

LlamaParse_GPT-4o_mode_10K_report

July 8, 2024

Hanane D

1- Parsing Amazon 10k Financial Report using:

- **LlamaParse** with **GPT-4o mode** to improve the parsing quality, particularly when financial charts are included in the report
- **SimpleDirectoryParser**: a standard way of parsing.

2- I used HuggingFace local **embedding** (using LlamaIndex), to store data in a VectorStore

3- I created a **query engine** using different LLMs and compare their results: **gpt-3.5-turbo**, **GPT-4o**, and **Claude Sonnet 3.5** .

1 Install Lib

```
[ ]: !pip install llama-index llama-index-core llama-parse openai llama_index.  
      ↳ embeddings.huggingface -q  
      !pip install llama-index-llms-anthropic -q
```

2 Specify API Keys

```
[2]: import nest_asyncio  
nest_asyncio.apply()  
  
from google.colab import userdata  
OPENAI_API_KEY = userdata.get('OPENAI_API_KEY')  
LLMAPARSE_API_KEY = userdata.get('LLMAPARSE_API_KEY')  
ANTHROPIC_API_KEY = userdata.get("CLAUDE_API_KEY")
```

3 Loading financial report: Amazon 2023 10K

```
[ ]: !wget "https://d18rn0p25nwr6d.cloudfront.net/CIK-0001018724/  
      ↳ c7c14359-36fa-40c3-b3ca-5bf7f3fa0b96.pdf" -O amzn_2023_10k.pdf
```

4 Standard Parsing with SimpleDirectoryReader:

4.1 VectorStore and specifying LLMs for the query_engine

```
[38]: # from llama_parse import LlamaParse
from llama_index.core import VectorStoreIndex, SimpleDirectoryReader
import nest_asyncio;
nest_asyncio.apply()

pdf_name = "amzn_2023_10k.pdf"
# use SimpleDirectoryReader to parse our file
documents = SimpleDirectoryReader(input_files=[pdf_name]).load_data()
from llama_index.core import Settings

embed_model = "local:BAAI/bge-small-en-v1.5" #https://huggingface.co/
↳collections/BAAI/bge-66797a74476eb1f085c7446d

vector_index_std = VectorStoreIndex(documents, embed_model = embed_model)

from llama_index.llms.openai import OpenAI

llm_gpt35 = OpenAI(model="gpt-3.5-turbo", api_key = OPENAI_API_KEY)
query_engine_gpt35 = vector_index_std.as_query_engine(similarity_top_k=3,↳
↳llm=llm_gpt35)

llm_gpt4o = OpenAI(model="gpt-4o", api_key = OPENAI_API_KEY)
query_engine_gpt4o = vector_index_std.as_query_engine(similarity_top_k=3,↳
↳llm=llm_gpt4o)

from llama_index.llms.anthropic import Anthropic
from llama_index.core import Settings

tokenizer = Anthropic().tokenizer
Settings.tokenizer = tokenizer

llm_claude = Anthropic(model="claude-3-5-sonnet-20240620",↳
↳api_key=ANTHROPIC_API_KEY)
query_engine_claude = vector_index_std.as_query_engine(similarity_top_k=3,↳
↳llm=llm_claude)
```

```
/usr/local/lib/python3.10/dist-packages/huggingface_hub/file_download.py:1132:
FutureWarning: `resume_download` is deprecated and will be removed in version
1.0.0. Downloads always resume when possible. If you want to force a new
download, use `force_download=True`.
warnings.warn(
```

```
[42]: print(documents[36].text)
```

Table of Contents

AMAZON.COM, INC.

CONSOLIDATED STATEMENTS OF CASH FLOWS

(in millions)

Year Ended December 31,

2021 2022 2023

CASH, CASH EQUIV ALENTS, AND RESTRICTED CASH, BEGINNING OF PERIOD \$ 42,377 \$ 36,477 \$ 54,253

OPERA TING ACTIVITIES:

Net income (loss) 33,364 (2,722) 30,425

Adjustments to reconcile net income (loss) to net cash from operating activities:

Depreciation and amortization of property and equipment and capitalized content costs, operating lease

assets, and other 34,433 41,921 48,663

Stock-based compensation 12,757 19,621 24,023

Non-operating expense (income), net (14,306) 16,966 (748)

Deferred income taxes (310) (8,148) (5,876)

Changes in operating assets and liabilities:

