related partner payments reduce card income when rewards costs are incurred or interchange income is earned, while new-account origination payments are deferred and recognized over 12 months.

58. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0767\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question**: What is the correlation between equity and commodity derivatives?
- **Ground truth**: UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL)**: No direct equity-commodity correlation is disclosed. Separately, the weighted-average inputs are **54% for equity correlation** and **1% for commodity correlation**.
- ✓ **VERITROOPER (PASS)**: The evidence does not provide a correlation between equity and commodity derivatives.

59. FAIL · Precision · `SEC\_10K\_multi\_0825\_01`

MISINTERPRETED

- **Question**: What is the total number of debt instruments?
- **Ground truth**: 2,178
- ✗ **Baseline (vanilla RAG) (FAIL)**: Total debt instruments were **\$2,182 million (\$2.182 billion)** as of December 31, 2025.
- ✓ **VERITROOPER (PASS)**: The total debt instruments were 2,178.

60. PASS · Calculation · `SEC\_10K\_multi\_0730\_01`

- **Question:** If the net asset value per share is \$178, what is the fair value of 100 shares?
- **Ground truth:** 17800
- ✓ **Baseline (vanilla RAG) (PASS):** 17800
- ✓ **VERITROOPER (PASS):** 17800

61. FAIL · Cross-Reference · `SEC\_10K\_multi\_1181\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the percentage of loan modifications to total retained Other loans for single modifications?
- **Ground truth:** 0.10
- ✗ **Baseline (vanilla RAG) (FAIL):** For 2025, single modifications represented **0.03%** of total retained Other loans.
- ✓ **VERITROOPER (PASS):** The percentage was **0.10%** for single modifications in 2023.

62. PASS · Precision · `SEC\_10K\_multi\_0108\_01`

- **Question:** What is the date of the signature of Jen-Hsun Huang?
- **Ground truth:** February 25, 2026
- ✓ **Baseline (vanilla RAG) (PASS):** February 25, 2026.
- ✓ **VERITROOPER (PASS):** February 25, 2026.

63. PASS · Cause & Effect · `SEC\_10K\_multi\_0013\_01`

- **Question:** What happens as a consequence of higher net sales of Services?
- **Ground truth:** Europe net sales increase
- ✓ **Baseline (vanilla RAG) (PASS):** Higher Services net sales contributed to an increase in Services gross margin in 2025 compared with 2024.
- ✓ **VERITROOPER (PASS):** Higher net sales of Services contributed to increased 2025 net sales in the Americas, Europe, Japan, and Rest of Asia Pacific compared with 2024.

64. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0516\_04`

- **Question:** What is the total amount of deposits that are held by JPMorgan Chase's non-bank subsidiaries as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, the Parent Company held **\$9.751 billion** in deposits with banking subsidiaries. The evidence does not separately report deposits held by non-bank subsidiaries.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the amount of deposits held by JPMorgan Chase's non-bank subsidiaries as of December 31, 2025.

65. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0135\_04`

- **Question:** What is the percentage of JPMorganChase's clients and customers that could be negatively impacted by extraordinary events?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a percentage of JPMorganChase's clients and customers that could be negatively impacted by extraordinary events.

- ✓ **VERITROOPER (PASS):** The evidence does not specify the percentage of JPMorganChase's clients and customers that could be negatively impacted by extraordinary events.

66. PASS · Precision · `SEC\_10K\_multi\_0072\_00`

- **Question:** What percentage of accounts receivable balance as of January 25, 2026, was accounted for by three direct customers?
- **Ground truth:** 25
- ✓ **Baseline (vanilla RAG) (PASS):** The three direct customers accounted for **25%, 18%, and 13%**, totaling **56%** of the accounts receivable balance as of January 25, 2026.
- ✗ **VERITROOPER (FAIL):** The three direct customers accounted for a combined 56% of accounts receivable as of January 25, 2026.

67. PASS · Conditional · `SEC\_10K\_multi\_1538\_01`

- **Question:** If we are unable to maintain good relationships with our bottling partners, what could happen to our business?
- **Ground truth:** It could suffer.
- ✓ **Baseline (vanilla RAG) (PASS):** Our business could suffer. Bottling partners might prioritize other products or short-term profits, reduce support for our brands, or take actions that harm brand reputation, consumer perception, profitability, and financial results.
- ✓ **VERITROOPER (PASS):** Our business could suffer, including potential harm to our brands and an adverse long-term effect on profitability.

68. FAIL · Precision · `SEC\_10K\_multi\_1159\_00`

MISINTERPRETED

- **Question:** What is the total retained Secured by real estate loans for 2025?
- **Ground truth:** 105,130
- ✗ **Baseline (vanilla RAG) (FAIL):** The total retained loans secured by real estate for 2025 were **\$165.545 billion**.
- ✗ **VERITROOPER (FAIL):** The total retained secured by real estate loans for 2025 was **\$165.545 billion**.

69. FAIL · Precision · `SEC\_10K\_multi\_1350\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total income tax expense and effective tax rate for the first period?
- **Ground truth:** 15,547
- ✗ **Baseline (vanilla RAG) (FAIL):** The total income tax expense was **\$15.547 billion**, with an effective tax rate of **21.4%**.
- ✓ **VERITROOPER (PASS):** The total income tax expense was **\$15,547**, with an effective tax rate of **21.4%**.

70. PASS · Precision · `SEC\_10K\_multi\_1600\_00`

- **Question:** What is the amount of loss carryforwards available to reduce future taxable income as of December 31, 2025?
- **Ground truth:** 1,600 million
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, loss carryforwards available to reduce future taxable income were **\$1.6 billion**.
- ✗ **VERITROOPER (FAIL):** As of December 31, 2025, the Company had **\$1.6 billion** in loss carryforwards available to reduce future taxable income.

71. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0581\_04`

- **Question:** What is the total value of loans that are both investment-grade and held for more than 5 years?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The exact combined value is not provided. For 2025, retained loans totaled **\$189.819 billion** with maturities after five years, while **\$541.364 billion** were investment-grade, but the overlap between these categories is not disclosed.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the value of loans that are both investment-grade and held for more than five years.

72. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_04`

BAD QUESTION

- **Question:** What is the total number of notes issued by Apple Inc. in 2022?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Apple Inc. issued **4 notes** in 2022.
- ✗ **VERITROOPER (FAIL):** Apple Inc. issued four notes in 2022.

73. PASS · Cross-Reference · `SEC\_10K\_multi\_0498\_00`

- **Question:** What is the date by which the benefit from the CECL capital transition provision had been fully phased out?
- **Ground truth:** January 1, 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The benefit was fully phased out as of **January 1, 2025**.
- ✓ **VERITROOPER (PASS):** January 1, 2025.

74. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0714\_04`

HALLUCINATION

- **Question:** What is the name of the specific firm or organization responsible for issuing the new guidance?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The Financial Accounting Standards Board (FASB).
- ✓ **VERITROOPER (PASS):** The evidence does not identify the specific firm or organization that issued the guidance.

75. PASS · Cross-Reference · `SEC\_10K\_multi\_0901\_00`

- **Question:** Given the credit risk concentrations in the consumer portfolio and the methods used to manage them, what is the primary focus of the Firm's credit risk management?
- **Ground truth:** trends and concentrations at the portfolio level
- ✓ **Baseline (vanilla RAG) (PASS):** The primary focus is monitoring consumer credit trends and concentrations by product and U.S. geographic region at the portfolio level, then mitigating risks through changes to underwriting policies and portfolio guidelines.
- ✓ **VERITROOPER (PASS):** The Firm primarily focuses on product and U.S. geographic-region trends and concentrations at the portfolio level, managing risks through changes in underwriting policies and portfolio guidelines.

76. PASS · Calculation · `SEC\_10K\_multi\_0730\_00`

- **Question:** If the net asset value per share is \$178, what is the fair value of a single share?
- **Ground truth:** 178
- ✓ **Baseline (vanilla RAG) (PASS):** 178
- ✓ **VERITROOPER (PASS):** 178

77. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_02`

- **Question:** What is the average annual loss due to data breaches for the Company?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose an average annual loss due to data breaches. It only states that past security incidents have not had a material adverse effect on the Company's business.
- ✓ **VERITROOPER (PASS):** The evidence does not state the Company's average annual loss due to data breaches.

78. FAIL · Cross-Reference · `SEC\_10K\_multi\_1115\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the total amount of sales for the year ended December 31, 2025?
- **Ground truth:** 55,583
- ✗ **Baseline (vanilla RAG) (FAIL):** The total amount of sales for the year ended December 31, 2025 was **\$55.583 billion** (\$55,583 million).
- ✓ **VERITROOPER (PASS):** The total amount of sales for the year ended December 31, 2025 was **\$55,583 million**.

79. PASS · Cause & Effect · `SEC\_10K\_multi\_1045\_01`

- **Question:** What happens due to the adoption of updates to the Accounting for Investments in Tax Credit Structures Using the Proportional Amortization Method?
- **Ground truth:** The amortization of certain of the Firm's alternative energy tax-oriented investments is now recognized in income tax expense.
- ✓ **Baseline (vanilla RAG) (PASS):** Effective January 1, 2024, adoption of the updates:
 - Expanded the tax-oriented investments eligible for proportional amortization beyond affordable housing, elected program by program.
 - Reclassified amortization of certain alternative-energy investments from other income to income tax expense, aligning it with related tax benefits.
 - Changed the timing of that amortization.
 - Decreased retained earnings by \$161 million.
 - Increased first-quarter 2024 income tax expense by about \$450 million and the effective tax rate by roughly two percentage points.
 - Had no material impact on net income.
- ✓ **VERITROOPER (PASS):** Effective January 1, 2024, amortization of certain alternative energy tax-oriented investments shifted from other income to income tax expense, aligning it with the associated tax credits and other tax benefits.

80. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0825\_03`

HALLUCINATION

- **Question:** What is the total value of equity securities held by individuals?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, exposure to **Individuals and Individual Entities** totaled **\$167.7 billion**.
- **✓ VERITROOPER (PASS):** The evidence does not specify the value of equity securities held by individuals.

81. PASS · Precision · `SEC\_10K\_multi\_0034\_01`

- **Question:** What is the filing date for Exhibit 10.1?
- **Ground truth:** 12/28/24
- **✓ Baseline (vanilla RAG) (PASS):** The filing date for Exhibit 10.1 is not provided in the evidence.
- **✓ VERITROOPER (PASS):** The filing date for Exhibit 10.1 was December 28, 2024.

82. PASS · Cause & Effect · `SEC\_10K\_multi\_0135\_00`

- **Question:** What could happen if JPMorganChase's management fails to develop and execute effective business strategies?
- **Ground truth:** JPMorganChase's results or competitive standing could suffer.
- **✓ Baseline (vanilla RAG) (PASS):** JPMorganChase's competitive standing and business value could suffer, its growth prospects could weaken, and its earnings could decline. It may also struggle to control expenses, develop resources, return capital to shareholders, and achieve its strategic objectives.
- **✓ VERITROOPER (PASS):** JPMorganChase's competitive standing and business values or growth prospects could suffer, and its earnings could decline.

83. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0576\_04`

- **Question:** What is the total value of derivative receivables in the wholesale credit portfolio at December 31, 2026?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not provide a December 31, 2026 figure. At December 31, 2025, wholesale derivative receivables totaled **\$57.777 billion**.
- **✓ VERITROOPER (PASS):** The evidence does not provide derivative receivables as of December 31, 2026; it only provides values for 2025 and 2024.

84. PASS · Precision · `SEC\_10K\_multi\_0023\_00`

- **Question:** What is the threshold for Level 2?
- **Ground truth:** 3
- **✓ Baseline (vanilla RAG) (PASS):** The threshold award level was approximately **996,000 performance share units** as of December 31, 2025.
- **✓ VERITROOPER (PASS):** The evidence does not specify a threshold for Level 2.

85. PASS · Exception · `SEC\_10K\_multi\_0713\_00`

- **Question:** What does not apply to the adoption of Income Taxes: Improvements to Income Tax Disclosures?
- **Ground truth:** Adoption before January 1, 2025
- **✓ Baseline (vanilla RAG) (PASS):** A **prospective transition method does not apply** to JPMorgan Chase's adoption. The firm adopted the guidance **retrospectively** for its 2025 annual consolidated financial

statements.

- ✓ **VERITROOPER (PASS):** A required effective date of annual financial statements for the year ending December 31, 2027, does not apply; that date relates to Expense Disaggregation Disclosures.

86. FAIL · Cross-Reference · `SEC\_10K\_multi\_0572\_00`

BAD QUESTION

- **Question:** What is the total amount of consumer loans excluding credit card loans, and what is the total amount of consumer loans including credit card loans?
- **Ground truth:** The total amount of consumer loans excluding credit card loans is \$402,258, and the total amount of consumer loans including credit card loans is \$247,797.
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025:
- **Consumer loans excluding credit card loans: \$402.258 billion**
- **Total consumer loans including credit card loans: \$650.055 billion**
- ✓ **VERITROOPER (PASS):** Consumer loans excluding credit card loans totaled **\$402,258**, while total consumer loans including credit card loans were **\$650,055**.

87. PASS · Precision · `SEC\_10K\_multi\_1347\_00`

- **Question:** What was the net pre-tax gain reclassified in 2024?
- **Ground truth:** 7 million
- ✓ **Baseline (vanilla RAG) (PASS):** The net pre-tax gain reclassified in 2024 was **\$7 million**.
- ✓ **VERITROOPER (PASS):** The net pre-tax gain reclassified in 2024 was **\$7 million**.

88. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0137\_01`

- **Question:** How many employees does JPMorganChase have in the United States?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not state how many JPMorganChase employees are based in the United States.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many employees JPMorganChase has in the United States.

89. PASS · Cross-Reference · `SEC\_10K\_multi\_1598\_00`

- **Question:** What is the percentage of total U.S. pension plan assets invested in the Company's common stock as of December 31, 2025?
- **Ground truth:** 3%
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, **3%** of total U.S. pension plan assets were invested in the Company's common stock.
- ✓ **VERITROOPER (PASS):** As of December 31, 2025, 3% of total U.S. pension plan assets were invested in the Company's common stock.

90. PASS · Exception · `SEC\_10K\_multi\_1595\_01`

- **Question:** What does not change to underlying benefits or plan administration in the settlement of U.S. other postretirement benefit obligations?
- **Ground truth:** underlying benefits or plan administration
- ✓ **Baseline (vanilla RAG) (PASS):** Nothing changed regarding the underlying benefits or plan administration; only the future financing shifted to an insurance company.
- ✓ **VERITROOPER (PASS):** The settlement changed only the future financing of the benefits; the underlying benefits and plan administration did not change.

91. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1265\_02`

- **Question:** What is the average annual increase in goodwill for the Commercial & Investment Bank segment?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide goodwill figures for the Commercial & Investment Bank segment, so its average annual increase cannot be calculated.
- ✗ **VERITROOPER (FAIL):** The average annual increase in goodwill was **\$4 million** from 2023 to 2025.

92. FAIL · Cross-Reference · `SEC\_10K\_multi\_0822\_01`

TABLE MISREAD

- **Question:** What is the fair value at December 31, 2023, and what is the change in unrealized gains/(losses) related to financial instruments held at December 31, 2023?
- **Ground truth:** 34,564, 738
- ✗ **Baseline (vanilla RAG) (FAIL):** For **long-term debt**, the fair value at December 31, 2023 was **\$27.726 billion**, and the change in unrealized gains/(losses) on instruments held at that date was **\$2.870 billion**.
- ✓ **VERITROOPER (PASS):** The fair value at December 31, 2023 was **\$34,564 million**, and the change in unrealized gains/(losses) was **\$1,212 million**.

93. FAIL · Precision · `SEC\_10K\_multi\_0060\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the amount of cash used in financing activities for fiscal year 2026?
- **Ground truth:** 48,474
- ✗ **Baseline (vanilla RAG) (FAIL):** Cash used in financing activities for fiscal year 2026 was **\$48.474 billion**.
- ✗ **VERITROOPER (FAIL):** Net cash used in financing activities for fiscal year 2026 was \$48.474 billion.

94. PASS · Conditional · `SEC\_10K\_multi\_0728\_01`

- **Question:** Only if the valuation is based on operating performance, what additional inputs are relevant?
- **Ground truth:** Additional available inputs relevant to the investment
- ✓ **Baseline (vanilla RAG) (PASS):** Additional relevant inputs may include:
 - Transaction prices
 - Trading multiples of comparable public companies
 - Adjustments for differences from comparable companies and company-specific issues, including lack of liquidity
 - Any other available inputs relevant to the investment
- ✓ **VERITROOPER (PASS):** Additional inputs may include transaction prices, trading multiples of comparable public companies, company-specific adjustments such as lack of liquidity, and other available inputs relevant to the investment.

95. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1274\_01`

- **Question:** What is the average contractual service fee included in income?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide the average contractual servicing fee included in income.
- ✗ **VERITROOPER (FAIL):** The average contractual service fee included in income was **1,610.33**.

96. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0135\_02`

- **Question:** What is the total amount of potential losses that JPMorganChase could incur from inadequate risk management?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a total monetary amount. It only states that inadequate risk management could cause unexpected or significant losses and materially adversely affect JPMorganChase's financial condition or results.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a total amount of potential losses from inadequate risk management.

97. PASS · Exception · `SEC\_10K\_multi\_1585\_00`

- **Question:** Under what circumstance does the Company NOT evaluate whether shipping and handling activities are services promised to customers?
- **Ground truth:** If revenue is recognized for the related goods before the shipping and handling activities occur.
- ✓ **Baseline (vanilla RAG) (PASS):** The Company does not evaluate them as promised services when shipping and handling activities occur after the customer obtains control of the goods; instead, it treats them as fulfillment costs.
- ✓ **VERITROOPER (PASS):** The Company does not evaluate this when shipping and handling activities occur after the customer obtains control of the goods, because it elects to treat them as fulfillment costs.

98. PASS · Precision · `SEC\_10K\_multi\_1169\_01`

- **Question:** What is the reduction in weighted-average contractual interest for multiple modifications in 2024?
- **Ground truth:** 162
- ✓ **Baseline (vanilla RAG) (PASS):** The weighted-average contractual interest rate was reduced by **162 basis points**.
- ✓ **VERITROOPER (PASS):** The weighted-average contractual interest was reduced by **162 basis points (bps)** in 2024.

99. FAIL · Cross-Reference · `SEC\_10K\_multi\_1403\_01`

UNSUPPORTED NUMBER

- **Question:** How do the unsettled repurchase and securities loaned agreements compare to the unsettled resale and securities borrowed agreements?
- **Ground truth:** 52,308 is less than 136,841.
- ✗ **Baseline (vanilla RAG) (FAIL):** - **Unsettled repurchase and securities loaned agreements: \$52.9 billion** at year-end 2025, down from **\$67.0 billion** in 2024. At settlement, the Firm receives cash and provides securities collateral.
- **Unsettled resale and securities borrowed agreements: \$137.1 billion** in 2025, up from **\$115.9 billion** in 2024. At settlement, the Firm advances cash and receives securities collateral.

Thus, the resale and securities borrowed commitments were about **\$84.2 billion larger** in 2025.

- ✗ **VERITROOPER (FAIL):** Unsettled repurchase and securities loaned agreements were lower than unsettled resale and securities borrowed agreements: 52,895 versus 137,072 in Col. 5, and 66,986 versus 115,939 in Col. 6.

100. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_01`

- **Question:** How many employees are hired by competitors annually?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify how many employees are hired by competitors annually.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many employees are hired by competitors annually.

101. FAIL · Cross-Reference · `SEC\_10K\_multi\_0754\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total amount of beneficial interests issued by consolidated VIEs and long-term debt combined?
 - **Ground truth:** $1 + 100,780 = 100,781$
 - ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025:
 - Beneficial interests issued by consolidated VIEs: **\$27.951 billion**
 - Long-term debt: **\$435.206 billion**
- Combined total: \$463.157 billion.**
- ✓ **VERITROOPER (PASS):** The combined total is **\$100,781** ($\$1 + \$100,780$).

102. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1169\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the percentage of loans that were extended by more than 21 months in 2024?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** In 2024, **0.07% of retained "Other" loans** had term extensions averaging 28 months, which was more than 21 months.
- ✓ **VERITROOPER (PASS):** The evidence does not state the percentage of loans extended by more than 21 months; it only reports a weighted-average extension of 21 months.

103. PASS · Exception · `SEC\_10K\_multi\_1105\_00`

- **Question:** What type of loans are not placed on nonaccrual status at the time of modification?
- **Ground truth:** credit card loans
- ✓ **Baseline (vanilla RAG) (PASS):** Loans that are current at the modification date, in certain limited and well-defined circumstances, are not placed on nonaccrual status.
- ✓ **VERITROOPER (PASS):** In limited, well-defined circumstances, loans that are current at the modification date are not placed on nonaccrual status.

104. PASS · Cause & Effect · `SEC\_10K\_multi\_0049\_01`

- **Question:** What happens as a consequence of failure to implement new or updated controls?
- **Ground truth:** Our operating results may be harmed, and we may fail to meet our reporting obligations.
- ✓ **Baseline (vanilla RAG) (PASS):** Failure to implement new or updated controls could harm operating results or cause the company to fail to meet its reporting obligations.
- ✓ **VERITROOPER (PASS):** It could harm operating results or cause the company to fail to meet its reporting obligations.

105. FAIL · Precision · `SEC\_10K\_multi\_0725\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total liabilities at December 31, 2024?
- **Ground truth:** 27,777
- **✗ Baseline (vanilla RAG) (FAIL):** Total liabilities at December 31, 2024 were **\$3,658,056 million** (approximately **\$3.66 trillion**).
- **✓ VERITROOPER (PASS):** Total liabilities at December 31, 2024 were **\$27,777 million**.

106. PASS · Cross-Reference · `SEC\_10K\_multi\_0064\_01`

- **Question:** What is the net income for the year ended January 25, 2026, and the year ended January 26, 2025?
- **Ground truth:** 120,067 and 72,880
- **✓ Baseline (vanilla RAG) (PASS):** - Year ended January 25, 2026: **\$120.067 billion**
- **Year ended January 26, 2025: \$72.880 billion**
- **✓ VERITROOPER (PASS):** Net income was **\$120,067 million** for the year ended January 25, 2026, and **\$72,880 million** for the year ended January 26, 2025.

107. PASS · Cross-Reference · `SEC\_10K\_multi\_0863\_00`

- **Question:** What factors affect the fair value of a loan but do not affect its carrying value?
- **Ground truth:** Changes in market interest rates, including credit spreads, market liquidity premiums, and other factors.
- **✓ Baseline (vanilla RAG) (PASS):** Changes in market interest rates-including credit spreads-market liquidity premiums, and other market-related factors affect a loan's fair value but not its carrying value.
- **✓ VERITROOPER (PASS):** Changes in market interest rates, including credit spreads and market liquidity premiums, as well as other unspecified factors.

108. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0137\_03`

- **Question:** How many countries does JPMorganChase conduct business in?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify how many countries JPMorganChase conducts business in.
- **✓ VERITROOPER (PASS):** The evidence does not specify how many countries JPMorganChase conducts business in.

109. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1602\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average annual return on the Company's pension plan assets?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The Company's expected long-term annual return on pension plan assets for 2025 was **6.75%**.
- **✓ VERITROOPER (PASS):** The evidence does not provide the average annual return on the Company's pension plan assets.

110. PASS · Precision · `SEC\_10K\_multi\_0112\_00`

- **Question:** What is the year in which the Corporate Sustainability Reporting Directive (CSRD) phased-in requirements started?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The CSRD phased-in requirements started in fiscal year **2024**.
- ✓ **VERITROOPER (PASS):** The CSRD phased-in requirements started in fiscal year 2024.

111. FAIL · Cross-Reference · `SEC\_10K\_multi\_1115\_00`

BAD QUESTION

- **Question:** What is the total amount of retained loans reclassified to held-for-sale for the year ended December 31, 2025?
- **Ground truth:** 55,583
- ✗ **Baseline (vanilla RAG) (FAIL):** The total retained loans reclassified to held-for-sale in 2025 was **\$1.552 billion**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total amount of retained loans reclassified to held-for-sale for 2025.

112. PASS · Cause & Effect · `SEC\_10K\_multi\_1569\_00`

- **Question:** What happens as a result of the Company's issuances of debt?
- **Ground truth:** The Company's issuances of debt can result in a lower overall cost of borrowing.
- ✓ **Baseline (vanilla RAG) (PASS):** Debt issuances increase the Company's outstanding debt and interest expense and require compliance with debt covenants or other restrictions.
- ✓ **VERITROOPER (PASS):** The Company's debt financing lowers its overall cost of capital and increases return on shareowners' equity, but exposes it to adverse interest-rate changes.

113. PASS · Conditional · `SEC\_10K\_multi\_1538\_00`

- **Question:** If our bottling partners' financial condition deteriorates, what could happen to our business and financial results?
- **Ground truth:** They could be affected.
- ✓ **Baseline (vanilla RAG) (PASS):** A deterioration in one or more major bottling partners' financial condition could:
 - Reduce net operating revenues from concentrate and syrup sales.
 - Limit our ability to raise prices if bottlers cannot pass increases to customers.
 - Decrease equity income from bottling partners accounted for under the equity method.
 - Cause impairments of our equity-method investments.
 - Ultimately adversely affect our business and financial results.
- ✓ **VERITROOPER (PASS):** A deterioration could reduce net operating revenues from concentrate and syrup sales. If it involves equity-method investee bottlers, it could also decrease equity income and/or impair those investments.

114. PASS · Cross-Reference · `SEC\_10K\_multi\_0569\_01`

- **Question:** What is the basis for the Firm's policy regarding credit card loans?
- **Ground truth:** regulatory guidance
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm's policy is based on regulatory guidance permitting credit card loans to remain exempt from nonaccrual status. Accordingly, interest and fees continue accruing until the loans are charged off or paid in full.

- ✓ **VERITROOPER (PASS):** The Firm generally exempts credit card loans from nonaccrual status, as permitted by regulatory guidance.

115. PASS · Exception · `SEC\_10K\_multi\_0897\_01`

- **Question:** What type of notes do not obligate the Firm to return a stated amount of principal at maturity?
- **Ground truth:** nonprincipal-protected structured notes and long-term beneficial interests
- ✓ **Baseline (vanilla RAG) (PASS):** Nonprincipal-protected structured notes and long-term beneficial interests.
- ✓ **VERITROOPER (PASS):** Nonprincipal-protected structured notes and long-term beneficial interests.

116. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1095\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average credit risk exposure for resale and securities borrowed agreements as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The average net credit risk exposure was **\$31.967 billion**, calculated as:
 - Resale agreements: **\$12.209 billion**
 - Securities borrowed: **\$51.725 billion**

$$\frac{(\$12.209B + \$51.725B)}{2} = \$31.967B$$

- ✓ **VERITROOPER (PASS):** The evidence does not provide enough information to calculate the average credit risk exposure for resale and securities borrowed agreements as of December 31, 2025.

117. PASS · Cause & Effect · `SEC\_10K\_multi\_0636\_01`

- **Question:** What results in a net addition to the allowance for credit losses?
- **Ground truth:** The Firm's weighted-average macroeconomic outlook and qualitative adjustments, including the impact of changes in the Firm's weighted-average macroeconomic outlook and a reduction due to the impact of charge-offs.
- ✓ **Baseline (vanilla RAG) (PASS):** The 2025 net addition to the allowance for credit losses was **\$4.4 billion**, consisting of:
 - **\$3.3 billion consumer:** driven by **\$2.2 billion related to the Apple Card transaction**, Card Services loan growth, and changes in the weighted-average macroeconomic outlook, partly offset by reduced borrower uncertainty.
 - **\$1.1 billion wholesale:** driven by portfolio and lending-commitment growth, changes in client-specific credit quality, revised loss assumptions for certain leveraged loans, and estimated fraud losses in secured lending facilities, partly offset by macroeconomic-outlook changes.
- ✓ **VERITROOPER (PASS):** The \$4.4 billion net addition resulted from a \$3.3 billion increase in consumer allowances-driven mainly by the Apple Card transaction, Card Services loan growth, and macroeconomic changes-and a \$1.1 billion increase in wholesale allowances from portfolio growth, updated leveraged-loan loss assumptions, and changes in client-specific credit quality.

118. FAIL · Precision · `SEC\_10K\_multi\_0072\_01`

VALIDATOR OVER-FIRED

- **Question:** What was the total inventories as of January 26, 2025?
- **Ground truth:** 10,080

- ✖ **Baseline (vanilla RAG) (FAIL):** Total inventories as of January 26, 2025 were **\$10.080 billion**.
- ✖ **VERITROOPER (FAIL):** Total inventories as of January 26, 2025 were **\$10.08 billion**.

119. FAIL · Precision · `SEC\_10K\_multi\_1241\_01`

UNSUPPORTED NUMBER

- **Question:** What is the amount of amortization losses from investments in 2023?
- **Ground truth:** zero
- ✖ **Baseline (vanilla RAG) (FAIL):** Amortization losses from investments were **\$1.561 billion in 2023**.
- ✓ **VERITROOPER (PASS):** The amortization losses from investments in 2023 were zero.

120. PASS · Precision · `SEC\_10K\_multi\_0766\_00`

- **Question:** What is the interest rate spread volatility in bps?
- **Ground truth:** 49
- ✓ **Baseline (vanilla RAG) (PASS):** The interest rate spread volatility ranges from **44 bps to 59 bps**, with a weighted average of **49 bps**.
- ✓ **VERITROOPER (PASS):** The interest rate spread volatility ranges from 44 bps to 59 bps, with a weighted average of 49 bps.

121. PASS · Cause & Effect · `SEC\_10K\_multi\_0886\_00`

- **Question:** What happens as a result of the fair value option being elected for loans?
- **Ground truth:** Changes in value are attributed to instrument-specific credit risk.
- ✓ **Baseline (vanilla RAG) (PASS):** When the fair value option is elected, loans are measured at fair value and changes in fair value are recognized in income, while contractual interest is recorded as interest income. For floating-rate loans, all value changes are attributed to instrument-specific credit risk; for fixed-rate loans, changes are allocated between interest-rate and credit-related effects.
- ✓ **VERITROOPER (PASS):** Loans are measured at fair value, with contractual interest recorded separately in interest income or expense. Fair-value changes are attributed to instrument-specific credit risk for floating-rate loans and allocated between interest-rate and credit effects for fixed-rate loans.

122. PASS · Conditional · `SEC\_10K\_multi\_1154\_01`

- **Question:** If a loan is classified as non-investment grade, what is its PD compared to non-criticized loans?
- **Ground truth:** higher
- ✓ **Baseline (vanilla RAG) (PASS):** Non-investment-grade loans may be either noncriticized or criticized. If criticized, the loan has a higher probability of default (PD) than a noncriticized loan.
- ✓ **VERITROOPER (PASS):** Non-investment-grade status alone does not determine this because such loans may be noncriticized or criticized; criticized loans have a higher PD than noncriticized loans.

123. PASS · Precision · `SEC\_10K\_multi\_1611\_00`

- **Question:** What is the date of the adoption of the 2024 Equity Plan?
- **Ground truth:** May 1, 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide an adoption date for a "2024 Equity Plan."
- ✓ **VERITROOPER (PASS):** The 2024 Equity Plan was adopted on May 1, 2024.

124. PASS · Precision · `SEC\_10K\_multi\_0767\_00`

- **Question:** What is the percentage for Equity volatility?
- **Ground truth:** 4
- ✓ **Baseline (vanilla RAG) (PASS):** For net equity derivatives, Equity volatility ranged from **4% to 130%**, with a weighted average of **32%**.
- ✓ **VERITROOPER (PASS):** Equity volatility percentages are 4%, 130%, and 32% across the three columns.

125. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0159\_00`

- **Question:** What is the total number of shares available for issuance under JPMorganChase's employee share-based incentive plans as of December 31, 2024?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide the number as of December 31, 2024. It reports **77,131,051 shares** available for future issuance as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total number of shares available for issuance as of December 31, 2024.

126. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0042\_04`

- **Question:** What is the specific date by which the company expects to return to competing in China's data center market?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The company does not provide a specific return date. It states that returning depends on developing a competitive product approved by both the U.S. and Chinese governments.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a date for returning to competition in China's data center market.

127. PASS · Exception · `SEC\_10K\_multi\_1071\_00`

- **Question:** Under what circumstance does the general rule NOT apply to the recognition of compensation expense for RSUs and PSUs?
- **Ground truth:** Employees who will become full-career eligible during the vesting period
- ✓ **Baseline (vanilla RAG) (PASS):** The general grant-to-vesting recognition rule does not apply when an employee becomes full-career eligible during the vesting period. In that case, expense is recognized through the earlier of the full-career eligibility date or the tranche's vesting date. Awards with no future substantive service requirement are accrued as of the grant date.
- ✓ **VERITROOPER (PASS):** The general rule does not apply when employees will become full-career eligible during the vesting period or when awards have no future substantive service requirement.

128. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0728\_03`

- **Question:** What is the maximum allowed adjustment for company-specific issues in private equity valuations?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a maximum adjustment. It only states that adjustments may be made for company-specific issues, including lack of liquidity.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a maximum allowed adjustment for company-specific issues in private equity valuations.

129. PASS · Precision · `SEC\_10K\_multi\_0646\_01`

- **Question:** What is the credit ratio of 1.05?
- **Ground truth:** 1.05
- ✓ **Baseline (vanilla RAG) (PASS):** The **1.05% credit ratio** is the **90+ day delinquency rate for Card Services in 2023**.
- ✓ **VERITROOPER (PASS):** The 1.05% credit ratio is the allowance for loan losses to retained loans (Column 3).

130. PASS · Conditional · `SEC\_10K\_multi\_1349\_01`

- **Question:** If the Firm adopted updates to the Accounting for Investments in Tax Credit Structures Using the Proportional Amortization Method guidance under the modified retrospective method, what method was used for the adoption?
- **Ground truth:** modified retrospective method
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm used the **modified retrospective method**.
- ✓ **VERITROOPER (PASS):** The Firm used the modified retrospective method.

131. FAIL · Cross-Reference · `SEC\_10K\_multi\_0318\_01`

BAD QUESTION

- **Question:** How does the bank's revenue from card income compare to its total net revenue in 2025?
- **Ground truth:** The bank's revenue from card income was \$2,436 million, which is part of the total net revenue of \$53,766 million in 2025.
- ✗ **Baseline (vanilla RAG) (FAIL):** In 2025, Consumer & Community Banking generated **\$2.23 billion in card income**, compared with **\$76.03 billion in total net revenue**. Card income represented approximately 2.9% of total net revenue.
- ✗ **VERITROOPER (FAIL):** In 2025, card income was \$2.436 billion, compared with total net revenue of \$78.454 billion, or approximately 3.1% of the total.

132. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1279\_03`

- **Question:** How many days does the Firm have to stop payment to investors after an advance?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a number of days. It only states that the Firm may stop payment if the collateral is insufficient to cover the advance.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many days the Firm has to stop payment to investors after an advance.

133. PASS · Precision · `SEC\_10K\_multi\_0062\_01`

- **Question:** What is the number of the Table of Contents?
- **Ground truth:** 47
- ✓ **Baseline (vanilla RAG) (PASS):** 5
- ✓ **VERITROOPER (PASS):** 47

134. FAIL · Precision · `SEC\_10K\_multi\_0725\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total assets at December 31, 2025?
- **Ground truth:** 43,295

- ✖ **Baseline (vanilla RAG) (FAIL):** Total assets at December 31, 2025 were **\$4,424.9 billion**.
- ✔ **VERITROOPER (PASS):** Total assets at December 31, 2025 were \$43,295 million.

135. PASS · Precision · `SEC\_10K\_multi\_1347\_01`

- **Question:** What was the net pre-tax loss reclassified in 2023?
- **Ground truth:** -3 million
- ✔ **Baseline (vanilla RAG) (PASS):** The net pre-tax loss reclassified in 2023 was **\$35 million**.
- ✔ **VERITROOPER (PASS):** The net pre-tax loss reclassified in 2023 was \$3 million.

136. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0852\_00`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of level 3 assets measured at fair value on a recurring basis as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✖ **Baseline (vanilla RAG) (FAIL):** The total value was **\$25.107 billion** as of December 31, 2025.
- ✔ **VERITROOPER (PASS):** The evidence does not provide the total value of Level 3 assets measured on a recurring basis as of December 31, 2025.

137. PASS · Exception · `SEC\_10K\_multi\_1393\_00`

- **Question:** What does not apply to the calculation of adjusted average assets?
- **Ground truth:** Tier 1 capital deductions for goodwill and other intangible assets
- ✔ **Baseline (vanilla RAG) (PASS):** Off-balance-sheet exposures do not apply to the calculation of adjusted average assets; they are added separately to determine total leverage exposure.
- ✔ **VERITROOPER (PASS):** Off-balance-sheet lending-related financial instruments, guarantees, and other commitments do not apply to the calculation of adjusted average assets.

138. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_07`

- **Question:** How many countries have enacted data localization laws?
- **Ground truth:** UNANSWERABLE
- ✔ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a number of countries. It only states that "some countries" have enacted data localization or residency laws.
- ✔ **VERITROOPER (PASS):** The evidence does not specify how many countries have enacted data localization laws.

139. PASS · Cross-Reference · `SEC\_10K\_multi\_0716\_00`

- **Question:** What is the basis on which the guidance may be applied, and what is the planned date of adoption?
- **Ground truth:** The guidance may be applied on a prospective, modified, or retrospective transition basis, and the Firm's planned date of adoption is not specified.
- ✔ **Baseline (vanilla RAG) (PASS):** The guidance may be applied on a **prospective, modified, or retrospective transition basis**. The Firm is still evaluating its planned adoption date; the required effective date is **January 1, 2028**, with early adoption permitted.
- ✔ **VERITROOPER (PASS):** The guidance may be applied on a prospective, modified, or retrospective transition basis. The Firm is still evaluating its planned adoption date; the required effective date is January 1, 2027.

140. PASS · Precision · `SEC\_10K\_multi\_0253\_01`

- **Question:** What is the net addition to the allowance for lending-related commitments recognized in other liabilities?
- **Ground truth:** \$3.0 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The net addition was **\$3.0 billion**.
- ✓ **VERITROOPER (PASS):** The net addition was **\$3.0 billion**.

141. PASS · Cross-Reference · `SEC\_10K\_multi\_1093\_00`

- **Question:** What is excluded from the effective yield calculation for average yield computation?
- **Ground truth:** unscheduled principal prepayments
- ✓ **Baseline (vanilla RAG) (PASS):** The effective yield excludes unscheduled principal prepayments.
- ✓ **VERITROOPER (PASS):** Unscheduled principal prepayments are excluded from the effective yield calculation.

142. PASS · Precision · `SEC\_10K\_multi\_1537\_00`

- **Question:** What is the number of the section that discusses risks related to consumer demand for our products?
- **Ground truth:** 15
- ✓ **Baseline (vanilla RAG) (PASS):** Section 15.
- ✓ **VERITROOPER (PASS):** Section 16.

143. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1255\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total number of loans sold to the U.S. GSEs in 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** In 2025, the carrying value of loans sold to U.S. GSEs was **\$30.496 billion**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the number of loans sold to U.S. GSEs in 2025; it only reports their carrying value of \$30.496 billion.

144. FAIL · Precision · `SEC\_10K\_multi\_1125\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total retained loans for 2024?
- **Ground truth:** 376,334
- ✗ **Baseline (vanilla RAG) (FAIL):** Total retained loans for 2024 were **\$690.396 billion (\$690,396 million)**.
- ✗ **VERITROOPER (FAIL):** The total retained loans for 2024 were **\$376.334 billion**.

145. PASS · Precision · `SEC\_10K\_multi\_0139\_01`

- **Question:** What is the total for Asia-Pacific, Latin America and Canada?
- **Ground truth:** 11.6
- ✓ **Baseline (vanilla RAG) (PASS):** The total for Asia-Pacific, Latin America and Canada was **11.6 million square feet** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total for Asia-Pacific, Latin America and Canada is **11.6**.

146. PASS · Exception · `SEC\_10K\_multi\_1534\_01`

- **Question:** What is excluded from the Proposition 65 warning requirement?
- **Ground truth:** the result of necessary cooking
- ✓ **Baseline (vanilla RAG) (PASS):** Proposition 65 exempts a product from warning requirements if the manufacturer demonstrates that consumer exposure to the listed substance is:
 - Below an established "safe harbor" threshold;
 - Naturally occurring;
 - The result of necessary cooking; or
 - Subject to another applicable exemption.
- ✓ **VERITROOPER (PASS):** A product is exempt if exposure to the listed substance is below a safe-harbor threshold, naturally occurring, results from necessary cooking, or qualifies for another applicable exemption.

147. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1169\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total number of loans modified in 2023?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** In 2023, total retained credit card loan modifications were **\$648 million** in amortized cost basis.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total number of loans modified in 2023; it reports only their amortized cost basis.

148. FAIL · Precision · `SEC\_10K\_multi\_0744\_01`

BAD QUESTION

- **Question:** What is the total liabilities measured at fair value on a recurring basis?
- **Ground truth:** 550,409
- ✗ **Baseline (vanilla RAG) (FAIL):** Total liabilities measured at fair value on a recurring basis were **\$586.175 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total liabilities measured at fair value on a recurring basis were **\$770,827 million**.

149. PASS · Cause & Effect · `SEC\_10K\_multi\_0924\_01`

- **Question:** What happens if the hedge relationship is terminated?
- **Ground truth:** The adjustment to the hedged item continues to be reported as part of the basis of the hedged item and, for interest-bearing financial instruments, is amortized to earnings as a yield adjustment.
- ✓ **Baseline (vanilla RAG) (PASS):** If a hedge relationship is terminated:
 - **Fair value hedge:** The adjustment remains part of the hedged item's basis and, for interest-bearing instruments, is amortized into earnings as a yield adjustment.
 - **Cash flow hedge:** The derivative amount in AOCI is recognized in earnings when the hedged cash flows affect earnings. If the forecasted transaction is no longer expected to occur, the related AOCI amount is recognized immediately in earnings.
- ✓ **VERITROOPER (PASS):** For a fair value hedge, the hedged-item adjustment remains in its basis and, for interest-bearing instruments, is amortized to earnings as a yield adjustment. For a cash flow hedge, amounts in AOCI are recognized in earnings when the hedged cash flows affect earnings, or immediately if the forecasted transaction is no longer expected to occur.

150. PASS · Precision · `SEC\_10K\_multi\_0720\_00`

- **Question:** What year is the end of the period for the financial statements?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The financial statements cover the period ended **December 31, 2025**.
- ✓ **VERITROOPER (PASS):** The period ended in 2025.

151. PASS · Precision · `SEC\_10K\_multi\_0719\_01`

- **Question:** What is the date on which the report of the Independent Registered Public Accounting Firm was issued?
- **Ground truth:** February 13, 2026
- ✓ **Baseline (vanilla RAG) (PASS):** February 20, 2026.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the issuance date of the Independent Registered Public Accounting Firm's report.

152. PASS · Precision · `SEC\_10K\_multi\_0108\_00`

- **Question:** What is the filing date of the report?
- **Ground truth:** February 25, 2026
- ✓ **Baseline (vanilla RAG) (PASS):** The report was filed on October 31, 2025.
- ✓ **VERITROOPER (PASS):** The report was filed on February 13, 2026.

153. PASS · Precision · `SEC\_10K\_multi\_1492\_01`

- **Question:** What is the value of 3,749,498?
- **Ground truth:** 3,749,498
- ✓ **Baseline (vanilla RAG) (PASS):** \$3,749,498 million, representing total liabilities for 2024.
- ✓ **VERITROOPER (PASS):** 3,749,498.

154. PASS · Conditional · `SEC\_10K\_multi\_0128\_00`

- **Question:** If JPMorganChase is subject to extensive supervision and regulation, what could happen to its business?
- **Ground truth:** It could be negatively affected.
- ✓ **Baseline (vanilla RAG) (PASS):** Extensive supervision and regulation could require JPMorganChase to modify or restrict its operations, limit products and services, raise prices, hold more capital or liquidity, dispose of assets, and forgo opportunities. This could increase compliance and operational costs, reduce revenues and competitiveness, and expose it to greater scrutiny, enforcement actions, penalties, litigation, and reputational harm for noncompliance.
- ✓ **VERITROOPER (PASS):** It could be required to modify its business and operations, incur higher costs or reduced revenues, and face scrutiny, enforcement actions, penalties, litigation, or reputational harm for non-compliance.

155. PASS · Conditional · `SEC\_10K\_multi\_0898\_00`

- **Question:** If the fair value option was elected for lending-related commitments at December 31, 2025, what was the corresponding fair value?
- **Ground truth:** \$42 million
- ✓ **Baseline (vanilla RAG) (PASS):** The corresponding fair value was **\$42 million** at December 31, 2025.
- ✓ **VERITROOPER (PASS):** The corresponding fair value was **\$42 million**.

156. PASS · Precision · `SEC\_10K\_multi\_1539\_00`

- **Question:** What is the threshold for Pillar One?
- **Ground truth:** €20 billion
- ✓ **Baseline (vanilla RAG) (PASS):** Pillar One applies to multinational businesses with annual global revenue exceeding **€20 billion** and a **profit margin above 10%**.
- ✓ **VERITROOPER (PASS):** Pillar One applies to multinational businesses with annual global revenue exceeding €20 billion and a profit margin above 10%.

157. FAIL · Precision · `SEC\_10K\_multi\_0494\_01`

VALIDATOR OVER-FIRED

- **Question:** What was the Tier 1 capital as of December 31, 2025?
- **Ground truth:** 307,630
- ✗ **Baseline (vanilla RAG) (FAIL):** Tier 1 capital was **\$307.63 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** Tier 1 capital was **\$307,630 million** as of December 31, 2025.

158. PASS · Exception · `SEC\_10K\_multi\_0677\_01`

- **Question:** What circumstance does the general rule of stress testing NOT apply to?
- **Ground truth:** tail events in specific asset classes and geographies
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not identify an exception to a general rule of stress testing. It does state that the usual **consolidation** rule-majority voting ownership-does not apply when an entity has a variable interest created through nonvoting arrangements, resulting in disproportionate economic risk or reward exposure.
- ✓ **VERITROOPER (PASS):** The evidence does not identify any circumstance in which the general rule of stress testing does not apply.

159. PASS · Precision · `SEC\_10K\_multi\_0705\_00`

- **Question:** What is the confidence level for the Loss Distribution Approach (LDA) model?
- **Ground truth:** 99.9%
- ✓ **Baseline (vanilla RAG) (PASS):** The LDA model uses a **99.9% confidence level** over a one-year time horizon.
- ✓ **VERITROOPER (PASS):** The LDA model uses a 99.9% confidence level.

160. PASS · Conditional · `SEC\_10K\_multi\_0866\_00`

- **Question:** If the Firm elects to measure certain instruments at fair value, what is one reason for this election?
- **Ground truth:** to mitigate income statement volatility
- ✓ **Baseline (vanilla RAG) (PASS):** One reason is to reduce income statement volatility caused by differences between an instrument's measurement basis and its associated risk-management arrangements.
- ✓ **VERITROOPER (PASS):** One reason is to mitigate income statement volatility caused by differences between the instruments' measurement basis and their associated risk management arrangements.

161. FAIL · Precision · `SEC\_10K\_multi\_1341\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the net unrealized gain/(loss) arising during the period for 2023?
- **Ground truth:** 3,891

- **✗ Baseline (vanilla RAG) (FAIL):** For 2023, the net unrealized gain arising during the period was **\$3.891 billion pre-tax**, or **\$2.969 billion after tax**.
- **✓ VERITROOPER (PASS):** For 2023, the net unrealized gain was \$3,891 million pre-tax and \$2,969 million after-tax.

162. PASS · Cause & Effect · `SEC\_10K\_multi\_0886\_01`

- **Question:** What is the consequence of the requirement for collateral in securities financing agreements?
- **Ground truth:** There would be no adjustment or an immaterial adjustment for instrument-specific credit risk.
- **✓ Baseline (vanilla RAG) (PASS):** Collateral requirements reduce counterparty credit exposure. The Firm monitors collateral values, requests additional collateral or returns excess collateral as values change, and may liquidate or set off collateral upon default. Exchanged collateral generally does not qualify for net presentation on the balance sheet.
- **✓ VERITROOPER (PASS):** Because collateral must have a market value equal to or exceeding the principal loaned, any instrument-specific credit risk adjustment is generally nonexistent or immaterial.

163. PASS · Cross-Reference · `SEC\_10K\_multi\_1531\_00`

- **Question:** What percentage of the Company's worldwide unit case volume in 2025 was attributable to sparkling soft drinks?
- **Ground truth:** 69%
- **✓ Baseline (vanilla RAG) (PASS):** 69% of the Company's worldwide unit case volume in 2025 was attributable to sparkling soft drinks.
- **✓ VERITROOPER (PASS):** 69%.

164. PASS · Cross-Reference · `SEC\_10K\_multi\_1530\_00`

- **Question:** What is the aggregate market value of the common equity held by non-affiliates of the Registrant as of June 27, 2025?
- **Ground truth:** \$300,552,956,545
- **✓ Baseline (vanilla RAG) (PASS):** The aggregate market value was **\$300,552,956,545**.
- **✓ VERITROOPER (PASS):** The aggregate market value was **\$300,552,956,545** as of June 27, 2025.

165. PASS · Cross-Reference · `SEC\_10K\_multi\_0235\_01`

- **Question:** What is the expected net charge-off rate in Card Services for 2026, according to the outlook information provided on January 13, 2026?
- **Ground truth:** approximately 3.4%
- **✓ Baseline (vanilla RAG) (PASS):** The expected 2026 Card Services net charge-off rate is approximately **3.4%**.
- **✓ VERITROOPER (PASS):** The expected net charge-off rate in Card Services for 2026 is approximately 3.4%.

166. PASS · Cross-Reference · `SEC\_10K\_multi\_1171\_01`

- **Question:** What was the weighted-average reduction in contractual interest for loans with multiple modifications, and how many months were extended?
- **Ground truth:** 1060 bps, 16 months
- **✓ Baseline (vanilla RAG) (PASS):** The weighted-average contractual interest rate was reduced by **1,060 basis points**, and the loans were extended by a weighted average of **16 months**.
- **✓ VERITROOPER (PASS):** For loans with payment deferral, interest-rate reduction, and term extension, contractual interest was reduced by 1,060 bps and terms extended by 16 months. For interest-rate reduction and term extension only, the figures were 672 bps and 15 months.

167. PASS · Cross-Reference · `SEC\_10K\_multi\_0956\_01`

- **Question:** How do gains/(losses) on hedging derivatives relate to the Consolidated statements of income?
- **Ground truth:** The Firm includes gains/(losses) on the hedging derivative in the same line item in the Consolidated statements of income as the related hedged item.
- ✓ **Baseline (vanilla RAG) (PASS):** Gains/(losses) on hedging derivatives are recorded in the same line item in the Consolidated statements of income as the related hedged item-or, for cash flow hedges, as the change in cash flows on that hedged item. Amounts initially recorded in AOCI are subsequently reclassified into income as the hedged cash flows affect earnings.
- ✓ **VERITROOPER (PASS):** Gains/(losses) on hedging derivatives are included in the same line item in the Consolidated statements of income as the related hedged item.

168. PASS · Calculation · `SEC\_10K\_multi\_0040\_01`

- **Question:** What is the total amount of money invested by NVIDIA in research and development since its inception?
- **Ground truth:** 76.7 billion
- ✓ **Baseline (vanilla RAG) (PASS):** 76.7 billion
- ✓ **VERITROOPER (PASS):** \$76.7 billion

169. PASS · Cross-Reference · `SEC\_10K\_multi\_0543\_01`

- **Question:** How does the Firm's credit risk management framework address potential credit losses?
- **Ground truth:** The Firm estimates credit losses for its exposures based on factors such as the probability of default, loss severity, and exposure at default, and reflects these estimates in the allowance for loan losses and lending-related commitments.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm addresses potential credit losses by:
 - Estimating losses based on probability of default, loss severity given default, and exposure at default.
 - Maintaining allowances for loan losses, lending-related commitments, and investment securities.
 - Allocating credit risk capital for potential and unexpected losses beyond established allowances.
 - Stress testing portfolios under centrally defined alternative macroeconomic and business scenarios to identify credit migration, delinquency trends, and potential losses, supplemented by industry- and country-specific stresses.
- Using stress results to set Firmwide and line-of-business risk appetite and assess individual counterparties.
- Regularly monitoring credit risk at transaction and portfolio levels through approval authorities, concentration limits, risk ratings, portfolio reviews, and distressed-exposure guidelines, with independent validation of relevant models and assumptions.
- ✓ **VERITROOPER (PASS):** The Firm estimates credit losses using probability of default, loss severity and exposure at default, and reflects them in allowances for loans, lending-related commitments and investment securities. It also allocates credit risk capital for potential unexpected losses and uses stress testing under alternative economic scenarios to assess impacts and inform risk appetite.

