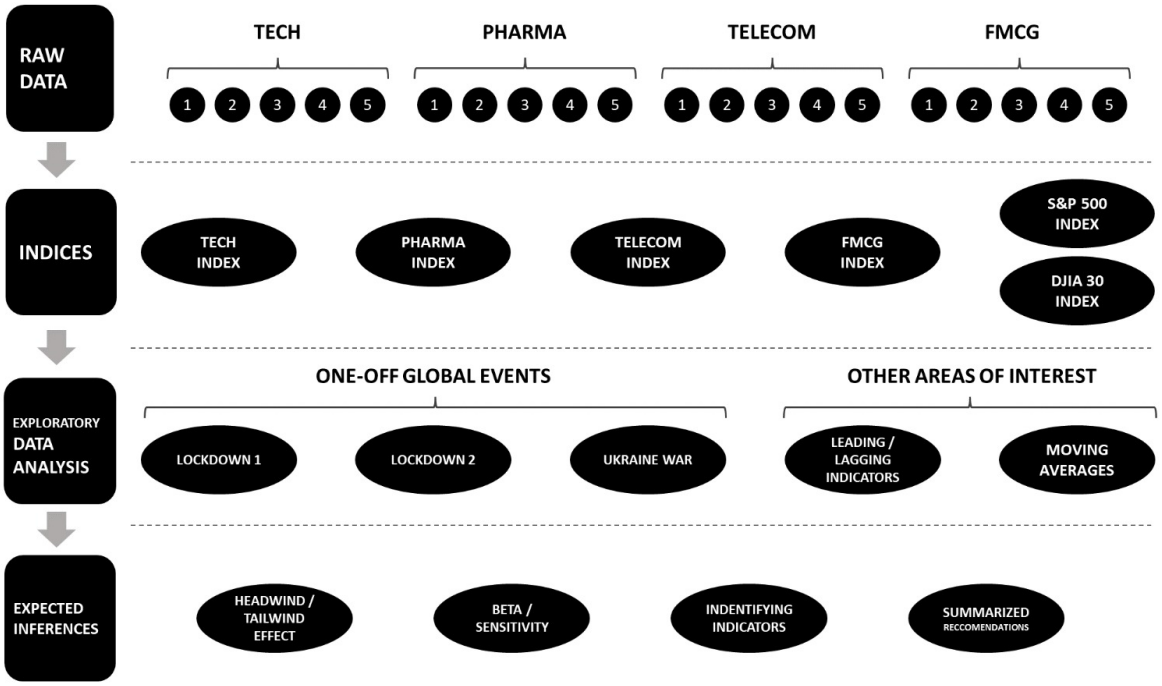
BUDT704 Project Proposal: Beyond the Headlines: Exploring Insights from Stocks

Group 19 (Stockbusters): Brigid Sax, Delvin Eluvathingal, Nan Liu, Navneeth Oruganti, Rakshanda Hedawoo, Saketh Gangavaram, Sumedh Tirodkar

1. Introduction and Summary:

The covid-19 pandemic is a one off global event. It has sent shockwaves to economies across the globe. In this project, we aim to perform an in-depth analysis of the impact of the COVID-19 pandemic on the performance of the top 5 stocks within four sectors that were severely affected by the crisis. The four sectors we are focusing on are Tech, Pharma, Telecom and FMCG (Fast moving consumer goods) and only on US stocks for this. This analysis is pivotal in understanding how economic crises impact stock performance and the broader financial markets. The study will be conducted in a Jupyter Notebook environment using statistical models and is expected to span a duration of 40-45 days.

The overview of the project:-



2. Questions of Interest:

Our research endeavors to address the following key questions:

- What were the variations in the price movements of the top 5 stocks in sectors most adversely affected by major world events like COVID-19 pandemic and the Ukraine-Russia war?
- Can we discern the factors acting as headwinds (Negative impact due to above events) and tailwinds (positive benefits due to above events) affecting these sectors during the crisis?
- What could be the possible reasons for this trend?
- How is the performance of these stocks related to the overall global economic indicators and trends?
- What is the level of sensitivity and exposure that these 4 sectors have to one off global events?
- To identify leading and lagging indicators driving these stock prices (Tentative)

3. Dataset Description:

We plan to utilize historical stock price data from reputable financial data sources, such as Yahoo Finance or Bloomberg, in a structured and time-series format, from over the last 4 years for this exercise. The dataset will cover a specific time frame, aligning with the onset and duration of the COVID-19 pandemic, allowing us to capture the relevant period.
The rationale behind selecting these datasets lies in their availability, accuracy, and reliability in tracking financial market movements during significant events.