Inventories (9,487) (2,592) 1,449

Accounts receivable, net and other (9,145) (8,622) (8,348)

Other assets (9,018) (13,275) (12,265)

Accounts payable 3,602 2,945 5,473

Accrued expenses and other 2,123 (1,558) (2,428)

Unearned revenue 2,314 2,216 4,578

Net cash provided by (used in) operating activities 46,327 46,752 84,946

INVESTING ACTIVITIES:

Purchases of property and equipment (61,053) (63,645) (52,729)

Proceeds from property and equipment sales and incentives 5,657 5,324 4,596

Acquisitions, net of cash acquired, non-marketable investments, and other (1,985) (8,316) (5,839)

Sales and maturities of marketable securities 59,384 31,601 5,627

Purchases of marketable securities (60,157) (2,565) (1,488)

Net cash provided by (used in) investing activities (58,154) (37,601) (49,833)

FINANCING ACTIVITIES:

Common stock repurchased - (6,000) -

Proceeds from short-term debt, and other 7,956 41,553 18,129

Repayments of short-term debt, and other (7,753) (37,554) (25,677)

Proceeds from long-term debt 19,003 21,166 -

Repayments of long-term debt (1,590) (1,258) (3,676)

Principal repayments of finance leases (11,163) (7,941) (4,384)

Principal repayments of financing obligations (162) (248) (271)

Net cash provided by (used in) financing activities 6,291 9,718 (15,879)

Foreign currency ef fect on cash, cash equivalents, and restricted cash (364) (1,093) 403

Net increase (decrease) in cash, cash equivalents, and restricted cash (5,900)
17,776 19,637
CASH, CASH EQUIV ALENTS, AND RESTRICTED CASH, END OF PERIOD \$ 36,477 \$ 54,253 \$
73,890
See accompanying notes to consolidated financial statements.
37

4.2 Chatting with the LLMs: GPT-3.5-Turbo, GPT-4o, Claude 3.5 Sonnet

```
[43]: query1 = "What was the net income in 2023?"
response = query_engine_gpt35.query(query1)
print("GPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query1)
print("\nGPT-4o")
print(str(resp))

resp = query_engine_claude.query(query1)
print("\nClaude 3.5 Sonnet")
print(str(resp))

# print(resp.source_nodes[0].get_content()) # to get the source_node used to
↪ answer
```

GPT-3.5 Turbo

The net income for 2023 is \$36,813 million.

GPT-4o

The net income for 2023 is not explicitly provided in the given context. However, you can calculate it using the provided information. The income before income taxes for 2023 is \$37,557 million, and the provision for income taxes is \$7,120 million. Therefore, the net income for 2023 would be \$37,557 million minus \$7,120 million, which equals \$30,437 million.

Claude 3.5 Sonnet

Based on the information provided, Amazon's net income for 2023 can be calculated as \$30,437 million. This is derived from the "Income (loss) before income taxes" of \$37,557 million for 2023, minus the "Provision (benefit) for income taxes, net" of \$7,120 million for the same year.

```
[44]: query1 = "What was the net income in 2022?"
response = query_engine_gpt35.query(query1)
print("GPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query1)
```

```

print("\nGPT-4o")
print(str(resp))

resp = query_engine_claude.query(query1)
print("\nClaude 3.5 Sonnet")
print(str(resp))

```

GPT-3.5 Turbo

The net income for 2022 is a provision (benefit) for income taxes of \$(3.2) billion.

GPT-4o

The net income for 2022 is not explicitly provided in the given context. However, you can calculate it using the provided information. The income (loss) before income taxes for 2022 is \$(5,936) million, and the provision (benefit) for income taxes is \$(3,217) million. Therefore, the net income for 2022 would be:

Net Income = Income (Loss) Before Income Taxes - Provision (Benefit) for Income Taxes

Net Income = \$(5,936) million - \$(3,217) million

Net Income = \$(5,936) million + \$3,217 million

Net Income = \$(2,719) million

So, the net income for 2022 is \$(2,719) million.

