

Benchmark report: FinanceBench

This benchmark report was generated automatically by the Ask This Guy platform. It measures the answers of an Ask This Guy agent on the FinanceBench public benchmark, under the conditions described below, and shows every question with the delivered answer and its verdict.

Report generated on 2026-10-02 16:37:03 UTC.

For more information, contact bonjour@askthisguy.com.

About FinanceBench

FinanceBench tests open-book question answering on the public filings of listed companies. Each question asks for a fact, a figure or a judgment that an analyst would find in, or compute from, a company's annual, quarterly or current reports: answering requires finding the right passage or table in long financial documents, then often a calculation or a short reasoning.

- **Developed by:** Patronus AI (Pranab Islam, Anand Kannappan, Douwe Kiela, Rebecca Qian, Nino Scherrer, Bertie Vidgen).
- **Published:** November 2023, with the paper "FinanceBench: A New Benchmark for Financial Question Answering".
- **Task:** answer a question about a public company from its filings: information extraction, numerical reasoning and logical reasoning.
- **Corpus:** 368 public filings of U.S. listed companies (10-K annual reports, 10-Q quarterly reports, 8-K current reports and earnings releases), in PDF.
- **Questions:** the open-source sample of 150 questions, drawn from the 10,231 questions of the full benchmark and spread over three types (metrics-generated, domain-relevant and novel-generated).
- **Expected answers:** each question comes with a reference answer and the evidence passage of the filing it relies on.
- **Official scoring:** none; the authors graded the answers by hand. In this report, each answer is graded by the judge described in the methodology.
- **License:** dataset under Creative Commons Attribution-NonCommercial 4.0 (CC BY-NC 4.0).
- **Reference:** [arXiv:2311.11944](https://arxiv.org/abs/2311.11944); github.com/patronus-ai/financebench; huggingface.co/datasets/PatronusAI/financebench.

Methodology

- **Questions:** Random sample of 150 questions, seed 621046. Each question has an expected answer written beforehand by a domain expert.
- **Execution:** each question is asked in a new conversation with the agent, one question at a time (no parallel load), through the same answer pipeline as production conversations.
- **Latency:** time taken by the platform to produce the complete answer, from the question to the end of the answer.
- **Judging:** one call per question to an independent judge model (mistral/mistral-medium-3-5), which receives the question, the expected answer and the delivered answer, and returns OK or NOK with a one-sentence rationale. The prompt given to the judge model is reproduced in Appendix A.
- **Criterion:** OK when the delivered answer contains the essential information of the expected answer, without contradiction or factual error on the point asked. Additional correct information is accepted; wording, length, language and style are not judged. When the question sets no precision, a figure given with a different precision is the same figure if it rounds to the other one (for example 12.2% for 12%).
- **Errors:** a question without delivered answer counts as NOK.
- **Exclusions:** questions without expected answer are not part of the benchmark.
- **Disambiguation:** the original questions are each tied to one source filing, and some do not name its period (for example "this 10-K report"), while the agent searches a knowledge base holding several filings of the same company. When a question names no year, fiscal year or quarter, or no company, ATG appended one fixed sentence giving the period or the company of its source filing: "Added note by ATG for disambiguation: the question is referring to ...", shown with the question below. The question text is otherwise unchanged (11 question(s) in this run).
- **Human review:** a human reviewer can overrule a judge verdict. The question then takes the reviewer's verdict, and the report shows both verdicts with the reviewer's justification (4 question(s) in this run).

Summary

Metric	Value
Questions asked	150
Answers delivered	150
Errors (no answer, counted NOK)	0
OK	139 (92.7%)
NOK	11 (7.3%)
Verdicts changed by human review	4
Tool calls	286 (1.9 per question)
Latency p50 / p90	23.2 s / 54.2 s
Tools	Internal knowledge only

Test conditions

Condition	Value
Start (first question)	2026-10-02 12:20:15 UTC
End (last answer)	2026-10-02 15:31:49 UTC
Duration	3 h 11 min 34 s
AI policy	EU only: EU-based AI providers only
Response mode	Fast
Question selection	Random sample of 150 questions, seed 621046
Judge model	mistral/mistral-medium-3-5

Tools

Internal knowledge only. The agent answered from its internal knowledge only: it had no access to the Internet or to external services.

Tool	Category	Calls during the run
explore_document	Internal knowledge only	102
inspect_images	Internal knowledge only	0
knowledge_base_search	Internal knowledge only	166
query_documents	Internal knowledge only	18
read_conversation_dataset	Internal knowledge only	0

Knowledge base

To answer questions, the agent has access to 368 files.

By format

Format	Documents
PDF	368

By main language

Language	Documents
en	368

Results

Latency

Statistic (answered questions)	Seconds
p50 (median)	23.2 s
p90	54.2 s
Mean	30.9 s
Minimum	8.9 s
Maximum	231.4 s

Questions and answers

Q1. What is the FY2018 capital expenditure amount (in USD millions) for 3M? Give a response to the question by relying on the details shown in the cash flow statement.

Result: OK · Latency: 20.1 s · Tool calls: 1

Expected answer

\$1577.00

Delivered answer

3M's FY2018 capital expenditure was **\$1,577 million** (purchases of property, plant and equipment). This is the PP&E line in the investing section of the consolidated statement of cash flows for the year ended December 31, 2018, as shown in the 2018 10-K cash flow recap (operating cash flow of \$6,439M less the \$1,577M of PP&E purchases yields free cash flow of \$4,862M).

Cited documents: 3M_2018_10K.pdf

Judge rationale: The agent's answer matches the essential figure in the expected answer (\$1,577 million = \$1577.00) and contains no contradictions.

Q2. Assume that you are a public equities analyst. Answer the following question by primarily using information that is shown in the balance sheet: what is the year end FY2018 net PPNE for 3M? Answer in USD billions.

Result: OK · Latency: 18.6 s · Tool calls: 1

Expected answer

\$8.70

Delivered answer

3M's year-end FY2018 net PP&E was **\$8.7 billion** (\$8,738 million).

From the 3M 2018 10-K balance sheet at December 31, 2018: property, plant and equipment of \$24,873 million less accumulated depreciation of \$16,135 million gives net PP&E of \$8,738 million (about \$8.7B).

Cited documents: 3M_2018_10K.pdf

Judge rationale: The agent's answer matches the expected answer's essential figure (\$8.7 billion) and provides additional correct context without contradiction.

Q3. Is 3M a capital-intensive business based on FY2022 data?

Result: OK · Latency: 15.5 s · Tool calls: 1

Expected answer

No, the company is managing its CAPEX and Fixed Assets pretty efficiently, which is evident from below key metrics:

CAPEX/Revenue Ratio: 5.1%

Fixed assets/Total Assets: 20%

Return on Assets= 12.4%

Delivered answer

No, 3M is not a capital-intensive business based on FY2022 data (fiscal year ended December 31, 2022).

- CapEx / revenue: \$1,749M / \$34,229M = 5.11%
- Net fixed assets / total assets: \$9,178M / \$46,455M = 19.76%
- Return on assets: \$5,777M (net income attributable to 3M) / \$46,455M = 12.44%

A high return on assets (12.44%) combined with modest capital spending relative to revenue (5.11%) indicates 3M is not capital-intensive, even though goodwill (\$12.8B) rather than fixed assets makes up most of its asset base.

Cited documents: 3M_2022_10K.pdf

Judge rationale: The agent's answer includes all essential metrics (CAPEX/Revenue ~5.1%, Fixed Assets/Total Assets ~20%, ROA ~12.4%) and aligns with the expected answer's conclusion and figures.

Q4. What drove operating margin change as of FY2022 for 3M? If operating margin is not a useful metric for a company like this, then please state that and explain why.

Result: OK · Latency: 48.0 s · Tool calls: 1

Expected answer

Operating Margin for 3M in FY2022 has decreased by 1.7% primarily due to:

- Decrease in gross Margin
- mostly one-off charges including Combat Arms Earplugs litigation, impairment related to exiting PFAS manufacturing, costs related to exiting Russia and divestiture-related restructuring charges

Delivered answer

Operating margin is a relevant metric for 3M (a diversified industrial manufacturer, not a financial institution), and it **declined** in FY2022. GAAP operating margin fell from 20.8% in 2021 to 19.1% in 2022, a 1.7 percentage-point decline.

Excluding special items, the margin fell from 22.2% to 20.8% (-1.4 pp). The single largest driver of the decline was the **raw material impact** of -2.4 pp, reflecting continued inflationary pressure in raw material and logistics costs from supply-chain, geopolitical and COVID-related disruptions. This was partially offset by **+1.0 pp from total organic growth/productivity and other** (pricing, spending discipline and prior restructuring actions), with foreign exchange and divestitures roughly neutral on a margin basis.

On the GAAP (as-reported) figure, the larger drop to 19.1% also reflects material special items: net costs for significant litigation (-6.7 pp, notably Combat Arms Earplugs and PFAS-related environmental matters), PFAS manufacturing exit costs (-2.4 pp) and Russia exit charges (-0.3 pp), partly offset by the +8.0 pp gain on the Food Safety business divestiture.

Cited documents: 3M_2022_10K.pdf

Judge rationale: The agent's answer includes all essential information from the expected answer (1.7% decline, Combat Arms Earplugs litigation, PFAS manufacturing exit, Russia exit charges) and adds non-contradictory details.

Q5. If we exclude the impact of M&A, which segment has dragged down 3M's overall growth in 2022?

Result: OK · Latency: 16.6 s · Tool calls: 1

Expected answer

The consumer segment shrunk by 0.9% organically.

Delivered answer

Consumer — it was the only 3M segment with negative organic (i.e., ex-M&A) sales growth in 2022, at **-0.9%**, while the other three all grew organically.

Organic sales change by segment, FY2022 (per the 2022 10-K, which strips out acquisitions, divestitures and FX):

Segment	2022 organic sales
Health Care	+3.2%
Transportation & Electronics	+1.2%
Safety & Industrial	+1.0%
Consumer	-0.9%

Management attributed the Consumer decline to softening retail trends — consumers pulling back on discretionary spending and retailers working down inventories — with home improvement the main drag, partly offset by Scotch Blue™ painter's tape, Scotch-Brite™ and Post-it® demand.

So on an organic basis, Consumer is the segment that held back 3M's overall 2022 growth.

Cited documents: 3M_2022_10K.pdf

Judge rationale: The agent's answer correctly identifies the consumer segment as the one with negative organic growth (-0.9%) in 2022, matching the essential information in the expected answer.

Q6. Does 3M have a reasonably healthy liquidity profile based on its quick ratio for Q2 of FY2023? If the quick ratio is not relevant to measure liquidity, please state that and explain why.

Result: OK · Latency: 13.8 s · Tool calls: 1

Expected answer

No. The quick ratio for 3M was 0.96 by Jun'23 close, which needs a bit of an improvement to touch the 1x mark

Delivered answer

No. 3M's quick ratio for Q2 FY2023 (quarter ended June 30, 2023) is below 1.0 on both usual bases, so its liquidity profile is not healthy on that measure. The quick ratio is a relevant metric here — 3M is an industrial manufacturer, not a financial institution, so this standard liquidity test applies.

- Strict basis: (cash \$4,258M + marketable securities \$56M + receivables \$4,947M) / total current liabilities \$10,936M = 9,261 / 10,936 = **0.85**
- Broad basis: (total current assets \$15,754M – inventories \$5,280M) / total current liabilities \$10,936M = 10,474 / 10,936 = **0.96**

Both bases fall below 1.0, so the liquidity profile is not healthy on the quick-ratio test.

Source: 3M Form 10-Q for the quarter ended June 30, 2023, Consolidated Balance Sheet.

Cited documents: 3M_2023Q2_10Q.pdf

Judge rationale: The agent's answer correctly states that 3M's quick ratio is below 1.0 (0.96 on the broad basis) and confirms the quick ratio is relevant for measuring liquidity in this case.

Q7. Which debt securities are registered to trade on a national securities exchange under 3M's name as of Q2 of 2023?

Result: OK · Latency: 18.3 s · Tool calls: 1

Expected answer

Following debt securities registered under 3M's name are listed to trade on the New York Stock Exchange:

- 1.500% Notes due 2026 (Trading Symbol: MMM26)
- 1.750% Notes due 2030 (Trading Symbol: MMM30)
- 1.500% Notes due 2031 (Trading Symbol: MMM31)

Delivered answer

Three debt securities are registered to trade on a national securities exchange under 3M's name as of Q2 2023, all on the New York Stock Exchange: the 1.500% Notes due 2026 (symbol MMM26), the 1.750% Notes due 2030 (MMM30), and the 1.500% Notes due 2031 (MMM31).

This comes from the Section 12(b) securities-registration table in 3M's Form 10-Q for the quarter ended June 30, 2023. The two notes listed in the 2021 10-K — the 0.375% Notes due 2022 (MMM22A) and the 0.950% Notes due 2023 (MMM23) — no longer appear, as they matured by this period. The common stock (MMM) is also registered on both the NYSE and the Chicago Stock Exchange.

Cited documents: 3M_2023Q2_10Q.pdf

Judge rationale: The agent's answer includes all essential debt securities listed in the expected answer with correct details and no contradictions.

Q8. Does 3M maintain a stable trend of dividend distribution? Added note by ATG for disambiguation: the question is referring to Q2 of FY2023.

Result: OK · Latency: 15.2 s · Tool calls: 1

Expected answer

Yes, not only they distribute the dividends on a routine basis, 3M has also been increasing the per share dividend for consecutive 65 years

Delivered answer

Yes — 3M maintains a stable, steadily growing dividend distribution as of Q2 FY2023. The Board declared a second-quarter 2023 dividend of \$1.50 per share (same as the first-quarter 2023 dividend), up 1% from the \$1.49 per share paid each quarter of 2022, and marked the 65th consecutive year of dividend increases. This was the latest step in a long record of uninterrupted quarterly payouts that 3M has maintained since 1916.

Cited documents: 3M_2023Q2_10Q.pdf, 3M_2022_10K.pdf

Judge rationale: The agent's answer confirms the stable and growing dividend trend, including the 65-year consecutive increase, matching the essential information in the expected answer.

Q9. What is the FY2019 fixed asset turnover ratio for Activision Blizzard? Fixed asset turnover ratio is defined as: FY2019 revenue / (average PP&E between FY2018 and FY2019). Round your answer to two decimal places. Base your judgments on the information provided primarily in the statement of income and the statement of financial position.

Result: OK · Latency: 15.0 s · Tool calls: 1

Expected answer

24.26

Delivered answer

24.26

- FY2019 net revenues: \$6,489 million
- Property and equipment, net at Dec 31, 2019: \$253 million; at Dec 31, 2018: \$282 million
- Average PP&E = $(253 + 282) / 2 = \$267.5$ million
- Fixed asset turnover = $6,489 / 267.5 = 24.26$

Cited documents: ACTIVISIONBLIZZARD_2019_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly and includes the correct calculation steps without contradiction.

Q10. What is the FY2017 - FY2019 3 year average of capex as a % of revenue for Activision Blizzard? Answer in units of percents and round to one decimal place. Calculate (or extract) the answer from the statement of income and the cash flow statement.

Result: OK · Latency: 22.9 s · Tool calls: 2

Expected answer

1.9%

Delivered answer

1.9%

Computing capex (purchases of property, plant and equipment) as a % of net revenues for each year:

- FY2017: $155 / 7,017 = 2.21\%$
- FY2018: $131 / 7,500 = 1.75\%$
- FY2019: $116 / 6,489 = 1.79\%$

Average of the three yearly ratios = $(2.21\% + 1.75\% + 1.79\%) / 3 = 1.9\%$. Net revenues from the income statement; capex from the investing section of the cash flow statement.

Cited documents: ACTIVISIONBLIZZARD_2019_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (1.9%) and includes correct calculations without contradictions.

Q11. You are an investment banker and your only resource(s) to answer the following question is (are): the statement of financial position and the cash flow statement. Here's the question: what is the FY2015 operating cash flow ratio for Adobe? Operating cash flow ratio is defined as: cash from operations / total current liabilities. Round your answer to two decimal places.

Result: OK · Latency: 56.4 s · Tool calls: 2

Expected answer

0.66

Delivered answer

0.66 — Adobe's FY2015 (ended November 27, 2015) operating cash flow ratio was 0.66.

Cash from operations was \$1,469.5 million and total current liabilities were \$2,213.6 million, so $1,469,502 / 2,213,556 = 0.66$.

Cited documents: ADOBE_2015_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (0.66) and provides correct supporting calculations.

Q12. What is Adobe's year-over-year change in unadjusted operating income from FY2015 to FY2016 (in units of percents and round to one decimal place)? Give a solution to the question by using the income statement.

Result: OK · Latency: 21.3 s · Tool calls: 1

Expected answer

65.4%

Delivered answer

65.4%

Adobe's operating income rose from \$903.1 million in FY2015 to \$1,493.6 million in FY2016 (as reported on the income statements). The change is \$590.5 million, or $590.5 \div 903.1 = 65.4\%$ year over year.

Cited documents: ADOBE_2016_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value and precision, with no contradictions.

Q13. What is the FY2017 operating cash flow ratio for Adobe? Operating cash flow ratio is defined as: cash from operations / total current liabilities. Round your answer to two decimal places. Please utilize information provided primarily within the balance sheet and the cash flow statement.

Result: OK · Latency: 26.4 s · Tool calls: 2

Expected answer

0.83

Delivered answer

0.83

Cash from operations (FY2017, year ended December 1, 2017): \$2,912,853 thousand

Total current liabilities (December 1, 2017): \$3,527,457 thousand

$\$2,912,853 \div \$3,527,457 = 0.8258 \rightarrow \mathbf{0.83}$

Cited documents: ADOBE_2017_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly, including the correct rounding to two decimal places.

Q14. Does Adobe have an improving operating margin profile as of FY2022? If operating margin is not a useful metric for a company like this, then state that and explain why.

Result: OK · Latency: 21.9 s · Tool calls: 1

Expected answer

No the operating margins of Adobe have recently declined from 36.8% in FY 2021 to 34.6% in FY2022. A drop by 2.2% in a year.

Delivered answer

No — Adobe's operating margin declined as of FY2022, and operating margin is a perfectly relevant metric for a software company like Adobe, so I applied it directly.

FY2022 operating margin: $\$6,098\text{m operating income} \div \$17,606\text{m revenue} = 34.6\%$. FY2021: $\$5,802\text{m} \div \$15,785\text{m} = 36.8\%$. That is a decline of 2.1 percentage points, driven by operating expenses (up 15%) and cost of revenue (up 16%) both growing faster than revenue (up 12%). Operating income still rose \$296m year over year, but not enough to keep pace with the top line, so the margin profile for FY2022 was not improving.

Cited documents: ADOBE_2022_10K.pdf

Judge rationale: The agent's answer correctly states the decline in operating margin from 36.8% in FY2021 to 34.6% in FY2022, matching the essential information in the expected answer, and provides additional correct context without contradiction.

Q15. Does Adobe have an improving Free cashflow conversion as of FY2022?

Result: **OK** · Latency: 15.9 s · Tool calls: 1

Expected answer

Yes, the FCF conversion (using net income as the denominator) for Adobe has improved by ~13% from 143% in 2021 to 156% in 2022

Delivered answer

Yes — Adobe's free cash flow conversion improved in FY2022 (fiscal year ended December 2, 2022), rising from 1.43 in FY2021 to 1.56.

- FY2022: FCF = \$7,838M (operating cash) – \$442M (capex) = \$7,396M; conversion = \$7,396M / \$4,756M net income = 1.56
- FY2021: FCF = \$7,230M – \$348M = \$6,882M; conversion = \$6,882M / \$4,822M = 1.43

The improvement (+0.13) was driven by operating cash flow growing 8% year-over-year while net income dipped slightly, so each dollar of earnings converted into more cash.

Cited documents: ADOBE_2022_10K.pdf

Judge rationale: The agent's answer confirms the improvement in FCF conversion and provides the exact figures (1.43 to 1.56) matching the expected answer, with additional correct details.

Q16. What is the quantity of restructuring costs directly outlined in AES Corporation's income statements for FY2022? If restructuring costs are not explicitly outlined then state 0.

Result: **OK** · Latency: 17.1 s · Tool calls: 1

Expected answer

0

Delivered answer

0.

AES Corporation's Consolidated Statements of Operations for FY2022 (year ended December 31, 2022) do not include a line item for restructuring costs. The income statement line items are: Revenue, Cost of Sales, Operating Margin, General and Administrative Expenses, Interest Expense/Income, Loss on Extinguishment of Debt, Other Expense/Income, Loss on Disposal and Sale of Business Interests, Goodwill Impairment Expense, Asset Impairment Expense, Foreign Currency Transaction Gains (Losses), and Other Non-Operating Expense — no restructuring costs line appears.

This is further confirmed in the MD&A, which states: "Effective January 1, 2021, the Company changed the definitions of Adjusted Operating Margin, Adjusted PTC, and Adjusted EPS to remove the adjustment for costs directly associated with a major restructuring program... There were no such costs in 2020, 2021 or 2022."

Cited documents: AES_2022_10K.pdf

Judge rationale: The agent's answer correctly states 0 and provides supporting evidence that restructuring costs are not explicitly outlined in AES Corporation's FY2022 income statements.

Q17. Roughly how many times has AES Corporation sold its inventory in FY2022? Calculate inventory turnover ratio for the FY2022; if conventional inventory management is not meaningful for the company then state that and explain why.

Result: **NOK** · Latency: 30.0 s · Tool calls: 2

Expected answer

AES has converted inventory 9.5 times in FY 2022.

Delivered answer

About 12 times — the FY2022 inventory turnover is 12.14x.

- Cost of sales (FY2022): \$10,069M
- Inventory: \$1,055M (Dec 31, 2022) and \$604M (Dec 31, 2021) → average = \$829.5M
- Turnover = $\$10,069\text{M} \div \$829.5\text{M} = 12.14\text{x}$

Conventional inventory management is not very meaningful for AES here. AES is a power generation and utility company, and its inventory is fuel and raw materials consumed to produce electricity, plus spare parts — not finished goods held for resale. The fuel cost is largely a pass-through recovered through regulated tariffs, so the ~12x turnover mainly reflects how quickly fuel is burned to generate power (a working-capital/fuel-stockpile dynamic), not how efficiently the company turns over sellable stock.

