**Project1- Stock Data Analysis**

**This project talks about**

1. Stock price analysis within various sectors as well as industries and visualize the data.
2. Furthermore, it gives volume of stock traded in each sector and each industry.
3. And find any correlation within sectors to check if one stock was affected by another stock within the sector.
4. As well as calculates moving average in each sector and find the most traded stock in each year.

**This report answers the following questions:**

1. [**Which stocks are best performing in their respective sectors? (Sourav)**](#_Question_1:_Which)
2. [**Which ticker performed the worst/best by year?** (Sourav)](#_Question_2:_Which)
3. [**How did each sector perform for each year regarding average volume?** (Kunal)](#_Question3:_How_did)
4. [**Find a correlation between industries within each sector.** (Kunal)](#_Question4:_Find_a)
5. [**What does the moving average look like for each industry?** (Alison)](#_Question5:_What_does)
6. [**Which stock was most traded in each year?** (Alison)](#_Question_6:_)
7. **[What is the allocation of portfolio for each Sector?](#_Question7:_What_is)** [(Sakina)](#_Question7:_What_is)
8. [**What is the allocation of portfolio for each Industry?** (Sakina)](#_Question8:_What_is)

**Members of the group**

The members in this group are:

1. Alison Andrade (@alisonands2000)
2. Sourav Kumar (@s0uravk)
3. Sakina Jaffri (@SakinaJaffri)
4. Kunal Khirwar (@kunalkhirwar)

**Work breakdown structure**

Each member will perform data collection for given sectors: Alison: Tech Sourav: Banking/Finance Sakina: Healthcare Kunal: Automobiles