Claude 3.5 Sonnet

Based on the information provided, Amazon reported a net loss in 2022. The company had a loss before income taxes of \$5,936 million in 2022. After accounting for an income tax benefit of \$3,217 million that year, the net loss for 2022 would be approximately \$2,719 million.

```

[46]: query2 = "What was the net income in 2023 compared to 2022?"
print("query2:", query2)
response = query_engine_gpt35.query(query2)
print("\nGPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query2)
print("\nGPT-4o")
print(str(resp))

resp = query_engine_claude.query(query2)
print("\nClaude 3.5 Sonnet")
print(str(resp))

```

query2: What was the net income in 2023 compared to 2022?

GPT-3.5 Turbo

The net income in 2023 was \$37.557 billion, while in 2022, it was \$(5.936) billion.

GPT-4o

The net income in 2023 was higher compared to 2022. In 2023, the income before income taxes was \$37,557 million, and the provision for income taxes was \$7,120 million, resulting in a net income of \$30,437 million. In contrast, in 2022, the loss before income taxes was \$5,936 million, and the benefit for income taxes was \$3,217 million, resulting in a net loss of \$2,719 million.

Claude 3.5 Sonnet

Based on the financial information provided, Amazon's net income situation changed significantly from 2022 to 2023:

In 2022, Amazon had a loss before income taxes of \$5,936 million and received an income tax benefit of \$3,217 million.

In 2023, Amazon had income before income taxes of \$37,557 million and incurred an income tax provision of \$7,120 million.

While the exact net income figure isn't directly stated, we can infer that Amazon went from a net loss position in 2022 to a substantial net profit in 2023. The dramatic increase in pre-tax income of over \$43 billion year-over-year, combined with returning to paying taxes rather than receiving a tax benefit, indicates a major improvement in Amazon's profitability in 2023 compared to 2022.

```
[47]: print(resp.source_nodes[0].get_content())
```

Table of Contents

The components of the provision (benefit) for income taxes, net are as follows (in millions):

Year Ended December 31,
2021 2022 2023

U.S. Federal:

Current \$ 2,129 \$ 2,175 \$ 8,652

Deferred 155 (6,686) (5,505)

Total 2,284 (4,511) 3,147

U.S. State:

Current 763 1,074 2,158

Deferred (178) (1,302) (498)

Total 585 (228) 1,660

International:

Current 2,209 1,682 2,186

Deferred (287) (160) 127

Total 1,922 1,522 2,313

Provision (benefit) for income taxes, net \$ 4,791 \$ (3,217) \$ 7,120

U.S. and international components of income (loss) before income taxes are as follows (in millions):

Year Ended December 31,	2021	2022	2023
U.S.	\$ 35,879	\$ (8,225)	\$ 32,328
International	2,272	2,289	5,229
Income (loss) before income taxes	\$ 38,151	\$ (5,936)	\$ 37,557

The items accounting for differences between income taxes computed at the federal statutory rate and the provision (benefit) recorded for income taxes are as follows (in millions):

Year Ended December 31,	2021	2022	2023
Income taxes computed at the federal statutory rate	\$ 8,012	\$ (1,246)	\$ 7,887
Effect of:			
Tax impact of foreign earnings and losses	(1,349)	(370)	594
State taxes, net of federal benefits	465	(173)	1,307
Tax credits	(1,136)	(1,006)	(2,362)
Stock-based compensation	(1)	(1,094)	612
Foreign income deduction	(2)	(301)	(1,258)
Other, net	194	224	76
Total	\$ 4,791	\$ (3,217)	\$ 7,120

(1)Includes non-deductible stock-based compensation and excess tax benefits or shortfalls from stock-based compensation. Our tax provision includes \$1.9 billion of excess tax benefits from stock-based compensation for 2021, and \$33 million and \$519 million of tax shortfalls from stock-based compensation for 2022 and 2023.

(2)U.S. companies are eligible for a deduction that lowers the effective tax rate on certain foreign income. This regime is referred to as the Foreign-Derived

Intangible Income deduction and is dependent on the amount of our U.S. taxable income.

We generated an income tax benefit in 2022 as compared to a provision for income taxes in 2021 primarily due to a decrease in pretax income and an increase in the foreign income deduction. This was partially offset by a reduction in excess tax benefits from stock-based compensation and a decrease in the

tax impact of foreign earnings and losses driven by a decline in the favorable effects of corporate restructuring transactions. The foreign income deduction benefit recognized in 2022 reflects a change in our application of tax regulations related to the computation of qualifying foreign income and includes a tax

benefit of approximately \$655 million related to years prior to 2022.