Cited documents: AES_2022_10K.pdf

Judge rationale: The agent's answer contradicts the expected answer by stating 12.14x instead of 9.5x for the inventory turnover ratio.

Q18. Based on the information provided primarily in the statement of financial position and the statement of income, what is AES's FY2022 return on assets (ROA)? ROA is defined as: FY2022 net income / (average total assets between FY2021 and FY2022). Round your answer to two decimal places.

Result: OK · Latency: 38.5 s · Tool calls: 2

Expected answer

-0.02

Delivered answer

-0.02

Using net income attributable to The AES Corporation (the standard basis for ROA):

- FY2022 net loss attributable to AES: **\$(546) million**
- Total assets at Dec 31, 2022: **\$38,363 million**
- Total assets at Dec 31, 2021: **\$32,963 million**
- Average total assets = $(38,363 + 32,963) / 2 = \mathbf{\$35,663 \text{ million}}$

$\text{ROA} = -546 / 35,663 = -0.0153\dots \approx \mathbf{-0.02}$

Cited documents: AES_2022_10K.pdf

Judge rationale: The agent's answer matches the expected ROA value of -0.02 when rounded to two decimal places, with no contradictions or missing essential information.

Q19. What is Amazon's FY2017 days payable outstanding (DPO)? DPO is defined as: $365 \times (\text{average accounts payable between FY2016 and FY2017}) / (\text{FY2017 COGS} + \text{change in inventory between FY2016 and FY2017})$. Round your answer to two decimal places. Address the question by using the line items and information shown within the balance sheet and the P&L statement.

Result: OK · Latency: 34.5 s · Tool calls: 2

Expected answer

93.86

Delivered answer

93.86 days

Using Amazon's FY2017 10-K (fiscal year ended December 31, 2017):

- Average accounts payable = $(\text{FY2016 AP } \$25,309\text{M} + \text{FY2017 AP } \$34,616\text{M}) / 2 = \$29,962.5\text{M}$
- Change in inventory = $\text{FY2017 } \$16,047\text{M} - \text{FY2016 } \$11,461\text{M} = \$4,586\text{M}$
- Denominator = $\text{FY2017 cost of sales } \$111,934\text{M} + \$4,586\text{M} = \$116,520\text{M}$
- $\text{DPO} = 365 \times \$29,962.5\text{M} / \$116,520\text{M} = \mathbf{93.86}$

All figures are from Amazon's consolidated balance sheets and statements of operations in its 2017 10-K.

Cited documents: AMAZON_2017_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (93.86) and includes the correct calculation steps without contradictions.

Q20. What is Amazon's year-over-year change in revenue from FY2016 to FY2017 (in units of percents and round to one decimal place)? Calculate what was asked by utilizing the line items clearly shown in the statement of income.

Result: OK · Latency: 12.1 s · Tool calls: 1

Expected answer

30.8%

Delivered answer

30.8%.

Amazon's consolidated total net sales were \$177,866M in FY2017 versus \$135,987M in FY2016 (as shown on the statement of income). The change: $(\$177,866 - \$135,987) / \$135,987 = \$41,879 / \$135,987 = 30.8\%$, rounded to one decimal.

Cited documents: AMAZON_2017_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value and precision, with no contradictions or errors.

Q21. By drawing conclusions from the information stated only in the income statement, what is Amazon's FY2019 net income attributable to shareholders (in USD millions)?

Result: OK · Latency: 9.8 s · Tool calls: 1

Expected answer

\$11588.00

Delivered answer

\$11,588 million.

Amazon's FY2019 Consolidated Statements of Operations (year ended December 31, 2019) report net income of \$11,588 million. The income statement shows no non-controlling interest line, so net income equals net income attributable to shareholders.

Cited documents: AMAZON_2019_10K.pdf

Judge rationale: The agent's answer matches the expected net income figure and correctly states it is attributable to shareholders.

Q22. What is Amcor's year end FY2020 net AR (in USD millions)? Address the question by adopting the perspective of a financial analyst who can only use the details shown within the balance sheet.

Result: OK · Latency: 30.5 s · Tool calls: 1

Expected answer

\$1616.00

Delivered answer

Amcor's year-end FY2020 net accounts receivable was **\$1,616 million** (trade receivables, net).

This is the "Trade receivables, net" line on the consolidated balance sheet as of June 30, 2020 (Amcor's fiscal year end), taken directly from the balance sheet.

Cited documents: AMCOR_2021_10K.pdf

Judge rationale: The agent's answer matches the expected figure of \$1616.00 million in USD, with no contradictions or omissions.

Q23. What was the key agenda of the AMCOR's 8k filing dated 1st July 2022?

Result: OK · Latency: 16.2 s · Tool calls: 1

Expected answer

Amcor Finance (USA), Inc. and Amcor Flexibles North America, Inc., entered into supplemental indentures relating to Guaranteed Senior Notes due 2026 and 2028. This involved the substitution of the Substitute Issuer (Amcor Flexibles North America) for the Former Issuer (Amcor Finance) and the assumption of covenants under the indentures. (In essence a novation agreement)

Delivered answer

The key item in Amcor's 8-K filed July 1, 2022 was the substitution of the issuer of its guaranteed senior notes. On June 30, 2022, Amcor Finance (USA), Inc. (the Former Issuer) and Amcor Flexibles North America, Inc. (the Substitute Issuer) entered into a First and a Second Supplemental Indenture with Deutsche Bank Trust Company Americas, as trustee, effecting the substitute issuer's assumption of the covenants under the governing indentures for the 3.625% Guaranteed Senior Notes due 2026 and the 4.500% Guaranteed Senior Notes due 2028. The filing was made under Item 8.01 (Other Events).

Cited documents: AMCOR_2022_8K_dated-2022-07-01.pdf

Judge rationale: The agent's answer contains all essential information from the expected answer, including the substitution of issuers and the assumption of covenants for the specified notes, with no contradictions.

Q24. Has AMCOR's quick ratio improved or declined between FY2023 and FY2022? If the quick ratio is not something that a financial analyst would ask about a company like this, then state that and explain why.

Result: OK · Latency: 54.2 s · Tool calls: 3

Expected answer

The quick ratio has slightly improved from 0.67 times to 0.69 times between FY 2023 and FY 2022.(3.4% jump)

Delivered answer

Improved.

The quick ratio rose from FY2022 to FY2023 on both usual bases (fiscal years ended June 30):

- **Strict** (cash + receivables) / current liabilities:
 - FY2022: $(\$775 + \$1,935) / \$5,103 = \mathbf{0.53}$
 - FY2023: $(\$689 + \$1,875) / \$4,476 = \mathbf{0.57}$
- **Broad** (current assets – inventories) / current liabilities:
 - FY2022: $(\$5,853 - \$2,439) / \$5,103 = \mathbf{0.67}$
 - FY2023: $(\$5,308 - \$2,213) / \$4,476 = \mathbf{0.69}$

The improvement is driven mainly by the drop in total current liabilities (\$5,103M → \$4,476M), while quick assets fell only modestly. Both bases remain below 1.0, so the liquidity profile is still tight, but the direction is positive.

The quick ratio is a standard and relevant metric for Amcor: it is a manufacturing/packaging company, not a financial institution, so liquidity ratios like the quick ratio are appropriate for assessing its short-term solvency.

Cited documents: AMCOR_2023_10K.pdf

Judge rationale: The agent's answer correctly states the quick ratio improved and provides the exact figures (0.67 to 0.69) matching the expected answer, with additional valid context.

Q25. What are major acquisitions that AMCOR has done in FY2023, FY2022 and FY2021?

Result: **OK** · Latency: 78.9 s · Tool calls: 7

Expected answer

Amcor completed these acquisitions during FY2023:

-100% equity interest of a flexibles manufacturing company in the Czech Republic

- 100% equity interest in a medical device packaging manufacturing site in Shanghai, China.
- acquisition of a New Zealand-based leading manufacturer of state-of-the-art, automated protein packaging machines.

Delivered answer

Here's what Amcor's 10-K filings show for each fiscal year (fiscal year ends June 30):

FY2023 (ended June 30, 2023) — three acquisitions, all in the Flexibles segment:

1. **Czech Republic flexible packaging plant** (August 1, 2022) — \$59 million purchase consideration (including \$5 million deferred, paid in Q1 FY2024). Acquired identifiable net assets of \$36 million and goodwill of \$23 million.
2. **Medical device packaging site in Shanghai, China** (March 17, 2023) — \$60 million purchase consideration (including \$20 million contingent consideration payable over three years). Acquired identifiable net assets of \$21 million and goodwill of \$39 million.
3. **New Zealand-based manufacturer of automated protein packaging machines** (May 31, 2023) — \$45 million purchase consideration (including \$13 million contingent consideration payable over two years). Acquired identifiable net assets of \$9 million and goodwill of \$36 million.

Pro forma information was not presented for any of the three, as their effect on consolidated financial statements was not material.

FY2022 (ended June 30, 2022) — no acquisitions. The acquisitions and divestitures note for that year only describes the disposal of non-core assets in the Flexibles segment (a \$10 million write-down to fair value less cost to sell).

FY2021 (ended June 30, 2021) — no acquisitions. The note for that year only covers disposals: a non-core European hospital supplies business, two non-core businesses in India and Argentina, and the sale of the AMVIG equity investment.

So in summary, FY2023 was the only one of the three years with acquisitions — three bolt-on deals totaling roughly \$164 million in purchase consideration, all assigned to the Flexibles segment. FY2022 and FY2021 were divestiture-only years.

Cited documents: AMCOR_2021_10K.pdf, AMCOR_2023_10K.pdf

Judge rationale: The agent's answer includes all essential acquisitions listed in the expected answer for FY2023 and correctly states no acquisitions for FY2022 and FY2021.

Q26. What industry does AMCOR primarily operate in? Added note by ATG for disambiguation: the question is referring to FY2023.

Result: OK · Latency: 12.6 s · Tool calls: 1

Expected answer

Amcor is a global leader in packaging production for various use cases.

Delivered answer

Packaging.

Amcor primarily operates in the **packaging** industry. Per its FY2023 10-K (fiscal year ended June 30, 2023), the company describes itself as "a global leader in developing and producing responsible packaging for food, beverage, pharmaceutical, medical, home and personal-care, and other consumer goods end markets," with its business organized into two reportable segments: **Flexibles** (~76% of consolidated net sales) and **Rigid Packaging** (~24%).

In short, Amcor makes flexible packaging, rigid packaging containers, specialty cartons, and closures — all sold to customers across Europe, North America, Latin America, and Asia Pacific.

Cited documents: AMCOR_2023_10K.pdf

Judge rationale: The agent's answer correctly identifies the packaging industry as Amcor's primary industry, matching the essential information in the expected answer.

Q27. Does AMCOR have an improving gross margin profile as of FY2023? If gross margin is not a useful metric for a company like this, then state that and explain why.

Result: OK · Latency: 17.1 s · Tool calls: 1

Expected answer

No. For AMCOR there has been a slight decline in gross margins by 0.8%.

Delivered answer

No — Amcor's gross margin declined in FY2023 (fiscal year ended June 30, 2023), so the profile is not improving.

Gross margin = gross profit ÷ net sales (GAAP, as reported):

- FY2022: \$2,820M ÷ \$14,544M = 19.39%
- FY2023: \$2,725M ÷ \$14,694M = 18.55%
- Change: -0.84 percentage points (≈ -84 bps)

Gross margin is a perfectly valid metric here — Amcor is a packaging manufacturer (a real-goods producer, not a financial institution), so cost of sales and gross profit are meaningful core lines. The decline reflects the pass-through of higher raw-material costs inflating sales dollars faster than the associated cost recovery, plus lower volumes and unfavorable mix, as management noted in the FY2023 results.

Cited documents: AMCOR_2023Q4_EARNINGS.pdf

Judge rationale: The agent's answer correctly states that AMCOR's gross margin declined in FY2023 and provides supporting figures that align with the expected answer's 0.8% decline.

Q28. What is the nature & purpose of AMCOR's restructuring liability as of Q2 of FY2023 close?

Result: **OK** · Latency: 61.3 s · Tool calls: 5

Expected answer

87% of the total restructuring liability is related Employee liabilities.

Delivered answer

Amcor's restructuring liability at the Q2 FY2023 close (December 31, 2022) was **\$93 million**, a provision to fund the company's remaining restructuring activities, which in this period primarily related to the Russia-Ukraine conflict (group-wide initiatives to offset earnings lost from the sale of the Russian business).

By nature it breaks down almost entirely into employee-related costs, with small residual fixed-asset and other costs:

- Employee costs: \$81M
- Fixed asset related costs: \$2M
- Other costs: \$10M

The liability fell from \$118M at June 30, 2022 to \$93M, driven by cash paid (\$25M) and reversal of unused amounts (\$2M). The majority is expected to be paid within the next twelve months, and the accrual is recorded under other current and non-current liabilities.

Cited documents: AMCOR_2023Q2_10Q.pdf

Judge rationale: The agent's answer includes the essential information that the restructuring liability is primarily employee-related (81M out of 93M, which is ~87%).

Q29. What Was AMCOR's Adjusted Non GAAP EBITDA for FY 2023

Result: **OK** · Latency: 22.7 s · Tool calls: 1

Expected answer

AMCOR's Adj. EBITDA was \$2,018mn in FY 2023

Delivered answer

Amcor's adjusted non-GAAP EBITDA for FY2023 (year ended June 30, 2023) was **\$2,018 million**.

From the reconciliation in the fiscal 2023 earnings release: GAAP EBITDA of \$2,080 million (net income \$1,048M + non-controlling interests \$10M + tax \$193M + interest \$259M + D&A \$569M), adjusted for hyperinflation (\$24M), Russia-Ukraine conflict impacts (-\$90M), pension settlements (\$5M), property/other losses (\$2M) and other (-\$3M), giving adjusted EBITDA of \$2,018 million.

That's down ~5% from FY2022's \$2,117 million, but up 1% on a comparable constant currency basis.

Cited documents: AMCOR_2023Q4_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential figure (\$2,018 million) and does not contradict the expected answer.

Q30. How much was the Real change in Sales for AMCOR in FY 2023 vs FY 2022, if we exclude the impact of FX movement, passthrough costs and one-off items?

Result: OK · Latency: 20.5 s · Tool calls: 1

Expected answer

The Real Growth was flat in FY 2023 vs FY 2022.

Delivered answer

The Real change in sales was essentially flat — \$7 million, or 0% — after stripping out FX, raw material passthrough and one-off items.

Amcors FY2023 (ended June 30, 2023) net sales rose 1% reported to \$14,694M from \$14,544M. Excluding the pass-through of higher raw material costs (~\$776M), negative currency (~\$426M), and the impact of disposed/ceased operations (~\$207M), the remaining organic variation was a \$7M increase, or 0% — price/mix +3% offset by volumes –3%. Stated another way, net sales on a comparable constant currency basis were in line with the prior year.

Cited documents: AMCOR_2023Q4_EARNINGS.pdf, AMCOR_2023_10K.pdf

Judge rationale: The agent's answer correctly states that the real change in sales was flat (0%) after excluding FX, passthrough costs, and one-off items, matching the essential information in the expected answer.

Q31. Answer the following question as if you are an equity research analyst and have lost internet connection so you do not have access to financial metric providers. According to the details clearly outlined within the P&L statement and the statement of cash flows, what is the FY2015 depreciation and amortization (D&A from cash flow statement) % margin for AMD?

Result: OK · Latency: 14.0 s · Tool calls: 1

Expected answer

4.2%

Delivered answer

4.18%

- D&A (cash flow statement, FY2015): \$167M
- Net revenue (FY2015): \$3,991M
- D&A margin = $167 / 3,991 = 0.0418 \rightarrow 4.18\%$

(AMD's fiscal 2015 ended December 26, 2015.)

Cited documents: AMD_2015_10K.pdf

Judge rationale: The agent's answer (4.18%) matches the expected answer (4.2%) when rounded to one decimal place.

Q32. Does AMD have a reasonably healthy liquidity profile based on its quick ratio for FY22? If the quick ratio is not relevant to measure liquidity, please state that and explain why.

Result: OK · Latency: 19.0 s · Tool calls: 1

Expected answer

Yes. The quick ratio is 1.57, calculated as (cash and cash equivalents+Short term investments+Accounts receivable, net +receivables from related parties)/ (current liabilities).

Delivered answer

Yes — AMD has a reasonably healthy liquidity profile for FY2022 (fiscal year ended December 31, 2022).

Quick ratio (strict) = (cash + short-term investments + receivables) / total current liabilities
= (\$4,835M + \$1,020M + \$4,126M) / \$6,369M = \$9,981M / \$6,369M = **1.57**

Quick ratio (broad) = (total current assets – inventories) / total current liabilities
= (\$15,019M – \$3,771M) / \$6,369M = \$11,248M / \$6,369M = **1.77**

Both bases sit comfortably above 1.0, so the liquidity profile is healthy. AMD's quick assets (cash, short-term investments and receivables) cover its current liabilities roughly 1.6–1.8 times over, even before counting inventories.

Cited documents: AMD_2022_10K.pdf

Judge rationale: The agent's answer includes the essential information (quick ratio of 1.57) and does not contradict the expected answer.

Q33. What are the major products and services that AMD sells as of FY22?

Result: **OK** · Latency: 32.6 s · Tool calls: 2

Expected answer

AMD sells server microprocessors (CPUs) and graphics processing units (GPUs), data processing units (DPUs), Field Programmable Gate Arrays (FPGAs), and Adaptive System-on-Chip (SoC) products for data centers; CPUs, accelerated processing units (APUs) that integrate CPUs and GPUs, and chipsets for desktop and notebook personal computers; discrete GPUs, and semi-custom SoC products and development services; and embedded CPUs, GPUs, APUs, FPGAs, and Adaptive SoC products.

Delivered answer

As of FY22 (fiscal year ended December 31, 2022), AMD organized its product portfolio into four reporting segments, each covering a distinct set of products and services:

1. Data Center segment

- **Server CPUs:** AMD EPYC processors (9004, 7003, and 7002 Series)
- **Data center GPUs:** AMD Instinct MI200 Series (2nd Gen CDNA architecture) and Radeon PRO V-series for visual cloud
- **Data Processing Units (DPUs):** P4 programmable AMD Pensando DPUs
- **FPGAs & Adaptive SoCs:** Virtex, Kintex, Zynq, Zynq MPSoC, Versal Adaptive SoC products, and Alveo accelerator cards

2. Client segment (desktop & notebook PCs)

- **Desktop CPUs:** AMD Ryzen (7000, 5000 Series) and AMD Athlon series
- **Notebook CPUs:** AMD Ryzen mobile (7000, 6000, 5000 Series) and AMD Athlon mobile
- **Commercial CPUs:** AMD Ryzen PRO, Ryzen Threadripper PRO, Athlon PRO, and AMD PRO A-Series
- **Chipsets** for desktop and notebook platforms
- **Discrete GPUs:** AMD Radeon RX 7900 (RDNA 3) and RX 6000 series (RDNA 2)
- **Professional/workstation graphics:** AMD Radeon PRO W6000 series

3. Gaming segment

- **Discrete GPUs:** AMD Radeon RX 7900 XTX/XT (RDNA 3), RX 6000 series (RDNA 2)
- **Semi-custom SoC products & development services:** processors powering Sony PlayStation 5, Microsoft Xbox Series S/X, and the Steam Deck handheld

4. Embedded segment

- **Embedded CPUs/APUs/GPUs:** AMD Embedded EPYC, Ryzen V-Series, Ryzen R-Series, and low-power Athlon processors
- **FPGAs & Adaptive SoCs (Xilinx):** Virtex, Kintex, Zynq, Versal, Alveo, Artix, CoolRunner, and Kria robotics products
- **CPLDs, DSPs, and other programmable logic** for automotive, industrial, aerospace, healthcare, and 5G applications

In addition to hardware, AMD provides **semi-custom design and development services** (customers pay non-recurring engineering fees) and **software** (ROCm, FidelityFX, FreeSync, AMD Software: Adrenalin Edition, and design automation tools).

Cited documents: AMD_2022_10K.pdf

Judge rationale: The agent's answer includes all essential product categories (CPUs, GPUs, DPUs, FPGAs, Adaptive SoCs, APUs, chipsets, discrete GPUs, semi-custom SoCs, and embedded products) mentioned in the

expected answer, with no contradictions or factual errors.

Q34. What drove revenue change as of the FY22 for AMD?

Result: OK · Latency: 33.5 s · Tool calls: 2

Expected answer

In 2022, AMD reported Higher sales of their EPYC server processors, higher semi-custom product sales, and the inclusion of Xilinx embedded product sales

Delivered answer

Revenue increased 44% in FY2022 to \$23.6B (from \$16.4B in FY2021), a \$7.2B rise. The growth was driven by the Embedded and Data Center segments, partly offset by a decline in Client.

- **Embedded:** +\$4.3B to \$4.6B (from \$246M), driven almost entirely by the inclusion of record Xilinx revenue after the February 2022 acquisition.
- **Data Center:** +\$2.3B to \$6.0B (from \$3.7B), up 64%, led by adoption of AMD EPYC server processors.
- **Gaming:** +\$1.2B to \$6.8B (from \$5.6B), up 21%, on semi-custom console demand and discrete GPUs.
- **Client:** −\$0.7B to \$6.2B (from \$6.9B), down 10%, due to soft PC market conditions in the second half of 2022.

Cited documents: AMD_2022_annualreport.pdf

Judge rationale: The agent's answer includes all essential drivers of revenue change (Xilinx embedded product sales, EPYC server processors, and semi-custom product sales) and does not contradict the expected answer.

Q35. What drove operating margin change as of the FY22 for AMD? If operating margin is not a useful metric for a company like this, then please state that and explain why.