170. FAIL · Cross-Reference · `SEC\_10K\_multi\_0080\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total amount of current income taxes for the year ended January 25, 2026, and how does it compare to the year ended January 26, 2025?
- **Ground truth:** The total current income taxes for the year ended January 25, 2026, is \$22,807, which is \$7,184 more than the \$15,623 for the year ended January 26, 2025.

- **✗ Baseline (vanilla RAG) (FAIL):** Total current income taxes were **\$22.807 billion** for the year ended January 25, 2026, compared with **\$15.623 billion** for the year ended January 26, 2025. This was an **increase of \$7.184 billion**, or approximately **46%**.
- **✗ VERITROOPER (FAIL):** Total current income taxes were \$22.807 billion for 2026, compared with \$15.623 billion for 2025-an increase of \$7.184 billion, or approximately 46%.

171. PASS · Cause & Effect · `SEC\_10K\_multi\_1567\_00`

- **Question:** What is the impact of a stronger U.S. dollar on the Company's Latin America operating segment?
- **Ground truth:** unfavorable
- **✓ Baseline (vanilla RAG) (PASS):** A stronger U.S. dollar-particularly against the Mexican peso, Argentine peso, and Brazilian real-negatively affected Latin America. Foreign currency exchange rates reduced the segment's 2025 operating income by 21% and also unfavorably impacted net operating revenues.
- **✓ VERITROOPER (PASS):** A stronger U.S. dollar adversely affected the Latin America segment, contributing to an unfavorable foreign currency exchange rate impact of 21% on operating income in 2025.

172. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0767\_00`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average price of oil in BBL?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The average oil price is **\$202 per barrel (BBL)**.
- **✗ VERITROOPER (FAIL):** The average oil commodity forward price was **\$202 per BBL**.

173. FAIL · Cross-Reference · `SEC\_10K\_multi\_1155\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the total noninvestment-grade loans for 2025 and 2024, and what percentage of total retained loans are criticized nonaccrual loans for each year?
- **Ground truth:** The total noninvestment-grade loans for 2025 is \$46,670 and for 2024 is \$48,152. The percentage of total retained loans that are criticized nonaccrual loans for 2025 is 1.01% and for 2024 is 0.89%.
- **✗ Baseline (vanilla RAG) (FAIL):** - **2025:** Total noninvestment-grade loans were **\$251.0 billion**; criticized nonaccrual loans were **0.56%** of total retained loans.
- **2024:** Total noninvestment-grade loans were **\$218.7 billion**; criticized nonaccrual loans were **0.57%** of total retained loans.
- **✗ VERITROOPER (FAIL):** Total noninvestment-grade loans were \$251,003 in 2025 and \$218,726 in 2024. Criticized nonaccrual loans represented 0.6% of total retained loans in both years.

174. PASS · Precision · `SEC\_10K\_multi\_1320\_01`

- **Question:** What is the fixed rate for Series MM?
- **Ground truth:** 420.00
- **✓ Baseline (vanilla RAG) (PASS):** The fixed rate for Series MM is **4.20%**.
- **✓ VERITROOPER (PASS):** The fixed rate for Series MM is **4.200%**.

175. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_05`

- **Question:** How many natural disasters has Apple experienced in the past five years?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify how many natural disasters Apple experienced in the past five years.

- ✓ **VERITROOPER (PASS)**: The evidence does not state how many natural disasters Apple experienced in the past five years.

176. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question**: What is the total value of the notes issued by Apple Inc. in 2025?
- **Ground truth**: UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL)**: Apple issued **\$4.5 billion** in notes in 2025.
- ✓ **VERITROOPER (PASS)**: The evidence does not provide the total value of notes issued by Apple Inc. in 2025.

177. FAIL · Precision · `SEC\_10K\_multi\_0520\_00`

UNIT/FORMAT MISMATCH

- **Question**: What is the total short-term unsecured funding for 2024?
- **Ground truth**: 24,985
- ✗ **Baseline (vanilla RAG) (FAIL)**: The total short-term unsecured funding for 2024 was **\$28.517 billion**.
- ✓ **VERITROOPER (PASS)**: The total short-term unsecured funding for 2024 was **\$28,517 million**.

178. PASS · Calculation · `SEC\_10K\_multi\_0929\_00`

- **Question**: If the gross long notional derivative contracts are \$100 and the gross short notional derivative contracts are \$200, what is the total?
- **Ground truth**: 300
- ✓ **Baseline (vanilla RAG) (PASS)**: 300
- ✓ **VERITROOPER (PASS)**: 300

179. FAIL · Cross-Reference · `SEC\_10K\_multi\_1409\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question**: What was the total amount for other letters of credit as of December 31, 2025?
- **Ground truth**: \$514 million
- ✗ **Baseline (vanilla RAG) (FAIL)**: The total contractual amount for other letters of credit was **\$4.529 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS)**: The total amount for other letters of credit as of December 31, 2025 was **\$514 million**.

180. PASS · Precision · `SEC\_10K\_multi\_0574\_00`

- **Question**: What is the carrying value of retained home equity lines of credit outstanding at December 31, 2025?
- **Ground truth**: 13.2 billion
- ✓ **Baseline (vanilla RAG) (PASS)**: The carrying value was **\$13.2 billion** at December 31, 2025.
- ✓ **VERITROOPER (PASS)**: The carrying value was **\$13.2 billion** at December 31, 2025.

181. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1595\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question**: What is the total value of the Company's other postretirement benefit plans' assets as of December 31, 2025?

- **Ground truth:** UNANSWERABLE

- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, the other postretirement benefit plans' assets totaled \$67 million.

- **✗ VERITROOPER (FAIL):** The fair value of the Company's other postretirement benefit plan assets was \$67 million as of December 31, 2025.

182. PASS · Cause & Effect · `SEC\_10K\_multi\_0083\_01`

- **Question:** What happens as a consequence of the realization of deferred tax assets becoming more likely than not?

- **Ground truth:** We would recognize such deferred tax assets as income tax benefits during the period.

- **✓ Baseline (vanilla RAG) (PASS):** The company would recognize the deferred tax assets as income tax benefits in that period, reversing any related valuation allowance as appropriate.

- **✓ VERITROOPER (PASS):** The deferred tax assets would be recognized as income tax benefits during the period.

183. PASS · Cause & Effect · `SEC\_10K\_multi\_0134\_00`

- **Question:** What happens as a consequence of a cybersecurity breach?

- **Ground truth:** significant disruption of or loss of access to JPMorganChase's operational systems and those of its clients, customers and counterparties

- **✓ Baseline (vanilla RAG) (PASS):** A cybersecurity breach can cause:

- Disruption, damage, or loss of access to operational systems and services

- Theft, manipulation, destruction, or loss of confidential data

- Fraudulent transactions, financial losses, or ransom demands

- Delayed or incomplete recovery of affected systems and data

- Privacy or other legal violations

- Higher operational and remediation costs

- Litigation, regulatory investigations, and penalties

- Reduced customer access, satisfaction, and confidence

- Reputational harm and loss of trust in cybersecurity and resilience measures

- Wider disruptions affecting interconnected clients, vendors, counterparties, and financial markets

- **✓ VERITROOPER (PASS):** A cybersecurity breach could disrupt systems, compromise confidential information, delay data recovery, enable fraudulent transactions, trigger ransom demands, and cause financial losses. It could also lead to legal and regulatory penalties, loss of confidence, and reputational harm.

184. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1531\_00`

- **Question:** What is the total number of employees in the Company's bottling and distribution operations?

- **Ground truth:** UNANSWERABLE

- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify the number of employees in bottling and distribution operations. It only reports approximately **65,900 total Company employees** as of December 31, 2025.

- **✓ VERITROOPER (PASS):** The evidence does not state the total number of employees in the Company's bottling and distribution operations.

185. PASS · Cause & Effect · `SEC\_10K\_multi\_1202\_00`

- **Question:** What is the consequence of the Apple Card transaction on the allowance for credit losses?

- **Ground truth:** a net addition of \$2.2 billion

- ✓ **Baseline (vanilla RAG) (PASS):** The Apple Card transaction increased JPMorgan Chase's allowance for credit losses by approximately **\$2.2 billion** in 2025, recorded within the consumer allowance, primarily as an allowance for lending-related commitments.
- ✓ **VERITROOPER (PASS):** The Apple Card transaction drove a \$2.2 billion increase in the consumer allowance for credit losses in 2025.

186. PASS · Cause & Effect · `SEC\_10K\_multi\_1208\_01`

- **Question:** What happens as a consequence of the Firm retaining an undivided seller's interest in the receivables?
- **Ground truth:** The Firm consolidates the assets and liabilities of its sponsored credit card trusts.
- ✓ **Baseline (vanilla RAG) (PASS):** The retained undivided seller's interest is eliminated in consolidation. It also contributes to the Firm's continuing involvement and potential exposure to significant benefits and losses from the credit card securitization trusts.
- ✓ **VERITROOPER (PASS):** Retaining an undivided seller's interest obligates the Firm to absorb losses and gives it the right to receive potentially significant benefits from the securitization trusts.

187. FAIL · Cross-Reference · `SEC\_10K\_multi\_0368\_00`

VALIDATOR OVER-FIRED

- **Question:** What were the total assets under custody in Fixed Income and Equity Markets as of December 31, 2025?
- **Ground truth:** 18,322 billion and 17,954 billion
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, assets under custody were:
- **Fixed Income:** \$18.322 trillion
- **Equity:** \$17.954 trillion
- **Combined total:** **\$36.276 trillion**
- ✗ **VERITROOPER (FAIL):** Total assets under custody were **\$36,276 billion** as of December 31, 2025 (\$18,322 billion Fixed Income plus \$17,954 billion Equity).

188. PASS · Cause & Effect · `SEC\_10K\_multi\_0535\_01`

- **Question:** What happens as a consequence of net issuances of structured notes in Markets due to client demand and an increase in the fair value of such instruments?
- **Ground truth:** The increase in unsecured other borrowed funds for the average year ended December 31, 2025, compared to the prior year.
- ✓ **Baseline (vanilla RAG) (PASS):** Long-term debt increases as a result of the net issuances of structured notes and the rise in their fair value.
- ✓ **VERITROOPER (PASS):** Unsecured other borrowed funds and structured notes increased for the average year ended December 31, 2025, compared with the prior year.

189. PASS · Conditional · `SEC\_10K\_multi\_0509\_00`

- **Question:** If the Firm's liquidity risk management function is responsible for defining and monitoring liquidity risk metrics, does it also independently establish and monitor limits and indicators?
- **Ground truth:** Yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. The Liquidity Risk Management function independently establishes and monitors limits and indicators, including the Firm's liquidity risk appetite.
- ✓ **VERITROOPER (PASS):** Yes. The Liquidity Risk Management function independently establishes and monitors limits and indicators, including liquidity risk appetite.

190. PASS · Conditional · `SEC\_10K\_multi\_0717\_01`

- **Question:** If the hedging instrument is highly effective against each hedged risk in the portfolio, can different risks be pooled in the same portfolio for cash flow hedging?
- **Ground truth:** yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. The amended hedge accounting guidance permits different risks to be pooled in the same cash flow hedging portfolio when the hedging instrument is highly effective against each hedged risk.
- ✓ **VERITROOPER (PASS):** Yes. Different risks can be pooled in the same portfolio for cash flow hedging if the hedging instrument is highly effective against each hedged risk.

191. PASS · Cause & Effect · `SEC\_10K\_multi\_0086\_00`

- **Question:** What happens as a consequence of higher U.S.-based Compute & Networking segment demand?
- **Ground truth:** The increase in revenue to the United States for fiscal years 2026 and 2025.
- ✓ **Baseline (vanilla RAG) (PASS):** It primarily increased revenue attributed to the United States in fiscal years 2026 and 2025.
- ✓ **VERITROOPER (PASS):** Higher U.S.-based Compute & Networking segment demand primarily increased NVIDIA's revenue from the United States in fiscal years 2026 and 2025.

192. FAIL · Cross-Reference · `SEC\_10K\_multi\_0320\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What was the return on equity for the years ended December 31, 2023, 2024, and 2025?
- **Ground truth:** 14%, 18%, and 18%
- ✗ **Baseline (vanilla RAG) (FAIL):** - **2023:** 38%
- **2024:** 32%
- **2025:** 32%
- ✓ **VERITROOPER (PASS):** Return on equity was 14% in 2023, 18% in 2024, and 18% in 2025.

193. PASS · Cross-Reference · `SEC\_10K\_multi\_0936\_01`

- **Question:** What is the net amount of derivative receivables as of December 31, 2025, for interest rate contracts that are exchange-traded?
- **Ground truth:** 179 million
- ✓ **Baseline (vanilla RAG) (PASS):** The net derivative receivables were **\$55 million** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the 2025 net derivative receivables amount for exchange-traded interest rate contracts.

194. PASS · Precision · `SEC\_10K\_multi\_1548\_00`

- **Question:** What is the percentage of ownership interest in BA Sports Nutrition, LLC that the Company acquired in November 2021?
- **Ground truth:** 85%
- ✓ **Baseline (vanilla RAG) (PASS):** The Company acquired the remaining **85% ownership interest** in BA Sports Nutrition, LLC in November 2021.
- ✓ **VERITROOPER (PASS):** The Company acquired the remaining 85% ownership interest in BA Sports Nutrition, LLC in November 2021.

195. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0109\_01`

- **Question:** What is the average salary of a JPMorgan Chase & Co. employee?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The average employee salary cannot be determined from the provided evidence. It does not include total employee salaries or the number of employees needed for the calculation.
- ✓ **VERITROOPER (PASS):** The evidence does not provide information about JPMorgan Chase & Co. employees' average salary.

196. PASS · Conditional · `SEC\_10K\_multi\_0778\_00`

- **Question:** If the yield increases, what effect does it have on a fair value measurement?
- **Ground truth:** It results in a decrease.
- ✓ **Baseline (vanilla RAG) (PASS):** An increase in yield, in isolation, decreases the fair value measurement.
- ✓ **VERITROOPER (PASS):** An increase in yield, in isolation, would decrease the fair value measurement.

197. PASS · Precision · `SEC\_10K\_multi\_0719\_00`

- **Question:** What is the year in which JPMorganChase's internal control over financial reporting was effective?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** 2025.
- ✓ **VERITROOPER (PASS):** 2025.

198. PASS · Conditional · `SEC\_10K\_multi\_1604\_00`

- **Question:** If the net change in operating assets and liabilities in 2024 was a decrease, what was the net change in trade accounts receivable?
- **Ground truth:** A decrease of \$295 million.
- ✓ **Baseline (vanilla RAG) (PASS):** The net change in trade accounts receivable was a **\$295 million decrease** in 2024.
- ✓ **VERITROOPER (PASS):** The net change in trade accounts receivable in 2024 was a decrease of \$295 million.

199. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1255\_00`

- **Question:** What is the total proceeds received from loan sales in 2026?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide 2026 data. Total proceeds received from loan sales were **\$30.354 billion in 2025**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide total proceeds received from loan sales for 2026.

200. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1095\_04`

- **Question:** How many securities financing agreements are carried at fair value as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify how many securities financing agreements were carried at fair value as of December 31, 2025; it only states that certain agreements were carried at fair value.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many securities financing agreements were carried at fair value as of December 31, 2025.

201. PASS · Exception · `SEC\_10K\_multi\_1049\_00`

- **Question:** What is excluded from fair value measurement for financial instruments?
- **Ground truth:** derivatives and certain financial instruments containing embedded derivatives
- ✓ **Baseline (vanilla RAG) (PASS):** Level 1 and Level 2 securities and derivative positions used to risk-manage the observable components of Level 3 instruments are excluded from the Level 3 fair value rollforward tables.
- ✓ **VERITROOPER (PASS):** Derivatives and certain financial instruments containing embedded derivatives are excluded from this interest accrual treatment for instruments measured at fair value.

202. PASS · Cross-Reference · `SEC\_10K\_multi\_0854\_01`

- **Question:** How does the Firm measure equity securities without readily determinable fair values?
- **Ground truth:** at cost less impairment, plus or minus observable price changes
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm uses the **measurement alternative**, measuring these equity securities at **cost less impairment, if any, plus or minus observable price changes from orderly transactions involving an identical or similar investment of the same issuer**. Any resulting changes are recognized in income.
- ✓ **VERITROOPER (PASS):** The Firm measures them at cost less impairment, if any, plus or minus observable price changes for an identical or similar investment of the same issuer, with changes recognized in other income.

203. FAIL · Precision · `SEC\_10K\_multi\_1455\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the net cash provided by/(used in) operating activities for 2024?
- **Ground truth:** -21,981
- ✗ **Baseline (vanilla RAG) (FAIL):** Net cash used in operating activities for 2024 was **\$42.012 billion** (reported as **\$(42,012) million**).
- ✓ **VERITROOPER (PASS):** The net cash used in operating activities for 2024 was **\$21,981 million**.

204. PASS · Precision · `SEC\_10K\_multi\_0699\_00`

- **Question:** What is the total exposure for Germany?
- **Ground truth:** \$83.9
- ✓ **Baseline (vanilla RAG) (PASS):** Germany's total exposure was **\$100.3 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** Germany's total exposure is **\$100.3**.

205. PASS · Precision · `SEC\_10K\_multi\_1607\_01`

- **Question:** What is the year for the Consolidated Balance Sheets?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The Consolidated Balance Sheets are for **December 31, 2025 and 2024**.
- ✓ **VERITROOPER (PASS):** The Consolidated Balance Sheets are for December 31, 2025 and 2024.

206. FAIL · Precision · `SEC\_10K\_multi\_0551\_00`

TABLE MISREAD

- **Question:** What was the total nonaccrual loans at December 31, 2025?
- **Ground truth:** 9,857

- ✖ **Baseline (vanilla RAG) (FAIL):** Total nonaccrual loans at December 31, 2025 were **\$5.159 billion**.
- ✖ **VERITROOPER (FAIL):** Total nonaccrual loans at December 31, 2025 were **\$9.857 billion**.

207. FAIL · Cross-Reference · `SEC\_10K\_multi\_0055\_00`

WRONG PERIOD

- **Question:** What is the total value of shares withheld for tax withholding obligations in fiscal year 2026, and how does it compare to the total value of shares repurchased during the same period?
- **Ground truth:** 51 million shares for \$7.9 billion, which is less than the \$1.5 billion spent on repurchasing 8 million shares.
- ✖ **Baseline (vanilla RAG) (FAIL):** In fiscal 2026, NVIDIA withheld shares valued at **\$7.948 billion** for tax obligations, compared with **\$40.388 billion** of shares repurchased. Thus, tax-withheld shares were **\$32.44 billion lower**, or about **20%** of the repurchase value.
- ✓ **VERITROOPER (PASS):** Shares withheld in fiscal 2026 totaled \$7.9 billion. The evidence does not provide fiscal 2026 repurchase value; it only states \$1.5 billion was repurchased from January 26 to February 20, 2026.

208. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_05`

- **Question:** What is the total amount of damages paid by the Company in patent infringement cases?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose any amount paid by the Company in patent infringement cases. It only states that plaintiffs seek substantial damages.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total amount of damages paid by the Company in patent infringement cases.

209. FAIL · Conditional · `SEC\_10K\_multi\_0942\_00`

UNIT/FORMAT MISMATCH

- **Question:** If the derivative receivables are not nettable on the Consolidated balance sheets, what is the total amount?
- **Ground truth:** 57,777
- ✖ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, collateral related to derivative receivables not nettable on the Consolidated balance sheets totaled **\$28.891 billion**.
- ✓ **VERITROOPER (PASS):** The total collateral not nettable on the Consolidated balance sheets was **\$28,891 million in 2025**.

210. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_01`

- **Question:** What is the average salary of a software engineer at NVIDIA?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not state the average salary of a software engineer at NVIDIA.
- ✓ **VERITROOPER (PASS):** The provided evidence does not state the average salary of a software engineer at NVIDIA.

211. PASS · Cause & Effect · `SEC\_10K\_multi\_0007\_00`

- **Question:** What happens as a result of the Company's failure to comply with evolving laws and regulations?
- **Ground truth:** Materially adverse effects on the Company's business, reputation, results of operations, financial condition, and stock price.

- ✓ **Baseline (vanilla RAG) (PASS):** Failure to comply can result in litigation, government investigations or enforcement actions, civil or criminal penalties and fines, corrective-action costs, production or distribution suspensions, operational interruptions, product bans or import delays, reputational damage, and material harm to the Company's business, financial condition, results of operations and stock price.
- ✓ **VERITROOPER (PASS):** Failure to comply can trigger legal and regulatory proceedings, significant penalties or liability, reputational harm, and material adverse effects on the Company's business, results of operations, financial condition, and stock price.

212. PASS · Precision · `SEC\_10K\_multi\_0985\_01`

- **Question:** What is the year for which the pre-tax gains/(losses) are recorded for the interest rate contract type?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The pre-tax gains/(losses) for the interest rate contract type are reported for the years ended December 31, **2025, 2024, and 2023**.
- ✓ **VERITROOPER (PASS):** The year is 2025.

213. PASS · Cross-Reference · `SEC\_10K\_multi\_0502\_00`

- **Question:** What factors influence management's decision on common share repurchases, and how does this relate to the Firm's capital position?
- **Ground truth:** Management determines the amount and timing of common share repurchases based on various factors, including market conditions, legal and regulatory considerations, the Firm's capital position, organic capital generation, and other investment opportunities.
- ✓ **Baseline (vanilla RAG) (PASS):** Management determines the amount and timing of common share repurchases based on:
 - Market conditions
 - Legal and regulatory considerations
 - The Firm's capital position, including goodwill and intangible assets
 - Organic capital generation
 - Current and proposed future capital requirements
 - Other investment opportunities

Thus, repurchases are discretionary and depend partly on whether the Firm has sufficient capital after considering existing and potential regulatory requirements. Repurchases may vary substantially or be suspended at any time; failure to meet minimum TLAC requirements and applicable buffers would also limit repurchases and other capital distributions.

- ✓ **VERITROOPER (PASS):** Management considers market conditions, legal and regulatory constraints, organic capital generation, current and future capital requirements, other investment opportunities, and the Firm's capital position. The capital-position assessment specifically takes goodwill and intangible assets into account.

214. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_08`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average interest rate of the notes issued by Apple Inc. in 2022?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The average interest rate of Apple's four notes issued in 2022 was **3.6625%**, or approximately **3.66%**.

$$\frac{(3.250\% + 3.350\% + 3.950\% + 4.100\%)}{4} = 3.6625\%$$

- ✖ **VERITROOPER (FAIL):** The average interest rate of Apple's notes issued in 2022 was **3.6625%**.

215. FAIL · Precision · `SEC\_10K\_multi\_0747\_00`

BAD QUESTION

- **Question:** What is the total value of equity securities?
- **Ground truth:** 417,428
- ✖ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, total equity securities were **\$2.633 billion** (\$2.589 billion at fair value plus \$44 million under the measurement alternative).
- ✖ **VERITROOPER (FAIL):** The total value of equity securities is **131,728**.

216. FAIL · Cross-Reference · `SEC\_10K\_multi\_1354\_00`

BAD QUESTION

- **Question:** What was the total income tax expense for the year ended December 31, 2025, and what portion of it was due to tax benefits?
- **Ground truth:** 15,547, 9,936
- ✖ **Baseline (vanilla RAG) (FAIL):** Total income tax expense was **\$15.547 billion** in 2025. It included **\$629 million of tax benefits** from resolved tax audits, equivalent to approximately **4.0%** of the total expense.
- ✖ **VERITROOPER (FAIL):** Total income tax expense was **\$15.547 billion** in 2025, including **\$629 million in tax benefits**, or approximately **4.0%** of the total.

217. PASS · Precision · `SEC\_10K\_multi\_0539\_01`

- **Question:** What is the rating agency mentioned in the text?
- **Ground truth:** Moody's Investors Service
- ✓ **Baseline (vanilla RAG) (PASS):** The rating agencies mentioned are Standard & Poor's and Moody's.
- ✓ **VERITROOPER (PASS):** The rating agency mentioned is Moody's Investors Service.

218. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1274\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of contractual service fees included in income for the year?
- **Ground truth:** UNANSWERABLE
- ✖ **Baseline (vanilla RAG) (FAIL):** Contractual service fees, late fees, and other ancillary fees included in income totaled **\$1.635 billion for 2025**.
- ✓ **VERITROOPER (PASS):** The evidence does not specify which year. Contractual service fees and related fees included in income are listed as \$1,635 million, \$1,606 million, and \$1,590 million for the three unspecified periods.

219. FAIL · Precision · `SEC\_10K\_multi\_0020\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the threshold for a fiscal year?
- **Ground truth:** 52 or 53 weeks
- ✖ **Baseline (vanilla RAG) (FAIL):** The threshold is **10% or more of total revenue in a fiscal year**.
- ✓ **VERITROOPER (PASS):** Apple's fiscal year is a 52- or 53-week period ending on the last Saturday of September.

220. PASS · Cross-Reference · `SEC\_10K\_multi\_0498\_01`

- **Question:** What year included adjustments for the CECL capital transition provisions?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The year ended December 31, 2024.
- ✓ **VERITROOPER (PASS):** The year ended December 31, 2024 included adjustments for the CECL capital transition provisions.

221. FAIL · Cross-Reference · `SEC\_10K\_multi\_0518\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total deposits in U.S. offices in the second year, given the total deposits in the first year and the growth rate?
- **Ground truth:** 1,918,734
- ✗ **Baseline (vanilla RAG) (FAIL):** Total deposits in U.S. offices in the second year were **\$1,991,661 million** (about **\$1.992 trillion**), up approximately **3.8%** from \$1,918,734 million.
- ✓ **VERITROOPER (PASS):** The total deposits in U.S. offices in the second year were **\$1,918,734**.

222. PASS · Exception · `SEC\_10K\_multi\_0708\_00`

- **Question:** What is excluded from the modeled estimates sensitivity analysis?
- **Ground truth:** The Apple Card transaction
- ✓ **Baseline (vanilla RAG) (PASS):** The forecasted Apple credit card portfolio is excluded while the Firm integrates the Apple Card transaction into its allowance model.
- ✓ **VERITROOPER (PASS):** The forecasted Apple credit card portfolio is excluded from the modeled estimates sensitivity analysis.

223. PASS · Conditional · `SEC\_10K\_multi\_0128\_01`

- **Question:** If JPMorganChase fails to comply with applicable law or meet supervisory expectations, what could result?
- **Ground truth:** Increased regulatory scrutiny, enforcement actions, penalties, increased exposure to litigation, or reputational harm.
- ✓ **Baseline (vanilla RAG) (PASS):** Failure to comply with applicable law or meet supervisory expectations could result in:
 - Increased regulatory scrutiny
 - Enforcement actions by governmental authorities
 - Penalties
 - Increased exposure to litigation
 - Reputational harm
- ✓ **VERITROOPER (PASS):** It could result in increased regulatory scrutiny, enforcement actions by governmental authorities, penalties, increased exposure to litigation, or reputational harm.

224. PASS · Cause & Effect · `SEC\_10K\_multi\_1261\_01`

- **Question:** What is the result of the Firm electing to repurchase delinquent loans from Ginnie Mae loan pools?
- **Ground truth:** Such loans continue to be insured or guaranteed.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm continues servicing the delinquent loans and/or managing foreclosure. Once the repurchase option is exercisable, it records the loans on its Consolidated balance sheet with a corresponding liability; the loans remain insured or guaranteed.

- ✓ **VERITROOPER (PASS):** The Firm continues servicing the delinquent loans and/or managing foreclosure, while the loans remain insured or guaranteed. Once the repurchase option is exercisable, the loans and corresponding liabilities are recognized on the Consolidated balance sheets.

225. PASS · Exception · `SEC\_10K\_multi\_1585\_01`

- **Question:** What is excluded from the line item trade accounts receivable in the consolidated balance sheet?
- **Ground truth:** Trade receivables sold under the trade accounts receivable factoring program.
- ✓ **Baseline (vanilla RAG) (PASS):** Trade accounts receivables sold under the Company's factoring program are excluded.
- ✓ **VERITROOPER (PASS):** Trade receivables sold under the Company's factoring program are excluded from trade accounts receivable.

226. FAIL · Precision · `SEC\_10K\_multi\_0388\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the allowance for loan losses to period-end loans?
- **Ground truth:** 0.20
- ✗ **Baseline (vanilla RAG) (FAIL):** The allowance for loan losses to period-end loans retained was **1.37% in 2025**.
- ✓ **VERITROOPER (PASS):** The allowance for loan losses to period-end loans was 0.20%, 0.23%, and 0.28% for Columns 1, 2, and 3, respectively.

227. PASS · Cause & Effect · `SEC\_10K\_multi\_1280\_00`

- **Question:** What happens as a result of excess MSRs being transferred to agency-sponsored trusts?
- **Ground truth:** Excess MSRs are exchanged for stripped mortgage-backed securities.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm receives stripped mortgage-backed securities (SMBS) in exchange. Some SMBS are acquired by third parties at the transaction date, while the Firm retains the remainder as trading securities.
- ✓ **VERITROOPER (PASS):** In exchange for the transferred excess MSRs, stripped mortgage-backed securities (SMBS) are received; third parties acquire a portion, while the Firm acquires the remaining balance as trading securities.

228. PASS · Cause & Effect · `SEC\_10K\_multi\_1218\_01`

- **Question:** What is the consequence of the Firm being the Residual holder?
- **Ground truth:** The Firm has the right to make decisions that significantly impact the economic performance of the municipal bond vehicle.
- ✓ **Baseline (vanilla RAG) (PASS):** As the Residual holder, the Firm consolidates non-customer TOB trusts because it can make decisions significantly affecting the vehicles' economic performance and can receive potentially significant benefits or bear losses.
- ✓ **VERITROOPER (PASS):** As the Residual holder, the Firm can make decisions significantly affecting the municipal bond vehicle's economic performance and receive significant benefits or bear significant losses; consequently, it consolidates non-customer TOB trusts.

229. PASS · Conditional · `SEC\_10K\_multi\_0008\_01`

- **Question:** If the Company is unable to monitor and limit exposure to credit risk on its trade and vendor non-trade receivables, what could happen?

- **Ground truth:** The Company's results of operations, financial condition and stock price could be materially adversely affected.
- ✓ **Baseline (vanilla RAG) (PASS):** The Company's results of operations, financial condition, and stock price could be materially adversely affected.
- ✓ **VERITROOPER (PASS):** The Company's results of operations, financial condition, and stock price could be materially adversely affected.

230. PASS · Cross-Reference · `SEC\_10K\_multi\_1022\_00`

- **Question:** What type of protection is represented by (e) and how is it used?
- **Ground truth:** Protection represented by (e) is credit protection, which is used against certain loans and lending-related commitments.
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not include item or footnote (e), so the type of protection and its use cannot be determined from the excerpt.
- ✓ **VERITROOPER (PASS):** (e) represents protection purchased on single-name, portfolio, or index instruments where the Firm has not sold protection on the identical instrument. It also includes credit protection used to hedge certain loans and lending-related commitments in the retained lending portfolio through credit derivatives and credit-related notes.

231. PASS · Cross-Reference · `SEC\_10K\_multi\_1415\_00`

- **Question:** What happens to deferred amounts for revolving commitments?
- **Ground truth:** They are amortized to lending- and deposit-related fees on a straight line basis over the commitment period.
- ✓ **Baseline (vanilla RAG) (PASS):** Deferred amounts for revolving commitments are amortized to lending- and deposit-related fees on a straight-line basis over the commitment period.
- ✓ **VERITROOPER (PASS):** They are amortized to lending- and deposit-related fees on a straight-line basis over the commitment period.

232. FAIL · Cross-Reference · `SEC\_10K\_multi\_0518\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total interest-bearing deposits in non-U.S. offices in the third year, given the total interest-bearing deposits in the first year and the growth rate?
- **Ground truth:** 408,419
- ✗ **Baseline (vanilla RAG) (FAIL):** The total interest-bearing deposits in non-U.S. offices in the third year (2023) were **\$408,419 million**, or approximately **\$408.4 billion**.
- ✗ **VERITROOPER (FAIL):** The total interest-bearing deposits in non-U.S. offices in the third year were **\$482,735**.

233. PASS · Precision · `SEC\_10K\_multi\_0263\_00`

- **Question:** What is the year in which the Firm adopted updates to the Accounting for Investments in Tax Credit Structures guidance?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm adopted the updates effective January 1, 2024.
- ✓ **VERITROOPER (PASS):** The Firm adopted the updates effective January 1, 2024.

234. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1531\_02`

- **Question:** How many countries does the Company have a direct presence in for alcohol beverage sales?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a number; it only states that the Company directly sells alcohol beverages in "numerous markets outside the United States."
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many countries have a direct Company presence in alcohol beverage sales.

235. PASS · Cross-Reference · `SEC\_10K\_multi\_1200\_00`

- **Question:** What were the lending-related commitments excluded for the consumer segment at December 31, 2023 and 2024?
- **Ground truth:** 17.2 billion and 19.2 billion
- ✓ **Baseline (vanilla RAG) (PASS):** Lending-related commitments excluded for the consumer segment, excluding credit cards, were:
- **December 31, 2023:** \$17.2 billion
- **December 31, 2024:** \$19.2 billion
- ✓ **VERITROOPER (PASS):** The excluded lending-related commitments for the consumer segment were \$17.2 billion in 2023 and \$19.2 billion in 2024.

236. PASS · Precision · `SEC\_10K\_multi\_0441\_01`

- **Question:** What is the number of the page where Reputation Risk Management is discussed?
- **Ground truth:** 108
- ✓ **Baseline (vanilla RAG) (PASS):** Reputation Risk Management is discussed on **page 108**.
- ✓ **VERITROOPER (PASS):** Reputation Risk Management is discussed on page 108.

237. PASS · Cross-Reference · `SEC\_10K\_multi\_1544\_00`

- **Question:** What was the total amount of insurance coverage purchased by the Company for Aqua-Chem during its ownership?
- **Ground truth:** over \$400 million
- ✓ **Baseline (vanilla RAG) (PASS):** The Company purchased over **\$400 million** in insurance coverage for Aqua-Chem.
- ✓ **VERITROOPER (PASS):** The Company purchased over \$400 million in insurance coverage for Aqua-Chem during its ownership from 1970 to 1981.

238. PASS · Cross-Reference · `SEC\_10K\_multi\_1181\_01`

- **Question:** How many months were loans extended by a weighted-average for single modifications?
- **Ground truth:** 23
- ✓ **Baseline (vanilla RAG) (PASS):** Loans were extended by a weighted average of **14 months**.
- ✓ **VERITROOPER (PASS):** Loans with single modifications were extended by a weighted average of 23 months.

239. FAIL · Cross-Reference · `SEC\_10K\_multi\_1265\_01`

BAD QUESTION

- **Question:** What is the net liquidation loss for securitized loans in 2025, and how does it compare to the net liquidation loss in 2024?

- **Ground truth:** The net liquidation loss for securitized loans in 2025 is \$5,233 million, which is higher than the net liquidation loss in 2024 of \$2,329 million.
- **✗ Baseline (vanilla RAG) (FAIL):** The net liquidation loss for securitized loans was **\$301 million in 2025**, compared with **\$89 million in 2024**-an increase of **\$212 million**.
- **✗ VERITROOPER (FAIL):** The net liquidation loss for securitized loans was \$301 million in 2025, compared with \$89 million in 2024-an increase of \$212 million.

240. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_04`

- **Question:** What is the percentage of Apple's revenue that comes from the European Union?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence provided does not state what percentage of Apple's revenue comes from the European Union.
- **✓ VERITROOPER (PASS):** The evidence does not specify what percentage of Apple's revenue comes from the European Union.

241. PASS · Cause & Effect · `SEC\_10K\_multi\_0006\_01`

- **Question:** What happens as a consequence of information technology system failures or network disruptions?
- **Ground truth:** Adverse impact on the Company's business, reputation, results of operations, financial condition, and stock price.
- **✓ Baseline (vanilla RAG) (PASS):** Information technology system failures or network disruptions can:
 - Prevent access to online services and account information.
 - Disrupt customer transactions, payments, communications, and other services.
 - Impede manufacturing, shipping, distribution, invoicing, and collections.
 - Delay financial reporting and transaction settlement.
 - Cause data loss, unauthorized disclosure, or misuse of confidential information.
 - Result in lost revenue, higher costs, litigation, fines, and other financial losses.
 - Damage business operations, customer confidence, reputation, financial condition, and stock price.
- **✓ VERITROOPER (PASS):** They can prevent access to online services, interfere with customer transactions, and impede product manufacturing and shipping, materially harming the Company's business, reputation, results of operations, financial condition, and stock price.

242. FAIL · Precision · `SEC\_10K\_multi\_0231\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the total credit provided and capital raised for U.S. small businesses?
- **Ground truth:** \$2.9 trillion
- **✗ Baseline (vanilla RAG) (FAIL):** The total credit provided and capital raised for U.S. small businesses was **\$33 billion**.
- **✗ VERITROOPER (FAIL):** The total credit provided and capital raised for U.S. small businesses was **\$33 billion** in 2025.

243. FAIL · Precision · `SEC\_10K\_multi\_0574\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the total number of government guaranteed loans at December 31, 2025?
- **Ground truth:** 1,159
- **✗ Baseline (vanilla RAG) (FAIL):** The total government-guaranteed loans at December 31, 2025 were **\$1.159 billion**.

- ✖ **VERITROOPER (FAIL):** The total government guaranteed loans at December 31, 2025, were \$1.159 billion.

244. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_02`

- **Question:** What is the total market value of NVIDIA's shares held by institutional investors?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The total market value of NVIDIA shares held by institutional investors cannot be determined from the provided evidence. It does not disclose institutional shareholdings or the market price needed for the calculation.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total market value of NVIDIA shares held by institutional investors.

245. FAIL · Cross-Reference · `SEC\_10K\_multi\_0911\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total amount for Insurance and what is the total amount for Central Govt, and how do these two totals compare?
- **Ground truth:** The total amount for Insurance is \$25,031 and the total amount for Central Govt is \$15,164. These two totals are not comparable as they are different categories.
- ✖ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025:
 - **Insurance:** \$25.031 billion
 - **Central Govt:** \$15.164 billion

Insurance exposure was **\$9.867 billion higher**, approximately **65% more** than Central Govt exposure.

- ✖ **VERITROOPER (FAIL):** Insurance totaled 25,031, while Central Govt totaled 15,164. Insurance was higher by 9,867.

246. PASS · Precision · `SEC\_10K\_multi\_1125\_01`

- **Question:** What is the threshold for delinquency rates?
- **Ground truth:** 30
- ✓ **Baseline (vanilla RAG) (PASS):** The primary delinquency threshold is **30 or more days past due**. For residential real estate loans, **more than 150 days past due** is considered late-stage delinquency.
- ✓ **VERITROOPER (PASS):** Loans more than 30 days past due are considered delinquent.

247. PASS · Conditional · `SEC\_10K\_multi\_0129\_01`

- **Question:** When resolving an investigation or enforcement action, what typically must JPMorganChase satisfy?
- **Ground truth:** New or enhanced regulatory requirements or restrictions.
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase typically must satisfy new or enhanced regulatory requirements or restrictions.
- ✓ **VERITROOPER (PASS):** JPMorganChase typically must satisfy new or enhanced regulatory requirements or restrictions.

248. PASS · Precision · `SEC\_10K\_multi\_0766\_01`

- **Question:** What is the IR-FX correlation percentage?
- **Ground truth:** 35
- ✓ **Baseline (vanilla RAG) (PASS):** The IR-FX correlation is **5%**.
- ✓ **VERITROOPER (PASS):** The IR-FX correlation is 5%.

249. PASS · Cross-Reference · `SEC\_10K\_multi\_0503\_00`

- **Question:** What is the percentage of RWA that the external TLAC represents as of December 31, 2025, and how does it compare to the regulatory requirements?
- **Ground truth:** 27.6% of RWA, which is greater than the regulatory requirements of 23.0%.
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, external TLAC represented **27.6% of RWA**, compared with the **23.0% regulatory requirement**-a surplus of **4.6 percentage points**, or **\$93.3 billion**.
- ✓ **VERITROOPER (PASS):** As of December 31, 2025, external TLAC represented 27.6% of RWA, exceeding the 23.0% regulatory requirement by 4.6 percentage points.

250. PASS · Exception · `SEC\_10K\_multi\_0713\_01`

- **Question:** What does not apply to the requirement for additional disaggregation of specific types of expenses?
- **Ground truth:** Interim financial statements before December 31, 2027
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify an exclusion. The additional expense-disaggregation requirement applies to disclosures in the notes for both annual and interim financial statements.
- ✓ **VERITROOPER (PASS):** It does not apply to the 2025 annual financial statements; the requirement is effective for the year ending December 31, 2027.

251. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0442\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of regulatory capital that must be held by the Firm's principal IDI subsidiary, JPMorgan Chase Bank, N.A.?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** JPMorgan Chase Bank, N.A.'s total capital requirement was **10.5% of risk-weighted assets**. Using its 2025 Standardized RWA of **\$1.928 trillion**, this equates to approximately **\$202.4 billion** in required regulatory capital. It actually held **\$317.7 billion** of Total capital under the Standardized approach.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the total amount of regulatory capital JPMorgan Chase Bank, N.A. must hold.

252. PASS · Cross-Reference · `SEC\_10K\_multi\_0850\_00`

- **Question:** What is the total amount of level 3 assets at December 31, 2025, and the amount of level 3 assets at December 31, 2024?
- **Ground truth:** \$25.1 billion and \$23.8 billion
- ✓ **Baseline (vanilla RAG) (PASS):** Level 3 assets totaled **\$25.1 billion at December 31, 2025**, an increase of **\$1.3 billion** from **\$23.8 billion at December 31, 2024**.
- ✓ **VERITROOPER (PASS):** Level 3 assets totaled \$25.1 billion at December 31, 2025, and \$23.8 billion at December 31, 2024.

253. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0646\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average allowance for loan losses to retained nonaccrual loans?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The average allowance for loan losses to retained nonaccrual loans was **325%**, based on **311% in 2025** and **339% in 2024**.

- ✓ **VERITROOPER (PASS):** The evidence does not provide the values for allowance for loan losses to retained nonaccrual loans, so the average cannot be determined.

254. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0714\_01`

- **Question:** What is the total number of contracts that will no longer require derivative accounting treatment?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify the number of contracts; it only states that "certain contracts" will no longer require derivative accounting treatment.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the number of contracts that will no longer require derivative accounting treatment.

255. PASS · Cause & Effect · `SEC\_10K\_multi\_0293\_01`

- **Question:** What is the consequence of changes in FTP in Banking & Wealth Management?
- **Ground truth:** Net interest income is affected, with a partial offset by lower average deposit balances.
- ✓ **Baseline (vanilla RAG) (PASS):** Changes in funds transfer pricing (FTP) increased Banking & Wealth Management's deposit margin, contributing to higher net interest income, partially offset by lower average deposit balances.
- ✓ **VERITROOPER (PASS):** Changes in FTP drove a higher deposit margin in Banking & Wealth Management, contributing to higher net interest income, partially offset by lower average deposit balances.

256. PASS · Cause & Effect · `SEC\_10K\_multi\_1202\_01`

- **Question:** What is the consequence of the Firm's qualitative adjustments and macroeconomic outlook on the allowance for credit losses?
- **Ground truth:** additional weight placed on adverse scenarios
- ✓ **Baseline (vanilla RAG) (PASS):** They increased the allowance for credit losses by placing additional-and during 2025, greater-weight on adverse scenarios to reflect geopolitical and macroeconomic uncertainties and downside risks. The allowance therefore includes losses beyond those estimated under the central scenario.
- ✓ **VERITROOPER (PASS):** The outlook changes contributed to the \$3.3 billion consumer allowance increase but partially offset the \$1.1 billion wholesale increase; the Firm also increased weighting on more punitive adverse scenarios to reflect ongoing uncertainties and downside risks.

257. PASS · Exception · `SEC\_10K\_multi\_1547\_00`

- **Question:** Under what circumstance does the general rule of consolidation not apply?
- **Ground truth:** when an entity holds an interest in another business enterprise through arrangements that do not involve voting interests
- ✓ **Baseline (vanilla RAG) (PASS):** The usual majority-voting-interest rule does not apply when an entity has a variable interest in another enterprise through arrangements not involving voting interests. Consolidation is then required if the entity is the VIE's primary beneficiary-having both power over its most significant activities and potentially significant economic exposure.
- ✓ **VERITROOPER (PASS):** The general rule of consolidation based on majority voting ownership does not apply when arrangements not involving voting interests create a variable interest and the company is the VIE's primary beneficiary.

258. PASS · Cause & Effect · `SEC\_10K\_multi\_0392\_00`

- **Question:** What happens as a result of the centralization of certain functions?
- **Ground truth:** 130 employees were transferred to Corporate

- ✓ **Baseline (vanilla RAG) (PASS):** As a result of centralizing certain functions in Q1 2025, 768 employees were transferred from the lines of business to Corporate: 419 from CCB, 219 from CIB, and 130 from AWM.
- ✓ **VERITROOPER (PASS):** As a result, 130 employees were transferred to Corporate in the first quarter of 2025.

259. FAIL · Cross-Reference · `SEC\_10K\_multi\_0642\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the net charge-off for the consumer credit card segment in the year ended December 31, 2024?
- **Ground truth:** 7,142
- ✗ **Baseline (vanilla RAG) (FAIL):** The consumer credit card segment's net charge-offs were **\$7.142 billion** for the year ended December 31, 2024.
- ✓ **VERITROOPER (PASS):** The net charge-off for the consumer credit card segment in 2024 was \$7,142 million.

260. FAIL · Precision · `SEC\_10K\_multi\_0803\_01`

MISINTERPRETED

- **Question:** What is the total number of debt instruments?
- **Ground truth:** 2,138
- ✗ **Baseline (vanilla RAG) (FAIL):** Total debt instruments were **\$2,182 million (\$2.182 billion)** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total debt instruments were 2,138.

261. FAIL · Precision · `SEC\_10K\_multi\_1157\_00`

TABLE MISREAD

- **Question:** What is the total retained loans for 2022?
- **Ground truth:** \$14,917
- ✗ **Baseline (vanilla RAG) (FAIL):** The total retained loans for the 2022 origination year were **\$5.069 billion**.
- ✓ **VERITROOPER (PASS):** The total retained loans for 2022 were **\$14,917 million**.

262. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0075\_00`

- **Question:** What is the total amount of customer advances as of January 27, 2026?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide an amount for January 27, 2026. As of **January 25, 2026**, customer advances totaled **\$160 million**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide customer advances as of January 27, 2026; it reports \$160 million as of January 25, 2026.

263. FAIL · Precision · `SEC\_10K\_multi\_0751\_01`

TABLE MISREAD

- **Question:** What is the total U.S. Treasury and government agencies?
- **Ground truth:** 234,779
- ✗ **Baseline (vanilla RAG) (FAIL):** The total U.S. Treasury and government agencies available-for-sale securities was **\$315.822 billion in fair value** as of December 31, 2025 (**\$313.470 billion amortized cost**).
- ✓ **VERITROOPER (PASS):** The total U.S. Treasury and government agencies is **234,779**.

264. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0744\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of accounts payable and other liabilities measured at fair value on a recurring basis?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The total was **\$5.893 billion (\$5,893 million)**.
- ✗ **VERITROOPER (FAIL):** The total accounts payable and other liabilities measured at fair value on a recurring basis was **\$6.660 billion**.

265. PASS · Exception · `SEC\_10K\_multi\_0039\_01`

- **Question:** What circumstance does the obligation to update forward-looking statements NOT apply to?
- **Ground truth:** Even if new information becomes available in the future
- ✓ **Baseline (vanilla RAG) (PASS):** The obligation does not apply when new information becomes available or future events occur, except where an update is required by law.
- ✓ **VERITROOPER (PASS):** The obligation does **not** apply even if new information becomes available in the future, except when an update is required by law.

266. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0539\_02`

HALLUCINATION

- **Question:** How many credit rating downgrades occurred for JPMorgan Chase & Co. in 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Zero credit rating downgrades occurred for JPMorgan Chase & Co. in 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many credit rating downgrades occurred in 2025.

267. PASS · Cross-Reference · `SEC\_10K\_multi\_0180\_01`

- **Question:** What is the reference for the Employee Stock Purchase Plan and the Performance-Based Incentive Compensation Plan?
- **Ground truth:** Exhibit 10.1 to the Quarterly Report on Form 10-Q and Exhibit 10.23 to the Annual Report on Form 10-K
- ✓ **Baseline (vanilla RAG) (PASS):** - **Employee Stock Purchase Plan:** Exhibit **10.18**, incorporated by reference to Exhibit **10.1** of JPMorgan Chase & Co.'s Form **10-Q** for the quarter ended **September 30, 2019**.
- **Performance-Based Incentive Compensation Plan:** Exhibit **10.19**, incorporated by reference to Exhibit **10.23** of JPMorgan Chase & Co.'s Form **10-K** for the year ended **December 31, 2021**.
- ✓ **VERITROOPER (PASS):** The Employee Stock Purchase Plan is referenced as Exhibit 10.18, and the Performance-Based Incentive Compensation Plan is referenced as Exhibit 10.19.

268. PASS · Cause & Effect · `SEC\_10K\_multi\_1145\_01`

- **Question:** What is the consequence of extending the weighted-average life of the loans by 19 years?
- **Ground truth:** The weighted-average contractual interest rate is reduced from 6.94% to 6.08%.
- ✓ **Baseline (vanilla RAG) (PASS):** The loans' repayment period was lengthened by an average of 19 years; alongside this modification, the weighted-average contractual interest rate fell from 6.94% to 6.08%.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a consequence of the 19-year extension; it only notes that the weighted-average contractual interest rate decreased from 6.94% to 6.08%.

269. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0109\_06`

HALLUCINATION

- **Question:** What is the name of the chairman of the board of directors of JPMorgan Chase & Co.?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** James Dimon.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the name of JPMorgan Chase & Co.'s chairman of the board.

270. PASS · Cross-Reference · `SEC\_10K\_multi\_0163\_00`

- **Question:** What is the total number of certificates of designation listed under Item 3, and how does it relate to the number of exhibits listed under Item 15?
- **Ground truth:** There are 13 certificates of designation listed under Item 3, and 3 exhibits listed under Item 15.
- ✓ **Baseline (vanilla RAG) (PASS):** There are **9 certificates of designation** (Exhibits **3.2 through 3.10**). They represent **9 of the 10 exhibits listed under Item 15**; Exhibit 3.1 is the Restated Certificate of Incorporation.
- ✓ **VERITROOPER (PASS):** There are 11 certificates of designation (Exhibits 3.2-3.12), representing 11 of the 13 Item 3 exhibits listed under Item 15.

271. PASS · Cause & Effect · `SEC\_10K\_multi\_0442\_01`

- **Question:** What is the effect of the Federal Reserve's proposal to reduce SCB volatility?
- **Ground truth:** The proposal aims to reduce SCB volatility by using the average of supervisory stress results from the previous two annual stress tests to calculate the SCB.
- ✓ **Baseline (vanilla RAG) (PASS):** The proposal would reduce Stress Capital Buffer (SCB) volatility by averaging supervisory stress-test results from the prior two annual tests. It would also shift the SCB's annual effective date from October 1 to January 1 and streamline related reporting requirements. The proposal has not been finalized.
- ✓ **VERITROOPER (PASS):** The proposal would calculate the SCB using the average supervisory stress results from the prior two annual stress tests, shift its annual effective date from October 1 to January 1, and streamline data collection through targeted reporting changes.

272. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1255\_03`

- **Question:** What is the percentage of loans sold that are insured by another U.S. government agency?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** Cannot be determined from the provided evidence. The disclosed loan-sale totals combine loans sold to U.S. GSEs with Ginnie Mae securitizations and do not separately quantify loans insured or guaranteed by another U.S. government agency.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the amount or percentage of loans sold that are insured by another U.S. government agency.

273. PASS · Exception · `SEC\_10K\_multi\_0651\_01`

- **Question:** What is not included in the measures used to capture market risk?
- **Ground truth:** there is no single measure
- ✓ **Baseline (vanilla RAG) (PASS):** Peak, Derivative Risk Equivalent (DRE), and Average Exposure (AVG) are **not** market-risk measures; they measure potential derivatives-related credit loss.
- ✓ **VERITROOPER (PASS):** Allowance for loan losses is not listed among the measures used to capture market risk.

274. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1602\_02`

- **Question:** What is the total value of the Company's long-term debt as of December 31, 2023?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide the Company's total long-term debt as of December 31, 2023. It only reports \$43.941 billion for 2025 and \$43.023 billion for 2024.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the Company's long-term debt as of December 31, 2023.

275. PASS · Exception · `SEC\_10K\_multi\_1540\_00`

- **Question:** Under what circumstance does the general rule NOT apply regarding compliance with legal requirements?
- **Ground truth:** Improper conduct by our employees or agents
- ✓ **Baseline (vanilla RAG) (PASS):** Under California Proposition 65, the warning requirement does not apply when exposure is below an established safe-harbor level or another exemption applies.
- ✓ **VERITROOPER (PASS):** The general expectation does not apply when the company's policies, procedures, and training programs fail to ensure full compliance by employees or agents with applicable legal requirements.

276. FAIL · Precision · `SEC\_10K\_multi\_0082\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the amount of cash paid for income taxes (net of refunds) for the fiscal year ended January 25, 2026?
- **Ground truth:** \$20,288
- ✗ **Baseline (vanilla RAG) (FAIL):** The amount of cash paid for income taxes, net of refunds, was **\$20.288 billion** for the fiscal year ended January 25, 2026.
- ✗ **VERITROOPER (FAIL):** The cash paid for income taxes, net of refunds, was **\$20.288 billion** for the fiscal year ended January 25, 2026.

277. PASS · Cross-Reference · `SEC\_10K\_multi\_0320\_00`

- **Question:** What were the total taxable-equivalent adjustments from income tax credits for the years ended December 31, 2023, 2024, and 2025?
- **Ground truth:** \$4.0 billion, \$2.8 billion, and \$2.9 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The taxable-equivalent adjustments were:
 - **2023:** \$4.0 billion
 - **2024:** \$2.8 billion
 - **2025:** \$2.9 billion

Total: \$9.7 billion.

- ✗ **VERITROOPER (FAIL):** The total taxable-equivalent adjustments for 2023, 2024, and 2025 were **\$9.7 billion** (\$4.0 billion, \$2.8 billion, and \$2.9 billion, respectively).

278. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0135\_05`

- **Question:** What is the name of the specific model used by JPMorganChase for calculating market risk RWA?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not identify a specific model by name for calculating market risk RWA.

- ✓ **VERITROOPER (PASS)**: The evidence does not specify the name of the model used to calculate market risk RWA.

279. PASS · Cause & Effect · `SEC\_10K\_multi\_1218\_00`

- **Question**: What happens as a consequence of the remarketing agent's inability to remarket the floaters?

- **Ground truth**: The liquidity provider provides a loan to the TOB trust or directly purchases the tendered floaters.

- ✓ **Baseline (vanilla RAG) (PASS)**: If the remarketing agent cannot remarket tendered floaters, the liquidity provider either lends funds to the TOB trust so it can purchase them or directly purchases the floaters.

- ✓ **VERITROOPER (PASS)**: The liquidity provider either lends to the TOB trust so it can purchase the floaters or directly purchases the tendered floaters.

280. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0137\_02`

- **Question**: What is the cost of the cybersecurity measures implemented by JPMorganChase?

- **Ground truth**: UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not disclose the cost of JPMorganChase's cybersecurity measures. It only states that the firm devotes "significant resources" to cybersecurity and may need to make significant future technology investments.

- ✓ **VERITROOPER (PASS)**: The evidence does not state the cost of JPMorganChase's cybersecurity measures.

281. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1537\_03`

- **Question**: What is the percentage of our revenue that comes from digital transformation initiatives?

- **Ground truth**: UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not disclose what percentage of revenue comes from digital transformation initiatives. It only describes digital transformation as important for adapting to changing consumer shopping preferences.

- ✓ **VERITROOPER (PASS)**: The evidence does not state what percentage of revenue comes from digital transformation initiatives.

282. FAIL · Conditional · `SEC\_10K\_multi\_0109\_00`

DETERMINISTIC:DIRECTION CONFLICT

- **Question**: If JPMorganChase is a large accelerated filer, does it have to file reports pursuant to Section 13 or 15(d) of the Act?

- **Ground truth**: No

- ✗ **Baseline (vanilla RAG) (FAIL)**: Yes. JPMorganChase is a large accelerated filer and indicated that it is required to file reports pursuant to Section 13 or 15(d) of the Act.

- ✗ **VERITROOPER (FAIL)**: Yes. JPMorganChase is a large accelerated filer and indicated that it is required to file reports pursuant to Section 13 or 15(d) of the Act.

283. FAIL · Cross-Reference · `SEC\_10K\_multi\_1160\_00`

MISINTERPRETED

- **Question**: What is the total retained loans amount in the U.S. and the total retained loans amount in non-U.S., respectively?

- **Ground truth**: \$165,545 and \$3,167

- ✗ **Baseline (vanilla RAG) (FAIL)**: As of December 31, 2025, total retained loans were **\$626.060 billion in the U.S.** and **\$166.307 billion in non-U.S.** markets, respectively.

- ✓ **VERITROOPER (PASS):** The total retained loans were **\$626,060 in the U.S.** and **\$166,307 in non-U.S. markets.**

284. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_05`

HALLUCINATION

- **Question:** What is the average maturity date of the notes issued by Apple Inc. in 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The notes mature in 2028, 2030, 2032, and 2035. Their simple average maturity year is **2031.25**, or approximately **2031**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide exact maturity dates, only maturity years (2028, 2030, 2032, and 2035), so an average maturity date cannot be determined.

285. PASS · Cross-Reference · `SEC\_10K\_multi\_0516\_01`

- **Question:** What is the percentage of deposits as a percentage of total liabilities as of December 31, 2025, and how does it compare to the percentage as of December 31, 2024?
- **Ground truth:** 63% and 66%, respectively, with a decrease of 3% compared to the previous year.
- ✓ **Baseline (vanilla RAG) (PASS):** Deposits were **63% of total liabilities** as of December 31, 2025, down **3 percentage points** from **66%** as of December 31, 2024.
- ✓ **VERITROOPER (PASS):** Deposits were 63% of total liabilities as of December 31, 2025, down 3 percentage points from 66% as of December 31, 2024.

286. PASS · Cause & Effect · `SEC\_10K\_multi\_0132\_00`

- **Question:** What happens as a consequence of a significant deterioration in the credit quality of a counterparty?
- **Ground truth:** concerns about the creditworthiness of other parties in similar, related or dependent industries
- ✓ **Baseline (vanilla RAG) (PASS):** It could raise concerns about the creditworthiness of other parties in similar, related, or dependent industries, exacerbating JPMorganChase's credit, liquidity, and market risk exposure and potentially causing losses.
- ✓ **VERITROOPER (PASS):** It could raise concerns about the creditworthiness of parties in similar, related, or dependent industries, exacerbating JPMorganChase's credit, liquidity, and market risk exposure and potentially causing losses.

287. PASS · Cause & Effect · `SEC\_10K\_multi\_0048\_01`

- **Question:** What happens as a result of our inability to attract, retain, and motivate our executives and key employees?
- **Ground truth:** our business may be harmed
- ✓ **Baseline (vanilla RAG) (PASS):** Our business may be harmed, hindering strategic planning, execution, and long-term success. It could also increase compensation costs, deplete institutional knowledge, erode competitiveness, and disrupt operations.
- ✓ **VERITROOPER (PASS):** The evidence states only that the business may be harmed; it does not specify further consequences.

288. PASS · Cross-Reference · `SEC\_10K\_multi\_1535\_00`

- **Question:** What are the potential risks to our business due to unfavorable economic conditions, and how might they impact our profitability?

- **Ground truth:** Unfavorable general economic and geopolitical conditions could negatively impact our financial results, leading to reduced consumer demand and profitability.
- **✓ Baseline (vanilla RAG) (PASS):** Unfavorable economic conditions-such as inflation, high interest rates, unemployment, recession, weak credit markets, currency fluctuations, commodity and energy price increases, and declining consumer confidence-could:
 - Reduce consumers' purchasing power and demand for our beverages.
 - Cause consumers to forgo purchases or switch to lower-priced competitors and private-label brands.
 - Increase transportation, energy, packaging, raw material, and other input costs.
 - Disrupt logistics, supply chains, labor availability, and operations.
 - Create trade, tariff, tax, capital-transfer, price-control, or sanctions-related restrictions.

These effects could lower sales, increase costs, compress margins, and thereby reduce profitability and overall financial performance.

- **✓ VERITROOPER (PASS):** Unfavorable economic conditions may reduce consumer purchasing power and demand, shift purchases toward lower-priced or private-label products, disrupt supply chains and operations, and increase transportation, energy, packaging, raw-material, and other input costs. Trade restrictions, tariffs, sanctions, price controls, and limits on cross-border capital transfers could further reduce profitability and negatively affect financial performance.

289. PASS · Cross-Reference · `SEC\_10K\_multi\_1601\_01`

- **Question:** What is the total amount of liabilities measured at fair value on a recurring basis at December 31, 2025, according to Note 110?
- **Ground truth:** 95
- **✓ Baseline (vanilla RAG) (PASS):** The total liabilities measured at fair value on a recurring basis were **\$95 million** at December 31, 2025.
- **✓ VERITROOPER (PASS):** The total liabilities measured at fair value on a recurring basis at December 31, 2025 were **\$95 million**.

290. PASS · Cross-Reference · `SEC\_10K\_multi\_0911\_01`

- **Question:** What is the total amount for Transportation and what is the total amount for Financial Markets Infrastructure, and how do these two totals compare?
- **Ground truth:** The total amount for Transportation is \$20,861 and the total amount for Financial Markets Infrastructure is \$5,734. These two totals are not comparable as they are different categories.
- **✓ Baseline (vanilla RAG) (PASS):** The provided evidence does not state amounts for either **Transportation** or **Financial Markets Infrastructure**, so their totals and comparison cannot be determined.
- **✗ VERITROOPER (FAIL):** Transportation totals 20,861, while Financial Markets Infrastructure totals 5,734. Transportation is 15,127 higher, or about 3.6 times as much.

291. PASS · Cause & Effect · `SEC\_10K\_multi\_1045\_00`

- **Question:** What is the consequence of the agreement between the Firm and the FDIC?
- **Ground truth:** A gain of \$588 million recorded in other income in the first quarter of 2025.
- **✓ Baseline (vanilla RAG) (PASS):** The Firm paid \$609 million to the FDIC on January 31, 2025 and reduced its additional payable to the FDIC, resulting in a \$588 million gain recorded in other income in the first quarter of 2025. By June 30, 2025, all outstanding purchase-price settlement matters related to the First Republic acquisition were resolved.
- **✓ VERITROOPER (PASS):** The Firm paid the FDIC \$609 million, reduced its additional payable, and recorded a \$588 million gain in other income in the first quarter of 2025.

292. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0727\_02`

- **Question:** What is the maximum amount of valuation adjustments that can be made to a position?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a maximum amount. Valuation adjustments are determined as necessary to ensure positions are recorded at fair value, based on factors such as liquidity, unobservable inputs, net open risk, credit quality, and funding.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a maximum amount of valuation adjustments that can be made to a position.

293. PASS · Cause & Effect · `SEC\_10K\_multi\_0707\_01`

- **Question:** What is the potential consequence of incorrect or misused estimation outputs?
- **Ground truth:** Adverse consequences
- ✓ **Baseline (vanilla RAG) (PASS):** Incorrect or misused estimation outputs can lead to adverse consequences, including unexpected losses and poor decisions.
- ✓ **VERITROOPER (PASS):** Incorrect or misused estimation outputs can lead to adverse consequences in decision-making.

294. FAIL · Cross-Reference · `SEC\_10K\_multi\_1297\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the net rental expense for the year ended December 31, 2025, and how does it compare to the gross rental expense?
- **Ground truth:** Net rental expense: \$2,360, Gross rental expense: \$2,388
- ✗ **Baseline (vanilla RAG) (FAIL):** For the year ended December 31, 2025, net rental expense was **\$2.360 billion**, compared with gross rental expense of **\$2.388 billion**. The **\$28 million difference** reflects sublease rental income.
- ✗ **VERITROOPER (FAIL):** Net rental expense was \$2.360 billion in 2025, compared with gross rental expense of \$2.388 billion-a \$28 million reduction from sublease rental income.

295. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_09`

- **Question:** What is the total number of notes issued by Apple Inc. in 2024?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not state the total number of notes Apple issued in 2024.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total number of notes issued by Apple Inc. in 2024.

296. FAIL · Precision · `SEC\_10K\_multi\_0388\_00`

UNSUPPORTED NUMBER

- **Question:** What is the allowance for lending-related commitments?
- **Ground truth:** 43
- ✗ **Baseline (vanilla RAG) (FAIL):** At December 31, 2025, the allowance for lending-related commitments was **\$5.071 billion**: **\$83 million** for consumer excluding credit card, **\$2.2 billion** for credit card, and **\$2.788 billion** for wholesale.
- ✓ **VERITROOPER (PASS):** The allowance for lending-related commitments was \$43, \$35, and \$28 for Columns 1, 2, and 3, respectively.

297. PASS · Cause & Effect · `SEC\_10K\_multi\_0760\_00`

- **Question:** What happens as a consequence of the lack of observability of significant inputs?
- **Ground truth:** Management must assess relevant empirical data in deriving valuation inputs.
- ✓ **Baseline (vanilla RAG) (PASS):** Management must assess relevant empirical data to derive valuation inputs, such as transaction details, yield curves, interest rates, default rates, volatilities, correlations, market prices, comparable-instrument valuations, foreign exchange rates, and credit curves. This requires greater management judgment in estimating fair value.
- ✓ **VERITROOPER (PASS):** Management must assess relevant empirical data to derive valuation inputs, and the judgments required to estimate fair value become more significant.

298. PASS · Cause & Effect · `SEC\_10K\_multi\_1185\_00`

- **Question:** What happens as a result of the Firm's allowance for credit losses being reviewed at least quarterly?
- **Ground truth:** Significant changes in the allowance for credit losses in future periods.
- ✓ **Baseline (vanilla RAG) (PASS):** Subsequent evaluations may result in significant changes to the allowance for credit losses as macroeconomic conditions, forecasts, and other factors evolve.
- ✓ **VERITROOPER (PASS):** Subsequent evaluations may result in significant changes to the allowance for credit losses in future periods.

299. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0254\_04`

HALLUCINATION

- **Question:** What is the name of the specific non-GAAP financial measure used by the Firm to compare with competitors?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The non-GAAP financial measure is **managed basis**.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the name of a particular non-GAAP financial measure used to compare the Firm with competitors.

300. PASS · Exception · `SEC\_10K\_multi\_0073\_01`

- **Question:** What is excluded from the expected remaining lease term for amortization?
- **Ground truth:** the asset's estimated useful life
- ✓ **Baseline (vanilla RAG) (PASS):** Extension or termination option periods are excluded unless the company is reasonably certain it will exercise the option.
- ✓ **VERITROOPER (PASS):** The evidence does not specify anything excluded from the expected remaining lease term for amortization.

301. PASS · Cross-Reference · `SEC\_10K\_multi\_0569\_00`

- **Question:** What is the amount of mortgage loans 90 or more days past due and insured by U.S. government agencies at December 31, 2024, and at December 31, 2025?
- **Ground truth:** \$121 million and \$198 million
- ✓ **Baseline (vanilla RAG) (PASS):** Mortgage loans 90 or more days past due and insured by U.S. government agencies were:
 - **December 31, 2024:** \$121 million
 - **December 31, 2025:** \$198 million
- ✓ **VERITROOPER (PASS):** At December 31, 2024, the amount was \$121 million; at December 31, 2025, it was \$198 million.

302. PASS · Cross-Reference · `SEC\_10K\_multi\_0587\_01`

- **Question:** How does the Firm obtain credit protection, and what type of notes are used for this purpose?
- **Ground truth:** The Firm obtains credit protection through the issuance of credit-related notes.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm obtains credit protection by purchasing credit derivatives, primarily credit default swaps (CDS), referencing identical or similar entities or indices. It also uses **credit-related notes**, which are funded derivatives containing a credit-risk component, including notes issued to protect certain retained loans.
- ✓ **VERITROOPER (PASS):** The Firm purchases protection through credit derivatives and obtains protection against certain retained loan portfolio loans by issuing credit-related notes.

303. PASS · Cause & Effect · `SEC\_10K\_multi\_0002\_00`

- **Question:** What happens because the Company competes in highly competitive markets?
- **Ground truth:** The Company faces aggressive price competition, downward pressure on gross margins, and continual improvement in product performance.
- ✓ **Baseline (vanilla RAG) (PASS):** Because the Company operates in highly competitive markets, it faces aggressive pricing, pressure on gross margins, rapid technological change, frequent product introductions, and short product life cycles. It must continually invest in R&D and successfully introduce innovative products and services. Failure to compete effectively could materially harm its business, reputation, operating results, financial condition, and stock price.
- ✓ **VERITROOPER (PASS):** The Company faces aggressive price competition, downward pressure on gross margins, frequent product introductions, short product life cycles, evolving standards and rapid technological advances.

304. PASS · Conditional · `SEC\_10K\_multi\_1035\_01`

- **Question:** If brokerage commissions are collected upon occurrence of the client transaction, what is reported net against commission revenue?
- **Ground truth:** Certain costs paid to third-party clearing houses and exchanges.
- ✓ **Baseline (vanilla RAG) (PASS):** Certain costs paid to third-party clearing houses and exchanges are reported net against commission revenue.
- ✓ **VERITROOPER (PASS):** Certain costs paid to third-party clearing houses and exchanges are reported net against commission revenue.

305. PASS · Precision · `SEC\_10K\_multi\_0304\_01`

- **Question:** What was the value of mortgage loans 90 or more days past due and insured by U.S. government agencies at December 31, 2024?
- **Ground truth:** 84 million
- ✓ **Baseline (vanilla RAG) (PASS):** At December 31, 2024, the value was **\$121 million**.
- ✓ **VERITROOPER (PASS):** \$84 million.

306. FAIL · Cross-Reference · `SEC\_10K\_multi\_0055\_01`

BAD QUESTION

- **Question:** What is the percentage increase in revenue for fiscal year 2026 compared to the previous year, and how does it relate to the growth in data center revenue?
- **Ground truth:** 65% increase in revenue, which is higher than the 68% growth in data center revenue.
- ✗ **Baseline (vanilla RAG) (FAIL):** Fiscal year 2026 revenue increased **65%** to **\$215.9 billion** from \$130.5 billion. Data Center revenue grew slightly faster, rising **68% year over year**, driven by accelerated computing, AI, and strong Blackwell demand.

- ✓ **VERITROOPER (PASS)**: Revenue increased 65% in fiscal year 2026 compared with the prior year, while Data Center revenue grew slightly faster at 68%.

307. PASS · Precision · `SEC\_10K\_multi\_1451\_00`

- **Question**: What is the income from interest income from subsidiaries in 2023?
- **Ground truth**: 1,166
- ✓ **Baseline (vanilla RAG) (PASS)**: Interest income from subsidiaries in 2023 was **\$1,166 million**.
- ✓ **VERITROOPER (PASS)**: Interest income from subsidiaries in 2023 was **\$1,166 million**.

308. FAIL · Calculation · `SEC\_10K\_multi\_0929\_01`

BAD QUESTION

- **Question**: If the gross long notional derivative contracts are \$500 and the gross short notional derivative contracts are \$300, what is the total?
- **Ground truth**: 200
- ✗ **Baseline (vanilla RAG) (FAIL)**: 800
- ✗ **VERITROOPER (FAIL)**: 800

309. FAIL · Precision · `SEC\_10K\_multi\_0551\_01`

TABLE MISREAD

- **Question**: What was the total nonaccrual loans at December 31, 2024?
- **Ground truth**: 8,837
- ✗ **Baseline (vanilla RAG) (FAIL)**: Total nonaccrual loans at December 31, 2024 were **\$4.760 billion**.
- ✓ **VERITROOPER (PASS)**: Total nonaccrual loans at December 31, 2024 were **\$8,837 million**.

310. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0714\_00`

- **Question**: What is the specific date by which the Firm must finalize its evaluation of the new guidance?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not specify a date by which the Firm must finalize its evaluation of the new guidance. It only provides the applicable required effective dates and states that the Firm is still evaluating the impact and planned adoption date.
- ✓ **VERITROOPER (PASS)**: The evidence does not specify a date by which the Firm must finalize its evaluation.

311. PASS · Precision · `SEC\_10K\_multi\_0678\_00`

- **Question**: What is the year in which the JPMorgan Chase & Co. Form 10-K is filed?
- **Ground truth**: 2025
- ✓ **Baseline (vanilla RAG) (PASS)**: 2026.
- ✓ **VERITROOPER (PASS)**: The JPMorgan Chase & Co. Form 10-K is for 2025.

312. PASS · Precision · `SEC\_10K\_multi\_0062\_00`

- **Question**: What is the number of the Table of Contents?
- **Ground truth**: 46
- ✓ **Baseline (vanilla RAG) (PASS)**: 5
- ✓ **VERITROOPER (PASS)**: The Table of Contents number is 47.

313. PASS · Precision · `SEC\_10K\_multi\_1073\_01`

- **Question:** What is the total investment securities, net of allowance for credit losses value?
- **Ground truth:** 777,360
- ✓ **Baseline (vanilla RAG) (PASS):** The total investment securities portfolio, net of allowance for credit losses, was **\$774.0 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total investment securities, net of allowance for credit losses value is **\$777,360**.

314. PASS · Cross-Reference · `SEC\_10K\_multi\_1528\_01`

- **Question:** What type of securities are presented in managed basis results on a level comparable to taxable investments and securities?
- **Ground truth:** tax-exempt securities
- ✓ **Baseline (vanilla RAG) (PASS):** Tax-exempt securities.
- ✓ **VERITROOPER (PASS):** Tax-exempt securities.

315. PASS · Cross-Reference · `SEC\_10K\_multi\_0587\_00`

- **Question:** What is the basis for the notional amounts presented, and what is the nature of the counterparties used in credit derivatives?
- **Ground truth:** The notional amounts are presented on a net basis by underlying reference entity, and the counterparties used are predominantly investment-grade.
- ✓ **Baseline (vanilla RAG) (PASS):** The notional amounts are presented on a net basis by underlying reference entity, considering net protection purchased or sold for each reference entity or index. Predominantly all credit derivatives for which the Firm purchased protection for credit portfolio management activities are executed with investment-grade counterparties.
- ✓ **VERITROOPER (PASS):** The notional amounts are presented on a net basis by underlying reference entity. Predominantly all credit derivatives where the Firm purchased protection for credit portfolio management were executed with investment-grade counterparties.

316. PASS · Precision · `SEC\_10K\_multi\_1077\_01`

- **Question:** What is the allowance for credit losses on investment securities as of December 31, 2025?
- **Ground truth:** \$106 million
- ✓ **Baseline (vanilla RAG) (PASS):** The allowance for credit losses on investment securities was **\$106 million** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The allowance for credit losses on investment securities was **\$106 million** as of December 31, 2025.

317. FAIL · Cross-Reference · `SEC\_10K\_multi\_1497\_01`

VALIDATOR OVER-FIRED

- **Question:** What was the average interest rate earned on securities borrowed in U.S. operations for the year ended December 31, 2025, and what was the average balance of securities borrowed in U.S. operations for the same year?
- **Ground truth:** 3.92% and \$185,559
- ✗ **Baseline (vanilla RAG) (FAIL):** The average interest rate was **3.92%**, and the average balance was **\$185.559 billion** (or **\$185,559 million**).
- ✓ **VERITROOPER (PASS):** The average interest rate was 3.92%, and the average balance was \$185,559 million.

318. PASS · Precision · `SEC\_10K\_multi\_0860\_00`

- **Question:** What is the carrying value of cash and due from banks at December 31, 2025?
- **Ground truth:** 21.7
- ✓ **Baseline (vanilla RAG) (PASS):** The carrying value was **\$21.7 billion** at December 31, 2025.
- ✓ **VERITROOPER (PASS):** The carrying value of cash and due from banks at December 31, 2025 was **\$21.7 billion**.

319. PASS · Exception · `SEC\_10K\_multi\_0366\_01`

- **Question:** What does the principal transactions revenue include, excluding?
- **Ground truth:** inventory-related revenue
- ✓ **Baseline (vanilla RAG) (PASS):** Principal transactions revenue includes:
 - Revenue recognized from executing new transactions with market participants;
 - Inventory-related realized and unrealized gains and losses on derivatives, debt and equity instruments, and commodities;
 - Gains and losses on qualifying hedge-accounting derivatives and derivatives used to manage credit, foreign-exchange and interest-rate risks; and
 - Fair-value changes on instruments elected under the fair value option.

In the fair-value-option table, related risk-management instruments required to be measured at fair value are excluded.

- ✓ **VERITROOPER (PASS):** Principal transactions revenue includes amounts recognized from executing new transactions and inventory-related gains and losses. The evidence does not specify an explicit exclusion.

320. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0134\_00`

- **Question:** What is the maximum number of employees that can be affected by a cybersecurity breach?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a maximum number of employees who could be affected by a cybersecurity breach.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a maximum number of employees who can be affected by a cybersecurity breach.

321. PASS · Cause & Effect · `SEC\_10K\_multi\_0693\_01`

- **Question:** What is the result of the Firm continuing to monitor its exposure to Russia?
- **Ground truth:** The Firm currently believes that its remaining exposure to Russia is not material
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm believes its remaining exposure to Russia-primarily cash placed with the central bank-is not material.
- ✓ **VERITROOPER (PASS):** The Firm believes its remaining exposure to Russia is not material.

322. PASS · Cause & Effect · `SEC\_10K\_multi\_0013\_00`

- **Question:** What happens as a consequence of higher net sales of iPhone?
- **Ground truth:** Americas net sales increase
- ✓ **Baseline (vanilla RAG) (PASS):** Higher iPhone net sales contributed to increased 2025 net sales in the Americas, Europe, Japan, and Rest of Asia Pacific compared with 2024.
- ✓ **VERITROOPER (PASS):** Higher iPhone net sales contributed to increased net sales in the Americas, Europe, Japan, and Rest of Asia Pacific during 2025.

323. PASS · Cause & Effect · `SEC\_10K\_multi\_0136\_01`

- **Question:** What could result from JPMorganChase's failure to address changing regulatory and market environments?
- **Ground truth:** JPMorganChase's growth prospects could be negatively impacted.
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase's business values and growth prospects could suffer, and its earnings could decline.
- ✓ **VERITROOPER (PASS):** It could adversely affect JPMorganChase's growth prospects and its ability to develop resources, control expenses, and return capital to shareholders.

324. FAIL · Precision · `SEC\_10K\_multi\_1501\_00`

UNSUPPORTED NUMBER

- **Question:** What is the amount of noninterest-bearing liabilities funding interest-earning assets?
- **Ground truth:** 284,913
- ✗ **Baseline (vanilla RAG) (FAIL):** The amount of noninterest-bearing liabilities funding interest-earning assets was **\$670.426 billion** in 2025.
- ✗ **VERITROOPER (FAIL):** Noninterest-bearing liabilities funding interest-earning assets were \$701.189 billion in 2024 and \$731.939 billion in 2023.

325. FAIL · Precision · `SEC\_10K\_multi\_0487\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the amount of the allowance for credit losses in 2025?
- **Ground truth:** 468 million
- ✗ **Baseline (vanilla RAG) (FAIL):** The allowance for credit losses was **\$31.2 billion as of December 31, 2025**.
- ✓ **VERITROOPER (PASS):** The allowance for credit losses was **\$468 million** as of December 31, 2025.

326. PASS · Precision · `SEC\_10K\_multi\_1610\_01`

- **Question:** What is the date of the 2024 Equity Plan adoption?
- **Ground truth:** May 1, 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The adoption date of the 2024 Equity Plan is not provided in the evidence.
- ✓ **VERITROOPER (PASS):** The 2024 Equity Plan was adopted on May 1, 2024.

327. PASS · Cross-Reference · `SEC\_10K\_multi\_1471\_01`

- **Question:** What is the estimated potential future payment for the shared-loss agreements?
- **Ground truth:** reflected as contingent consideration as part of the purchase price consideration
- ✓ **Baseline (vanilla RAG) (PASS):** The estimated potential future payment was **\$15 million**, recorded as contingent consideration for the shared-loss agreements.
- ✓ **VERITROOPER (PASS):** The evidence does not provide an amount for the estimated potential future payment; it only states that it is reflected as contingent consideration in the purchase price.

328. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1537\_00`

HALLUCINATION

- **Question:** What is the maximum number of employees covered by collective bargaining agreements at our key manufacturing locations?

- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Approximately 400 employees in North America were covered by collective bargaining agreements as of December 31, 2025.
- **✓ VERITROOPER (PASS):** The evidence does not specify the maximum number of employees covered by collective bargaining agreements.

329. PASS · Cross-Reference · `SEC\_10K\_multi\_1265\_00`

- **Question:** What is the total amount of goodwill attributed to the Consumer & Community Banking segment as of December 31, 2025, and how does it compare to the amount in 2024?
- **Ground truth:** The total amount of goodwill attributed to the Consumer & Community Banking segment as of December 31, 2025, is \$32,116 million, which is the same as the amount in 2024.
- **✓ Baseline (vanilla RAG) (PASS):** As of December 31, 2025, goodwill attributed to Consumer & Community Banking was **\$32.116 billion**, unchanged from **\$32.116 billion in 2024**.
- **✓ VERITROOPER (PASS):** Consumer & Community Banking goodwill was \$32,116 million as of December 31, 2025, unchanged from \$32,116 million in 2024.

330. PASS · Cause & Effect · `SEC\_10K\_multi\_0135\_01`

- **Question:** What could result from a significant inadequacy in disclosure or financial reporting controls?
- **Ground truth:** JPMorganChase's business, operations, and reputation could be negatively affected.
- **✓ Baseline (vanilla RAG) (PASS):** A significant inadequacy could lead to inaccurate financial reporting, which may:
 - Materially harm business, operations, financial condition, or results
 - Restrict access to capital markets
 - Require significant remediation resources
 - Expose JPMorganChase to litigation and penalties
 - Cause reputational harm
- **✓ VERITROOPER (PASS):** It could cause inaccurate financial reporting, adversely affect business and financial condition, restrict capital-market access, require significant remediation costs, expose JPMorganChase to litigation and penalties, and harm its reputation.

331. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_00`

- **Question:** What is the total number of notes issued by Apple Inc. in 2025?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The provided evidence does not specify the total number of notes issued by Apple Inc. in 2025.
- **✗ VERITROOPER (FAIL):** Apple Inc. issued four notes in 2025.

332. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0159\_02`

- **Question:** How many shares were issued to non-employee directors in 2025?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify how many shares were issued to non-employee directors in 2025.
- **✓ VERITROOPER (PASS):** The evidence does not state how many shares were issued to non-employee directors in 2025.

333. PASS · Cause & Effect · `SEC\_10K\_multi\_0693\_00`

- **Question:** What happens as a consequence of increased investment and market-making securities activities?
- **Ground truth:** An increase in exposure to the United Kingdom
- ✓ **Baseline (vanilla RAG) (PASS):** They lead to higher holdings of government debt securities and, consequently, increased exposure to the United Kingdom.
- ✓ **VERITROOPER (PASS):** Higher holdings of U.K. government debt securities increased the Firm's exposure to the United Kingdom.

334. PASS · Cross-Reference · `SEC\_10K\_multi\_1530\_01`

- **Question:** How many shares of the Registrant's Common Stock were outstanding as of February 18, 2026?
- **Ground truth:** 4,300,723,069
- ✓ **Baseline (vanilla RAG) (PASS):** As of February 18, 2026, **4,300,723,069 shares** of the Registrant's Common Stock were outstanding.
- ✓ **VERITROOPER (PASS):** As of February 18, 2026, 4,300,723,069 shares were outstanding.

335. PASS · Exception · `SEC\_10K\_multi\_0366\_00`

- **Question:** What does the global investment banking fees exclude?
- **Ground truth:** money market, short-term debt and shelf securities
- ✓ **Baseline (vanilla RAG) (PASS):** Global investment banking fees exclude money market, short-term debt, and shelf securities.
- ✓ **VERITROOPER (PASS):** Global investment banking fees exclude money market, short-term debt, and shelf securities.

336. PASS · Cross-Reference · `SEC\_10K\_multi\_0653\_01`

- **Question:** What types of assets does CIB trade?
- **Ground truth:** Trading assets/liabilities-debt and marketable equity instruments, and derivatives, including hedges of the retained loan portfolio
- ✓ **Baseline (vanilla RAG) (PASS):** CIB trades fixed-income, foreign-exchange, commodities, and equity assets.
- ✓ **VERITROOPER (PASS):** CIB trades debt and marketable equity instruments, derivatives, certain securities under resale agreements, and borrowed securities.

337. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** How many lawsuits are pending against the company at any given time?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Approximately 15,000 active claims are pending against Aqua-Chem.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many lawsuits are pending against the company at any given time.

338. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_06`

- **Question:** What is the total amount of fines paid by the company in the past year?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose any fines paid by the company in the past year.

- ✓ **VERITROOPER (PASS):** The evidence does not state the total amount of fines paid by the company in the past year.

339. PASS · Conditional · `SEC\_10K\_multi\_0633\_01`

- **Question:** If the Firm's risk management process for derivatives exposures takes into consideration the potential impact of wrong-way risk, does the Firm adjust the CVA associated with a particular counterparty's AVG?
- **Ground truth:** Yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. When wrong-way risk correlations are identified, the Firm may adjust the CVA associated with a particular counterparty's average exposure (AVG).
- ✓ **VERITROOPER (PASS):** Yes. When wrong-way risk correlations are identified, the Firm may adjust the CVA associated with a particular counterparty's AVG.

340. PASS · Conditional · `SEC\_10K\_multi\_0717\_00`

- **Question:** If the Firm adopts the new hedge accounting guidance, when will it be effective?
- **Ground truth:** January 1, 2027
- ✓ **Baseline (vanilla RAG) (PASS):** The new hedge accounting guidance will be effective January 1, 2027. Early adoption is permitted.
- ✓ **VERITROOPER (PASS):** The required effective date is January 1, 2027.

341. PASS · Cross-Reference · `SEC\_10K\_multi\_0721\_01`

- **Question:** What is the principal consideration for our determination that performing procedures relating to the fair value of certain level 3 financial instruments is a critical audit matter?
- **Ground truth:** significant judgment by management when developing the fair value estimate of certain level 3 financial instruments
- ✓ **Baseline (vanilla RAG) (PASS):** The principal considerations were:
 1. Significant management judgment in developing the fair value estimates.
 2. A high degree of auditor judgment, subjectivity, and effort in evaluating the significant unobservable inputs.
 3. The need for professionals with specialized skills and knowledge.
- ✓ **VERITROOPER (PASS):** The principal considerations were management's significant judgment in estimating fair value, the high degree of auditor judgment, subjectivity, and effort in evaluating unobservable inputs, and the need for professionals with specialized skills and knowledge.

342. PASS · Cross-Reference · `SEC\_10K\_multi\_1062\_01`

- **Question:** What is the approved asset allocation range for debt securities in the Firm's principal defined benefit plan as of December 31, 2025?
- **Ground truth:** 41% - 100%
- ✓ **Baseline (vanilla RAG) (PASS):** The approved asset allocation range for debt securities was **41%-100%**.
- ✓ **VERITROOPER (PASS):** The approved asset allocation range for debt securities was 41% to 100%.

343. PASS · Precision · `SEC\_10K\_multi\_1492\_00`

- **Question:** What is the value of 88,726?
- **Ground truth:** 88,726
- ✓ **Baseline (vanilla RAG) (PASS):** 88,726 million equals **\$88.726 billion**.
- ✓ **VERITROOPER (PASS):** The value is 88,726.

344. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0576\_01`

- **Question:** What is the percentage of credit card loans that are retained in the Firm's portfolio at December 31, 2026?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide data for December 31, 2026. At December 31, 2025, all credit card loans were retained: \$247.797 billion retained out of \$247.797 billion total, or 100%.
- ✓ **VERITROOPER (PASS):** The evidence does not provide information for December 31, 2026.

345. FAIL · Precision · `SEC\_10K\_multi\_0060\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the amount of cash and cash equivalents as of January 25, 2026?
- **Ground truth:** 10,605
- ✗ **Baseline (vanilla RAG) (FAIL):** Cash and cash equivalents were **\$10.605 billion** as of January 25, 2026.
- ✗ **VERITROOPER (FAIL):** Cash and cash equivalents were \$10.605 billion as of January 25, 2026.

346. PASS · Precision · `SEC\_10K\_multi\_0592\_00`

- **Question:** What is the threshold for Transportation?
- **Ground truth:** 285
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a threshold for Transportation.
- ✓ **VERITROOPER (PASS):** The evidence does not identify which column represents the threshold for Transportation.

347. PASS · Exception · `SEC\_10K\_multi\_1533\_00`

- **Question:** What circumstance does the general rule of compliance with all applicable laws NOT apply?
- **Ground truth:** In countries and territories where our Company does not do business.
- ✓ **Baseline (vanilla RAG) (PASS):** The Proposition 65 warning requirement does not apply when exposure is below an established safe-harbor level or another exemption applies.
- ✓ **VERITROOPER (PASS):** The evidence states no exception; the Company's compliance policy applies in all countries and territories where it does business.

348. PASS · Exception · `SEC\_10K\_multi\_0045\_00`

- **Question:** What circumstance may lead to a shortage of necessary resources impacting future revenue and financial performance?
- **Ground truth:** The availability of data centers, energy, and capital to support the buildout of NVIDIA AI infrastructure by our customers and partners.
- ✓ **Baseline (vanilla RAG) (PASS):** Delays in expanding energy capacity-due to regulatory, technical, and construction challenges-or constrained access to capital for infrastructure projects may cause shortages of data centers, energy, and capital, impacting future revenue and financial performance.
- ✓ **VERITROOPER (PASS):** Delays in expanding energy capacity or constrained access to capital-especially for less-capitalized companies-may create resource shortages that delay deployments and impact future revenue and financial performance.

349. PASS · Conditional · `SEC\_10K\_multi\_0729\_00`

- **Question:** If observable market information is available for beneficial interests issued by consolidated VIEs, what valuation method is used?
- **Ground truth:** Valued using observable market information.
- ✓ **Baseline (vanilla RAG) (PASS):** They are valued using observable market information.
- ✓ **VERITROOPER (PASS):** They are valued using observable market information.

350. PASS · Cross-Reference · `SEC\_10K\_multi\_0864\_01`

- **Question:** What is the carrying value of wholesale lending-related commitments at December 31, 2025?
- **Ground truth:** 3.2
- ✓ **Baseline (vanilla RAG) (PASS):** The carrying value was **\$3.2 billion** at December 31, 2025.
- ✓ **VERITROOPER (PASS):** The carrying value was **\$3.2 billion** at December 31, 2025.

351. PASS · Conditional · `SEC\_10K\_multi\_0242\_01`

- **Question:** If the net addition to the allowance for credit losses in consumer was \$3.3 billion, and this was driven by \$2.2 billion related to the Apple Card transaction, what was the remaining amount driven by?
- **Ground truth:** loan growth in Card Services and the impact of changes in the Firm's weighted-average macroeconomic outlook
- ✓ **Baseline (vanilla RAG) (PASS):** The remaining **\$1.1 billion** was driven by **loan growth in Card Services and changes in the firm's weighted-average macroeconomic outlook**, partially offset by reduced borrower uncertainty.
- ✓ **VERITROOPER (PASS):** The remaining \$1.1 billion was driven by loan growth in Card Services and changes in the firm's weighted-average macroeconomic outlook, partially offset by reduced borrower uncertainty.

352. FAIL · Precision · `SEC\_10K\_multi\_1025\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total notional amount?
- **Ground truth:** 628,130
- ✗ **Baseline (vanilla RAG) (FAIL):** The total derivative notional amount was **\$50.642 trillion (\$50,642 billion)** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total notional amount was **\$628,130 million** as of December 31, 2024.

353. PASS · Cross-Reference · `SEC\_10K\_multi\_1269\_01`

- **Question:** How does the fair value of MSRs change in response to interest rate changes, and what risk management strategies does the Firm employ?
- **Ground truth:** The fair value of MSRs typically decreases when interest rates decline, due to increased prepayments. The Firm uses combinations of derivatives and securities to manage this risk, aiming to offset changes in MSR fair value with changes in related risk management instruments.
- ✓ **Baseline (vanilla RAG) (PASS):** MSR fair value generally **decreases when interest rates decline**, because lower rates tend to accelerate mortgage prepayments and shorten the expected life of servicing cash flows. Conversely, MSR values generally benefit when rates rise and prepayments slow.

The Firm manages this interest-rate risk using a combination of **securities**, such as mortgage-backed securities, and **derivatives**, including instruments that receive fixed-rate interest payments. These positions are intended to offset interest-rate-driven changes in MSR fair value. The Firm measures MSRs at fair value using an option-adjusted spread model incorporating multiple interest-rate scenarios and prepayment

assumptions.

- ✓ **VERITROOPER (PASS):** MSR fair value typically decreases when interest rates decline because prepayments rise, shortening expected servicing cash flows. The Firm manages this exposure using combinations of derivatives and securities whose fair values increase when rates decline, aiming to offset interest-rate-driven changes in MSR value.

354. PASS · Precision · `SEC\_10K\_multi\_0057\_00`

- **Question:** What is the percentage of revenue from sales to customers headquartered outside of the United States for fiscal year 2026?
- **Ground truth:** 31
- ✓ **Baseline (vanilla RAG) (PASS):** 31% of total revenue.
- ✓ **VERITROOPER (PASS):** 31%

355. FAIL · Precision · `SEC\_10K\_multi\_0520\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the total short-term unsecured funding for 2025?
- **Ground truth:** 28,668
- ✗ **Baseline (vanilla RAG) (FAIL):** The total short-term unsecured funding for 2025 was **\$27.341 billion**.
- ✗ **VERITROOPER (FAIL):** The total short-term unsecured funding for 2025 was **\$27.341 billion**.

356. PASS · Precision · `SEC\_10K\_multi\_1136\_01`

- **Question:** At what date were Chapter 7 residential real estate loans evaluated?
- **Ground truth:** December 31, 2025
- ✓ **Baseline (vanilla RAG) (PASS):** December 31, 2025.
- ✓ **VERITROOPER (PASS):** December 31, 2025.

357. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_06`

- **Question:** How many years does the Company's cybersecurity insurance coverage last?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify how many years the cybersecurity insurance coverage lasts.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many years the Company's cybersecurity insurance coverage lasts.

358. PASS · Cause & Effect · `SEC\_10K\_multi\_0136\_00`

- **Question:** What could happen if JPMorganChase fails to devise effective business plans and strategies?
- **Ground truth:** JPMorganChase's ability to develop its resources, control expenses, and return capital to shareholders could be adversely affected.
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase's growth prospects, competitive standing, and earnings could suffer. It may also struggle to develop resources, control expenses, return capital to shareholders, and achieve its strategic objectives.
- ✓ **VERITROOPER (PASS):** It could adversely affect JPMorganChase's ability to develop resources, control expenses, and return capital to shareholders.

359. FAIL · Cross-Reference · `SEC\_10K\_multi\_0781\_01`

TABLE MISREAD

- **Question:** What is the total value of debt instruments and equity securities combined?
- **Ground truth:** $2,144 + 62 = 2,206$
- **✗ Baseline (vanilla RAG) (FAIL):** The combined value was **\$549.156 billion** (\$417.428 billion in debt instruments plus \$131.728 billion in equity securities).
- **✓ VERITROOPER (PASS):** The total value of debt instruments and equity securities combined is \$2,206 ($2,144 + 62$).

360. FAIL · Precision · `SEC\_10K\_multi\_0597\_00`

BAD QUESTION

- **Question:** What is the threshold for Chemicals & Plastics?
- **Ground truth:** 8,152
- **✗ Baseline (vanilla RAG) (FAIL):** No specific threshold is stated for Chemicals & Plastics. Its 2025 wholesale credit exposure was **\$23.79 billion**, including **\$2.091 billion criticized performing** and **\$93 million criticized nonperforming**.
- **✓ VERITROOPER (PASS):** The evidence does not identify which column represents the threshold for Chemicals & Plastics.

361. PASS · Precision · `SEC\_10K\_multi\_1255\_01`

- **Question:** What is the carrying value of loans sold in 2024?
- **Ground truth:** 25,765
- **✓ Baseline (vanilla RAG) (PASS):** The carrying value of loans sold in 2024 was **\$25.765 billion**.
- **✓ VERITROOPER (PASS):** The carrying value of loans sold in 2024 was **\$25,765 million**.

362. PASS · Conditional · `SEC\_10K\_multi\_0133\_01`

- **Question:** If JPMorganChase does not prioritize or complete enhancements to address emerging threats, what could happen?
- **Ground truth:** it could become increasingly vulnerable to cyber attacks
- **✓ Baseline (vanilla RAG) (PASS):** JPMorganChase could become increasingly vulnerable to cyberattacks, potentially exposing its systems and data to unauthorized access, misuse, or compromise.
- **✓ VERITROOPER (PASS):** JPMorganChase could become increasingly vulnerable to cyber attacks.

363. PASS · Precision · `SEC\_10K\_multi\_1550\_00`

- **Question:** What is the percent change in concentrate sales volume worldwide from 2024 to 2025?
- **Ground truth:** 1
- **✓ Baseline (vanilla RAG) (PASS):** Worldwide concentrate sales volume increased **1%** from 2024 to 2025.
- **✓ VERITROOPER (PASS):** Worldwide concentrate sales volume increased by 1% from 2024 to 2025.

364. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0254\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average annual revenue growth rate on a managed basis?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Managed-basis total revenue was:

- **2023:** \$162.366 billion
- **2025:** \$185.581 billion

Average annual growth rate (CAGR), 2023-2025:

$$\sqrt[3]{\frac{185.581}{162.366}} - 1 \approx \mathbf{6.9\%}$$

- **✓ VERITROOPER (PASS):** The evidence does not provide managed-basis revenue amounts, so the average annual revenue growth rate cannot be calculated.

365. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1554\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What was the percentage of the net operating revenue increase attributed to acquisitions and divestitures in the Asia Pacific operating segment?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Acquisitions and divestitures had a **1% unfavorable impact** on Asia Pacific net operating revenues in 2025 versus 2024.
- **✓ VERITROOPER (PASS):** The evidence does not provide the percentage attributed to acquisitions and divestitures for the Asia Pacific operating segment.

366. PASS · Exception · `SEC\_10K\_multi\_0277\_00`

- **Question:** What does not apply when calculating TCE?
- **Ground truth:** tax-deductible goodwill and identifiable intangibles created in nontaxable transactions
- **✓ Baseline (vanilla RAG) (PASS):** Mortgage servicing rights (MSRs) are not deducted when calculating tangible common equity (TCE).
- **✓ VERITROOPER (PASS):** The evidence does not specify what does not apply when calculating TCE.

367. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0727\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the specific percentage of loans carried at fair value that are classified as level 2?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Approximately **95.7%** of loans carried at fair value are classified as Level 2:

$$\frac{\$70.684\text{B} - \$3.062\text{B}}{\$70.684\text{B}} \times 100 \approx 95.7\%$$

- **✓ VERITROOPER (PASS):** The evidence does not specify the percentage of loans carried at fair value classified as Level 2.

368. PASS · Exception · `SEC\_10K\_multi\_1049\_01`

- **Question:** What is reported in principal transactions revenue for financial instruments?
- **Ground truth:** all changes in fair value including any interest elements
- **✓ Baseline (vanilla RAG) (PASS):** Principal transactions revenue reports:
- Realized and unrealized gains and losses on financial instruments and commodities transactions, including instruments accounted for under the fair value option.

- Gains and losses on derivatives used in qualifying hedge relationships or for specific risk-management purposes.
- Realized gains and losses from instrument-specific credit risk for fair-value-option liabilities; related unrealized gains and losses are reported in OCI.
- ✓ **VERITROOPER (PASS)**: For derivatives and certain financial instruments containing embedded derivatives, all fair-value changes, including interest elements, are primarily reported in principal transactions revenue.

369. PASS · Exception · `SEC\_10K\_multi\_0110\_00`

- **Question**: What circumstance does the Federal Reserve's authority to limit an FHC's ability NOT apply to?
- **Ground truth**: if the FHC or any of its depository institution subsidiaries ceases to meet applicable eligibility requirements
- ✓ **Baseline (vanilla RAG) (PASS)**: It does not apply when the FHC and all its depository institution subsidiaries continue to meet the applicable eligibility requirements.
- ✓ **VERITROOPER (PASS)**: It does not apply when the FHC and all its depository institution subsidiaries continue to meet applicable eligibility requirements.

370. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0042\_02`

- **Question**: What is the name of the specific inspection process required for H200 products before shipment to China?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not specify a name; it only states that H200 products must undergo "an inspection process in the United States" before shipment.
- ✓ **VERITROOPER (PASS)**: The evidence does not specify the name of the inspection process.

371. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1531\_05`

- **Question**: How many distributors does the Company have in the United States?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not specify how many distributors the Company has in the United States.
- ✓ **VERITROOPER (PASS)**: The evidence does not state how many distributors the Company has in the United States.

372. PASS · Conditional · `SEC\_10K\_multi\_1058\_00`

- **Question**: If the net gain exceeds 10% of the greater of the projected benefit obligation or the fair value of the plan assets, what is the amortization period?
- **Ground truth**: Generally over the average expected remaining lifetime of plan participants.
- ✓ **Baseline (vanilla RAG) (PASS)**: The net gain is generally amortized over the average expected remaining lifetime of plan participants, given the frozen status of most plans.
- ✓ **VERITROOPER (PASS)**: The net gain is generally amortized over the average expected remaining lifetime of plan participants, given the frozen status of most plans.

373. PASS · Precision · `SEC\_10K\_multi\_0037\_00`

- **Question**: What is the filing date/period end date for Exhibit 101?
- **Ground truth**: 9/28/24
- ✓ **Baseline (vanilla RAG) (PASS)**: Exhibit 101 was filed with the 2025 Form 10-K on **October 31, 2025**.
- ✓ **VERITROOPER (PASS)**: The evidence does not provide a filing date or period end date for Exhibit 101.

374. PASS · Exception · `SEC\_10K\_multi\_0507\_00`

- **Question:** What is the exception to the effective date of the Basel III standards?
- **Ground truth:** market risk aspects
- ✓ **Baseline (vanilla RAG) (PASS):** The exception is the **market risk aspects**, whose effective date is **January 1, 2027**, rather than January 1, 2025.
- ✓ **VERITROOPER (PASS):** The exception is the market risk aspects, which take effect on January 1, 2027.

375. PASS · Cause & Effect · `SEC\_10K\_multi\_1541\_00`

- **Question:** What happens as a result of a cybersecurity incident?
- **Ground truth:** Unauthorized or accidental access to or disclosure of material confidential information or regulated personal data.
- ✓ **Baseline (vanilla RAG) (PASS):** A cybersecurity incident can cause unauthorized access or disclosure of confidential information, data destruction, system or service disruptions, and other operational damage. It may also lead to financial and reputational harm, legal action, regulatory investigations and fines, costly remediation, and potentially a material adverse effect on business operations, strategy, financial condition, and results.
- ✓ **VERITROOPER (PASS):** A cybersecurity incident could disrupt operations, delay financial reporting, reduce revenue and profits, expose confidential or personal data, damage reputation, trigger legal and regulatory actions and fines, and require significant remediation costs.

376. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1614\_01`

HALLUCINATION

- **Question:** What is the date of the next annual report submission?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The next annual submission is due **April 6, 2026**.
- ✓ **VERITROOPER (PASS):** The evidence does not state the date of the next annual report submission.

377. PASS · Cross-Reference · `SEC\_10K\_multi\_0543\_00`

- **Question:** What is the relationship between credit risk management and the stability of JPMorgan Chase's credit ratings?
- **Ground truth:** Credit risk management is critical in maintaining high credit ratings, which are influenced by factors such as a stable and diverse earnings stream, strong capital and liquidity ratios, and diverse funding sources.
- ✓ **Baseline (vanilla RAG) (PASS):** Strong credit risk management supports stable, high credit ratings. Rating agencies consider JPMorganChase's risk management practices, credit quality, capital and liquidity ratios, earnings stability, and funding diversity. Deteriorating credit quality-such as increased delinquencies or client downgrades-or weaker risk controls could increase risk-weighted assets and stress losses and lead to a ratings downgrade. Such a downgrade could raise funding costs, restrict capital-market access, trigger collateral requirements, and reduce willing counterparties.
- ✓ **VERITROOPER (PASS):** Strong credit quality and risk management controls are critical to maintaining high credit ratings. Because rating agencies evaluate risk management practices, weaknesses in credit risk management could lead to rating changes.

378. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1149\_03`

- **Question:** What percentage of the credit card portfolio is comprised of prime consumer borrowers?
- **Ground truth:** UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify what percentage of the credit card portfolio comprises prime consumer borrowers. It only states that the Firm generally originates new credit card accounts to prime consumers.
- ✓ **VERITROOPER (PASS):** The evidence does not specify what percentage of the credit card portfolio comprises prime consumer borrowers.

379. PASS · Cause & Effect · `SEC\_10K\_multi\_1554\_00`

- **Question:** What happens as a consequence of the refranchising of bottling operations in the Philippines, Bangladesh, and certain territories in India?
- **Ground truth:** Bottling Investments' unit case volume decreases by 8%.
- ✓ **Baseline (vanilla RAG) (PASS):** The refranchising primarily caused Bottling Investments' unit case volume to decline 8%. It also contributed to the increase in gross profit margin from 61.1% in 2024 to 61.6% in 2025.
- ✓ **VERITROOPER (PASS):** Unit case volume for Bottling Investments decreased 8%.

380. PASS · Precision · `SEC\_10K\_multi\_0852\_01`

- **Question:** What is the amount of level 3 assets related to equity securities adjusted based on the measurement alternative?
- **Ground truth:** 721 million
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, **\$721 million** of Level 3 assets related to equity securities adjusted under the measurement alternative.
- ✓ **VERITROOPER (PASS):** \$721 million.

381. PASS · Cause & Effect · `SEC\_10K\_multi\_1593\_01`

- **Question:** What is the potential impact on the Company's financial position if the IRS' positions are upheld?
- **Ground truth:** The Company would not receive a refund of the applicable portion or all of the \$6.0 billion it paid, and it would likely be subject to significant additional liabilities for subsequent years.
- ✓ **Baseline (vanilla RAG) (PASS):** If the IRS' positions are upheld, the Company could:
 - Lose some or all of the \$6.0 billion tax payment made in September 2024 and the related \$385 million interest receivable.
 - Face approximately \$14 billion in additional tax and interest liabilities for 2010-2025, with further interest continuing to accrue.
 - Incur additional liabilities for subsequent years; continued application in 2026 could increase its effective tax rate by approximately 3.5%.

These outcomes could materially adversely affect the Company's financial position, results of operations, and cash flows.

- ✓ **VERITROOPER (PASS):** If upheld, the Company could lose some or all of its \$6.0 billion tax payment and related \$385 million interest receivable, while facing significant additional liabilities for subsequent years-potentially about \$14 billion for 2010-2025-materially adversely affecting its financial position, results, and cash flows.

382. PASS · Exception · `SEC\_10K\_multi\_1186\_00`

- **Question:** What is excluded from the consumer loans?
- **Ground truth:** credit card
- ✓ **Baseline (vanilla RAG) (PASS):** Consumer loans exclude operating lease assets-\$20.0 billion at December 31, 2025 and \$12.8 billion at December 31, 2024-which are reported in other assets.
- ✓ **VERITROOPER (PASS):** Credit card loans are excluded from consumer loans.

383. FAIL · Cross-Reference · `SEC\_10K\_multi\_0503\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the net capital of J.P. Morgan Securities as of December 31, 2025, and how does it compare to the minimum requirements?
- **Ground truth:** \$27,196 million, which is greater than the minimum of \$6,559 million.
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, J.P. Morgan Securities had **net capital of \$27.196 billion**, compared with a **minimum requirement of \$6.559 billion**-an excess of **\$20.637 billion**.
- **✗ VERITROOPER (FAIL):** As of December 31, 2025, J.P. Morgan Securities had net capital of \$27.196 billion, compared with the minimum requirement of \$6.559 billion-\$20.637 billion above the minimum.

384. FAIL · Cross-Reference · `SEC\_10K\_multi\_1588\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the total revenue from concentrate operations in the United States for the year ended December 31, 2025, and how does it compare to the total revenue from finished product operations in the same period?
- **Ground truth:** The total revenue from concentrate operations in the United States for the year ended December 31, 2025, is \$8,956 million, and the total revenue from finished product operations is \$10,171 million.
- **✗ Baseline (vanilla RAG) (FAIL):** For 2025 in the United States:
- **Concentrate operations revenue: \$8,956 million**
- **Finished product operations revenue: \$10,171 million**

Finished product operations generated **\$1,215 million more** revenue than concentrate operations.

- **✗ VERITROOPER (FAIL):** For 2025, U.S. concentrate operations generated \$8,956 million, compared with \$10,171 million from finished product operations. Finished product revenue was \$1,215 million higher.

385. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1554\_00`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What was the total concentrate sales volume growth rate for the Asia Pacific operating segment?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The Asia Pacific operating segment's concentrate sales volume was **even (0%)** in 2025 versus 2024. After structural changes, it grew **1%**.
- **✓ VERITROOPER (PASS):** The evidence does not state the total rate; it only says Asia Pacific concentrate sales volume grew 1% after considering structural changes.

386. FAIL · Cross-Reference · `SEC\_10K\_multi\_1566\_01`

VALIDATOR OVER-FIRED

- **Question:** How did the operating margin in 2025 compare to the operating margin in 2024 for the EMEA operating segment?
- **Ground truth:** The operating margin in 2025 was 39.7%, while in 2024 it was 41.4%. This indicates a decrease in operating margin for the EMEA operating segment from 2024 to 2025.
- **✗ Baseline (vanilla RAG) (FAIL):** The EMEA operating margin decreased from 41.4% in 2024 to 39.7% in 2025, a decline of 1.7 percentage points.
- **✗ VERITROOPER (FAIL):** EMEA's operating margin decreased to 39.7% in 2025 from 41.4% in 2024, a decline of 1.7 percentage points.

387. FAIL · Cross-Reference · `SEC\_10K\_multi\_1286\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total amortization expense for the years ended December 31, 2025 and 2024, and how does it relate to the net carrying values of other intangible assets?
- **Ground truth:** The total amortization expense was \$292 million for 2025 and \$339 million for 2024, which relates to the net carrying values of \$1.3 billion and \$1.7 billion, respectively.
- **✗ Baseline (vanilla RAG) (FAIL):** Amortization expense totaled **\$631 million: \$292 million in 2025 plus \$339 million in 2024.**

This expense applies to finite-lived intangible assets, whose net carrying values were **\$1.3 billion in 2025** and **\$1.7 billion in 2024**. Indefinite-lived intangible assets-**\$1.3 billion and \$1.2 billion**, respectively-are not amortized. Total net other intangible assets were therefore **\$2.6 billion in 2025** and **\$2.9 billion in 2024**.

- **✗ VERITROOPER (FAIL):** Total amortization expense was \$631 million-\$292 million in 2025 and \$339 million in 2024. Net other intangible assets were \$2.6 billion in 2025 and \$2.9 billion in 2024, comprising finite-lived assets subject to amortization and indefinite-lived assets not subject to amortization.

388. FAIL · Precision · `SEC\_10K\_multi\_1156\_01`

BAD QUESTION

- **Question:** What is the total retained loans for noninvestment-grade risk ratings as of December 31, 2024?
- **Ground truth:** 162,432
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2024, total retained loans rated noninvestment-grade were **\$218.726 billion**.
- **✗ VERITROOPER (FAIL):** As of December 31, 2024, noninvestment-grade retained loans totaled \$48.152 billion for loans secured by real estate and \$95.928 billion for commercial and industrial loans.

389. PASS · Exception · `SEC\_10K\_multi\_1583\_00`

- **Question:** What does not apply to the Company's hedging activities?
- **Ground truth:** Trading purposes
- **✓ Baseline (vanilla RAG) (PASS):** Hedge accounting does not apply to all derivatives used in the Company's risk-management and hedging activities.
- **✓ VERITROOPER (PASS):** Entering into derivative financial instruments for trading purposes does not apply to the Company's hedging activities.

390. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1274\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of servicer advances at December 31?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Servicer advances totaled **\$493 million at December 31, 2025**, net of an allowance for uncollectible amounts.
- **✓ VERITROOPER (PASS):** The evidence does not provide the total value of servicer advances at December 31.

391. PASS · Cross-Reference · `SEC\_10K\_multi\_1140\_00`

- **Question:** What is the frequency at which current property values are estimated?
- **Ground truth:** quarterly
- **✓ Baseline (vanilla RAG) (PASS):** Current property values are estimated at least quarterly.

- ✓ **VERITROOPER (PASS):** Current property values are estimated at least quarterly.

392. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_06`

- **Question:** What is the average age of NVIDIA's board of directors?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide the ages of NVIDIA's full board of directors, so their average age cannot be determined. It states that director information will appear in the 2026 Proxy Statement.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the ages of NVIDIA's directors, so the average age cannot be determined.

393. PASS · Cross-Reference · `SEC\_10K\_multi\_1587\_01`

- **Question:** What is the total amount of acquisitions of businesses, equity method investments, and nonmarketable securities in 2025, and how does it compare to the total in 2024?
- **Ground truth:** The total amount of acquisitions in 2025 is \$461 million, which is \$146 million more than the total in 2024, which was \$315 million.
- ✓ **Baseline (vanilla RAG) (PASS):** Acquisitions totaled **\$461 million in 2025**, compared with **\$315 million in 2024**-an increase of **\$146 million**, or approximately **46%**.
- ✓ **VERITROOPER (PASS):** The total was \$461 million in 2025, compared with \$315 million in 2024-an increase of \$146 million.

394. PASS · Cross-Reference · `SEC\_10K\_multi\_0653\_00`

- **Question:** What types of risks are associated with the activities of CIB?
- **Ground truth:** Risk of loss from adverse movements in market prices and implied volatilities across interest rate, foreign exchange, credit, commodity and equity instruments, and Basis and correlation risk from changes in the way asset values move relative to one another.
- ✓ **Baseline (vanilla RAG) (PASS):** CIB's activities are exposed to four broad risk types:
 - **Strategic risk:** Risk to earnings, capital, liquidity or reputation from failed business plans or inadequate responses to operating-environment changes.
 - **Credit and investment risk:** Counterparty or client default, changes in credit profiles, loss of principal or reduced investment returns.
 - **Market risk:** Changes in interest and foreign-exchange rates, equity and commodity prices, credit spreads and implied volatility.
 - **Operational risk:** Failures in processes, systems or human factors, and external events-including cybersecurity, compliance, conduct, legal, estimation and model risks.
- ✓ **VERITROOPER (PASS):** CIB faces risks from adverse movements in market prices and implied volatilities across interest rate, foreign exchange, credit, commodity, and equity instruments. It also faces basis and correlation risk, interest rate risk, and prepayment risk.

395. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0678\_00`

- **Question:** What is the average deposit beta for consumer deposits?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose an average deposit beta for consumer deposits. It only states that consumer deposit betas are a significant assumption in earnings-at-risk sensitivity analyses.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the average deposit beta for consumer deposits.

396. PASS · Cause & Effect · `SEC\_10K\_multi\_1134\_00`

- **Question:** What happens as a consequence of revolving loans being converted to term loans?
- **Ground truth:** They have higher delinquency rates.
- ✓ **Baseline (vanilla RAG) (PASS):** Converting revolving loans to term loans generally results in higher required, fully amortizing payments and consequently higher delinquency rates than loans still within the revolving period.
- ✓ **VERITROOPER (PASS):** They have higher delinquency rates because the required fully amortizing payment is generally higher than the minimum payment options available during the revolving period.

397. PASS · Precision · `SEC\_10K\_multi\_1029\_01`

- **Question:** What is the total investment banking fee for 2024?
- **Ground truth:** 8,910
- ✓ **Baseline (vanilla RAG) (PASS):** The total Investment Banking revenue for 2024 was **\$9.636 billion**.
- ✗ **VERITROOPER (FAIL):** The total investment banking fees for 2024 were **\$8.910 billion**.

398. FAIL · Cross-Reference · `SEC\_10K\_multi\_0583\_01`

BAD QUESTION

- **Question:** What is the total amount of derivative receivables, net of collateral, and what percentage of total derivative receivables is that?
- **Ground truth:** 11,515, 75%
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, derivative receivables net of all collateral were **\$27.937 billion**, equal to approximately **48.4%** of total net derivative receivables of \$57.777 billion.
- ✗ **VERITROOPER (FAIL):** Derivative receivables net of collateral totaled **\$32.807 billion**, representing approximately **53.8%** of total derivative receivables of \$60.967 billion.

399. FAIL · Precision · `SEC\_10K\_multi\_1255\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total proceeds received from loan sales in 2023?
- **Ground truth:** 19,389
- ✗ **Baseline (vanilla RAG) (FAIL):** The total proceeds received from loan sales in 2023 were **\$19.689 billion**.
- ✗ **VERITROOPER (FAIL):** Total proceeds received from loan sales in 2023 were **\$19,689 million**.

400. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0080\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the estimated range of loss for the legal proceedings described in the passage?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The estimated aggregate range of reasonably possible losses is **\$0 to approximately \$1.2 billion**, in excess of established reserves, as of December 31, 2025.
- ✗ **VERITROOPER (FAIL):** The possible loss or range of loss cannot be reasonably estimated at this time.

401. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the average cost of an industrial accident at one of Apple's suppliers?

- **Ground truth:** UNANSWERABLE

- **✗ Baseline (vanilla RAG) (FAIL):** Apple does not disclose an average cost. It states that an industrial accident at a supplier could cause injuries or loss of life, business disruption, reputational harm, recovery expenditures, and lost sales.
- **✓ VERITROOPER (PASS):** The evidence does not state the average cost of an industrial accident at an Apple supplier.

402. PASS · Precision · `SEC\_10K\_multi\_0629\_00`

- **Question:** What is the net change in nonaccrual loan exposure for the year ended December 31, 2025?
- **Ground truth:** 273
- **✓ Baseline (vanilla RAG) (PASS):** The net increase in nonaccrual loan exposure was **\$273 million** for the year ended December 31, 2025.
- **✓ VERITROOPER (PASS):** The net change in nonaccrual loan exposure was an increase of \$273 million in 2025.

403. PASS · Cross-Reference · `SEC\_10K\_multi\_1243\_00`

- **Question:** What is the maximum exposure as a liquidity provider to customer TOB trusts at December 31, 2025, and what is the fair value of assets held by such VIEs at the same date?
- **Ground truth:** The maximum exposure was \$7.7 billion, and the fair value of assets was \$10.5 billion.
- **✓ Baseline (vanilla RAG) (PASS):** At December 31, 2025:
 - **Maximum exposure as liquidity provider:** **\$7.7 billion**
 - **Fair value of assets held by the VIEs:** **\$10.5 billion**
- **✓ VERITROOPER (PASS):** At December 31, 2025, the maximum exposure was \$7.7 billion, and the fair value of assets held by such VIEs was \$10.5 billion.

404. FAIL · Cross-Reference · `SEC\_10K\_multi\_1409\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What was the total amount for standby letters of credit and other financial guarantees as of December 31, 2024?
- **Ground truth:** \$9.5 billion
- **✗ Baseline (vanilla RAG) (FAIL):** The total contractual amount was **\$28.676 billion** as of December 31, 2024.
- **✓ VERITROOPER (PASS):** As of December 31, 2024, standby letters of credit and other financial guarantees totaled **\$9.5 billion**.

405. PASS · Precision · `SEC\_10K\_multi\_0927\_00`

- **Question:** What is the total notional amount of commodity contracts?
- **Ground truth:** 698
- **✓ Baseline (vanilla RAG) (PASS):** The total notional amount of commodity contracts was **\$698 billion as of December 31, 2025 (\$601 billion in 2024)**.
- **✓ VERITROOPER (PASS):** The total notional amount of commodity contracts is **\$698 million** in Col. 1 and **\$601 million** in Col. 2.

406. PASS · Precision · `SEC\_10K\_multi\_1077\_00`

- **Question:** What is the percentage of HTM securities rated at least AA+ at December 31, 2025?
- **Ground truth:** 99%

- ✓ **Baseline (vanilla RAG) (PASS):** Approximately 99% of HTM securities were rated at least AA+ at December 31, 2025.
- ✓ **VERITROOPER (PASS):** Approximately 99% of HTM securities were rated at least AA+ at December 31, 2025.

407. PASS · Exception · `SEC\_10K\_multi\_1007\_00`

- **Question:** Under what circumstance does the Firm NOT use notional amounts of credit derivatives as the primary measure of risk management?
- **Ground truth:** The circumstance where the notional amount does not take into account the probability of the occurrence of a credit event, the recovery value of the reference obligation, or related cash instruments and economic hedges.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm does not use notional amounts as the primary risk-management measure because they do not reflect the probability of a credit event, the reference obligation's recovery value, or related cash instruments and economic hedges that reduce risk.
- ✓ **VERITROOPER (PASS):** The Firm does not use notional amounts as the primary risk measure because they do not account for credit-event probability, recovery value, or related cash instruments and economic hedges.

408. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1614\_03`

- **Question:** What is the total compensation of the directors who signed the report?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose the directors' compensation, so their total compensation cannot be determined.
- ✓ **VERITROOPER (PASS):** The evidence does not provide compensation amounts for the directors who signed the report.

409. PASS · Cross-Reference · `SEC\_10K\_multi\_0712\_01`

- **Question:** What is the process for determining fair value, as described in Note 2, and what is the significance of the judgments used in estimating fair value for instruments classified within level 3 of the fair value hierarchy?
- **Ground truth:** Estimating fair value requires the application of judgment, and the type and level of judgment required is largely dependent on the amount of observable market information available to the Firm. For instruments valued using internally developed valuation models and other valuation techniques that use significant unobservable inputs and are therefore classified within level 3 of the fair value hierarchy, judgments used to estimate fair value are more significant than those required when estimating the fair value of instruments classified within levels 1 and 2.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm uses well-structured valuation processes and various methodologies and assumptions to determine fair value. For a Level 3 instrument, management:
 1. Selects the appropriate valuation model or other valuation technique.
 2. Assesses relevant empirical data to derive significant unobservable inputs, such as transaction details, yield curves, interest and default rates, prepayment speeds, volatilities, correlations, market prices, comparable-instrument valuations, foreign exchange rates, and credit curves.
 3. Applies valuation adjustments, where appropriate, for counterparty credit quality, the Firm's own creditworthiness, market funding rates, liquidity, unobservable parameters, and certain net open risk positions.

Judgment is especially significant for Level 3 instruments because observable market information is limited and the valuation relies materially on internally developed models and unobservable inputs. These judgments can be further affected by product terms, market liquidity, and heightened volatility or uncertainty. Imprecision in assumptions may change recorded gains or losses, and different reasonable methodologies or assumptions could produce different fair-value estimates at the reporting date.

- ✓ **VERITROOPER (PASS):** For Level 3 instruments, management first selects an appropriate valuation model or technique, then assesses empirical data to derive significant unobservable inputs and applies valuation adjustments. These judgments are more significant than for Levels 1 and 2 because observable market information is limited, and different assumptions or methodologies could materially change the estimated fair value and related gains or losses.

410. FAIL · Precision · `SEC\_10K\_multi\_0932\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total fair value of trading assets and liabilities?
- **Ground truth:** 606,382
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, the total fair value was:
- **Trading assets:** \$606.382 billion gross; **\$57.777 billion net**
- **Trading liabilities:** \$596.738 billion gross; **\$46.329 billion net**
- ✓ **VERITROOPER (PASS):** The total fair value was \$606,382 million for derivative receivables and \$596,738 million for derivative payables.

411. PASS · Precision · `SEC\_10K\_multi\_0151\_01`

- **Question:** What is the year in which the Firm used the COSO 2013 framework to assess internal control?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** 2025.
- ✓ **VERITROOPER (PASS):** The Firm used the COSO 2013 framework to assess internal control as of December 31, 2025.

412. PASS · Exception · `SEC\_10K\_multi\_0045\_01`

- **Question:** What factor may cause us to underestimate or overestimate demand, impacting the timing and volume of our revenue?
- **Ground truth:** Changes in product development cycles and time to market.
- ✓ **Baseline (vanilla RAG) (PASS):** Factors include:
 - Changes in product development cycles and time to market
 - Competing technologies and competitors' product releases, announcements, or other actions
 - Changes in business and economic conditions
 - Sudden or sustained government lockdowns or public health issues
 - Rapidly changing technology or customer requirements
 - Availability of sufficient data-center capacity or energy
- ✓ **VERITROOPER (PASS):** Changes in product development cycles and time to market may cause demand to be underestimated or overestimated.

413. PASS · Precision · `SEC\_10K\_multi\_1612\_01`

- **Question:** What is the year in which the Company filed its Current Report on Form 8-K for the quarter ended April 3, 2015?
- **Ground truth:** 2015
- ✓ **Baseline (vanilla RAG) (PASS):** 2015.

- ✓ **VERITROOPER (PASS)**: The evidence does not identify a Current Report on Form 8-K for the quarter ended April 3, 2015; it identifies a Quarterly Report on Form 10-Q.

414. PASS · Precision · `SEC\_10K\_multi\_0066\_00`

- **Question**: What is the useful life of the developed technology intangible asset?
- **Ground truth**: five years
- ✓ **Baseline (vanilla RAG) (PASS)**: The developed technology intangible asset has a **five-year useful life**.
- ✓ **VERITROOPER (PASS)**: The developed technology intangible asset has a useful life of five years.

415. FAIL · Precision · `SEC\_10K\_multi\_1025\_00`

BAD QUESTION

- **Question**: What is the fair value of receivables?
- **Ground truth**: 190,662
- ✗ **Baseline (vanilla RAG) (FAIL)**: As of December 31, 2025, derivative receivables had a **gross fair value of \$606.382 billion** and a **net fair value of \$57.777 billion** after netting adjustments.
- ✓ **VERITROOPER (PASS)**: The fair value of receivables was **\$8,408 million** as of December 31, 2024.

416. FAIL · Precision · `SEC\_10K\_multi\_1029\_00`

VALIDATOR OVER-FIRED

- **Question**: What is the total underwriting fee for 2023?
- **Ground truth**: 3,759
- ✗ **Baseline (vanilla RAG) (FAIL)**: The total underwriting fee for 2023 was **\$3.759 billion**.
- ✓ **VERITROOPER (PASS)**: The total underwriting fee for 2023 was **\$3,759 million**.

417. FAIL · Precision · `SEC\_10K\_multi\_0630\_00`

UNSUPPORTED NUMBER

- **Question**: What is the total for wholesale loans?
- **Ground truth**: 843,374
- ✗ **Baseline (vanilla RAG) (FAIL)**: Total wholesale loans were **\$843.374 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS)**: The total for wholesale loans is **\$843,374**.

418. PASS · Conditional · `SEC\_10K\_multi\_0071\_01`

- **Question**: If the net realized gains on investments in publicly-held equity securities sold were not significant for fiscal years 2026, 2025, and 2024, what was the implication for the fiscal year 2026?
- **Ground truth**: The difference between the sale proceeds and the carrying value of the equity securities at the beginning of the period or the purchase date, if later
- ✓ **Baseline (vanilla RAG) (PASS)**: For fiscal 2026, the difference between sale proceeds and the securities' carrying value at the beginning of the period-or purchase date, if later-was immaterial. Thus, sales of publicly held equity securities had no significant impact on fiscal 2026 earnings.
- ✓ **VERITROOPER (PASS)**: For fiscal 2026, net realized gains were not significant, indicating little difference between sale proceeds and the securities' applicable carrying value.

419. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0109\_03`

HALLUCINATION

- **Question:** What is the name of the CEO of JPMorgan Chase & Co.?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** James Dimon.
- **✓ VERITROOPER (PASS):** The evidence does not provide the name of JPMorgan Chase & Co.'s CEO.

420. FAIL · Cross-Reference · `SEC\_10K\_multi\_1584\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total amount of cash and cash equivalents at the end of 2025, considering the cash and cash equivalents at the beginning of the year and the net increase (decrease) in cash, cash equivalents, restricted cash and restricted cash equivalents during the year?
- **Ground truth:** 10,270
- **✗ Baseline (vanilla RAG) (FAIL):** Beginning balance of \$11.488 billion minus the \$478 million decrease equals **\$11.010 billion** in total cash, cash equivalents, and restricted cash. After subtracting \$740 million of restricted cash, **cash and cash equivalents at the end of 2025 were \$10.270 billion.**
- **✗ VERITROOPER (FAIL):** At the end of 2025, cash, cash equivalents, restricted cash and restricted cash equivalents totaled **\$11.010 billion** (\$11.488 billion – \$0.478 billion).

421. PASS · Precision · `SEC\_10K\_multi\_0082\_00`

- **Question:** What is the U.S. federal statutory rate for income tax?
- **Ground truth:** 21%
- **✓ Baseline (vanilla RAG) (PASS):** The U.S. federal statutory income tax rate is **21.0%**.
- **✓ VERITROOPER (PASS):** The U.S. federal statutory income tax rate is **21%**.

422. PASS · Cross-Reference · `SEC\_10K\_multi\_0726\_00`

- **Question:** What is the primary consideration for assessing whether the Firm has the power to direct the activities of a VIE, and how does it relate to the Firm's role in establishing the VIE?
- **Ground truth:** The primary consideration is identifying the activities that most significantly impact the VIE's economic performance and determining which party has power over those activities, including the Firm's role in establishing the VIE.
- **✓ Baseline (vanilla RAG) (PASS):** The primary consideration is identifying which activities most significantly affect the VIE's economic performance and determining which party has power over those activities. The Firm considers all facts and circumstances, including its role in establishing the VIE and its ongoing rights and responsibilities; involvement in establishing the VIE alone does not establish control if decision-making power is shared or held by another party.
- **✓ VERITROOPER (PASS):** The Firm considers all facts and circumstances, including its role in establishing the VIE and its ongoing rights and responsibilities. It identifies the activities that most significantly affect the VIE's economic performance and determines which party has power over them.

423. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0581\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the percentage of total loans that are noninvestment-grade and held for more than 5 years?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Noninvestment-grade loans originated prior to 2021 totaled:

- Real estate: \$10,978 million
- Commercial and industrial: \$825 million
- Other: \$2,563 million

Total: **\$14,366 million**

$$\left[\frac{14,366}{792,367} \times 100 \approx 1.81\% \right]$$

Therefore, **approximately 1.8% of total retained loans** were noninvestment-grade and held for more than five years.

- ✓ **VERITROOPER (PASS)**: The evidence does not specify the amount of loans that are both noninvestment-grade and held for more than five years, so the percentage cannot be calculated.

424. PASS · Cause & Effect · `SEC\_10K\_multi\_0707\_00`

- **Question**: What happens as a consequence of the failure to meet legal obligations?
- **Ground truth**: Legal risk
- ✓ **Baseline (vanilla RAG) (PASS)**: Failure to meet legal obligations can cause financial loss, litigation, investigations, enforcement actions, significant penalties or judgments, reputational harm, higher compliance costs, and restrictions or changes to business operations or management.
- ✓ **VERITROOPER (PASS)**: Failure to meet legal obligations can primarily cause loss to the Firm, constituting legal risk.

425. PASS · Exception · `SEC\_10K\_multi\_0708\_01`

- **Question**: What does the Firm's process not consider?
- **Ground truth**: Changes in other adjustments to the quantitative calculation
- ✓ **Baseline (vanilla RAG) (PASS)**: The modeled-estimates sensitivity analysis does not consider the forecasted Apple credit card portfolio while the Firm integrates the Apple Card transaction into its allowance model.
- ✓ **VERITROOPER (PASS)**: The sensitivity analysis does not consider offsetting or correlated effects in other qualitative components of the allowance for credit losses.

426. FAIL · Cross-Reference · `SEC\_10K\_multi\_1403\_00`

UNSUPPORTED NUMBER

- **Question**: What is the total number of derivatives qualifying as guarantees, and how does it compare to the unsettled resale and securities borrowed agreements?
- **Ground truth**: 1,168, which is less than 136,841.
- ✗ **Baseline (vanilla RAG) (FAIL)**: As of December 31, 2025:
- **Derivatives qualifying as guarantees: \$49.031 billion**
- **Unsettled resale and securities borrowed agreements: \$137.072 billion**

The unsettled agreements were **\$88.041 billion higher**, or approximately **2.8 times** the derivatives qualifying as guarantees.

- ✗ **VERITROOPER (FAIL)**: Derivatives qualifying as guarantees total 49,031, compared with 137,072 for unsettled resale and securities borrowed agreements-88,041 lower.

427. PASS · Precision · `SEC\_10K\_multi\_1454\_00`

- **Question:** What is the total liabilities?
- **Ground truth:** 254,633
- ✓ **Baseline (vanilla RAG) (PASS):** The total liabilities amount is not provided in the evidence.
- ✗ **VERITROOPER (FAIL):** Total liabilities are **\$347,606** for Col 1 and **\$324,856** for Col 2.

428. PASS · Exception · `SEC\_10K\_multi\_1540\_01`

- **Question:** What could lead to the erosion of our products' competitiveness?
- **Ground truth:** Failure to adequately protect our intellectual property rights
- ✓ **Baseline (vanilla RAG) (PASS):** Failure to adequately protect intellectual property rights-or legal changes that diminish or remove protections for trademarks, formulas, and other intellectual property-could erode the competitiveness of the products.
- ✓ **VERITROOPER (PASS):** All except which? Please provide the answer choices.

429. PASS · Precision · `SEC\_10K\_multi\_0075\_00`

- **Question:** What was the balance of customer advances as of January 26, 2025?
- **Ground truth:** 81 million
- ✓ **Baseline (vanilla RAG) (PASS):** The balance of customer advances as of January 26, 2025 was **\$81 million**.
- ✓ **VERITROOPER (PASS):** The balance of customer advances as of January 26, 2025 was **\$81 million**.

430. PASS · Cross-Reference · `SEC\_10K\_multi\_1609\_01`

- **Question:** What is the maximum percentage of total assets that the Company agrees to omit from filing certain instruments defining the rights of holders of long-term debt, and what is the corresponding exhibit number for the Amended and Restated Indenture?
- **Ground truth:** 10%, and 4.3
- ✓ **Baseline (vanilla RAG) (PASS):** The maximum is **10% of total assets**, and the Amended and Restated Indenture is **Exhibit 4.3**.
- ✓ **VERITROOPER (PASS):** The maximum is 10% of total assets, and the Amended and Restated Indenture is Exhibit 4.3.

431. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0767\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the volatility of natural gas commodity forward?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The commodity volatility ranged from **2% to 36%**, with a weighted average of **6%**.
- ✗ **VERITROOPER (FAIL):** The commodity volatility values are 2%, 36%, and 6%.

432. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_03`

- **Question:** What percentage of the Company's revenue is allocated to cybersecurity measures?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose any percentage of the Company's revenue allocated to cybersecurity measures.
- ✓ **VERITROOPER (PASS):** The evidence does not specify what percentage of the Company's revenue is allocated to cybersecurity measures.

433. PASS · Cause & Effect · `SEC\_10K\_multi\_0007\_01`

- **Question:** What happens due to the Company's noncompliance with laws and regulations?
- **Ground truth:** Significant penalties or legal liability.
- ✓ **Baseline (vanilla RAG) (PASS):** Noncompliance may result in:
 - Significant penalties, fines, legal liability, litigation, and government investigations or enforcement actions.
 - Investigatory expenses, mandatory corrective actions, audits, and costly changes to business practices, products, services, or operations.
 - Adverse publicity, reputational harm, business disruption, and data loss.
 - Material adverse effects on the Company's business, financial condition, operating results, and stock price.
- ✓ **VERITROOPER (PASS):** Noncompliance can result in significant penalties, legal liability, fines, product bans or import delays, required business-practice changes, and material harm to the Company's business, reputation, operations, financial condition, and stock price.

434. PASS · Cross-Reference · `SEC\_10K\_multi\_1200\_01`

- **Question:** What were the lending-related commitments excluded for the credit card portfolio segment at December 31, 2023 and 2024?
- **Ground truth:** 915.7 billion and 1.0 trillion
- ✓ **Baseline (vanilla RAG) (PASS):** The excluded lending-related commitments for the credit card portfolio segment were:
 - **December 31, 2023:** \$915.7 billion
 - **December 31, 2024:** \$1.0 trillion
- ✓ **VERITROOPER (PASS):** The excluded lending-related commitments for the credit card portfolio segment were \$915.7 billion in 2023 and \$1.0 trillion in 2024.

435. PASS · Precision · `SEC\_10K\_multi\_1039\_01`

- **Question:** What is the loss on tax-oriented investments for the year ended December 31, 2024?
- **Ground truth:** 97
- ✓ **Baseline (vanilla RAG) (PASS):** The loss on tax-oriented investments for 2024 was **\$97 million**.
- ✓ **VERITROOPER (PASS):** The loss on tax-oriented investments for 2024 was **\$97 million**.

436. FAIL · Cross-Reference · `SEC\_10K\_multi\_0404\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total assets under management in North America compared to the total international assets under management?
- **Ground truth:** 3,582 vs 1,209
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, assets under management were **\$3.582 trillion in North America**, compared with **\$1.209 trillion internationally**. North America had **\$2.373 trillion more** in assets under management.
- ✓ **VERITROOPER (PASS):** North America assets under management were \$3,582, \$3,033, and \$2,534, compared with total international assets under management of \$1,209, \$1,012, and \$888, respectively.

437. PASS · Cross-Reference · `SEC\_10K\_multi\_1544\_01`

- **Question:** What was the outcome of the Wisconsin Court of Appeals regarding the summary judgment ruling in favor of the Company and Aqua-Chem?
- **Ground truth:** affirmed

- ✓ **Baseline (vanilla RAG) (PASS):** On October 29, 2013, the Wisconsin Court of Appeals affirmed the summary judgment in favor of the Company and Aqua-Chem.
- ✓ **VERITROOPER (PASS):** On October 29, 2013, the Wisconsin Court of Appeals affirmed the summary judgment in favor of the Company and Aqua-Chem.

438. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1265\_01`

HALLUCINATION

- **Question:** How many business combinations occurred in 2025 that resulted in goodwill?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Zero business combinations occurred in 2025 that resulted in goodwill.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many business combinations occurred in 2025 that resulted in goodwill.

439. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_06`

VALIDATOR OVER-FIRED

- **Question:** What is the total cost of implementing new tariffs on imports to the U.S.?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The total cost is not quantified. Apple states that the impact of the new U.S. tariffs remains uncertain and depends on their scope, duration, future changes, and retaliatory measures.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total cost of implementing the new U.S. tariffs.

440. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0607\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of the All other category as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The "All other commissions and fees" category totaled **\$2.049 billion** for 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total value of the All other category as of December 31, 2025.

441. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_03`

- **Question:** What percentage of our revenue is allocated to data privacy and security?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify what percentage of revenue is allocated to data privacy and security.
- ✓ **VERITROOPER (PASS):** The evidence does not state what percentage of revenue is allocated to data privacy and security.

442. PASS · Precision · `SEC\_10K\_multi\_0993\_01`

- **Question:** What is the approximate amount of net losses related to cash flow hedges that the Firm expects to recognize in income over the next 12 months?
- **Ground truth:** \$926 million
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm expects to recognize approximately **\$926 million in net losses, after tax**, over the next 12 months.
- ✓ **VERITROOPER (PASS):** Approximately \$926 million after tax.

443. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1614\_02`

HALLUCINATION

- **Question:** What is the name of the attorney who certified the report?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The NVIDIA report appoints **Jen-Hsun Huang and Colette M. Kress** as attorneys-in-fact.
- **✓ VERITROOPER (PASS):** No attorney is identified as certifying the report; Jennifer D. Manning signed as attorney-in-fact.

444. PASS · Calculation · `SEC\_10K\_multi\_1335\_00`

- **Question:** What is the total number of shares of common stock repurchased in 2025?
- **Ground truth:** 114.4
- **✓ Baseline (vanilla RAG) (PASS):** 114.4
- **✓ VERITROOPER (PASS):** 114.4 million

445. PASS · Cause & Effect · `SEC\_10K\_multi\_1279\_00`

- **Question:** What happens because reimbursement of the advances is typically senior to all cash payments to investors?
- **Ground truth:** The Firm's credit risk associated with these servicer advances is minimal
- **✓ Baseline (vanilla RAG) (PASS):** The Firm's credit risk associated with servicer advances is minimal.
- **✓ VERITROOPER (PASS):** The Firm's credit risk associated with these servicer advances is minimal.

446. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0717\_02`

- **Question:** What is the expected increase in adoption rate for hedge accounting under the new guidance?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify an expected increase in the adoption rate for hedge accounting under the new guidance. It only states that the guidance expands eligibility and flexibility and that JPMorganChase is evaluating its potential impact.
- **✓ VERITROOPER (PASS):** The evidence does not state an expected increase in the adoption rate for hedge accounting.

447. FAIL · Precision · `SEC\_10K\_multi\_0861\_01`

UNSUPPORTED NUMBER

- **Question:** What is the value of Short-term borrowings?
- **Ground truth:** 32.3
- **✗ Baseline (vanilla RAG) (FAIL):** Short-term borrowings were **\$0 million** as of December 31, 2025 and 2024.
- **✓ VERITROOPER (PASS):** Short-term borrowings were **\$32.3**.

448. PASS · Exception · `SEC\_10K\_multi\_1583\_01`

- **Question:** What does not reflect the actual impact of fluctuations in foreign currency exchange rates on operating results?
- **Ground truth:** The percentages in the table above
- **✓ Baseline (vanilla RAG) (PASS):** The percentages in the table, because they exclude the effects of hedging activities.

- ✓ **VERITROOPER (PASS)**: The percentages in the table above do not reflect the actual impact because they exclude the effects of hedging activities.

449. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0539\_04`

HALLUCINATION

- **Question**: What is the credit rating of J.P. Morgan Securities plc according to Standard & Poor's?
- **Ground truth**: UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL)**: J.P. Morgan Securities plc is rated **AA-** for long-term issuer credit and **A-1+** for short-term issuer credit by Standard & Poor's, with a **Stable** outlook.
- ✓ **VERITROOPER (PASS)**: The evidence does not provide Standard & Poor's credit rating for J.P. Morgan Securities plc.

450. PASS · Conditional · `SEC\_10K\_multi\_0669\_00`

- **Question**: If the refined time series was effective at the beginning of each year presented, would the average Total VaR be lower?
- **Ground truth**: Yes
- ✓ **Baseline (vanilla RAG) (PASS)**: Yes. Average Total VaR would have been lower by \$1 million in 2025 and \$2 million in 2024.
- ✓ **VERITROOPER (PASS)**: Yes. The average Total VaR would have been lower for each year presented.

451. PASS · Conditional · `SEC\_10K\_multi\_0570\_01`

- **Question**: If the loan maturities are categorized by interest rate type, and the distribution is between fixed and floating interest rates, what is the basis for the distribution?
- **Ground truth**: based on the stated terms of the loan agreements
- ✓ **Baseline (vanilla RAG) (PASS)**: The distribution is based on the stated terms of the loan agreements.
- ✓ **VERITROOPER (PASS)**: The distribution between fixed and floating interest rates is based on the stated terms of the loan agreements.

452. PASS · Exception · `SEC\_10K\_multi\_0051\_01`

- **Question**: What scenario may encourage customers outside China to purchase from competitors rather than NVIDIA?
- **Ground truth**: export controls
- ✓ **Baseline (vanilla RAG) (PASS)**: Export controls may prompt overseas governments to request that customers outside China purchase from NVIDIA's competitors rather than from NVIDIA or other U.S. firms.
- ✓ **VERITROOPER (PASS)**: Export controls may prompt overseas governments to request that customers outside China purchase from NVIDIA's competitors rather than NVIDIA or other U.S. firms.

453. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0109\_00`

- **Question**: What is the total number of employees at JPMorgan Chase & Co.?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not provide JPMorgan Chase & Co.'s total employee count.
- ✓ **VERITROOPER (PASS)**: The evidence does not state JPMorgan Chase & Co.'s total number of employees.

454. PASS · Conditional · `SEC\_10K\_multi\_0727\_01`

- **Question:** If an instrument is classified within level 1 of the fair value hierarchy, are any adjustments to quoted prices applied?
- **Ground truth:** no
- ✓ **Baseline (vanilla RAG) (PASS):** No. Instruments classified within Level 1 use unadjusted quoted prices, and no adjustments are applied.
- ✓ **VERITROOPER (PASS):** No. Instruments classified within Level 1 receive no adjustments to quoted prices.

455. FAIL · Precision · `SEC\_10K\_multi\_1113\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total loan balance at December 31, 2024?
- **Ground truth:** 1,347,988
- ✗ **Baseline (vanilla RAG) (FAIL):** The total retained loan balance at December 31, 2024 was **\$1.300 trillion (\$1,299.590 billion)**.
- ✓ **VERITROOPER (PASS):** The total loan balance at December 31, 2024 was **\$1,347,988 million**.

456. PASS · Cause & Effect · `SEC\_10K\_multi\_1284\_01`

- **Question:** What happens because of a 20% adverse change in the weighted-average prepayment speed assumption?
- **Ground truth:** The fair value decreases by \$353 million.
- ✓ **Baseline (vanilla RAG) (PASS):** A 20% adverse change in the weighted-average prepayment speed assumption reduces MSR fair value by **\$353 million in 2025** and **\$406 million in 2024**.
- ✓ **VERITROOPER (PASS):** A 20% adverse change reduces the fair value of the Firm's MSRs by \$353 million in 2025 and \$406 million in 2024.

457. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0714\_03`

- **Question:** What is the estimated cost savings for companies adopting the new guidance?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide an estimated cost savings for companies adopting the new guidance.
- ✓ **VERITROOPER (PASS):** The evidence does not provide an estimated cost savings for companies adopting the new guidance.

458. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1279\_01`

- **Question:** What percentage of servicer advances are typically recoverable?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** No specific percentage is provided. The filing states that servicer advances are generally reimbursed and associated credit risk is minimal, implying that most are typically recoverable.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a percentage of servicer advances that are typically recoverable.

459. FAIL · Cross-Reference · `SEC\_10K\_multi\_1525\_01`

UNSUPPORTED NUMBER

- **Question:** How do the assets under management (AUM) and the assets under custody (AUC) relate to the client assets?
- **Ground truth:** Both AUM and AUC represent assets held on behalf of clients, but AUM includes assets managed by AWM, while AUC includes assets held directly or indirectly under safekeeping.
- **✗ Baseline (vanilla RAG) (FAIL):** Client assets are the broader measure:
- **AUM:** Assets managed by Asset & Wealth Management on behalf of clients.
- **AUC:** Assets held on clients' behalf under custody, safekeeping, and servicing arrangements.
- **Client assets:** AUM plus assets in custody, brokerage, administration, and deposit accounts.

For AWM at December 31, 2025, **\$7.118 trillion of client assets** consisted of **\$4.791 trillion of AUM** and **\$2.327 trillion in custody, brokerage, administration, and deposits**. Firmwide AUC was **\$41.172 trillion**, so it is not equivalent to, or wholly included in, AWM client assets.

- **✓ VERITROOPER (PASS):** Client assets include both assets under management (AUM) and assets under custody (AUC), along with brokerage, administration, and deposit accounts.

460. FAIL · Cross-Reference · `SEC\_10K\_multi\_0084\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the total revenue for the United States and Taiwan for fiscal year 2026, and how does it compare to the total revenue for the same regions in fiscal year 2025?
- **Ground truth:** The total revenue for the United States for fiscal year 2026 is \$149,617 million, and for Taiwan is \$42,345 million. For fiscal year 2025, the total revenue for the United States is \$77,482 million, and for Taiwan is \$23,600 million.
- **✗ Baseline (vanilla RAG) (FAIL):** - **Fiscal 2026:** United States (\$149,617 million) + Taiwan (\$42,345 million) = **\$191,962 million**
- **Fiscal 2025:** United States (\$77,482 million) + Taiwan (\$23,600 million) = **\$101,082 million**

The combined revenue increased by **\$90,880 million**, or approximately **89.9%**, in fiscal 2026.

- **✗ VERITROOPER (FAIL):** United States and Taiwan revenue totaled \$191.962 billion in fiscal 2026, versus \$101.082 billion in fiscal 2025—an increase of \$90.880 billion.

461. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1265\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of goodwill allocated to the Corporate segment in 2025 that was not allocated in 2024?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Corporate goodwill increased by **\$30 million** in 2025, from \$692 million in 2024 to \$722 million.
- **✗ VERITROOPER (FAIL):** The Corporate segment's goodwill increased by **\$30 million** in 2025, from \$692 million to \$722 million.

462. PASS · Precision · `SEC\_10K\_multi\_0322\_01`

- **Question:** What is the number?
- **Ground truth:** 69
- **✓ Baseline (vanilla RAG) (PASS):** The question is ambiguous because the evidence contains many numbers; please specify which figure you mean.

- ✓ **VERITROOPER (PASS):** 69

463. PASS · Precision · `SEC\_10K\_multi\_0974\_01`

- **Question:** What is the year of the second date mentioned?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** 2025
- ✓ **VERITROOPER (PASS):** 2024

464. PASS · Conditional · `SEC\_10K\_multi\_0729\_01`

- **Question:** If observable market information is not available for beneficial interests issued by consolidated VIEs, what valuation method is used?
- **Ground truth:** Valuations are based on the fair value of the underlying assets held by the VIE.
- ✓ **Baseline (vanilla RAG) (PASS):** They are valued based on the fair value of the underlying assets held by the VIE.
- ✓ **VERITROOPER (PASS):** Valuations are based on the fair value of the underlying assets held by the VIE.

465. PASS · Precision · `SEC\_10K\_multi\_0943\_01`

- **Question:** What is the OTC-cleared exchange contracts?
- **Ground truth:** 772
- ✓ **Baseline (vanilla RAG) (PASS):** OTC-cleared derivative contracts are negotiated and traded bilaterally with a counterparty, then transferred (novated) to a central counterparty clearinghouse (CCP), which becomes the counterparty for clearing and settlement.
- ✓ **VERITROOPER (PASS):** OTC-cleared foreign exchange contracts were 897 (850 and 47) in the first period and 772 (678 and 94) in the second period.

466. PASS · Exception · `SEC\_10K\_multi\_1186\_01`

- **Question:** What is excluded from the allowance for loan losses breakdown?
- **Ground truth:** Cumulative effect of a change in accounting principle
- ✓ **Baseline (vanilla RAG) (PASS):** The breakdown excludes allowances for lending-related commitments and investment securities, as well as \$288 million in 2025 and \$268 million in 2024 related to certain CIB accounts receivable.
- ✓ **VERITROOPER (PASS):** Credit card loans are excluded from the consumer category and presented separately.

467. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1537\_04`

- **Question:** What is the name of the specific government agency that has made public statements about alleged risks associated with processing or ingredients in our products?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not identify a specific government agency; it refers only generally to "government officials."
- ✓ **VERITROOPER (PASS):** The evidence does not name a specific government agency.

468. FAIL · Precision · `SEC\_10K\_multi\_1039\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the operating lease income for the year ended December 31, 2023?
- **Ground truth:** 2,843

- ✗ **Baseline (vanilla RAG) (FAIL):** The operating lease income for 2023 was **\$2.843 billion**.
- ✓ **VERITROOPER (PASS):** The operating lease income for 2023 was **\$2,843 million**.

469. PASS · Exception · `SEC\_10K\_multi\_1532\_00`

- **Question:** Under what circumstance does the general rule NOT apply to bottler's agreements?
- **Ground truth:** Outside the United States
- ✓ **Baseline (vanilla RAG) (PASS):** The general rule limiting bottlers to an identified territory may not apply when local law requires otherwise; territorial restrictions can vary based on applicable local law.
- ✓ **VERITROOPER (PASS):** The general territorial restriction may not apply where applicable local law provides otherwise.

470. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_03`

- **Question:** How many notes were issued by Apple Inc. in 2023?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not state how many notes Apple issued in 2023. It only reports **\$5.228 billion** in net proceeds from term debt issuance.
- ✗ **VERITROOPER (FAIL):** Apple Inc. issued five notes in 2023.

471. PASS · Precision · `SEC\_10K\_multi\_1215\_00`

- **Question:** What is the value of senior securities at December 31, 2024?
- **Ground truth:** 256 million
- ✓ **Baseline (vanilla RAG) (PASS):** The value of senior securities at December 31, 2024 was **\$256 million**.
- ✓ **VERITROOPER (PASS):** \$256 million.

472. PASS · Exception · `SEC\_10K\_multi\_1217\_01`

- **Question:** Under what circumstance does the Firm NOT consolidate re-securitization VIEs?
- **Ground truth:** when the Firm does not have the unilateral ability to direct the most significant activities of the re-securitization trust
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm does **not** consolidate re-securitization VIEs when it lacks unilateral power to direct their most significant activities-for example, when decision-making power is shared with clients or an independent third-party sponsor, or when the Firm was not involved in the trust's initial design.
- ✓ **VERITROOPER (PASS):** The Firm does not consolidate re-securitization VIEs when decision-making power is shared with clients or an independent third-party sponsor, so it lacks unilateral ability to direct the trust's most significant activities.

473. PASS · Precision · `SEC\_10K\_multi\_0069\_00`

- **Question:** What is the total carrying amount of goodwill as of January 26, 2025?
- **Ground truth:** 5.2 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The total carrying amount of goodwill as of January 26, 2025 was **\$5.2 billion**.
- ✓ **VERITROOPER (PASS):** The total carrying amount of goodwill as of January 26, 2025 was **\$5.2 billion**.

474. PASS · Cross-Reference · `SEC\_10K\_multi\_1598\_01`

- **Question:** What is the target allocation for alternative investments in the Company's pension plans?
- **Ground truth:** 31%

- ✓ **Baseline (vanilla RAG) (PASS):** The target allocation for alternative investments in the Company's U.S. pension plan is **31%**.
- ✓ **VERITROOPER (PASS):** The target allocation for alternative investments is 31%.

475. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1609\_04`

BAD QUESTION

- **Question:** What is the name of the trustee for the notes issued by the Company?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Deutsche Bank Trust Company Americas.
- ✗ **VERITROOPER (FAIL):** The trustee is Deutsche Bank Trust Company Americas.

476. PASS · Precision · `SEC\_10K\_multi\_1612\_00`

- **Question:** What is the date of the letter from the Company to Kathy N. Waller?
- **Ground truth:** March 22, 2017
- ✓ **Baseline (vanilla RAG) (PASS):** The letters from the Company to Kathy N. Waller are dated April 24, 2014; March 22, 2017; and October 17, 2018.
- ✓ **VERITROOPER (PASS):** The listed letters from the Company to Kathy N. Waller are dated March 22, 2017, and October 17, 2018.