1. Pulling Data from AV
2. Cleaning Data
3. Merging Data
4. EDA
5. Data Visualization
6. Analysis

# Question 1: Which stocks are best performing in their respective sectors? (Sourav)

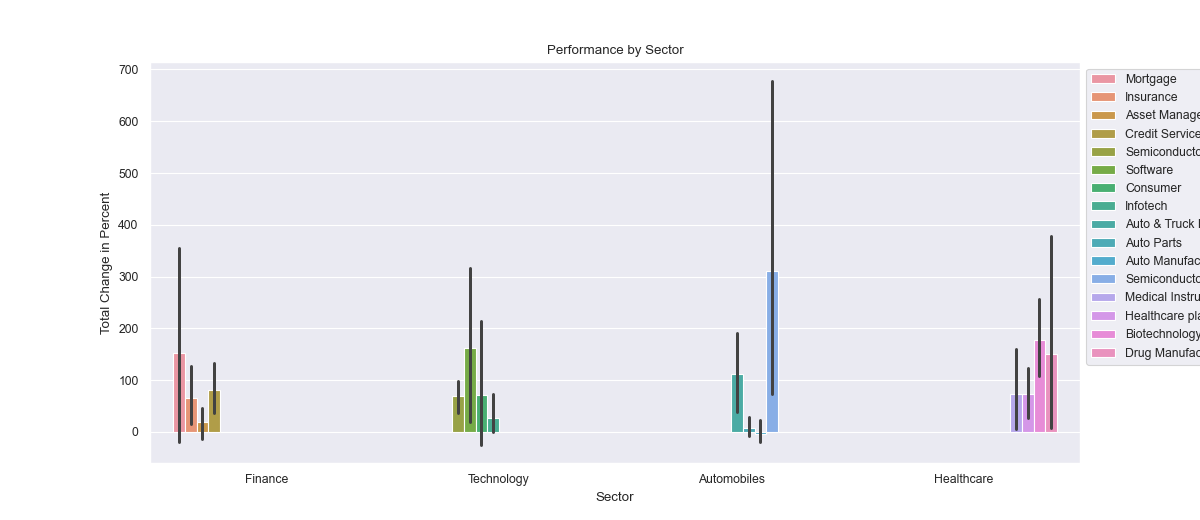
**Answer:** AMD- Automobile

COOP- Finance

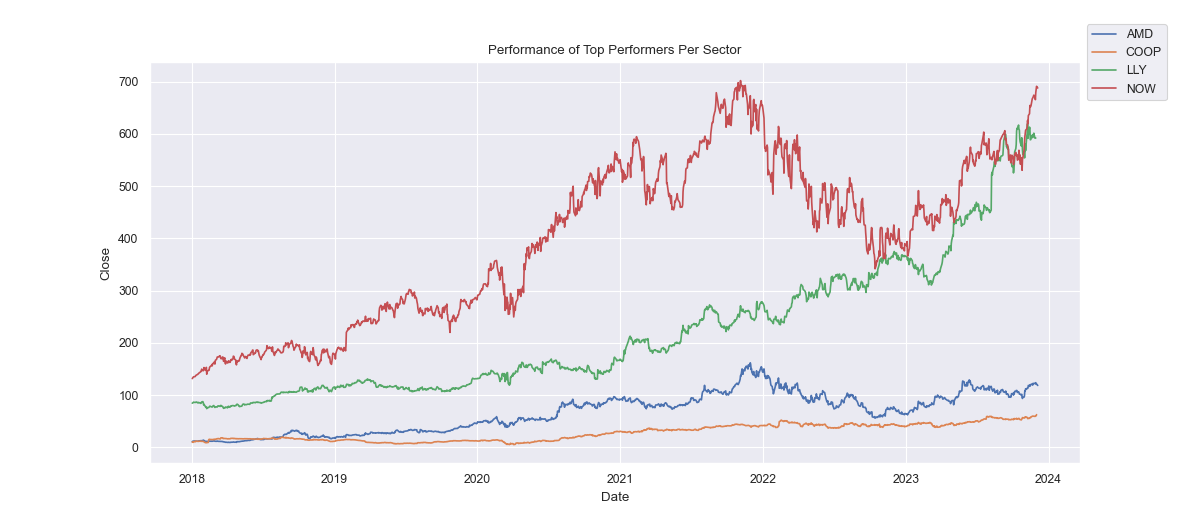
LLY- Healthcare

NOW- Technology

**Performance By Sector**



First all the stock data was processed and overall percent change was calculated for each stock in available data. Then, each Sector was plotted with the average change occurred in each industry using grouped bar graph. Where height of bar represents average of total change in percent for each industry in each sector and line in between the bar represent the amount of change occurred. In Finance Sector, Mortgage has the most change while in Technology, Software has most change and going further, Semiconductors in Automobiles and Drug Manufacturers in Healthcare has most change occurred in prices.

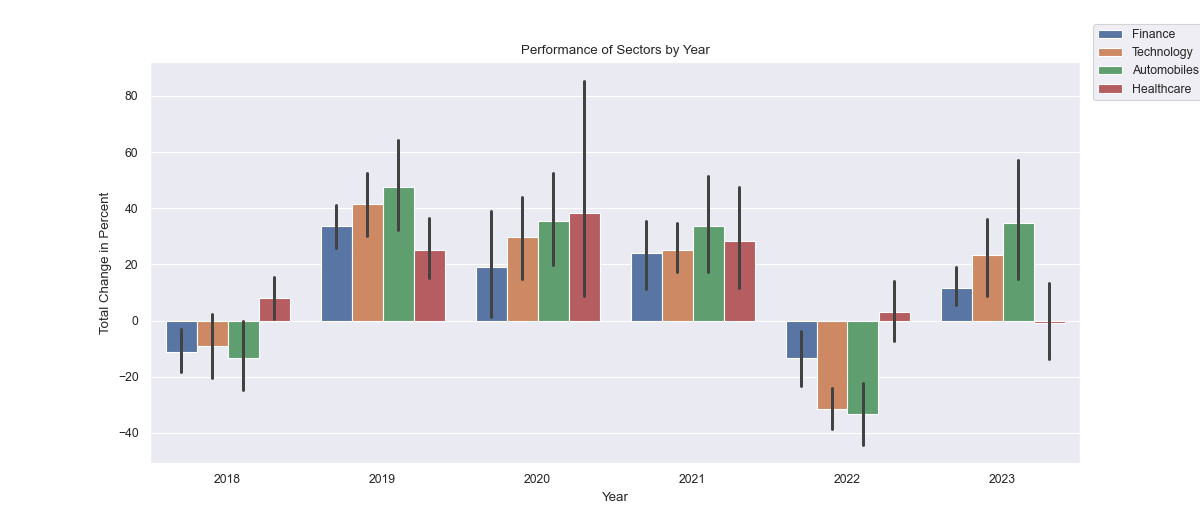
**Performance of Top Performers Per Sector**

As per the previous bar graph. We have the industry with most change in price for each sector. Then we found the best performing stocks in those industries and plotted the line graph for closing price of those stocks over the years available in dataset. 1037% of growth was seen in AMD where price went from 10.42 to 118.570000, while LLY had 600%, COOP had 507% and NOW had 424% increase in close price.

