Meta Platforms, Inc.

Financial Analysis Report

ACCT 745.01

Chaturya Valli Pithani

Table of Contents

I. Company Background

Introduction

Meta Platforms, Inc. (META) was founded in 2004 as TheFacebook, Inc. Their flagship service, Facebook, was a social networking platform that grew popular as internet access and mobile devices became more accessible in the mid- to late-2000s. TheFacebook, Inc. rebranded as Facebook, Inc. in 2005, and in 2021, rebranded again as Meta Platforms, Inc. in a shift to convey that as a company, Meta's focus was not just on their original Facebook platform but also on other services like Instagram, WhatsApp, and the Meta Quest line of virtual reality headsets.

Meta Platforms, Inc. filed for IPO in 2012. The analysis of Meta throughout this report will cover Meta's financial history from between their 2012 IPO and 2022.

Year-by-Year History of Meta Platforms, Inc. (2012-2022)

2012

- Facebook, Inc. files for IPO
- Facebook, Inc. acquires Instagram for $1 billion in cash and stock
- Facebook (the platform) reaches 1 billion active users

2013

- Facebook, Inc. launches internet.org (later renamed to Free Basics) initiative/service to bring affordable internet access to developing countries

2014

- Facebook, Inc. acquires WhatsApp for $16 billion in cash and stock
- Facebook, Inc. acquires virtual reality headset company Oculus for $2 billion in cash and stock

- Facebook, Inc. faced criticism when it was discovered to be conducting psychological tests on unconsenting Facebook users by adding/removing certain words from user feeds to test how posts affected the moods of users
- Facebook mobile app was the most downloaded mobile app of the year

2015

- Facebook mobile app was the most downloaded mobile app of the year

2016

- Facebook, Inc. faced criticism for the spread of misinformation regarding the 2016 U.S. presidential election on Facebook (the platform) after reports demonstrated how false articles spread and garnered more clicks than legitimate articles.
- Facebook mobile app was the most downloaded mobile app of the year and Facebook Messenger was the second most downloaded mobile app of the year

2017

- Facebook, Inc. acquires “tbh”, a small social media app targeted towards teens
- WhatsApp was the most downloaded app of the year (excluding mobile game apps)

2018

- Facebook, Inc. faced criticism for a 2016 scandal involving data analytics firm Cambridge Analytica, who assisted in targeting advertisements for U.S. presidential candidate Donald Trump’s campaign by improperly using the data of millions of U.S. Facebook users
- Facebook, Inc. CEO Mark Zuckerberg testified in front of the U.S. Congress about privacy rights, banning Cambridge Analytica from the Facebook platform, and other scandals

2019

- Facebook, Inc. faced a record-breaking fine of $5 billion over user data privacy violations from the FTC

2020

- Advertisers on Facebook, Inc.'s platforms boycotted after CEO Mark Zuckerberg made it clear he would not remove posts (including potentially dangerous ones) from noteworthy figures and that Facebook would not fact-check political ads leading up to the 2020 U.S. presidential election
- Facebook, Inc., Amazon, Apple, and Alphabet were ruled by U.S. lawmakers to be monopolies that should be regulated after U.S. Congressional hearing

2021

- Facebook (the platform) had its user data breached, with over 530 million of its users' data being exposed
- Facebook, Inc.'s stock price hit its all time high on September 6th, at $378.69 per share
- Leaked internal documents from Facebook, Inc. revealed that company management knowingly prioritized profitability over public good after research concluded the company's profound effects on society and politics in ways that could incite violence and harm
- Facebook, Inc. rebranded to Meta Platforms, Inc.

2022

- The average number of active daily users of Facebook (the platform) dropped for the first time
- Meta Platforms, Inc.'s stock price hit the lowest price it's been since 2015, on October 31st at $90.79 per share

II. Data and Analysis Methods Used

Two types of data were used to create the visualizations and analyses of this report: structured data and unstructured data. The structured data consisted of dollar amounts from Meta's financial statements, dating between 2012 and 2022. Structured data was obtained through XBRL integration in Google Sheets. Once imported into Google Sheets, the XBRL data

was brought into Tableau to create visualizations that will be seen throughout this report to convey increases and decreases in Meta's financial standing.

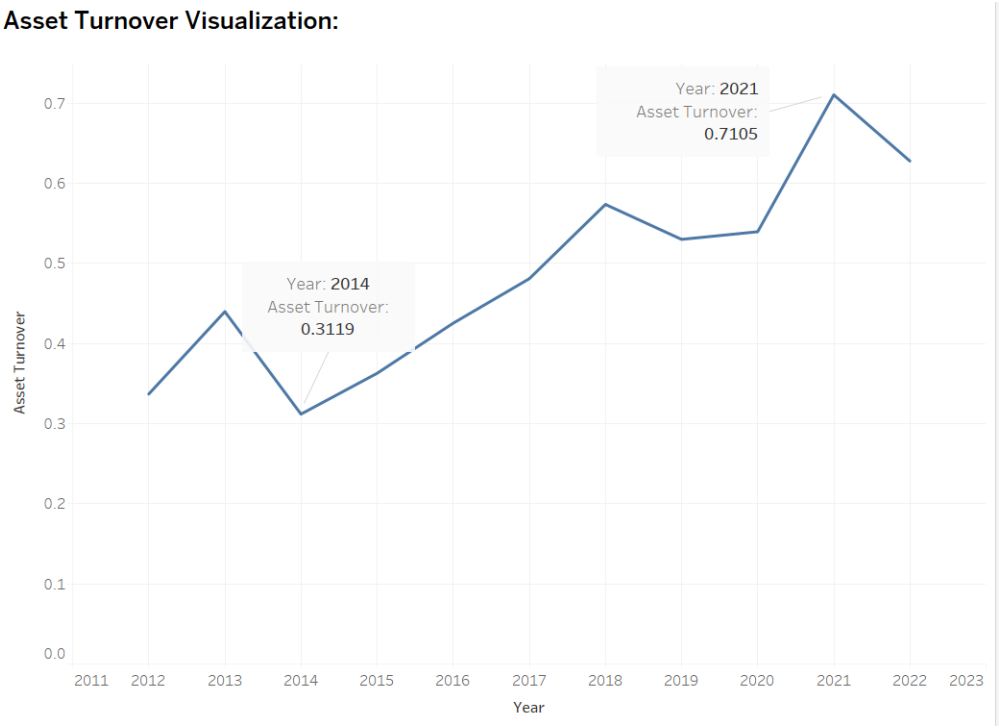
The unstructured data consisted of text mined from the Management's Discussion & Analysis sections from Meta's 10K filings between 2012 and 2022. The text from the filings was mined through use of Python scripts that obtained and cleaned text from the filings and analyzed the sentiment of the filing's text. Once the Python script has been run, the program will output a WordCloud containing the most frequently used words in the desired year's 10K MD&A, and a bar chart that measures the percent of words in the MD&A that were positive, negative, or neutral sentiment. Two Python scripts were used to create the output visuals: One that would output visualizations for a single desired year's 10K MD&A, and one that would output visualizations for all 10K MD&A sections combined. The Python scripts used to create the visualizations using the unstructured data can be found in the appendix of this report.

III. Overview of Historical Financial Health

Meta Platforms Inc.'s financial journey from 2011 to 2023, as captured in our report and illustrated through various data visualizations, reflects a company that has navigated its market with agility and strategic foresight. This section weaves the narrative of Meta's historical financial data with the graphical analysis to present a comprehensive picture of its financial evolution.