5 LlamaParse: Amazon 10K Financial report

```
[17]: from llama_parse import LlamaParse
      from llama_index.core import VectorStoreIndex

      pdf_name = "amzn_2023_10k.pdf"
      # set up parser
      # parser = LlamaParse(api_key=LLAMAPARSE_API_KEY, result_type="markdown",
      #                    ↪gpt4o_mode = True)
      parser = LlamaParse(api_key=LLAMAPARSE_API_KEY, result_type="markdown",
      #                    ↪gpt4o_mode = True)
      documents = parser.load_data(pdf_name)
```

Started parsing the file under job_id cac11eca-3194-4b60-a73f-76fb55d0bbaa

```
[29]: parser
```

```
[29]: LlamaParse(is_remote=False, api_key='llx-
      IGUjt4CbLaOtNr62qME76pw7QTPZv1dVL5toWNrFarwD9NP6',
      base_url='https://api.cloud.llamaindex.ai', result_type=<ResultType.MD:
      'markdown'>, num_workers=4, check_interval=1, max_timeout=2000, verbose=True,
      show_progress=True, language=<Language.ENGLISH: 'en'>, parsing_instruction='',
      skip_diagonal_text=False, invalidate_cache=False, do_not_cache=False,
      fast_mode=False, do_not_unroll_columns=False, page_separator=None,
      gpt4o_mode=True, gpt4o_api_key=None, bounding_box=None, target_pages=None,
      ignore_errors=True, split_by_page=True)
```

```
[18]: print(documents[43].text)
```

```
# AMAZON.COM, INC.
## CONSOLIDATED STATEMENTS OF CASH FLOWS
### (in millions)
#### Year Ended December 31,
```

	2021	2022	2023
CASH, CASH EQUIVALENTS, AND RESTRICTED CASH, BEGINNING OF PERIOD	\$ 42,377	\$ 36,477	\$ 54,253
OPERATING ACTIVITIES:			
Net income (loss)	33,364	(2,722)	30,425
Adjustments to reconcile net income (loss) to net cash from operating activities:			
Depreciation and amortization of property and equipment and capitalized content costs, operating lease assets, and other	34,433	41,921	48,663
Stock-based compensation	12,757	19,621	24,023
Non-operating expense (income), net	(14,306)	16,966	(748)

Deferred income taxes	(310)	(8,145)	(5,876)	
Changes in operating assets and liabilities:				
Inventories	(9,487)	(2,592)	1,449	
Accounts receivable, net and other	(9,145)	(8,822)	2,345	
Other assets	(9,018)	(13,275)	(12,365)	
Accounts payable	3,602	2,945	5,473	
Accrued expenses and other	3,123	(1,558)	(2,473)	
Unearned revenue	2,314	2,210	4,798	
Net cash provided by (used in) operating activities	46,327	46,752		
84,946				
INVESTING ACTIVITIES:				
Purchases of property and equipment	(61,053)	(63,454)	(52,729)	
Proceeds from property and equipment sales and incentives	5,657	5,324		
4,596				
Acquisitions, net of cash acquired, non-marketable investments, and other	(1,985)	(8,316)	(5,839)	
Sales and maturities of marketable securities	59,384	31,101	43,527	
Purchases of marketable securities	(60,157)	(2,565)	(1,488)	
Net cash provided by (used in) investing activities	(58,154)	(37,901)		
(49,933)				
FINANCING ACTIVITIES:				
Common stock repurchased	-	(6,000)	-	
Proceeds from short-term debt, and other	7,956	41,553	18,129	
Repayments of short-term debt, and other	(7,753)	(37,554)	(25,877)	
Proceeds from long-term debt	19,003	21,166	18,129	
Repayments of long-term debt	(1,590)	(1,258)	(3,766)	
Principal repayments of finance leases	(11,163)	(7,941)	(4,384)	
Principal repayments of financing obligations	(162)	(248)	(271)	
Net cash provided by (used in) financing activities	6,291	9,718		
(15,979)				
Foreign currency effect on cash, cash equivalents, and restricted cash	(364)			
(1,093)	403			
Net increase (decrease) in cash, cash equivalents, and restricted cash	(5,900)	17,576	19,637	
CASH, CASH EQUIVALENTS, AND RESTRICTED CASH, END OF PERIOD	\$ 36,477	\$ 54,253	\$ 73,890	

See accompanying notes to consolidated financial statements.