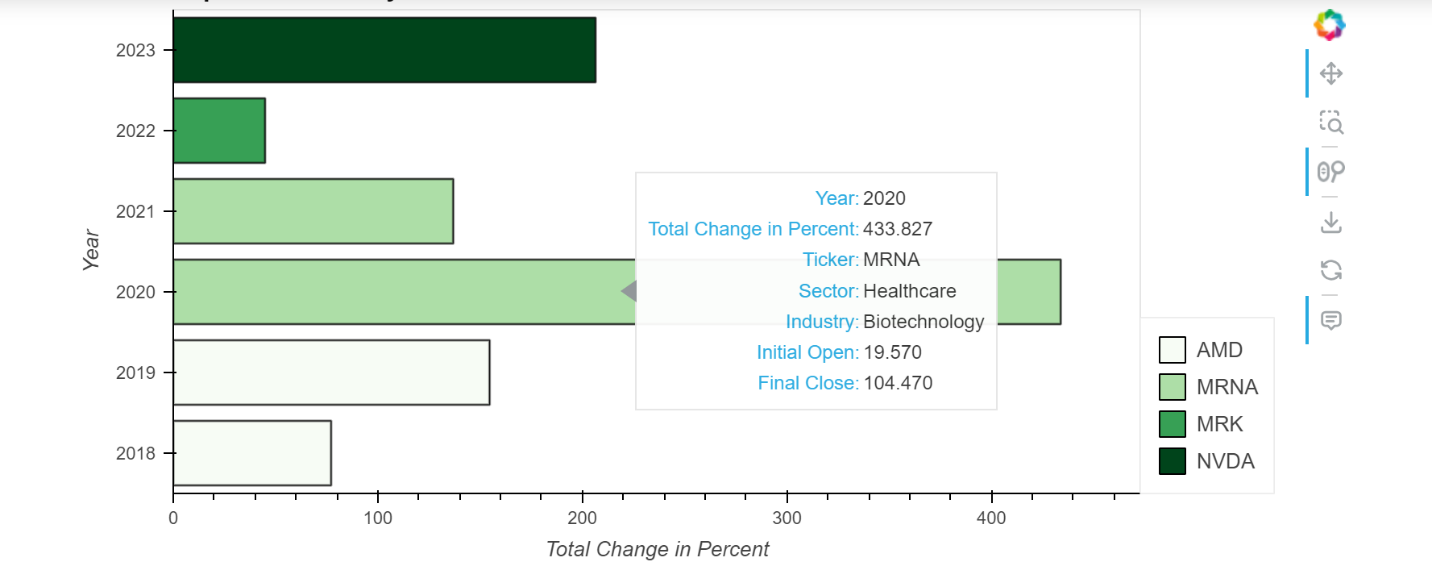
# Question 2: Which ticker performed the worst/best by year?