Asset Efficiency and Revenue Generation

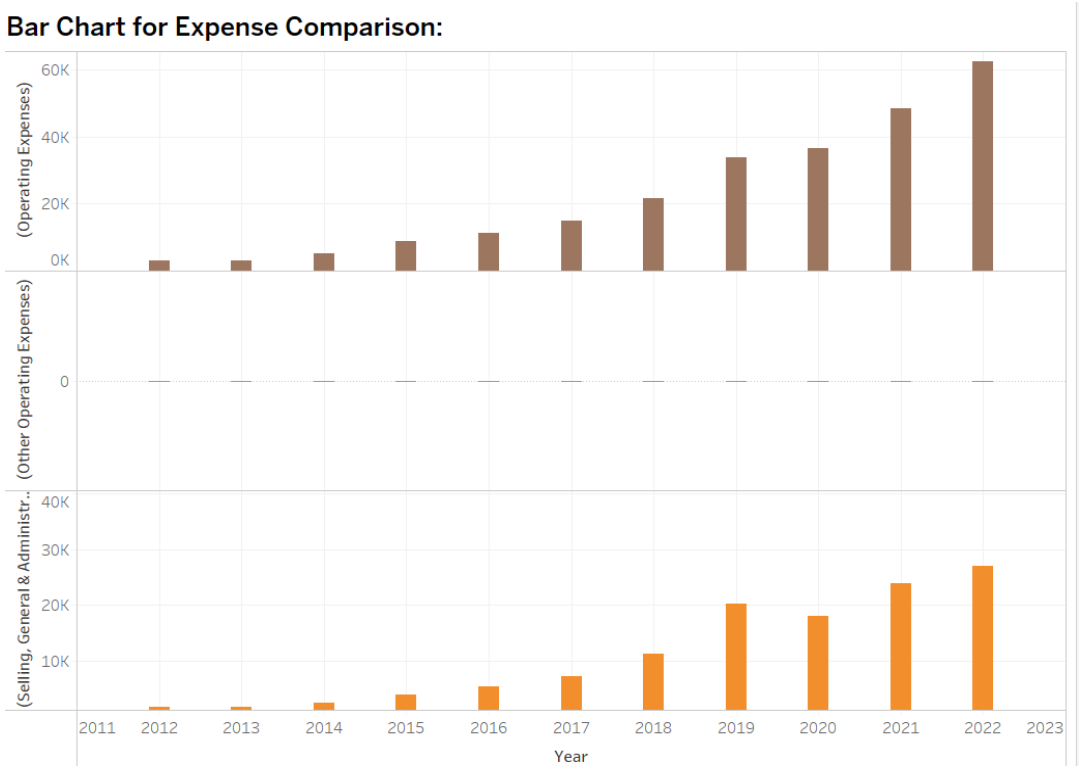
Asset Turnover Visualization:



The Asset Turnover Visualization graph underscores periods of intense investment in technology and acquisitions, which initially resulted in lower asset turnover rates. However, these strategic decisions have laid the groundwork for future revenue growth, which is evident from the subsequent increase in the asset turnover ratio, particularly the peak observed in 2021. This trajectory highlights how Meta's investments have matured, translating into more efficient use of assets to generate higher sales.

Scaling Operations and Expense Dynamics

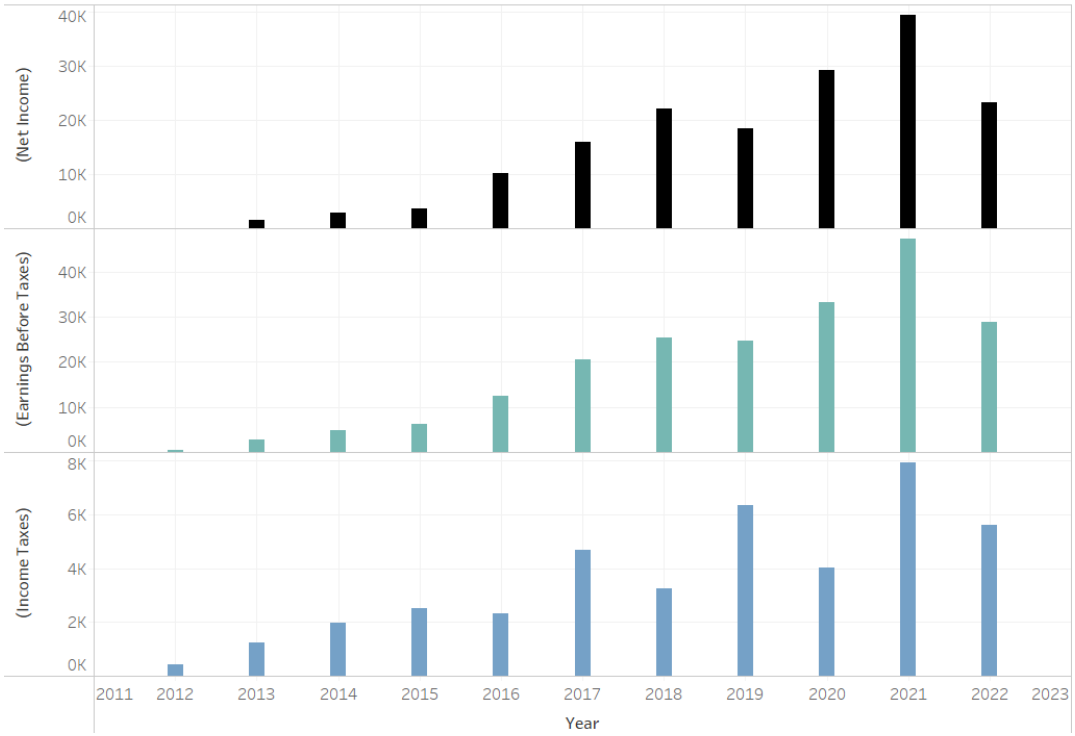
Bar Chart for Expense Comparison:



The Bar Chart for Expense Comparison graph provides a granular view of the operational expenses, which have escalated as Meta expanded its user base and product offerings. The increasing trend in SG&A expenses aligns with our report's findings that underscore Meta's aggressive investment in research and development as well as its marketing efforts. Despite the growing expenses, these investments are seen as strategic initiatives to cement Meta's position in the competitive landscape.

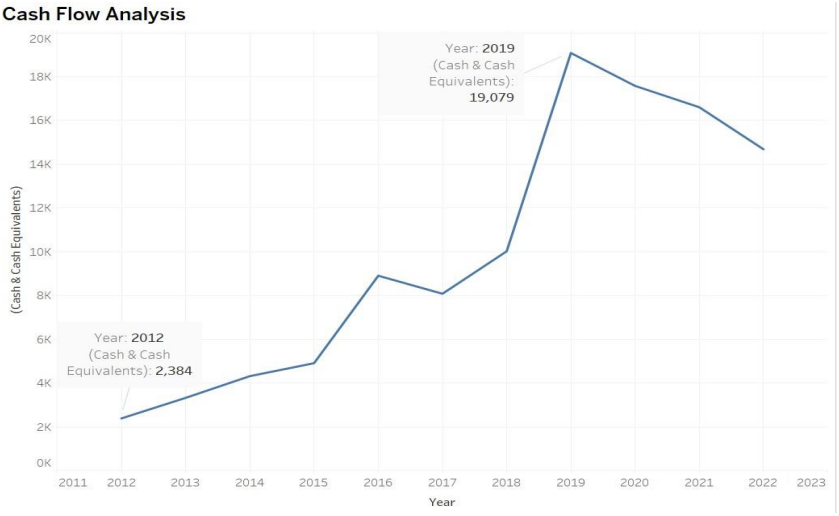
Robust Profitability amidst Cost Management

Bar Chart for Net Income Comparison:



Our analysis is complemented by the Bar Chart for Net Income Comparison graph, which reveals Meta's sustained profitability. Despite the uptick in expenses, net income has shown an upward trend, affirming the company's ability to manage costs and enhance profitability through diverse revenue channels. This is indicative of a well-executed balance between spending on growth and maintaining healthy profit margins.

Financial Stability and Liquidity



The liquidity of Meta is graphically represented in the Cash Flow Analysis graph, which shows a substantial increase in cash and cash equivalents over the years. This robust liquidity profile empowers Meta to pursue strategic acquisitions, invest in innovation, and weather economic fluctuations without compromising its operational momentum.

The integrated analysis of Meta Platforms Inc.'s historical financial health, supported by both the report and the illustrative graphs, paints a picture of a company that has not only grown but also optimized its financial operations over the years. Meta has demonstrated a consistent ability to generate revenue, manage expenses, maintain profitability, and ensure liquidity. As we look ahead, the company's historical financial health bodes well for its ability to seize new opportunities and tackle the evolving challenges in the digital arena.