477. PASS · Precision · `SEC\_10K\_multi\_1167\_01`

- **Question:** What was the weighted-average of months that extended loans by for single modifications?
- **Ground truth:** 14 months
- ✓ **Baseline (vanilla RAG) (PASS):** The loans were extended by a weighted average of **14 months**.
- ✓ **VERITROOPER (PASS):** For single modifications, loans were extended by a weighted average of 14 months.

478. PASS · Cross-Reference · `SEC\_10K\_multi\_0655\_00`

- **Question:** What types of investments are made by AWM alongside third-party investors?
- **Ground truth:** Privately distributed collective vehicles managed by AWM (i.e., co-investments), as well as in third-party funds
- ✓ **Baseline (vanilla RAG) (PASS):** AWM invests capital alongside third-party investors through co-investments, typically in privately distributed collective vehicles managed by AWM, as well as in third-party funds.
- ✓ **VERITROOPER (PASS):** AWM invests capital alongside third-party investors in privately distributed collective vehicles managed by AWM (co-investments), as well as in third-party funds.

479. PASS · Exception · `SEC\_10K\_multi\_0073\_00`

- **Question:** What is excluded from depreciation?
- **Ground truth:** Land
- ✓ **Baseline (vanilla RAG) (PASS):** Long-lived assets or disposal groups classified as **held for sale** are excluded from depreciation during that period, even if they continue generating revenue.
- ✓ **VERITROOPER (PASS):** Land and construction in process are excluded from depreciation.

480. PASS · Cause & Effect · `SEC\_10K\_multi\_1420\_00`

- **Question:** What happens as a consequence of the Firm's maximum possible exposure under membership agreements being difficult to estimate?
- **Ground truth:** The Firm's maximum possible exposure under these membership agreements cannot be disclosed in the table on page 296.
- ✓ **Baseline (vanilla RAG) (PASS):** When the maximum exposure cannot be estimated, no amount is disclosed; only estimable exposures are included in the "Exchange & clearing house guarantees and commitments" line. Management nevertheless expects the risk of loss to be remote based on historical experience.
- ✓ **VERITROOPER (PASS):** The Firm discloses the maximum possible exposure only where it can be estimated; otherwise, no estimate is provided.

481. PASS · Precision · `SEC\_10K\_multi\_0607\_01`

- **Question:** What was the decrease in Criticized exposure from December 31, 2024 to December 31, 2025?
- **Ground truth:** 714 million
- ✓ **Baseline (vanilla RAG) (PASS):** Criticized exposure decreased by **\$714 million**, from **\$12.4 billion** at December 31, 2024, to **\$11.7 billion** at December 31, 2025.
- ✓ **VERITROOPER (PASS):** Criticized exposure decreased by \$714 million.

482. FAIL · Cross-Reference · `SEC\_10K\_multi\_1158\_00`

TABLE MISREAD

- **Question:** What is the total amount of retained loans for the year 2022, and how does it compare to the total retained loans for the year 2023?
- **Ground truth:** The total retained loans for 2022 is \$19,190 million, and for 2023 is \$24,287 million.
- ✗ **Baseline (vanilla RAG) (FAIL):** Total retained loans were **\$8.967 billion for 2022**, compared with **\$18.794 billion for 2023**. Thus, 2022 was **\$9.827 billion lower**, a decrease of approximately **52.3%**.
- ✗ **VERITROOPER (FAIL):** Total retained loans were \$19,190 million in 2022, compared with \$24,287 million in 2023-an increase of \$5,097 million in 2023.

483. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0591\_01`

- **Question:** What is the average salary in the Technology, Media & Telecommunications sector?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide average salary data for the Technology, Media & Telecommunications sector.
- ✓ **VERITROOPER (PASS):** The evidence does not provide average salary information for the Technology, Media & Telecommunications sector.

484. PASS · Exception · `SEC\_10K\_multi\_1316\_00`

- **Question:** What does the Firm's unsecured debt NOT contain?
- **Ground truth:** requirements that would call for an acceleration of payments, maturities or changes in the structure of the existing debt
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm's unsecured debt does **not** contain provisions that, based on unfavorable changes in credit ratings, financial ratios, earnings, or stock price, would:
 - Accelerate payments or maturities
 - Change the existing debt structure
 - Limit future borrowings

- Require additional collateral
- ✓ **VERITROOPER (PASS):** The Firm's unsecured debt does **not** contain requirements triggered by unfavorable changes in credit ratings, financial ratios, earnings, or stock price that would accelerate payments or maturities, change existing debt structures, limit future borrowings, or require additional collateral.

485. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_00`

HALLUCINATION

- **Question:** What is the total number of countries where Apple's suppliers have manufacturing facilities?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Apple identifies supplier manufacturing facilities in **7 countries/regions:** China mainland, India, Japan, South Korea, Taiwan, Vietnam, and the United States.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total number of countries where Apple's suppliers have manufacturing facilities.

486. FAIL · Cross-Reference · `SEC\_10K\_multi\_1568\_01`

BAD QUESTION

- **Question:** What is the estimated potential aggregate remaining incremental tax and interest liability for the tax years 2010 through 2025, and how does it compare to the total amount of unused backup lines of credit?
- **Ground truth:** \$14 billion
- ✗ **Baseline (vanilla RAG) (FAIL):** The estimated potential aggregate remaining incremental tax and interest liability is approximately **\$14 billion**. This exceeds the **\$6.2 billion** in unused backup lines of credit by about **\$7.8 billion**, or roughly **2.3 times** the available backup credit.
- ✗ **VERITROOPER (FAIL):** The estimated liability is approximately \$14 billion, compared with \$6.2 billion in unused backup credit lines-\$7.8 billion higher, or about 2.3 times the credit lines.

487. PASS · Conditional · `SEC\_10K\_multi\_1590\_00`

- **Question:** If the Company uses cash flow hedges to minimize the variability in cash flows of assets or liabilities, what is the maximum length of time for which the Company hedges its exposure to the variability in future cash flows?
- **Ground truth:** three years
- ✓ **Baseline (vanilla RAG) (PASS):** The maximum length of time is typically **three years**.
- ✓ **VERITROOPER (PASS):** The maximum length of time is typically three years.

488. PASS · Cross-Reference · `SEC\_10K\_multi\_0130\_00`

- **Question:** What are the potential consequences for JPMorganChase if the Federal Reserve and the FDIC were to determine that a resolution plan submitted by JPMorganChase has deficiencies?
- **Ground truth:** JPMorganChase could face more stringent capital, leverage, or liquidity requirements, or restrictions on its growth, activities, or operations.
- ✓ **Baseline (vanilla RAG) (PASS):** The Federal Reserve and FDIC could jointly:
 - Impose more stringent capital, leverage, or liquidity requirements;
 - Restrict JPMorganChase's growth, activities, or operations; and
 - Require it to restructure, reorganize, or divest assets or businesses.

These actions could increase funding or operational costs, hamper client service, and materially and adversely affect JPMorganChase's operations and strategy.

- ✓ **VERITROOPER (PASS):** They could jointly impose stricter capital, leverage, or liquidity requirements; restrict JPMorganChase's growth, activities, or operations; or require restructuring, reorganization, or divestiture of assets or businesses, potentially materially harming its operations and strategy.

489. PASS · Cause & Effect · `SEC\_10K\_multi\_1527\_00`

- **Question:** What happens as a consequence of loans being placed on nonaccrual status?
- **Ground truth:** Loans for which interest income is not recognized on an accrual basis.
- ✓ **Baseline (vanilla RAG) (PASS):** When a loan is placed on nonaccrual status:
 - Accrual of interest is suspended.
 - All accrued but uncollected interest is reversed against interest income.
 - Amortization of deferred fees, costs, premiums, or discounts is suspended.
 - Subsequent cash interest may be recognized as income if the loan balance is considered fully collectible; otherwise, cash receipts reduce the loan's carrying value under the cost-recovery method.
- ✓ **VERITROOPER (PASS):** Interest income on those loans is no longer recognized on an accrual basis.

490. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1554\_02`

- **Question:** What was the percentage of unit case volume growth in the Latin America operating segment attributed to energy drinks?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence states only that energy drinks grew in Latin America; it does not provide a specific percentage.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a percentage for energy drinks' contribution to unit case volume growth in Latin America.

491. PASS · Cross-Reference · `SEC\_10K\_multi\_0180\_00`

- **Question:** What is the effective date of the Employee Stock Purchase Plan and the Performance-Based Incentive Compensation Plan?
- **Ground truth:** January 1, 2019 and January 1, 2021
- ✓ **Baseline (vanilla RAG) (PASS):** - **Employee Stock Purchase Plan:** January 1, 2019
- **Performance-Based Incentive Compensation Plan:** January 1, 2021
- ✓ **VERITROOPER (PASS):** The Employee Stock Purchase Plan was effective January 1, 2019, and the Performance-Based Incentive Compensation Plan was effective January 1, 2021.

492. PASS · Precision · `SEC\_10K\_multi\_0139\_00`

- **Question:** What is the total for New York City, New York?
- **Ground truth:** 7.1
- ✓ **Baseline (vanilla RAG) (PASS):** The total for New York City, New York was approximately **9.1 million square feet** as of December 31, 2025.
- ✗ **VERITROOPER (FAIL):** The total for New York City, New York is **9.1**.

493. FAIL · Precision · `SEC\_10K\_multi\_0803\_00`

WRONG PERIOD

- **Question:** What is the total number of mortgage-backed securities?
- **Ground truth:** 775
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, total mortgage-backed securities classified as trading assets were **\$162.085 billion**.
- ✓ **VERITROOPER (PASS):** The total mortgage-backed securities are **775** (Col 1).

494. FAIL · Precision · `SEC\_10K\_multi\_0943\_00`

TABLE MISREAD

- **Question:** What is the total interest rate contracts?
- **Ground truth:** 256,851
- **✗ Baseline (vanilla RAG) (FAIL):** Total interest rate contracts were **\$29.536 trillion in 2025** and **\$29.660 trillion in 2024** (notional amounts).
- **✓ VERITROOPER (PASS):** The total interest rate contracts are 256,851 for the first period and 273,133 for the second period.

495. PASS · Cause & Effect · `SEC\_10K\_multi\_1261\_00`

- **Question:** What happens as a consequence of the Firm's repurchase option becoming exercisable?
- **Ground truth:** Loans must be reported on the Consolidated balance sheets as a loan with a corresponding liability.
- **✓ Baseline (vanilla RAG) (PASS):** When the repurchase option becomes exercisable, the loans must be recognized on the Firm's Consolidated balance sheets as loans, with a corresponding liability.
- **✓ VERITROOPER (PASS):** The loans must be reported on the Consolidated balance sheets as loans with corresponding liabilities.

496. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0322\_01`

- **Question:** How many hedging activities are associated with the valuation adjustments?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify a number; it only states that CVA and FVA are presented net of associated hedging activities.
- **✓ VERITROOPER (PASS):** The evidence does not specify the number of associated hedging activities.

497. PASS · Precision · `SEC\_10K\_multi\_1610\_00`

- **Question:** What is the date of the 2014 Equity Plan amendment?
- **Ground truth:** February 17, 2016
- **✓ Baseline (vanilla RAG) (PASS):** The 2014 Equity Plan was amended and restated on **February 17, 2016**.
- **✓ VERITROOPER (PASS):** The 2014 Equity Plan was amended and restated as of February 17, 2016.

498. PASS · Exception · `SEC\_10K\_multi\_0111\_00`

- **Question:** Under what circumstance does the general rule NOT apply to the payment of dividends by national banks?
- **Ground truth:** if, in the banking regulator's opinion, payment of a dividend would constitute an unsafe or unsound practice in light of the financial condition of the bank
- **✓ Baseline (vanilla RAG) (PASS):** The general rule does not apply if the OCC or Federal Reserve determines that paying the dividend would be an unsafe or unsound practice given the bank's financial condition; they may then prohibit or limit the payment.
- **✓ VERITROOPER (PASS):** The general rule does not apply if the OCC or Federal Reserve determines that paying the dividend would be an unsafe or unsound practice given the bank's financial condition.

499. PASS · Cause & Effect · `SEC\_10K\_multi\_0046\_01`

- **Question:** What is the result of a cyber-attack on our operations?
- **Ground truth:** Disruption of operations, reduction in expected revenue, increase in expenses, and significant harm to our business and reputation.
- ✓ **Baseline (vanilla RAG) (PASS):** A cyber-attack could disrupt or block access to operational systems, delay data recovery, enable fraudulent transactions, expose or destroy confidential data, reduce revenue, increase remediation, litigation and regulatory costs, and damage customer confidence and reputation. It could also cause broader industry-wide disruption through interconnected third-party systems.
- ✓ **VERITROOPER (PASS):** A cyber-attack could disrupt operations, reduce expected revenue, increase expenses, and significantly harm the business and its reputation.

500. PASS · Precision · `SEC\_10K\_multi\_0705\_01`

- **Question:** What is the time horizon for the Loss Distribution Approach (LDA) model?
- **Ground truth:** one-year
- ✓ **Baseline (vanilla RAG) (PASS):** The LDA model uses a **one-year time horizon**.
- ✓ **VERITROOPER (PASS):** The LDA model uses a one-year time horizon.

501. PASS · Precision · `SEC\_10K\_multi\_0768\_00`

- **Question:** What is the threshold for Bermudan switch value?
- **Ground truth:** 0
- ✓ **Baseline (vanilla RAG) (PASS):** The Bermudan switch value ranges from **0% to 48%**, with a **17% weighted average**.
- ✓ **VERITROOPER (PASS):** The threshold for the Bermudan switch value is 17%.

502. FAIL · Precision · `SEC\_10K\_multi\_0029\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the total net sales for the Americas in 2025?
- **Ground truth:** 178,353
- ✗ **Baseline (vanilla RAG) (FAIL):** The Americas' total net sales in 2025 were **\$178.353 billion**.
- ✗ **VERITROOPER (FAIL):** The Americas' total net sales in 2025 were **\$178.353 billion**.

503. PASS · Exception · `SEC\_10K\_multi\_1547\_01`

- **Question:** What is excluded from the definition of a primary beneficiary?
- **Ground truth:** the obligation to absorb losses or the right to receive benefits from the VIE that could potentially be significant to the VIE
- ✓ **Baseline (vanilla RAG) (PASS):** An entity is not a primary beneficiary if it lacks either:
 1. Power to direct the VIE's most economically significant activities; or
 2. An obligation to absorb potentially significant losses or a right to receive potentially significant benefits.

Thus, financial exposure alone-such as equity investments, guarantees, or subordinated support-does not make an entity the primary beneficiary without directing power.

- ✓ **VERITROOPER (PASS):** Merely holding variable interests is excluded as sufficient qualification; the entity must also have power over key VIE activities and significant loss exposure or benefit rights.

504. PASS · Precision · `SEC\_10K\_multi\_1476\_00`

- **Question:** What is the total interest-earning assets?
- **Ground truth:** 3,834,359
- ✓ **Baseline (vanilla RAG) (PASS):** Total interest-earning assets were \$3,834,359 million (\$3.834 trillion) in 2025.
- ✓ **VERITROOPER (PASS):** The total interest-earning assets are \$3,834,359.

505. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0717\_04`

- **Question:** What is the total cost of implementing the new guidance for a typical firm?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose the total implementation cost for a typical firm. It only states that firms are evaluating the potential financial-statement impact and adoption timing.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total implementation cost for a typical firm.

506. PASS · Calculation · `SEC\_10K\_multi\_1337\_00`

- **Question:** What is the net income applicable to common stockholders for the year ended December 31, 2025?
- **Ground truth:** 55,681
- ✓ **Baseline (vanilla RAG) (PASS):** 55681
- ✓ **VERITROOPER (PASS):** 55,681

507. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0322\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the percentage of revenue attributed to derivatives in the Fixed Income Markets?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The filing does not separately disclose derivatives revenue. If principal transactions revenue is used as a proxy, Fixed Income Markets' 2025 percentage was approximately **54.7%**:

$$\frac{\$12.327 \text{ billion}}{\$22.532 \text{ billion}} \times 100 \approx 54.7\%$$

Principal transactions also include non-derivative financial instruments and commodities.

- ✓ **VERITROOPER (PASS):** The evidence does not provide enough information to determine the percentage of revenue attributed to derivatives in the Fixed Income Markets.

508. PASS · Precision · `SEC\_10K\_multi\_1210\_01`

- **Question:** What is the date of the financial statement?
- **Ground truth:** December 31, 2024
- ✓ **Baseline (vanilla RAG) (PASS):** December 31, 2025.
- ✓ **VERITROOPER (PASS):** The financial statement is dated December 31, 2024.

509. PASS · Exception · `SEC\_10K\_multi\_0248\_00`

- **Question:** What circumstance does the income tax expense NOT apply to?
- **Ground truth:** the absence of a \$1.0 billion contribution of Visa shares to the JPMorgan Chase Foundation

- ✓ **Baseline (vanilla RAG) (PASS):** Income tax expense does not apply to non-income-tax-related gains, losses, and other returns; these are recognized in **Other income**.
- ✓ **VERITROOPER (PASS):** The evidence does not specify any circumstance to which income tax expense does not apply.

510. PASS · Precision · `SEC\_10K\_multi\_0079\_01`

- **Question:** What is the maximum gross exposure under all agreements for facility lease guarantees?
- **Ground truth:** 3.5 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The maximum gross exposure is **\$3.5 billion**.
- ✓ **VERITROOPER (PASS):** The maximum gross exposure is **\$3.5 billion**.

511. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0043\_01`

- **Question:** How many employees are engaged in sales, marketing, operations, and administrative positions in countries outside of Asia?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide a geographic breakdown of the 11,000 employees in sales, marketing, operations, and administrative positions, so the number outside Asia cannot be determined.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many of the 11,000 employees in those positions are located outside Asia.

512. PASS · Precision · `SEC\_10K\_multi\_1453\_00`

- **Question:** What is the year for the December 31 balance sheet?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The balance sheet date is December 31, 2025.
- ✓ **VERITROOPER (PASS):** The balance sheets are dated December 31, 2025 and 2024.

513. FAIL · Precision · `SEC\_10K\_multi\_1210\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total principal amount outstanding?
- **Ground truth:** 306,687
- ✗ **Baseline (vanilla RAG) (FAIL):** The total principal amount outstanding was **\$306.687 billion as of December 31, 2025**, compared with **\$266.202 billion in 2024**.
- ✓ **VERITROOPER (PASS):** The total principal amount outstanding was **\$306,687 million**.

514. PASS · Exception · `SEC\_10K\_multi\_0328\_01`

- **Question:** What does not apply to the net charge-offs?
- **Ground truth:** \$617 million
- ✓ **Baseline (vanilla RAG) (PASS):** Net charge-offs do **not include gains or losses from sales of nonaccrual loans recognized in noninterest revenue**.
- ✗ **VERITROOPER (FAIL):** The \$1.1 billion amount does not apply to net charge-offs; it was the net addition to the allowance for credit losses.

515. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0322\_04`

CERTAINTY MISMATCH

- **Question:** What is the name of the specific form number for the hedging activities?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The specific form is **Form 10-K**.
- **✓ VERITROOPER (PASS):** The evidence does not specify a form number for the hedging activities.

516. PASS · Cause & Effect · `SEC\_10K\_multi\_0636\_00`

- **Question:** What causes earnings volatility in the Firm's view?
- **Ground truth:** The asymmetry in accounting treatment between loans and lending-related commitments and the credit derivatives used in credit portfolio management activities.
- **✓ Baseline (vanilla RAG) (PASS):** The accounting asymmetry between loans and lending-related commitments, which are accounted for on an accrual basis, and credit derivatives used to manage them, which are reported at fair value with gains and losses recognized in principal transactions revenue, causes earnings volatility.
- **✓ VERITROOPER (PASS):** Earnings volatility is caused by the accounting asymmetry between accrual-based loans and lending-related commitments and fair-value credit derivatives, whose gains and losses are recognized in principal transactions revenue.

517. PASS · Exception · `SEC\_10K\_multi\_1526\_00`

- **Question:** What is excluded from the definition of investment-grade?
- **Ground truth:** BBB-/Baa3 or lower
- **✓ Baseline (vanilla RAG) (PASS):** Ratings below BBB-/Baa3 are excluded from investment-grade and are classified as noninvestment-grade.
- **✓ VERITROOPER (PASS):** Ratings below BBB-/Baa3 are excluded from JPMorganChase's definition of investment-grade.

518. PASS · Cause & Effect · `SEC\_10K\_multi\_1420\_01`

- **Question:** What happens as a consequence of the Firm cooperating with the inquiry into the Amrapali matter?
- **Ground truth:** The Firm is cooperating with the inquiry.
- **✓ Baseline (vanilla RAG) (PASS):** The Enforcement Directorate fined J.P. Morgan India Private Limited approximately \$31.5 million in August 2021. The Firm is appealing the fine.
- **✓ VERITROOPER (PASS):** The evidence does not state any consequence resulting from the Firm's cooperation with the Amrapali inquiry.

519. PASS · Conditional · `SEC\_10K\_multi\_0004\_01`

- **Question:** If the Company is unable to obtain components in sufficient quantities on commercially reasonable terms, what could happen to its business?
- **Ground truth:** It could be materially adversely affected.
- **✓ Baseline (vanilla RAG) (PASS):** Supply shortages and price increases could materially adversely affect the Company's business, results of operations, financial condition, and stock price.
- **✓ VERITROOPER (PASS):** Its business, results of operations, financial condition, and stock price could be materially adversely affected.

520. PASS · Exception · `SEC\_10K\_multi\_0141\_01`

- **Question:** What does NOT apply to the Firm's employee share-based incentive plans?
- **Ground truth:** Securities authorized for issuance under the Firm's employee share-based incentive plans
- ✓ **Baseline (vanilla RAG) (PASS):** They do **not apply to non-employees**, except non-employee directors.
- ✓ **VERITROOPER (PASS):** The evidence does not specify what does not apply to the Firm's employee share-based incentive plans.

521. FAIL · Cross-Reference · `SEC\_10K\_multi\_1286\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the estimated future amortization expense for the year 2030, and how does it compare to the gross carrying values of other intangible assets?
- **Ground truth:** The estimated future amortization expense for 2030 is \$100 million, which is less than the gross carrying values of \$3.5 billion as of December 31, 2025.
- ✗ **Baseline (vanilla RAG) (FAIL):** The estimated future amortization expense for 2030 is **\$100 million**. This is substantially below the gross carrying values of other intangible assets: **\$3.5 billion in 2025** and **\$3.8 billion in 2024**-about **2.9%** and **2.6%**, respectively.
- ✓ **VERITROOPER (PASS):** The estimated amortization expense for 2030 is **\$100 million**, compared with gross carrying values of other intangible assets of **\$3.5 billion in 2025** and **\$3.8 billion in 2024**.

522. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1393\_01`

- **Question:** What percentage of off-balance sheet lending-related financial instruments are expected to be utilized at the same time?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** No specific percentage is provided. The Firm states that it does not anticipate all available credit lines will be utilized simultaneously.
- ✓ **VERITROOPER (PASS):** The evidence does not specify what percentage is expected to be utilized at the same time.

523. PASS · Conditional · `SEC\_10K\_multi\_0866\_01`

- **Question:** If the Firm elects to measure certain instruments at fair value, does it include loans purchased or originated as part of securitization warehousing activity?
- **Ground truth:** yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. The fair value election includes loans purchased or originated as part of securitization warehousing activity, including related lending commitments.
- ✓ **VERITROOPER (PASS):** Yes. The fair value election includes loans purchased or originated as part of securitization warehousing activity, including lending-related commitments.

524. FAIL · Cross-Reference · `SEC\_10K\_multi\_1591\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total notional value of derivatives related to the Company's foreign currency economic hedges as of December 31, 2025, and how does it compare to the value as of December 31, 2024?
- **Ground truth:** The total notional value of derivatives related to the Company's foreign currency economic hedges was \$9,744 million as of December 31, 2025, and \$8,620 million as of December 31, 2024.
- ✗ **Baseline (vanilla RAG) (FAIL):** The total notional value was **\$9.744 billion** as of December 31, 2025, compared with **\$8.620 billion** as of December 31, 2024-an increase of **\$1.124 billion**, or approximately **13%**.

- ✖ **VERITROOPER (FAIL):** The total notional value was \$9,744 million as of December 31, 2025, compared with \$8,620 million in 2024—an increase of \$1,124 million.

525. PASS · Cross-Reference · `SEC\_10K\_multi\_1416\_00`

- **Question:** What is the nature of the commitments under multipurpose facilities, as described in section (k), and how do they relate to the issuance of standby letters of credit?
- **Ground truth:** The commitments under multipurpose facilities could be drawn upon in several forms, including the issuance of a standby letter of credit.
- ✓ **Baseline (vanilla RAG) (PASS):** Commitments under multipurpose facilities are flexible credit commitments that may be drawn in several forms. One permitted form is the issuance of a standby letter of credit, under which JPMorgan guarantees a client's performance or payment obligation to a third party.
- ✓ **VERITROOPER (PASS):** Multipurpose facilities are unfunded commitments that may be drawn in several forms, including through the issuance of standby letters of credit. These letters conditionally guarantee a client's or customer's performance to a third party.

526. FAIL · Precision · `SEC\_10K\_multi\_0487\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the amount of the allowance for credit losses in 2024?
- **Ground truth:** 541 million
- ✖ **Baseline (vanilla RAG) (FAIL):** The allowance for credit losses in 2024 was **\$26.8 billion**.
- ✓ **VERITROOPER (PASS):** The allowance for credit losses in 2024 was **\$541 million**.

527. PASS · Cause & Effect · `SEC\_10K\_multi\_0006\_00`

- **Question:** What happens as a result of the purchasing preferences and behaviors of consumers changing?
- **Ground truth:** Materially adverse impact on the Company's business, results of operations, financial condition, and stock price.
- ✓ **Baseline (vanilla RAG) (PASS):** Resellers may reduce orders or marketing of the Company's products, potentially materially harming its business, operating results, financial condition, and stock price.
- ✓ **VERITROOPER (PASS):** Resellers may reduce their ordering and marketing of the Company's products, potentially materially harming its business, results of operations, financial condition, and stock price.

528. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0080\_04`

VALIDATOR OVER-FIRED

- **Question:** What is the percentage of NVIDIA's operating results that could be impacted by the outcome of the legal proceedings?
- **Ground truth:** UNANSWERABLE
- ✖ **Baseline (vanilla RAG) (FAIL):** NVIDIA did not quantify a percentage. It stated that an unfavorable outcome could have a **material adverse impact** on operating results, but any possible loss or range of loss could not be reasonably estimated.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a percentage of operating results that could be impacted by the legal proceedings.

529. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0576\_00`

- **Question:** What is the total value of mortgage loans past due and insured by U.S. government agencies at December 31, 2026?
- **Ground truth:** UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide a value for December 31, 2026. At December 31, 2025, the total was **\$198 million**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the value as of December 31, 2026; it only reports \$198 million for 2025 and \$121 million for 2024.

530. PASS · Conditional · `SEC\_10K\_multi\_1035\_00`

- **Question:** If investment management fees are based on the value of assets under management, when are they collected and recognized?
- **Ground truth:** At the end of each period over which the management services are provided and the value of the managed assets is known.
- ✓ **Baseline (vanilla RAG) (PASS):** They are collected and recognized at the end of each period over which management services are provided, once the value of the managed assets is known.
- ✓ **VERITROOPER (PASS):** They are collected and recognized at the end of each service period, once the value of the managed assets is known.

531. FAIL · Precision · `SEC\_10K\_multi\_1320\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the fixed rate for Series DD?
- **Ground truth:** 169,625
- ✗ **Baseline (vanilla RAG) (FAIL):** The fixed rate for Series DD is **5.75%**.
- ✗ **VERITROOPER (FAIL):** The fixed rate for Series DD is **5.750%**.

532. FAIL · Precision · `SEC\_10K\_multi\_0259\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total noninterest expense?
- **Ground truth:** 95,640
- ✗ **Baseline (vanilla RAG) (FAIL):** Total noninterest expense was **\$95.64 billion in 2025**.
- ✓ **VERITROOPER (PASS):** Total noninterest expense was \$95,640, \$91,797, and \$87,172 for the three periods shown.

533. PASS · Exception · `SEC\_10K\_multi\_0382\_01`

- **Question:** What share classes are used as the primary share class for European funds?
- **Ground truth:** C share class
- ✓ **Baseline (vanilla RAG) (PASS):** The C share class is used as the primary share class for European funds.
- ✓ **VERITROOPER (PASS):** The C share class is used as the primary share class for European funds.

534. PASS · Cross-Reference · `SEC\_10K\_multi\_0712\_00`

- **Question:** What is the impact on the rewards liability if the redemption rate assumption is increased by 25 basis points and the cost per point assumption is increased by 1 basis point?
- **Ground truth:** approximately \$512 million
- ✓ **Baseline (vanilla RAG) (PASS):** The rewards liability would increase by approximately **\$512 million**.
- ✓ **VERITROOPER (PASS):** The combined increases would raise the rewards liability by approximately \$512 million as of December 31, 2025.

535. FAIL · Cross-Reference · `SEC\_10K\_multi\_1419\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the maximum exposure of stable value contracts with contractually limited exposure, and how does it compare to the notional amount?
- **Ground truth:** 1,312 million, which is a substantially lower percentage of the notional amount
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, the maximum exposure was **\$1.312 billion**, compared with a **\$35.462 billion notional amount**-about **3.7% of notional**. In 2024, it was **\$1.740 billion** versus **\$32.939 billion**, or about **5.3%**.
- **✗ VERITROOPER (FAIL):** At December 31, 2025, maximum exposure was \$1.312 billion, compared with a \$35.462 billion notional amount-about 3.7% of notional.

536. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0581\_00`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of loans held-for-sale?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, total loans held-for-sale were **\$13.84 billion**.
- **✓ VERITROOPER (PASS):** The evidence does not provide the total value of loans held-for-sale.

537. PASS · Cross-Reference · `SEC\_10K\_multi\_1171\_00`

- **Question:** What percentage of loan modifications were single modifications, and how many months were extended on average?
- **Ground truth:** 0.75%, 19 months
- **✓ Baseline (vanilla RAG) (PASS):** For secured-by-real-estate loans in 2025, single term-extension modifications represented **0.44% of retained loans**, with terms extended by a weighted average of **14 months**.
- **✓ VERITROOPER (PASS):** Single term-extension modifications represented 0.75% of retained commercial and industrial loans, with an average extension of 19 months.

538. PASS · Cause & Effect · `SEC\_10K\_multi\_1421\_00`

- **Question:** What happens as a result of the European Commission's decision against JPMorganChase?
- **Ground truth:** a fine is imposed
- **✓ Baseline (vanilla RAG) (PASS):** The European Commission found JPMorganChase infringed EU antitrust rules relating to EURIBOR and imposed a fine. Although the General Court annulled the original fine in December 2023, it reimposed an identical fine. JPMorganChase appealed; after a January 2026 hearing, judgment remains pending.
- **✓ VERITROOPER (PASS):** JPMorganChase appealed the decision; the European General Court annulled but re-imposed the same fine, prompting a further appeal to the EU Court of Justice.

539. PASS · Conditional · `SEC\_10K\_multi\_1545\_01`

- **Question:** If the Company's tax positions are not sustained on appeal, what would happen to the Company?
- **Ground truth:** The Company would not receive a refund of the applicable portion or all of the \$6.0 billion it paid in response to the IRS invoices issued in September 2024 and the related accrued interest receivable of \$385 million as of December 31, 2025.
- **✓ Baseline (vanilla RAG) (PASS):** If the Company's tax positions are not sustained on appeal:

- It would not receive a refund of some or all of the \$6.0 billion paid to the IRS, nor the related \$385 million accrued interest receivable.
- It could face approximately \$14 billion in additional tax and interest liabilities for 2010-2025, with interest continuing to accrue until paid.
- These liabilities could materially adversely affect its financial position, results of operations, and cash flows.
- ✓ **VERITROOPER (PASS)**: The Company could lose some or all of its \$6.0 billion tax payment and related \$385 million accrued interest receivable, and likely face significant additional liabilities for subsequent years, materially harming its financial position, operating results, and cash flows.

540. PASS · Precision · `SEC\_10K\_multi\_1350\_01`

- **Question**: What is the effective tax rate for the second period?
- **Ground truth**: 22.1
- ✓ **Baseline (vanilla RAG) (PASS)**: The effective tax rate for the second period, 2024, was **22.1%**.
- ✓ **VERITROOPER (PASS)**: The effective tax rate for the second period is **22.1%**.

541. PASS · Exception · `SEC\_10K\_multi\_1595\_00`

- **Question**: Under what circumstance do recipients not forfeit restricted stock, restricted stock units, and matching share awards?
- **Ground truth**: death or disability
- ✓ **Baseline (vanilla RAG) (PASS)**: Recipients may avoid forfeiture upon termination of employment in certain circumstances, such as **death or disability**.
- ✓ **VERITROOPER (PASS)**: Recipients may avoid forfeiture upon termination of employment in certain circumstances, such as death or disability.

542. PASS · Conditional · `SEC\_10K\_multi\_0509\_01`

- **Question**: If the Firm's liquidity management objectives include ensuring the Firm's core businesses can operate during stress events, does it analyze the liquidity characteristics of its assets and liabilities?
- **Ground truth**: Yes
- ✓ **Baseline (vanilla RAG) (PASS)**: Yes. The Firm analyzes the liquidity characteristics of its assets and liabilities across lines of business, legal entities, and currencies, considering legal, regulatory, and operational restrictions.
- ✓ **VERITROOPER (PASS)**: Yes. The Firm analyzes the liquidity characteristics of its assets and liabilities, including across business lines, legal entities, and currencies.

543. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0135\_06`

- **Question**: How many employees of JPMorganChase are involved in the development and execution of business strategies?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not specify how many JPMorganChase employees are involved in developing and executing business strategies.
- ✓ **VERITROOPER (PASS)**: The evidence does not state how many employees are involved in developing and executing business strategies.

544. FAIL · Cross-Reference · `SEC\_10K\_multi\_0080\_01`

UNIT/FORMAT MISMATCH

- **Question**: What is the income tax expense for the year ended January 25, 2026, and how does it relate to the U.S. federal statutory tax rate?

- **Ground truth:** The income tax expense for the year ended January 25, 2026, is \$21,383, which is \$8,257 less than the \$29,704 computed by applying the U.S. federal statutory tax rate of 21.0% to income before income taxes.
- **✗ Baseline (vanilla RAG) (FAIL):** For the year ended January 25, 2026, income tax expense was **\$21.383 billion**, representing an effective tax rate of **15.1%**. This was **5.9 percentage points below** the U.S. federal statutory tax rate of **21.0%**, primarily due to benefits from foreign-derived deduction eligible income, stock-based compensation, lower-tax foreign earnings, and tax credits.
- **✗ VERITROOPER (FAIL):** For the year ended January 25, 2026, income tax expense was \$21.383 billion, representing an effective tax rate of 15.1%, compared with the U.S. federal statutory rate of 21.0%.

545. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0781\_01`

HALLUCINATION

- **Question:** How many U.S. government agency securities are listed?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** None. U.S. government agency securities are excluded from the tables.
- **✗ VERITROOPER (FAIL):** 488 U.S. GSE and government agency securities are listed.

546. FAIL · Cross-Reference · `SEC\_10K\_multi\_1601\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total amount of assets measured at fair value on a recurring basis at December 31, 2025, according to Note 110?
- **Ground truth:** 4,451
- **✗ Baseline (vanilla RAG) (FAIL):** The total assets measured at fair value on a recurring basis were **\$4.451 billion** as of December 31, 2025.
- **✓ VERITROOPER (PASS):** The total assets measured at fair value on a recurring basis at December 31, 2025 were **\$4,451 million**.

547. PASS · Conditional · `SEC\_10K\_multi\_0669\_01`

- **Question:** Only if the refined time series was effective at the beginning of each year presented, would the CIB trading VaR by risk type: Fixed income be lower?
- **Ground truth:** Yes
- **✓ Baseline (vanilla RAG) (PASS):** Yes. Average CIB trading VaR for fixed income would have been lower by **\$1 million in 2025** and **\$3 million in 2024**.
- **✓ VERITROOPER (PASS):** Yes. If effective at the beginning of each year presented, average CIB trading VaR for fixed income would have been lower, though the amount is not provided.

548. PASS · Precision · `SEC\_10K\_multi\_1183\_00`

- **Question:** What is the total for defaults of FDMs in 2023?
- **Ground truth:** 31
- **✓ Baseline (vanilla RAG) (PASS):** The total defaults of retained wholesale FDMs in 2023 were **\$87 million** (\$6 million + \$50 million + \$31 million).
- **✓ VERITROOPER (PASS):** The total defaults of FDMs in 2023 were **\$87 million**.

549. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0027\_04`

- **Question:** What is the total value of Apple's foreign tax credit carryforwards in 2023?
- **Ground truth:** UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide Apple's foreign tax credit carryforwards for 2023. It only states that Apple had **\$4.7 billion** in Irish foreign tax credit carryforwards as of September 27, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide Apple's foreign tax credit carryforwards for 2023.

550. FAIL · Precision · `SEC\_10K\_multi\_0378\_00`

WRONG ENTITY

- **Question:** What is the revenue for 2023?
- **Ground truth:** 11,826
- ✗ **Baseline (vanilla RAG) (FAIL):** Apple's 2023 total net sales (revenue) were **\$383.285 billion**.
- ✓ **VERITROOPER (PASS):** Asset management fee revenue for 2023 was **\$11,826 million**.

551. PASS · Cross-Reference · `SEC\_10K\_multi\_0956\_00`

- **Question:** What is the purpose of the tables provided in the evidence, and how do they relate to the Consolidated statements of income?
- **Ground truth:** The tables provide information related to gains and losses recorded on derivatives based on their hedge accounting designation or purpose, and they are used to present derivative instruments in the Consolidated statements of income.
- ✓ **Baseline (vanilla RAG) (PASS):** The tables quantify gains and losses from derivative instruments according to their hedge-accounting designation or purpose. For fair value hedges, they show the derivatives by contract type, the pre-tax gains or losses on those derivatives, and the corresponding gains or losses on the hedged items for 2025, 2024, and 2023.

These amounts are reflected in the Consolidated statements of income, with gains or losses on the hedging derivative recorded in the same income-statement line item as the related hedged item.

- ✓ **VERITROOPER (PASS):** The tables show gains and losses recorded on derivatives based on their hedge accounting designation or purpose. Gains or losses on hedging derivatives are included in the same line item in the Consolidated statements of income as the related hedged item.

552. PASS · Conditional · `SEC\_10K\_multi\_0898\_01`

- **Question:** If the fair value option was elected for lending-related commitments at December 31, 2024, what was the corresponding fair value?
- **Ground truth:** \$50 million
- ✓ **Baseline (vanilla RAG) (PASS):** The corresponding fair value was **\$50 million** at December 31, 2024.
- ✓ **VERITROOPER (PASS):** The corresponding fair value was **\$50 million**.

553. FAIL · Cause & Effect · `SEC\_10K\_multi\_0086\_01`

BAD QUESTION

- **Question:** What is the result of operating leases for our offices and data centers?
- **Ground truth:** Our lease obligations primarily consist of operating leases.
- ✗ **Baseline (vanilla RAG) (FAIL):** As of January 25, 2026, operating leases for offices and data centers resulted in:
 - **\$3.624 billion** of future minimum lease obligations
 - **\$2.944 billion** present value after \$680 million of imputed interest
 - **\$372 million** of short-term liabilities
 - **\$2.572 billion** of long-term operating lease liabilities

The leases had a weighted-average remaining term of **8.8 years** and a weighted-average discount rate of **4.38%**.

- ✖ **VERITROOPER (FAIL):** As of January 25, 2026, office and data center operating leases resulted in \$3.624 billion of future minimum obligations, with a present value of \$2.944 billion and long-term liabilities of \$2.572 billion.

554. PASS · Precision · `SEC\_10K\_multi\_0724\_00`

- **Question:** What is the number of shares of preferred stock authorized?

- **Ground truth:** 200,000,000

- ✓ **Baseline (vanilla RAG) (PASS):** JPMorgan Chase was authorized to issue **200 million shares of preferred stock**.

- ✓ **VERITROOPER (PASS):** 200,000,000 shares.

555. PASS · Precision · `SEC\_10K\_multi\_0066\_01`

- **Question:** What is the weighted average expected life for ESPP shares?

- **Ground truth:** 0.1 - 2.0 years

- ✓ **Baseline (vanilla RAG) (PASS):** The weighted-average expected life for ESPP shares is **0.1 to 2.0 years**

- ✓ **VERITROOPER (PASS):** The weighted average expected life for ESPP shares was **0.1-2.0 years** for each period presented.

556. PASS · Conditional · `SEC\_10K\_multi\_1151\_01`

- **Question:** Only if the borrower makes the contractual payments when due under the modified payment terms, does the Firm reinstate the borrower's line of credit?

- **Ground truth:** No

- ✓ **Baseline (vanilla RAG) (PASS):** No. In most cases, the Firm does not reinstate the borrower's line of credit, even if contractual payments are made when due under the modified terms.

- ✓ **VERITROOPER (PASS):** No. The evidence states that, in most cases, the Firm does not reinstate the borrower's line of credit, but it does not tie reinstatement to making contractual payments.

557. PASS · Conditional · `SEC\_10K\_multi\_1590\_01`

- **Question:** If the Company has entered into commodity futures contracts and other derivative instruments on various commodities, what is the objective of this hedging program?

- **Ground truth:** to reduce the variability of cash flows associated with future purchases of certain commodities

- ✓ **Baseline (vanilla RAG) (PASS):** The objective is to reduce the variability of cash flows associated with forecasted purchases of commodities used in the Company's manufacturing process.

- ✓ **VERITROOPER (PASS):** The objective is to reduce the variability of cash flows associated with future purchases of certain commodities.

558. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0744\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of derivative netting adjustments?

- **Ground truth:** UNANSWERABLE

- ✖ **Baseline (vanilla RAG) (FAIL):** The total derivative netting adjustments were **\$1.357 billion** in 2025, comprising **\$403 million for assets** and **\$954 million for liabilities**.

- ✖ **VERITROOPER (FAIL):** The total derivative netting adjustments were \$(550,409) million.

559. PASS · Precision · `SEC\_10K\_multi\_0057\_01`

- **Question:** What is the amount of the charge associated with H2O excess inventory and purchase obligations in the first quarter of fiscal year 2026?
- **Ground truth:** 4.5 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The charge was **\$4.5 billion**.
- ✓ **VERITROOPER (PASS):** The charge was \$4.5 billion.

560. PASS · Cross-Reference · `SEC\_10K\_multi\_0009\_00`

- **Question:** What is the maximum amount the Company may be fined if the Commission makes a final determination in the Article 6(4) Investigation?
- **Ground truth:** 10% of the Company's annual worldwide net sales
- ✓ **Baseline (vanilla RAG) (PASS):** The Commission may impose a fine of up to **10% of the Company's annual worldwide net sales**.
- ✓ **VERITROOPER (PASS):** The Commission may impose a fine of up to 10% of the Company's annual worldwide net sales.

561. PASS · Cause & Effect · `SEC\_10K\_multi\_0003\_00`

- **Question:** What happens as a consequence of adverse macroeconomic conditions?
- **Ground truth:** Materially adversely affect the Company's business, results of operations, financial condition, and stock price.
- ✓ **Baseline (vanilla RAG) (PASS):** Adverse macroeconomic conditions can:
 - Reduce consumer confidence, disposable income, spending, and demand for products and services.
 - Increase unemployment, delinquencies, credit losses, allowances, and net charge-offs.
 - Lower transaction volumes, advisory activity, refinancing, asset values, and fee revenue, while causing investment outflows or losses on credit commitments.
 - Disrupt suppliers, manufacturers, logistics providers, distributors, and other business partners.
 - Ultimately reduce revenue and earnings and adversely affect business operations, financial condition, and stock price.
- ✓ **VERITROOPER (PASS):** Adverse macroeconomic conditions can reduce consumer confidence, spending, and demand for Apple's products and services, materially harming its business, results of operations, financial condition, and stock price. They can also increase credit and collectibility risks, reduce liquidity, limit debt issuance, and destabilize suppliers and other partners.

562. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1135\_01`

- **Question:** What is the average FICO score for all retained loans?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The weighted-average FICO score for retained residential real estate loans was **775 as of December 31, 2025** (774 in 2024). The evidence does not provide an average FICO score across all retained loan classes.
- ✗ **VERITROOPER (FAIL):** The weighted-average FICO scores were 775 in Col. 1 and 774 in Col. 2.

563. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0744\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of beneficial interests issued by consolidated VIEs at Level 1 of the fair value hierarchy?
- **Ground truth:** UNANSWERABLE

- ✖ **Baseline (vanilla RAG) (FAIL):** The Level 1 amount for beneficial interests issued by consolidated VIEs was \$0.
- ✖ **VERITROOPER (FAIL):** No amount was reported at Level 1 (shown as "-").

564. PASS · Precision · `SEC\_10K\_multi\_0588\_01`

- **Question:** What was the total criticized exposure as of December 31, 2025?
- **Ground truth:** 48.5
- ✓ **Baseline (vanilla RAG) (PASS):** The total criticized exposure was **\$48.5 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total criticized exposure was **\$48.5 billion** as of December 31, 2025.

565. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0727\_03`

- **Question:** How many levels of management review are required for positions classified as level 3?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a fixed number. It only states that more significant or complex positions receive additional levels of management review.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many levels of management review are required for Level 3 positions.

566. PASS · Precision · `SEC\_10K\_multi\_0597\_01`

- **Question:** What is the threshold for Transportation?
- **Ground truth:** 7,135
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a threshold for Transportation.
- ✓ **VERITROOPER (PASS):** The evidence does not identify which column represents the threshold for Transportation.

567. PASS · Exception · `SEC\_10K\_multi\_0576\_00`

- **Question:** What type of loans are excluded from the general rule?
- **Ground truth:** mortgage loans past due and insured by U.S. government agencies
- ✓ **Baseline (vanilla RAG) (PASS):** Credit card loans and certain consumer loans insured by U.S. government agencies.
- ✓ **VERITROOPER (PASS):** Mortgage loans that are past due and insured by U.S. government agencies are excluded because of the government guarantee.

568. PASS · Cross-Reference · `SEC\_10K\_multi\_0715\_01`

- **Question:** When does software cost capitalization begin, and what is the required effective date?
- **Ground truth:** Software cost capitalization begins when management has authorized and committed to funding the software project, and it is probable that the software will be completed and used to perform its intended function, with a required effective date of January 1, 2028.
- ✓ **Baseline (vanilla RAG) (PASS):** Software cost capitalization begins when:
 1. Management has authorized and committed funding for the software project; and
 2. It is probable that the software will be completed and used for its intended function.

The required effective date is **January 1, 2028**.

- ✓ **VERITROOPER (PASS):** Software cost capitalization begins when management has authorized and committed funding and it is probable the software will be completed and used for its intended function. The required effective date is January 1, 2028.

569. PASS · Precision · `SEC\_10K\_multi\_0380\_00`

- **Question:** What is the total net revenue for the year?
- **Ground truth:** 24,073
- ✓ **Baseline (vanilla RAG) (PASS):** The total net revenue for 2025 was **\$24.073 billion**.
- ✓ **VERITROOPER (PASS):** The total net revenue for the year was \$24,073.

570. PASS · Precision · `SEC\_10K\_multi\_0974\_00`

- **Question:** What is the year of the first date mentioned?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** 2025
- ✓ **VERITROOPER (PASS):** 2025

571. PASS · Precision · `SEC\_10K\_multi\_0473\_00`

- **Question:** What is the net benefit from the CECL capital transitions in 2024?
- **Ground truth:** 720 million
- ✓ **Baseline (vanilla RAG) (PASS):** The net benefit from the CECL capital transition provisions in 2024 was **\$720 million**.
- ✓ **VERITROOPER (PASS):** The net benefit from the CECL capital transitions in 2024 was **\$720 million**.

572. PASS · Cross-Reference · `SEC\_10K\_multi\_1243\_01`

- **Question:** For loan securitizations, what are the three accounting criteria for a sale that the Firm must meet?
- **Ground truth:** The criteria are: (1) the transferred financial assets are legally isolated from the Firm's creditors; (2) the transferee or beneficial interest holder can pledge or exchange the transferred financial assets; and (3) the Firm does not maintain effective control over the transferred financial assets.
- ✓ **Baseline (vanilla RAG) (PASS):** For a loan securitization to be accounted for as a sale:
 1. The transferred financial assets must be legally isolated from the Firm's creditors.
 2. The transferee or beneficial interest holder must be able to pledge or exchange the transferred financial assets.
 3. The Firm must not maintain effective control over the assets-for example, it cannot repurchase them before maturity or unilaterally require their return.
- ✓ **VERITROOPER (PASS):** The three criteria are: (1) the transferred financial assets are legally isolated from the Firm's creditors; (2) the transferee or beneficial interest holder can pledge or exchange the assets; and (3) the Firm does not maintain effective control over the assets.

573. PASS · Precision · `SEC\_10K\_multi\_0254\_00`

- **Question:** What is the year for which the results are reported?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The results are reported for 2025.
- ✓ **VERITROOPER (PASS):** The results are reported for the year ended December 31, 2025.

574. PASS · Precision · `SEC\_10K\_multi\_1157\_01`

- **Question:** What is the gross charge-offs for 2023?
- **Ground truth:** \$32
- ✓ **Baseline (vanilla RAG) (PASS):** Gross charge-offs for 2023 were **\$11 million**.
- ✓ **VERITROOPER (PASS):** Gross charge-offs for 2023 were **\$32 million**.

575. FAIL · Precision · `SEC\_10K\_multi\_0023\_01`

UNSUPPORTED NUMBER

- **Question:** What is the value of the total?
- **Ground truth:** 132,420
- **✗ Baseline (vanilla RAG) (FAIL):** The total net sales for 2025 were **\$416.161 billion** (\$416,161 million).
- **✗ VERITROOPER (FAIL):** The total is \$134,711.

576. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1274\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the percentage change in valuation due to inputs and assumptions?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** From 2024 to 2025, changes in valuation due to inputs and assumptions fell from **\$762 million to \$48 million**:

$$\left[\frac{48-762}{762} \right] \times 100 \approx -93.7\%$$

Thus, the valuation change **decreased by approximately 93.7%**.

- **✓ VERITROOPER (PASS):** The evidence does not specify the periods or baseline needed to calculate the percentage change in valuation due to inputs and assumptions.

577. FAIL · Cross-Reference · `SEC\_10K\_multi\_1062\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the expected long-term rate of return on plan assets for the year ended December 31, 2025, and how does it compare to the rate for 2024?
- **Ground truth:** 5.44% for 2025 and 6.15% for 2024.
- **✗ Baseline (vanilla RAG) (FAIL):** The expected long-term rate of return on plan assets was **5.44% in 2025**, down from **6.15% in 2024**, a decrease of **0.71 percentage points**.
- **✗ VERITROOPER (FAIL):** The expected long-term rate of return was 5.44% in 2025, down from 6.15% in 2024-a decrease of 0.71 percentage points.

578. PASS · Cause & Effect · `SEC\_10K\_multi\_0706\_01`

- **Question:** What is the consequence of the Firm not being successful in preventing or mitigating a future cybersecurity incident?
- **Ground truth:** a material adverse effect on the Firm or its business strategy, results of operations or financial condition
- **✓ Baseline (vanilla RAG) (PASS):** A future cybersecurity incident could materially adversely affect the Firm, its business strategy, results of operations, or financial condition.
- **✓ VERITROOPER (PASS):** A future cybersecurity incident could materially adversely affect the Firm, its business strategy, results of operations, or financial condition.

579. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0744\_02`

BAD QUESTION

- **Question:** What is the total amount of long-term debt measured at fair value on a recurring basis?
- **Ground truth:** UNANSWERABLE

- ✖ **Baseline (vanilla RAG) (FAIL):** The total long-term debt measured at fair value on a recurring basis was **\$134.559 billion** as of December 31, 2025.
- ✖ **VERITROOPER (FAIL):** The total long-term debt measured at fair value on a recurring basis was **\$134,559 million**.