**Answer:**

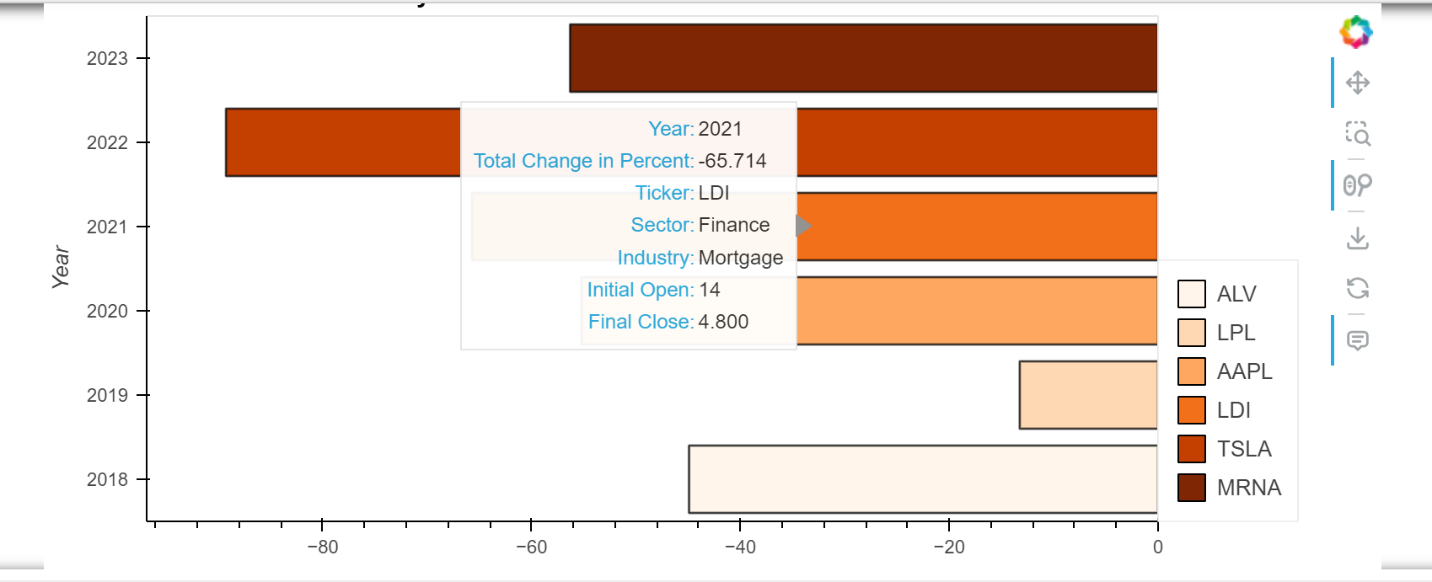
|  |  |  |
| --- | --- | --- |
| **Year** | **Best** | **Worst** |
| **2018** | AMD | ALV |
| **2019** | AMD | LPL |
| **2020** | MRNA | AAPL |
| **2021** | MRNA | LDI |
| **2022** | MRK | TSLA |
| **2023** | NVDA | MRNA |

**Performance of Sectors by Year**

This grouped bar graph shows the performance of each sector during different years. Based on the data we have, Only healthcare performed well in 2018 and 2022 while other Sectors were going down. Whereas, Automobiles was best performing sector in 2019, 2021 and 2023. While for 2020, Healthcare was the best performing sector with astonishing change in close price which could be due to covid-19.

**Best Performer Per Year**

Plotting the Best performing stock per Year based on the available data, for 2018 and 2019 AMD was best stock while checking 2020 MRNA was not only the top performer of that year but also the growth was radical as compared to other stocks performances in other years. It could be the effect of Covid-19 and Moderna was one of the major vaccines producing organization. MRNA was also top performer in 2021 but growth in price was lesser than previous years. Then MRK shows an increase of 44% and NVDA shows increase of more than 200% in 2022 and 2023 respectively.

**Worst Performer Per Year**

Plotting Worst performers by years, Worst performer for 2018 was ALV that decreased around 44%. for 2019, it was LPL but with just 13% decrease which is least of them all. In 2020, AAPL whose prices drop almost by half making it worst performer. Moving forward, LPL and TSLA were worst performer in 2021 and 2022 respectively. With TSLA being the stock with greatest decrease. In 2023, MRNA's price had a major drop compared to the previous horizontal bar chart where MRNA was Best performer for two consecutive years.

Checking both of these graphs, MRNA was common in both as MRNA's price increase maybe influenced by Covid-19 and as the vaccines were administered during 2022, MRNA experienced a subtle decrease in 2023 making it the worst performer for 2023.

# Question3: How did each sector perform for each year regarding average volume?

**Answer:** The bar graph below clearly shows that the Automobile sector was traded the most throughout the date range 2016-2023, while Finance was traded the least in terms of average volume.

**A graph with different colored bars

Description automatically generated**

# Question4: Find a correlation between industries within each sector.

**Answer:**

**Correlation within the Automobile sector**

**A chart of a car sales chart

Description automatically generated with medium confidence**

The above heat map shows a positive correlation of 0.83 between ‘Auto & Truck Dealerships’ and ‘Semiconductors’.

**Correlation within the Finance sector**

**A close-up of a chart

Description automatically generated**

The Finance sector shows several positive correlations between industries. The first correlation is between ‘Mortgage’ and ‘Asset Management’ which has a correlation coefficient of 0.9. While ‘Mortgage’ has a correlation of 0.91 with ‘Credit Services’.

There’s also a correlation between ‘Insurance’ and ‘Credit Services’ with a value of 0.83. ‘Asset Management’ and ‘Mortgage’ show a correlation of 0.9. ‘Asset Management’ also correlates 0.86 with ‘Credit Services’.