IV. Historical Financial Ratios

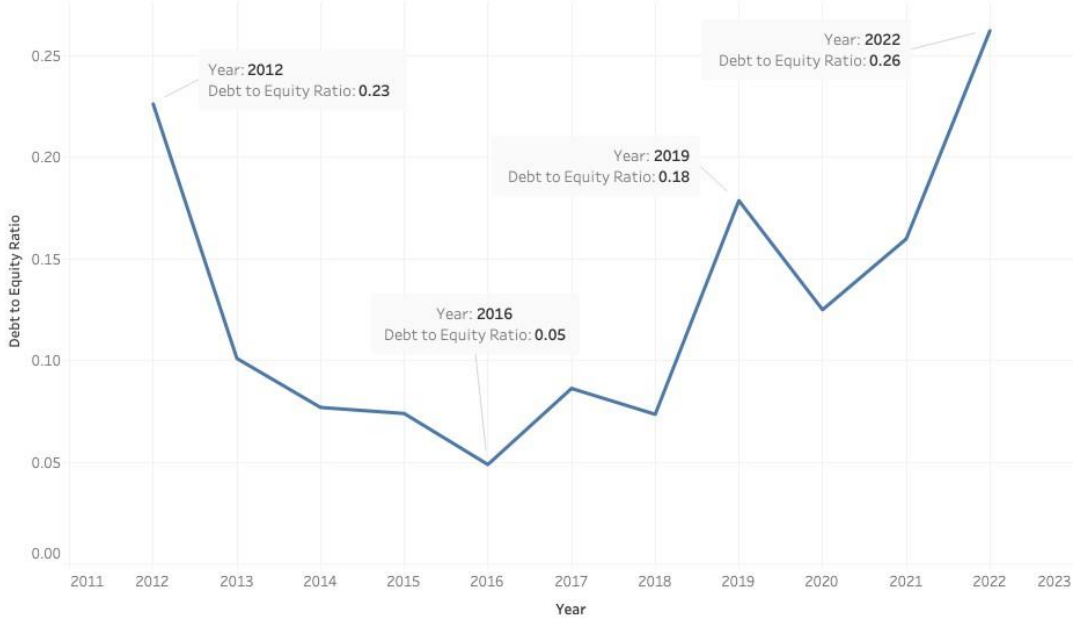
To further understand Meta's finances, we looked at their Debt to Equity ratio, Profitability Ratio, and their Current Ratio while correlating it with their history over the years to get a bigger picture.

Although a fairly good ratio, Meta's Debt to Equity Ratio shows an initial downward trend until around 2016. In these years Meta was acquiring many assets, expanding, and growing and having lower debt. In the early years, Meta was in a growing phase. Meta's current ratio helps to explain this as well. In those first 4-5 years, Meta's current ratio also soared to their highest ratio, 12.92 in 2017. If you also take a look into their profitability ratios, you will also notice that this also sees an upward trend in these first few years of Meta. The years from 2012-2017 could be considered growth/expansion stages for them.

If we take a look at the latter half of the years, roughly from 2017-2022, we can see that Meta is facing more of an up/down trend. They are beginning to acquire more debt, more assets, and face more public situations that have put the company at risk. We begin to see more fluctuation as the years go by. In 2019, we see an increase in their debt to equity ratio, a huge drop in their current ratio, and a large decrease in their profitability as well. We can tie this to the presidential scandal involving Facebook and the 2016 presidential election, as well as the $5 million fine faced pertaining to user data privacy. In 2021 we saw their data breach where nearly 5 million users were impacted, as well as a domino effect into 2022. These can also be seen

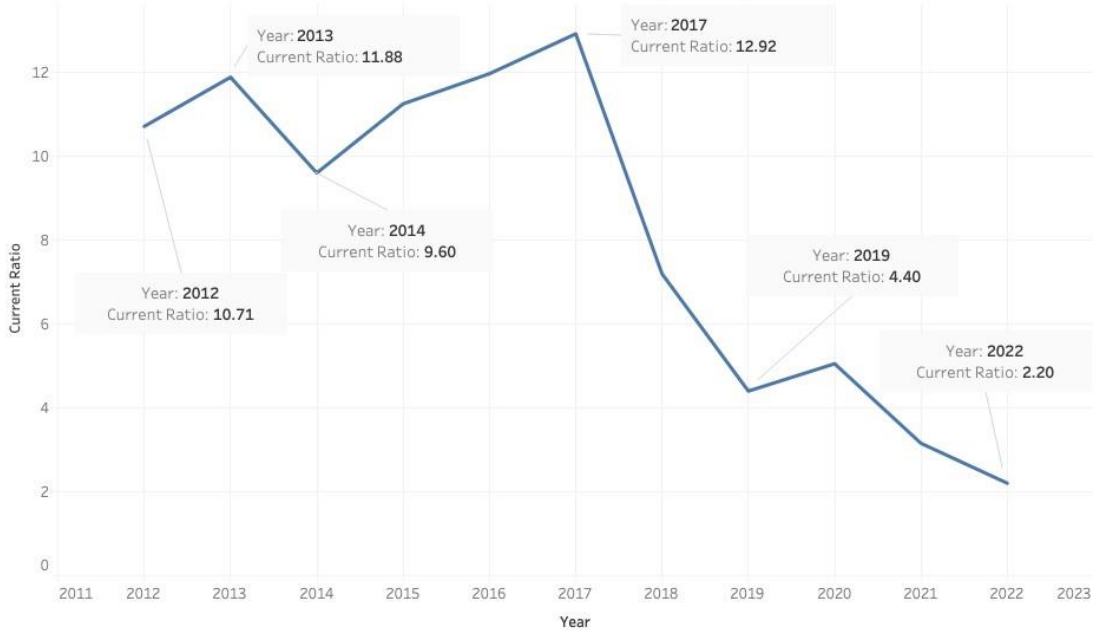
with their ratios, debt to equity reached their highest, current ratio seen an all time low, and profitability also seen a large decrease.

Debt to Equity Ratio Visualization

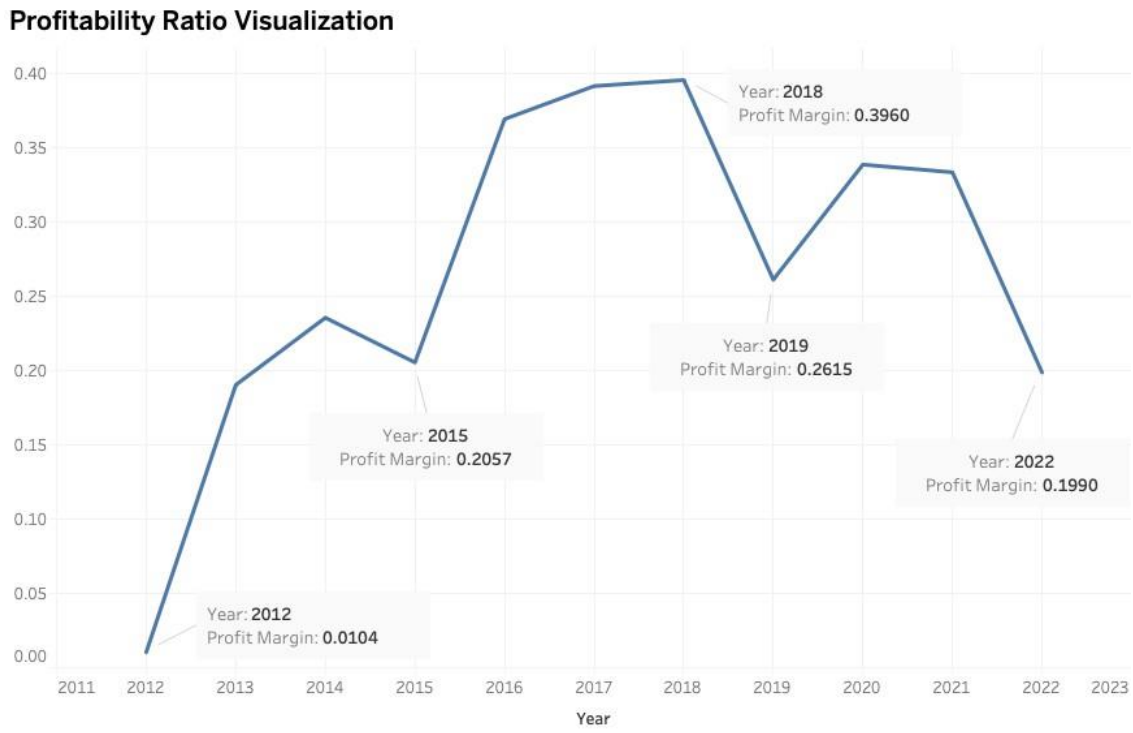


The trend of Debt to Equity Ratio for Year.

Current Ratio Visualization:



The trend of sum of Current Ratio for Year.