580. PASS · Cross-Reference · `SEC\_10K\_multi\_0250\_01`

- **Question:** What drove the decrease in the effective tax rate in 2025, and what was the income tax benefit related to it?
- **Ground truth:** The decrease was driven by a \$774 million income tax benefit, which was related to the resolution of certain tax audits and the impact of tax regulations related to foreign currency translation gains and losses.
- ✓ **Baseline (vanilla RAG) (PASS):** The 2025 effective tax rate decreased primarily due to the State Aid Decision and changes in unrecognized tax benefits. The State Aid Decision resulted in a **\$10.7 billion year-over-year decrease in income tax provision**.
- ✓ **VERITROOPER (PASS):** The decrease was driven by resolving certain tax audits, the impact of foreign-currency translation tax regulations effective in 2025, and higher tax benefits from vesting employee share-based awards. The related income tax benefit recorded in Corporate was \$774 million.

581. PASS · Precision · `SEC\_10K\_multi\_0937\_00`

- **Question:** What is the number of OTC-cleared foreign exchange contracts?
- **Ground truth:** 904
- ✓ **Baseline (vanilla RAG) (PASS):** OTC-cleared foreign exchange contracts were **\$904 million in 2025**, compared with **\$685 million in 2024** (gross derivative receivables).
- ✓ **VERITROOPER (PASS):** OTC-cleared foreign exchange contracts were 904 in Col 1 and 685 in Col 4.

582. PASS · Conditional · `SEC\_10K\_multi\_1154\_00`

- **Question:** If a borrower misses two consecutive contractual payments, what is deemed to have occurred?
- **Ground truth:** payment default
- ✓ **Baseline (vanilla RAG) (PASS):** A payment default is deemed to have occurred.
- ✓ **VERITROOPER (PASS):** A payment default is deemed to have occurred.

583. PASS · Precision · `SEC\_10K\_multi\_0164\_00`

- **Question:** What is the year in which the Deferred Compensation Plan was amended and restated?
- **Ground truth:** 2001
- ✓ **Baseline (vanilla RAG) (PASS):** The Deferred Compensation Plan was amended and restated in **2024**.
- ✓ **VERITROOPER (PASS):** The Deferred Compensation Plan was amended and restated in 2001.

584. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average interest rate of the notes issued by Apple Inc. in 2025?
- **Ground truth:** UNANSWERABLE
- ✖ **Baseline (vanilla RAG) (FAIL):** Apple's 2025 notes had fixed rates ranging from **4.00% to 4.75%**, giving a simple average of **4.375%**.
- ✖ **VERITROOPER (FAIL):** The average interest rate of Apple's four notes issued in 2025 was 4.3625%.

585. PASS · Exception · `SEC\_10K\_multi\_0382\_00`

- **Question:** What funds are excluded from the rankings?
- **Ground truth:** Funds with fewer than three years of history
- ✓ **Baseline (vanilla RAG) (PASS):** Funds with fewer than three years of performance history are not rated and are excluded from the rankings.
- ✓ **VERITROOPER (PASS):** Funds with fewer than three years of history are excluded from the rankings.

586. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0539\_03`

- **Question:** What is the specific impact of a credit rating downgrade on the Firm's funding requirements for VIEs?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** A credit rating downgrade may adversely affect the Firm's funding requirements for VIEs and other third-party commitments. The evidence does not quantify the specific impact.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the exact impact; it only states that funding requirements for VIEs may be adversely affected by a credit rating decline.

587. PASS · Precision · `SEC\_10K\_multi\_0473\_01`

- **Question:** What is the net reduction for deferred tax assets related to tax attribute carryforwards in 2025?
- **Ground truth:** 1.8 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The net reduction was **\$1.8 billion** in 2025.
- ✓ **VERITROOPER (PASS):** The net reduction was **\$1.8 billion** as of December 31, 2025.

588. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1609\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of the notes issued by the Company with a maturity date of 2030?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The notes maturing in 2030 total **\$1.5 billion**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the principal values of the notes maturing in 2030, so their total value cannot be determined.

589. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_01`

- **Question:** How many employees are involved in the Company's cybersecurity efforts?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify how many employees are involved; it only describes a team of cybersecurity, IT, engineering, operations, and legal professionals.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many employees are involved in the Company's cybersecurity efforts.

590. PASS · Precision · `SEC\_10K\_multi\_1358\_00`

- **Question:** What is the percentage limit for covered transactions between a banking subsidiary and its affiliates?
- **Ground truth:** 20
- ✓ **Baseline (vanilla RAG) (PASS):** Covered transactions are limited to:
 - **10%** of the banking subsidiary's total capital for any single affiliate.
 - **20%** of total capital in aggregate for all affiliates.

- ✓ **VERITROOPER (PASS):** Covered transactions with any particular affiliate are generally limited to 10% of the banking subsidiary's total capital, while aggregate transactions with all affiliates are limited to 20%.

591. FAIL · Cross-Reference · `SEC\_10K\_multi\_1471\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total value of lease right-of-use assets and the related tax-oriented investment liabilities?
- **Ground truth:** \$683 million and \$669 million
- ✗ **Baseline (vanilla RAG) (FAIL):** The total is **\$1.352 billion**: \$683 million in lease right-of-use assets plus \$669 million in tax-oriented investment liabilities.
- ✗ **VERITROOPER (FAIL):** The total is **\$1.352 billion** (\$683 million + \$669 million).

592. PASS · Precision · `SEC\_10K\_multi\_1323\_01`

- **Question:** What is the SOFR rate for Series HH?
- **Ground truth:** 3.125
- ✓ **Baseline (vanilla RAG) (PASS):** Series HH's floating rate was **three-month term SOFR plus 3.125% annually**.
- ✓ **VERITROOPER (PASS):** The SOFR rate for Series HH is **SOFR + 3.125%**.

593. PASS · Precision · `SEC\_10K\_multi\_0925\_01`

- **Question:** What is the year for which the notional amount of free-standing derivative contracts is reported in the previous year?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The previous year reported is **2024**.
- ✓ **VERITROOPER (PASS):** The previous year reported is 2024.

594. PASS · Cross-Reference · `SEC\_10K\_multi\_1549\_01`

- **Question:** What is the percentage of the Company's consolidated projected benefit obligation represented by the U.S. qualified pension plan, and what is the percentage of the Company's pension plan assets represented by this plan?
- **Ground truth:** 63% and 60%, respectively
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, the U.S. qualified pension plan represented **63%** of the consolidated projected benefit obligation and **60%** of pension plan assets.
- ✓ **VERITROOPER (PASS):** As of December 31, 2025, the U.S. qualified pension plan represented 63% of the Company's consolidated projected benefit obligation and 60% of its pension plan assets.

595. PASS · Cross-Reference · `SEC\_10K\_multi\_0492\_00`

- **Question:** What types of changes are considered for Credit risk RWA?
- **Ground truth:** Changes in book size, composition, credit quality, market movements, and deductions for excess eligible allowances for credit losses.
- ✓ **Baseline (vanilla RAG) (PASS):** For Credit risk RWA, changes considered include:
 - Changes in book size, including the Apple Card transaction
 - Changes in portfolio composition and credit quality
 - Market movements
 - Deductions for excess eligible allowances for credit losses that are not eligible for inclusion in Tier 2 capital
 - Model and data changes arising from revised methodologies or regulatory guidance

- ✓ **VERITROOPER (PASS)**: Credit risk RWA considers changes in book size (including the Apple Card transaction), composition and credit quality, market movements, and deductions for excess eligible allowances for credit losses not eligible for inclusion in Tier 2 capital.

596. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1537\_05`

- **Question**: What is the total number of strikes or work stoppages that have occurred at our major manufacturing facilities or bottling operations in the past year?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not specify any strikes or work stoppages that occurred in the past year; it only discusses them as potential risks.
- ✓ **VERITROOPER (PASS)**: The evidence does not state how many strikes or work stoppages occurred in the past year.

597. FAIL · Cross-Reference · `SEC\_10K\_multi\_0822\_00`

UNIT/FORMAT MISMATCH

- **Question**: What is the total realized/unrealized gains/(losses) for the year ended December 31, 2023, and what is the fair value at January 1, 2023?
- **Ground truth**: 738, 27,726
- ✗ **Baseline (vanilla RAG) (FAIL)**: For **deposits**, total realized/unrealized gains/(losses) were **\$95 million**, and fair value at January 1, 2023 was **\$2.162 billion**.
- ✓ **VERITROOPER (PASS)**: The total realized/unrealized gains were \$1,475 million, and the fair value at January 1, 2023, was \$27,726 million.

598. FAIL · Precision · `SEC\_10K\_multi\_0646\_00`

UNSUPPORTED NUMBER

- **Question**: What is the average retained loans?
- **Ground truth**: 371,238
- ✗ **Baseline (vanilla RAG) (FAIL)**: The average retained residential real estate FDM loans for 2023-2025 were approximately **\$455.3 million**:

$$\frac{(\$1,024M + \$206M + \$136M)}{3} = \$455.3M$$

- ✓ **VERITROOPER (PASS)**: Average retained loans were \$371,238, \$231,644, \$732,793, \$1,335,675, \$384,001, \$214,033, \$673,310, and \$1,271,344 across Columns 1-8, respectively.

599. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_00`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question**: What is the total number of employees at NVIDIA Corporation?
- **Ground truth**: UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL)**: As of December 31, 2025, NVIDIA Corporation had approximately **65,900 employees**.
- ✓ **VERITROOPER (PASS)**: The provided evidence does not state NVIDIA Corporation's total number of employees.

600. PASS · Precision · `SEC\_10K\_multi\_0151\_00`

- **Question:** What is the date of the report of PricewaterhouseCoopers LLP?
- **Ground truth:** February 13, 2026
- ✓ **Baseline (vanilla RAG) (PASS):** February 13, 2026.
- ✓ **VERITROOPER (PASS):** February 13, 2026.

601. FAIL · Conditional · `SEC\_10K\_multi\_1290\_01`

DETERMINISTIC:DIRECTION CONFLICT

- **Question:** If the lease term is 7 years and the estimated useful life of leasehold improvements is 10 years, does the Firm depreciate leasehold improvements over the lease term?
- **Ground truth:** No
- ✗ **Baseline (vanilla RAG) (FAIL):** Yes. The Firm depreciates leasehold improvements over the lesser of the remaining lease term or estimated useful life. Since 7 years is shorter than 10 years, it would depreciate them over the 7-year lease term.
- ✗ **VERITROOPER (FAIL):** Yes. The Firm uses the lesser of the remaining lease term or estimated useful life, so it would depreciate the improvements over the 7-year lease term.

602. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0607\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What percentage of the total credit exposure is attributed to loans and lending-related commitments?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Loans and lending-related commitments represent approximately **98.3%** of total Consumer & Retail credit exposure:

$$\frac{\$131.710\text{B}}{\$133.945\text{B}} \times 100 \approx 98.3\%$$

- ✓ **VERITROOPER (PASS):** The evidence does not provide the amounts needed to calculate the percentage of total credit exposure attributable to loans and lending-related commitments.

603. PASS · Cause & Effect · `SEC\_10K\_multi\_0003\_01`

- **Question:** What happens as a consequence of restrictions on international trade?
- **Ground truth:** Materially adversely affect the Company's business and supply chain.
- ✓ **Baseline (vanilla RAG) (PASS):** Restrictions on international trade can:
 - Increase costs or limit the availability of products, components, rare earths, and other raw materials.
 - Force companies to change suppliers, restructure operations and business relationships, withdraw affected products, services, or applications, and raise prices.
 - Disrupt supply chains and make restructuring expensive and time-consuming.
 - Reduce consumer spending and demand and worsen consumer sentiment.
 - Restrict cross-border transfers of earnings or capital and negatively affect profitability and overall business results.
- ✓ **VERITROOPER (PASS):** Restrictions on international trade can raise costs or limit product and material availability, disrupt supply chains, force operational or supplier changes, reduce offerings, and increase prices. These effects can materially harm Apple's business and results of operations.

604. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1595\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the average annual return on the Company's other postretirement benefit plans' assets for the year 2025?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The 2025 annual return was approximately **13.0%**, calculated as:

$$\frac{\$9\text{ billion}}{\$69\text{ billion}} \times 100 \approx 13.0\%$$

- **✓ VERITROOPER (PASS):** The evidence does not state the average annual return on other postretirement benefit plan assets for 2025.

605. FAIL · Cross-Reference · `SEC\_10K\_multi\_1587\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total amount of proceeds from disposals of businesses, equity method investments, and nonmarketable securities in 2025, and how does it compare to the total in 2024?
- **Ground truth:** The total amount of proceeds in 2025 is \$3,567 million, which is \$82 million more than the total in 2024, which was \$3,485 million.
- **✗ Baseline (vanilla RAG) (FAIL):** Proceeds totaled **\$3.567 billion in 2025**, compared with **\$3.485 billion in 2024**-an increase of **\$82 million**, or approximately **2.4%**.
- **✗ VERITROOPER (FAIL):** Proceeds totaled \$3.567 billion in 2025, which was \$82 million higher than the \$3.485 billion reported in 2024.

606. PASS · Cross-Reference · `SEC\_10K\_multi\_1274\_00`

- **Question:** What is the total change in valuation due to inputs and assumptions, and how does it compare to the total change due to other inputs and assumptions?
- **Ground truth:** 92, which is greater than 43
- **✓ Baseline (vanilla RAG) (PASS):** For 2025, the total change in valuation due to inputs and assumptions was **\$48 million**, compared with **\$0** from other inputs and assumptions. Thus, the total was **\$48 million higher**, entirely attributable to market interest rates and other factors.
- **✓ VERITROOPER (PASS):** The total change in valuation due to inputs and assumptions was \$48 million, compared with no reported change due to other inputs and assumptions.

607. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1585\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total value of tax credits received from limited partnerships in 2025?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The total tax credits and other income tax benefits received from limited partnerships in 2025 were **\$241 million**.
- **✓ VERITROOPER (PASS):** The evidence does not state the total value of tax credits received from limited partnerships in 2025.

608. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1149\_00`

- **Question:** What is the average FICO score of the credit card portfolio?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The average FICO score is not provided. The portfolio is reported by FICO bands: 84.6% had scores of 660 or higher, 15.2% below 660, and 0.2% had no FICO score as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the average FICO score of the credit card portfolio.

609. PASS · Cause & Effect · `SEC\_10K\_multi\_0044\_00`

- **Question:** What happens as a consequence of failing to meet evolving customer and industry needs?
- **Ground truth:** Our financial results may be adversely impacted.
- ✓ **Baseline (vanilla RAG) (PASS):** Failure to meet evolving customer and industry needs could harm financial results, including by reducing product adoption or revenue and negatively affecting the business, operating results, and financial condition.
- ✓ **VERITROOPER (PASS):** Failing to meet evolving customer and industry needs could harm the company's financial results.

610. FAIL · Precision · `SEC\_10K\_multi\_1453\_01`

UNSUPPORTED NUMBER

- **Question:** What is the amount of cash and due from banks for 2024?
- **Ground truth:** 38
- ✗ **Baseline (vanilla RAG) (FAIL):** Cash and due from banks and deposits with banks at December 31, 2024 was \$469.317 billion.
- ✓ **VERITROOPER (PASS):** Cash and due from banks for 2024 was \$38 million.

611. PASS · Precision · `SEC\_10K\_multi\_0033\_00`

- **Question:** What is the date of the Company's assessment of internal control over financial reporting?
- **Ground truth:** September 27, 2025
- ✓ **Baseline (vanilla RAG) (PASS):** September 27, 2025.
- ✓ **VERITROOPER (PASS):** September 27, 2025.

612. FAIL · Precision · `SEC\_10K\_multi\_0301\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the net charge-off/(recovery) rate for Auto?
- **Ground truth:** 0.49
- ✗ **Baseline (vanilla RAG) (FAIL):** The Auto net charge-off/(recovery) rate was 0.47% in 2025.
- ✓ **VERITROOPER (PASS):** The Auto net charge-off/(recovery) rates are 0.47%, 0.59%, and 0.49% for Columns 1, 2, and 3, respectively.

613. FAIL · Precision · `SEC\_10K\_multi\_0075\_01`

UNIT/FORMAT MISMATCH

- **Question:** What was the total of other long-term liabilities as of January 26, 2025?
- **Ground truth:** 4,245

- ✖ **Baseline (vanilla RAG) (FAIL):** The total other long-term liabilities as of January 26, 2025 were **\$4.245 billion**.
- ✖ **VERITROOPER (FAIL):** The total other long-term liabilities as of January 26, 2025 was **\$4.245 billion**.

614. PASS · Cause & Effect · `SEC\_10K\_multi\_1279\_01`

- **Question:** What is the consequence of the Firm not making advances in accordance with applicable rules and agreements?
- **Ground truth:** The advances may not be recoverable
- ✓ **Baseline (vanilla RAG) (PASS):** Certain servicer advances may not be recoverable.
- ✓ **VERITROOPER (PASS):** Certain servicer advances may not be recoverable.

615. FAIL · Cross-Reference · `SEC\_10K\_multi\_0019\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the total net sales for the years ended September 27, 2025, September 28, 2024, and September 30, 2023, and how do they compare to each other?
- **Ground truth:** \$416,161, \$391,035, and \$383,285
- ✖ **Baseline (vanilla RAG) (FAIL):** Total net sales were:
 - **2025:** \$416.161 billion
 - **2024:** \$391.035 billion
 - **2023:** \$383.285 billion

Net sales increased by **\$25.126 billion (6.4%)** from 2024 to 2025 and by **\$7.750 billion (2.0%)** from 2023 to 2024. Overall, sales increased **\$32.876 billion (8.6%)** from 2023 to 2025.

- ✖ **VERITROOPER (FAIL):** Total net sales were \$416.161 billion in 2025, \$391.035 billion in 2024, and \$383.285 billion in 2023. Sales increased by \$7.750 billion from 2023 to 2024 and by \$25.126 billion from 2024 to 2025.

616. PASS · Cause & Effect · `SEC\_10K\_multi\_0048\_00`

- **Question:** What happens as a consequence of losing a large customer?
- **Ground truth:** harm our financial condition and results of operations
- ✓ **Baseline (vanilla RAG) (PASS):** Losing a large customer could reduce demand and materially adversely affect future revenue and financial performance.
- ✓ **VERITROOPER (PASS):** Losing a large customer would likely harm the company's financial condition and results of operations.

617. PASS · Precision · `SEC\_10K\_multi\_0539\_00`

- **Question:** What is the year in which the credit ratings of the Parent Company and certain of its principal subsidiaries were reported?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The credit ratings were reported as of December 31, **2025**.
- ✓ **VERITROOPER (PASS):** The credit ratings were reported as of December 31, 2025.

618. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1136\_00`

- **Question:** What percentage of Chapter 7 residential real estate loans were 90 days or more past due at December 31, 2025?
- **Ground truth:** UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The percentage is not provided. The evidence only states that approximately 9% of Chapter 7 residential real estate loans were **30 days or more past due** at December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the percentage that were 90 days or more past due; it only states that 9% were 30 days or more past due.

619. PASS · Cross-Reference · `SEC\_10K\_multi\_1529\_00`

- **Question:** What is the filing status of Coca-Cola's report, and what is the relevant section of the Securities Exchange Act?
- **Ground truth:** Annual Report pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934.
- ✓ **Baseline (vanilla RAG) (PASS):** Coca-Cola filed an **Annual Report on Form 10-K** for the year ended December 31, 2025, pursuant to **Section 13 or 15(d) of the Securities Exchange Act of 1934**.
- ✓ **VERITROOPER (PASS):** Coca-Cola filed an Annual Report on Form 10-K under Section 13 or 15(d) of the Securities Exchange Act of 1934.

620. PASS · Precision · `SEC\_10K\_multi\_1299\_00`

- **Question:** What is the credit card rewards liability at December 31, 2024?
- **Ground truth:** 14.4 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The credit card rewards liability was **\$14.4 billion** at December 31, 2024.
- ✓ **VERITROOPER (PASS):** The credit card rewards liability was **\$14.4 billion** at December 31, 2024.

621. PASS · Cause & Effect · `SEC\_10K\_multi\_0924\_00`

- **Question:** What happens as a result of fluctuations in foreign currencies?
- **Ground truth:** The U.S. dollar-equivalent values of the foreign currency-denominated assets and liabilities or the forecasted revenues or expenses increase or decrease.
- ✓ **Baseline (vanilla RAG) (PASS):** Foreign-currency fluctuations can change U.S.-dollar-reported revenues, earnings, gross margins, and balance-sheet values. A stronger U.S. dollar generally reduces the value of foreign sales and earnings and may force price increases that weaken demand; stronger foreign currencies may raise component costs or cause derivative losses. Hedging can offset only part of these effects.
- ✓ **VERITROOPER (PASS):** The U.S. dollar-equivalent values of foreign currency-denominated assets, liabilities, forecasted revenues, or expenses increase or decrease.

622. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0134\_04`

- **Question:** What is the name of the specific cybersecurity framework used by JPMorganChase?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not name a specific cybersecurity framework used by JPMorganChase.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the name of a cybersecurity framework used by JPMorganChase.

623. PASS · Precision · `SEC\_10K\_multi\_0020\_01`

- **Question:** What is the amount of cash paid for income taxes, net, in 2023?
- **Ground truth:** \$18,679
- ✓ **Baseline (vanilla RAG) (PASS):** Cash paid for income taxes, net of refunds, was **\$9.9 billion (\$9.908 billion)** in 2023.
- ✓ **VERITROOPER (PASS):** Cash paid for income taxes, net, in 2023 was **\$18,679 million**.

624. PASS · Exception · `SEC\_10K\_multi\_0061\_00`

- **Question:** Under what circumstance does the general rule NOT apply to the adoption of new and recently issued accounting pronouncements?
- **Ground truth:** Refer to Note 1 of the Notes to the Consolidated Financial Statements in Part IV, Item 15 of this Annual Report on Form 10-K for a discussion.
- ✓ **Baseline (vanilla RAG) (PASS):** The general adoption timeline does not apply if the registrant is an emerging growth company that elects to use the extended transition period for new or revised accounting standards.
- ✓ **VERITROOPER (PASS):** The evidence does not state any exception to the general rule regarding adoption of new and recently issued accounting pronouncements.

625. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_00`

- **Question:** What is the total cost of the Company's cybersecurity insurance coverage?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The total cost of the Company's cybersecurity insurance coverage is not disclosed in the provided evidence.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total cost of the Company's cybersecurity insurance coverage.

626. FAIL · Cross-Reference · `SEC\_10K\_multi\_0024\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total value of U.S. Treasury securities and U.S. agency securities combined?
- **Ground truth:** $16,150 + 5,431 = 21,581$
- ✗ **Baseline (vanilla RAG) (FAIL):** As of September 27, 2025, their combined **fair value was \$20.968 billion** (\$15.848 billion + \$5.120 billion).
- ✓ **VERITROOPER (PASS):** The combined total value is **\$21,581** ($16,150 + 5,431$).

627. PASS · Exception · `SEC\_10K\_multi\_0111\_01`

- **Question:** What is excluded from the claims of a receiver of an insured depository institution for administrative expense?
- **Ground truth:** the claims of other unsecured creditors of the institution
- ✓ **Baseline (vanilla RAG) (PASS):** Deposits in non-U.S. branches are excluded, unless they are dually payable in both the U.S. and a non-U.S. branch. Public noteholders' claims are also not given priority.
- ✓ **VERITROOPER (PASS):** Depositors in non-U.S. branches and public noteholders are excluded from priority and treated as other unsecured creditors.

628. FAIL · Cross-Reference · `SEC\_10K\_multi\_1155\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total retained loans for 2025 and 2024, and what percentage of investment-grade loans is that for each year?
- **Ground truth:** The total retained loans for 2025 is \$165,545 and for 2024 is \$162,432. The percentage of investment-grade loans for 2025 is 71.81% and for 2024 is 70.36%.
- ✗ **Baseline (vanilla RAG) (FAIL):** - **2025:** Total retained loans were **\$792.367 billion**, of which **68.32%** were investment-grade.
- **2024:** Total retained loans were **\$690.396 billion**, of which **68.32%** were investment-grade.

- ✓ **VERITROOPER (PASS):** Total retained loans were \$792,367 in 2025 and \$690,396 in 2024; investment-grade loans represented 68.3% in both years.

629. PASS · Precision · `SEC\_10K\_multi\_1607\_00`

- **Question:** What is the year for the Consolidated Statements of Income?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The Consolidated Statements of Income cover the years ended December 31, 2025, 2024, and 2023.
- ✓ **VERITROOPER (PASS):** The Consolidated Statements of Income cover the years ended December 31, 2025, 2024, and 2023.

630. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1279\_04`

- **Question:** What is the minimum amount that must be advanced to investors before the Firm can stop payment?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a minimum amount. The Firm may stop payments to investors when collateral is insufficient to cover the servicer advance.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a minimum amount that must be advanced before the Firm can stop payment to investors.

631. PASS · Cause & Effect · `SEC\_10K\_multi\_1235\_00`

- **Question:** What happens as a result of the Firm not having the power to direct the activities of the VIE that most significantly impact the VIE's economic performance?
- **Ground truth:** The Firm generally does not consolidate the VIE.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm generally does not consolidate the VIE; instead, it records and reports its positions on the consolidated balance sheets like other third-party transactions.
- ✓ **VERITROOPER (PASS):** The Firm generally does not consolidate the VIE, but records and reports the positions on its Consolidated balance sheets like any other third-party transaction.

632. PASS · Cross-Reference · `SEC\_10K\_multi\_0580\_00`

- **Question:** What is the purpose of the net notional amount of protection purchased and sold through credit derivatives?
- **Ground truth:** Represents the net notional amount of protection purchased and sold through credit derivatives and credit-related notes used to manage both performing and nonperforming wholesale credit exposures
- ✓ **Baseline (vanilla RAG) (PASS):** The net notional amount of protection purchased and sold through credit derivatives is used to manage the firm's credit exposures, including loans, lending-related commitments, and derivative receivables.
- ✓ **VERITROOPER (PASS):** It is used to manage both performing and nonperforming wholesale credit exposures.

633. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0080\_00`

- **Question:** What is the total amount of damages sought in the Horanic v. Huang lawsuit?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The plaintiffs seek unspecified damages; no total amount is stated.
- ✓ **VERITROOPER (PASS):** The total amount is not specified; the plaintiffs seek unspecified damages and other relief.

634. PASS · Cross-Reference · `SEC\_10K\_multi\_1531\_01`

- **Question:** What percentage of the Company's worldwide unit case volume in 2025 was accounted for by Trademark Coca-Cola?
- **Ground truth:** 47%
- ✓ **Baseline (vanilla RAG) (PASS):** Trademark Coca-Cola accounted for **47%** of the Company's worldwide unit case volume in 2025.
- ✓ **VERITROOPER (PASS):** Trademark Coca-Cola accounted for 47% of the Company's worldwide unit case volume in 2025.

635. FAIL · Cross-Reference · `SEC\_10K\_multi\_1591\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the difference between the calculated fair value and the carrying value of the Company's investment in Monster Beverage Corporation as of December 31, 2025?
- **Ground truth:** The difference between the calculated fair value and the carrying value of the Company's investment in Monster Beverage Corporation was \$10,066 million as of December 31, 2025.
- ✗ **Baseline (vanilla RAG) (FAIL):** The difference was **\$10,066 million** (approximately **\$10.1 billion**).
- ✓ **VERITROOPER (PASS):** The difference was **\$10,066 million** as of December 31, 2025.

636. FAIL · Cross-Reference · `SEC\_10K\_multi\_1160\_01`

TABLE MISREAD

- **Question:** What is the total retained loans amount in the U.S. and the total loan delinquency amount in the 30-89 days past due category, respectively?
- **Ground truth:** \$165,545 and 918
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025:
 - **Total retained U.S. loans: \$626.060 billion**
 - **Loans 30-89 days past due: \$2.913 billion**
- ✗ **VERITROOPER (FAIL):** The total retained loans in the U.S. were **\$626,060**, and loans 30-89 days past due totaled **\$2,913**.

637. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0006\_04`

- **Question:** How many patent claims has the Company faced in total?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide a total number. It only states that the Company has faced "a significant number" of patent claims.
- ✓ **VERITROOPER (PASS):** The evidence does not specify the total number of patent claims.

638. PASS · Cross-Reference · `SEC\_10K\_multi\_0009\_01`

- **Question:** How many shareholders of record were there as of October 17, 2025?
- **Ground truth:** 22,429
- ✓ **Baseline (vanilla RAG) (PASS):** As of October 17, 2025, there were **22,429 shareholders of record**.
- ✓ **VERITROOPER (PASS):** There were 22,429 shareholders of record as of October 17, 2025.

639. PASS · Exception · `SEC\_10K\_multi\_1071\_01`

- **Question:** What is excluded from the general rule for the amortization of premiums on callable debt securities?
- **Ground truth:** premiums on certain callable debt securities are amortized to the earliest call date
- ✓ **Baseline (vanilla RAG) (PASS):** Premiums on certain callable debt securities are excluded; they are amortized to the earliest call date rather than over the security's contractual life.
- ✓ **VERITROOPER (PASS):** Premiums on certain callable debt securities are excluded; they are amortized to the earliest call date.

640. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0071\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the estimated fair value of non-marketable equity securities as of January 26, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The reported carrying value of non-marketable equity securities as of January 26, 2025 was **\$3.387 billion**.
- ✓ **VERITROOPER (PASS):** The estimated fair value of non-marketable equity securities as of January 26, 2025 is not provided.

641. FAIL · Cross-Reference · `SEC\_10K\_multi\_0285\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the total revenue for the year ended December 31, 2024, as reported in the selected income statement data, and how does it compare to the year ended December 31, 2025?
- **Ground truth:** The total revenue for the year ended December 31, 2024, is \$3,387 million, and for the year ended December 31, 2025, it is \$3,669 million.
- ✗ **Baseline (vanilla RAG) (FAIL):** Total net revenue for 2024 was **\$177.556 billion**, compared with **\$182.447 billion** in 2025-an increase of **\$4.891 billion, or 3%**, in 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide total revenue for 2024 or 2025, so the comparison cannot be determined.

642. PASS · Exception · `SEC\_10K\_multi\_1534\_00`

- **Question:** Under what circumstance does the general rule NOT apply regarding Proposition 65?
- **Ground truth:** unless the conditions of an exemption can be met
- ✓ **Baseline (vanilla RAG) (PASS):** The Proposition 65 warning requirement does not apply if the manufacturer demonstrates that exposure is below an established "safe harbor" threshold or qualifies for another exemption, such as being naturally occurring or resulting from necessary cooking.
- ✓ **VERITROOPER (PASS):** A Proposition 65 warning is not required if the manufacturer shows exposure is below a safe-harbor threshold, naturally occurring, caused by necessary cooking, or covered by another applicable exemption.

643. PASS · Cause & Effect · `SEC\_10K\_multi\_1433\_01`

- **Question:** What is the impact of the Firm updating its FTP with respect to consumer deposits?
- **Ground truth:** An increase in the funding benefit reflected within CCB's net interest income that is fully offset in Corporate, with no effect on the Firm's net interest income.
- ✓ **Baseline (vanilla RAG) (PASS):** Effective in Q4 2024, the FTP update increased the funding benefit reflected in CCB's net interest income. This increase was fully offset in Corporate, so it had no impact on the Firm's total net interest income.

- ✓ **VERITROOPER (PASS)**: The update increased the funding benefit reflected in CCB's net interest income, fully offset by Corporate, with no effect on the Firm's total net interest income.

644. PASS · Cause & Effect · `SEC\_10K\_multi\_1208\_00`

- **Question**: What happens as a consequence of the Firm being the primary beneficiary of the securitization trusts?

- **Ground truth**: The Firm consolidates the assets and liabilities of its sponsored credit card trusts.

- ✓ **Baseline (vanilla RAG) (PASS)**: As the primary beneficiary, the Firm consolidates the securitization trusts' assets and liabilities in its financial statements.

- ✓ **VERITROOPER (PASS)**: As the primary beneficiary, the Firm consolidates the assets and liabilities of its sponsored credit card securitization trusts.

645. FAIL · Cross-Reference · `SEC\_10K\_multi\_0084\_00`

VALIDATOR OVER-FIRED

- **Question**: What is the total depreciation and amortization expense for the Compute & Networking segment for fiscal years 2026 and 2025, and how does it compare to the Graphics segment for the same years?

- **Ground truth**: The total depreciation and amortization expense for the Compute & Networking segment for fiscal years 2026 and 2025 is \$1.6 billion and \$732 million, respectively. For the Graphics segment, the total depreciation and amortization expense for fiscal years 2026 and 2025 is \$590 million and \$372 million, respectively.

- ✗ **Baseline (vanilla RAG) (FAIL)**: - **Compute & Networking**: \$1.6 billion (2026) + \$732 million (2025) = **\$2.332 billion**

- **Graphics**: \$590 million (2026) + \$372 million (2025) = **\$962 million**

Compute & Networking's total was **\$1.370 billion higher**, about **2.4 times** the Graphics segment's total.

- ✗ **VERITROOPER (FAIL)**: Compute & Networking's total depreciation and amortization expense was \$2.332 billion for fiscal 2026 and 2025, compared with \$962 million for Graphics-\$1.370 billion higher.

646. PASS · Exception · `SEC\_10K\_multi\_1105\_01`

- **Question**: What type of loans are not subject to the Firm's allowance for loan losses and charge-off policies?

- **Ground truth**: loans held-for-sale and loans at fair value

- ✓ **Baseline (vanilla RAG) (PASS)**: Loans held-for-sale and loans measured at fair value are not subject to the Firm's allowance for loan losses and charge-off policies.

- ✓ **VERITROOPER (PASS)**: Loans held-for-sale and loans measured at fair value are not subject to the Firm's allowance for loan losses and charge-off policies.

647. FAIL · Cause & Effect · `SEC\_10K\_multi\_1567\_01`

VALIDATOR OVER-FIRED

- **Question**: What is the effect of lower other operating charges on the Company's Corporate segment?

- **Ground truth**: decrease in operating loss

- ✗ **Baseline (vanilla RAG) (FAIL)**: Lower other operating charges substantially reduced the Corporate segment's operating loss-from \$5.263 billion in 2024 to \$1.816 billion in 2025-primarily because the prior year included a large fairlife contingent-consideration remeasurement charge.

- ✓ **VERITROOPER (PASS)**: Lower other operating charges primarily reduced the Corporate segment's operating loss in 2025.

648. PASS · Precision · `SEC\_10K\_multi\_0985\_00`

- **Question:** What is the year for which the pre-tax gains/(losses) are recorded?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The pre-tax gains/(losses) were recorded for the years ended December 31, 2025, 2024, and 2023.
- ✓ **VERITROOPER (PASS):** The pre-tax gains/(losses) are recorded for the year ended December 31, 2025.

649. PASS · Precision · `SEC\_10K\_multi\_1167\_00`

- **Question:** What was the interest income on nonaccrual loans recognized on a cash basis for the year ended December 31, 2024?
- **Ground truth:** \$51 million
- ✓ **Baseline (vanilla RAG) (PASS):** The interest income recognized on a cash basis was **\$51 million** for the year ended December 31, 2024.
- ✓ **VERITROOPER (PASS):** The interest income was **\$51 million** for the year ended December 31, 2024.

650. PASS · Exception · `SEC\_10K\_multi\_1175\_01`

- **Question:** What circumstance does the financial effect of loan modifications NOT apply to?
- **Ground truth:** NM
- ✓ **Baseline (vanilla RAG) (PASS):** The financial effect is not applicable or meaningful for modifications categorized as "Other," including certain single or multiple loan modifications.
- ✓ **VERITROOPER (PASS):** The financial effect does not meaningfully apply to loans with both an interest rate reduction and a term extension.

651. FAIL · Precision · `SEC\_10K\_multi\_0932\_01`

BAD QUESTION

- **Question:** What is the total derivative receivables?
- **Ground truth:** 2,609
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, total derivative receivables were **\$606.382 billion** gross, or **\$57.777 billion** after netting adjustments.
- ✓ **VERITROOPER (PASS):** Total derivative receivables were **\$606,382 million**.

652. FAIL · Precision · `SEC\_10K\_multi\_0629\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the net charge-off/(recovery) rate for the year ended December 31, 2025?
- **Ground truth:** 0.22
- ✗ **Baseline (vanilla RAG) (FAIL):** The net charge-off/(recovery) rate for 2025 was **0.29%**.
- ✓ **VERITROOPER (PASS):** The net charge-off/(recovery) rate for 2025 was 0.22%.

653. PASS · Cross-Reference · `SEC\_10K\_multi\_0999\_01`

- **Question:** What was the net pre-tax loss reclassified to other revenue in the year ended December 31, 2023?
- **Ground truth:** -35 million
- ✓ **Baseline (vanilla RAG) (PASS):** The net pre-tax loss reclassified to other revenue in 2023 was **\$35 million**.
- ✓ **VERITROOPER (PASS):** The net pre-tax loss reclassified to other revenue in 2023 was **\$35 million**.

654. PASS · Conditional · `SEC\_10K\_multi\_1594\_01`

- **Question:** If the certified results equal the predefined performance criteria, what is the number of performance share units earned?
- **Ground truth:** The target award.
- ✓ **Baseline (vanilla RAG) (PASS):** The number of performance share units earned will equal the target award.
- ✓ **VERITROOPER (PASS):** The number of performance share units earned will equal the target award.

655. PASS · Precision · `SEC\_10K\_multi\_1341\_00`

- **Question:** What is the after-tax net unamortized unrealized loss for 2023?
- **Ground truth:** 895
- ✓ **Baseline (vanilla RAG) (PASS):** The after-tax net unamortized unrealized loss for 2023 was **\$895 million**.
- ✓ **VERITROOPER (PASS):** The after-tax net unamortized unrealized loss for 2023 was **\$895 million**.

656. PASS · Cross-Reference · `SEC\_10K\_multi\_1592\_01`

- **Question:** What is the amount of consolidated reinvested earnings that represents undistributed earnings of investments accounted for under the equity method as of December 31, 2025?
- **Ground truth:** 9,067 million.
- ✓ **Baseline (vanilla RAG) (PASS):** The amount was **\$9,067 million** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The amount was **\$9,067 million** as of December 31, 2025.

657. PASS · Precision · `SEC\_10K\_multi\_1476\_01`

- **Question:** What is the total assets?
- **Ground truth:** 4,431,624
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide a total assets figure.
- ✓ **VERITROOPER (PASS):** Total assets were \$4,431,624.

658. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0034\_06`

- **Question:** What is the total value of the notes issued by Apple Inc. in 2022?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify the principal amounts of Apple's 2022 notes, so their total value cannot be determined from the information provided.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total value of Apple Inc.'s notes issued in 2022.

659. PASS · Cause & Effect · `SEC\_10K\_multi\_1235\_01`

- **Question:** What happens due to significant modifications or events that result in a change in the nature of the investment or a change in the Firm's relationship with the underlying project?
- **Ground truth:** A reevaluation of eligible investments is required.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm must reevaluate the investment's eligibility for the proportional amortization method. During the period, no such significant modifications or events occurred.
- ✓ **VERITROOPER (PASS):** Such significant modifications or events require the Firm to reevaluate the investment's eligibility for the proportional amortization method.

660. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1149\_01`

- **Question:** How many credit card accounts were originated in 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify how many credit card accounts were originated in 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many credit card accounts were originated in 2025.

661. PASS · Conditional · `SEC\_10K\_multi\_0043\_01`

- **Question:** If we are unable to provide contractual warranty or other extended service obligations, will our competitive position be harmed?
- **Ground truth:** Yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. The evidence states that the company's competitive position and future results will be further harmed if it cannot provide contractual warranty or other extended service obligations.
- ✓ **VERITROOPER (PASS):** Yes. The evidence states that NVIDIA's competitive position and future results will be further harmed over the long term if it cannot provide contractual warranty or other extended service obligations.

662. FAIL · Precision · `SEC\_10K\_multi\_0537\_00`

UNSUPPORTED NUMBER

- **Question:** What is the amount of FHLB advances in 2024?
- **Ground truth:** 6,000
- ✗ **Baseline (vanilla RAG) (FAIL):** FHLB advances were **\$29.257 billion as of December 31, 2024**. The 2024 average balance was **\$35.040 billion**.
- ✓ **VERITROOPER (PASS):** FHLB advances issuance in 2024 was **\$6,000 million**.

663. PASS · Precision · `SEC\_10K\_multi\_0380\_01`

- **Question:** What is the return on equity percentage?
- **Ground truth:** 40
- ✓ **Baseline (vanilla RAG) (PASS):** The return on tangible common equity was **20% in 2025**.
- ✓ **VERITROOPER (PASS):** Return on equity was 40% in Col 1, 34% in Col 2, and 31% in Col 3.

664. PASS · Conditional · `SEC\_10K\_multi\_0242\_00`

- **Question:** If the provision for credit losses in 2025 was \$14.2 billion, and the net charge-offs were \$9.8 billion, what was the net addition to the allowance for credit losses?
- **Ground truth:** \$4.4 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The net addition to the allowance for credit losses was **\$4.4 billion** (\$14.2 billion – \$9.8 billion).
- ✓ **VERITROOPER (PASS):** The net addition to the allowance for credit losses was **\$4.4 billion**.

665. PASS · Cross-Reference · `SEC\_10K\_multi\_1415\_01`

- **Question:** How are deferred amounts treated for commitments not expected to fund?
- **Ground truth:** They are amortized to lending- and deposit-related fees on a straight line basis over the commitment period.

- ✓ **Baseline (vanilla RAG) (PASS):** Deferred amounts for commitments not expected to fund are amortized to lending- and deposit-related fees on a straight-line basis over the commitment period.
- ✓ **VERITROOPER (PASS):** They are amortized to lending- and deposit-related fees on a straight-line basis over the commitment period.

666. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_02`

- **Question:** What is the average cost of responding to an infringement claim?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide an average cost. It only states that responding to, defending, and resolving infringement claims can involve "significant costs."
- ✓ **VERITROOPER (PASS):** The evidence does not provide an average cost of responding to an infringement claim.

667. PASS · Cause & Effect · `SEC\_10K\_multi\_1433\_00`

- **Question:** What happens as a result of lower average interest rates in the current year?
- **Ground truth:** The cost of funding for assets and the funding benefit earned for liabilities generally decreased compared with the prior year.
- ✓ **Baseline (vanilla RAG) (PASS):** Lower average interest rates reduced both the cost of funding assets and the funding benefit earned on liabilities. This lowered funding costs for loans and Markets activities, while deposit margins compressed because the FTP benefit on deposits fell more than the rates paid to depositors.
- ✓ **VERITROOPER (PASS):** Lower average interest rates reduced funding costs for assets and the funding benefit for liabilities, resulting in lower funding costs for loans and Markets activities and overall deposit margin compression.

668. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0134\_01`

- **Question:** How much time does JPMorganChase typically spend on investigating a cyber attack?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a typical duration. It states that an investigation "could take a significant amount of time" to complete.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a typical investigation time; it only states that an investigation could take a significant amount of time.

669. PASS · Precision · `SEC\_10K\_multi\_0851\_00`

- **Question:** What is the total liabilities measured at fair value on a nonrecurring basis?
- **Ground truth:** 5
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, total liabilities measured at fair value on a nonrecurring basis were **\$5 million**.
- ✗ **VERITROOPER (FAIL):** At December 31, 2024, total liabilities measured at fair value on a nonrecurring basis were \$0 million.

670. PASS · Cause & Effect · `SEC\_10K\_multi\_1593\_00`

- **Question:** What happens as a result of the IRS' designation of the Company's matter for litigation?
- **Ground truth:** The Company is forced to either accept the IRS' newly imposed tax assessment and pay the full amount or litigate the matter in the federal courts.
- ✓ **Baseline (vanilla RAG) (PASS):** The litigation designation forecloses all alternative resolution methods and prevents settlement with the IRS. The Company must either accept and fully pay the IRS's assessment or litigate in federal court; consequently, it filed a Tax Court petition in December 2015.

- ✓ **VERITROOPER (PASS):** The designation prevents alternative resolution or settlement, forcing the Company either to pay the full assessed tax or litigate in federal court. Consequently, the Company filed a Tax Court petition in December 2015 challenging the adjustments.

671. PASS · Precision · `SEC\_10K\_multi\_0720\_01`

- **Question:** What year is the date of the financial statements?
- **Ground truth:** 2024
- ✓ **Baseline (vanilla RAG) (PASS):** The financial statements are dated **2025**.
- ✓ **VERITROOPER (PASS):** 2025.

672. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0159\_04`

- **Question:** What is the percentage of shares available for issuance under JPMorganChase's employee share-based incentive plans that are held by non-employee directors?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** Cannot be determined from the evidence. JPMorganChase had 77,131,051 shares available for future issuance, but it does not specify how many were allocated to or held by non-employee directors.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many shares are held by or available to non-employee directors, so the percentage cannot be determined.

673. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0052\_05`

- **Question:** What is the typical duration of a data breach investigation?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a typical duration. It states that a cyberattack investigation may take "a significant amount of time" to complete.
- ✓ **VERITROOPER (PASS):** The evidence does not state the typical duration of a data breach investigation.

674. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0618\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the dollar amount of investment-grade loans as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Investment-grade retained loans totaled **\$541.364 billion** (\$541,364 million) as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide enough information to determine the dollar amount of investment-grade loans as of December 31, 2025.

675. PASS · Calculation · `SEC\_10K\_multi\_1337\_01`

- **Question:** What is the net income per share for the year ended December 31, 2025?
- **Ground truth:** 20.05
- ✓ **Baseline (vanilla RAG) (PASS):** 20.02
- ✓ **VERITROOPER (PASS):** 20.05

676. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0727\_04`

- **Question:** What is the specific fee charged by the Firm for using valuation models?
- **Ground truth:** UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify any fee charged by the Firm for using valuation models.
- ✓ **VERITROOPER (PASS):** The evidence does not specify any fee charged by the Firm for using valuation models.

677. PASS · Cause & Effect · `SEC\_10K\_multi\_0018\_00`

- **Question:** What happens as a consequence of higher net sales of Pro models?
- **Ground truth:** iPhone net sales increased during 2025 compared to 2024
- ✓ **Baseline (vanilla RAG) (PASS):** Higher net sales of Pro models caused iPhone net sales to increase in 2025 compared with 2024.
- ✓ **VERITROOPER (PASS):** iPhone net sales increased in 2025 compared with 2024.

678. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1279\_02`

HALLUCINATION

- **Question:** What is the maximum amount the Firm can advance to investors?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The Firm can advance only up to the amount covered by the collateral; it may stop payments to investors if the collateral is insufficient.
- ✓ **VERITROOPER (PASS):** The evidence does not specify a maximum amount the Firm can advance to investors.

679. PASS · Precision · `SEC\_10K\_multi\_0069\_01`

- **Question:** What is the total amortization expense associated with intangible assets for fiscal year 2024?
- **Ground truth:** 614 million
- ✓ **Baseline (vanilla RAG) (PASS):** The total amortization expense associated with intangible assets for fiscal year 2024 was **\$614 million**.
- ✓ **VERITROOPER (PASS):** The amortization expense associated with intangible assets for fiscal year 2024 was **\$614 million**.

680. PASS · Conditional · `SEC\_10K\_multi\_0129\_00`

- **Question:** If JPMorganChase fails to meet the requirements of a resolution, what could happen?
- **Ground truth:** It could be required to enter into further resolutions, incur additional penalties or judgments, or accept material restrictions on, or changes in the management of, its businesses.
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase could be required to:
 - Enter into further resolutions
 - Incur additional penalties or judgments
 - Accept material restrictions on, or changes to, the management of its businesses
 - Face prosecution or civil litigation related to the underlying matters.
- ✓ **VERITROOPER (PASS):** It could face further resolutions, additional penalties or judgments, material business restrictions or management changes, and possible prosecution or civil litigation.

681. FAIL · Precision · `SEC\_10K\_multi\_1345\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total other comprehensive income/(loss)?
- **Ground truth:** 8,997
- ✗ **Baseline (vanilla RAG) (FAIL):** Apple's total other comprehensive income was **\$1.601 billion in 2025**.

- ✓ **VERITROOPER (PASS):** Total other comprehensive income/(loss) across Columns 1-9 is: \$8,997; \$(831); \$8,166; \$(1,631); \$(382); \$(2,013); \$8,567; \$(1,669); and \$6,898.

682. FAIL · Cross-Reference · `SEC\_10K\_multi\_1274\_01`

UNSUPPORTED NUMBER

- **Question:** What is the fair value at December 31, and how does it relate to the change in unrealized gains/(losses) included in income?
- **Ground truth:** \$9,167, which is greater than \$48
- ✗ **Baseline (vanilla RAG) (FAIL):** The question does not specify the financial instrument. For the Level 3 liabilities at December 31, 2023:
 - Deposits: fair value **\$1,833 million**; unrealized gain/loss included in income for instruments still held: **\$73 million**
 - Short-term borrowings: **\$1,758 million**; **\$14 million**
 - Trading liabilities: **\$37 million**; **\$0**
 - Accounts payable and other liabilities: **\$52 million**; **\$(4) million**
 - Long-term debt: **\$27,726 million**; **\$2,870 million**

The unrealized amount is the portion of the period's fair-value change attributable to instruments still held at year-end; it is not the total ending fair value.

- ✓ **VERITROOPER (PASS):** Fair value at December 31 was \$9,167, \$9,121, and \$8,522 across the three columns, respectively. The corresponding unrealized gains/(losses) included in income were \$48, \$762, and \$467.

683. PASS · Exception · `SEC\_10K\_multi\_1393\_01`

- **Question:** What does JPMorgan Chase & Co. not anticipate regarding available lines of credit for commitments?
- **Ground truth:** That all available lines of credit will be utilized at the same time
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorgan Chase does not anticipate that all available lines of credit for these commitments will be used at the same time.
- ✓ **VERITROOPER (PASS):** JPMorgan Chase does not anticipate that all available lines of credit for these commitments will be utilized at the same time.

684. PASS · Cross-Reference · `SEC\_10K\_multi\_0235\_00`

- **Question:** What is the expected net interest income excluding Markets for 2026, according to the outlook information provided on January 13, 2026?
- **Ground truth:** approximately \$95 billion
- ✓ **Baseline (vanilla RAG) (PASS):** Management expects 2026 net interest income excluding Markets to be approximately **\$95 billion**, market dependent.
- ✓ **VERITROOPER (PASS):** The expected net interest income excluding Markets for 2026 is approximately **\$95 billion**, market dependent.

685. FAIL · Cross-Reference · `SEC\_10K\_multi\_0019\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the maximum one-day loss in fair value of the Company's foreign currency derivative positions due to adverse movements in exchange rates, as of September 27, 2025, and how does it compare to the loss in 2024?
- **Ground truth:** \$590 million and \$538 million

- **✗ Baseline (vanilla RAG) (FAIL):** As of September 27, 2025, the maximum one-day loss was estimated at **\$590 million** with 95% confidence, compared with **\$538 million** as of September 28, 2024-an increase of **\$52 million**, or approximately **9.7%**.
- **✗ VERITROOPER (FAIL):** The maximum one-day loss was \$590 million as of September 27, 2025, compared with \$538 million in 2024-an increase of \$52 million.

686. PASS · Conditional · `SEC\_10K\_multi\_0757\_01`

- **Question:** If net realizable value is below cost, what is the carrying value of physical commodities?
- **Ground truth:** approximates fair value
- **✓ Baseline (vanilla RAG) (PASS):** When net realizable value is below cost, physical commodities are carried at net realizable value, which approximates fair value.
- **✓ VERITROOPER (PASS):** If net realizable value is below cost, the carrying value of physical commodities approximates fair value.

687. PASS · Cause & Effect · `SEC\_10K\_multi\_0424\_01`

- **Question:** What is the consequence of the resolution of certain tax audits?
- **Ground truth:** A \$774 million income tax benefit was recorded.
- **✓ Baseline (vanilla RAG) (PASS):** The resolution of certain tax audits resulted in tax benefits, including a \$774 million income tax benefit recorded in Corporate in the second quarter of 2025.
- **✓ VERITROOPER (PASS):** The resolution of certain tax audits contributed to a \$774 million income tax benefit recorded in the second quarter of 2025.