**Correlation within the Healthcare sector**

**A chart of a price

Description automatically generated with medium confidence**

In the Healthcare sector, the data showcased a correlation of 0.9 between ‘Healthcare Plans’ and ‘Biotechnology’. Another correlation exists between ‘Healthcare Plans’ and ‘Drug Manufacturers’ of a value of 0.91. Also, ‘Biotechnology’ correlates with ‘Drug Manufacturers’ with a 0.9 correlation coefficient. Interestingly, there’s a negative correlation of -0.017 between ‘Medical Instruments & Supplies’ and ‘Drug Manufacturers’. Though the negative correlation isn’t high, it draws attention.

**Correlation within the Tech sector**

**A chart of a software closing price

Description automatically generated with medium confidence**

In the Technology sector, we see a Correlation of 0.85 between ‘Semiconductors’ and ‘Software’. While ‘Semiconductors’ has a correlation of 0.91 with ‘Infotech’. Another pair of industries with a correlation is ‘Software’ and ‘Infotech’ with a value of 0.85.

# Question5: What does the moving average look like for each industry? (Alison)

**Answer:**

**Moving Average Analysis**

This method of moving averages is used to assess the overall sector performance during market declines, with a focus on understanding the impact of the pandemic and identifying sectors resilient to market factors. Utilizing Plotly and Matplotlib for visualizations, the period spanned the years Jan 1, 2012 to Dec 4, 2023. Moving averages were computed over 30, 60, and 150 days using the pandas.rolling() function, providing an overview of trends for each sector based on the cumulative closing prices.

|  |  |
| --- | --- |
|  |  |
|  |  |

**Technology**

Looking at the moving average for 150 days, it had a strong recovery after the covid decline during the pandemic and in the 2022 market decline.

**Finance**

Similar to the technology sector, finance experienced a strong recovery post-pandemic and in the 2022 market decline.

**Automotives**

The automotive sector exhibited higher volatility than finance and tech, with a significant upward trend post-pandemic and a steep decline during the 2022 market downturn.

**Healthcare**

Healthcare, resilient post-pandemic, faced a decline during the 2022 market downturn but recovered similarly to tech and finance.

In general, the finance and technology sectors recovered the best after market declines, followed by healthcare and automotives, which were the most volatile.

# Question 6: Which stock was most traded in each year? (Alison)

**Answer:**

**Highest Volume Traded Analysis**

This analysis focuses on identifying the most traded ticker in each year within different sectors, aiming to determine the impact of the pandemic on trading volumes. Using code, each sector is individually examined to determine the tickers with the highest annual trading volumes. This process involves grouping data by year, and summing the volume for each ticker over the entire year.

|  |  |
| --- | --- |
|  |  |
|  |  |

**Tech**

Apple witnessed a significant increase in volumes post-COVID, while Micron Tech had higher volumes than Apple before the pandemic.

**Finance**

Morgan Stanley and Visa were the most traded stocks, with volumes increasing during the pandemic and reverting to lower than pre-pandemic levels afterward.