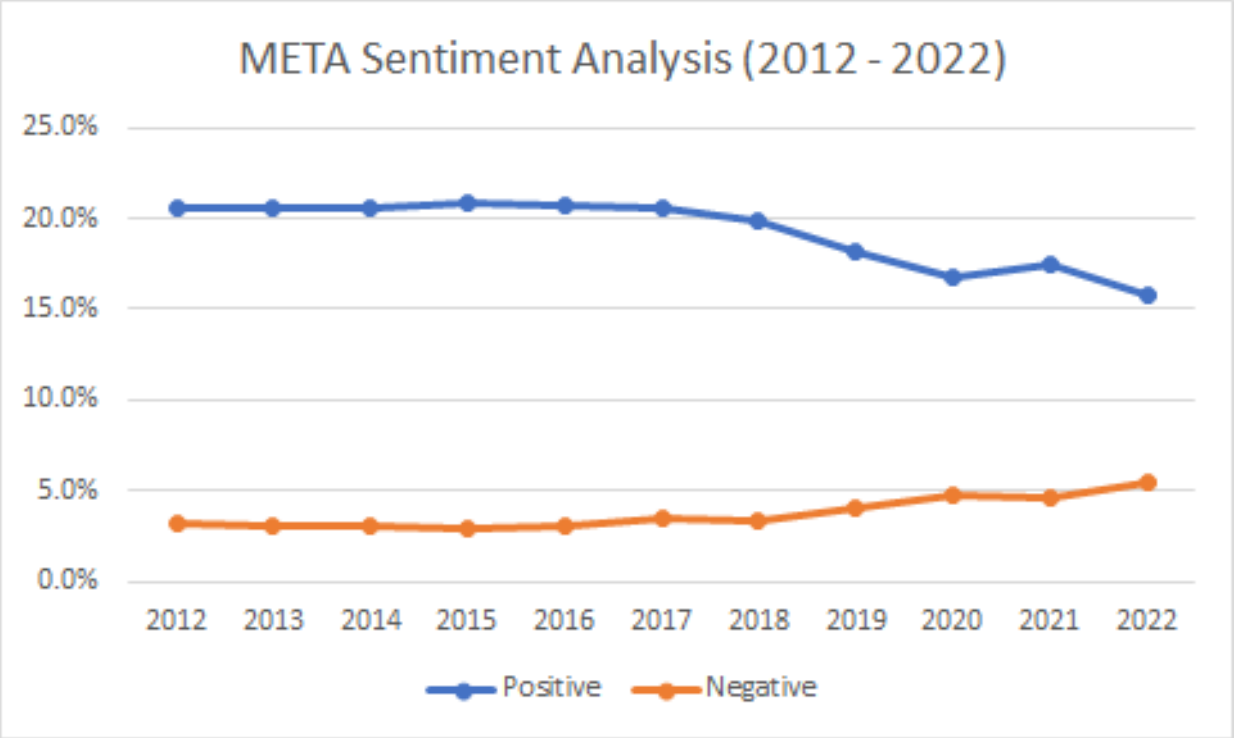
The trend of sum of Profit Margin for Year. The view is filtered on Year, which ranges from 2012 to 2022.

V. Analysis of Sentiment and WordClouds

Sentiment Analysis Overview

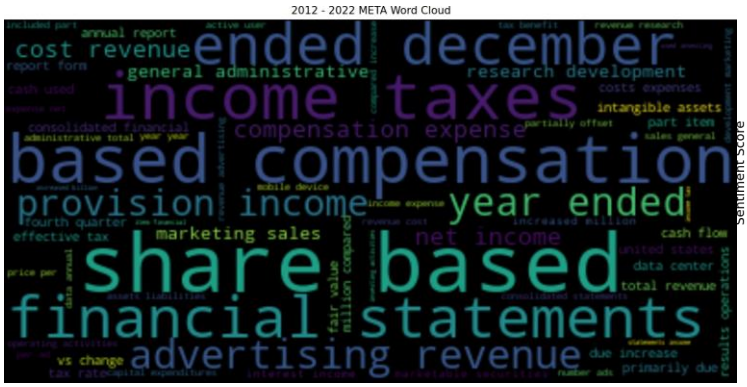
The upward and downward trends of Meta's financial ratios examined in the previous section of this analysis go hand-in-hand with the trends of the sentiment analysis we performed on the MD&A sections of Meta's 10K filings from 2012 - 2022. Between 2012 and 2016, the positive sentiment of the words being used in their MD&As steadily increased with a peak in 2015 where 20.8% of words used had positive sentiment. Similarly, the negative sentiment of the words being used in Meta's MD&As stayed relatively stable, with their 2015 MD&A having the least amount of negative sentiment with 2.9% of words having negative connotations. Much like Meta's financial ratios, the sentiment of their MD&As started a negative trend after 2017. Between 2017 and 2022, the positive sentiment of the words being used in Meta's MD&As

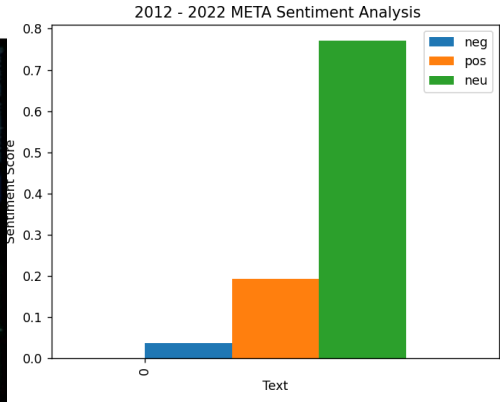
trended downwards, with an all-time low in 2022 with only 15.7% of words having positive connotations. Similarly, the negative sentiment of the MD&As in this time period increased, with an all-time high of negative sentiment being seen in the 2022 MD&A where 5.4% of words had negative connotations. The graph below shows the changes in positive and negative sentiment of Meta's MD&A sections over time, and shows the shift in trends around 2017.



The rest of the sentiment analysis section of this report will examine the sentiment and WordClouds extracted from Meta's 10K MD&A sections for each individual year between 2012 - 2022, along with reasoning behind increases and decreases in sentiment based off of financial performance and/or events and controversies surrounding Meta and their operations.

2012 - 2022 (WordCloud & Sentiment Analysis for All Years)



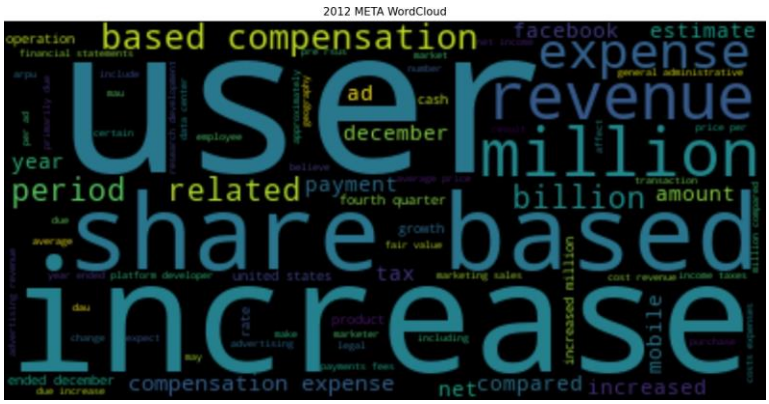


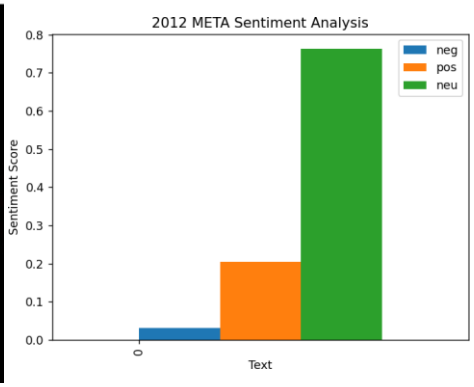
Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.7%	19.3%	77.1%

Overall between 2012 and 2022, Meta's MD&A sections of their 10K filings had on average 3.7% negative sentiment, 19.3% positive sentiment, and 77.1% neutral sentiment. These averages will be used as a baseline for comparison as each individual year's MD&A sentiment is examined throughout the remainder of the sentiment analysis section of this report.

Three of the most prominent words and phrases on the WordCloud generated from all of Meta's MD&As between 2012 and 2022 are "share based", "based compensation", and "advertising revenue". The first two phrases "share based" and "based compensation" likely refer to Meta's acquisitions of platforms like Instagram, WhatsApp, and Oculus, all of which heavily relied on financing through stocks (See Section I. Company Background). The third phrase, "advertising revenue", likely refers to the fact that advertising is Meta's greatest source of revenue on their social media platforms Facebook and Instagram.