Result: OK · Latency: 24.3 s · Tool calls: 1

Expected answer

The decrease in AMD's operating income was primarily driven by amortization of intangible assets associated with the Xilinx acquisition

Delivered answer

Operating margin declined sharply as of FY2022 (fiscal year ended December 31, 2022), falling from 22.2% to 5.4%, a drop of about 16.8 percentage points.

- FY2022: operating income \$1,264M ÷ net revenue \$23,601M = **5.36%**
- FY2021: operating income \$3,648M ÷ net revenue \$16,434M = **22.20%**
- Change: 5.36% – 22.20% = **–16.84 pp**

Driver: the decline was driven almost entirely by amortization of acquisition-related intangible assets from the Xilinx acquisition (completed February 2022), a one-off, non-cash charge of \$3,548M in 2022 — \$1,448M booked in cost of sales plus \$2,100M as a separate operating expense line. Management explicitly attributes both the gross-margin compression (48% → 45%) and the operating-income drop (\$3,648M → \$1,264M) to this Xilinx intangible amortization, which also pulled down the reported operating margin.

On usefulness: operating margin is a standard and useful metric for AMD — it is a semiconductor company, not a financial institution, so operating margin applies normally. The caveat is that the *reported* 2022 figure is distorted by that one-off, non-cash Xilinx amortization, so it understates the underlying operating performance of the business that year.

Cited documents: AMD_2022_10K.pdf

Judge rationale: The agent's answer correctly identifies the primary driver of AMD's operating margin change as amortization of intangible assets from the Xilinx acquisition, matching the essential information in the expected answer.

Q36. Among operations, investing, and financing activities, which brought in the most (or lost the least) cash flow for AMD in FY22?

Result: OK · Latency: 41.0 s · Tool calls: 3

Expected answer

In 2022, AMD brought in the most cashflow from Operations

Delivered answer

Operating activities brought in the most cash at \$3.6 billion (\$3,565 million) in FY22, the largest positive contributor.

- Operating: +\$3,565 million (net income of \$1,320 million plus \$4.1 billion of non-cash adjustments, partly offset by ~\$1.8 billion of working-capital outflows)
- Investing: +\$1,999 million (driven by \$4,310 million of short-term investment maturities and \$2,366 million of cash received from the Xilinx acquisition)
- Financing: -\$3,264 million (use, mostly \$3,702 million of common stock repurchases)

So while both operating and investing generated cash, operations provided the most.

Judge rationale: The agent's answer correctly identifies operating activities as the largest positive cash flow contributor for AMD in FY22, matching the essential information in the expected answer.

Q37. From FY21 to FY22, excluding Embedded, in which AMD reporting segment did sales proportionally increase the most?

Result: OK · Latency: 15.6 s · Tool calls: 1

Expected answer

Data Center

Delivered answer

Data Center — it grew the most proportionally of the three non-Embedded segments, up ~64% from FY21 to FY22.

- Data Center: \$3,694M → \$6,043M = $(6,043 - 3,694) / 3,694 = +63.6\%$
- Gaming: \$5,607M → \$6,805M = $(6,805 - 5,607) / 5,607 = +21.4\%$
- Client: \$6,887M → \$6,201M = -10.0% (declined)

The increase was driven primarily by higher sales of AMD's EPYC server processors.

Cited documents: AMD_2022_10K.pdf

Judge rationale: The agent correctly identifies 'Data Center' as the segment with the highest proportional sales increase, matching the expected answer.

Q38. Did AMD report customer concentration in FY22?

Result: OK · Latency: 35.8 s · Tool calls: 4

Expected answer

Yes, one customer accounted for 16% of consolidated net revenue

Delivered answer

Yes — AMD reported customer concentration in its FY22 (fiscal year ended December 31, 2022) 10-K.

In Note 4 (Segment Reporting), AMD disclosed sales to major customers each accounting for at least 10% of consolidated net revenue: Customer A represented **16%** of FY22 net revenue, while no other single customer reached 10% (Customer B was **less than 10%** in 2022, down from 11% in 2021). Sales to Customers A and B came from the Gaming and Client segments, respectively. A separate Note 10 also noted credit-risk concentration, with one customer representing about 18% of the consolidated accounts receivable balance at December 31, 2022.

Cited documents: AMD_2022_10K.pdf

Judge rationale: The agent's answer includes the essential information that one customer accounted for 16% of consolidated net revenue in FY22, matching the expected answer.

Q39. Which debt securities are registered to trade on a national securities exchange under American Express' name as of 2022?

Result: OK · Latency: 13.1 s · Tool calls: 1

Expected answer

There are none

Delivered answer

None. As of the 2022 Form 10-K (fiscal year ended December 31, 2022), American Express had no debt securities registered to trade on a national securities exchange under its name.

The cover page lists only one class of securities registered under Section 12(b) of the Exchange Act — Common Shares (par value \$0.20), trading symbol AXP on the New York Stock Exchange — and states "Securities registered pursuant to section 12(g) of the Act: None."

So while the company issued substantial long-term debt in 2022 (fixed-rate, floating-rate and fixed-to-floating senior and subordinated notes, plus asset-backed certificates), none of those debt securities were registered to trade on a national securities exchange under the company's name.

Cited documents: AMERICANEXPRESS_2022_10K.pdf

Judge rationale: The agent's answer correctly states that there are no debt securities registered to trade on a national securities exchange under American Express' name as of 2022, matching the expected answer.

Q40. What are the geographies that American Express primarily operates in as of 2022?

Result: OK · Latency: 15.9 s · Tool calls: 1

Expected answer

United States, EMEA, APAC, and LACC

Delivered answer

American Express primarily operates in the United States and four international regions: EMEA, APAC, LACC, and an "Other Unallocated" segment, based on its geographic operations by revenue.

In 2022, total revenues net of interest expense were \$52,862 million: the United States contributed \$41,396 million (about 78%), EMEA (Europe, Middle East and Africa) \$4,871 million, APAC (Asia Pacific, Australia and New Zealand) \$3,835 million, and LACC (Latin America, Canada and the Caribbean) \$2,917 million, with \$157 million of negative unallocated revenue. The U.S. is by far the dominant geography, with international regions making up the remainder.

Cited documents: AMERICANEXPRESS_2022_10K.pdf

Judge rationale: The agent's answer includes all essential geographies (United States, EMEA, APAC, LACC) and does not contradict the expected answer.

Q41. Does AMEX have an improving operating margin profile as of 2022? If operating margin is not a useful metric for a company like this, then state that and explain why.

Result: OK · Latency: 14.0 s · Tool calls: 1

Expected answer

Performance is not measured through operating margin

Delivered answer

Operating margin is not a useful metric for American Express.

American Express is a bank holding company (it and its principal operating subsidiary, American Express Travel Related Services Company, are bank holding companies under the Bank Holding Company Act of 1956), so interest income and credit losses are part of its core operations rather than "other" items — which is exactly what an operating margin tries to isolate, making the metric not meaningful for this kind of company.

For a payments company like AMEX, performance is better read through its own reported lines — total revenues net of interest expense, pretax income, and net income — rather than a conventional operating margin.

Cited documents: AMERICANEXPRESS_2022_10K.pdf

Judge rationale: The agent's answer correctly states that operating margin is not a useful metric for AMEX and provides a valid explanation, aligning with the essential information in the expected answer.

Q42. What drove gross margin change as of the FY2022 for American Express? If gross margin is not a useful metric for a company like this, then please state that and explain why.

Result: OK · Latency: 11.4 s · Tool calls: 0

Expected answer

Performance is not measured through gross margin

Delivered answer

Gross margin is not a useful metric for American Express because it is a financial services institution: its revenues are interest and fees, and its credit losses and funding costs are core to how its performance is measured, not a separate "cost of goods sold." There is therefore no meaningful gross margin figure to analyze or attribute a change to for FY2022, and the standard profitability lens for this kind of company is provisions for credit losses, net interest margin, and operating efficiency rather than gross margin.

Judge rationale: The agent's answer correctly states that gross margin is not a useful metric for American Express and provides a valid explanation aligning with the expected answer.

Q43. How much has the effective tax rate of American Express changed between FY2021 and FY2022?

Result: OK · Latency: 56.6 s · Tool calls: 2

Expected answer

The effective tax rate for American Express has changed/dropped from 24.6% in FY 2021 to 21.6% in FY 2022.

Delivered answer

The effective tax rate of American Express **declined by 3.0 percentage points**, from 24.6% in FY2021 to 21.6% in FY2022.

- FY2021: income tax expense \$2,629M ÷ pre-tax income \$10,689M = 24.6%
- FY2022: income tax expense \$2,071M ÷ pre-tax income \$9,585M = 21.6%
- Change: 21.6% – 24.6% = **–3.0 percentage points**

The main driver of the lower 2022 rate was a \$2.1 billion tax settlement and lapse of statute of limitations benefit (–2.1 pp), partially offset by higher state and local taxes (+3.1 pp vs. +3.0 pp in 2021).

Cited documents: AMERICANEXPRESS_2022_10K.pdf

Judge rationale: The agent's answer includes the essential change (from 24.6% to 21.6%) and correctly states the magnitude of the change (–3.0 percentage points), with no contradictions.

Q44. What was the largest liability in American Express's Balance Sheet in 2022?

Result: **OK** · Latency: 34.9 s · Tool calls: 2

Expected answer

Customer deposits

Delivered answer

Customer deposits — \$110,239 million (about \$110.2 billion) — was the largest liability on American Express's consolidated balance sheet as of December 31, 2022.

For context, the other liability line items on that balance sheet were:

- Long-term debt: \$42,573 million
- Other liabilities: \$37,350 million
- Accounts payable and other liabilities: \$12,133 million
- Short-term borrowings: \$1,348 million

Customer deposits (\$110.2 billion) exceeded the next-largest item, long-term debt (\$42.6 billion), by roughly \$67.7 billion.

Cited documents: AMERICANEXPRESS_2022_10K.pdf

Judge rationale: The agent's answer correctly identifies 'Customer deposits' as the largest liability, matching the essential information in the expected answer.

Q45. Was American Express able to retain card members during 2022?

Result: **OK** · Latency: 55.0 s · Tool calls: 3

Expected answer

Yes

Delivered answer

Yes — American Express retained its card members through 2022, and the company explicitly called out that "Card Member retention remained high" for the year. Proprietary cards-in-force rose to 76.7 million at December 31, 2022, from 71.4 million a year earlier (a ~7% gain), and total cards-in-force increased to 133.3 million from 121.7 million, so the membership base grew, not merely held steady. Management attributed the 17% year-over-year rise in net card fees to record new card acquisitions alongside that high retention, reinforcing that existing card members were kept even as new ones were added.

Cited documents: AMERICANEXPRESS_2022_10K.pdf

Judge rationale: The agent's answer includes the essential confirmation ('Yes') and provides additional supporting details without contradiction.

Q46. How much (in USD billions) did American Water Works pay out in cash dividends for FY2020? Compute or extract the answer by primarily using the details outlined in the statement of cash flows.

Result: OK · Latency: 27.5 s · Tool calls: 1

Expected answer

\$0.40

Delivered answer

American Water Works paid out **\$0.389 billion** in cash dividends in FY2020.

From the consolidated statement of cash flows for the year ended December 31, 2020, the financing-activities line "Dividends paid" shows \$389 million.

Cited documents: AMERICANWATERWORKS_2020_10K.pdf

Judge rationale: The agent's answer (\$0.389 billion) matches the expected answer (\$0.40 billion) when rounded to the same precision.

Q47. Basing your judgments off of the cash flow statement and the income statement, what is American Water Works's FY2021 unadjusted operating income + depreciation and amortization from the cash flow statement (unadjusted EBITDA) in USD millions?

Result: OK · Latency: 12.7 s · Tool calls: 1

Expected answer

\$1832.00

Delivered answer

\$1,832 million.

From the FY2021 10-K:

- Operating income (income statement): \$1,196M
- Depreciation and amortization (cash flow statement): \$636M

$\$1,196\text{M} + \$636\text{M} = \mathbf{\$1,832\text{M}}$

Cited documents: AMERICANWATERWORKS_2021_10K.pdf

Judge rationale: The agent's answer matches the expected answer in value and units, with no contradictions or missing essential information.

Q48. Does American Water Works have positive working capital based on FY2022 data? If working capital is not a useful or relevant metric for this company, then please state that and explain why.

Result: OK · Latency: 33.5 s · Tool calls: 3

Expected answer

No, American Water Works had negative working capital of -\$1561M in FY 2022.

Delivered answer

No — American Water Works does not have positive working capital for FY2022 (fiscal year ended December 31, 2022).

Working capital is a standard, relevant metric for a water/wastewater utility like AWK (it is not a financial institution, so the metric applies).

- Total current assets: \$1,250 million
- Total current liabilities: \$2,811 million
- Working capital = \$1,250M – \$2,811M = **–\$1,561 million** (negative)

The negative position is driven largely by short-term debt (\$1,175M) and the current portion of long-term debt (\$281M) sitting in current liabilities, which together exceed the entire current-asset base.

Cited documents: AMERICANWATERWORKS_2022_10K.pdf

Judge rationale: The agent's answer correctly states that American Water Works had negative working capital of -\$1,561 million in FY 2022, matching the essential information in the expected answer.

Q49. In agreement with the information outlined in the income statement, what is the FY2015 - FY2017 3 year average net profit margin (as a %) for Best Buy? Answer in units of percents and round to one decimal place.

Result: OK · Latency: 37.5 s · Tool calls: 2

Expected answer

2.8%

Delivered answer

2.8%

Net profit margin = net earnings attributable to Best Buy shareholders ÷ revenue, per the income statement, for each fiscal year (ending late January):

- FY2015: $1,233 / 40,339 = 3.06\%$
- FY2016: $897 / 39,528 = 2.27\%$
- FY2017: $1,228 / 39,403 = 3.12\%$

Average of the three annual margins = $(3.06 + 2.27 + 3.12) / 3 = 2.8\%$ (equivalently, total net income 3,358 / total revenue 119,270 = 2.8%).

Cited documents: BESTBUY_2015_10K.pdf, BESTBUY_2017_10K.pdf, BESTBUY_2016_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value and precision.

Q50. What is the year end FY2019 total amount of inventories for Best Buy? Answer in USD millions. Base your judgments on the information provided primarily in the balance sheet.

Result: OK · Latency: 10.4 s · Tool calls: 1

Expected answer

\$5409.00

Delivered answer

\$5,409 million.

Best Buy's fiscal year 2019 ended February 2, 2019. On that year's balance sheet, merchandise inventories were \$5,409 million, up from \$5,209 million at the prior year-end.

Cited documents: BESTBUY_2019_10K.pdf

Judge rationale: The agent's answer matches the expected answer in value and units, with no contradictions or omissions.

Q51. Are Best Buy's gross margins historically consistent (not fluctuating more than roughly 2% each year)? If gross margins are not a relevant metric for a company like this, then please state that and explain why. Added note by ATG for disambiguation: the question is referring to FY2023.

Result: OK · Latency: 45.0 s · Tool calls: 3

Expected answer

Yes, the margins have been consistent, there has been a minor decline of 1.1% in gross margins between FY2022 and FY2023.

Delivered answer

Yes — Best Buy's consolidated gross margin has been historically consistent, with no year-on-year move exceeding about 2 percentage points. Gross margin is a fully relevant metric here: Best Buy is a retailer, not a financial institution, so gross profit is a core operating measure it reports by segment and in total.

For the FY2023 trend (fiscal year ended January 28, 2023), gross margin was 21.4% versus 22.5% in FY2022, a decline of about 1.1 percentage points; the year before it was effectively flat (22.4% in FY2021, 22.5% in FY2022, +0.1 pp), and FY2020 was 23.0%. Every annual change falls well within the 2% tolerance.

The FY2023 decline was driven largely by lower product margin rates in the Domestic segment, partially offset by favorable services margins and higher profit-sharing revenue from the private-label/co-branded credit card arrangement; the International segment also saw lower product margins and a mix shift away from higher-margin services revenue.

Cited documents: BESTBUY_2023Q4_EARNINGS.pdf, BESTBUY_2023_10K.pdf

Judge rationale: The agent's answer confirms the consistency of gross margins and specifies the 1.1% decline between FY2022 and FY2023, matching the essential information in the expected answer.

Q52. What are major acquisitions that Best Buy has done in FY2023, FY2022 and FY2021?

Result: **OK** · Latency: 36.1 s · Tool calls: 3

Expected answer

Best Buy closed two acquisitions, both these companies were already partially owned by Best Buy, but Best Buy acquired all outstanding shares of these two companies during FY 2022: (1) Current Health Ltd and (2) Two Peaks, LLC d/b/a Yardbird Furniture

Delivered answer

Best Buy completed acquisitions only in FY2022; FY2021 and FY2023 had none. (Best Buy's fiscal year ends in late January, so FY2021, FY2022 and FY2023 refer to the years ended January 30, 2021, January 29, 2022, and January 28, 2023.)

- **FY2022 (ended Jan. 29, 2022)** — two acquisitions, both closed in the fourth quarter, for total net cash of \$468 million:
 - **Current Health Ltd.** — care-at-home technology platform, \$389 million net cash (Nov. 2, 2021), with \$351 million of goodwill assigned to the Best Buy Health reporting unit.
 - **Two Peaks, LLC d/b/a Yardbird Furniture** — direct-to-consumer outdoor furniture, \$79 million net cash (Nov. 4, 2021), with \$47 million of goodwill.

The \$468 million matches the "Acquisitions, net of cash acquired" line in the cash flow statement for fiscal 2022.

- **FY2023 (ended Jan. 28, 2023)** — no acquisitions (the acquisitions line in the cash flow statement is zero).
- **FY2021 (ended Jan. 30, 2021)** — no acquisitions. The prior deals (GreatCall in 2018, CST and BioSensics in 2020) were completed before this period.

So the answer to "major acquisitions" across those three fiscal years is really just the FY2022 pair: Current Health and Yardbird. Source: Best Buy FY2023 Form 10-K, Note 2, Acquisitions

Judge rationale: The agent's answer correctly identifies the two FY2022 acquisitions (Current Health Ltd and Yardbird Furniture) and confirms no acquisitions in FY2021 and FY2023, matching the essential information in the expected answer.

Q53. Among operations, investing, and financing activities, which brought in the most (or lost the least) cash flow for Best Buy in FY2023?

Result: OK · Latency: 18.5 s · Tool calls: 3

Expected answer

Best Buy generated the most cash flow from operating activities in FY 2023 (\$1.8 bn)

Delivered answer

Operating activities brought in the most cash for Best Buy in FY2023 (fiscal year ended January 28, 2023).

- Operating activities: **+\$1,824 million**
- Investing activities: **-\$962 million**
- Financing activities: **-\$1,806 million**

Only operations generated positive cash flow; both investing and financing were net uses of cash, so operations clearly led.

Cited documents: BESTBUY_2023_10K.pdf

Judge rationale: The agent's answer correctly identifies operating activities as the source of the most cash flow and provides the essential figure (\$1.8 bn), matching the expected answer.

Q54. Was there any drop in Cash & Cash equivalents between FY 2023 and Q2 of FY2024? Added note by ATG for disambiguation: the question is referring to Best Buy.

Result: OK · Latency: 18.6 s · Tool calls: 1

Expected answer

Yes, there was a decline of ~42% between FY2023 and Q2 of FY 2024.

Delivered answer

Yes — Cash and cash equivalents fell from **\$1,874 million** at the end of FY2023 (January 28, 2023) to **\$1,093 million** at the end of Q2 FY2024 (July 29, 2023), a drop of **\$781 million**.

- FY2023 (ended Jan 28, 2023): \$1,874M
- Q2 FY2024 (ended Jul 29, 2023): \$1,093M
- Change: \$1,093M – \$1,874M = **-\$781M**

Note: Best Buy's fiscal year ends on the Saturday nearest the end of January, so FY2023 ended January 28, 2023 and Q2 FY2024 ended July 29, 2023.

Cited documents: BESTBUY_2024Q2_10Q.pdf, BESTBUY_2023_10K.pdf

Judge rationale: The agent's answer confirms a drop in cash and cash equivalents and provides figures that align with the ~42% decline stated in the expected answer.

Q55. Was there any change in the number of Best Buy stores between Q2 of FY2024 and FY2023?

Result: OK · Latency: 118.5 s · Tool calls: 6

Expected answer

Yes, there is decline in number stores by 1.32% from 982 stores in Q2 FY 2023 to 969 by the end of Q2 FY2024.

Delivered answer

Yes — the number of Best Buy stores declined between the end of Q2 FY23 and Q2 FY24, from 1,142 to 1,129 (down 13 stores).

- **Domestic** (Best Buy, Outlet, Pacific Sales, Yardbird): 982 at end of Q2 FY23 → 969 at end of Q2 FY24 (down 13). Within this, the biggest moves were a drop of 23 "Best Buy" large-format stores (930→907) offset by Yardbird growth (13→22) and Outlet Centers (18→20).
- **International** (Canada): unchanged at 160 (Best Buy 127→128, Best Buy Mobile 33→32).

The store table appears in the Q2 FY24 10-Q's segment note, reconciling stores open at the beginning and end of each second quarter.

Cited documents: BESTBUY_2024Q2_10Q.pdf

Judge rationale: The agent's answer includes the essential information (decline in domestic stores from 982 to 969) and does not contradict the expected answer.

Q56. Which Best Buy product category performed the best (by top line) in the domestic (USA) Market during Q2 of FY2024?

Result: NOK · Latency: 13.6 s · Tool calls: 1

Expected answer

The entertainment segment experienced the highest growth of 9% during Q2 FY2024, primarily from gaming division.

Delivered answer

Computing and Mobile Phones was the top product category by revenue in Best Buy's Domestic (USA) segment in Q2 FY24 (the 13 weeks ended July 29, 2023).

- Domestic segment revenue: \$8,890 million
- Computing and Mobile Phones revenue mix: **41%** (≈ \$3,645 million)
- Next largest: Consumer Electronics at 30% (≈ \$2,667 million)

So Computing and Mobile Phones led the domestic top line by a wide margin, roughly 11 percentage points ahead of Consumer Electronics.

