Introduction

Financial data, especially stock data, is crucial to investors and analysts because it provides valuable insights into an industry's performance and market trends. Effective data visualization can help these stakeholders regconize key patterns in stock performance, thus help them making more informed decisions and understand more about the market behaviour. For that reason, out team project attempt to effectively visualize financial data of mainly about big tech companies from the year 2010 to 2023. This project includes four main datasets, which we will briefly introduce:

1. Big Tech Stock Stock Price: is the data we got from the following link in Github

```
path1 <- "data/big_tech_stock_prices.csv"
dataset1 <- read.csv(path1)
head(dataset1, 5)
```

```
##   stock_symbol       date     open     high      low    close adj_close
## 1         AAPL 2010-01-04 7.622500 7.660714 7.585000 7.643214  6.515213
## 2         AAPL 2010-01-05 7.664286 7.699643 7.616071 7.656429  6.526476
## 3         AAPL 2010-01-06 7.656429 7.686786 7.526786 7.534643  6.422664
## 4         AAPL 2010-01-07 7.562500 7.571429 7.466071 7.520714  6.410790
## 5         AAPL 2010-01-08 7.510714 7.571429 7.466429 7.570714  6.453412
##      volume
## 1 493729600
## 2 601904800
## 3 552160000
## 4 477131200
## 5 447610800
```

2. COVID-19 Cases Datasets from Kaggle link

```
dataset2 <- read.csv("data/day_wise.csv")
head(dataset2, 5)
```

```
##         Date Confirmed Deaths Recovered Active New.cases New.deaths
## 1 2020-01-22       555     17        28    510         0          0
## 2 2020-01-23       654     18        30    606        99          1
## 3 2020-01-24       941     26        36    879       287          8
## 4 2020-01-25      1434     42        39   1353       493         16
## 5 2020-01-26      2118     56        52   2010       684         14
##   New.recovered Deaths...100.Cases Recovered...100.Cases Deaths...100.Recovered
## 1             0               3.06                  5.05                  60.71
## 2             2               2.75                  4.59                  60.00
## 3             6               2.76                  3.83                  72.22
## 4             3               2.93                  2.72                 107.69
## 5            13               2.64                  2.46                 107.69
##   No..of.countries
## 1                6
## 2                8
## 3                9
## 4               11
## 5               13
```

3. Walmart Stock Historical Data from Kaggle [link](link

```
dataset3 <- read.csv("data/WMT.csv")
head(dataset3, 5)
```

```
##         Date  Open  High   Low Close Adj.Close   Volume
## 1 2011-11-16 57.10 57.42 56.64 56.68  44.89946 11780800
## 2 2011-11-17 56.54 57.19 56.26 56.73  44.93906 10223800
## 3 2011-11-18 57.03 57.36 56.61 57.23  45.33513  8982300
## 4 2011-11-21 56.93 57.29 56.38 56.66  44.88361  9932200
## 5 2011-11-22 56.56 57.13 56.50 56.85  45.03411  7497300
```

4. Pfizer Stock Historical Prices from Kaggle link

```
dataset4 <- read.csv("data/pfizer.csv")
head(dataset4, 5)
```

```
##        Date     Open     High      Low    Close Adj.Close   Volume
## 1 1/22/2020 38.25427 38.33966 37.92220 38.13093  32.17976 18097812
## 2 1/23/2020 38.13093 38.73814 38.07400 38.62429  32.59610 27148510
## 3 1/24/2020 38.84251 38.87097 37.60911 37.77988  31.88349 34143698
## 4 1/27/2020 37.39089 38.35863 37.23909 38.10247  32.15573 31964026
## 5 1/28/2020 37.30550 37.46679 36.00569 36.18596  30.53834 70202408
```

Both dataset 1, 3, and 4 are financial data about stock prices such as open price, close price, high price, low price, adjusted close price and volume of stock daily (with data column). As dataset 1 is just about big tech stock companies so we include dataset 3 and 4 which are about stock price of companies in different sectors for more wholistic view. We also integrated these financial datasets with dataset three about the COVID-19 cases to see how the pandemic have affect the companies. For more detailed view about the datasets, please view the head function of each dataset.