2012



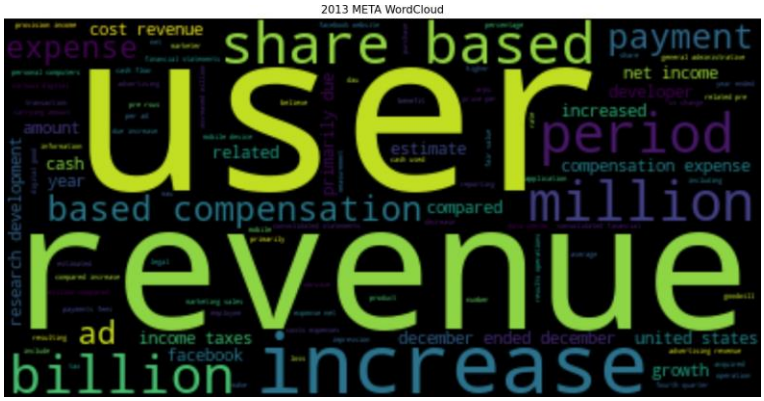


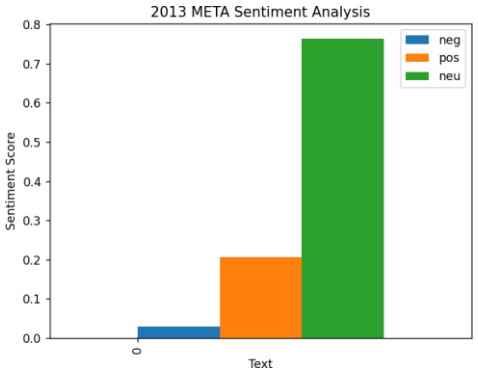
Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.2%	20.5%	76.3%

In 2012, Meta's MD&A section of their 10K filing was made up of 3.2% words with negative sentiment (lower than the 3.7% average) and 20.5% words with positive sentiment (higher than 19.3% average). Being their first year as a public company, there is no baseline to compare Meta to, although on paper Meta was in good shape. In 2012, Meta filed for their IPO (under their former company name Facebook, Inc.), purchased Instagram for $1 billion in cash and stock, and their flagship Facebook social media platform reached 1 billion active users.

Two of the most prominent words and phrases on the WordCloud generated from Meta's 2012 MD&A are "share based" and "user". The first phrase "share based" likely refers to Meta's acquisition of Instagram, which was purchased mostly with stock and some cash (See Section I. Company Background). The second phrase, "user", likely refers to user metrics discussed in the MD&A, including user growth of their Facebook and Instagram platforms.

2013



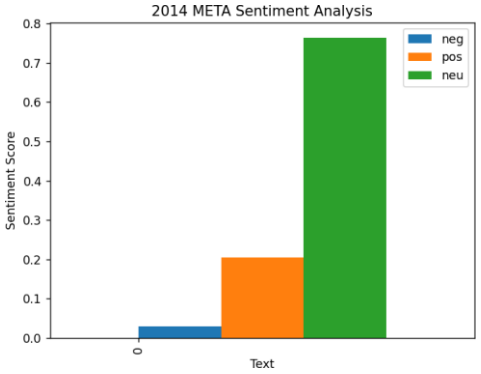


Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.0%	20.6%	76.4%

In 2013, Meta's MD&A section of their 10K filing was made up of 3.0% words with negative sentiment (lower than the 3.7% average) and 20.6% words with positive sentiment (higher than 19.3% average). Compared to the previous year, 2013 had less negative sentiment and more positive sentiment. This more positive outlook is likely due to the much lower DE ratio, higher current ratio, and much greater profitability seen by Meta in 2013 when compared to 2012 (see Section IV. Historical Financial Ratios).

Two of the most prominent words and phrases on the WordCloud generated from Meta's 2013 MD&A are "revenue" and "user". The first phrase "revenue" likely refers to high increases in revenue which can be seen in Meta's financial ratios between 2012 and 2013 (see Section IV. Historical Financial Ratios). The second phrase, "user", likely refers to user metrics discussed in the MD&A, including user growth of their Facebook and Instagram platforms and potential user growth after the 2013 implementation of Meta's "internet.org" initiative/service (See Section I. Company Background).

2014

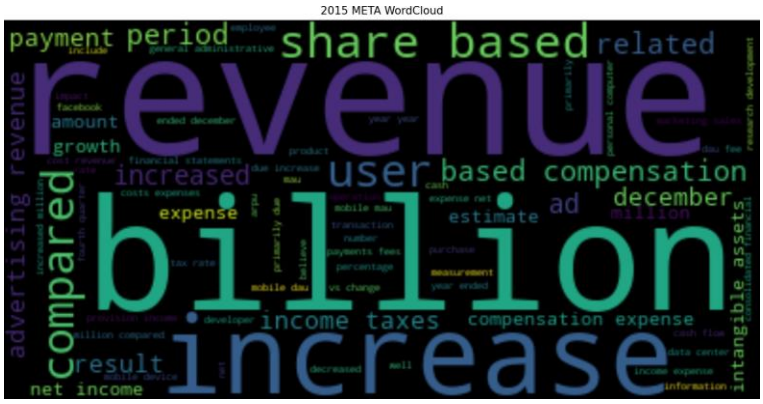


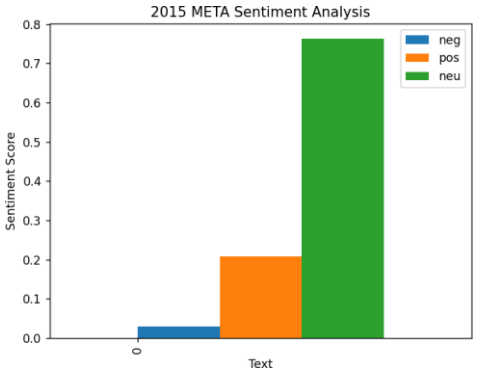
Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.0%	20.5%	76.4%

In 2014, Meta's MD&A section of their 10K filing was made up of 3.0% words with negative sentiment (lower than the 3.7% average) and 20.5% words with positive sentiment (higher than 19.3% average). Compared to the previous year, 2014's MD&A section was fairly comparable as the amount of negative sentiment stayed the same and positive sentiment only decreased by 0.1%. This is likely due to Meta's 2014 operations being similarly strong as they were in 2013, since Meta's DE ratio continued dropping, profitability kept increasing, and more major acquisitions were made including the acquisitions of WhatsApp and Oculus.

Two of the most prominent words and phrases on the WordCloud generated from Meta's 2014 MD&A are "revenue" and "share based". The first phrase "revenue" likely refers to more increases in revenue (beyond the impressive increases seen in 2013) which can be seen in Meta's financial ratios between 2013 and 2014 (see Section IV. Historical Financial Ratios). The second phrase, "share based", likely refers to Meta's acquisition of WhatsApp and Oculus, which were both purchased mostly with stock and some cash (See Section I. Company Background).

2015



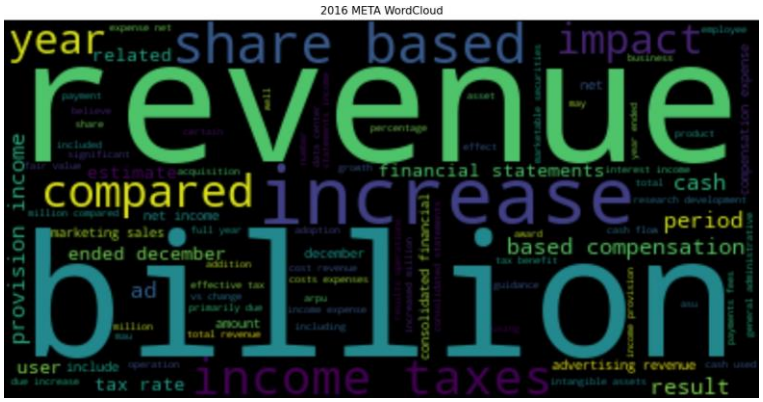


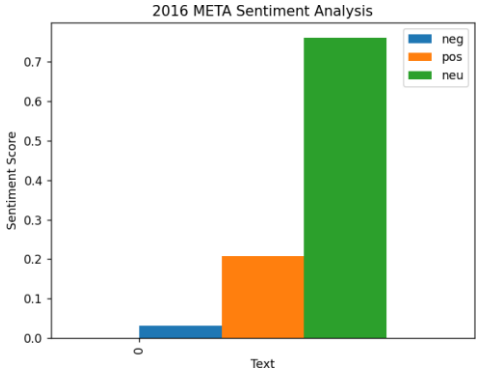
Sentiment Analysis - % Output		
Negative	Positive	Neutral
2.9%	20.8%	76.3%

In 2015, Meta's MD&A section of their 10K filing was made up of 2.9% words with negative sentiment (lower than the 3.7% average) and 20.8% words with positive sentiment (higher than 19.3% average). Compared to the previous year, 2015's MD&A section contained slightly more positive words and slightly less negative words. This is likely due to Meta's 2015 operations being similarly strong as they were in 2014, since Meta's DE ratio continued dropping, their current ratio increased, and Facebook's mobile app became the most downloaded mobile app of the year. Despite this, Meta's profitability had a slight drop.