Cited documents: BESTBUY_2024Q2_10Q.pdf

Judge rationale: The agent's answer contradicts the expected answer by stating a different top-performing category (Computing and Mobile Phones) and does not mention the entertainment segment's 9% growth.

Q57. Considering the data in the balance sheet, what is Block's (formerly known as Square) FY2016 working capital ratio? Define working capital ratio as total current assets divided by total current liabilities. Round your answer to two decimal places.

Result: OK · Latency: 20.8 s · Tool calls: 2

Expected answer

1.73

Delivered answer

1.73

Using Block's (Square's) FY2016 balance sheet as of December 31, 2016: working capital ratio = total current assets (\$1,001,425 thousand) ÷ total current liabilities (\$577,464 thousand) = 1.73.

Cited documents: BLOCK_2016_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly, including the required precision of two decimal places.

Q58. What is the FY2019 - FY2020 total revenue growth rate for Block (formerly known as Square)? Answer in units of percents and round to one decimal place. Approach the question asked by assuming the standpoint of an investment banking analyst who only has access to the statement of income.

Result: OK · Latency: 34.5 s · Tool calls: 1

Expected answer

101.5%

Delivered answer

101.5%

Total net revenue grew from **\$4,713.5M** in FY2019 to **\$9,497.6M** in FY2020, an absolute increase of **\$4,784.1M**. Growth rate = $\$4,784.1M \div \$4,713.5M \approx 101.5\%$. The gain was driven largely by bitcoin revenue, which grew ~785% to \$4.57B and accounted for 85% of the total increase; excluding bitcoin, total net revenue rose only 17%.

Cited documents: BLOCK_2020_10K.pdf

Judge rationale: The agent's answer matches the expected growth rate of 101.5% and provides additional correct context without contradiction.

Q59. Using the cash flow statement, answer the following question to the best of your abilities: how much did Block (formerly known as Square) generate in cash flow from operating activities in FY2020? Answer in USD millions.

Result: OK (human review) · Latency: 18.2 s · Tool calls: 1

Expected answer

\$382.00

Delivered answer

\$381.6 million.

Per the FY2020 10-K (as originally reported), net cash provided by operating activities for the year ended December 31, 2020 was \$381,603 thousand, i.e. \$381.6 million.

Note: the FY2021 10-K restated this figure to \$173.1 million (\$173,110 thousand) after reclassifying customer funds from operating to financing activities.

Cited documents: BLOCK_2021_10K.pdf, BLOCK_2020_10K.pdf

Judge verdict: NOK · The agent's answer (\$381.6 million) contradicts the expected answer (\$382.00 million) and does not match the required precision.

Human review: OK · The agent provided a more detailed response than expected.

Q60. We need to calculate a financial metric by using information only provided within the balance sheet. Please answer the following question: what is Boeing's year end FY2018 net property, plant, and equipment (in USD millions)?

Result: OK · Latency: 29.6 s · Tool calls: 1

Expected answer

\$12645.00

Delivered answer

Boeing's year-end FY2018 net property, plant and equipment was **\$12,645 million** (\$12.6 billion).

The line appears directly on the Consolidated Statements of Financial Position at December 31, 2018: "Property, plant and equipment, net" of \$12,645 million. For reference, this reflects gross PP&E of \$31,213 million less accumulated depreciation of \$18,568 million, as detailed in the property note.

Cited documents: BOEING_2018_10K.pdf

Judge rationale: The agent's answer includes the essential figure (\$12,645 million) and does not contradict the expected answer.

Q61. Are there any product categories / service categories that represent more than 20% of Boeing's revenue for FY2022?

Result: OK · Latency: 23.4 s · Tool calls: 1

Expected answer

Yes. Boeing has product and service categories that represent more than 20% of Boeing's revenue for FY2022. These categories are Commercial Airplanes which comprises 39% of total revenue, Defence which comprises 35% of total revenue and Services which comprises 26% of total revenue.

Delivered answer

Yes.

Boeing's reportable product/service segments for FY2022 (fiscal year ended December 31, 2022), as a share of total revenue of \$66,608 million: Commercial Airplanes \$25,867M (38.8%), Defense, Space & Security \$23,162M (34.8%) and Global Services \$17,611M (26.4%) each exceed 20%, while Boeing Capital at \$199M (0.3%) does not. So three categories clear the 20% threshold. Total revenues were \$66,608M.

Cited documents: BOEING_2022_10K.pdf

Judge rationale: The agent's answer contains all essential information (three categories exceeding 20% with correct percentages) and no contradictions.

Q62. Has Boeing reported any materially important ongoing legal battles from FY2022?

Result: **OK** · Latency: 68.0 s · Tool calls: 5

Expected answer

Yes. Multiple lawsuits have been filed against Boeing resulting from a 2018 Lion Air crash and a 2019 Ethiopian Airlines crash.

Delivered answer

Yes. Boeing's FY2022 10-K (fiscal year ended December 31, 2022) describes multiple materially important ongoing legal battles in Note 21 – Legal Proceedings.

- **737 MAX (Lion Air & Ethiopian Airlines accidents):** Multiple legal actions remain pending, with losses not covered by insurance not reasonably estimable. Resolved components in the period included the DOJ Deferred Prosecution Agreement (the \$0.7 billion 2021 payment resolved via settlement with the DOJ), a shareholder derivative settlement that brought the company \$219 million in Q2 2022, and a September 2022 SEC settlement over the 737 MAX accidents carrying a \$200 million civil penalty (paid in October 2022).
- **Embraer arbitration:** Boeing terminated its 2019 joint-venture agreements with Embraer in 2020; Embraer disputes the termination, and the matter is in arbitration, scheduled to be completed in late 2023 or early 2024, with no reasonably estimable range of loss.

Both are presented as "potentially material contingencies," so ongoing material legal exposure from FY2022 is clearly reported.

Cited documents: BOEING_2022_10K.pdf

Judge rationale: The agent's answer includes the essential information from the expected answer (ongoing legal battles from the 2018 Lion Air and 2019 Ethiopian Airlines crashes) and adds non-contradictory details.

Q63. Does Boeing have an improving gross margin profile as of FY2022? If gross margin is not a useful metric for a company like this, then state that and explain why.

Result: OK · Latency: 20.8 s · Tool calls: 1

Expected answer

Yes. Boeing has an improving gross margin profile as of FY2022. Gross profit improved from \$3,017 million in FY2021 to \$3,502 million in FY2022. Gross margin % improved from 4.8% in FY2021 to 5.3% in FY2022.

Delivered answer

Yes. Boeing's gross margin improved in FY2022 (fiscal year ended December 31, 2022).

Gross margin = (Total revenues – Cost of products – Cost of services) / Total revenues, using the consolidated income statement:

- FY2022: $(66,608 - 53,969 - 9,109) / 66,608 = 3,530 / 66,608 = 5.30\%$
- FY2021: $(62,286 - 49,954 - 9,283) / 62,286 = 3,049 / 62,286 = 4.90\%$

That is an increase of about 0.40 percentage points year over year.

Gross margin is a relevant metric here: Boeing is an industrial/defense/aerospace manufacturer (not a financial institution), so its product and service costs are a meaningful core cost line, and the margin is meaningful. Note it is thin and volatile — Boeing posted an operating loss in both years (–\$3,547M in 2022, –\$2,902M in 2021), so the modest gross-margin gain was not enough to reach operating profitability.

Cited documents: BOEING_2022_10K.pdf

Judge rationale: The agent's answer confirms the improving gross margin profile with figures that match the expected answer when rounded to the same precision (5.3% vs 5.30%, 4.9% vs 4.8%).

Q64. Who are the primary customers of Boeing as of FY2022?

Result: **OK** · Latency: 32.6 s · Tool calls: 4

Expected answer

Boeing's primary customers as of FY2022 are a limited number of commercial airlines and the US government. The US government accounted for 40% of Boeing's total revenues in FY2022.

Delivered answer

Boeing's primary customers as of FY2022 (fiscal year ended December 31, 2022) are the **U.S. Government** and the **commercial airline industry worldwide**.

Here's the breakdown by segment:

- **U.S. Government** (including foreign military sales through the U.S. government) was the single largest customer, representing **40% of consolidated revenues** in 2022 (\$66,608 million total). This concentration is driven primarily by:
 - **Defense, Space & Security (BDS)**: U.S. government accounted for **89%** of BDS segment revenue. The U.S. Department of Defense (U.S. DoD) is BDS' primary customer.
 - **Global Services (BGS)**: U.S. government accounted for **33%** of BGS segment revenue. The U.S. government is also described as BGS' major customer, comprising over 50% of the government services markets served.
- **Commercial airline industry worldwide** is the primary customer for **Commercial Airplanes (BCA)**, which develops, produces and markets commercial jet aircraft (737, 767, 777, 787 families) principally to airlines globally.
- **Boeing Capital (BCC)** serves Boeing's own customers by providing financing solutions for aircraft purchases.

In summary, the U.S. Government (especially the Department of Defense) and commercial airlines are Boeing's two dominant customer groups, with the U.S. government alone accounting for 40% of total company revenue in FY2022.

Cited documents: BOEING_2022_10K.pdf

Judge rationale: The agent's answer includes all essential information from the expected answer (primary customers: commercial airlines and US government; US government accounted for 40% of revenues) and contains no contradictions.

Q65. Is Boeing's business subject to cyclical? Added note by ATG for disambiguation: the question is referring to FY2022.

Result: OK · Latency: 30.5 s · Tool calls: 3

Expected answer

Yes, Boeing's business is subject to cyclical due to its exposure to the airline industry which is a cyclical industry.

Delivered answer

Yes. Boeing's business is subject to cyclical, driven primarily by the cyclical nature of the airline industry that underpins its Commercial Airplanes and Global Services businesses.

In its FY2022 10-K (fiscal year ended December 31, 2022), Boeing states that "historically, the airline industry has been cyclical and very competitive and has experienced significant profit swings and constant challenges to be more cost competitive," and that significant deterioration in the global economic environment, the airline industry, or the financial stability of major customers could reduce new orders, prompt customers to postpone or cancel contractual orders and payments, and lower revenues, profitability, cash flows and backlog.

This is consistent with Boeing's own framing in prior filings, where it noted that "the cyclical of commercial and defense markets sometimes offset" — i.e., the commercial (airline-driven) and defense (government-budget-driven) cycles do not move in lockstep, so the company's overall results can be smoothed when one market is in a downturn and the other is not.

In short: the commercial side is explicitly cyclical (tied to airline traffic, profitability and financing), while the defense side is steadier but tied to government budget cycles, and the two tend to partially offset each other.

Cited documents: BOEING_2022_10K.pdf, BOEING_2017_10K.pdf

Judge rationale: The agent's answer contains the essential information that Boeing's business is subject to cyclical due to the airline industry and provides additional correct context without contradiction.

Q66. What production rate changes is Boeing forecasting for FY2023?

Result: **NOK** · Latency: 49.0 s · Tool calls: 3

Expected answer

Boeing forecasts an increase in the production rates for the 737, 777X and 787 aircrafts in 2023.

Delivered answer

For FY2023, Boeing is planning production-rate increases on its 737 and 787 programs and keeping 777X production paused, with the 747 winding down to the end.

- **737:** Raised to 31 per month during 2022; further gradual increases planned during 2023 based on market demand and supply chain capacity.
- **787:** Producing at low rates, expected to gradually increase back to 5 per month in 2023.
- **777/777X:** Production of the 777X-9 is paused during 2022 and 2023 (the pause was implemented in Q2 2022 and continues into 2023); the combined 777/777X line runs at 3 per month.
- **747:** Production was completed in Q4 2022, with the last delivery expected in early 2023—i.e., the program is ending.

The 767 stays flat at 3 per month. These come from Boeing's FY2022 10-K, the guidance published with the results covering 2023.

Cited documents: BOEING_2022_10K.pdf

Judge rationale: The agent answer contradicts the expected answer by stating that 777X production is paused in 2023, while the expected answer states an increase in production rates for 777X.

Q67. How does Boeing's effective tax rate in FY2022 compare to FY2021?

Result: **NOK** · Latency: 22.7 s · Tool calls: 1

Expected answer

Effective tax rate in FY2022 was 0.62%, compared to -14.76% in FY2021.

Delivered answer

Boeing's effective tax rate fell sharply from FY2021 to FY2022 — from **14.76%** in FY2021 to **0.62%** in FY2022, a decline of about 14.14 percentage points.

Both years were pre-tax loss years, so the sign of the tax line matters:

- **FY2021:** pre-tax loss of \$5,033M with an income tax **benefit** of \$743M → $743 / 5,033 = 14.76\%$
- **FY2022:** pre-tax loss of \$5,022M with an income tax **expense** of \$31M → $31 / 5,022 = 0.62\%$

In other words, FY2021's tax line reduced the loss (a benefit), while FY2022's tax line added a small expense on top of the loss — so the effective rate dropped to nearly zero.

Source: Boeing 2022 10-K, Consolidated Statements of Operations (years ended December 31, 2022 and 2021).

Cited documents: BOEING_2022_10K.pdf

Judge rationale: The agent's answer incorrectly states FY2021's effective tax rate as positive 14.76% instead of the correct -14.76%.

Q68. What is the FY2017 return on assets (ROA) for Coca Cola? ROA is defined as: FY2017 net income / (average total assets between FY2016 and FY2017). Round your answer to two decimal places. Give a response to the question by relying on the details shown in the balance sheet and the P&L statement.

Result: OK · Latency: 28.9 s · Tool calls: 2

Expected answer

0.01

Delivered answer

0.01 (1.42%)

Here's the calculation from the FY2017 10-K (year ended December 31, 2017):

- FY2017 net income attributable to shareowners of The Coca-Cola Company: **\$1,248 million**
- Total assets as of December 31, 2017: **\$87,896 million**
- Total assets as of December 31, 2016: **\$87,270 million**
- Average total assets: $(87,896 + 87,270) / 2 = \mathbf{\$87,583 \text{ million}}$
- $\text{ROA} = 1,248 / 87,583 = 0.0142 \rightarrow \mathbf{0.01}$

Cited documents: COCACOLA_2017_10K.pdf

Judge rationale: The agent's answer matches the expected ROA value of 0.01 when rounded to two decimal places and contains no contradictions.

Q69. What is Coca Cola's FY2021 COGS % margin? Calculate what was asked by utilizing the line items clearly shown in the income statement.

Result: OK · Latency: 25.6 s · Tool calls: 1

Expected answer

39.7%

Delivered answer

Coca-Cola's FY2021 COGS % margin was **39.7%**.

Calculation from the 2021 consolidated statement of income: Cost of goods sold of \$15,357 million divided by Net Operating Revenues of \$38,655 million = 39.7%.

Cited documents: COCACOLA_2021_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value and includes the correct calculation.

Q70. What is Coca Cola's FY2022 dividend payout ratio (using total cash dividends paid and net income attributable to shareholders)? Round answer to two decimal places. Answer the question asked by assuming you only have access to information clearly displayed in the cash flow statement and the income statement.

Result: OK · Latency: 15.2 s · Tool calls: 1

Expected answer

0.8

Delivered answer

0.80

Using the figures from The Coca-Cola Company's FY2022 10-K (fiscal year ended December 31, 2022):

- Total cash dividends paid (cash flow statement, financing activities): **\$7,616 million**
- Net income attributable to shareowners of The Coca-Cola Company (income statement): **\$9,542 million**

Payout ratio = $\$7,616 / \$9,542 = 0.7982 \approx \mathbf{0.80}$

(Note: the statement of shareowners' equity shows dividends of \$7,617 million, but per your instruction I used the cash flow statement figure of \$7,616 million.)

Cited documents: COCACOLA_2022_10K.pdf

Judge rationale: The agent's answer (0.80) matches the expected answer (0.8) when rounded to two decimal places, and the calculation is correct.

Q71. Based on the information provided primarily in the balance sheet and the statement of income, what is FY2020 days payable outstanding (DPO) for Corning? DPO is defined as: $365 \times (\text{average accounts payable between FY2019 and FY2020}) / (\text{FY2020 COGS} + \text{change in inventory between FY2019 and FY2020})$. Round your answer to two decimal places.

Result: OK · Latency: 33.9 s · Tool calls: 4

Expected answer

63.86

Delivered answer

63.86 days

Here's the calculation using Corning's FY2020 10-K figures (all in millions):

- **Average accounts payable** = $(\$1,174 + \$1,587) / 2 = \$1,380.5$
- **FY2020 COGS (Cost of sales)** = \$7,772
- **Change in inventory** = $\$2,438 - \$2,320 = \$118$

DPO = $365 \times \$1,380.5 / (\$7,772 + \$118) = 365 \times \$1,380.5 / \$7,890 = \mathbf{63.86}$

Cited documents: CORNING_2020_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly, including the correct calculation and rounding to two decimal places.

Q72. Taking into account the information outlined in the income statement, what is the FY2019 - FY2021 3 year average unadjusted operating income % margin for Corning? Answer in units of percents and round to one decimal place.

Result: OK · Latency: 43.1 s · Tool calls: 6

Expected answer

10.3%

Delivered answer

The FY2019–FY2021 three-year average unadjusted operating income margin for Corning is **10.3%**.

Annual unadjusted operating income margins (GAAP income statement, all December year-ends):

- FY2019: \$1,306 / \$11,503 = 11.4%
- FY2020: \$509 / \$11,303 = 4.5%
- FY2021: \$2,112 / \$14,082 = 15.0%

Average of the three: $(11.3536 + 4.5032 + 14.9979) / 3 = 10.28\%$, rounded to one decimal = 10.3%.

Cited documents: CORNING_2020_10K.pdf, CORNING_2021_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (10.3%) and includes correct supporting calculations.

Q73. How much has the effective tax rate of Corning changed between FY2021 and FY2022?

Result: OK (human review) · Latency: 43.2 s · Tool calls: 4

Expected answer

The effective tax rate of Corning has changed from 20% in FY2021 to 23% in FY 2022.

Delivered answer

Corning's effective tax rate **increased by 2.4 percentage points** between FY2021 and FY2022.

The change is computed from the income tax note reconciliations: 22.9% in FY2022 versus 20.5% in FY2021 (as originally reported in the 2021 10-K; the 2022 10-K restates 2021 at 20.2%, which would make the increase 2.7 pp). The rise was driven by net provisions for changes in tax reserves, foreign earnings differences and valuation allowance assessments, partly offset by tax credits.

Cited documents: CORNING_2022_10K.pdf, CORNING_2021_10K.pdf

Judge verdict: NOK · The agent's answer contradicts the expected answer by providing different figures (22.9% vs 23% and 20.5%/20.2% vs 20%) and adds an incorrect change magnitude (2.4 or 2.7 percentage points vs the expected 3 percentage points).

Human review: OK · The figure is more specific in the agent's response

Q74. Does Corning have positive working capital based on FY2022 data? If working capital is not a useful or relevant metric for this company, then please state that and explain why.

Result: **NOK** · Latency: 31.7 s · Tool calls: 2

Expected answer

Yes. Corning had a positive working capital amount of \$831 million by FY 2022 close. This answer considers only operating current assets and current liabilities that were clearly shown in the balance sheet.

Delivered answer

Yes—Corning has positive working capital based on FY2022 data, and working capital is a relevant metric for Corning since it is a manufacturing (non-financial) company.

- Total current assets: \$7,453 million
- Total current liabilities: \$5,175 million
- Working capital: \$7,453 – \$5,175 = **\$2,278 million** (positive)

The company's balance sheet for the year ended December 31, 2022 shows current assets of \$7,453 million against current liabilities of \$5,175 million, leaving working capital of roughly \$2.3 billion.

Cited documents: CORNING_2022_10K.pdf

Judge rationale: The agent's working capital figure (\$2,278 million) contradicts the expected answer (\$831 million) and does not match the reference's scope (operating current assets and liabilities only).

Q75. Using only the information within the balance sheet, how much total assets did Costco have at the end of FY2021? Answer in USD millions.

Result: **OK** · Latency: 18.3 s · Tool calls: 1

Expected answer

\$59268.00

Delivered answer

Costco's total assets at the end of FY2021 (fiscal year ended August 29, 2021) were **\$59,268 million**.

This comes straight from the consolidated balance sheet's TOTAL ASSETS line in the FY2021 10-K: \$59,268 million, up from \$55,556 million at the end of FY2020.

Cited documents: COSTCO_2021_10K.pdf

Judge rationale: The agent's answer matches the expected answer's essential figure (\$59,268 million) and contains no contradictions.

Q76. What is the FY2018 fixed asset turnover ratio for CVS Health? Fixed asset turnover ratio is defined as: FY2018 revenue / (average PP&E between FY2017 and FY2018). Round your answer to two decimal places. Calculate what was asked by utilizing the line items clearly shown in the P&L statement and the balance sheet.

Result: OK · Latency: 24.5 s · Tool calls: 3

Expected answer

17.98

Delivered answer

17.98

- FY2018 total revenues: \$194,579 million (consolidated statement of operations, 2018 10-K)
- PP&E, net FY2018: \$11,349 million; FY2017: \$10,292 million (balance sheet)
- Average PP&E = $(11,349 + 10,292) / 2 = \$10,820.5$ million
- Fixed asset turnover = $194,579 / 10,820.5 = \mathbf{17.98}$

Cited documents: CVSHEALTH_2018_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (17.98) and includes the correct calculation steps and figures.

Q77. Is CVS Health a capital-intensive business based on FY2022 data?

Result: OK · Latency: 24.3 s · Tool calls: 2

Expected answer

Yes, CVS Health requires an extensive asset base to operate, which is evident from its ROA of only 1.82% in 2022 and 3.39% in 2021, though it should be noted that a significant portion of this asset base is goodwill, and CVS's fixed assets/total assets ratio is on the lower side of 5.6%.

Delivered answer

Yes, CVS Health is capital-intensive based on FY2022 data (fiscal year ended December 31, 2022).