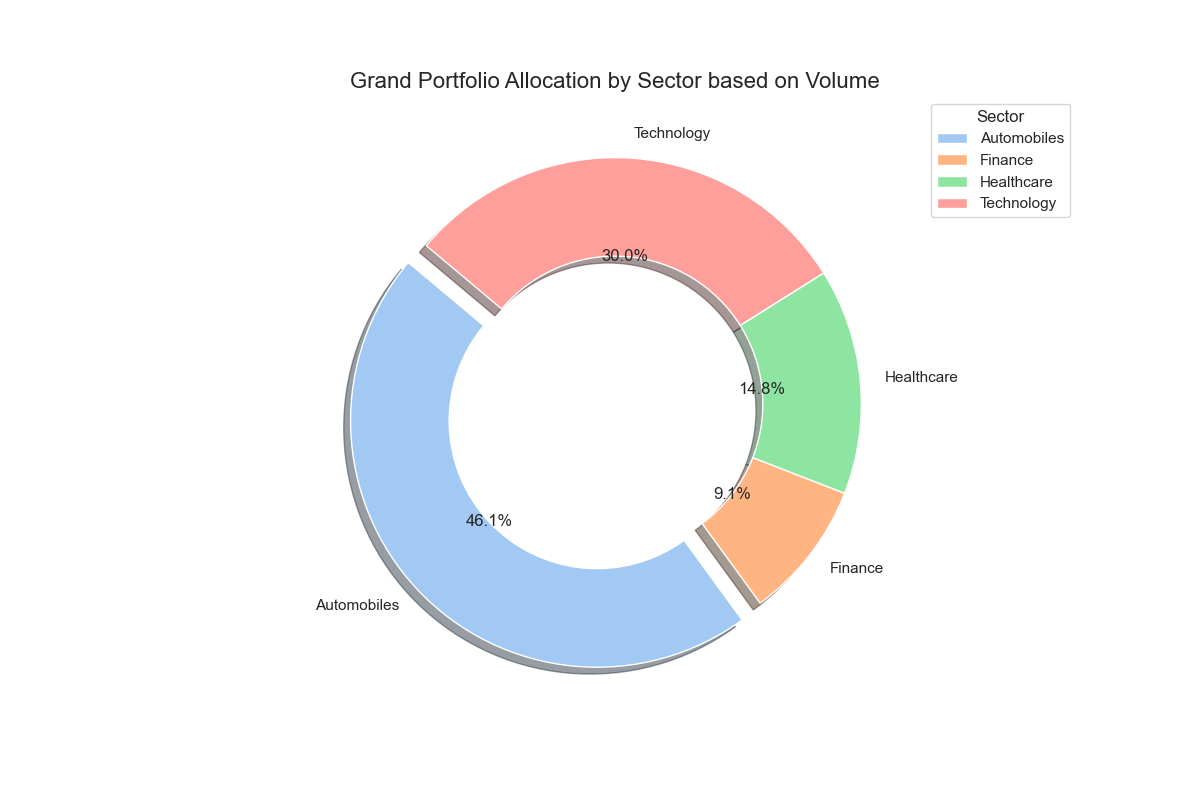
**Automobiles**

Ford, AMD, and Tesla dominated trading volumes, with Tesla experiencing a substantial increase in production and volumes in 2023.

**Healthcare**

Pfizer and Moderna emerged as the highest traded healthcare stocks, with Moderna experiencing a notable increase in production during the pandemic.

# Question7: What is the allocation of portfolio for each Sector? (Sakina)

**Answer:**

**This analysis provides a comprehensive view of how the portfolio's volume is distributed among key sectors, offering insights into the relative importance of each sector in the overall portfolio.**

The grand portfolio's volume is distributed across diverse sectors, each playing a crucial role in the overall composition.

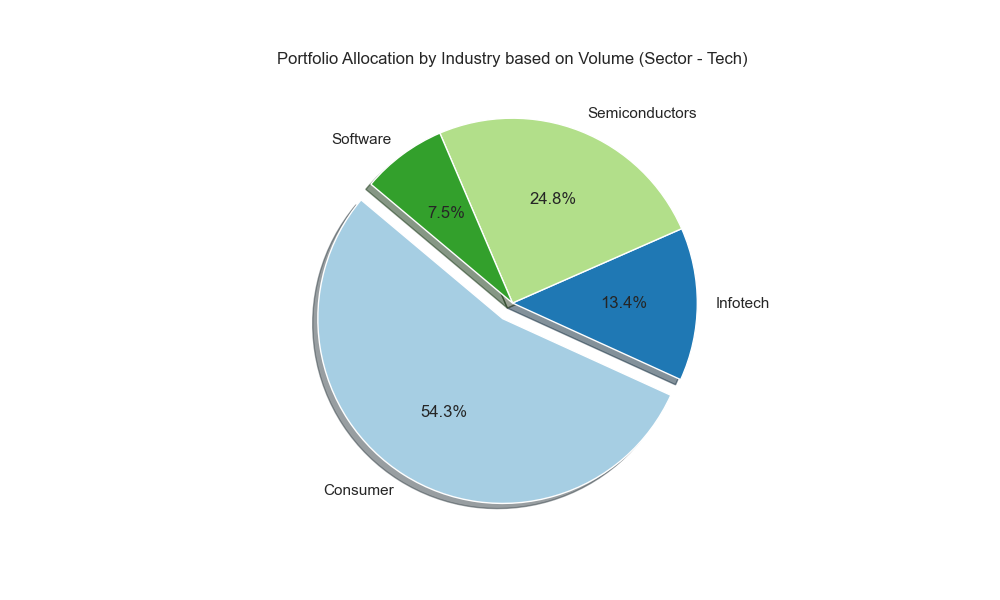
The dominance of the Automobiles sector suggests a significant focus on activities related to the automotive industry.