The most prominent word on the WordCloud generated from Meta's 2015 MD&A was "revenue", likely referring to revenue sources in addition to a drop in profitability for the first time.

2016



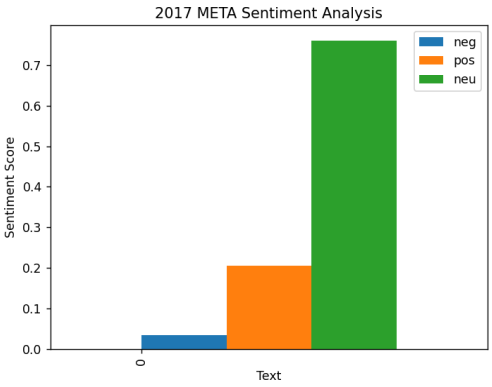


Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.1%	20.7%	76.1%

In 2016, Meta's MD&A section of their 10K filing was made up of 3.1% words with negative sentiment (lower than the 3.7% average) and 20.7% words with positive sentiment (higher than 19.3% average). Compared to the previous year, 2016's MD&A section contained slightly fewer positive words and slightly more negative words, making for an overall more negative MD&A. The more negative nature of 2016's MD&A compared to 2015 may not make sense at first glance, considering Meta's profitability and current ratio was the highest it had been up to that point, and their DE ratio hit an all time low. However, the greater amount of negativity found in Meta's 2016 MD&A could be attributed to the amount of controversy and impact on advertisers due to the role that Meta's Facebook platform played in the spread of misinformation regarding the 2016 U.S. Presidential election.

The most prominent word on the WordCloud generated from Meta's 2016 MD&A was "revenue", likely referring to the very strong year that Meta had financially, considering their high profitability and low DE ratio.

2017

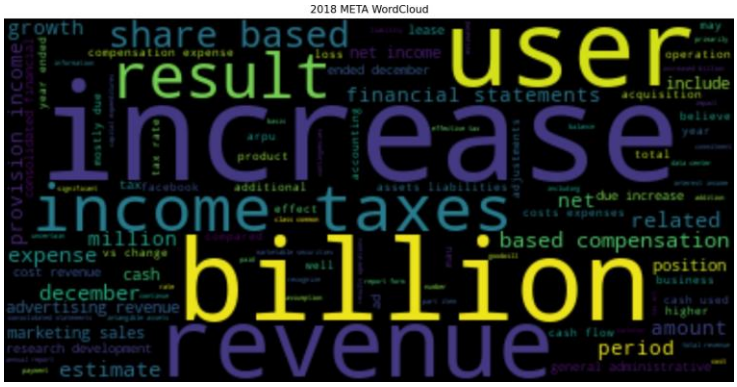


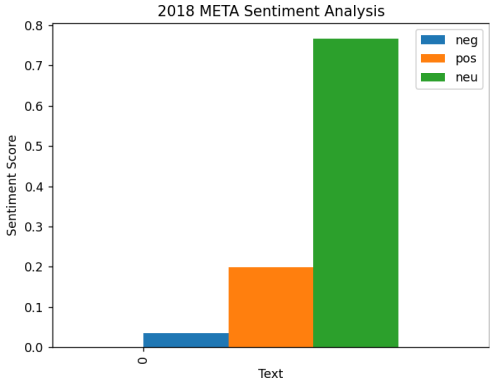
Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.5%	20.5%	76.0%

In 2017, Meta's MD&A section of their 10K filing was made up of 3.5% words with negative sentiment (lower than the 3.7% average) and 20.5% words with positive sentiment (higher than 19.3% average). Compared to the previous year, the sentiment of 2017's MD&A continued a negative trend of more negative sentiment and less positive sentiment. Although the sentiment of Meta's 2017 MD&A remains above their average positive sentiment levels, they are closer than ever to crossing the line to becoming below average positivity. The decrease in positive sentiment and increase in negative sentiment is likely due to an increase in Meta's DE ratio compared to 2016 and ongoing controversy following the 2016 U.S. Presidential election.

The most prominent word on the WordCloud generated from Meta's 2017 MD&A was "revenue", likely referring to the very strong year that Meta had in regards to profitability, as in 2017 Meta had strong revenue growth and had a higher profitability ratio than they had ever had.

2018



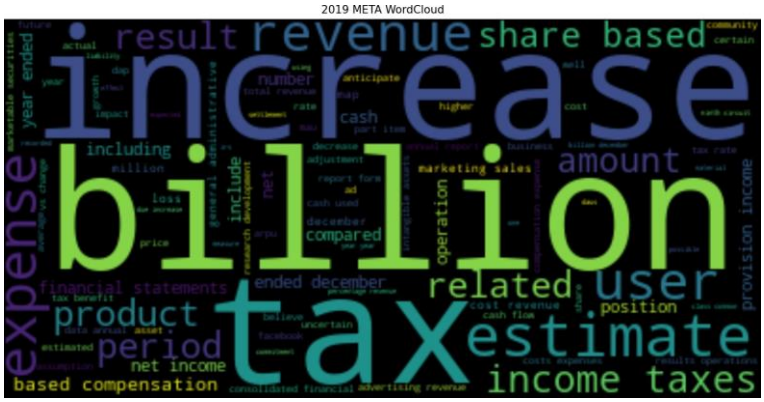


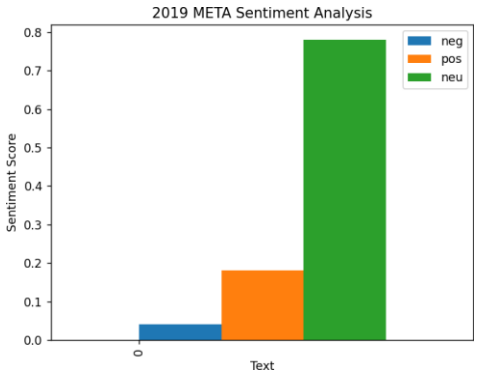
Sentiment Analysis - % Output		
Negative	Positive	Neutral
3.4%	19.9%	76.7%

In 2018, Meta's MD&A section of their 10K filing was made up of 3.4% words with negative sentiment (lower than the 3.7% average) and 19.9% words with positive sentiment (higher than 19.3% average). Compared to the previous year, the sentiment of 2018's MD&A continued the negative trend of more negative sentiment and less positive sentiment. Although the sentiment of Meta's 2018 MD&A remains above their average positive sentiment levels, they are closer than ever to crossing the line to becoming below average positivity. The decrease in positive sentiment and increase in negative sentiment is likely due to a sharp decrease in Meta's quick ratio, as well as several controversies surrounding Meta including the Cambridge Analytica scandal and CEO Mark Zuckerberg's testimony to the U.S. Congress about privacy rights (see Section I. Company Background).

Two of the most prominent words and phrases on the WordCloud generated from Meta's 2018 MD&A are "revenue" and "user". The first phrase "revenue" likely refers to high increases in revenue, and an increase in Meta's profitability ratio (which is at an all-time high in 2018). The second phrase, "user", likely refers to discussion on user metrics and potentially user privacy, as such was a major topic for Meta in 2018 considering the Cambridge Analytica scandal and testimony to the U.S. Congress about user data.

2019



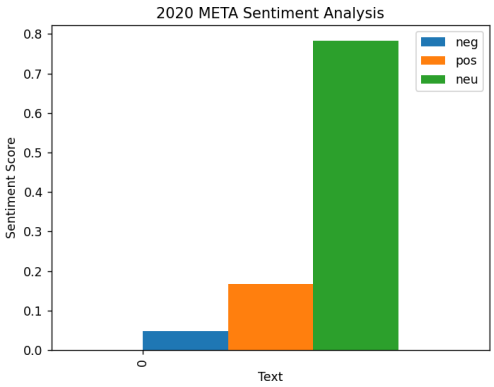


Sentiment Analysis - % Output		
Negative	Positive	Neutral
4.0%	18.1%	78.0%

In 2019, Meta's MD&A section of their 10K filing was made up of 4.0% words with negative sentiment (higher than the 3.7% average) and 18.1% words with positive sentiment (lower than 19.3% average). Compared to all other years, 2019 is the most negative year in regards to the sentiment in Meta's MD&As and is the first time where the sentiment levels drop below Meta's all time average. The large drop in positivity can be attributed to an increase in expenses, sharp increase in DE ratio, drop in current ratio, and sharp drop in profitability. In addition to a poor year financially when compared to previous years, in 2019 Meta also faced a
$5 billion fine from the FTC over user data privacy violations.