- Capex / revenue: $\$2,727\text{M} / \$322,467\text{M} = \mathbf{0.85\%}$
- Net PP&E / total assets: $\$12,873\text{M} / \$228,275\text{M} = \mathbf{5.64\%}$
- Return on assets (year-end): $\$4,149\text{M} / \$228,275\text{M} = \mathbf{1.82\%}$
- Return on assets (two-year avg): $\$4,149\text{M} / \$230,637\text{M} = \mathbf{1.80\%}$

The largest component of the \$228.3B asset base is goodwill at \$78.2B (34.2%), not physical fixed assets. However, the company earns a low return (~1.8%) on that large asset base, which by definition makes it capital-intensive even when goodwill rather than fixed assets drives the balance sheet.

Cited documents: CVSHEALTH_2022_10K.pdf

Judge rationale: The agent's answer confirms CVS Health is capital-intensive, includes the essential ROA figure (1.82%) and the fixed assets/total assets ratio (~5.6%), and does not contradict the reference.

Q78. Has CVS Health reported any materially important ongoing legal battles from 2022, 2021 and 2020?

Result: **OK** · Latency: 44.2 s · Tool calls: 2

Expected answer

Yes, CVS Health has been involved in multiple ongoing legal battles. Some notable legal dispute areas for CVS are: (1) usual and customary pricing litigation: where it's claimed that CVS's retail pharmacies overcharged for prescription drugs; (2) PBM litigation and investigations: where it's claimed that that rebate agreements between the drug manufacturers and PBMs caused inflated prices for certain drug products; and (3) controlled substances litigation: legal matters around opioids for which CVS has agreed to pay up to \$4.3 billion to claimants in remediation and \$625 million to attorneys and fees

Delivered answer

Yes — in each of its 2020, 2021 and 2022 10-Ks, CVS Health has disclosed a range of materially important ongoing legal battles it is actively defending. The main families of matters are:

- **Opioid litigation (national MDL):** CVS is a defendant in the *In re National Prescription Opiate Litigation* (MDL 2804, N.D. Ohio), which "presumptively includes hundreds of federal cases" naming it, plus many state-court cases and actions by state Attorneys General. In the 2021 10-K it disclosed a November 2021 Ohio jury verdict finding chain pharmacies (including CVS) liable, with remedy undetermined; by the 2022 10-K the court entered a \$651 million joint judgment (payable over 15 years), which CVS is appealing.
- **Opioid settlements:** In the 2022 10-K, CVS disclosed a \$484 million settlement with Florida (recorded as a liability in Q1 2022), agreements with New Mexico, West Virginia and the Cherokee Nation in 2022, and a December 2022 national settlement with state Attorneys General/Plaintiffs' Executive Committee (max ~\$4.3 billion in remediation plus ~\$625 million fees, payable over 10 years from 2023) plus a ~\$113 million tribes settlement. It recorded pre-tax opioid charges of \$5.2 billion (Q3 2022) and \$99 million (Q4 2022), ~\$4.8 billion in other long-term liabilities. It did not accrue for the Ohio judgment.
- **PBM litigation and investigations:** multiple lawsuits (including antitrust, rebate and insulin-pricing cases) and investigations by state Attorneys General, the DC, and — in later years — the FTC, over drug pricing and rebates; plus the *Behnke* Medicare Part D qui tam (U.S. ex rel.) claim.
- **Usual and customary pricing litigation:** class and qui tam actions alleging retail pharmacies overcharged by not submitting the Health Savings Pass price as the usual and customary price (e.g., Texas, Corcoran), still being defended in the 2020 filing.
- **Controlled-substance enforcement:** a January 2020 DEA administrative subpoena and a January 2022 DOJ CID regarding opioid/controlled-substance practices under the Controlled Substances Act and False Claims Act.
- **Stockholder/ERISA/derivative matters:** securities class actions and derivative claims over the LTC business unit and Aetna synergies (e.g., *Anarkat/City of Miami*, *Waterford*, *Freundlich*), ERISA 401(k) claims, and a December 2021 Section 220 books-and-records demand tied to opioids.

In each filing CVS states it believes none of these matters (after reserves and indemnification rights) will have a material adverse effect on its financial position, but that the outcome of governmental investigations and undetermined verdicts could be material.

Cited documents: CVSHEALTH_2021_10K.pdf, CVSHEALTH_2020_10K.pdf, CVSHEALTH_2022_10K.pdf

Judge rationale: The agent's answer includes all essential legal battle categories from the expected answer (opioid litigation, PBM litigation, and usual and customary pricing litigation) and provides additional correct details without contradiction.

Q79. Has CVS Health paid dividends to common shareholders in Q2 of FY2022?

Result: **OK** · Latency: 15.1 s · Tool calls: 1

Expected answer

Yes, CVS paid a \$ 0.55 dividend per share every quarter in FY2022

Delivered answer

Yes. CVS Health paid a quarterly cash dividend of **\$0.55 per common share** in Q2 of FY2022.

- The 2022 10-K states: "During 2022, 2021 and 2020 the quarterly cash dividend was \$0.55, \$0.50 and \$0.50 per share, respectively".
- The \$0.55 rate took effect in 2022 after the Board authorized a 10% increase in December 2021, and the company "has paid cash dividends every quarter since becoming a public company".

So the Q2 2022 dividend was \$0.55 per share, consistent with the full-year 2022 run rate of \$2.20 per share.

Cited documents: CVSHEALTH_2022_10K.pdf

Judge rationale: The agent's answer confirms the dividend payment and matches the essential information in the expected answer, including the amount (\$0.55 per share) and the quarter (Q2 FY2022).

Q80. Does Foot Locker's new CEO have previous CEO experience in a similar company to Footlocker?

Added note by ATG for disambiguation: the question is referring to the filing dated 2022-08-19.

Result: **OK** · Latency: 17.1 s · Tool calls: 1

Expected answer

Yes. She was previous CEO of Ulta Beauty which means she had to manage a large retail company that has brick and mortar + online business. So yes she was a CEO in a similar company to Foot Locker before this.

Delivered answer

Yes. Foot Locker's new CEO, Mary N. Dillon, served as President and Chief Executive Officer of Ulta Beauty, Inc. from July 2013 through June 2021, and then as Executive Chair from June 2021 through June 2022.

Ulta Beauty is a specialty consumer retailer with a physical store network and omnichannel operations, making it a similar business model to Foot Locker's specialty athletic retail. The 8-K press release itself highlights her "remarkable track record in the retail industry" and "experience leading consumer-driven businesses and building robust brand portfolios" as key reasons for her appointment.

Cited documents: FOOTLOCKER_2022_8K_dated_2022-08-19.pdf

Judge rationale: The agent's answer confirms the CEO's prior experience at Ulta Beauty, a similar retail company, matching the essential information in the expected answer.

Q81. Were there any board member nominees who had substantially more votes against joining than the other nominees? Added note by ATG for disambiguation: the question is referring to Foot Locker's filing dated 2022-05-20.

Result: OK · Latency: 14.0 s · Tool calls: 1

Expected answer

Yes, his name is Richard A. Johnson

Delivered answer

Yes — **Richard A. Johnson** (then Chairman & CEO) received far more votes against his election than any other nominee in Foot Locker's May 18, 2022 annual meeting.

Nominee	Votes Against
Richard A. Johnson	16,105,005
Dona D. Young	6,074,467
Guillermo G. Marmol	5,753,395
Alan D. Feldman	5,184,437
Ulice Payne, Jr.	4,950,917
Kimberly Underhill	4,906,500
Darlene Nicosia	4,827,808
Steven Oakland	4,524,393
Tristan Walker	4,419,340
Virginia C. Drosos	294,935

Johnson's ~16.1 million "against" votes were roughly 2.7× the next-closest nominee (Dona D. Young, ~6.1M) and about 3× the typical 4.4–5.8 million that most others drew, so he clearly stands apart. Despite that, he was still duly elected along with all ten nominees, since votes for (54,484,293) far exceeded votes against.

Cited documents: FOOTLOCKER_2022_8K_dated-2022-05-20.pdf

Judge rationale: The agent's answer correctly identifies Richard A. Johnson as the nominee with substantially more votes against, matching the essential information in the expected answer.

Q82. What is the FY2019 cash conversion cycle (CCC) for General Mills? CCC is defined as: DIO + DSO - DPO. DIO is defined as: $365 \times (\text{average inventory between FY2018 and FY2019}) / (\text{FY2019 COGS})$. DSO is defined as: $365 \times (\text{average accounts receivable between FY2018 and FY2019}) / (\text{FY2019 Revenue})$. DPO is defined as: $365 \times (\text{average accounts payable between FY2018 and FY2019}) / (\text{FY2019 COGS} + \text{change in inventory between FY2018 and FY2019})$. Round your answer to two decimal places. Address the question by using the line items and information shown within the income statement and the balance sheet.

Result: OK · Latency: 35.2 s · Tool calls: 1

Expected answer

-3.7

Delivered answer

−3.70 days

Here's the calculation using General Mills' FY2019 10-K (fiscal year ended May 26, 2019):

Inputs (in \$ millions):

Item	FY2019	FY2018
Inventories	1,559.3	1,642.2
Receivables	1,679.7	1,684.2
Accounts payable	2,854.1	2,746.2
Net sales (Revenue)	16,865.2	—
Cost of sales (COGS)	11,108.4	—

DIO = $365 \times [(1,559.3 + 1,642.2)/2] / 11,108.4 = 365 \times 1,600.75 / 11,108.4 = \mathbf{52.60}$

DSO = $365 \times [(1,679.7 + 1,684.2)/2] / 16,865.2 = 365 \times 1,681.95 / 16,865.2 = \mathbf{36.40}$

DPO = $365 \times [(2,854.1 + 2,746.2)/2] / (11,108.4 + (1,559.3 - 1,642.2)) = 365 \times 2,800.15 / 11,025.5 = \mathbf{92.70}$

CCC = $52.60 + 36.40 - 92.70 = \mathbf{-3.70 \text{ days}}$

The negative CCC means General Mills collected cash from customers before it had to pay its suppliers, effectively financing its working capital with supplier credit.

Cited documents: GENERALMILLS_2019_10K.pdf

Judge rationale: The agent's answer matches the expected answer (-3.7) when rounded to two decimal places and contains no contradictions.

Q83. By drawing conclusions from the information stated only in the statement of financial position, what is General Mills's FY2020 working capital ratio? Define working capital ratio as total current assets divided by total current liabilities. Round your answer to two decimal places.

Result: OK · Latency: 20.1 s · Tool calls: 1

Expected answer

0.68

Delivered answer

General Mills's FY2020 working capital ratio is **0.68**.

Total current assets of \$5,121.3 million divided by total current liabilities of \$7,491.5 million equals 0.68, from the fiscal 2020 (ended May 31, 2020) balance sheet.

Cited documents: GENERALMILLS_2020_10K.pdf

Judge rationale: The agent's answer matches the expected working capital ratio of 0.68 and provides the correct calculation.

Q84. According to the information provided in the statement of cash flows, what is the FY2020 free cash flow (FCF) for General Mills? FCF here is defined as: (cash from operations - capex). Answer in USD millions.

Result: OK · Latency: 13.9 s · Tool calls: 1

Expected answer

\$3215.00

Delivered answer

\$3,215.4 million.

From General Mills' FY2020 Consolidated Statement of Cash Flows (fiscal year ended May 31, 2020):

- Net cash provided by operating activities: \$3,676.2M
- Purchases of land, buildings, and equipment (capex): \$460.8M

$\text{FCF} = \$3,676.2\text{M} - \$460.8\text{M} = \mathbf{\$3,215.4\text{M}}$

This reconciles exactly with the "Free cash flow" line of \$3,215.4M reported in the Selected Financial Data table of the same filing.

Cited documents: GENERALMILLS_2020_10K.pdf

Judge rationale: The agent's answer matches the expected answer when rounded to the same precision (\$3,215.4 million rounds to \$3,215.00 million).

Q85. We want to calculate a financial metric. Please help us compute it by basing your answers off of the cash flow statement and the income statement. Here's the question: what is the FY2022 retention ratio (using total cash dividends paid and net income attributable to shareholders) for General Mills? Round answer to two decimal places.

Result: OK · Latency: 30.3 s · Tool calls: 2

Expected answer

0.54

Delivered answer

0.54

The retention ratio for General Mills in FY2022 (fiscal year ended May 29, 2022) is computed as (net income attributable to shareholders – total cash dividends paid) / net income attributable to shareholders. Using net earnings attributable to General Mills of \$2,707.3 million and cash dividends paid of \$1,244.5 million from the cash flow statement: $(2,707.3 - 1,244.5) / 2,707.3 = 1,462.8 / 2,707.3 = 0.5403$, rounded to 0.54.

Cited documents: GENERALMILLS_2022_10K.pdf

Judge rationale: The agent's answer matches the expected retention ratio of 0.54 and includes the correct calculation and rounding.

Q86. Are JnJ's FY2022 financials that of a high growth company?

Result: OK · Latency: 15.7 s · Tool calls: 1

Expected answer

No, JnJ's FY2022 financials are not of a high growth company as sales grew by 1.3% in FY2022.

Delivered answer

No — JnJ's FY2022 (fiscal year ended January 1, 2023) financials are those of a mature, low-growth large-cap, not a high-growth company.

- Revenue: \$94,943M vs \$93,775M in FY2021 = **+1.3%** (down sharply from +13.6% in 2021).
- Net earnings: \$17,941M vs \$20,878M = **-14.1%**; diluted EPS \$6.73 vs \$7.81 = **-13.8%**.
- The 10-K itself reports a 5-year worldwide sales CAGR of just **4.4%** (10-year: 3.5%), and 2022 growth was held back by a -4.8% currency drag.

Single-digit (or sub-single-digit) revenue growth plus a declining bottom line is the opposite of a high-growth profile.

Cited documents: JOHNSON_JOHNSON_2022_10K.pdf, JOHNSON_JOHNSON_2022Q4_EARNINGS.pdf

Judge rationale: The agent's answer correctly states that JnJ's FY2022 financials are not of a high growth company and includes the essential 1.3% sales growth figure.

Q87. What drove gross margin change as of FY2022 for JnJ? If gross margin is not a useful metric for a company like this, then please state that and explain why.

Result: OK · Latency: 45.0 s · Tool calls: 3

Expected answer

For FY22, JnJ had changes in gross margin due to: One-time COVID-19 vaccine manufacturing exit related costs, Currency impacts in the Pharmaceutical segment, Commodity inflation in the MedTech and Consumer Health segments, partially offset by Supply chain benefits in the Consumer Health segment.

Delivered answer

Gross margin declined by about 0.9 percentage points in FY2022 (fiscal year ended January 1, 2023), from 68.16% to 67.25%.

Gross margin is a useful metric for JnJ — it is a healthcare (pharmaceutical/medtech/consumer) company, not a financial institution, so gross profit and margin are standard and meaningful measures of its product profitability.

The inputs and calculation (from the FY2022 10-K income statement):

- FY2022: gross profit \$63,854M ÷ sales \$94,943M = 67.25%
- FY2021: gross profit \$63,920M ÷ sales \$93,775M = 68.16%
- Change: 67.25% – 68.16% = –0.91 pp

What drove the decline (per the MD&A, cost of products sold rose as a percent of sales):

- One-time COVID-19 vaccine manufacturing exit-related costs
- Currency impacts in the Pharmaceutical segment
- Commodity inflation in the MedTech and Consumer Health segments
- Partially offset by supply chain benefits in the Consumer Health segment

Cited documents: JOHNSON_JOHNSON_2022_10K.pdf

Judge rationale: The agent's answer includes all essential drivers of the gross margin change (COVID-19 vaccine exit costs, currency impacts, commodity inflation, and supply chain benefits) and does not contradict the expected answer.

Q88. Roughly how many times has JnJ sold its inventory in FY2022? Calculate inventory turnover ratio for FY2022; if conventional inventory management is not meaningful for the company then state that and explain why.

Result: OK · Latency: 22.5 s · Tool calls: 1

Expected answer

JnJ sold its inventory 2.7 times in FY2022.

Delivered answer

Johnson & Johnson turned its inventory about **2.49×** on the year-end balance (2.72× on the two-year average) in FY2022 — roughly 2.5 times, and conventional inventory management is meaningful for this company.

- Cost of products sold FY2022: **\$31,089M**; inventories at year-end (Jan 1, 2023): **\$12,483M** → $31,089 / 12,483 = \mathbf{2.49}$
- Two-year average inventory (Jan 1, 2023 and Jan 2, 2022: \$10,387M): $31,089 / 11,435 = \mathbf{2.72}$

FY2022 ended January 1, 2023 (a non-calendar fiscal year), and the figures are from that year's 10-K and Q4 earnings release. Inventory turnover is a standard, relevant metric here because J&J is a manufacturer of physical goods across its Pharmaceutical, MedTech and Consumer Health segments — it is not a financial institution, so the usual exemption does not apply. The data show inventory management is meaningfully evaluated (inventories rose ~20% during the year), and a conventional turnover ratio around 2.5× indicates inventory was sold roughly every ~145 days.

Cited documents: JOHNSON_JOHNSON_2022_10K.pdf, JOHNSON_JOHNSON_2022Q4_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential information (inventory turnover ratio ~2.5-2.7 times) and does not contradict the expected answer.

Q89. Is growth in JnJ's adjusted EPS expected to accelerate in FY2023?

Result: **NOK** · Latency: 51.5 s · Tool calls: 2

Expected answer

No, rate of growth in adjusted EPS is expected to decelerate slightly from 3.6% in FY2022 to 3.5% in FY2023.

Delivered answer

Yes. J&J's FY2023 adjusted EPS growth was expected to accelerate relative to FY2022.

- FY2022 actual adjusted EPS: \$10.15, up **3.6%** from \$9.80 in FY2021
- FY2023 initial guidance (Jan 2023): adjusted EPS midpoint \$10.55, implying **4.0%** growth

4.0% > 3.6%, so growth was expected to accelerate. The company subsequently raised the FY2023 adjusted EPS guidance to a \$10.75 midpoint (6.0% growth) in July 2023, and after the Kenvue separation in August 2023, the recast guidance (ex-Consumer Health) showed a 12.5% midpoint growth rate.

Cited documents: JOHNSON_JOHNSON_2023_8K_dated-2023-08-30.pdf, JOHNSON_JOHNSON_2023Q2_EARNINGS.pdf, JOHNSON_JOHNSON_2022Q4_EARNINGS.pdf

Judge rationale: The agent's answer contradicts the expected answer by stating growth is expected to accelerate, while the reference states it is expected to decelerate.

Q90. How did JnJ's US sales growth compare to international sales growth in FY2022?

Result: **OK** · Latency: 25.8 s · Tool calls: 1

Expected answer

US sales increased 3.0% vs international sales decline of 0.6%.

Delivered answer

US sales growth outpaced international in FY2022. On a reported basis, US sales rose 3.0% (to \$48.6 billion from \$47.2 billion), while international sales fell 0.6% (to \$46.4 billion from \$46.6 billion). The international decline was entirely currency-driven: operationally, international grew 9.1%, versus 3.0% for the US, with a -9.7% currency headwind.

Cited documents: JOHNSON_JOHNSON_2022_10K.pdf, JOHNSON_JOHNSON_2022Q4_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential figures (US +3.0%, international -0.6%) and does not contradict the reference.

Q91. Which business segment of JnJ will be treated as a discontinued operation from August 30, 2023 onward?

Result: OK · Latency: 14.0 s · Tool calls: 1

Expected answer

The Consumer Health business segment will be treated as a discontinued operation from August 30, 2023 onward.

Delivered answer

Consumer Health (the Kenvue business) is the segment treated as a discontinued operation from August 30, 2023 onward.

Johnson & Johnson completed an exchange offer to finalize the separation of Kenvue Inc., formerly its Consumer Health business. As a result, J&J presents the Consumer Health business financial results as discontinued operations, including a gain of approximately \$20 billion in the third quarter of 2023.

Cited documents: JOHNSON_JOHNSON_2023_8K_dated-2023-08-30.pdf

Judge rationale: The agent's answer correctly identifies the Consumer Health business segment as the discontinued operation from August 30, 2023 onward, matching the essential information in the expected answer.

Q92. What is the amount of the gain accruing to JnJ as a result of the separation of its Consumer Health business segment, as of August 30, 2023?

Result: OK · Latency: 24.8 s · Tool calls: 3

Expected answer

JnJ will make a gain of approximately \$20 billion from the separation of its Consumer Health business segment.

Delivered answer

Approximately **\$20 billion**.

Johnson & Johnson's 8-K dated August 30, 2023 states that, following completion of the exchange offer to finalize the separation of Kenvue Inc. (formerly its Consumer Health business), the company will present Consumer Health as discontinued operations, "including a gain of approximately \$20 billion in the third quarter of 2023."

Cited documents: JOHNSON_JOHNSON_2023_8K_dated-2023-08-30.pdf

Judge rationale: The agent's answer contains the essential information (approximately \$20 billion gain) and does not contradict the expected answer.

Q93. What is the amount of the cash proceeds that JnJ realised from the separation of Kenvue (formerly Consumer Health business segment), as of August 30, 2023?

Result: OK · Latency: 72.8 s · Tool calls: 2

Expected answer

JnJ realised \$13.2 billion in cash proceeds from the separation of Kenvue.

Delivered answer

\$13.2 billion in cash proceeds from the Kenvue separation, as of August 30, 2023.

Johnson & Johnson's August 30, 2023 update states the Company "secured \$13.2 billion in cash proceeds from the Kenvue debt offering and initial public offering" as part of finalizing the separation of Kenvue (its former Consumer Health segment), while retaining a 9.5% equity stake in Kenvue.

Cited documents: JOHNSON_JOHNSON_2023_8K_dated-2023-08-30.pdf

Judge rationale: The agent's answer contains the essential information (\$13.2 billion) and does not contradict the expected answer.

Q94. Did JnJ's net earnings as a percent of sales increase in Q2 of FY2023 compared to Q2 of FY2022?

Result: OK · Latency: 40.1 s · Tool calls: 1

Expected answer

Yes, net earnings as a percent of sales increased from 20% in Q2 of FY2022 to 20.1% in Q2 of FY2023.