5.1 Store documents: with embeddings for later retrieval

```
[20]: from llama_index.core import VectorStoreIndex
```

```
embed_model = "local:BAAI/bge-small-en-v1.5" #https://huggingface.co/
↳collections/BAAI/bge-66797a74476eb1f085c7446d
vector_index = VectorStoreIndex(documents, embed_model = embed_model)
```

```
/usr/local/lib/python3.10/dist-packages/huggingface_hub/file_download.py:1132:
FutureWarning: `resume_download` is deprecated and will be removed in version
1.0.0. Downloads always resume when possible. If you want to force a new
download, use `force_download=True`.
warnings.warn(
```

```
[22]: # vector_index.storage_context.persist(persist_dir=path)
```

5.2 Specify LLMs for chat

```
[23]: from llama_index.llms.openai import OpenAI

llm_gpt35 = OpenAI(model="gpt-3.5-turbo", api_key = OPENAI_API_KEY)
query_engine_gpt35 = vector_index.as_query_engine(similarity_top_k=3,
↳llm=llm_gpt35)

llm_gpt4o = OpenAI(model="gpt-4o", api_key = OPENAI_API_KEY)
query_engine_gpt4o = vector_index.as_query_engine(similarity_top_k=3,
↳llm=llm_gpt4o)

from llama_index.llms.anthropic import Anthropic
from llama_index.core import Settings

tokenizer = Anthropic().tokenizer
Settings.tokenizer = tokenizer

llm_claude = Anthropic(model="claude-3-5-sonnet-20240620",
↳api_key=ANTHROPIC_API_KEY)
query_engine_claude = vector_index.as_query_engine(similarity_top_k=3,
↳llm=llm_claude)
```

5.3 Chatting with the LLMs: GPT-3.5-Turbo, GPT-4o, Claude 3.5 Sonnet

```
[24]: query1 = "What is the net income on 2023?"
response = query_engine_gpt35.query(query1)
print("GPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query1)
print("\nGPT-4o")
print(str(resp))
```

```
resp = query_engine_claude.query(query1)
print("\nClaude 3.5 Sonnet")
print(str(resp))
```

GPT-3.5 Turbo

The net income for 2023 is \$16.9 billion.

GPT-4o

The net income for 2023 can be calculated by subtracting the provision for income taxes from the income before income taxes. For 2023, the income before income taxes is \$37,557 million, and the provision for income taxes is \$7,120 million. Therefore, the net income for 2023 is \$37,557 million - \$7,120 million = \$30,437 million.

Claude 3.5 Sonnet

The net income for 2023 can be calculated by subtracting the provision for income taxes from the income before income taxes.

Income before income taxes in 2023 was \$37,557 million.

Provision for income taxes in 2023 was \$7,120 million.

Therefore, the net income for 2023 was:

\$37,557 million - \$7,120 million = \$30,437 million

So, the net income for 2023 was \$30,437 million.

```
[25]: query2 = "What was the net income in 2023 compared to 2022?"
print("query2:", query2)
response = query_engine_gpt35.query(query2)
print("\nGPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query2)
print("\nGPT-4o")
print(str(resp))

resp = query_engine_claude.query(query2)
print("\nClaude 3.5 Sonnet")
print(str(resp))
```

query2: What was the net income in 2023 compared to 2022?

GPT-3.5 Turbo

The net income in 2023 was higher compared to 2022.

GPT-4o

The net income in 2023 was \$30.4 billion, compared to a net loss of \$2.7 billion in 2022.

Claude 3.5 Sonnet

The net income in 2023 was significantly higher compared to 2022. In 2023, the company reported income before income taxes of \$37.557 billion, and after accounting for income taxes of \$7.120 billion, the net income would be approximately \$30.437 billion.

In contrast, 2022 saw a loss before income taxes of \$5.936 billion. However, there was a tax benefit of \$3.217 billion, which would result in a net loss of approximately \$2.719 billion for 2022.

This represents a substantial turnaround, with the company moving from a net loss position in 2022 to a significant profit in 2023. The improvement was primarily driven by a large increase in income before taxes, reflecting better overall business performance in 2023.

```
[26]: query2 = "What was the reason of the net income loss in 2022?"
print("query2:", query2)
response = query_engine_gpt35.query(query2)
print("\nGPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query2)
print("\nGPT-4o")
print(str(resp))

resp = query_engine_claude.query(query2)
print("\nClaude 3.5 Sonnet")
print(str(resp))
```

query2: What was the reason of the net income loss in 2022?