Technology follows closely, underlining the importance of technological advancements in the overall portfolio.

Healthcare and Finance, while substantial, have comparatively smaller shares, contributing to a well-rounded and diversified portfolio.

# Question8: What is the allocation of portfolio for each Industry? (Sakina)

**Answer:**

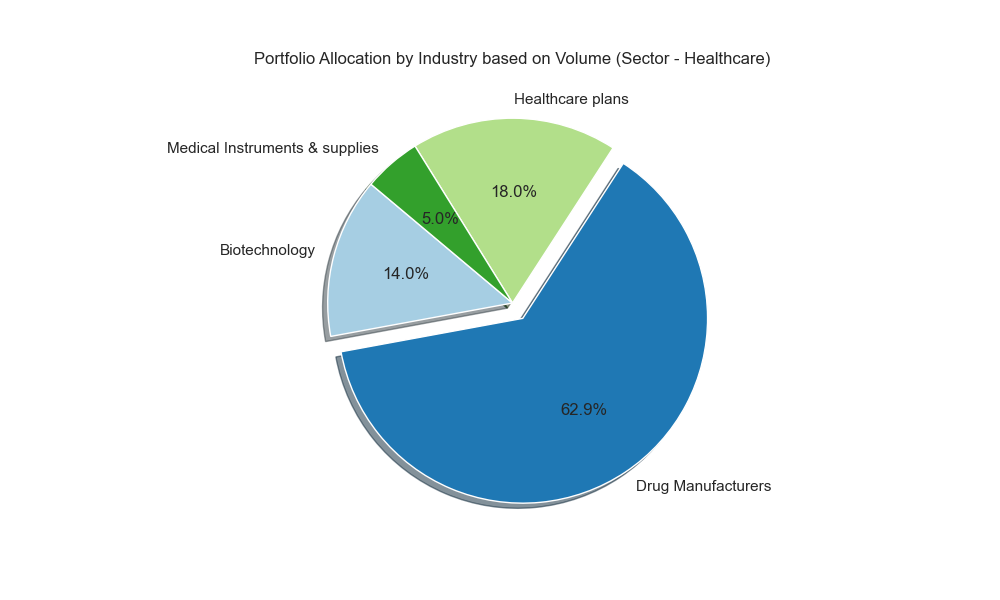
****

**This analysis provides insights into the distribution of volume across various technology industries, highlighting the sector's key areas of influence and the relative significance of each industry segment.**

The technology sector exhibits a diverse composition, with different industries contributing to the overall volume.

Semiconductors and Consumer technology emerge as major players, collectively representing a significant majority of the sector's volume.

Software and Infotech, while important, have a relatively smaller impact on the overall volume but contribute to the sector's diversity.

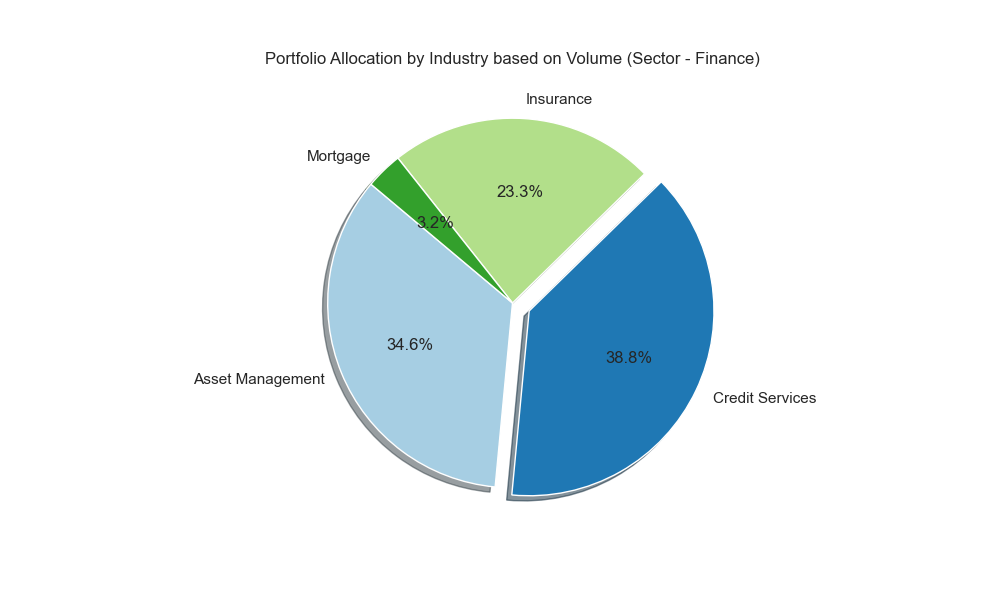


**This analysis provides a snapshot of how different segments within the healthcare sector contribute to the overall volume, shedding light on the sector's composition and key areas of activity.**

The healthcare sector is diverse, encompassing various sub-industries with distinct contributions to the total volume.