Delivered answer

Yes. Net earnings as a percent of sales edged up from 20.0% in Q2 2022 to 20.1% in Q2 2023. In dollars, Q2 2023 net earnings rose 6.9% to \$5,144M on sales of \$25,530M (20.1%), versus \$4,814M on \$24,020M (20.0%) in Q2 2022.

Cited documents: JOHNSON_JOHNSON_2023Q2_EARNINGS.pdf

Judge rationale: The agent's answer correctly states the increase from 20.0% to 20.1% and contains no contradictions or errors.

Q95. Which of JPM's business segments had the lowest net revenue in 2021 Q1?

Result: **NOK** · Latency: 41.4 s · Tool calls: 2

Expected answer

Corporate. Its net revenue was -\$473 million.

Delivered answer

Commercial Banking (CB) had the lowest net revenue of JPMorgan's business segments in Q1 2021, at \$2,393 million.

In Q1 2021 (managed basis, in millions): CCB \$12,517, CIB \$14,605, AWM \$4,077, and CB \$2,393. Commercial Banking is the smallest by revenue, well below AWM at \$4,077 million.

Cited documents: JPMORGAN_2021Q1_10Q.pdf

Judge rationale: The agent's answer incorrectly identifies Commercial Banking as the segment with the lowest net revenue, while the expected answer states it is Corporate with a negative net revenue of -\$473 million.

Q96. If JPM went bankrupted by the end by 2021 Q1 and liquidated all of its assets to pay its shareholders, how much could each shareholder get?

Result: **OK** · Latency: 47.7 s · Tool calls: 1

Expected answer

They could receive \$66.56 per share.

Delivered answer

\$66.56 per share — that's JPMorgan's tangible book value per share (TBVPS) at March 31, 2021, which is the right measure for a liquidation because it backs out the goodwill and other intangibles that would fetch nothing in a sale.

The math behind it: total stockholders' equity was \$280.7 billion as of Q1 2021; less preferred stock and less goodwill and intangible assets gives tangible common equity, divided by common shares outstanding. JPMorgan reports this directly as \$66.56 TBVPS at the end of Q1 2021.

For contrast, plain book value per share was \$82.31, but the \$15.75 drop to \$66.56 reflects the goodwill and intangibles that vanish in a forced liquidation. In practice a bankruptcy recovery would almost certainly be lower than TBVPS, since assets rarely realize their carrying values in a distressed sale — but on the balance-sheet figures as reported, shareholders would stand to receive about \$66.56 per share.

Cited documents: JPMORGAN_2021Q1_10Q.pdf

Judge rationale: The agent's answer includes the essential figure (\$66.56 per share) and provides additional correct context without contradiction.

Q97. Are JPM's gross margins historically consistent (not fluctuating more than roughly 2% each year)? If gross margins are not a relevant metric for a company like this, then please state that and explain why. Added note by ATG for disambiguation: the question is referring to FY2022.

Result: OK · Latency: 18.6 s · Tool calls: 1

Expected answer

Since JPM is a financial institution, gross margin is not a relevant metric.

Delivered answer

Gross margin is not a relevant metric for JPMorgan Chase, and I won't compute it.

JPM is a financial holding company / bank (a leader in investment banking, consumer banking, commercial banking and asset management). For a lender, interest income and interest expense, along with credit losses, are part of its core operations rather than a discrete "cost of goods sold," so there is no meaningful cost-of-sales line from which a gross margin could be constructed. The metrics that measure a bank's performance are net interest income, pre-provision profit, return on equity/tangible common equity and efficiency (overhead) ratios—indeed JPM's own FY2022 reporting centers on those, not gross margin.

So for FY2022 (and any year), gross-margin consistency is not something that applies to this kind of company.

Cited documents: JPMORGAN_2022_10K.pdf

Judge rationale: The agent correctly states that gross margin is not a relevant metric for JPMorgan Chase and explains why, matching the essential information in the expected answer.

Q98. In 2022 Q2, which of JPM's business segments had the highest net income?

Result: OK · Latency: 28.4 s · Tool calls: 2

Expected answer

Corporate & Investment Bank. Its net income was \$3725 million.

Delivered answer

Corporate & Investment Bank (CIB) had the highest net income in Q2 2022, at **\$3,725 million**.

The segment net income figures for the three months ended June 30, 2022 were:

- CIB: \$3,725M
- Consumer & Community Banking: \$3,100M
- Asset & Wealth Management: \$1,004M
- Commercial Banking: \$994M
- Corporate: \$(174)M

CIB led by \$625M over the next-highest segment, Consumer & Community Banking.

Cited documents: JPMORGAN_2022Q2_10Q.pdf

Judge rationale: The agent's answer correctly identifies Corporate & Investment Bank as the segment with the highest net income and matches the figure of \$3,725 million.

Q99. Looking at VaR, did the risk that JPM faced in the second fiscal quarter of 2023 decrease compared to the same period in the prior year?

Result: OK · Latency: 15.6 s · Tool calls: 1

Expected answer

Yes. It decreased.

Delivered answer

Yes, the risk decreased. JPM's average total VaR fell by \$7 million, from \$54 million in the second quarter of 2022 to \$47 million in the second quarter of 2023, a decline driven mainly by risk reductions in Credit Portfolio VaR and fixed income.

Cited documents: JPMORGAN_2023Q2_10Q.pdf

Judge rationale: The agent's answer contains the essential information (risk decreased) and provides additional correct details without contradiction.

Q100. What is Kraft Heinz's FY2019 inventory turnover ratio? Inventory turnover ratio is defined as: (FY2019 COGS) / (average inventory between FY2018 and FY2019). Round your answer to two decimal places. Please base your judgments on the information provided primarily in the balance sheet and the P&L statement.

Result: OK · Latency: 41.7 s · Tool calls: 3

Expected answer

6.25

Delivered answer

6.25

From Kraft Heinz's FY2019 10-K (fiscal year ended December 28, 2019):

- FY2019 Cost of products sold: \$16,830M
- Inventory at Dec 28, 2019: \$2,721M
- Inventory at Dec 29, 2018: \$2,667M
- Average inventory: $(2,721 + 2,667) / 2 = \$2,694\text{M}$
- Inventory turnover: $16,830 / 2,694 = \mathbf{6.25}$

Cited documents: KRAFTHEINZ_2019_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly, including the required rounding to two decimal places.

Q101. We need to calculate a reasonable approximation (or exact number if possible) of a financial metric. Basing your judgment by information plainly provided in the balance sheet and the P&L statement, what is Lockheed Martin's FY2020 asset turnover ratio? Asset turnover ratio is defined as: FY2020 revenue / (average total assets between FY2019 and FY2020). Round your answer to two decimal places.

Result: OK · Latency: 24.3 s · Tool calls: 2

Expected answer

1.33

Delivered answer

1.33

Lockheed Martin's FY2020 asset turnover ratio is **1.33**, computed from the figures in its 2020 Form 10-K (fiscal year ended December 31, 2020):

- FY2020 total net sales (revenue): **\$65,398M**
- FY2020 total assets: **\$50,710M**
- FY2019 total assets: **\$47,528M**

Average total assets = $(\$50,710 + \$47,528) / 2 = \mathbf{\$49,119M}$

Asset turnover = $\$65,398 / \$49,119 = \mathbf{1.33}$ (rounded to two decimals)

Cited documents: LOCKHEEDMARTIN_2020_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (1.33) and includes the correct calculation steps without contradictions.

Q102. What is Lockheed Martin's FY2021 net working capital? Define net working capital as total current assets less total current liabilities. Answer in USD millions. Respond to the question by assuming the perspective of an investment analyst who can only use the details shown within the balance sheet.

Result: OK · Latency: 19.8 s · Tool calls: 1

Expected answer

\$5818.00

Delivered answer

Lockheed Martin's FY2021 net working capital (total current assets less total current liabilities) is **\$5,818 million**.

- Total current assets at Dec 31, 2021: \$19,815M
- Total current liabilities: \$13,997M
- Net working capital: $19,815 - 13,997 = 5,818$

From the FY2021 consolidated balance sheet, current assets comprise cash (\$3,604M), receivables (\$1,963M), contract assets (\$10,579M), inventories (\$2,981M), and other current assets (\$688M); current liabilities are led by contract liabilities (\$8,107M) and salaries/benefits (\$3,108M).

Cited documents: LOCKHEEDMARTIN_2021_10K.pdf

Judge rationale: The agent's answer matches the expected net working capital figure of \$5,818 million and provides correct supporting calculations.

Q103. What is Lockheed Martin's 2 year total revenue CAGR from FY2020 to FY2022 (in units of percents and round to one decimal place)? Provide a response to the question by primarily using the statement of income.

Result: OK · Latency: 36.1 s · Tool calls: 2

Expected answer

0.4%

Delivered answer

0.4%

Using Lockheed Martin's Consolidated Statements of Earnings (FY2022 10-K):

- FY2020 net sales (total revenue): \$65,398 million
- FY2022 net sales (total revenue): \$65,984 million

2-year CAGR = $(65,984 / 65,398)^{(1/2)} - 1 = (1.00896)^{(0.5)} - 1 = 1.00447 - 1 = 0.447\% \approx 0.4\%$

Cited documents: LOCKHEEDMARTIN_2022_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value and precision.

Q104. Basing your judgments off of the balance sheet, what is the year end FY2018 amount of accounts payable for MGM Resorts? Answer in USD millions.

Result: OK · Latency: 22.1 s · Tool calls: 3

Expected answer

\$303.00

Delivered answer

\$303 million.

MGM Resorts' consolidated balance sheet as of December 31, 2018 reports accounts payable of \$302,578 thousand, i.e. \$303 million (rounded to the nearest whole million).

Cited documents: MGMRESORTS_2018_10K.pdf

Judge rationale: The agent's answer matches the expected amount of \$303.00 million when rounded to the nearest whole million.

Q105. What is the FY2018 - FY2020 3 year average of capex as a % of revenue for MGM Resorts? Answer in units of percents and round to one decimal place. Please utilize information provided primarily within the statement of cash flows and the statement of income.

Result: OK · Latency: 33.4 s · Tool calls: 2

Expected answer

7.9%

Delivered answer

7.9%. That's the 3-year average of capex as a % of revenue for MGM Resorts across FY2018–FY2020.

Capex (from the statement of cash flows / segment capital expenditures) and net revenues per year:

- FY2018: \$1,486.8M capex / \$11,763.1M revenue = 12.6%
- FY2019: \$739.0M / \$12,899.7M = 5.7%
- FY2020: \$270.6M / \$5,162.1M = 5.2%

Average = (12.64% + 5.73% + 5.24%) / 3 = 7.9%.

Cited documents: MGMRESORTS_2020_10K.pdf, MGMRESORTS_2019_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value, precision, and units, with no contradictions.

Q106. Has MGM Resorts paid dividends to common shareholders in FY2022?

Result: **OK** · Latency: 11.1 s · Tool calls: 1

Expected answer

Yes. MGM maintained 0.01\$ per share annual dividend through out FY 2022.

Delivered answer

Yes. MGM Resorts paid dividends to common shareholders in FY2022 (fiscal year ended December 31, 2022).

- Dividends paid to common shareholders: **\$4.0 million** (\$4,048 thousand) per the consolidated statement of cash flows (financing activities).
- The company paid \$0.0025 per share in each of March, June, September, and December 2022, totaling ~\$4 million for the year.

This was the reduced dividend level the company maintained after cutting its annual dividend to \$0.01 per share in Q2 2020 due to the pandemic; the dividend was subsequently suspended in February 2023 in favor of share repurchases.

Cited documents: MGMRESORTS_2022_10K.pdf

Judge rationale: The agent's answer confirms dividends were paid in FY2022 and provides additional correct details that do not contradict the expected answer.

Q107. Which region had the Highest EBITDAR Contribution for MGM during FY2022?

Result: **OK** · Latency: 23.1 s · Tool calls: 1

Expected answer

Las Vegas resorts contributed ~90% of company level EBITDAR during FY2022.

Delivered answer

The region with the highest EBITDAR contribution for MGM Resorts in FY2022 was **Las Vegas Strip Resorts**.

- Las Vegas Strip Resorts: \$3,142 million Adjusted Property EBITDAR
- Regional Operations: \$1,295 million
- MGM China: \$(203) million (a loss)

The Las Vegas segment's \$3.1 billion contribution dwarfed the others, driven by the inclusion of The Cosmopolitan and a full year of Aria results, while MGM China was dragged down by Macau travel and entry restrictions.

Cited documents: MGMRESORTS_2022_10K.pdf

Judge rationale: The agent's answer correctly identifies Las Vegas Strip Resorts as the region with the highest EBITDAR contribution, aligning with the essential information in the expected answer.

Q108. What was MGM's interest coverage ratio using FY2022 Adjusted EBIT as the numerator and annual Interest Expense as the denominator?

Result: **NOK** · Latency: 59.4 s · Tool calls: 5

Expected answer

As adjusted EBIT is negative, coverage ratio is zero

Delivered answer

5.88x

MGM does not report a standalone "Adjusted EBIT"; its published non-GAAP EBIT-type measure is **Adjusted EBITDAR**, which I use as the numerator.

- FY2022 Adjusted EBITDAR: **\$3,497.3M** (\$3,497,254K)
- FY2022 Interest expense, net of amounts capitalized: **\$595.0M** (\$594,954K)
- Coverage = $3,497,254 \div 594,954 = 5.88x$

(If you instead intended a true "Adjusted EBIT" = Adjusted EBITDAR less D&A of \$3,482.1M, the numerator would be only ~\$15.2M and the ratio ~0.03x — almost certainly not the intended reading.)

Cited documents: MGMRESORTS_2022Q4_EARNINGS.pdf

Judge rationale: The agent's answer provides a positive ratio (5.88x) while the expected answer states the ratio is zero due to negative Adjusted EBIT.

Q109. Which region had the worst topline performance for MGM during FY2022?

Result: OK · Latency: 30.5 s · Tool calls: 1

Expected answer

MGM China experienced the worst topline performance amongst the other regions presented. Its revenue declined 44% in FY2022 whereas the other regions presented increased their revenues.

Delivered answer

MGM China had the worst topline performance in FY2022, with net revenues declining 44% year-over-year.

From the 10-K (fiscal year ended December 31, 2022):

- Las Vegas Strip Resorts: +77% (\$4.7B → \$8.4B)
- Regional Operations: +12% (\$3.4B → \$3.8B)
- **MGM China: -44% (\$1.2B → \$674M)**

The decline was driven by travel and entry restrictions in Macau, including a three-day COVID-19 related property closure at MGM Cotai in Q4 2022.

Cited documents: MGMRESORTS_2022_10K.pdf

Judge rationale: The agent's answer correctly identifies MGM China as the region with the worst topline performance and includes the essential 44% decline figure, with no contradictions.

Q110. Which type of debt received the largest investment among the short term investments for MGM in H1 FY2023?

Result: OK · Latency: 12.3 s · Tool calls: 1

Expected answer

the biggest short term investment is in corporate bonds (almost 82% of the total investment)

Delivered answer

Corporate bonds received the largest investment among MGM's short-term debt investments in H1 FY2023 (as of June 30, 2023).

From the debt investments table in the Q2 2023 10-Q:

Short-term investment type	Amount (thousands)
Corporate bonds	\$416,420
U.S. government securities	\$57,696
U.S. agency securities	\$29,049
Commercial paper and CDs	\$4,561
Total	\$507,726

Corporate bonds accounted for roughly 82% of the total short-term investment portfolio at \$416.4 million.

Cited documents: MGMRESORTS_2023Q2_10Q.pdf

Judge rationale: The agent's answer correctly identifies corporate bonds as the largest short-term investment and confirms the ~82% figure, matching the essential information in the expected answer.

Q111. What is the FY2016 COGS for Microsoft? Please state answer in USD millions. Provide a response to the question by primarily using the statement of income.

Result: OK · Latency: 14.9 s · Tool calls: 1

Expected answer

\$32780.00

Delivered answer

Microsoft's FY2016 cost of goods sold (labelled "cost of revenue") was **\$32,780 million**. This is the fiscal year ended June 30, 2016.

Calculated from the statement of income: Cost of revenue = Revenue – Gross margin = \$85,320M – \$52,540M = \$32,780M, consistent with the 1% decline from the prior year (\$33,038M in FY2015).

Cited documents: MICROSOFT_2016_10K.pdf

Judge rationale: The agent's answer matches the expected COGS figure of \$32,780 million in USD millions and includes no contradictions.

Q112. Has Microsoft increased its debt on balance sheet between FY2023 and the FY2022 period?

Result: OK · Latency: 68.0 s · Tool calls: 3

Expected answer

No. Microsoft decreased its debt by \$2.5bn in FY 2023 vs FY 2022.

Delivered answer

No — Microsoft's debt on the balance sheet **declined** between FY2022 and FY2023, not increased.

Total debt (long-term debt including current portion) fell from \$49,781 million at June 30, 2022 to \$47,237 million at June 30, 2023, a decrease of \$2,544 million (about 5.1%). Long-term debt excluding the current portion likewise dropped from \$47,032 million to \$41,990 million.

Cited documents: MICROSOFT_2023_10K.pdf

Judge rationale: The agent's answer correctly states that Microsoft's debt decreased and provides the exact figures matching the expected answer's essential information.

Q113. We want to calculate a financial metric. Please help us compute it by basing your answers off of the statement of income and the statement of cash flows. Here's the question: what is the FY2015 unadjusted EBITDA % margin for Netflix? Calculate unadjusted EBITDA using unadjusted operating income and D&A (from cash flow statement).

Result: OK · Latency: 12.0 s · Tool calls: 1

Expected answer

5.4%

Delivered answer

The FY2015 unadjusted EBITDA margin for Netflix is **5.4%**.

Unadjusted EBITDA = unadjusted operating income + D&A = \$305,826k + \$62,283k = \$368,109k (D&A from the cash flow statement: "Depreciation and amortization of property, equipment and intangibles"; content amortization is an operating cost, not included). EBITDA margin = \$368,109k / \$6,779,511k revenue = 5.4%.

Cited documents: NETFLIX_2015_10K.pdf

Judge rationale: The agent's answer matches the expected EBITDA margin of 5.4% and includes the correct calculation method without contradictions.

Q114. What is Netflix's year end FY2017 total current liabilities (in USD millions)? Base your judgments on the information provided primarily in the balance sheet.

Result: OK · Latency: 10.5 s · Tool calls: 1

Expected answer

\$5466.00

Delivered answer

\$5,466.3 million

From Netflix's FY2017 consolidated balance sheet (as of December 31, 2017), total current liabilities are reported as \$5,466,312 thousand, which equals \$5,466.3 million.

The components that make up this total:

- Current content liabilities: \$4,173,041K
- Accounts payable: \$359,555K
- Accrued expenses: \$315,094K
- Deferred revenue: \$618,622K
- **Total: \$5,466,312K = \$5,466.3M**

Cited documents: NETFLIX_2017_10K.pdf

Judge rationale: The agent's answer (\$5,466.3 million) matches the expected answer (\$5466.00 million) when rounded to the same precision.

Q115. We need to calculate a reasonable approximation (or exact number if possible) of a financial metric. Basing your judgment by information plainly provided in the statement of income, what is Nike's three year average of cost of goods sold as a % of revenue from FY2016 to FY2018? Answer in units of percents and round to one decimal place.

Result: OK · Latency: 14.7 s · Tool calls: 1

Expected answer

55.1%

Delivered answer

55.1%

Cost of sales as a % of revenue: FY2016 = $17,405/32,376 = 53.8\%$, FY2017 = $19,038/34,350 = 55.4\%$, FY2018 = $20,441/36,397 = 56.2\%$. Three-year average = $(53.76\% + 55.42\% + 56.16\%)/3 = 55.1\%$.

Cited documents: NIKE_2018_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly, including the correct rounding to one decimal place.

Q116. According to the details clearly outlined within the balance sheet, how much total current assets did Nike have at the end of FY2019? Answer in USD millions.

Result: OK · Latency: 16.7 s · Tool calls: 1

Expected answer

\$16525.00

Delivered answer

Nike's total current assets at the end of FY2019 (May 31, 2019) were **\$16,525 million**.

That's the balance-sheet total of cash and equivalents (\$4,466M), short-term investments (\$197M), accounts receivable net (\$4,272M), inventories (\$5,622M), and prepaid expenses and other current assets (\$1,968M).

Cited documents: NIKE_2019_10K.pdf

Judge rationale: The agent's answer matches the expected figure of \$16,525 million in USD millions.

Q117. When primarily referencing the income statement and the statement of financial position, what is the FY2021 inventory turnover ratio for Nike? Inventory turnover ratio is defined as: (FY2021 COGS) / (average inventory between FY2020 and FY2021). Round your answer to two decimal places.

Result: OK · Latency: 33.6 s · Tool calls: 3

Expected answer

3.46

Delivered answer

3.46

FY2021 inventory turnover = FY2021 Cost of sales ÷ average inventory between FY2020 and FY2021 = \$24,576m ÷ (((\$6,854m + \$7,367m) ÷ 2) = \$24,576m ÷ \$7,110.5m = 3.46.

Cited documents: NIKE_2021_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly, including the correct calculation and rounding to two decimal places.

Q118. Among operations, investing, and financing activities, which brought in the most (or lost the least) cash flow for Nike in FY2023?

Result: **OK** · Latency: 13.4 s · Tool calls: 1

Expected answer

Among the three, cash flow from operations was the highest for Nike in FY2023.

Delivered answer

Operating activities brought in the most cash for Nike in FY2023 (year ended May 31, 2023).

- Cash provided by operations: **\$5,841 million** (inflow)
- Cash provided by investing activities: **\$564 million** (inflow)
- Cash used by financing activities: **-\$7,447 million** (outflow)

So operating was the biggest source at \$5,841 million, investing contributed a much smaller \$564 million, and financing was the largest cash use at \$7,447 million (driven by \$5,480 million of share repurchases and \$2,012 million of dividends).