The most prominent word on the WordCloud generated from Meta's 2019 MD&A was "billion", which could refer to billions in income and expenses, in addition to mentions of the $5 billion fine that Meta faced from the FTC over user data privacy violations.

2020

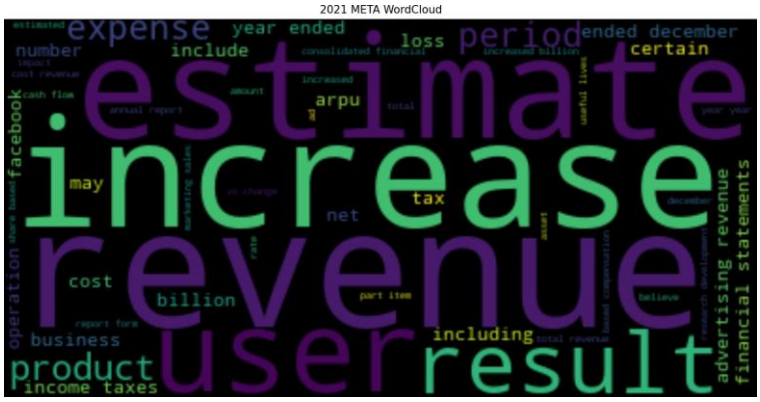


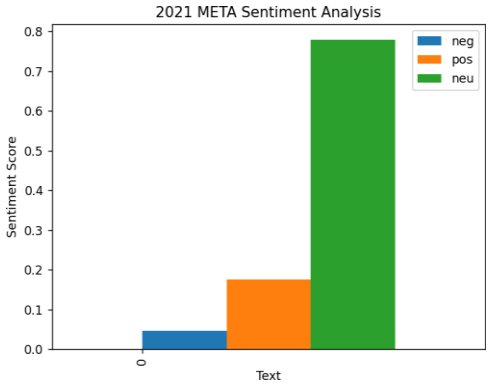
Sentiment Analysis - % Output		
Negative	Positive	Neutral
4.8%	16.8%	78.3%

In 2020, Meta's MD&A section of their 10K filing was made up of 4.8% words with negative sentiment (higher than the 3.7% average) and 16.8% words with positive sentiment (lower than 19.3% average). Compared to 2019 (their most negative year yet), Meta's 2020 MD&A continues a sharp downward trend in decrease of positive sentiment and increase in negative sentiment. Despite 2020 being financially better in almost every way compared to 2019 for Meta, the decrease in positivity could be attributed to economic and social effects of the COVID-19 pandemic, as well as more controversy surrounding Meta's operations including U.S. Congressional hearings involving Meta about tech monopolies, and controversy surrounding Facebook's role in the spread of misinformation in the 2020 U.S. Presidential election (see Section I. Company Background).

The two most prominent phrases on the WordCloud generated from Meta's 2020 MD&A were "revenue" and "increase", which could refer to the financial improvement that Meta saw in 2020 compared to 2019, despite the negative events that Meta faced in 2020.

2021



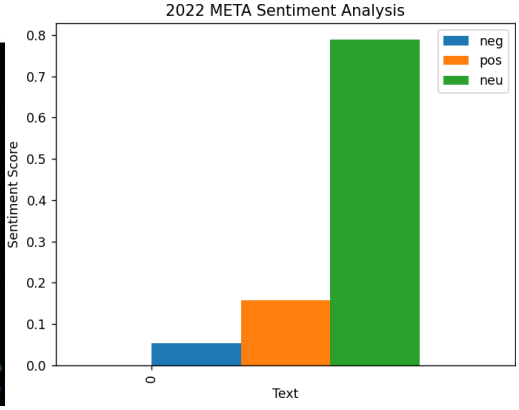


Sentiment Analysis - % Output		
Negative	Positive	Neutral
4.6%	17.5%	77.9%

In 2021, Meta's MD&A section of their 10K filing was made up of 4.6% words with negative sentiment (higher than the 3.7% average) and 17.5% words with positive sentiment (lower than 19.3% average). Compared to 2020 (their most negative year yet), Meta's 2021 MD&A showed slight improvement with slightly higher positive sentiment and slightly lower negative sentiment. While still below their average sentiment levels, the slight improvement of MD&A sentiment could be attributed to a record high stock price for Meta being hit on September 6th (at $378.69 per share) and record high net income. Despite this, the overall sentiment levels are still negative which is likely due to increased expenses, an increase in DE ratio, and a slight decrease in profitability. Additionally, in 2021 Meta faced further trouble as over 530 million Facebook users had their data leaked and exposed, and controversy spread as internal Meta documents were leaked revealing ethically questionable policies that were encouraged by management (see Section I. Company Background).

One of the most prominent phrases on the WordCloud for Meta's 2021 MD&A, "user", likely refers to the leaking of data of over 530 million Facebook users. Other words like "revenue" and "increase" may refer to the improved financials that Meta saw when compared to 2020.

2022



Sentiment Analysis - % Output		
Negative	Positive	Neutral
5.4%	15.7%	78.9%

In 2022, Meta's MD&A section of their 10K filing was made up of 5.4% words with negative sentiment (higher than the 3.7% average) and 15.7% words with positive sentiment (lower than 19.3% average). Compared to all other years, Meta's 2022 MD&A contained the most amount of negative sentiment and the least amount of positive sentiment. This is likely because 2022 was Meta's worst year financially, as their expenses hit a record high, net income dropped sharply, their DE ratio hit a record high, and profitability dropped significantly. Additionally, the average number of active daily users of Meta's flagship Facebook platform dropped for the first time, and Meta's stock price hit its lowest point since 2015 on October 31st (at $90.79 per share), which was a huge drop compared to the peak of Meta's stock price from 2021.

The two most prominent phrases on the WordCloud generated from Meta's 2022 MD&A were "revenue" and "product", which could refer to drops in revenue and discussion of their product performance in Meta's worst year to date.

VI. Conclusion and Future Work

There is a very strong correlation between the company's financial health and their sentiment. In the years 2012-2017, we see higher positivity in overall sentiment. Meta is in their growing/expansion stage, acquiring new assets, building, and making profit. The stakeholders and shareholders are very pleased to see this. It is reflected in their $10k as their more profitable years. While the company was expanding, they were not in the spotlight for much wrong doing. Facebook was a number one app many years in a row, and obtaining WhatsApp, Oculus, and Instagram were very big wins for the company and their growth.

As for the second half, we begin to see more negativity come across in the sentiment analysis, rightfully so with all of the challenges the company faced. Meta took on a large amount of liability and debt between the 2016 presidential scandal proceedings with Cambridge Analytica, a fine for data user privacy, controversy over fact checking posts leading up to the 2020 presidential election, and the boycott pertaining to the removal of certain information. Meta also faced a data breach where about 530 million users' information was impacted and documents were leaked with info on Meta's management groups knowledge of events. These events in recent years have really hurt Meta and had an impact on their sentiment, profitability, and image.

Looking into the future, it is very important for Meta to gain back the trust of their stakeholders. After more than one instance pertaining to users and their privacy, that is very important ground to make up. Limiting their liabilities will also be very important as they have already acquired many between their assets as well as the fines they faced. It will be important to leverage their assets to create more profit and lessen their liabilities. If Meta continues on the same track, they are likely to see a decline stage of their growth cycle. If Meta takes action now to address the events in the more recent years, with all the assets they have acquired and expanded with, they could make a good turnaround and be very profitable and successful.