Drug manufacturing emerges as the leading contributor, underscoring the importance of pharmaceutical production in the healthcare landscape.

The presence of healthcare plans and biotech highlights the multifaceted nature of the sector, with a focus on insurance services and cutting-edge biotechnological advancements.

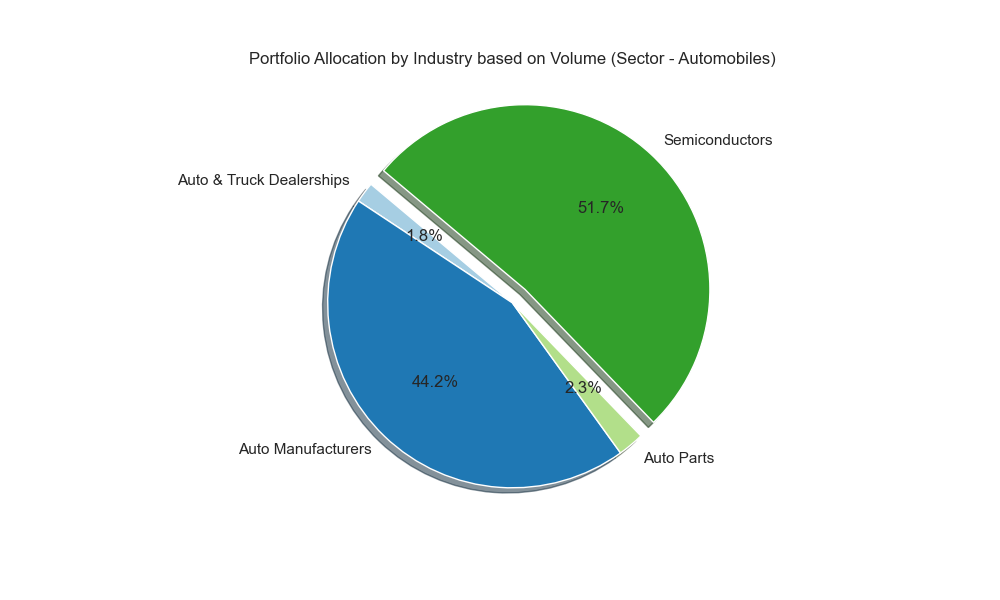


**This analysis provides insights into how different segments contribute to the overall volume in the finance sector, offering a snapshot of the sector's composition and key areas of influence.**

The finance sector is diverse, encompassing various segments with distinct roles in the overall volume composition.

Insurance, Asset Management, and Credit Services emerge as major contributors, collectively representing the majority of the sector's volume.

Mortgage services, while present, make up a smaller share, indicating that mortgage-related activities have a relatively smaller impact on the sector's overall volume.



**This analysis provides insights into how different segments contribute to the overall volume in the automobile sector, offering a snapshot of the sector's composition and key areas of influence.**

The automobile sector encompasses diverse segments, each playing a distinct role in the overall volume composition.

Semiconductors emerge as a crucial element, highlighting the growing reliance on electronic components in modern vehicles.

Auto manufacturers take a significant share, underlining the importance of vehicle production within the sector.

Auto & Truck Dealerships and Auto Parts, while contributing to the sector, represent smaller shares, emphasizing the dominance of manufacturing and semiconductor activities.

## Datasets used:

* [AlphaVantage package](https://www.alphavantage.co/documentation/)
* Yahoo Finance Module

**Limitations**

While our analysis provides valuable insights, it is essential to acknowledge certain limitations. The data used in the analysis represents a small fraction of the entire stock market, and should be acknowledged as a limitation, as this is not fully representative of the entire market.

Factors such as market sentiment, geopolitical events(i.e. The war in Ukraine, 2022), and unforeseen circumstances may also not be fully captured in our methodology.