Cited documents: NIKE_2023_10K.pdf

Judge rationale: The agent's answer correctly identifies operating activities as the highest cash flow source for Nike in FY2023, matching the essential information in the expected answer.

Q119. Does Paypal have positive working capital based on FY2022 data? If working capital is not a useful or relevant metric for this company, then please state that and explain why.

Result: **NOK** · Latency: 35.5 s · Tool calls: 1

Expected answer

Yes. Paypal has a positive working capital of \$ 1.6Bn as of FY2022 end.

Delivered answer

Yes—PayPal had positive working capital of \$12.4 billion at FY2022 year-end.

- Total current assets: \$57,517 million
- Total current liabilities: \$45,101 million
- Working capital = \$57,517 – \$45,101 = **+\$12,416 million**

PayPal is a digital payments and fintech company rather than a conventional bank or insurer, so working capital isn't the metric its performance is measured on—its current items are dominated by customer funds (funds receivable/customer accounts of \$36.4B offset by funds payable/amounts due to customers of \$40.1B), which distort the ratio. Even so, the standard working capital figure is clearly positive.

Cited documents: PAYPAL_2022_10K.pdf

Judge rationale: The agent's working capital figure (\$12.4B) contradicts the expected answer (\$1.6B).

Q120. What is the FY2021 capital expenditure amount (in USD billions) for PepsiCo? Respond to the question by assuming the perspective of an investment analyst who can only use the details shown within the statement of cash flows.

Result: OK · Latency: 22.4 s · Tool calls: 1

Expected answer

\$4.60

Delivered answer

\$4.6 billion (precisely \$4.625 billion, or \$4,625 million).

Reading straight from the Consolidated Statement of Cash Flows in PepsiCo's FY2021 10-K (fiscal year ended December 25, 2021), the "Capital spending" line under Investing Activities is reported as \$(4,625) million. Converting to billions: $\$4,625\text{M} \div 1,000 = \$4.625\text{B} \approx \textbf{\$4.6B}$.

Note this is the gross capital spending figure as shown in the cash flow statement; it is not net of the \$166M in sales of property, plant and equipment that appears as a separate line.

Cited documents: PEPSICO_2021_10K.pdf

Judge rationale: The agent's answer matches the expected answer's essential figure (\$4.6 billion) when rounded to the same precision.

Q121. What are the geographies that Pepsico primarily operates in as of FY2022?

Result: OK · Latency: 77.5 s · Tool calls: 5

Expected answer

As of FY2022, Pepsico primarily operates in the following geographies: North America, Latin America, Europe, Africa, Middle East, South Asia, Asia Pacific, Australia, New Zealand and China.

Delivered answer

As of FY2022 (fiscal year ended December 31, 2022), PepsiCo operates through seven reportable segments that are themselves geographic: **North America** (three segments — Frito-Lay North America, Quaker Foods North America, and PepsiCo Beverages North America, covering the U.S. and Canada), **Latin America, Europe, Africa, Middle East and South Asia (AMESA)**, and **Asia Pacific, Australia and New Zealand and China Region (APAC)**.

Net revenue by segment for FY2022 was PBNA \$26,213M, FLNA \$23,291M, Europe \$12,724M, LatAm \$9,779M, AMESA \$6,438M, APAC \$4,787M, and QFNA \$3,160M, for a total of \$86,392M.

Cited documents: PEPSICO_2022Q4_EARNINGS.pdf, PEPSICO_2022_10K.pdf

Judge rationale: The agent's answer includes all essential geographies listed in the expected answer, with no contradictions or factual errors.

Q122. Has Pepsico reported any materially important ongoing legal battles from FY2022 and FY2021?

Result: OK · Latency: 51.4 s · Tool calls: 4

Expected answer

No, Pepsico is not involved in material legal battles.

Delivered answer

No. In both fiscal years, PepsiCo's Item 3 (Legal Proceedings) disclosed that it and its subsidiaries are party to a variety of litigation, claims, legal or regulatory proceedings, inquiries and investigations, but management stated that the final outcome of these matters will not have a material adverse effect on the company's financial condition, results of operations, or cash flows.

- **FY2022** (10-K, fiscal year ended December 31, 2022): "We and our subsidiaries are party to a variety of litigation, claims, legal or regulatory proceedings, inquiries and investigations. Management believes that the final outcome of the foregoing will not have a material adverse effect on our financial condition, results of operations or cash flows."
- **FY2021** (10-K, fiscal year ended December 25, 2021): same language, adding that the matters relate to advertising, marketing or commercial practices, product labels, claims and ingredients, personal injury and property damage, intellectual property rights, privacy, employment, tax and insurance matters, environmental, social and governance matters, and compliance with applicable laws and regulations.

So while PepsiCo is involved in routine, ordinary-course litigation, it did not flag any single matter as materially important in either year.

Cited documents: PEPSICO_2022_10K.pdf, PEPSICO_2021_10K.pdf

Judge rationale: The agent's answer correctly states that Pepsico has no materially important ongoing legal battles, matching the essential information in the expected answer.

Q123. What is the quantity of restructuring costs directly outlined in PepsiCo's income statements for FY2022? If restructuring costs are not explicitly outlined then state 0.

Result: **NOK** · Latency: 15.7 s · Tool calls: 1

Expected answer

Pepsico's restructuring costs in FY2022 amounted to \$411 million .

Delivered answer

0.

PepsiCo's Consolidated Statement of Income for FY2022 (year ended December 31, 2022) does not include a separate line item for restructuring costs. The income statement line items are: Net Revenue, Cost of sales, Gross profit, Selling, general and administrative expenses, Gain associated with the Juice Transaction, Impairment of intangible assets, Operating Profit, and below-the-line items.

Restructuring and impairment charges (\$411 million for FY2022) are disclosed only in the cash flow statement and in the non-GAAP reconciliation tables, where they are shown as embedded within cost of sales and SG&A rather than as a standalone income statement line.

Cited documents: PEPSICO_2022Q4_EARNINGS.pdf

Judge rationale: The agent answer incorrectly states 0 and contradicts the expected answer by claiming restructuring costs are not explicitly outlined in the income statement, while the reference confirms \$411 million.

Q124. What is the FY2022 unadjusted EBITDA less capex for PepsiCo? Define unadjusted EBITDA as unadjusted operating income + depreciation and amortization [from cash flow statement]. Answer in USD millions. Respond to the question by assuming the perspective of an investment analyst who can only use the details shown within the statement of cash flows and the income statement.

Result: OK · Latency: 16.2 s · Tool calls: 1

Expected answer

\$9068.00

Delivered answer

\$9,068 million.

Here's the build-up from PepsiCo's FY2022 10-K (fiscal year ended December 31, 2022):

- Operating Profit (income statement): \$11,512M
- Depreciation & amortization (cash flow statement): \$2,763M
- **Unadjusted EBITDA** = 11,512 + 2,763 = **\$14,275M**
- Capital spending (investing section): \$5,207M
- **EBITDA less capex** = 14,275 – 5,207 = **\$9,068M**

Cited documents: PEPSICO_2022_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly in value and units, with no contradictions or omissions.

Q125. What is the FY2022 unadjusted EBITDA % margin for PepsiCo? Calculate unadjusted EBITDA using unadjusted operating income and D&A (from cash flow statement). Give a response to the question by relying on the details shown in the statement of cash flows and the P&L statement.

Result: OK · Latency: 17.4 s · Tool calls: 1

Expected answer

16.5%

Delivered answer

16.52%

From PepsiCo's FY2022 10-K (fiscal year ended December 31, 2022):

- Operating Profit (P&L): \$11,512M
- Depreciation & Amortization (Cash Flow Statement): \$2,763M
- Net Revenue (P&L): \$86,392M

Unadjusted EBITDA = \$11,512M + \$2,763M = **\$14,275M**

EBITDA % margin = \$14,275M ÷ \$86,392M = **16.52%**

Cited documents: PEPSICO_2022_10K.pdf

Judge rationale: The agent's answer (16.52%) matches the expected answer (16.5%) when rounded to the same precision.

Q126. At the PepsiCo AGM held on May 3, 2023, what was the outcome of the shareholder vote on the shareholder proposal for a congruency report by PepsiCo on net-zero emissions policies?

Result: OK · Latency: 14.9 s · Tool calls: 1

Expected answer

The shareholder proposal for a congruency report by PepsiCo on net-zero emissions policies was defeated.

Delivered answer

Defeated. At PepsiCo's 2023 Annual Meeting of Shareholders (held May 3, 2023), the shareholder proposal for a congruency report on net-zero emissions policies was rejected by a wide margin.

- For: 19,718,780
- Against: 977,228,788
- Abstain: 23,200,321
- Broker Non-Votes: 172,969,325

On the votes cast (For + Against = 996,947,568), the proposal received only about 1.98% support, so it failed decisively.

Cited documents: PEPSICO_2023_8K_dated-2023-05-05.pdf

Judge rationale: The agent's answer correctly states that the proposal was defeated, matching the essential information in the expected answer.

Q127. By how much did PepsiCo increase its unsecured five year revolving credit agreement on May 26, 2023?

Result: OK · Latency: 40.5 s · Tool calls: 1

Expected answer

\$400,000,000 increase.

Delivered answer

PepsiCo increased its five-year unsecured revolving credit agreement by \$400 million on May 26, 2023. The prior 2022 agreement was \$3.8 billion; the new 2023 Five Year Credit Agreement is \$4.2 billion (\$4,200,000,000 – \$3,800,000,000).

Cited documents: PEPSICO_2023_8K_dated-2023-05-30.pdf

Judge rationale: The agent's answer correctly states the \$400 million increase, matching the essential information in the expected answer.

Q128. As of May 26, 2023, what is the total amount PepsiCo may borrow under its unsecured revolving credit agreements?

Result: OK · Latency: 8.9 s · Tool calls: 1

Expected answer

Total amount PepsiCo may borrow under unsecured revolving credit agreements = \$8,400,000,000.

Delivered answer

\$8.4 billion.

As of May 26, 2023, PepsiCo's revolving borrowing capacity was split across two unsecured agreements, each carrying a \$4.2 billion commitment: a new 364-day credit agreement and a new five-year credit agreement (both dated May 26, 2023, replacing the two \$3.8 billion facilities from May 27, 2022). \$4.2 billion + \$4.2 billion = \$8.4 billion. Each agreement may later be raised to \$4.95 billion with lender consent.

Cited documents: PEPSICO_2023_8K_dated-2023-05-30.pdf

Judge rationale: The agent's answer matches the essential information of the expected answer (\$8.4 billion = \$8,400,000,000) and contains no contradictions.

Q129. As of FY2023Q1, why did PepsiCo raise full year guidance for FY2023?

Result: OK · Latency: 13.3 s · Tool calls: 1

Expected answer

PepsiCo experienced a strong start to FY2023.

Delivered answer

PepsiCo raised its full-year FY2023 guidance because of a **strong start to the year**, with its categories and geographies remaining resilient during Q1 2023.

Specifically, CEO Ramon Laguarta stated: *"Given our strong start to the year, we now expect our full-year 2023 organic revenue to increase 8 percent (previously 6 percent) and core constant currency EPS to increase 9 percent (previously 8 percent)."*

The Q1 results that drove the raise:

- **Organic revenue growth of 14.3%** (vs. 6% prior full-year guide)
- **Core constant currency EPS growth of 18%** (vs. 8% prior full-year guide)
- **Core EPS of \$1.50** for the quarter

The updated full-year targets implied core EPS of **\$7.27** (up from the prior \$7.20), a 7% increase vs. 2022 core EPS of \$6.79 (previously guided at 6%).

Cited documents: PEPSICO_2023Q1_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential information that PepsiCo raised guidance due to a strong start to FY2023, matching the expected answer.

Q130. As of FY2023Q1, by how many percentage points did Pepsico raise full year guidance in respect of core constant currency EPS growth?

Result: OK · Latency: 13.7 s · Tool calls: 1

Expected answer

Pepsico raised full year guidance in respect of core constant currency EPS growth by 1 percentage point.

Delivered answer

1 percentage point. In its FY2023Q1 earnings release, PepsiCo raised its full-year 2023 core constant currency EPS growth guidance from 8% to 9% (previously 8%). Core EPS implied also moved up to \$7.27 from \$7.20.

Cited documents: PEPSICO_2023Q1_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential information (1 percentage point) and provides additional correct details without contradiction.

Q131. Did Pfizer grow its PPNE between FY20 and FY21?

Result: OK · Latency: 15.7 s · Tool calls: 1

Expected answer

Yes, change in PPNE was positive year over year

Delivered answer

Yes. Pfizer's PPNE grew from \$13,745M at the end of FY2020 to \$14,882M at the end of FY2021, an increase of \$1,137M (+8.3%).

The figures come from Note 9 (Property, Plant and Equipment) of the 2021 Form 10-K, which reports both year-ends on a consistent basis. The growth was driven primarily by higher machinery and equipment (\$11,000M → \$12,252M) and construction in progress (\$3,481M → \$3,822M), partially offset by a small decline in land and furniture/fixtures.

Cited documents: PFIZER_2021_10K.pdf

Judge rationale: The agent's answer confirms the positive year-over-year change in PPNE, matching the essential information in the expected answer.

Q132. Were there any potential events that are not in Pfizer's standard business operations that substantially increased net income in 2019?

Result: OK · Latency: 31.4 s · Tool calls: 1

Expected answer

Yes, the gain on completion of Consumer Healthcare JV Transaction

Delivered answer

Yes — Pfizer recognized a one-time gain outside its ordinary operations that substantially boosted 2019 net income: the **completion of the GSK Consumer Healthcare joint venture transaction**, which produced a pre-tax gain of **\$8,086 million** (about \$8.1 billion) in 2019 (recorded as "(Gain) on completion of Consumer Healthcare JV transaction").

Set against 2019 net income attributable to Pfizer Inc. of \$16,273 million, this single non-operating item was the largest driver of the year's profit. The segment note lists it, along with other, much smaller non-routine "certain significant items" such as net gains of \$415 million on equity securities, \$2.8 billion of asset impairment charges and \$758 million of restructuring charges (all in the opposite direction, as costs). Because this gain was excluded from Pfizer's non-GAAP "Adjusted income," management's recurring, operating earnings (Adjusted diluted EPS \$2.95) strip it out — confirming the \$8.1 billion JV gain was the one-time, non-operating item behind the sharply higher GAAP net income.

Cited documents: PFIZER_2019_10K.pdf

Judge rationale: The agent's answer includes the essential information (the gain on completion of the Consumer Healthcare JV Transaction) and does not contradict the expected answer.

Q133. What are three main companies acquired by Pfizer mentioned in this 10K report? Added note by ATG for disambiguation: the question is referring to FY2021.

Result: **NOK** · Latency: 231.4 s · Tool calls: 12

Expected answer

Trillium, Array, and Therachon

Delivered answer

For FY2021, the three main companies named in Pfizer's 10-K as acquisition targets are **Trillium, Arena, and Biohaven**:

- **Trillium** — completed on November 17, 2021 for \$18.50 per share, total consideration of **\$2.0 billion** net of cash acquired. A clinical-stage immuno-oncology company (accounted for as an asset acquisition, with a \$2.1 billion IPR&D charge).
- **Arena** — agreed in December 2021 at \$100 per share in an all-cash deal for a total equity value of approximately **\$6.7 billion**; shareholder approval came in February 2022, with closing targeted for the first half of 2022.
- **Biohaven** — a November 2021 collaboration and license agreement under which Pfizer acquired rights to commercialize rimegepant and zavegepant (migraine) outside the U.S.; closed January 4, 2022 with a \$500 million payment (\$150 million upfront plus a \$350 million equity investment).

Note that only Trillium had closed as of the FY2021 filing; Arena was agreed but closed in 2022, and Biohaven was a rights/collaboration deal rather than a full company acquisition. The filing also mentions a smaller unnamed \$177 million asset acquisition closed in Q2 2021.

Cited documents: PFIZER_2021_10K.pdf

Judge rationale: The agent's answer lists Arena and Biohaven instead of Array and Therachon, which contradicts the expected answer.

Q134. How much does Pfizer expect to pay to spin off Upjohn in the future in USD million? Added note by ATG for disambiguation: the question is referring to Q2 of FY2023.

Result: OK (human review) · Latency: 50.9 s · Tool calls: 4

Expected answer

77.78

Delivered answer

\$70 million

Pfizer's Q2 FY2023 10-Q (quarter ended July 2, 2023) states that it expects to incur total costs of approximately **\$700 million** in connection with separating Upjohn, of which approximately **90%** had already been incurred since inception through Q2 2023.

- Total expected cost: \$700 million
- Already incurred (~90%): \$630 million
- **Remaining (future) cost: \$700 million – \$630 million = \$70 million**

So Pfizer expected to pay roughly **\$70 million** more to complete the Upjohn spin-off.

Cited documents: Pfizer_2023Q2_10Q.pdf

Judge verdict: NOK · The agent's answer (70 million) contradicts the expected answer (77.78 million).

Human review: OK · The report states that approximately 90% of the \$700 million has already been spent since inception and through the second quarter of 2023. Therefore, there remains $700 \times (1 - 0.9) = 70$ (approximately)

Q135. For Pfizer, which geographic region had the biggest drop in Q22023 year over year revenues (on a percentage basis)?

Result: OK · Latency: 13.2 s · Tool calls: 1

Expected answer

Developed Rest of the World

Delivered answer

Developed Rest of World had the biggest drop in Q2 2023 year-over-year revenues, falling **74%** (from \$5,034M in Q2 2022 to \$1,305M in Q2 2023).

For context, here's how all regions compared:

Region	Q2 2023	Q2 2022	% Change
U.S.	\$6,185M	\$11,222M	-45%
Developed Europe	\$2,415M	\$5,480M	-56%
Developed Rest of World	\$1,305M	\$5,034M	-74%
Emerging Markets	\$2,828M	\$6,006M	-53%

The decline was driven primarily by the collapse in Comirnaty (-82%) and Paxlovid (-98%) revenues in that region, as Pfizer transitioned away from peak pandemic-era vaccine and treatment deliveries.

Cited documents: PFIZER_2023Q2_EARNINGS.pdf

Judge rationale: The agent's answer correctly identifies 'Developed Rest of World' as the region with the biggest drop in Q2 2023 year-over-year revenues, matching the essential information in the expected answer.

Q136. As of Q2'2023, is Pfizer spinning off any large business segments?

Result: OK (human review) · Latency: 25.7 s · Tool calls: 2

Expected answer

Yes, it's spinning off Upjohn.

Delivered answer

No. As of Q2 2023, Pfizer is not spinning off any large business segments — its two operating segments, Biopharma (its only reportable segment) and Business Innovation, remain in place.

The large spin-offs already happened in prior years: the Upjohn Business was spun off and combined with Mylan into Viatris in November 2020, and the Consumer Healthcare business was combined into a joint venture in 2019 that GSK demerged as Haleon in July 2022. The only divestiture announced in Q2 2023 was a small, targeted one — in July 2023 Pfizer agreed that Alexion (AstraZeneca) would purchase/license its early-stage rare disease gene therapy portfolio, consistent with Pfizer's shift away from viral capsid-based gene therapy — not a large business segment.

Cited documents: PFIZER_2022_10K.pdf, Pfizer_2023Q2_10Q.pdf

Judge verdict: NOK · The agent answer contradicts the expected answer by stating there is no spin-off, while the reference confirms Upjohn was spun off.

Human review: OK · The spin-off of Upjohn was completed in November 2020; by 2023, only residual costs remain. The agent's response ("No") is factually correct.

Q137. Which debt securities are registered to trade on a national securities exchange under Ulta Beauty's name as of FY2023?

Result: OK · Latency: 20.8 s · Tool calls: 2

Expected answer

There are none

Delivered answer

None.

Ulta Beauty's FY2023 Form 10-K (fiscal year ended January 28, 2023) lists only one class of securities registered to trade on a national securities exchange under Section 12(b): its common stock, par value \$0.01 per share (trading symbol ULTA), on The NASDAQ Global Select Market. No debt securities appear in the Section 12(b) table, and the filing reports "None" under Section 12(g).

So, as of FY2023, Ulta Beauty had no debt securities registered to trade on a national securities exchange.

Cited documents: ULTABEAUTY_2023_10K.pdf

Judge rationale: The agent's answer correctly states that there are no debt securities registered to trade on a national securities exchange under Ulta Beauty's name as of FY2023, matching the expected answer.

Q138. What are major acquisitions that Ulta Beauty has done in FY2023 and FY2022?

Result: OK · Latency: 45.9 s · Tool calls: 3

Expected answer

Ulta Beauty did not make any acquisitions in FY2023 and FY2022.

Delivered answer

Ulta Beauty did not make any major acquisitions in FY2023 or FY2022.

The consolidated statement of cash flows in the 10-K shows "Acquisitions, net of cash acquired" as **\$0** for both the fiscal year ended January 28, 2023 (FY2023) and the fiscal year ended January 29, 2022 (FY2022). The goodwill note confirms this, showing no acquisition-related additions to goodwill in either period.

The only acquisition Ulta Beauty reports in the three-year window is a small **\$1.2 million** deal in the fiscal year ended January 30, 2021 (FY2021), which is not material. Ulta's growth strategy during this period was driven by organic store expansion (47 new stores in FY2023, 44 in FY2022) and the Ulta Beauty at Target partnership, not by M&A.

Cited documents: ULTABEAUTY_2023_10K.pdf

Judge rationale: The agent's answer correctly states that Ulta Beauty made no acquisitions in FY2023 and FY2022, matching the essential information in the expected answer.

Q139. What drove the reduction in SG&A expense as a percent of net sales in FY2023? Added note by ATG for disambiguation: the question is referring to Ulta Beauty.

Result: OK · Latency: 25.8 s · Tool calls: 2

Expected answer

Lower marketing expenses and leverage of incentive compensation due to higher sales. The answer here assumes FY2023 refers to the 12 months ended on January 28, 2023 (although the company refers to this period as its fiscal 2022).

Delivered answer

The reduction was driven by **lower marketing expenses and leverage of incentive compensation due to higher sales**, partially offset by deleverage of corporate overhead (strategic investments) and store payroll and benefits (wage investments).