688. PASS · Precision · `SEC\_10K\_multi\_0301\_00`

- **Question:** What is the net charge-off/(recovery) for Home Lending?
- **Ground truth:** -106
- **✓ Baseline (vanilla RAG) (PASS):** Home Lending net charge-offs/(recoveries) were **\$(122) million in 2025**, indicating a net recovery.
- **✓ VERITROOPER (PASS):** Home Lending net charge-offs/(recoveries) were \$(122), \$(106), and \$(56) across Columns 1, 2, and 3, respectively.

689. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1554\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What was the net operating revenue increase for the EMEA operating segment?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The EMEA operating segment's net operating revenues increased **5% in 2025 versus 2024**.
- **✓ VERITROOPER (PASS):** The evidence does not provide the net operating revenue increase for the EMEA operating segment.

690. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1149\_02`

- **Question:** What is the total amount of gross charge-offs for the year ended December 31, 2023?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not provide total gross charge-offs for the year ended December 31, 2023. It reports \$1.787 billion for 2025 and \$1.022 billion for 2024.
- **✓ VERITROOPER (PASS):** The evidence does not provide gross charge-offs for the year ended December 31, 2023.

691. PASS · Precision · `SEC\_10K\_multi\_1605\_00`

- **Question:** What is the date of the auditor's report?
- **Ground truth:** February 20, 2026
- ✓ **Baseline (vanilla RAG) (PASS):** The auditor's report is dated **February 20, 2026**.
- ✓ **VERITROOPER (PASS):** February 20, 2026.

692. PASS · Cross-Reference · `SEC\_10K\_multi\_0502\_01`

- **Question:** How does the Firm's common share repurchase program compare to the amounts repurchased in prior periods?
- **Ground truth:** The amount of common shares that the Firm repurchases in any period may be substantially more or less than the amounts estimated or actually repurchased in prior periods.
- ✓ **Baseline (vanilla RAG) (PASS):** The amount repurchased in any period may be substantially higher or lower than amounts estimated or actually repurchased in prior periods. Management determines repurchases dynamically based on factors such as market conditions, regulatory considerations, capital position and requirements, capital generation, and other investment opportunities.
- ✓ **VERITROOPER (PASS):** The amount repurchased in any period may be substantially more or less than amounts estimated or actually repurchased in prior periods, reflecting management's dynamic decision-making process.

693. FAIL · Precision · `SEC\_10K\_multi\_1299\_01`

BAD QUESTION

- **Question:** What is the total senior debt at December 31, 2025?
- **Ground truth:** 226,071
- ✗ **Baseline (vanilla RAG) (FAIL):** Total senior debt at December 31, 2025 was **\$347.5 billion**:
- Parent company: **\$234.5 billion**
- Subsidiaries: **\$113.0 billion**
- ✗ **VERITROOPER (FAIL):** Total senior debt at December 31, 2025 was **\$234.532 billion** (\$226.071 billion fixed-rate plus \$8.461 billion variable-rate).

694. PASS · Cause & Effect · `SEC\_10K\_multi\_0003\_02`

- **Question:** What happens as a consequence of natural disasters?
- **Ground truth:** Make it difficult or impossible for the Company to manufacture and deliver products to its customers, create delays and inefficiencies in the Company's supply and manufacturing chain, result in slowdowns and outages to the Company's service offerings, increase the Company's costs, and negatively impact consumer spending and demand in affected areas.
- ✓ **Baseline (vanilla RAG) (PASS):** Natural disasters can disrupt operations and supply chains, delay manufacturing and deliveries, cause service outages, reduce product and raw-material availability, increase costs and prices, weaken consumer demand, and adversely affect financial results and credit quality.
- ✓ **VERITROOPER (PASS):** Natural disasters can disrupt manufacturing and deliveries, delay supply chains, cause service outages, increase costs, and reduce consumer spending and demand in affected areas.

695. FAIL · Precision · `SEC\_10K\_multi\_0202\_00`

UNIT/FORMAT MISMATCH

- **Question:** What is the income before income tax expense?
- **Ground truth:** 72,595

- ✖ **Baseline (vanilla RAG) (FAIL):** JPMorgan Chase's income before income tax expense was **\$72.595 billion in 2025** (2024: \$75.081 billion; 2023: \$61.612 billion).
- ✔ **VERITROOPER (PASS):** Income before income tax expense was 72,595 in Col 1, 75,081 in Col 2, and 61,612 in Col 3.

696. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1265\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of goodwill allocated to the Asset & Wealth Management segment in 2024 that was not allocated in 2023?
- **Ground truth:** UNANSWERABLE
- ✖ **Baseline (vanilla RAG) (FAIL):** Asset & Wealth Management goodwill decreased by **\$61 million**, from **\$8.582 billion in 2023** to **\$8.521 billion in 2024**.
- ✖ **VERITROOPER (FAIL):** There was no additional goodwill allocated in 2024; Asset & Wealth Management goodwill decreased by \$61 million, from \$8.582 billion in 2023 to \$8.521 billion in 2024.

697. PASS · Cross-Reference · `SEC\_10K\_multi\_0053\_01`

- **Question:** What could delay or prevent a change in control of NVIDIA?
- **Ground truth:** The anti-takeover provisions of the Delaware General Corporation Law and provisions in our certificate of incorporation and bylaws.
- ✔ **Baseline (vanilla RAG) (PASS):** A change in control of NVIDIA could be delayed or prevented by:
 - Anti-takeover provisions of Delaware law.
 - Provisions in NVIDIA's certificate of incorporation and bylaws, including:
 - The Board's ability to issue preferred stock, change the number of directors, and amend bylaws without prior shareholder approval.
 - Shareholders' inability to act by written consent.
 - Advance-notice requirements for director nominations and shareholder proposals.
 - Supermajority voting requirements for certain amendments.
 - NVIDIA's Xbox agreement with Microsoft, which gives Microsoft first and last rights of refusal if someone offers to acquire at least 30% of NVIDIA's outstanding common stock.
- ✔ **VERITROOPER (PASS):** Delaware anti-takeover law, provisions in NVIDIA's certificate of incorporation and bylaws, and Microsoft's first and last rights of refusal under the Xbox agreement could delay or prevent a change in control.

698. PASS · Conditional · `SEC\_10K\_multi\_1349\_00`

- **Question:** If the Firm adopted the Income Taxes: Improvement to Income Tax Disclosures accounting standard under the retrospective method, what method was used for the adoption?
- **Ground truth:** retrospective method
- ✔ **Baseline (vanilla RAG) (PASS):** The Firm used the **retrospective method**.
- ✔ **VERITROOPER (PASS):** The Firm used the retrospective method.

699. PASS · Cross-Reference · `SEC\_10K\_multi\_1354\_01`

- **Question:** What was the income tax expense associated with the First Republic acquisition, and how did it affect the Firm's effective tax rate?
- **Ground truth:** reflected in the estimated bargain purchase gain, reduction in the Firm's effective tax rate
- ✔ **Baseline (vanilla RAG) (PASS):** The evidence does not specify the amount. It states that the income tax expense associated with the First Republic acquisition was reflected in the estimated bargain purchase gain, reducing the Firm's effective tax rate.

- ✓ **VERITROOPER (PASS):** The amount of income tax expense associated with the First Republic acquisition is not provided; it was reflected in the estimated bargain purchase gain and reduced the Firm's effective tax rate.

700. FAIL · Precision · `SEC\_10K\_multi\_1451\_01`

VALIDATOR OVER-FIRED

- **Question:** What is the income from dividends from subsidiaries and affiliates in 2025?
- **Ground truth:** 50,000
- ✗ **Baseline (vanilla RAG) (FAIL):** Income from dividends from subsidiaries and affiliates in 2025 was **\$50.0 billion**.
- ✓ **VERITROOPER (PASS):** Income from dividends from subsidiaries and affiliates in 2025 was **\$37,000 million**.

701. PASS · Conditional · `SEC\_10K\_multi\_1545\_00`

- **Question:** If the Company's tax positions are ultimately sustained on appeal, what amount would be refunded in full or in part?
- **Ground truth:** The amount, plus interest earned, would be refunded in full or in part.
- ✓ **Baseline (vanilla RAG) (PASS):** The \$6.0 billion IRS Tax Litigation Deposit, plus interest earned, would be refunded in full or in part.
- ✓ **VERITROOPER (PASS):** The Company would receive a refund of up to \$6.0 billion, plus interest earned.

702. FAIL · Precision · `SEC\_10K\_multi\_1159\_01`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the percentage of criticized to total retained loans secured by real estate for 2025?
- **Ground truth:** 4.43
- ✗ **Baseline (vanilla RAG) (FAIL):** The percentage was **6.37%** in 2025.
- ✗ **VERITROOPER (FAIL):** The percentage was **6.37%** in 2025.

703. PASS · Conditional · `SEC\_10K\_multi\_1151\_00`

- **Question:** If the borrower does not make the contractual payments when due under the modified payment terms, what happens to the credit card loan?
- **Ground truth:** The credit card loan continues to age and will be charged-off in accordance with the Firm's standard charge-off policy.
- ✓ **Baseline (vanilla RAG) (PASS):** The credit card loan continues to age and is charged off according to the Firm's standard charge-off policy.
- ✓ **VERITROOPER (PASS):** The credit card loan continues to age and is charged off according to the Firm's standard charge-off policy.

704. PASS · Exception · `SEC\_10K\_multi\_0065\_01`

- **Question:** What type of financial instruments are NOT measured and disclosed at fair value based on the observability of inputs used in the determination of fair value?
- **Ground truth:** Financial instruments measured and disclosed at fair value are classified and disclosed based on the observability of inputs used in the determination of fair value as follows: Level 1: Observable inputs such as quoted prices in active markets.
- ✓ **Baseline (vanilla RAG) (PASS):** Investments measured at fair value using net asset value (NAV) per share-or its equivalent-as a practical expedient are not classified based on input observability. These include certain hedge funds, private equity funds, real estate funds, and other funds.

- ✓ **VERITROOPER (PASS):** Cash equivalents, accounts receivable, accounts payable, and accrued liabilities are not measured at fair value; their carrying values approximate fair value due to short maturities.

705. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1531\_01`

- **Question:** What is the average annual salary of a bottling partner in the United States?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide information on the average annual salary of a bottling partner in the United States.
- ✓ **VERITROOPER (PASS):** The evidence does not provide information about the average annual salary of a bottling partner in the United States.

706. PASS · Exception · `SEC\_10K\_multi\_1177\_01`

- **Question:** What type of loan modifications are excluded from the calculation?
- **Ground truth:** Other-than-insignificant payment deferral and term extension
- ✓ **Baseline (vanilla RAG) (PASS):** Insignificant payment delays are excluded; only **other-than-insignificant payment delays** qualify as financial difficulty modifications (FDMs).
- ✓ **VERITROOPER (PASS):** The evidence does not specify any loan modification types excluded from the calculation; it states that loans with multiple modifications are included.

707. PASS · Conditional · `SEC\_10K\_multi\_0041\_00`

- **Question:** If we announced the NVIDIA Blackwell GeForce RTX 50 Series family in fiscal year 2025, did we launch and scale the Blackwell architecture for gaming and GeForce NOW in fiscal year 2026?
- **Ground truth:** Yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. NVIDIA announced the Blackwell GeForce RTX 50 Series in fiscal 2025 and launched and scaled the Blackwell architecture for gaming and GeForce NOW in fiscal 2026.
- ✓ **VERITROOPER (PASS):** Yes. NVIDIA launched and scaled the Blackwell architecture for gaming and GeForce NOW in fiscal year 2026.

708. PASS · Cross-Reference · `SEC\_10K\_multi\_0318\_00`

- **Question:** What is the total number of open credit cards, and how does it relate to the commercial and investment bank's services?
- **Ground truth:** The total number of open credit cards is 68, which is part of the commercial and investment bank's services.
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, JPMorgan Chase had **116.5 million open credit cards**, including primary cardholders and authorized users.

This metric belongs to **Consumer & Community Banking's Card Services**, not the Commercial & Investment Bank (CIB). CIB separately provides corporate and institutional clients with payments services, including commercial cards, liquidity, clearing, trade, and working-capital solutions.

- ✓ **VERITROOPER (PASS):** The total number of open credit cards is 68, including primary cardholders and authorized users. The evidence does not specify how this count relates to the Commercial & Investment Bank's services.

709. PASS · Cause & Effect · `SEC\_10K\_multi\_0018\_01`

- **Question:** What happens as a consequence of a lower effective tax rate on foreign earnings?
- **Ground truth:** The Company's effective tax rate for 2025 was lower than the statutory federal income tax rate

- ✓ **Baseline (vanilla RAG) (PASS):** A lower effective tax rate on foreign earnings reduces the company's overall effective tax rate and provision for income taxes relative to the statutory federal tax rate.
- ✓ **VERITROOPER (PASS):** A lower effective tax rate on foreign earnings contributed to Apple's 2025 effective tax rate being below the 21% statutory federal income tax rate.

710. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1095\_01`

- **Question:** How many master netting agreements has JPMorgan Chase & Co. entered into with its counterparties as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose the number of master netting agreements JPMorgan Chase had entered into as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not specify how many master netting agreements JPMorgan Chase & Co. had entered into as of December 31, 2025.

711. PASS · Precision · `SEC\_10K\_multi\_0112\_01`

- **Question:** What is the number of employees JPMorganChase had as of December 31, 2025?
- **Ground truth:** 318,512
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase had **318,512 employees globally** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** JPMorganChase had 318,512 employees as of December 31, 2025.

712. FAIL · Precision · `SEC\_10K\_multi\_0852\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the total amount of level 3 assets measured at fair value on a nonrecurring basis as of December 31, 2025?
- **Ground truth:** 863 million
- ✗ **Baseline (vanilla RAG) (FAIL):** The total amount was **\$1.392 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The total amount was **\$863 million** as of December 31, 2025.

713. PASS · Cause & Effect · `SEC\_10K\_multi\_0083\_00`

- **Question:** What happens as a consequence of the determination of unrecognized deferred tax liabilities for temporary differences related to investments not being practicable?
- **Ground truth:** The amount of unrecognized deferred tax liabilities for temporary differences related to these investments is not provided.
- ✓ **Baseline (vanilla RAG) (PASS):** The company does not provide or disclose an amount for those unrecognized deferred tax liabilities because it cannot practicably determine it.
- ✓ **VERITROOPER (PASS):** NVIDIA does not provide the amount of those unrecognized deferred tax liabilities.

714. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1531\_03`

- **Question:** What is the total revenue generated from the sale of Coca-Cola products in the United States in 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not state 2025 U.S. revenue. It only reports that the United States represented **16% of worldwide unit case volume**, which is insufficient to calculate revenue.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total revenue generated from U.S. Coca-Cola product sales in 2025.

715. PASS · Exception · `SEC\_10K\_multi\_0051\_00`

- **Question:** What circumstance may prevent NVIDIA from selling its inventory of products subject to export restrictions?
- **Ground truth:** change in export control rules
- ✓ **Baseline (vanilla RAG) (PASS):** New or expanded export controls imposed after products are manufactured may restrict their sale, leaving NVIDIA with excess inventory and purchase obligations.
- ✓ **VERITROOPER (PASS):** A change in export control rules that subjects more products to export restrictions and licensing requirements may prevent NVIDIA from selling its inventory.

716. FAIL · Cross-Reference · `SEC\_10K\_multi\_1497\_00`

BAD QUESTION

- **Question:** What was the average interest rate earned on loans in U.S. operations for the year ended December 31, 2025, and what was the average balance of loans in U.S. operations for the same year?
- **Ground truth:** 6.86% and \$1,271,393
- ✗ **Baseline (vanilla RAG) (FAIL):** For 2025, the average interest rate earned on loans in U.S. operations was **6.86%**, and the average loan balance was **\$1,271,393 million** (approximately **\$1.27 trillion**).
- ✓ **VERITROOPER (PASS):** The average interest rate earned on U.S. loans was 6.86%, and the average loan balance was \$1,271,393 million for 2025.

717. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0646\_02`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the allowance for loan losses to retained nonaccrual loans for the highest credit ratio?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The highest allowance-for-loan-losses-to-retained-nonaccrual-loans ratio was **437% in 2023**.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the allowance for loan losses to retained nonaccrual loans.

718. FAIL · Cross-Reference · `SEC\_10K\_multi\_0910\_00`

BAD QUESTION

- **Question:** What is the total amount for Asset Managers and Consumer & Retail combined?
- **Ground truth:** 302,693
- ✗ **Baseline (vanilla RAG) (FAIL):** Asset Managers and Consumer & Retail combined total **\$286.793 billion** (\$152.848 billion + \$133.945 billion).
- ✗ **VERITROOPER (FAIL):** The combined total for Asset Managers and Consumer & Retail is **286,793**.

719. PASS · Precision · `SEC\_10K\_multi\_1241\_00`

- **Question:** What is the amount of income tax credits recognized in 2023?
- **Ground truth:** zero
- ✓ **Baseline (vanilla RAG) (PASS):** The amount of income tax credits recognized in 2023 was **zero**.
- ✓ **VERITROOPER (PASS):** The Firm recognized zero income tax credits in 2023.

720. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0728\_01`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** How much is the funding adjustment (FVA) for a typical credit derivative?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** No fixed "typical" FVA is provided. FVA depends on expected funding requirements, collateral arrangements, market funding costs, and the firm's credit risk. Aggregate derivatives FVA was a **\$18 million loss in 2025**, versus gains of **\$99 million in 2024** and **\$114 million in 2023**.
- **✓ VERITROOPER (PASS):** The evidence does not specify the amount of the funding adjustment (FVA) for a typical credit derivative.

721. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0728\_02`

- **Question:** What is the minimum number of comparable public companies required for valuation?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify a minimum number of comparable public companies required for valuation.
- **✓ VERITROOPER (PASS):** The evidence does not specify a minimum number of comparable public companies required for valuation.

722. FAIL · Cross-Reference · `SEC\_10K\_multi\_1584\_01`

UNIT/FORMAT MISMATCH

- **Question:** What is the total amount of equity attributable to shareowners of The Coca-Cola Company at the end of 2025, considering the equity attributable to shareowners at the beginning of the year and the net income attributable to shareowners of The Coca-Cola Company?
- **Ground truth:** 32,169
- **✗ Baseline (vanilla RAG) (FAIL):** Equity attributable to Coca-Cola shareowners was **\$32.169 billion at the end of 2025**. It began at **\$24.856 billion**, and **\$13.107 billion** of net income was added, with dividends and other equity adjustments reducing the final balance.
- **✗ VERITROOPER (FAIL):** At the end of 2025, total equity attributable to shareowners of The Coca-Cola Company was **\$32.169 billion**.

723. PASS · Cross-Reference · `SEC\_10K\_multi\_0863\_01`

- **Question:** What is the difference between the estimated fair value and carrying value of a loan generally attributable to?
- **Ground truth:** Changes in market interest rates, including credit spreads, market liquidity premiums and other factors.
- **✓ Baseline (vanilla RAG) (PASS):** The difference is generally attributable to changes in market interest rates-including credit spreads and market liquidity premiums-and other factors that affect a loan's fair value but not its carrying value.
- **✓ VERITROOPER (PASS):** It is generally attributable to changes in market interest rates, including credit spreads, market liquidity premiums, and other factors that affect fair value but not carrying value.

724. PASS · Exception · `SEC\_10K\_multi\_0576\_01`

- **Question:** What is the condition under which credit card loans remain on accrual status?
- **Ground truth:** until charged off
- **✓ Baseline (vanilla RAG) (PASS):** Credit card loans typically remain on accrual status until they are charged off.

- ✓ **VERITROOPER (PASS)**: Credit card loans typically remain on accrual status until they are charged off.

725. PASS · Exception · `SEC\_10K\_multi\_1434\_01`

- **Question**: What is excluded from the year-end figures?
- **Ground truth**: ratios
- ✓ **Baseline (vanilla RAG) (PASS)**: Mortgage loans past due and insured by U.S. government agencies are excluded from the year-end delinquency and nonaccrual figures because of the government guarantee.
- ✓ **VERITROOPER (PASS)**: Ratios are excluded from the figures stated in millions.

726. PASS · Calculation · `SEC\_10K\_multi\_0275\_00`

- **Question**: What is TCE for December 31, 2025?
- **Ground truth**: 289,002
- ✓ **Baseline (vanilla RAG) (PASS)**: 290018
- ✗ **VERITROOPER (FAIL)**: Insufficient information

727. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0134\_06`

- **Question**: How many days does it typically take for JPMorganChase to recover from a cybersecurity breach?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not specify a typical number of days for JPMorganChase to recover from a cybersecurity breach. It states that recovery and investigation may take a significant amount of time and could involve extended delays.
- ✓ **VERITROOPER (PASS)**: The evidence does not specify how many days recovery from a cybersecurity breach typically takes.

728. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0075\_02`

- **Question**: What is the amount of deferred revenue as of January 27, 2025?
- **Ground truth**: UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS)**: The deferred revenue balance was **\$1.813 billion as of January 26, 2025**. The evidence does not provide an amount for January 27, 2025.
- ✓ **VERITROOPER (PASS)**: The evidence does not provide deferred revenue as of January 27, 2025; it reports \$976 million as of January 26, 2025.

729. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1614\_04`

HALLUCINATION

- **Question**: What is the name of the company's auditor?
- **Ground truth**: UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL)**: Ernst & Young LLP.
- ✓ **VERITROOPER (PASS)**: The evidence does not provide the name of The Coca-Cola Company's auditor.

730. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_03`

- **Question**: What is the total value of Apple's trade receivables that are subject to credit and collectibility risk?
- **Ground truth**: UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide Apple's total trade receivables balance, so the value subject to credit and collectibility risk cannot be determined. It only states that cellular network carriers accounted for 34% of trade receivables and one customer accounted for 12% as of September 27, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total value of Apple's trade receivables subject to credit and collectibility risk.

731. PASS · Cause & Effect · `SEC\_10K\_multi\_1570\_00`

- **Question:** What happens as a result of the Company's self-insurance reserves?
- **Ground truth:** The Company's self-insurance reserves totaled \$155 million.
- ✓ **Baseline (vanilla RAG) (PASS):** The Company's self-insurance reserves cover estimated losses from self-insured claims. They totaled \$155 million as of December 31, 2025, and estimated payments related to them were excluded from the obligations table.
- ✓ **VERITROOPER (PASS):** As a result, estimated payments related to the Company's self-insurance reserves were not included in the contractual obligations table.

732. PASS · Precision · `SEC\_10K\_multi\_0441\_00`

- **Question:** What is the number of the page where Capital Risk Management is discussed?
- **Ground truth:** 89
- ✓ **Baseline (vanilla RAG) (PASS):** Capital Risk Management is discussed on pages **89-99**.
- ✓ **VERITROOPER (PASS):** Capital Risk Management is discussed on page 89.

733. PASS · Cause & Effect · `SEC\_10K\_multi\_1569\_01`

- **Question:** What happens if the credit ratings of certain bottlers in which the Company has equity method investments were to decline?
- **Ground truth:** The Company's equity income could be reduced as a result of the potential increase in interest expense for those bottlers.
- ✓ **Baseline (vanilla RAG) (PASS):** The bottlers' interest expense could increase, reducing the Company's equity income.
- ✓ **VERITROOPER (PASS):** The Company's equity income could be reduced because those bottlers could face higher interest expense.

734. FAIL · Cross-Reference · `SEC\_10K\_multi\_1549\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the estimated increase in net periodic pension cost if the discount rate decreases by 50 basis points, and how does this compare to the estimated increase if the expected long-term rate of return on plan assets decreases by 50 basis points?
- **Ground truth:** \$9 million and \$19 million, respectively
- ✗ **Baseline (vanilla RAG) (FAIL):** A 50-basis-point decrease in the discount rate would increase 2026 net periodic pension cost by an estimated **\$9 million**.

A 50-basis-point decrease in the expected long-term return on plan assets would increase it by **\$19 million-\$10 million more**, or about **2.1 times** the discount-rate impact.

- ✗ **VERITROOPER (FAIL):** A 50-basis-point decrease in the discount rate would increase 2026 net periodic pension cost by \$9 million, compared with a \$19 million increase from the same decrease in the expected long-term return on plan assets-\$10 million more.

735. PASS · Cause & Effect · `SEC\_10K\_multi\_0155\_01`

- **Question:** What happens as a result of PSUs vesting on March 25, 2026?
- **Ground truth:** 50% of the net issued shares are received
- ✓ **Baseline (vanilla RAG) (PASS):** Under the listed Rule 10b5-1 trading arrangements, 50% of the net shares issued from PSUs vesting on March 25, 2026 will be sold.
- ✓ **VERITROOPER (PASS):** For certain executives, 50% of the net issued shares received as a result of the PSU vesting are subject to the stated restriction through June 30, 2026.

736. FAIL · Precision · `SEC\_10K\_multi\_0253\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What is the percentage change in total liabilities from 2024 to 2025?
- **Ground truth:** 11%
- ✗ **Baseline (vanilla RAG) (FAIL):** Total liabilities increased by approximately **8.4%** from 2024 to 2025:

$$\frac{3,834,359 - 3,537,567}{3,537,567} \times 100 \approx 8.4\%$$

- ✓ **VERITROOPER (PASS):** Total liabilities increased by 11% from 2024 to 2025.

737. PASS · Precision · `SEC\_10K\_multi\_0607\_00`

- **Question:** What was the Real Estate exposure as of December 31, 2025?
- **Ground truth:** 224.9 billion
- ✓ **Baseline (vanilla RAG) (PASS):** Real Estate exposure was **\$224.9 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The Real Estate exposure was \$224.9 billion as of December 31, 2025.

738. PASS · Cross-Reference · `SEC\_10K\_multi\_0936\_00`

- **Question:** What is the net amount of derivative receivables as of December 31, 2025, for interest rate contracts that are over-the-counter (OTC) and have a legally enforceable master netting agreement?
- **Ground truth:** 24,193 million
- ✓ **Baseline (vanilla RAG) (PASS):** The net derivative receivables were **\$24.193 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The net derivative receivables for OTC interest rate contracts were **\$24.193 billion** as of December 31, 2025.

739. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1160\_00`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total number of loans that are 90 or more days past due and still accruing?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, loans 90 or more days past due and still accruing totaled **\$2.732 billion** in the retained credit card portfolio.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total number of loans that are 90 or more days past due and still accruing.

740. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0678\_03`

- **Question:** What is the percentage of deposits that are subject to the Firm's sensitivity analyses?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The percentage is not disclosed in the provided evidence. The Firm analyzes deposit betas and deposit-balance forecasts, particularly for consumer deposits, but gives no percentage of deposits subject to these analyses.
- ✓ **VERITROOPER (PASS):** The evidence does not state the percentage of deposits subject to the Firm's sensitivity analyses.

741. PASS · Precision · `SEC\_10K\_multi\_0993\_00`

- **Question:** What is the maximum length of time over which the derivative results recorded in AOCI will be recognized in earnings for cash flow hedges that have been terminated?
- **Ground truth:** ten years
- ✓ **Baseline (vanilla RAG) (PASS):** Approximately ten years.
- ✓ **VERITROOPER (PASS):** Approximately ten years.

742. PASS · Precision · `SEC\_10K\_multi\_1323\_00`

- **Question:** What is the SOFR rate for Series II?
- **Ground truth:** 2.745
- ✓ **Baseline (vanilla RAG) (PASS):** Series II's floating dividend rate is **three-month term SOFR plus 2.745%**.
- ✓ **VERITROOPER (PASS):** The SOFR rate for Series II is **SOFR + 2.745%**.

743. PASS · Cross-Reference · `SEC\_10K\_multi\_0864\_00`

- **Question:** What is the total estimated fair value of wholesale lending-related commitments at December 31, 2024, according to the fair value hierarchy?
- **Ground truth:** 4.4
- ✓ **Baseline (vanilla RAG) (PASS):** The total estimated fair value was **\$4.4 billion** at December 31, 2024.
- ✓ **VERITROOPER (PASS):** The total estimated fair value was **\$4.4 billion** at December 31, 2024.

744. PASS · Precision · `SEC\_10K\_multi\_1215\_01`

- **Question:** What is the value of subordinated securities at December 31, 2025?
- **Ground truth:** 56 million
- ✓ **Baseline (vanilla RAG) (PASS):** The subordinated securities were valued at **\$56 million** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The value of subordinated securities at December 31, 2025 was **\$56 million**.

745. PASS · Conditional · `SEC\_10K\_multi\_0728\_00`

- **Question:** If the valuation is based on a transaction price, what level is it classified under?
- **Ground truth:** Level 2
- ✓ **Baseline (vanilla RAG) (PASS):** It is classified as **Level 3**, because transaction-price-based valuation involves significant unobservable inputs.
- ✓ **VERITROOPER (PASS):** A valuation based on transaction prices for private equity direct investments is classified as Level 2 or Level 3.

746. PASS · Conditional · `SEC\_10K\_multi\_0041\_01`

- **Question:** If we provide a complete, end-to-end accelerated computing platform for AI, does it address both training and inferencing?
- **Ground truth:** Yes
- ✓ **Baseline (vanilla RAG) (PASS):** Yes. The complete, end-to-end accelerated computing platform for AI addresses both training and inferencing.
- ✓ **VERITROOPER (PASS):** Yes. NVIDIA's complete, end-to-end accelerated computing platform for AI addresses both training and inferencing.

747. PASS · Calculation · `SEC\_10K\_multi\_1589\_00`

- **Question:** What is the total fair value of the Company's equity securities as of December 31, 2025?
- **Ground truth:** 2589
- ✓ **Baseline (vanilla RAG) (PASS):** 2589
- ✗ **VERITROOPER (FAIL):** 2633

748. PASS · Cause & Effect · `SEC\_10K\_multi\_1356\_00`

- **Question:** What happens as a result of the Firm recognizing U.S. income tax expense on global intangible low tax income?
- **Ground truth:** The Firm recognizes the tax expense in the period in which the tax is incurred.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm recognizes U.S. income tax expense on global intangible low-tax income in the period in which the tax is incurred.
- ✓ **VERITROOPER (PASS):** The amount is recognized as income tax expense in the period in which the tax is incurred.

749. PASS · Cross-Reference · `SEC\_10K\_multi\_0681\_00`

- **Question:** What is the primary driver of the Firm's sensitivities as of December 31, 2025, compared to December 31, 2024?
- **Ground truth:** The net impact of Treasury and CIO actual and forecasted actions, including an increase in cash flow hedges of floating rate loans and in investment securities.
- ✓ **Baseline (vanilla RAG) (PASS):** The primary driver was the net impact of Treasury and CIO's actual and forecasted actions, including increased cash flow hedges of floating-rate loans and increased investment securities, both of which add duration.
- ✓ **VERITROOPER (PASS):** The primary driver was the net impact of Treasury and CIO actual and forecasted actions, including increased cash flow hedges of floating-rate loans and investment securities, both of which add duration.

750. PASS · Cross-Reference · `SEC\_10K\_multi\_1022\_01`

- **Question:** What is the basis for the maturity profile of credit derivatives?
- **Ground truth:** The maturity profile is based on the remaining contractual maturity of the credit derivative contracts.
- ✓ **Baseline (vanilla RAG) (PASS):** The maturity profile is based on the credit derivative contracts' remaining contractual maturity.
- ✓ **VERITROOPER (PASS):** The maturity profile is based on the remaining contractual maturity of the credit derivative contracts.

751. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_05`

- **Question:** What is the number of patents filed by NVIDIA in 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not state how many patents NVIDIA filed in 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not state the number of patents NVIDIA filed in 2025.

752. PASS · Precision · `SEC\_10K\_multi\_1436\_01`

- **Question:** What is the return on equity?
- **Ground truth:** 18%
- ✓ **Baseline (vanilla RAG) (PASS):** Return on equity (ROE) is net income divided by shareholders' equity. The evidence defines the term but does not provide a calculated ROE value.
- ✓ **VERITROOPER (PASS):** The evidence lists return on equity values of 32%, 32%, 38%, 18%, 18%, 14%, 40%, 34%, and 31%, depending on the column; no specific period or segment was provided.

753. PASS · Cause & Effect · `SEC\_10K\_multi\_1095\_00`

- **Question:** What happens as a result of the Firm's credit risk mitigation practices?
- **Ground truth:** The Firm does not hold any allowance for credit losses with respect to resale and securities borrowed arrangements.
- ✓ **Baseline (vanilla RAG) (PASS):** As a result of its credit risk mitigation practices, the Firm held no allowance for credit losses on resale and securities borrowed arrangements as of December 31, 2025 and 2024.
- ✓ **VERITROOPER (PASS):** As a result of its credit risk mitigation practices, the Firm held no allowance for credit losses on resale and securities borrowed arrangements as of December 31, 2025 and 2024.

754. PASS · Cross-Reference · `SEC\_10K\_multi\_1419\_00`

- **Question:** What is the minimum collateral required for securities lending, and how does it relate to the value of the securities on loan?
- **Ground truth:** 100% of the value of the securities on loan
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not specify a minimum collateral amount or percentage for securities lending, nor how it compares with the value of securities on loan. It only states that collateral and margin levels are monitored and additional collateral may be required if collateral values decline.
- ✓ **VERITROOPER (PASS):** The Firm requires collateral with a market value exceeding 100% of the securities on loan-i.e., collateral must be worth more than the lent securities.

755. PASS · Cause & Effect · `SEC\_10K\_multi\_1095\_01`

- **Question:** What is the consequence of the Firm obtaining an appropriate legal opinion with respect to a master netting agreement?
- **Ground truth:** The Firm nets, on the Consolidated balance sheets, the balances outstanding under its securities financing agreements with the same counterparty.
- ✓ **Baseline (vanilla RAG) (PASS):** Obtaining an appropriate legal opinion allows the Firm, when other U.S. GAAP netting criteria are met, to net eligible derivative and securities-financing balances with the same counterparty on its Consolidated balance sheets, including related eligible cash collateral. Without such an opinion, the balances are presented gross.

- **✓ VERITROOPER (PASS):** If U.S. GAAP netting criteria are also met, the Firm nets outstanding securities financing balances with the same counterparty on its Consolidated balance sheets. The opinion also allows related collateral and nonqualifying balances to be presented as amounts not nettable, reducing reported net amounts.

756. FAIL · Precision · `SEC\_10K\_multi\_1073\_00`

UNSUPPORTED NUMBER

- **Question:** What is the total held-to-maturity securities value?
- **Ground truth:** 270,134
- **✗ Baseline (vanilla RAG) (FAIL):** The total held-to-maturity securities value is **\$270.199 billion** in amortized cost (**\$253.344 billion** fair value).
- **✓ VERITROOPER (PASS):** The total held-to-maturity securities value is **\$270,134**.

757. PASS · Cross-Reference · `SEC\_10K\_multi\_0285\_01`

- **Question:** What is the contribution of Visa shares to the JPMorgan Chase Foundation in 2024, and what is the reference for further details?
- **Ground truth:** The contribution is \$1.0 billion, and for additional information, refer to Note 6.
- **✓ Baseline (vanilla RAG) (PASS):** The 2024 contribution of Visa shares to the JPMorgan Chase Foundation was **\$1.0 billion**. For further details, refer to **Note 6**.
- **✓ VERITROOPER (PASS):** The contribution was \$1.0 billion in 2024. Further details are provided in Note 6.

758. PASS · Precision · `SEC\_10K\_multi\_1539\_01`

- **Question:** What is the minimum tax rate under Pillar Two?
- **Ground truth:** 15%
- **✓ Baseline (vanilla RAG) (PASS):** Pillar Two establishes a global per-country minimum tax rate of **15%**.
- **✓ VERITROOPER (PASS):** The minimum tax rate under Pillar Two is 15%.

759. PASS · Cause & Effect · `SEC\_10K\_multi\_0155\_00`

- **Question:** What happens as a result of PSUs vesting on March 25, 2026?
- **Ground truth:** 50% of the net issued shares are received
- **✓ Baseline (vanilla RAG) (PASS):** Under the listed Rule 10b5-1 trading arrangements, 50% of the net shares issued from PSUs vesting on March 25, 2026 will be sold.
- **✓ VERITROOPER (PASS):** Certain executives receive net issued shares, 50% of which are subject to the stated restriction period through June 30, 2026.

760. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0727\_00`

- **Question:** What is the average time taken by the Valuation Control Group to verify fair value estimates?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify the average time taken by the Valuation Control Group to verify fair value estimates.
- **✓ VERITROOPER (PASS):** The evidence does not state the average time taken by the Valuation Control Group to verify fair value estimates.

761. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1554\_03`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What was the total net operating revenue increase for the Bottling Investments operating segment?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Bottling Investments' net operating revenues **decreased 8%** in 2025 versus 2024.
- **✓ VERITROOPER (PASS):** The evidence does not provide the total net operating revenue increase for the Bottling Investments operating segment.

762. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1169\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the total amount of payment deferrals provided in 2023?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** The total amount of payment deferrals provided in 2023 was **\$690 million**:
 - Secured by real estate: \$8 million
 - Commercial and industrial: \$437 million
 - Other loans: \$245 million
- **✗ VERITROOPER (FAIL):** The total amount of payment deferrals provided in 2023 was \$8 million.

763. PASS · Exception · `SEC\_10K\_multi\_0061\_01`

- **Question:** What is excluded from the disclosure controls and procedures?
- **Ground truth:** all error and all fraud
- **✓ Baseline (vanilla RAG) (PASS):** No specific information is identified as excluded. The controls cover information required to be disclosed in reports filed or submitted under the Exchange Act.
- **✓ VERITROOPER (PASS):** Disclosure controls and procedures do not provide absolute assurance or prevent all errors and fraud.

764. PASS · Cause & Effect · `SEC\_10K\_multi\_1599\_00`

- **Question:** What is the consequence of the Company's investment in limited partnerships that receive tax credits and other tax benefits?
- **Ground truth:** The Company received tax credits and other income tax benefits of \$241 million and recognized amortization expense of \$224 million related to these investments.
- **✓ Baseline (vanilla RAG) (PASS):** The investments generate returns through tax credits, other tax benefits, and cash distributions, thereby reducing the Company's income-tax payments. Under the proportional amortization method, their cost is amortized within income-tax expense in proportion to the tax benefits received. In 2025, the Company received \$241 million in tax credits and other benefits and recognized \$224 million of amortization expense.
- **✓ VERITROOPER (PASS):** In 2025, the investment generated \$241 million in tax credits and other income tax benefits, while the Company recognized \$224 million in related amortization expense.

765. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_07`

- **Question:** What is the total number of NVIDIA's research and development facilities worldwide?
- **Ground truth:** UNANSWERABLE

- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not state the total number of NVIDIA research and development facilities worldwide.
- ✓ **VERITROOPER (PASS):** The evidence does not state the total number of NVIDIA's research and development facilities worldwide.

766. PASS · Precision · `SEC\_10K\_multi\_0937\_01`

- **Question:** What is the number of OTC-cleared credit contracts?
- **Ground truth:** 3,233
- ✓ **Baseline (vanilla RAG) (PASS):** OTC-cleared credit contracts were **\$3,233 million** in gross derivative receivables at December 31, 2025.
- ✓ **VERITROOPER (PASS):** The number of OTC-cleared credit contracts is 3,233.

767. PASS · Conditional · `SEC\_10K\_multi\_0714\_00`

- **Question:** If the Firm adopts the new guidance on January 1, 2027, does it require derivative accounting treatment for contracts with both in scope and out of scope terms?
- **Ground truth:** No
- ✓ **Baseline (vanilla RAG) (PASS):** The guidance clarifies when derivative accounting applies to contracts containing both in-scope and out-of-scope terms; the evidence does not state that such contracts automatically require derivative accounting treatment.
- ✓ **VERITROOPER (PASS):** The evidence does not specify whether such contracts require derivative accounting treatment; it only states that the guidance clarifies its applicability.

768. PASS · Cross-Reference · `SEC\_10K\_multi\_0580\_01`

- **Question:** What criteria does the Firm use to determine the internal risk rating for each credit facility?
- **Ground truth:** The Firm generally considers internal ratings with qualitative characteristics equivalent to BBB-/Baa3 or higher as investment grade, and takes into consideration collateral and structural support
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm determines each credit facility's internal risk rating by considering:
 - Probability of default (PD)
 - Loss given default (LGD), including collateral and structural support
 - The obligor's debt capacity and financial flexibility
 - Earnings level
 - Amount and sources of repayment
 - Level and nature of contingencies
 - Management strength
 - The obligor's industry and geography

Ratings are regularly reviewed and updated based on information affecting the obligor's ability to meet its obligations.

- ✓ **VERITROOPER (PASS):** The Firm considers collateral and structural support when determining each credit facility's internal risk rating.

769. PASS · Conditional · `SEC\_10K\_multi\_0071\_00`

- **Question:** If the net unrealized gains on investments in publicly-held equity securities held at period end were not significant for fiscal years 2025 and 2024, what was the net unrealized gain for fiscal year 2026?
- **Ground truth:** \$6.6 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The net unrealized gain for fiscal year 2026 was **\$6.6 billion**.
- ✓ **VERITROOPER (PASS):** The net unrealized gain for fiscal year 2026 was \$6.6 billion.

770. PASS · Precision · `SEC\_10K\_multi\_0202\_01`

- **Question:** What is the diluted earnings per share?
- **Ground truth:** 20.02
- ✓ **Baseline (vanilla RAG) (PASS):** For 2025, diluted earnings per share was:
- **JPMorgan Chase:** \$20.02
- **Apple:** \$7.46
- ✓ **VERITROOPER (PASS):** The diluted earnings per share is **\$20.02**.

771. FAIL · Conditional · `SEC\_10K\_multi\_0570\_00`

UNSUPPORTED NUMBER

- **Question:** If the loan maturities are categorized by class, and the residential real estate loan class is considered, what is the estimated principal repayment amount for this class?
- **Ground truth:** calculated by weighted-average loan balance and interest rates for loan pools based on remaining loan term
- ✗ **Baseline (vanilla RAG) (FAIL):** The estimated principal repayment amount for residential real estate loans was **\$333.148 billion** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the estimated principal repayment amount for the residential real estate loan class.

772. PASS · Conditional · `SEC\_10K\_multi\_0757\_00`

- **Question:** If fair value hedging has been applied, what is the carrying value of physical commodities?
- **Ground truth:** approximates fair value
- ✓ **Baseline (vanilla RAG) (PASS):** When fair value hedging is applied, physical commodities are carried at fair value.
- ✓ **VERITROOPER (PASS):** The carrying value of physical commodities approximates fair value.

773. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1095\_03`

- **Question:** What is the total amount of cash collateral exchanged in securities financing agreements as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not separately disclose total **cash** collateral exchanged. It reports combined **securities and/or cash collateral**, so the cash-only amount cannot be determined.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the total amount of cash collateral exchanged in securities financing agreements as of December 31, 2025.

774. PASS · Cause & Effect · `SEC\_10K\_multi\_0137\_00`

- **Question:** What could result from a failure to effectively manage potential conflicts of interest or to satisfy fiduciary obligations?
- **Ground truth:** litigation and enforcement actions and reputational harm
- ✓ **Baseline (vanilla RAG) (PASS):** It could result in customer dissatisfaction, lost revenue, higher operating costs, litigation, penalties, heightened regulatory scrutiny and enforcement actions, and reputational harm.
- ✓ **VERITROOPER (PASS):** It could result in customer dissatisfaction, litigation, penalties, heightened regulatory scrutiny and enforcement actions, leading to lost revenue, higher operating costs, and reputational harm.

775. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1035\_03`

- **Question:** What is the average fee collected per client for administration fees in 2025?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** Cannot be determined from the provided evidence. Administration fees were **\$2.764 billion in 2025**, but the number of clients is not provided. Average fee per client would be:
\$2.764 billion ÷ number of clients.
- **✓ VERITROOPER (PASS):** The evidence does not provide the number of clients, so the average administration fee collected per client in 2025 cannot be calculated.

776. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0039\_03`

HALLUCINATION

- **Question:** What is the name of the CEO of NVIDIA Corporation?
- **Ground truth:** UNANSWERABLE
- **✗ Baseline (vanilla RAG) (FAIL):** Jen-Hsun Huang.
- **✓ VERITROOPER (PASS):** The evidence does not state the name of NVIDIA Corporation's CEO.

777. PASS · Precision · `SEC\_10K\_multi\_0079\_00`

- **Question:** What is the estimated fair value of debt as of January 25, 2026?
- **Ground truth:** 7.5 billion
- **✓ Baseline (vanilla RAG) (PASS):** The estimated fair value of debt as of January 25, 2026 was **\$7.5 billion**.
- **✓ VERITROOPER (PASS):** The estimated fair value of debt as of January 25, 2026 was **\$7.5 billion**.

778. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0717\_03`

- **Question:** What is the specific type of nonfinancial transaction that is eligible for hedge accounting under the new guidance?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify a particular type; it only states that the new guidance expands hedge-accounting eligibility for **hedges of nonfinancial transactions**.
- **✓ VERITROOPER (PASS):** The evidence does not specify the type of nonfinancial transaction eligible for hedge accounting.

779. FAIL · Precision · `SEC\_10K\_multi\_0592\_01`

META LEAK

- **Question:** What is the threshold for Metals & Mining?
- **Ground truth:** 406
- **✗ Baseline (vanilla RAG) (FAIL):** No specific "threshold" is stated for Metals & Mining. Its wholesale credit exposure was **\$17.767 billion in 2025**, compared with **\$15.860 billion in 2024**.
- **✓ VERITROOPER (PASS):** The evidence does not identify which column represents the threshold for Metals & Mining.

780. FAIL · Cross-Reference · `SEC\_10K\_multi\_0581\_00`

VALIDATOR OVER-FIRED

- **Question:** What percentage of total loans is comprised of investment-grade loans, given the total amount of investment-grade loans and the total amount of all loans?

- **Ground truth:** 68%

- **✗ Baseline (vanilla RAG) (FAIL):** Investment-grade loans comprise **68.32%** of total loans:

$$\frac{\$541.364 \text{ billion}}{\$792.367 \text{ billion}} \times 100 = 68.32\%$$

- **✓ VERITROOPER (PASS):** Investment-grade loans comprise **68%** of total loans (\$541,364 million ÷ \$792,367 million).

781. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0135\_01`

- **Question:** How many days does JPMorganChase have to correct a significant failure or disruption of its business and operations?

- **Ground truth:** UNANSWERABLE

- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not specify a number of days. It only states that JPMorganChase may need to expend significant resources to correct a failure or disruption.

- **✓ VERITROOPER (PASS):** The evidence does not specify how many days JPMorganChase has to correct such a failure or disruption.

782. FAIL · Precision · `SEC\_10K\_multi\_0861\_00`

UNSUPPORTED NUMBER

- **Question:** What is the value of Federal funds purchased and securities loaned or sold under repurchase agreements?

- **Ground truth:** 82.2

- **✗ Baseline (vanilla RAG) (FAIL):** Federal funds purchased and securities loaned or sold under repurchase agreements were **\$442.396 billion as of December 31, 2025**, compared with **\$296.835 billion in 2024**.

- **✓ VERITROOPER (PASS):** Federal funds purchased and securities loaned or sold under repurchase agreements were **\$82.2**.

783. PASS · Exception · `SEC\_10K\_multi\_0198\_01`

- **Question:** What financial information is not included in the XBRL interactive data files?

- **Ground truth:** Notes to Consolidated Financial Statements

- **✓ Baseline (vanilla RAG) (PASS):** The XBRL interactive data files do not include the Three-Year Summary of Consolidated Financial Highlights, Five-Year Stock Performance, or Management's Discussion and Analysis. They include only the consolidated financial statements and accompanying notes specified in the filing.

- **✓ VERITROOPER (PASS):** The Five-Year Stock Performance information is not listed as included in the XBRL interactive data files.

784. PASS · Cause & Effect · `SEC\_10K\_multi\_0132\_01`

- **Question:** What happens as a consequence of a downgrade in JPMorganChase's credit ratings?

- **Ground truth:** curtailment of business activities and profitability

- **✓ Baseline (vanilla RAG) (PASS):** A credit-rating downgrade could:

- Reduce JPMorganChase's access to capital markets and increase its funding and debt-servicing costs.
- Trigger additional collateral or funding requirements and termination payments on certain derivatives.
- Reduce the number of investors and counterparties willing or permitted to transact with or lend to it.
- Increase its credit spreads, lowering the value of its debt and other obligations.

- Curtail business activity, liquidity, and profitability.
- Adversely affect funding requirements for VIEs and other third-party commitments.

As of December 31, 2025, a one-notch downgrade could require \$28 million of additional derivative collateral and \$15 million in termination settlements; a two-notch downgrade could require \$124 million and \$96 million, respectively. Its unsecured debt would not be accelerated or require additional collateral solely because of a downgrade.

- ✓ **VERITROOPER (PASS):** A downgrade could reduce JPMorganChase's capital-market access, raise funding and servicing costs, trigger additional collateral or funding requirements, and reduce willing investors and counterparties. It could also widen credit spreads and lower the value of its debt and other obligations.

785. PASS · Cross-Reference · `SEC\_10K\_multi\_1535\_01`

- **Question:** How do we measure and support progress against our growth behaviors, and what tools do we use for this purpose?

- **Ground truth:** We use a variety of practices, including Performance Enablement and Culture & Engagement Survey platforms, to measure and support progress against our growth behaviors.

- ✓ **Baseline (vanilla RAG) (PASS):** We measure and support progress through:

- **Performance Enablement and Culture & Engagement Survey platforms**, which regularly gather employee feedback on how leaders, teammates, and workplace experiences support growth behaviors.

- **Anonymized questionnaire data**, compared with historical results to identify strengths and improvement opportunities across teams and functions.

- **Regular live communication**, including quarterly global town halls with senior leadership, employee Q&A sessions, and function-level town halls.

- ✓ **VERITROOPER (PASS):** We use a variety of practices, including the Performance Enablement and Culture & Engagement Survey platforms, to gather regular employee feedback on how leaders, teammates and work experiences support growth behaviors. Anonymized questionnaire data is compared with historical results to identify strengths and improvement opportunities.

786. PASS · Exception · `SEC\_10K\_multi\_1532\_01`

- **Question:** What is excluded from the general rule regarding bottler's agreements in the United States?

- **Ground truth:** Legacy bottler's agreements with no stated expiration date

- ✓ **Baseline (vanilla RAG) (PASS):** A small number of U.S. bottlers operate under legacy agreements with no stated expiration date for Trademark Coca-Cola and other cola-flavored Company Trademark Beverages.

- ✓ **VERITROOPER (PASS):** A small number of U.S. bottlers are excluded; they continue operating under legacy agreements with no stated expiration date for Trademark Coca-Cola and other cola-flavored Company Trademark Beverages.

787. PASS · Cause & Effect · `SEC\_10K\_multi\_1356\_01`

- **Question:** What is the consequence of the Firm having unrecognized tax benefits?

- **Ground truth:** The unrecognized tax benefits, if recognized, would reduce the annual effective tax rate.

- ✓ **Baseline (vanilla RAG) (PASS):** For JPMorganChase, changes in unrecognized tax benefits could result in a payment or recognition in the income statement. If recognized, \$4.6 billion of the \$5.6 billion balance at December 31, 2025 would reduce the Firm's annual effective tax rate.

- ✓ **VERITROOPER (PASS):** A change in unrecognized tax benefits would result in either a payment or recognition in the income statement. If recognized, certain benefits would reduce the annual effective tax rate.

788. PASS · Cross-Reference · `SEC\_10K\_multi\_1037\_01`

- **Question:** How are payments to co-brand partners based on new credit card account originations accounted for, and what is the period over which they are recognized?
- **Ground truth:** Payments to partners based on new credit card account originations are accounted for as direct loan origination costs and are deferred and recognized as a reduction of card income on a straight-line basis over a 12-month period.
- ✓ **Baseline (vanilla RAG) (PASS):** They are accounted for as direct loan origination costs, deferred, and recognized as a reduction of card income on a straight-line basis over 12 months.
- ✓ **VERITROOPER (PASS):** They are accounted for as direct loan origination costs, deferred, and recognized as a reduction of card income on a straight-line basis over 12 months.

789. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0516\_00`

BAD QUESTION

- **Question:** What is the average interest rate on non-interest-bearing deposits at JPMorgan Chase's U.S. offices as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** Not applicable (N/A). Noninterest-bearing deposits do not earn interest.
- ✗ **VERITROOPER (FAIL):** Not applicable (NA); non-interest-bearing deposits do not have an average interest rate.

790. PASS · Exception · `SEC\_10K\_multi\_0039\_00`

- **Question:** What circumstance does the forward-looking statement NOT apply to?
- **Ground truth:** Except as required by law
- ✓ **Baseline (vanilla RAG) (PASS):** Forward-looking statements do not apply to strictly historical or current facts.
- ✓ **VERITROOPER (PASS):** The forward-looking statement disclaimer does not apply when an update is required by law.

791. PASS · Cause & Effect · `SEC\_10K\_multi\_0760\_01`

- **Question:** What happens as a consequence of the determination to classify an instrument within level 3?
- **Ground truth:** The level 3 financial instruments typically include observable components in addition to the unobservable components.
- ✓ **Baseline (vanilla RAG) (PASS):** Once classified within Level 3 based on significant unobservable inputs, the instrument may still include observable components. Accordingly, reported gains and losses can partly reflect observable factors, while related Level 1 or 2 risk-management positions are excluded from the Level 3 amounts.
- ✓ **VERITROOPER (PASS):** Once classified within level 3, the instrument typically includes observable components alongside significant unobservable components; its level 1 and level 2 inputs are excluded from the level 3 inputs table.

792. PASS · Precision · `SEC\_10K\_multi\_0860\_01`

- **Question:** What is the carrying value of deposits with banks at December 31, 2025?
- **Ground truth:** 321.6
- ✓ **Baseline (vanilla RAG) (PASS):** The carrying value of deposits with banks at December 31, 2025 was \$321.6 billion.
- ✓ **VERITROOPER (PASS):** The carrying value of deposits with banks at December 31, 2025 was \$321.6 billion.

793. FAIL · Precision · `SEC\_10K\_multi\_0934\_01`

TABLE MISREAD

- **Question:** What is the total fair value of trading assets and liabilities before rounding?
- **Ground truth:** 663,566
- **✗ Baseline (vanilla RAG) (FAIL):** The total fair value was **\$1,018,892 million (\$1.018892 trillion)**, comprising:
 - Trading assets: **\$802,873 million**
 - Trading liabilities: **\$216,019 million**
- **✓ VERITROOPER (PASS):** The total fair value was \$665,587 for trading assets and \$60,967 for trading liabilities.

794. FAIL · Conditional · `SEC\_10K\_multi\_0942\_01`

BAD QUESTION

- **Question:** If the derivative payables are U.S. GAAP nettable and interest rate contracts, what is the net amount?
- **Ground truth:** 7,840
- **✗ Baseline (vanilla RAG) (FAIL):** The net derivative payables for U.S. GAAP-nettable interest rate contracts were **\$6.729 billion as of December 31, 2025**.
- **✓ VERITROOPER (PASS):** Net derivative payables were \$6,581 million for OTC and \$145 million for OTC-cleared interest rate contracts in 2025.

795. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1585\_01`

- **Question:** What is the average cost of shipping and handling activities per unit?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The average cost per unit cannot be determined from the evidence. It provides neither total shipping and handling costs nor the number of units shipped.
- **✓ VERITROOPER (PASS):** The evidence does not provide shipping and handling costs or unit volumes, so the average cost per unit cannot be determined.

796. FAIL · Cross-Reference · `SEC\_10K\_multi\_0467\_00`

VALIDATOR OVER-FIRED

- **Question:** What is the total stockholders' equity as of December 31, 2025, and what is the common stockholders' equity as of the same date?
- **Ground truth:** The total stockholders' equity is \$362,438, and the common stockholders' equity is \$342,393.
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025:
 - **Total stockholders' equity: \$362.438 billion**
 - **Common stockholders' equity: \$342.393 billion**
- **✓ VERITROOPER (PASS):** As of December 31, 2025, total stockholders' equity was **\$362,438 million**, and common stockholders' equity was **\$342,393 million**.

797. PASS · Cause & Effect · `SEC\_10K\_multi\_0442\_00`

- **Question:** What happens as a consequence of the Firm's capital management framework being established?

- **Ground truth:** The capital management framework specifies the principles underlying the Firm's approach towards capital management in normal economic conditions and in stressed environments.
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not specify the consequence of establishing the Firm's capital management framework.
- ✓ **VERITROOPER (PASS):** The Firm can calibrate targeted capital levels, meet minimum capital requirements, monitor planned capital distributions, and identify contingency actions to take or consider as capital is depleted during stress.

798. PASS · Cause & Effect · `SEC\_10K\_multi\_0131\_01`

- **Question:** What happens as a consequence of a CCP through which JPMorganChase executes contracts suffers a financial or operational failure?
- **Ground truth:** JPMorganChase would be required to replace the relevant contracts, which would increase its operational costs and potentially result in losses.
- ✓ **Baseline (vanilla RAG) (PASS):** JPMorganChase would have to replace the relevant contracts, increasing its operational costs and potentially resulting in losses.
- ✓ **VERITROOPER (PASS):** JPMorganChase would have to replace the relevant contracts, increasing operational costs and potentially resulting in losses.

799. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0539\_01`

- **Question:** What is the average credit rating of the subsidiaries listed?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The average credit rating cannot be determined from the provided evidence because no specific credit ratings for the subsidiaries are listed.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the subsidiaries' credit ratings, so an average cannot be calculated.

800. PASS · Cause & Effect · `SEC\_10K\_multi\_0535\_00`

- **Question:** What happens as a consequence of higher financing requirements in Markets?
- **Ground truth:** The increases in secured other borrowed funds at December 31, 2025, from December 31, 2024, and for the average year ended December 31, 2025, compared to the prior year.
- ✓ **Baseline (vanilla RAG) (PASS):** Higher financing requirements in Markets increased short-term borrowings and secured other borrowed funds.
- ✓ **VERITROOPER (PASS):** Secured other borrowed funds increased at December 31, 2025, and on average for the year compared with 2024.

801. FAIL · Precision · `SEC\_10K\_multi\_0744\_00`

BAD QUESTION

- **Question:** What is the total derivative payables?
- **Ground truth:** 14,806
- ✗ **Baseline (vanilla RAG) (FAIL):** Total derivative payables were **\$596.738 billion** as of December 31, 2025, or **\$46.329 billion after netting adjustments**.
- ✓ **VERITROOPER (PASS):** Total derivative payables were **\$46,329 million**.

802. PASS · Cause & Effect · `SEC\_10K\_multi\_1421\_01`

- **Question:** What is the consequence of the Russian courts' decisions against JPMorganChase?
- **Ground truth:** the Firm's assets in Russia could be seized

- ✓ **Baseline (vanilla RAG) (PASS):** Certain judgments are enforceable against JPMorganChase's Russian assets. Its onshore assets have already been used to satisfy one judgment, and further enforcement could seize all its assets in Russia, potentially seize client assets, or prevent the Firm from meeting its obligations. Courts have also frozen assets and allowed plaintiffs to withhold client dividends.
- ✓ **VERITROOPER (PASS):** Certain judgments are enforceable against JPMorganChase's Russian assets, potentially allowing all its assets in Russia-and some client assets-to be seized or preventing the Firm from meeting its obligations.

803. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1531\_06`

- **Question:** What is the average cost of goods sold for a bottling partner in the United States?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The average cost of goods sold for a U.S. bottling partner cannot be determined from the provided evidence. It does not disclose U.S.-specific bottling partner costs or the number of U.S. bottling partners.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the cost of goods sold for a U.S. bottling partner, so an average cannot be calculated.

804. PASS · Cross-Reference · `SEC\_10K\_multi\_0159\_00`

- **Question:** What is the total number of shares available for issuance under JPMorganChase's employee share-based incentive plans as of December 31, 2025, and how does it relate to the number of shares to be issued upon exercise of outstanding stock appreciation rights?
- **Ground truth:** 2,250,000
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, **77,131,051 shares** were available for future issuance under JPMorganChase's employee share-based incentive plans. Separately, **2,250,000 shares** were issuable upon exercise of outstanding stock appreciation rights, meaning the available pool was about **34.3 times** the shares underlying those rights.
- ✓ **VERITROOPER (PASS):** The evidence shows 2,250,000 shares to be issued upon exercise of outstanding stock appreciation rights, but it does not provide the shares remaining for future issuance; therefore, the total available cannot be determined.

805. PASS · Cross-Reference · `SEC\_10K\_multi\_1140\_01`

- **Question:** How are current property values estimated?
- **Ground truth:** using nationally recognized home price index valuation estimates incorporating actual and forecasted data
- ✓ **Baseline (vanilla RAG) (PASS):** Current property values are estimated at least quarterly using home valuation models based on nationally recognized home price indexes, incorporating actual data when available and forecasted data otherwise.
- ✓ **VERITROOPER (PASS):** Current property values are estimated at least quarterly using home valuation models based on nationally recognized home price index estimates, incorporating actual data where available and forecasted data otherwise.

806. PASS · Cross-Reference · `SEC\_10K\_multi\_0716\_01`

- **Question:** What is the required effective date for the new allowance framework, and what is the nature of the loans it excludes?
- **Ground truth:** The required effective date is January 1, 2027, and the new allowance framework excludes credit cards.
- ✓ **Baseline (vanilla RAG) (PASS):** The new allowance framework is effective **January 1, 2027**, and it excludes **credit card loans**.

- ✓ **VERITROOPER (PASS):** The required effective date is January 1, 2027. The framework excludes purchased, seasoned held-for-investment credit card loans.

807. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0134\_05`

- **Question:** What percentage of JPMorganChase's employees are responsible for managing cybersecurity risks?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not disclose what percentage of JPMorganChase employees are responsible for managing cybersecurity risks.
- ✓ **VERITROOPER (PASS):** The evidence does not state what percentage of JPMorganChase's employees are responsible for managing cybersecurity risks.

808. PASS · Precision · `SEC\_10K\_multi\_1606\_01`

- **Question:** What is the year in which the fiscal year ended?
- **Ground truth:** 2025
- ✓ **Baseline (vanilla RAG) (PASS):** The fiscal year ended in **2026**.
- ✓ **VERITROOPER (PASS):** The fiscal year ended in **2025**.

809. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0137\_04`

- **Question:** What is the average salary of JPMorganChase employees in the United States?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide salary or compensation data for JPMorganChase's U.S. employees, so the average salary cannot be determined.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the average salary of JPMorganChase employees in the United States.

810. PASS · Exception · `SEC\_10K\_multi\_0328\_00`

- **Question:** What does not apply to the provision for credit losses?
- **Ground truth:** the impact of changes in the Firm's weighted-average macroeconomic outlook
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not specify what does not apply to the provision for credit losses.
- ✓ **VERITROOPER (PASS):** The evidence does not provide answer choices, so it is insufficient to identify which statement does not apply.

811. PASS · Cause & Effect · `SEC\_10K\_multi\_1570\_01`

- **Question:** What is the consequence of the Company's guarantees?
- **Ground truth:** The maximum potential future payments that we could be required to make under the guarantees is \$786 million.
- ✓ **Baseline (vanilla RAG) (PASS):** The guarantees create a contingent liability of up to **\$786 million** as of December 31, 2025, including **\$61 million related to VIEs**. However, management considers the likelihood of any significant payment under these guarantees to be remote.
- ✓ **VERITROOPER (PASS):** The guarantees could require maximum future payments of \$786 million, including \$61 million related to VIEs; however, management considers significant payments remote.

812. FAIL · Precision · `SEC\_10K\_multi\_0825\_00`

TABLE MISREAD

- **Question:** What is the total number of mortgage-backed securities?
- **Ground truth:** 771
- **✗ Baseline (vanilla RAG) (FAIL):** As of December 31, 2025, total mortgage-backed securities classified as trading assets were **\$162.085 billion**.
- **✓ VERITROOPER (PASS):** The total mortgage-backed securities were **771**.

813. PASS · Conditional · `SEC\_10K\_multi\_0008\_00`

- **Question:** If the Company fails to adequately comply with rules and requirements regarding sensitive data, what can happen?
- **Ground truth:** The Company can be subject to litigation or government investigations, can be liable for associated investigatory expenses, and can incur significant fees or fines.
- **✓ Baseline (vanilla RAG) (PASS):** The Company may face litigation or government investigations, be liable for related investigatory expenses, and incur significant fees or fines.
- **✓ VERITROOPER (PASS):** The Company can face litigation or government investigations, incur associated investigatory expenses, and be liable for significant fees or fines.

814. FAIL · Cross-Reference · `SEC\_10K\_multi\_1173\_00`

BAD QUESTION

- **Question:** What is the percentage of loan modifications to total retained Commercial and industrial loans, and how does it compare to the number of single modifications?
- **Ground truth:** 0.71% of loan modifications to total retained Commercial and industrial loans, with 1,180 single modifications.
- **✗ Baseline (vanilla RAG) (FAIL):** For 2025, Commercial and industrial loan modifications represented **1.42%** of total retained C&I loans. Clearly identified single modifications totaled **\$1.997 billion**-\$1.308 billion of term extensions plus \$689 million of payment deferrals-representing **1.15%** of retained C&I loans and about **81%** of the total \$2.461 billion modified balance.
- **✗ VERITROOPER (FAIL):** Loan modifications totaled 1.12% of retained commercial and industrial loans. Single modifications accounted for \$1.644 billion, comprising 0.99% of retained loans.

815. PASS · Cross-Reference · `SEC\_10K\_multi\_0516\_00`

- **Question:** What is the total amount of available cash and securities as of December 31, 2025, and how does it compare to the total amount as of December 31, 2024?
- **Ground truth:** \$1.5 trillion and \$1.4 trillion, respectively, with a decrease of \$100 billion compared to the previous year.
- **✓ Baseline (vanilla RAG) (PASS):** As of December 31, 2025, available cash and securities totaled approximately **\$1.5 trillion**, up from **\$1.4 trillion** as of December 31, 2024-an increase of about **\$100 billion**.
- **✗ VERITROOPER (FAIL):** The Firm had approximately \$1.5 trillion of available cash and securities as of December 31, 2025, up \$0.1 trillion from \$1.4 trillion as of December 31, 2024.

816. PASS · Cross-Reference · `SEC\_10K\_multi\_0854\_00`

- **Question:** What were the total net gains/losses from the measurement alternative for the years ended December 31, 2023, 2024, and 2025?
- **Ground truth:** \$197 million, \$232 million, and \$122 million, respectively
- **✓ Baseline (vanilla RAG) (PASS):** Net gains/(losses) from the measurement alternative were:

- **2023: \$(232) million loss**
- **2024: \$(197) million loss**
- **2025: \$122 million gain**
- **✗ VERITROOPER (FAIL):** The total was a net loss of **\$307 million**: losses of \$232 million in 2023 and \$197 million in 2024, offset by a \$122 million gain in 2025.

817. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0576\_02`

- **Question:** What is the total amount of nonperforming exposure in the wholesale credit portfolio at December 31, 2027?
- **Ground truth:** UNANSWERABLE
- **✓ Baseline (vanilla RAG) (PASS):** The evidence does not provide data for December 31, 2027. At December 31, 2025, total nonperforming exposure in the wholesale credit portfolio was **\$6.477 billion**.
- **✓ VERITROOPER (PASS):** The evidence does not provide nonperforming exposure for December 31, 2027.

818. PASS · Conditional · `SEC\_10K\_multi\_0004\_00`

- **Question:** If the Company is unable to develop and market new products and services successfully, what impact could this have on its competitive advantage?
- **Ground truth:** A negative impact.
- **✓ Baseline (vanilla RAG) (PASS):** It could materially adversely affect the Company's ability to maintain a competitive advantage.
- **✓ VERITROOPER (PASS):** Its ability to maintain a competitive advantage could be materially adversely affected.

819. PASS · Cause & Effect · `SEC\_10K\_multi\_1280\_01`

- **Question:** What is the consequence of changes in MSR asset fair value due to market interest rates and other factors?
- **Ground truth:** Risk management activities are impacted.
- **✓ Baseline (vanilla RAG) (PASS):** Changes in MSR fair value due to market interest rates and related prepayment effects are recognized in income, affecting net mortgage servicing revenue. They produced gains of \$48 million in 2025, \$670 million in 2024, and \$424 million in 2023. Generally, declining rates reduce MSR value by increasing prepayments and shortening expected servicing cash flows.
- **✓ VERITROOPER (PASS):** The evidence does not specify the consequence of changes in MSR asset fair value due to market interest rates and other factors.

820. PASS · Cross-Reference · `SEC\_10K\_multi\_1528\_00`

- **Question:** What is the basis for presenting total net revenue for reportable business segments and Corporate, and the Firm as a whole?
- **Ground truth:** tax-equivalent basis
- **✓ Baseline (vanilla RAG) (PASS):** Total net revenue is presented on a **managed, fully taxable-equivalent (FTE) basis**. This starts with reported U.S. GAAP results and reclassifies revenue from tax-credit investments and tax-exempt securities to make it comparable with taxable investments and securities.
- **✓ VERITROOPER (PASS):** Total net revenue is presented on a tax-equivalent basis.

821. PASS · Cross-Reference · `SEC\_10K\_multi\_1609\_00`

- **Question:** What is the earliest date mentioned for the filing of the Company's Current Report on Form 8-K, and what is the corresponding exhibit number for the By-Laws of the Company?
- **Ground truth:** July 27, 2012, and 3.2
- ✓ **Baseline (vanilla RAG) (PASS):** The earliest filing date mentioned is **April 18, 2005**, and the By-Laws correspond to **Exhibit 3.2**.
- ✓ **VERITROOPER (PASS):** The earliest filing date mentioned is May 25, 2017, and the By-Laws are Exhibit 3.2.

822. FAIL · Precision · `SEC\_10K\_multi\_0724\_01`

UNSUPPORTED NUMBER

- **Question:** What is the total liabilities and stockholders' equity?
- **Ground truth:** \$4,424,900
- ✗ **Baseline (vanilla RAG) (FAIL):** Total liabilities and stockholders' equity were **\$4,424,900 million** (approximately **\$4.425 trillion**) as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** Total liabilities and stockholders' equity are **\$4,424,900** for Col 1 and **\$4,002,814** for Col 2.

823. PASS · Cause & Effect · `SEC\_10K\_multi\_1145\_00`

- **Question:** What happens as a result of forbearances granted to certain borrowers impacted by the wildfires in Los Angeles County, California in January 2025?
- **Ground truth:** Other-than-insignificant payment deferrals were driven by forbearances.
- ✓ **Baseline (vanilla RAG) (PASS):** The forbearances caused retained residential real estate nonaccrual loans-and therefore consumer nonperforming assets-to increase. A second 90-day forbearance also resulted in \$882 million of loan modifications involving significant payment deferrals in 2025.
- ✓ **VERITROOPER (PASS):** The forbearances led to a second 90-day arrangement, resulting in \$882 million of residential real estate FDMs classified as other-than-insignificant payment deferrals.

824. PASS · Cross-Reference · `SEC\_10K\_multi\_1529\_01`

- **Question:** What is the state of incorporation for Coca-Cola, and what is the I.R.S. Employer Identification Number?
- **Ground truth:** Delaware, 58-0628465.
- ✓ **Baseline (vanilla RAG) (PASS):** Coca-Cola is incorporated in **Delaware**, and its I.R.S. Employer Identification Number is **58-0628465**.
- ✓ **VERITROOPER (PASS):** Coca-Cola is incorporated in Delaware, and its I.R.S. Employer Identification Number is 58-0628465.

825. PASS · Cause & Effect · `SEC\_10K\_multi\_0134\_01`

- **Question:** What happens as a consequence of failure to identify operational risks?
- **Ground truth:** hindering JPMorganChase's ability to operate its businesses
- ✓ **Baseline (vanilla RAG) (PASS):** Failure to identify and address operational risks can lead to:
 - Disruption of business operations
 - Liability to clients, customers, and counterparties
 - Impaired liquidity
 - Weaker competitive standing
 - Higher compliance, operational, integration, and litigation costs

- Regulatory intervention and penalties
- Losses from fraudulent transactions
- Reputational harm

These consequences could materially and adversely affect JPMorganChase's businesses and results of operations.

- ✓ **VERITROOPER (PASS)**: Failure to identify operational risks can hinder operations, create client or counterparty liability, impair liquidity, weaken competitiveness, raise compliance, integration and litigation costs, trigger regulatory intervention and penalties, cause fraud losses, and harm reputation and results.

826. FAIL · Precision · `SEC\_10K\_multi\_0630\_01`

BAD QUESTION

- **Question**: What is the total for loans due after one year at fixed interest rates?
- **Ground truth**: 14,356
- ✗ **Baseline (vanilla RAG) (FAIL)**: The total consumer loans due after one year at fixed interest rates was **\$192.066 billion**.
- ✗ **VERITROOPER (FAIL)**: The total loans due after one year at fixed interest rates were **\$86,224**.

827. PASS · Exception · `SEC\_10K\_multi\_1542\_00`

- **Question**: What circumstance could result in the Company's inability to sell products?
- **Ground truth**: negative impact on our reputation
- ✓ **Baseline (vanilla RAG) (PASS)**: U.S. or other countries' trade restrictions could prevent the Company from selling products to a customer.
- ✓ **VERITROOPER (PASS)**: The Company's ability to sell products could be negatively impacted if it is perceived as not responding appropriately to growing sustainability concerns.

828. PASS · Precision · `SEC\_10K\_multi\_0925\_00`

- **Question**: What is the year for which the notional amount of free-standing derivative contracts is reported?
- **Ground truth**: 2025
- ✓ **Baseline (vanilla RAG) (PASS)**: The notional amounts of free-standing derivative contracts are reported for **2025 and 2024**, as of December 31 of each year.
- ✓ **VERITROOPER (PASS)**: The notional amounts are reported as of December 31, 2025 and 2024.

829. PASS · Calculation · `SEC\_10K\_multi\_1602\_00`

- **Question**: What is the total of the impairment charges related to the BodyArmor trademark in North America for the years 2025 and 2024?
- **Ground truth**: 1,720
- ✓ **Baseline (vanilla RAG) (PASS)**: 1720
- ✓ **VERITROOPER (PASS)**: 1,720

830. PASS · Precision · `SEC\_10K\_multi\_0378\_01`

- **Question**: What is the revenue for 2024?
- **Ground truth**: 13,693
- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not provide total revenue for 2024. It only gives each operating segment's percentage of net operating revenues for 2024.
- ✓ **VERITROOPER (PASS)**: Asset management fee revenue for 2024 was \$13,693 million.

831. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0080\_02`

HALLUCINATION

- **Question:** What is the name of the regulatory agency conducting the inquiries and investigations mentioned in the passage?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The U.S. Securities and Exchange Commission (SEC).
- ✓ **VERITROOPER (PASS):** The passage does not name any regulatory agency conducting the inquiries and investigations.

832. PASS · Cross-Reference · `SEC\_10K\_multi\_0655\_01`

- **Question:** What types of risks does Corporate manage related to the Firm's liquidity?
- **Ground truth:** funding, capital, structural interest rate and foreign exchange risks
- ✓ **Baseline (vanilla RAG) (PASS):** Corporate manages strategic risk that could affect the Firm's liquidity, earnings, capital, or reputation. Through ALCO, it also oversees liquidity risk, interest-rate risk, and capital risk as part of asset and liability management.
- ✓ **VERITROOPER (PASS):** Corporate manages the Firm's liquidity, funding, capital, structural interest rate, and foreign exchange risks.

833. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1609\_01`

- **Question:** What is the average interest rate of the notes issued by the Company in 2020?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide the interest rates for the notes issued in 2020, so their average interest rate cannot be determined.
- ✗ **VERITROOPER (FAIL):** The average interest rate of the 15 notes issued in 2020 was approximately 2.15%.

834. PASS · Conditional · `SEC\_10K\_multi\_0133\_00`

- **Question:** If JPMorganChase experiences a failure in its operational systems, what could be a negative effect on its clients and customers?
- **Ground truth:** delays or other disruptions in providing services
- ✓ **Baseline (vanilla RAG) (PASS):** Clients and customers could face service disruptions, such as being unable to access account information or complete transactions through ATMs, internet, or mobile channels.
- ✓ **VERITROOPER (PASS):** Clients and customers could be unable to access account information or conduct transactions through ATMs, internet, or mobile channels.

835. PASS · Precision · `SEC\_10K\_multi\_1183\_01`

- **Question:** What is the total for defaults of FDMs in 2025?
- **Ground truth:** 173
- ✓ **Baseline (vanilla RAG) (PASS):** The total defaults of retained wholesale FDMs in 2025 were **\$241 million** (\$173 million + \$65 million + \$3 million).
- ✓ **VERITROOPER (PASS):** The total defaults of FDMs in 2025 were **\$241 million**.

836. FAIL · Cross-Reference · `SEC\_10K\_multi\_1158\_01`

MISINTERPRETED

- **Question:** What is the total amount of gross charge-offs for the year 2024, and how does it compare to the total gross charge-offs for the year 2023?
- **Ground truth:** The total gross charge-offs for 2024 is \$-, and for 2023 is \$38.
- **✗ Baseline (vanilla RAG) (FAIL):** Gross charge-offs were **\$43 million for 2024-originated loans**, compared with **\$64 million for 2023-originated loans**-a decrease of **\$21 million**, or approximately **32.8%**.
- **✓ VERITROOPER (PASS):** Total gross charge-offs for 2024 were \$173 million. The evidence does not provide total gross charge-offs for 2023, so a comparison cannot be made.

837. FAIL · Precision · `SEC\_10K\_multi\_0259\_01`

UNSUPPORTED NUMBER

- **Question:** What is the income tax expense?
- **Ground truth:** 15,547
- **✗ Baseline (vanilla RAG) (FAIL):** JPMorgan Chase's income tax expense was **\$15.547 billion for 2025**.
- **✓ VERITROOPER (PASS):** Income tax expense is 15,547 in Col 1; the table also lists 3,134, 18,681, and 16,610 in Cols 2-4.

838. PASS · Precision · `SEC\_10K\_multi\_0322\_00`

- **Question:** What is the year of the Form 10-K?
- **Ground truth:** 2025
- **✓ Baseline (vanilla RAG) (PASS):** 2025.
- **✓ VERITROOPER (PASS):** 2025.

839. PASS · Precision · `SEC\_10K\_multi\_0699\_01`

- **Question:** What is the total exposure for Brazil?
- **Ground truth:** \$20.9
- **✓ Baseline (vanilla RAG) (PASS):** Brazil's total exposure was **\$20.9 billion as of December 31, 2025**.
- **✓ VERITROOPER (PASS):** The total exposure for Brazil is \$20.9.

840. PASS · Precision · `SEC\_10K\_multi\_0231\_01`

- **Question:** What is the total credit provided and capital raised for consumers?
- **Ground truth:** \$33 billion
- **✓ Baseline (vanilla RAG) (PASS):** The total credit provided and capital raised for consumers was approximately **\$280 billion** in 2025.
- **✓ VERITROOPER (PASS):** The total credit provided and capital raised for consumers was **\$280 billion in 2025**.

841. PASS · Exception · `SEC\_10K\_multi\_0618\_01`

- **Question:** What is excluded from the amounts listed?
- **Ground truth:** ratios
- **✓ Baseline (vanilla RAG) (PASS):** The listed carrying amounts exclude:
 - Physical commodities subject to fair value hedge accounting; and
 - Hedged items for which foreign-currency risk is the only designated hedged risk.
- **✓ VERITROOPER (PASS):** The evidence does not specify what is excluded from the amounts listed.

842. PASS · Exception · `SEC\_10K\_multi\_0248\_01`

- **Question:** What circumstance does the income tax expense NOT apply to?
- **Ground truth:** restructuring and integration costs associated with First Republic
- ✓ **Baseline (vanilla RAG) (PASS):** Income tax expense does not apply to non-income-tax-related gains, losses, and other returns; these are recognized in **Other income**.
- ✓ **VERITROOPER (PASS):** The evidence does not specify any circumstance to which income tax expense does not apply.

843. PASS · Cross-Reference · `SEC\_10K\_multi\_0053\_00`

- **Question:** What could increase our worldwide effective tax rate and harm our financial position, results of operations, net income, and cash flows?
- **Ground truth:** An adverse outcome in tax audits or changes in tax laws or their interpretation by tax authorities.
- ✓ **Baseline (vanilla RAG) (PASS):** An adverse outcome from audits by tax authorities could increase the worldwide effective tax rate and harm the company's financial position, results of operations, net income, and cash flows.
- ✓ **VERITROOPER (PASS):** An adverse outcome from audits by tax authorities in different jurisdictions could increase the worldwide effective tax rate and harm the company's financial position, results of operations, net income, and cash flows.

844. PASS · Cross-Reference · `SEC\_10K\_multi\_1269\_00`

- **Question:** What is the method used by the Firm to estimate the fair value of its reporting units, and how does it compare to market-based trading and transaction multiples?
- **Ground truth:** The Firm uses the income approach, projecting cash flows for the forecast period and using the perpetuity growth method to calculate terminal values, which are then discounted using an appropriate discount rate. This is compared with market-based trading and transaction multiples for relevant competitors.
- ✓ **Baseline (vanilla RAG) (PASS):** The Firm primarily uses the **income approach**, projecting cash flows over a forecast period, calculating terminal values using the perpetuity growth method, and discounting both using a reporting-unit-specific cost of equity.

It then compares the resulting discounted cash flow valuations with **market-based trading and transaction multiples of relevant competitors** as general indicators of reasonableness. Precise comparisons are limited because of differences between the Firm's businesses and competitor institutions.

- ✓ **VERITROOPER (PASS):** The Firm primarily uses the income approach, projecting and discounting forecast-period cash flows and terminal values calculated using the perpetuity growth method. It compares these valuations with market-based trading and transaction multiples for relevant competitors as general reasonableness indicators.

845. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0618\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the dollar amount of net portfolio activity as of December 31, 2025?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** The net portfolio activity was **\$747 million** as of December 31, 2025.
- ✓ **VERITROOPER (PASS):** The evidence does not provide the dollar amount of net portfolio activity as of December 31, 2025.

846. PASS · Precision · `SEC\_10K\_multi\_1455\_00`

- **Question:** What is the net income for 2023?
- **Ground truth:** 49,552
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence provided does not state net income for 2023.
- ✓ **VERITROOPER (PASS):** The net income for 2023 was **\$49,552 million**.

847. PASS · Precision · `SEC\_10K\_multi\_1136\_00`

- **Question:** What percentage of Chapter 7 residential real estate loans were 30 days or more past due at December 31, 2025?
- **Ground truth:** 9%
- ✓ **Baseline (vanilla RAG) (PASS):** Approximately 9%.
- ✓ **VERITROOPER (PASS):** Approximately 9%.

848. FAIL · Precision · `SEC\_10K\_multi\_1436\_00`

BAD QUESTION

- **Question:** What is the total noninterest expense?
- **Ground truth:** 38,036
- ✗ **Baseline (vanilla RAG) (FAIL):** Total noninterest expense was **\$95.64 billion in 2025**.
- ✓ **VERITROOPER (PASS):** Total noninterest expense was **\$40,267 million**.

849. PASS · Precision · `SEC\_10K\_multi\_0164\_01`

- **Question:** What is the date of the year-end for the Deferred Compensation Plan?
- **Ground truth:** December 31, 2004
- ✓ **Baseline (vanilla RAG) (PASS):** The year-end date was **September 27, 2025**.
- ✓ **VERITROOPER (PASS):** The year-end date for the Deferred Compensation Plan is December 31, 2004.

850. PASS · Exception · `SEC\_10K\_multi\_0052\_00`

- **Question:** What circumstance could undermine public confidence in AI?
- **Ground truth:** d regulations relating to the responsible use of AI by us or others
- ✓ **Baseline (vanilla RAG) (PASS):** Failure to adequately address concerns and regulations related to the responsible use of AI by the company or others could undermine public confidence in AI.
- ✓ **VERITROOPER (PASS):** Failure by the company or others to comply with regulations on the responsible use of AI could undermine public confidence.

851. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0581\_02`

- **Question:** What is the total value of derivative receivables that are not collateralized?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** As of December 31, 2025, derivative receivables not covered by collateral totaled **\$27.937 billion** (\$27,937 million).
- ✗ **VERITROOPER (FAIL):** The total derivative receivables not collateralized were **\$28,886 million**.

852. PASS · Cause & Effect · `SEC\_10K\_multi\_1134\_01`

- **Question:** What happens as a consequence of revolving loans being within the revolving period?
- **Ground truth:** They have lower delinquency rates compared to those that have been converted to term loans.

- ✓ **Baseline (vanilla RAG) (PASS):** Revolving loans within the revolving period offer lower minimum payment options than the fully amortizing payments generally required after conversion to term loans. Consequently, they have lower delinquency rates.
- ✓ **VERITROOPER (PASS):** Revolving loans within the revolving period have lower delinquency rates because they offer minimum payment options that are generally lower than fully amortizing payments.

853. PASS · Calculation · `SEC\_10K\_multi\_1586\_01`

- **Question:** If the Company grants stock options with a vesting period of 4 years, and the fair value of the stock option awards is estimated using the Black-Scholes-Merton option-pricing model, what is the number of years over which the Company recognizes compensation expense?
- **Ground truth:** 4
- ✓ **Baseline (vanilla RAG) (PASS):** 4
- ✓ **VERITROOPER (PASS):** 4

854. PASS · Cross-Reference · `SEC\_10K\_multi\_0681\_01`

- **Question:** What is the significance of the Firm's economic value sensitivity (EVS) measure?
- **Ground truth:** EVS stress tests the longer-term economic value of equity by measuring the sensitivity of the Firm's current balance sheet under various interest rate scenarios.
- ✓ **Baseline (vanilla RAG) (PASS):** The EVS measure assesses the longer-term economic value of the Firm's equity by stress-testing how interest-rate scenarios affect its current balance sheet, primarily retained loans, deposits, debt, investment securities, and related hedges. It complements the 12-month earnings-at-risk measure and is managed against limits set as a percentage of tangible common equity (TCE). Deposit pricing and cash-flow assumptions, along with loan and securities prepayment assumptions, significantly influence EVS.
- ✓ **VERITROOPER (PASS):** EVS stress-tests the longer-term economic value of equity by measuring the sensitivity of the Firm's current balance sheet and related hedges under various interest-rate scenarios. The Firm manages this risk by setting EVS limits as a percentage of tangible common equity (TCE).

855. FAIL · Calculation · `SEC\_10K\_multi\_1335\_01`

BAD QUESTION

- **Question:** What is the aggregate purchase price of common stock repurchases in 2023?
- **Ground truth:** 69.5
- ✗ **Baseline (vanilla RAG) (FAIL):** 9898
- ✗ **VERITROOPER (FAIL):** Insufficient information

856. PASS · Precision · `SEC\_10K\_multi\_0537\_01`

- **Question:** What is the amount of credit card securitization in 2025?
- **Ground truth:** 1,498
- ✓ **Baseline (vanilla RAG) (PASS):** Credit card securitization issuance in 2025 was **\$1.498 billion**.
- ✓ **VERITROOPER (PASS):** Credit card securitization issuance in 2025 was **\$1,498 million**.

857. PASS · Cause & Effect · `SEC\_10K\_multi\_0706\_00`

- **Question:** What happens as a result of the Firm experiencing a higher volume and complexity of cyber attacks?
- **Ground truth:** a material adverse effect on the Firm or its business strategy, results of operations or financial condition

- ✓ **Baseline (vanilla RAG) (PASS):** The Firm implements measures and controls to address the evolving threat environment, including enhanced threat monitoring, and continues reviewing and strengthening capabilities such as administrative-access management.
- ✓ **VERITROOPER (PASS):** The Firm implements measures and controls, including enhanced threat monitoring, and continues reviewing and improving its cybersecurity capabilities.

858. PASS · Exception · `SEC\_10K\_multi\_0897\_00`

- **Question:** When is the remaining contractual principal not applicable?
- **Ground truth:** nonprincipal-protected structured notes and long-term beneficial interests
- ✓ **Baseline (vanilla RAG) (PASS):** Remaining contractual principal is not applicable to **nonprincipal-protected structured notes and long-term beneficial interests**, because these instruments do not obligate the firm to return a stated principal amount at maturity.
- ✓ **VERITROOPER (PASS):** Remaining contractual principal is not applicable to nonprincipal-protected structured notes and long-term beneficial interests.

859. PASS · Cross-Reference · `SEC\_10K\_multi\_1525\_00`

- **Question:** What is the relationship between the Combined loan-to-value (CLTV) and the loan modifications effective January 1, 2023, as defined in the FDM?
- **Ground truth:** CLTV is not directly related to FDM, as FDM refers to loan modifications, not loan-to-value ratios.
- ✓ **Baseline (vanilla RAG) (PASS):** There is no direct definitional relationship stated between CLTV and FDMs.
- **CLTV** measures the loan balance relative to property value, considering all lien positions and unused lines, particularly for junior-lien home equity loans.
- **FDM**, effective January 1, 2023, classifies modifications for borrowers experiencing financial difficulty, including principal forgiveness, interest-rate reductions, significant payment delays, term extensions, or combinations thereof.

Thus, CLTV is a collateral/leverage measure, while FDM describes qualifying changes to loan terms.

- ✓ **VERITROOPER (PASS):** CLTV means "Combined loan-to-value." The evidence defines FDM loan modifications effective January 1, 2023, but states no relationship between CLTV and those modifications.

860. PASS · Cause & Effect · `SEC\_10K\_multi\_0049\_00`

- **Question:** What happens as a consequence of changes in immigration and work permit regulations?
- **Ground truth:** Our ability to attract, employ, and retain qualified employees may be impaired.
- ✓ **Baseline (vanilla RAG) (PASS):** Changes in immigration and work permit regulations could impair companies' ability to attract, employ, and retain qualified, highly skilled employees, potentially harming operations, strategy, culture, and costs.
- ✓ **VERITROOPER (PASS):** Changes in immigration and work permit regulations, or their administration or interpretation, could impair the company's ability to attract, employ, and retain qualified employees.

861. PASS · Cross-Reference · `SEC\_10K\_multi\_0642\_01`

- **Question:** What is the total allowance for loan losses at the beginning of the year 2025?
- **Ground truth:** 24,345
- ✓ **Baseline (vanilla RAG) (PASS):** The total allowance for loan losses at January 1, 2025 was **\$24.345 billion**.
- ✓ **VERITROOPER (PASS):** The total allowance for loan losses at the beginning of 2025 was **\$24.345 billion**.

862. PASS · Exception · `SEC\_10K\_multi\_0067\_00`

- **Question:** Under what circumstance do PSUs not vest?
- **Ground truth:** subject to continued service and performance conditions
- ✓ **Baseline (vanilla RAG) (PASS):** PSUs do not vest if the specified, predetermined performance criteria are not achieved. They are also generally subject to continued-service requirements.
- ✓ **VERITROOPER (PASS):** PSUs do not vest if continued-service or performance conditions are not met.

863. PASS · Conditional · `SEC\_10K\_multi\_1058\_01`

- **Question:** If the net gain is attributable to higher than expected returns on plan assets, what is the impact on the net periodic benefit cost?
- **Ground truth:** A credit
- ✓ **Baseline (vanilla RAG) (PASS):** Higher-than-expected returns create a net gain recorded initially in OCI. If the gain exceeds the 10% corridor, it is amortized over subsequent periods, reducing net periodic benefit cost (or increasing the net periodic benefit credit).
- ✓ **VERITROOPER (PASS):** The net gain is recorded in OCI and, if it exceeds the 10% threshold, is amortized into net periodic benefit cost over subsequent periods, reducing that cost.

864. PASS · Cross-Reference · `SEC\_10K\_multi\_1592\_00`

- **Question:** What is the total amount of dividends received from equity method investees for the years ended December 31, 2025, 2024, and 2023?
- **Ground truth:** 993 million, 968 million, and 672 million, respectively.
- ✓ **Baseline (vanilla RAG) (PASS):** Dividends received from equity method investees totaled:
 - **2025:** \$993 million
 - **2024:** \$968 million
 - **2023:** \$672 million
- ✓ **VERITROOPER (PASS):** Dividends received from equity method investees totaled \$993 million in 2025, \$968 million in 2024, and \$672 million in 2023.

865. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_0003\_02`

- **Question:** How many employees does Apple have in its retail stores worldwide?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The provided evidence does not state how many employees Apple has in its retail stores worldwide.
- ✓ **VERITROOPER (PASS):** The evidence does not state how many employees Apple has in its retail stores worldwide.

866. PASS · Precision · `SEC\_10K\_multi\_1169\_00`

- **Question:** What is the weighted-average of months extended for single modifications in 2023?
- **Ground truth:** 14
- ✓ **Baseline (vanilla RAG) (PASS):** The loans were extended by a weighted average of **17 months**.
- ✓ **VERITROOPER (PASS):** The weighted-average term extension for single modifications in 2023 was **14 months**.

867. PASS · Cross-Reference · `SEC\_10K\_multi\_0492\_01`

- **Question:** What are the factors considered for Operational risk RWA?
- **Ground truth:** Updates to cumulative losses, macroeconomic model inputs, and other model parameters.
- ✓ **Baseline (vanilla RAG) (PASS):** Operational risk RWA considers:
 - Updates to cumulative losses
 - Macroeconomic model inputs
 - Other model parameters
- ✓ **VERITROOPER (PASS):** Operational risk RWA considers updates to cumulative losses, macroeconomic model inputs, and other model parameters.

868. PASS · Precision · `SEC\_10K\_multi\_1537\_01`

- **Question:** What is the number of the section that discusses public debate and concern about perceived negative health consequences of processing and of certain ingredients?
- **Ground truth:** 16
- ✓ **Baseline (vanilla RAG) (PASS):** Section 16.
- ✓ **VERITROOPER (PASS):** The section is numbered 16.

869. FAIL · Precision · `SEC\_10K\_multi\_0477\_00`

DETERMINISTIC:GT NUMBER CONFLICT

- **Question:** What was the amount of dividends declared on common stock for the year ended December 31, 2025?
- **Ground truth:** -16,060
- ✗ **Baseline (vanilla RAG) (FAIL):** The disclosed quarterly common stock dividend was **\$1.50 per share**, declared on December 9, 2025.
- ✓ **VERITROOPER (PASS):** Dividends declared on common stock were **\$16,060 million** for the year ended December 31, 2025.

870. PASS · Cause & Effect · `SEC\_10K\_multi\_0392\_01`

- **Question:** What happens due to higher market levels and continued net inflows?
- **Ground truth:** Assets under management were \$4.8 trillion, up 18%, and client assets were \$7.1 trillion, up 20%.
- ✓ **Baseline (vanilla RAG) (PASS):** Assets under management rose 18% to \$4.8 trillion, and client assets increased 20% to \$7.1 trillion.
- ✓ **VERITROOPER (PASS):** Assets under management and client assets increased, rising 18% to \$4.8 trillion and 20% to \$7.1 trillion, respectively.

871. FAIL · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1409\_04`

DETERMINISTIC:ANSWERED UNANSWERABLE

- **Question:** What is the percentage of risk participations for standby letters of credit compared to other letters of credit?
- **Ground truth:** UNANSWERABLE
- ✗ **Baseline (vanilla RAG) (FAIL):** In 2025, standby letters of credit had **approximately 1,790%** of the risk participations of other letters of credit:

$$\frac{\$9.2 \text{ billion}}{\$514 \text{ million}} \times 100 \approx 1,790\%$$

\]

That is about **17.9 times** as much.

- ✓ **VERITROOPER (PASS)**: The evidence does not provide the amounts of risk participations themselves; it only gives contractual amounts net of risk participations, so the requested percentage cannot be calculated.

872. PASS · Conditional · `SEC\_10K\_multi\_0778\_01`

- **Question**: If a correlation is positively correlated, what happens to the other parameter when one parameter increases?

- **Ground truth**: It also increases.

- ✓ **Baseline (vanilla RAG) (PASS)**: The other parameter increases.

- ✓ **VERITROOPER (PASS)**: When parameters are positively correlated, an increase in one parameter results in an increase in the other parameter.

873. PASS · Cause & Effect · `SEC\_10K\_multi\_1462\_01`

- **Question**: What happens as a result of the Firm reaching an agreement with the FDIC?

- **Ground truth**: The Firm made a payment of \$609 million to the FDIC and recorded a gain of \$588 million.

- ✓ **Baseline (vanilla RAG) (PASS)**: As a result of the FDIC agreement, the Firm:

- Paid \$609 million to the FDIC on January 31, 2025.

- Reduced its additional payable to the FDIC.

- Recorded a \$588 million gain in other income in the first quarter of 2025.

By June 30, 2025, all outstanding purchase-price settlement matters related to the First Republic acquisition had been resolved.

- ✓ **VERITROOPER (PASS)**: The Firm paid \$609 million to the FDIC, reduced its additional payable, and recorded a \$588 million gain in other income in the first quarter of 2025.

874. PASS · Precision · `SEC\_10K\_multi\_1606\_00`

- **Question**: What is the number of days after the end of the fiscal year to file the Proxy Statement?

- **Ground truth**: 120

- ✓ **Baseline (vanilla RAG) (PASS)**: The Proxy Statement must be filed within **120 days** after the end of the fiscal year.

- ✓ **VERITROOPER (PASS)**: 120 days.

875. PASS · Precision · `SEC\_10K\_multi\_1614\_00`

- **Question**: What is the year in which The Coca-Cola Company's Annual Report on Form 10-K was filed for the year ended December 31, 2025?

- **Ground truth**: 2026

- ✓ **Baseline (vanilla RAG) (PASS)**: 2026.

- ✓ **VERITROOPER (PASS)**: The Annual Report on Form 10-K was filed in 2026.

876. PASS · Precision · `SEC\_10K\_multi\_1454\_01`

- **Question**: What is the total assets?

- **Ground truth**: 710,044

- ✓ **Baseline (vanilla RAG) (PASS)**: The evidence does not provide a total assets figure.

- ✓ **VERITROOPER (PASS)**: Total assets were **\$710,044 in Col 1** and **\$669,614 in Col 2**.

877. PASS · Precision · `SEC\_10K\_multi\_0029\_00`

- **Question:** What is the total unrecognized compensation cost related to outstanding RSUs as of September 27, 2025?
- **Ground truth:** 21.8 billion
- ✓ **Baseline (vanilla RAG) (PASS):** The total unrecognized compensation cost related to outstanding RSUs was **\$21.8 billion** as of September 27, 2025.
- ✓ **VERITROOPER (PASS):** The total unrecognized compensation cost related to outstanding RSUs was **\$21.8 billion** as of September 27, 2025.

878. PASS · Cause & Effect · `SEC\_10K\_multi\_1284\_00`

- **Question:** What happens because of a 10% adverse change in the weighted-average prepayment speed assumption?
- **Ground truth:** The fair value decreases by \$181 million.
- ✓ **Baseline (vanilla RAG) (PASS):** A 10% adverse change reduces the fair value of JPMorgan's MSRs by **\$181 million in 2025 and \$209 million in 2024.**
- ✓ **VERITROOPER (PASS):** A 10% adverse change reduces fair value by \$181 million in 2025 and \$209 million in 2024.

879. PASS · Exception · `SEC\_10K\_multi\_1542\_01`

- **Question:** What could lead to a reconsideration of investment in the Company?
- **Ground truth:** investors' dissatisfaction with our sustainability practices
- ✓ **Baseline (vanilla RAG) (PASS):** A reconsideration could be prompted by material harm to the Company's business, reputation, operating results, financial condition, or stock price-for example, from failure to retain skilled personnel and maintain its culture, geopolitical and tariff disputes, or major supply-chain and operational disruptions.
- ✓ **VERITROOPER (PASS):** Certain investors may reconsider their investment if they are dissatisfied with the Company's sustainability practices, actions, or progress.

880. PASS · Calculation · `SEC\_10K\_multi\_0005\_00`

- **Question:** If the Company records a write-down for product and component inventories if cost exceeds net realizable value, and the cost is \$120 and the net realizable value is \$90, what is the write-down?
- **Ground truth:** 30
- ✓ **Baseline (vanilla RAG) (PASS):** 30
- ✓ **VERITROOPER (PASS):** 30

881. PASS · Exception · `SEC\_10K\_multi\_0065\_00`

- **Question:** Under what circumstance does the accounting for changes in the fair value of a derivative NOT depend on the intended use of the derivative?
- **Ground truth:** For derivative instruments designated as accounting hedges
- ✓ **Baseline (vanilla RAG) (PASS):** When the derivative is **not designated as an accounting hedge**, changes in its fair value are recognized directly in earnings, regardless of its intended use.
- ✓ **VERITROOPER (PASS):** The evidence does not identify any circumstance; it states that accounting for fair value changes depends on the derivative's intended use and resulting designation.

882. PASS · Precision · `SEC\_10K\_multi\_0768\_01`

- **Question:** What is the threshold for Interest rate volatility?
- **Ground truth:** 24
- ✓ **Baseline (vanilla RAG) (PASS):** Interest rate volatility ranges from **24 bps to 490 bps**, with a weighted average of **85 bps**.
- ✓ **VERITROOPER (PASS):** The interest rate volatility range is 24-490 bps.

883. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1585\_00`

- **Question:** What is the total amount of advertising costs incurred in 2026?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide advertising costs for 2026. It reports \$5.4 billion in 2025, \$5.1 billion in 2024, and \$5.0 billion in 2023.
- ✓ **VERITROOPER (PASS):** The evidence does not provide advertising costs for 2026.

884. PASS · Exception · `SEC\_10K\_multi\_1177\_00`

- **Question:** What type of loan modifications are excluded from the calculation?
- **Ground truth:** Single modifications
- ✓ **Baseline (vanilla RAG) (PASS):** Insignificant payment delays are excluded; only **other-than-insignificant payment delays** qualify as financial difficulty modifications (FDMs).
- ✓ **VERITROOPER (PASS):** The evidence does not specify any loan modification types excluded from the calculation; it states that loans with multiple modifications are included.

885. PASS · Cause & Effect · `SEC\_10K\_multi\_0046\_00`

- **Question:** What happens as a consequence of suppliers failing to provide high-quality products?
- **Ground truth:** Delays in product shipments, shortages, a decrease in product quality, and/or higher expenses.
- ✓ **Baseline (vanilla RAG) (PASS):** It could cause defects, reduced product quality, shipment delays, higher expenses, and harm to business operations, revenue, gross margin, and financial results.
- ✓ **VERITROOPER (PASS):** It can cause product defects, leading to significant remediation, warranty, support, repair or replacement costs, reputational damage, and loss of market share.

886. PASS · Conditional · `SEC\_10K\_multi\_0727\_00`

- **Question:** If prices or quotes are not available for an instrument, what is generally determined using valuation models?
- **Ground truth:** fair value
- ✓ **Baseline (vanilla RAG) (PASS):** The instrument's fair value is generally determined using valuation models.
- ✓ **VERITROOPER (PASS):** Fair value is generally determined using valuation models.

887. PASS · Negative (hallucination traps) · `SEC\_10K\_multi\_neg\_1151\_03`

- **Question:** What is the percentage of loan modifications that are reinstated in 2025?
- **Ground truth:** UNANSWERABLE
- ✓ **Baseline (vanilla RAG) (PASS):** The evidence does not provide a percentage of loan modifications reinstated in 2025; it only states that, in most cases, the borrower's line of credit is not reinstated.
- ✓ **VERITROOPER (PASS):** The evidence does not provide a percentage of loan modifications reinstated in 2025; it only states that, in most cases, the borrower's line of credit is not reinstated.

888. FAIL · Calculation · `SEC\_10K\_multi\_0275\_01`

BAD QUESTION

- **Question:** What is TBVPS for December 31, 2025?
- **Ground truth:** 0.845
- **✗ Baseline (vanilla RAG) (FAIL):** 107.56
- **✗ VERITROOPER (FAIL):** Insufficient information

889. PASS · Precision · `SEC\_10K\_multi\_0033\_01`

- **Question:** What is the year in which the Company's 2026 Proxy Statement will be filed?
- **Ground truth:** 2026
- **✓ Baseline (vanilla RAG) (PASS):** The Company's Proxy Statement will be filed in **2026**.
- **✓ VERITROOPER (PASS):** The Company's 2026 Proxy Statement will be filed in 2026.