We will be combining the stock price data from 20 different stocks from 4 sectors, as well as the S&P 500 and Dow Jones into one dataset. Overall, we will have about 1,460 rows, 110 columns and around 160,600 observations.

What variables are we looking for:

- Closing stock price
- Opening stock price
- Traded Volume
- Day High (Tentative)
- Day Low (Tentative)
- Date

What stocks and indices are we looking at:

- Economic Indicators
    - S&P 500 ()
    - Dow Jones (DJIA - 30)

- Tech Stocks:
    - Apple Inc. (AAPL)
    - Microsoft Corporation (MSFT)
    - Amazon.com Inc. (AMZN)
    - Alphabet Inc. (GOOGL) - Class A shares
    - Meta Platforms, Inc. (META) - formerly Facebook, Inc.
- Pharma:
    - Johnson & Johnson (JNJ)
    - Pfizer Inc. (PFE)
    - Merck & Co., Inc. (MRK)
    - Bristol-Myers Squibb Company (BMY)
    - Eli Lilly and Company (LLY)
- Telecom:
    - AT&T Inc. (T)
    - Verizon Communications Inc. (VZ)
    - T-Mobile US, Inc. (TMUS)
    - Comcast Corporation (CMCSA)
    - Charter Communications, Inc. (CHTR)
- FMCG (Fast-Moving Consumer Goods):
    - Procter & Gamble Co. (PG)
    - The Coca-Cola Company (KO)
    - PepsiCo, Inc. (PEP)
    - Walmart Inc. (WMT)
    - Colgate-Palmolive Company (CL)
- Leading lagging indicators
    - To be figured out and imported subject to results from exploratory analysis

4. Data Processing Tasks:

- Collecting and merging data

We'll be importing the required data from YahooFinance website into independent csv files.

- Data Cleaning

The initial data processing tasks will encompass data cleaning to rectify anomalies, missing values, and outliers. This ensures the integrity of our dataset.

- Data Transformation

Top 5 stocks within each sector are picked and a market cap weighted average value is calculated. This value is representative of how the sector is doing. This is a method that recognizes the influence of larger companies over smaller ones within an index.
The market cap-weighting method is chosen due to its reflection of the actual market dynamics where larger companies carry more influence over an index.

5. Data Analysis Explanation:

Our analytical framework comprises rigorous statistical methodologies, including time series analysis, correlation analysis, and possibly regression analysis.

We intend to juxtapose sector performance with global economic indicators, which may involve constructing an economic indicator index. This will allow us to assess the interactions and relationships between sector performance and the broader economy. We acknowledge the presence of potential assumptions and limitations in our analysis, including the assumption of efficient markets and the availability of reliable economic indicators.

6. Data Visualizations Covered in the Project:

Our project will produce various data visualizations, including but not limited to:

- **Line Charts:**
  - *Why*: Line charts are useful for displaying the historical stock price trends over time. They provide a clear view of how stock prices have changed during the COVID-19 pandemic.
  - Illustrating the performance of sector indices over the pandemic period
  - To visually represent the forces acting as headwinds and tailwinds in the sectors
- **Scatter Plots:**
  - *Why*: Scatter plots can be used to compare the returns of two stocks or an individual stock against a market index. They reveal how a stock's returns relate to another.
- **Heatmaps:**
  - *Why*: Heatmaps are effective for displaying correlations between stocks or sectors. They help identify which stocks or sectors are closely related and which are less correlated.
  - Correlation matrices to quantitatively depict the relationship between sector indices and economic indicators.
- **Bar Charts or Histograms:**
  - *Why*: Bar charts or histograms can be used to visualize the distribution of daily or weekly stock returns, allowing you to assess volatility and risk.
  - Use Cases: By creating a histogram of returns, you can see how often different levels of returns occur, helping to understand the risk profile of the stocks.

7. Expected Findings:

We expect to gain valuable insights from our research on how different sectors and individual stocks responded to the impacts of the pandemic and the war. Our analysis aims to reveal patterns that show which sectors and stocks demonstrated resilience in the face of these external shocks and which ones proved vulnerable. Additionally, we will identify potential indicators that could predict how these sectors would react, whether in advance or with a delay. We plan to validate these patterns across various timeframes, depending on the results of our initial exploratory analysis. Ultimately, our findings are intended to offer important insights into how external shocks affect financial markets. These insights will serve as a valuable resource for investors and policymakers, helping them make more informed decisions in response to such events.