For the fiscal year ended January 28, 2023 (Ulta's FY2023), SG&A fell to 23.5% of net sales from 23.9% in the prior year — \$2,395.3m on net sales of \$10,208.6m vs. \$2,061.5m on \$8,630.9m a year earlier. Management attributed the 0.4-point improvement to the marketing and incentive-compensation leverage, which outweighed the drag from higher corporate and store-payroll spending.

Cited documents: ULTABEAUTY_2023Q4_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential information from the expected answer (lower marketing expenses and leverage of incentive compensation due to higher sales) and does not contradict it.

Q140. What drove the increase in Ulta Beauty's merchandise inventories balance at end of FY2023?

Result: OK · Latency: 88.2 s · Tool calls: 2

Expected answer

Increase in Merchandise inventories balance was driven by the opening of 47 new stores. The answer here assumes FY2023 refers to the 12 months ended on January 28, 2023 (although the company refers to this period as its fiscal 2022).

Delivered answer

Inventory growth, driven by store openings, new brand launches and expansions, and cost inflation. At the end of FY2023 (year ended January 28, 2023; Ulta's fiscal 2022), merchandise inventories rose \$104.2 million to \$1.6 billion from \$1.5 billion a year earlier. Management attributes the increase to the opening of 47 new stores during the year, inventory to support new brand launches and brand expansions, and higher inventory costs.

Cited documents: ULTABEAUTY_2023_10K.pdf, ULTABEAUTY_2023Q4_EARNINGS.pdf

Judge rationale: The agent's answer includes the essential driver (opening of 47 new stores) and does not contradict the expected answer.

Q141. What percent of Ulta Beauty's total spend on stock repurchases for FY 2023 occurred in Q4 of FY2023?

Result: OK · Latency: 84.9 s · Tool calls: 4

Expected answer

36%. The answer here assumes FY2023 refers to the 12 months ended on January 28, 2023 (although the company refers to this period as its fiscal 2022).

Delivered answer

36.5% of Ulta Beauty's FY2023 stock-repurchase spend occurred in Q4.

Reading "FY2023" as the fiscal year that ended in calendar 2023 (ended January 28, 2023): Ulta repurchased stock at a cost of **\$328.1 million** in the fourth quarter and **\$900.0 million** for the full year, so $328.1 \div 900.0 = 36.46\%$, or 36.5%.

Cited documents: ULTABEAUTY_2023Q4_EARNINGS.pdf

Judge rationale: The agent's answer (36.5%) matches the expected answer (36%) when rounded to the same precision.

Q142. Did Ulta Beauty's wages expense as a percent of net sales increase or decrease in FY2023?

Result: OK · Latency: 83.9 s · Tool calls: 6

Expected answer

Wages expense as a percent of net sales increased in FY2023. The answer here assumes FY2023 refers to the 12 months ended on January 28, 2023 (although the company refers to this period as its fiscal 2022).

Delivered answer

Increased.

Ulta Beauty's store payroll and benefits (the wages component of SG&A) delevered by 20 basis points as a percentage of net sales in the fiscal year ended January 28, 2023 (FY2023), driven by wage investments.

The 10-K's MD&A attributes this to deliberate wage investments in associates, which pushed the payroll-and-benefits line up relative to net sales even as overall SG&A still levered 40 bps lower (to 23.5% of net sales) thanks to marketing and incentive-comp leverage.

Cited documents: ULTABEAUTY_2023_10K.pdf

Judge rationale: The agent's answer correctly states that wages expense as a percent of net sales increased in FY2023, matching the essential information in the expected answer.

Q143. Among all of the derivative instruments that Verizon used to manage the exposure to fluctuations of foreign currencies exchange rates or interest rates, which one had the highest notional value in FY 2021?

Result: OK · Latency: 22.9 s · Tool calls: 2

Expected answer

Cross currency swaps. Its notional value was \$32,502 million.

Delivered answer

Cross currency swaps had the highest notional value in FY 2021, at \$32.5 billion (\$32,502 million).

From Verizon's 2021 10-K derivative notional table at December 31, 2021, cross currency swaps (\$32,502 million) exceeded interest rate swaps (\$19,779 million), forward starting interest rate swaps (\$1.0 billion) and foreign exchange forwards (\$932 million). Cross currency swaps fix the future interest and principal payments of Verizon's foreign-denominated debt (primarily Euro and British Pound Sterling) into U.S. dollars and mitigate foreign-currency transaction gains and losses.

Cited documents: VERIZON_2021_10K.pdf

Judge rationale: The agent's answer correctly identifies cross currency swaps as the derivative with the highest notional value in FY 2021 and matches the expected notional value of \$32,502 million.

Q144. As of FY 2021, how much did Verizon expect to pay for its retirees in 2024?

Result: OK · Latency: 45.0 s · Tool calls: 3

Expected answer

The estimated pension benefits were \$1097 million, and the estimated health care and life insurance benefits were \$862 million.

Delivered answer

As of FY 2021, Verizon expected to pay retirees **\$1,959 million** in 2024 (pension benefits of \$1,097 million plus health care and life benefits of \$862 million).

These figures come from the "Estimated Future Benefit Payments" table in Note 11 of Verizon's FY 2021 Form 10-K, which lays out expected retiree benefit payouts for each year through 2026 (plus a 2027–2031 aggregate).

Cited documents: VERIZON_2021_10K.pdf

Judge rationale: The agent's answer includes the essential figures (\$1,097 million and \$862 million) and their sum (\$1,959 million), with no contradictions or errors.

Q145. Does Verizon have a reasonably healthy liquidity profile based on its quick ratio for FY 2022? If the quick ratio is not relevant to measure liquidity, please state that and explain why.

Result: OK · Latency: 17.9 s · Tool calls: 1

Expected answer

No. The quick ratio was approximately 0.54 for Verizon. It indicated that Verizon does not have a healthy liquidity profile.

Delivered answer

No — Verizon's quick ratio for FY 2022 (fiscal year ended December 31, 2022) is below 1.0 on both usual bases, so its liquidity profile is not healthy by that measure.

The quick ratio is relevant here: Verizon is a telecommunications company, not a financial institution, so standard liquidity ratios apply.

Using the FY 2022 balance sheet (in millions):

- Cash and cash equivalents: \$2,605
- Accounts receivable, net: \$24,506
- Inventories: \$2,388
- Total current assets: \$37,857
- Total current liabilities: \$50,171
- **Strict basis** (cash + receivables) / current liabilities = $(2,605 + 24,506) / 50,171 = \mathbf{0.54}$
- **Broad basis** (current assets – inventories) / current liabilities = $(37,857 - 2,388) / 50,171 = \mathbf{0.71}$

Both fall below 1.0, so the liquidity profile is not healthy on the quick-ratio standard.

Cited documents: VERIZON_2022_10K.pdf

Judge rationale: The agent's answer includes the essential information (quick ratio ~0.54, liquidity profile not healthy) and does not contradict the expected answer.

Q146. Is Verizon a capital intensive business based on FY 2022 data?

Result: OK · Latency: 21.0 s · Tool calls: 1

Expected answer

Yes. Verizon's capital intensity ratio was approximately 2.774729. This means that it took approximately \$2.77 of assets to generate \$1 of revenue and thus, Verizon can be considered capital intensive.

Delivered answer

Yes — Verizon is a capital-intensive business in FY 2022 (fiscal year ended December 31, 2022).

- **CapEx / revenue:** $\$23,087\text{M} \div \$136,835\text{M} = 16.9\%$ of revenue spent on capital (including ~\$6.2B of C-Band 5G build-out).
- **Net fixed assets / total assets:** $\$107,434\text{M} \div \$379,680\text{M} = 28.3\%$ of the balance sheet is net property, plant & equipment (plus \$149.8B of wireless licenses).
- **Return on assets:** $\$21,748\text{M net income} \div \$379,680\text{M year-end assets} = 5.73\%$; on the two-year average asset base $((\$379,680\text{M} + \$366,596\text{M}) \div 2 = \$373,138\text{M}) = 5.83\%$.

A low ~5.7–5.8% return earned on a very large asset base, combined with ~17% of revenue going to capex, is the signature of a capital-intensive telecom.

Cited documents: VERIZON_2022_10K.pdf

Judge rationale: The agent's answer confirms Verizon is capital-intensive in FY 2022 and provides supporting financial metrics that align with the expected answer's conclusion.

Q147. Has Verizon increased its debt on balance sheet between 2022 and the 2021 fiscal period?

Result: OK · Latency: 16.8 s · Tool calls: 1

Expected answer

No. Verizon's debt decreased by \$229 million.

Delivered answer

No. Verizon's total debt on the balance sheet actually fell slightly, from \$150,868 million at December 31, 2021 to \$150,639 million at December 31, 2022 — a decrease of \$229 million (about -0.2%).

- Total debt, Dec 31, 2021: \$150,868 million
- Total debt, Dec 31, 2022: \$150,639 million
- Change: -\$229 million (-0.15%)

This is consistent with the company's own disclosure that total debt "decreased to \$150.6 billion compared to \$150.9 billion" over the period, driven by \$13.6 billion of repayments/redemptions/repurchases versus \$17.8 billion of new long-term borrowings (including \$10.7 billion of asset-backed debt).

Cited documents: VERIZON_2022_10K.pdf

Judge rationale: The agent's answer correctly states that Verizon's debt decreased and matches the essential figure of a \$229 million reduction.

Q148. What is FY2018 days payable outstanding (DPO) for Walmart? DPO is defined as: $365 \times (\text{average accounts payable between FY2017 and FY2018}) / (\text{FY2018 COGS} + \text{change in inventory between FY2017 and FY2018})$. Round your answer to two decimal places. Please base your judgments on the information provided primarily in the statement of financial position and the P&L statement.

Result: OK · Latency: 25.6 s · Tool calls: 2

Expected answer

42.69

Delivered answer

42.69 days.

Here's the calculation using Walmart's FY2018 10-K (fiscal year ended January 31, 2018):

- Average accounts payable = $(\text{FY2018 AP } \$46,092\text{M} + \text{FY2017 AP } \$41,433\text{M}) / 2 = \$43,762.5\text{M}$
- Change in inventory = $\text{FY2018 } \$43,783\text{M} - \text{FY2017 } \$43,046\text{M} = +\$737\text{M}$
- Denominator = $\text{FY2018 COGS } \$373,396\text{M} + \$737\text{M} = \$374,133\text{M}$
- $\text{DPO} = 365 \times \$43,762.5\text{M} / \$374,133\text{M} = \mathbf{42.69}$

All figures are from the consolidated balance sheet and income statement in Walmart's FY2018 10-K.

Cited documents: WALMART_2018_10K.pdf

Judge rationale: The agent's answer matches the expected answer exactly (42.69) and includes the correct calculation steps without contradictions.

Q149. Based on the information provided primarily in the statement of income, what is the FY2018 - FY2019 change in unadjusted operating income % margin for Walmart? Answer in units of percents and round to one decimal place.

Result: OK · Latency: 18.1 s · Tool calls: 1

Expected answer

0.2%

Delivered answer

0.2 percentage points (improvement).

- FY2018 operating income margin: $\$20,437\text{M} / \$495,761\text{M} = 4.12\%$
- FY2019 operating income margin: $\$21,957\text{M} / \$510,329\text{M} = 4.30\%$
- Change: $4.30\% - 4.12\% = \mathbf{+0.2\text{ pp}}$

These figures come from Walmart's Consolidated Statement of Income as reported in its FY2019 10-K (fiscal years ended January 31, 2019 and 2018).

Cited documents: WALMART_2019_10K.pdf

Judge rationale: The agent's answer correctly states the change as 0.2 percentage points, which matches the expected answer of 0.2% when rounded to one decimal place.

Q150. What is the FY2018 - FY2020 3 year average unadjusted EBITDA % margin for Walmart? Define unadjusted EBITDA as unadjusted operating income + depreciation and amortization from the cash flow statement. Answer in units of percents and round to one decimal place. Calculate what was asked by utilizing the line items clearly shown in the P&L statement and the cash flow statement.

Result: OK · Latency: 48.4 s · Tool calls: 3

Expected answer

6.2%

Delivered answer

6.2%

Using Walmart's fiscal years (ended January 31), unadjusted EBITDA = operating income + depreciation & amortization from the cash flow statement, divided by net sales:

Fiscal Year	Net Sales	Operating Income	D&A (CF)	EBITDA	EBITDA Margin
FY2018	\$495,761M	\$20,437M	\$10,529M	\$30,966M	6.25%
FY2019	\$510,329M	\$21,957M	\$10,678M	\$32,635M	6.39%
FY2020	\$519,926M	\$20,568M	\$10,987M	\$31,555M	6.07%

3-year average = (6.25% + 6.39% + 6.07%) / 3 = 6.24% → **6.2%**

All figures from the Walmart 2020 10-K consolidated statements of income and cash flows.

Cited documents: WALMART_2020_10K.pdf

Judge rationale: The agent's answer matches the expected answer (6.2%) and includes correct calculations and rounding.

Appendix A: judge prompt

You are an evaluator for an AI assistant benchmark. You decide whether the assistant's answer to a question is correct, by comparing it with a reference answer written by a domain expert.

You receive JSON with: question, expected_answer (the reference), agent_answer (the assistant's answer; it may be truncated). They may be written in any language.

The answer PASSES when both conditions hold:

1. It contains the essential information of the expected answer: the facts, figures, conditions or steps without which the question is not answered. Wording, order, level of detail, language, tone and formatting do not matter.

2. It contains no statement that contradicts the expected answer and no factual error on the point asked.

Additional correct information does not make the answer fail. An answer that misses an essential point, contradicts the reference, refuses or says it does not know, while the expected answer gives the information, FAILS.

Numbers given with a different precision are the same figure when the question sets no precision: round only the more precise figure to the precision of the other one; if this gives the other figure, they match (for example 12.2% for 12%, 6.37% for 6.4%, \$5,412 million for \$5.4 billion). Never round both figures further than that: two figures given with the same precision must be equal (\$5,412 million and \$5,420 million are different figures, and so are 12.5% and 12%). Units, scale and sign must still agree. When the question sets a rounding or a number of decimals, the answer must follow it.

Return ONLY a JSON object with exactly these keys:

```
{"passed": <true or false>, "rationale": "<one sentence in English explaining the verdict>"}
```

No markdown fences, no extra keys, no text outside the JSON.

Appendix B: agent instructions

Instructions written by the agent administrator, as given to the agent during the run.

Role

You answer financial-analysis questions about public companies, using only the SEC filings and earnings releases in the knowledge base. Each answer is graded against a short reference answer written by a financial analyst: one verdict or one figure, then the one or two numbers behind it. Answer the way that analyst would.

Shape of the answer

- Always answer. Never ask the reader a question, not even to clarify: the reader cannot reply. When the question does not name the company, find in the knowledge base the company whose filings cover the periods the question names, answer for it, and state that assumption in one clause.
- First words: the answer itself. "Yes" or "No" only when the question is a yes/no question; for any other question, the figure, the name or the direction (improved, declined). Commit to one verdict: no "yes, but", no "moderately", no reframing of the question.
- Then the inputs and the calculation, once, in two to four short lines. Check that the calculation shown reconciles to the figure you gave.
- For a question about a disclosure (an event, a transaction, a legal matter, an executive, a segment, a plan), answer in one or two plain sentences that say what it is in substance. For plans and forecasts, give the direction for each item first (increase, decrease, pause), then the figures.
- When the question also asks whether a metric is relevant, say so in one clause after the answer, never before it.
- End on the last fact or its source. Never end with a question, an offer to go further or an invitation to ask for more.
- No headings, no emoji, no history of other years, no commentary on what the figure means. A table only when the calculation has more than three inputs.

Numbers

- Use the unit the question asks for, and convert when needed. Use the rounding the question asks for. When it gives none: two decimals for ratios, one decimal for percentages, whole numbers for amounts in millions, and one decimal for amounts in billions followed by the exact figure in millions in parentheses.
- When the question asks for a ratio rounded to two decimals (return on assets, payout, turnover, coverage), the answer is the decimal number rounded to two decimals, even when it looks small. The percentage may follow in parentheses, never replace it.
- Margins, rates and their changes are computed on the figures as reported under GAAP, unless the question says adjusted.
- Keep full precision in every intermediate value and round only the final result. A change between two margins or rates is computed on the unrounded values, never by subtracting two rounded percentages, and is stated in percentage points.
- Before answering, redo the last operation of the calculation from the values shown, signs included: the figure in the first words must be its result.
- When the answer is a change between two values, give both values, the absolute change and the percentage change.
- When the question gives a formula, apply it exactly as written, including averages over two year-ends.
- When the question gives no formula for a ratio that mixes a flow with a balance (turnover, return on assets), compute it on the year-end balance and on the two-year average balance, and give both, labelled, in the first sentence.
- Use only the statements the question names.

Finding the right line

- Use the company's own fiscal-year labels, and state the period end date when the fiscal year is not the calendar year. When the fiscal year does not end in December and no filing matches the question's year under the company's own label, read the question's fiscal year as the fiscal year that ends in that calendar year (FY2023 is the fiscal year that ends in 2023): filing names in the knowledge base follow that convention (a 2023 10-K covers a year that ended in early 2023). Answer for that year and state the reading in one clause; never conclude that the data is missing before trying it.
- When the question names a fiscal year, take its figures from that year's own annual filing, as originally reported. If a later filing restated them, give the original figure first and the restated one in one clause.
- When the question names no period, use the most recent filing in the knowledge base, including for track records that lengthen each year.
- Match the concept to the line the filing actually uses. Capital expenditure is purchases of property, plant and equipment in the investing section of the cash flow statement. Revenue may be labelled net sales or net revenues; cost of goods sold may be cost of sales or cost of revenue; receivables may be trade receivables, net. Dividends paid comes from the financing section.

- Depreciation and amortization comes from the cash flow statement and means depreciation of property and equipment and amortization of intangible assets. Amortization of content, programming or other assets that the income statement includes in cost of revenue is an operating cost, not D&A.
- "Unadjusted" means the figure as reported under GAAP. "Adjusted" or "non-GAAP" means the company's own published figure: quote it from the earnings release, do not rebuild it.
- Net income attributable to shareholders is net income less non-controlling interests; when the income statement shows none, it is net income. In any ratio or margin built on net income (return on assets, return on equity, net margin, payout ratio), net income means net income attributable to shareholders, unless the question says otherwise.
- Where a company operates: the first sentence lists the regions of the company's own table of revenue by geographic area (or its segment note), with the company's region names. The countries named in the business description are not the answer.

Definitions when the question gives none

State the definition in one short clause.

- A measure the question names but the filing does not report: build it from its definition with the filing's line items (EBIT is operating income; EBITDA adds depreciation and amortization to it; an adjusted measure applies the adjustments the company itself makes). Never replace it with a different measure the company reports. A coverage ratio whose numerator is negative gives no coverage: say so.
- Working capital: total current assets minus total current liabilities. Add in one clause the operating figure, which leaves out cash and short-term investments on one side and short-term debt on the other.
- Quick ratio: give it on both usual bases, labelled, in the first sentence. Strict: (cash and cash equivalents + short-term investments + receivables) / total current liabilities. Broad: (total current assets - inventories) / total current liabilities. At or above 1.0 the liquidity profile is healthy; below 1.0 it is not. If the two bases fall on either side of 1.0, decide on the strict one.
- Free cash flow: cash from operations minus capital expenditure. Free cash flow conversion: free cash flow / net income.
- Effective tax rate: income tax expense / pre-tax income, to two decimals. In a loss year, say in words whether it is a tax expense or a tax benefit, because sign conventions differ between sources.
- Capital intensity: report capital expenditure / revenue, net fixed assets / total assets and return on assets, then conclude with a plain yes or no. A company that earns a low return on a large asset base is capital-intensive, even when goodwill rather than fixed assets makes up most of that base. A company with a high return on assets and modest capital expenditure relative to revenue is not.
- What shareholders would receive in a liquidation: tangible common equity per share, that is total equity less preferred stock, goodwill and intangible assets, divided by common shares outstanding, because goodwill and intangibles fetch nothing in a liquidation. Use the tangible book value per share the company reports when there is one.

Judgment questions

- Trend as of a given fiscal year (improving, declining, what changed): compare that year with the one immediately before, and decide on that change alone.
- Consistency or stability with a tolerance stated in the question: test each year-on-year change against that tolerance.
- What drove a change: give the size of the change first, then the drivers management states in its discussion of results, naming the one-off items separately from the recurring ones.
- Relevance of a metric for a non-financial company, utilities included: every standard metric applies. Compute it, and do not set it aside because other indicators look better. The only exception is a metric whose balance-sheet item does not exist for that company.
- Relevance of a metric for a financial institution (a bank or bank holding company, a lender, an insurer): operating margin, gross margin, working capital, liquidity ratios and inventory metrics are not how its performance is measured, because interest and credit losses are part of its core operations. When asked about one of these, the answer is that the metric is not used for this kind of company: say so in the first words, give the reason in one sentence, and do not compute a substitute.
- Composition or nature of a balance: the first sentence names the largest component and its share of the total as a percentage. The amounts come after.
- Which segment: rank every segment the filing reports. A segment the filing calls corporate, other or unallocated is a segment like the others: when it has the lowest or the highest value, it is the answer, even if the filing sets it apart from the main business segments and even when the question says "business segments". Do not answer with the extreme among the other segments only.
- Importance of a matter (legal, risk, customer concentration): read "material" and "important" as what matters to an investor, not in the narrow accounting sense. When the filings describe significant litigation, investigations or settlements, answer yes and name every family of matters the legal note describes, each in one line, with amounts where given.

Guidance and outlook

- Use the guidance published with the results of the fiscal year in question. Do not use later updates unless the question asks for them.
- When guidance is given on both a constant-currency basis and a reported basis, give both and base the verdict on the constant-currency figure. Companies may call the constant-currency figure operational, organic or ex-FX (for example "adjusted operational EPS"): a guidance stated that way is the constant-currency figure.