GPT-3.5 Turbo

The net income loss in 2022 was primarily due to a decrease in pretax income, an increase in the foreign income deduction, a reduction in excess tax benefits from stock-based compensation, and a decrease in the tax impact of foreign earnings and losses driven by a decline in the favorable effects of corporate restructuring transactions.

GPT-4o

The net income loss in 2022 was primarily due to a decrease in pretax income and an increase in the foreign income deduction. This was partially offset by a reduction in excess tax benefits from stock-based compensation and a decrease in the tax impact of foreign earnings and losses driven by a decline in the favorable effects of corporate restructuring transactions.

Claude 3.5 Sonnet

The net income loss in 2022 was primarily due to a significant decrease in

pretax income. The company experienced a loss before income taxes of \$5,936 million in 2022, compared to an income of \$38,151 million in 2021. This substantial decline in pretax income was a major factor contributing to the net loss.

Additionally, there were other factors that influenced the tax situation in 2022:

1. An increase in the foreign income deduction, which lowered the effective tax rate on certain foreign income.
2. A reduction in excess tax benefits from stock-based compensation.
3. A decrease in the tax impact of foreign earnings and losses, driven by a decline in the favorable effects of corporate restructuring transactions.
4. The company also recognized a change in the application of tax regulations related to the computation of qualifying foreign income, which included a tax benefit of approximately \$655 million for years prior to 2022.

These factors combined to result in an income tax benefit of \$3,217 million in 2022, as opposed to a tax provision in the previous year. However, this tax benefit was not sufficient to offset the significant decrease in pretax income, leading to the overall net loss for 2022.

```
[28]: print(resp.source_nodes[0].get_content())
```

The components of the provision (benefit) for income taxes, net are as follows (in millions):

	Year Ended December 31,		
	2021	2022	2023
U.S. Federal:			
Current	\$ 2,129	\$ 2,175	\$ 8,652
Deferred	155	(6,686)	(5,505)
Total	2,284	(4,511)	3,147
U.S. State:			
Current	763	1,074	2,158
Deferred	(178)	(1,302)	(498)
Total	585	(228)	1,660
International:			
Current	2,209	1,682	2,186
Deferred	(287)	(160)	127
Total	1,922	1,522	2,313
Provision (benefit) for income taxes, net	\$ 4,791	\$ (3,217)	\$ 7,120

U.S. and international components of income (loss) before income taxes are as follows (in millions):

	Year Ended December 31,		
	2021	2022	2023
U.S.	\$ 35,579	\$ (8,225)	\$ 32,328
International	2,272	2,289	5,229
Income (loss) before income taxes	\$ 38,151	\$ (5,936)	\$ 37,557

The items accounting for differences between income taxes computed at the federal statutory rate and the provision (benefit) recorded for income taxes are as follows (in millions):

	Year Ended December 31,		
	2021	2022	2023
Income taxes computed at the federal statutory rate	\$ 8,012	\$ (1,246)	\$ 7,857
Effect of:			
Tax impact of foreign earnings and losses	(1,349)	(370)	594
State taxes, net of federal benefits	465	173	1,307
Tax credits	(1,136)	(1,006)	(2,362)
Stock-based compensation (1)	(1,094)	612	1,047
Foreign income deduction (2)	(301)	(1,258)	(1,429)
Other, net	194	224	76
Total	\$ 4,791	\$ (3,217)	\$ 7,120

1. Includes non-deductible stock-based compensation and excess tax benefits or shortfalls from stock-based compensation. Our tax provision includes \$1.9 billion of excess tax benefits from stock-based compensation for 2021, and \$33 million and \$519 million of tax shortfalls from stock-based compensation for 2022 and 2023.

2. U.S. companies are eligible for a deduction that lowers the effective tax rate on certain foreign income. This regime is referred to as the Foreign-Derived Intangible Income deduction and is dependent on the amount of our U.S. taxable income.