FMCG: The majority of supermarkets and retail chains were closed during lockdown, which was anticipated to cause a dramatic decline in the share price.

PHARMA: By the time COVID was in place, many pharma companies were already working on vaccines. During the second lockdown, vaccine development had advanced, and so this sector will respond the quickest of all the sectors.

TECH: It took some time for them to comprehend that the price of digital goods would rise so they didn't hire immediately, but eventually the tech stocks did rise.

TELECOM: The lockdown caused these sectors to perform well once 5G was released to the public due to the increased demand for data and high speeds owing to the lockdown.

EDA to be done to infer more insights by digging deeper into our findings from the data.

We anticipate a potential significant decline in the FMCG (Fast-Moving Consumer Goods) stocks, while pharmaceutical stocks may have experienced an extraordinary surge. We intend to substantiate this hypothesis by analyzing the patterns that emerged during the lockdown period and exploring other prevalent trends both before and after the lockdown.

8. Project Timeline (Expected to take 52 days):

The project timeline is divided into well-defined phases:

- Data collection and preprocessing (6 days): Gathering and cleaning the stock price and economic data.
- Index calculation and dataset creation (7 days): Creating sector indices and constructing a combined dataset.
- Data analysis and visualization (27 days): Conducting statistical analysis and generating data visualizations.

- Report writing and Jupyter Notebook development (8 days): Compiling the results and code into a clear and well-documented format.
- Review and finalization (4 days): A final review, revision, and preparation for publication.

This comprehensive proposal underscores the technical aspects and methods that will be employed in your stock analysis project, showcasing the depth of your research plan and the clarity of your approach.

Task	**Task Lead**	**Due Date**
Decide on Project Topic	Team	10/10
Complete Overview of Project Topic	Saketh	10/12
Create Questions of Interest	Navneeth	10/12
Find and Describe Dataset	Brigid	10/12
Describe Data Processing Tasks	Nan	10/12
Describe Data Analysis Tasks	Rakshanda	10/12
Discuss Expected Findings	Delvin	10/12
Discuss Expected Inferences	Sumedh	10/12
Submit Project Proposal	Team	**10/13**
Data Collection and Preprocessing	Brigid	**10/19**
Data Gathering	Brigid	10/15
Figuring out the Libraries required	Sumedh	10/15
Data Cleaning	Nan	10/18
Data Merging	Nan	10/19
Data Transformation	Sumedh	**10/26**
Index calculation	Saketh	10/23
Dataset creation	Brigid	10/26
Data Analysis and Visualization	Saketh	**11/22**
Analysis 1 - Impact of pandemic on stock prices	Navneeth	10/31
Creating Visualization for analysis 1	Rakshanda/ Sumedh	11/01

Analysis 2 - Factors for Headwinds and Tailwinds	Delvin	11/07
Creating Visualization for analysis 2	Rakshanda/ Sumedh	11/08
Compiling Project Update	Rakshanda/Nan	**11/09**
Submit Project Update	Team	**11/10**
Analysis 3 - Detecting reasons for the identified trends	Navneeth	11/14
Creating Visualization for analysis 3	Rakshanda/ Sumedh	11/15
Analysis 4 - Sector analysis	Sumedh	11/18
Creating Visualization for analysis 4	Rakshanda/ Sumedh	11/19
Analysis 5 - Volatility analysis	Brigid	11/22
Creating Visualization for analysis 5	Rakshanda/ Sumedh	11/22
Report writing and Jupyter Notebook development	Navneeth/Delvin	**11/30**
Introductory Markdown cells - brief writeup on the project	Navneth	11/25
Displaying datasets and outputs	Delvin	11/27
Explaining the outputs	Saketh	11/28
Proper comments in the code	Delvin	11/29
Mentioning the inferences and conclusions	Saketh	11/30
Recommendations (MAYBE)	Saketh	11/30
Review and finalization	Saketh	**12/4**
Compose Presentation Slides	Rakshanda	11/29
Compose Project Report	Nan	12/1
Practice Presentation	Team	12/1
Submit Presentation	Team	**12/4**
Submit Report	Team	**12/4**