VII. Appendix

Python Script 1: *Text Mining for a Single Year's MD&A*

```
# --------------
# Libraries
# --------------
import pandas as pd
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import re
import numpy as np
from sec_api import ExtractorApi
from nltk.sentiment import SentimentIntensityAnalyzer
import nltk
from string import digits
from nltk.tokenize import word_tokenize
from wordcloud import WordCloud, STOPWORDS
from nltk.corpus import stopwords
nltk.download('stopwords')
nltk.download('punkt')
nltk.download('vader_lexicon')



# ----------------
# Extracting Text
# ----------------

# Year of Filing - Input by user in terminal. Changes which filing is used
from the below links and the year label in the visuals
year_10k = int(input('Input Desired Year: '))

# Meta 10K SEC File Links
```

```
url_2022 =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680123000013/meta-20221231.htm"
url_2021 =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680122000018/fb-20211231.htm"
url_2020 =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680121000014/fb-20201231.htm"
url_2019 =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680120000013/fb-12312019x10k.htm"
url_2018 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680119000009/fb-12312018x10k.htm"
url_2017 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680118000009/fb-12312017x10k.htm"
url_2016 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680117000007/fb-12312016x10k.htm"
url_2015 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680116000043/fb-12312015x10k.htm"
url_2014 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680115000006/fb-12312014x10k.htm"
url_2013 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680114000007/fb-12312013x10k.htm"
url_2012 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680113000003/fb-12312012x10k.htm"
test_url =
"https://www.sec.gov/Archives/edgar/data/1318605/000156459021004599/tsla-10k_20201231.htm"
```

```
# API Key to connect to the SEC API
APIKEY = "INSERT YOUR API KEY HERE"

extractorApi = ExtractorApi(APIKEY)

# Meta 10K URL
if year_10k == 2022:
    url_10k = url_2022
if year_10k == 2021:
    url_10k = url_2021
if year_10k == 2020:
    url_10k = url_2020
if year_10k == 2019:
    url_10k = url_2019
if year_10k == 2018:
    url_10k = url_2018
if year_10k == 2017:
    url_10k = url_2017
if year_10k == 2016:
    url_10k = url_2016
if year_10k == 2015:
    url_10k = url_2015
if year_10k == 2014:
    url_10k = url_2014
if year_10k == 2013:
    url_10k = url_2013
if year_10k == 2012:
    url_10k = url_2012
if year_10k == 2001:
    url_10k = test_url

# get the standardized and cleaned text of section 7 (MD&A)
text_format_10k = extractorApi.get_section(url_10k, "7", "text")

# ------------------
```

```
# Cleaning Data
# -----------------

#Remove unwanted characters
clean_10k = re.sub('\W+',' ', text_format_10k)

# Convert data to lower case
clean_10k=clean_10k.lower()

# Remove all the digits
remove_digits = str.maketrans('', '', digits)
clean_10k = clean_10k.translate(remove_digits)

# Remove stop words
pattern = re.compile(r'\b(' + r'|'.join(stopwords.words('english')) +
r')\b\s*')
clean_10k = pattern.sub('', clean_10k)

# Remove single letters
tmp = re.sub(r'\b\w\b', ' ', clean_10k)
clean_10k = re.sub(r'\s{2,}', ' ', tmp).strip()

# Remove table words and roman numbers
clean_10k = clean_10k.replace('table_start', '')
clean_10k = clean_10k.replace('table_end', '')
clean_10k = clean_10k.replace('ii', '')

# Tokenization of filing
token_10k = word_tokenize(clean_10k)

# Checks if any data is found in the filing (bug checking)
if clean_10k == "":
    print("No data found in filing")
else:
    print("Data Found!")#clean_10k)
```

```
# Calculate most frequent words in the filing
frequentWords = nltk.FreqDist(token_10k)
frequentWords.tabulate(10)

# Get stopwords
stopwords = set(STOPWORDS)

# WordCloud Generation
wordcloud = WordCloud(stopwords=stopwords, background_color="black",
max_words=100, ).generate(clean_10k)

# WordCloud Visualization
fig=plt.figure(figsize=(15, 8))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.title(str(year_10k) + ' META WordCloud')
plt.show()

#Sentiment Score for 10K Sec 7
sentiment = SentimentIntensityAnalyzer()
sentiment.polarity_scores(clean_10k)

# Save sentiment scores in dataframe and plot it
df_10k = sentiment.polarity_scores(clean_10k)
df_10k = pd.DataFrame(df_10k, index=[0])
df_10k = df_10k[["neg", "pos", "neu"]]
df_10k.plot.bar(align='edge', width=0.7)
plt.title(str(year_10k) + ' META Sentiment Analysis')
plt.xlabel('Text')
plt.ylabel('Sentiment Score')
print(df_10k)
plt.show()
```

Python Script 2: *Text Mining for All MD&As Combined*

```
# --------------
# Libraries
# --------------
import pandas as pd
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import re
from sec_api import ExtractorApi
from nltk.sentiment import SentimentIntensityAnalyzer
import nltk
from string import digits
from nltk.tokenize import word_tokenize
from wordcloud import WordCloud, STOPWORDS
from nltk.corpus import stopwords
nltk.download('stopwords')
nltk.download('punkt')
nltk.download('vader_lexicon')



# -------------------
# Extracting Text
# -------------------

# Meta 10K SEC File Links
url_2022                                                                =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680123000
0013/meta-20221231.htm"
url_2021                                                                =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680122000
0018/fb-20211231.htm"
url_2020                                                                =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680121000
0014/fb-20201231.htm"
```

```
url_2019 =
"https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680120000013/fb-12312019x10k.htm"
url_2018 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680119000009/fb-12312018x10k.htm"
url_2017 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680118000009/fb-12312017x10k.htm"
url_2016 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680117000007/fb-12312016x10k.htm"
url_2015 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680116000043/fb-12312015x10k.htm"
url_2014 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680115000006/fb-12312014x10k.htm"
url_2013 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680114000007/fb-12312013x10k.htm"
url_2012 =
"https://www.sec.gov/Archives/edgar/data/1326801/000132680113000003/fb-12312012x10k.htm"

# API Key to connect to the SEC API
APIKEY = "INSERT YOUR API KEY HERE"

extractorApi = ExtractorApi(APIKEY)

# get the standardized and cleaned text of section 7 (MD&A)
text_format_10k = extractorApi.get_section(url_2012, "7", "text") +
extractorApi.get_section(url_2013, "7", "text") +
extractorApi.get_section(url_2014, "7", "text") +
extractorApi.get_section(url_2015, "7", "text") +
extractorApi.get_section(url_2016, "7", "text") +
```

```
extractorApi.get_section(url_2017,          "7",          "text")          +
extractorApi.get_section(url_2018,          "7",          "text")          +
extractorApi.get_section(url_2019,          "7",          "text")          +
extractorApi.get_section(url_2020,          "7",          "text")          +
extractorApi.get_section(url_2021,          "7",          "text")          +
extractorApi.get_section(url_2022,  "7", "text")

# --------------------
# Cleaning Data
# --------------------

#Remove unwanted characters
clean_10k = re.sub('\W+',' ', text_format_10k)

# Convert data to lower case
clean_10k=clean_10k.lower()

# Remove all the digits
remove_digits = str.maketrans('', '', digits)
clean_10k = clean_10k.translate(remove_digits)

# Remove stop words
pattern  =  re.compile(r'\b('  +  r'|'.join(stopwords.words('english'))  +
r')\b\s*')
clean_10k = pattern.sub('', clean_10k)

# Remove single letters
tmp = re.sub(r'\b\w\b', ' ', clean_10k)
clean_10k = re.sub(r'\s{2,}', ' ', tmp).strip()

# Remove table words and roman numbers
clean_10k = clean_10k.replace('table_start', '')
clean_10k = clean_10k.replace('table_end', '')
clean_10k = clean_10k.replace('ii', '')

# Tokenization of filing
```

```
token_10k = word_tokenize(clean_10k)

# Checks if any data is found in the filing (bug checking)
if clean_10k == "":
    print("No data found in filing")
else:
    print(clean_10k)

# Calculate most frequent words in the filing
frequentWords = nltk.FreqDist(token_10k)
frequentWords.tabulate(10)

# Get stopwords
stopwords = set(STOPWORDS)

# WordCloud Generation
wordcloud = WordCloud(stopwords=stopwords, background_color="black",
max_words=100, ).generate(clean_10k)

# WordCloud Visualization
fig=plt.figure(figsize=(15, 8))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.title('2012 - 2022 META Word Cloud')
plt.show()

#Sentiment Score for 10K Sec 7
sentiment = SentimentIntensityAnalyzer()
sentiment.polarity_scores(clean_10k)

# Save sentiment scores in dataframe and plot it
df_10k = sentiment.polarity_scores(clean_10k)
df_10k = pd.DataFrame(df_10k, index=[0])
df_10k = df_10k[["neg", "pos", "neu"]]
df_10k.plot.bar(align='edge', width=0.7)
plt.title(' 2012 - 2022 META Sentiment Analysis')
```

```
plt.xlabel('Text')
plt.ylabel('Sentiment Score')
print(df_10k)
plt.show()
```