We generated an income tax benefit in 2022 as compared to a provision for income taxes in 2021 primarily due to a decrease in pretax income and an increase in the foreign income deduction. This was partially offset by a reduction in excess tax benefits from stock-based compensation and a decrease in the tax impact of foreign earnings and losses driven by a decline in the favorable effects of corporate restructuring transactions. The foreign income deduction benefit recognized in 2022 reflects a change in our application of tax regulations related to the computation of qualifying foreign income and includes a tax benefit of approximately \$655 million related to years prior to 2022.

```
[30]: query3 = "What are the most important takeaways from the report?"
print("query3:", query3)
response = query_engine_gpt35.query(query3)
print("\nGPT-3.5 Turbo")
print(str(response))

resp = query_engine_gpt4o.query(query3)
print("\nGPT-4o")
print(str(resp))

resp = query_engine_claude.query(query3)
print("\nClaude 3.5 Sonnet")
print(str(resp))
```

query3: What are the most important takeaways from the report?

GPT-3.5 Turbo

The report includes the consolidated financial statements, notes to the financial statements, and the independent auditor's report. Additionally, it lists the documents filed as part of the report, including the index to consolidated financial statements, financial statement schedules, and exhibits. The forward-looking statements section highlights the uncertainties and factors that could impact future results.

GPT-4o

The most important takeaways from the report include the detailed financial statements and supplementary data, such as the consolidated statements of cash flows, operations, comprehensive income (loss), balance sheets, and stockholders' equity. Additionally, the report includes the independent auditor's report from Ernst & Young LLP. The management's discussion and analysis section highlights the company's forward-looking statements, emphasizing the inherent uncertainties and various factors that could affect future results, such as fluctuations in foreign exchange rates, changes in global economic conditions, customer demand, and supply chain disruptions. The report also lists several exhibits, including the company's amended and restated certificate of incorporation and bylaws, as well as various indentures and officers' certificates related to the company's notes.

Claude 3.5 Sonnet

The key takeaways from this report include:

1. The company's financial statements are audited by Ernst & Young LLP, an independent registered public accounting firm.
2. The report contains consolidated financial statements covering cash flows, operations, comprehensive income/loss, balance sheets, and stockholders' equity for multiple years.

3. There are several exhibits referenced, including amended and restated certificates of incorporation and bylaws, as well as various indentures and notes.
4. The company acknowledges that its future results and financial position may be affected by various factors, including foreign exchange rates, economic conditions, customer demand, inflation, and supply chain disruptions.
5. The company expects variability in its accounts payable days due to factors like product sales mix, use of third-party sellers, inventory management strategies, and investments in new markets and product lines.
6. The report contains forward-looking statements that reflect management's expectations but are inherently uncertain and subject to various risks and factors that could cause actual results to differ.
7. The company's financial performance and operations are influenced by a complex mix of factors, including its product and service offerings, sales channels, fulfillment strategies, and global expansion efforts.

```
[ ]: # For this question, the cost went from $0.07 to $0.16 ==> $0.09
```

```
[31]: print(resp.source_nodes[0].get_content())
```

```
# Item 8. Financial Statements and Supplementary Data
```

```
## INDEX TO CONSOLIDATED FINANCIAL STATEMENTS
```

```
|
| Page |
|-----|
|-----|-----|
| Report of Ernst & Young LLP, Independent Registered Public Accounting Firm
(PCAOB ID: 42) | 35 |
| Consolidated Statements of Cash Flows
| 37 |
| Consolidated Statements of Operations
| 38 |
| Consolidated Statements of Comprehensive Income (Loss)
| 39 |
| Consolidated Balance Sheets
| 40 |
| Consolidated Statements of Stockholders' Equity
| 41 |
| Notes to Consolidated Financial Statements
| 42 |
```



```
[ ]: print(resp.source_nodes[1].get_content())
```

```
[ ]: print(resp.source_nodes[2].get_content())
```

6 Key Takeaways

1- **In terms of parsing**, since the 10k report includes only tables, using LlamaParse with mark-down is sufficient (and also visually clear). There's no need for GPT-4o mode. However, as you don't know if there are charts in the PDFs, using GPT-4o mode could be beneficial. However, I found that Claude 3.5 Sonnet (in a previous post) was better than GPT-4o on parsing charts as images.

It could be great to have Claude 3.5 Sonnet mode on LlamaParse. It can also be more expensive.

The simple parsing (SimpleDirectoryReader), was good too for these kind of simple questions.

2- **In the chat comparison**, Claude 3.5 Sonnet was slightly better than GPT-4o. It provides accurate answers and offers valuable details on how some calculations are performed, as well as justifications for certain losses. However, GPT-3.5-Turbo lags significantly behind.