

Project1

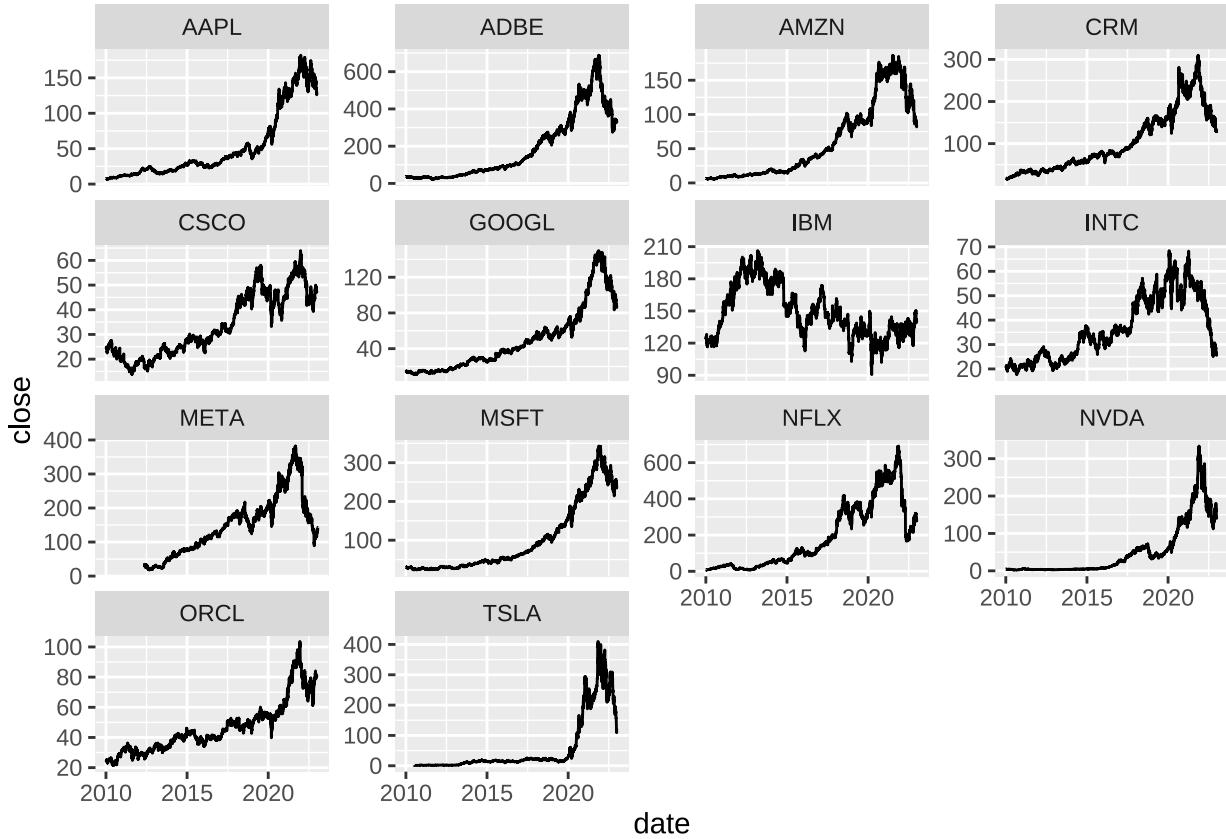
2024-04-02

2. How did the COVID-19 Pandemic affect the stock prices of Big Tech Companies?

2.1. Introduction:

- We would like to answer this question because we observed a significant decrease in stock prices of major tech groups from late 2019 to mid-2020 which **overlapped with the dramatic expansion of COVID-19**. We'll check how stock prices changed during this time and see which companies were hit the hardest. By detecting less affected companies, we can focus on exploring their potential strategies dealt with the pandemic which saves time from general research. Therefore, we might be able to learn how to handle similar situations better in the future. ### Data preparation:

```
ggplot(df,aes(y = close , x = date, group = 1)) +  
  geom_line () +  
  scale_x_date(labels = date_format("%Y"))+  
  facet_wrap(~ stock_symbol, scales = "free_y")
```



- To answer that question, from the original dataset, we need the **close price**, which can be considered the primary price of the stock, from late 2019 to mid-2020.
- Moreover, to visualize the correlation between covid 19 spread and stock price, we use an external dataset, which contains **daily global statistics about covid 19 cases** from 22 Jan 2020 to 27 Jul 2020, published on Kaggle.
- Finally, to compare technology companies with other sectors, for example, groceries or medical products, we use 2 more datasets about the **historical stock price of Walmart and Pfizer** in the corresponding period.

2.2 Approaches

Plot choice explanation:

- **Firstly**, because COVID-19 affecting tech companies' stock prices is just our assumption, we have to qualify whether COVID-19 actually had that influence by plotting out stock prices and the global increase percentage of COVID-19 cases in the same plot. We use **bar plots**, which is good for describing the distribution of low-dimension data, to show the changes in stock price over time. Moreover, by visualizing with bar plot, we can easily see the change in stock price, compare it to past or future prices, and classify them at the same time (We use different colors for *Increase, Decrease or No change* samples compared to the base price).
- The growth speed of COVID-19 cases will be visualized with a **line plot** because we want to reduce the plot's density (lines save more space than bars) and emphasize the trend, which is the advantage of line charts, of COVID-19 spread.

- **Secondly**, after we verify the assumption and visualize some patterns, we want to clearly visualize which company suffers most from COVID-19. In other words, we want to visualize their loss ($loss = (expected\ price - actual\ price) * volume$) and rank them in descending order. We chose a bar plot for this purpose because this type of chart is optimized for rank illustration. We can see what companies did the best policies and their differences with tech giants who performed worse by comparing bars.
- **Finally**, different from the 2 tasks mentioned above, our final sub-question requires more data points to intuitively show the trend in the stock price of each tech company and 2 external representatives from the groceries/ medical fields. In this situation, a line chart would be an optimized consideration to reduce the density caused by a huge number of data points. Moreover, line charts can also show the trend over time better compared to other types of charts, so that we can clearly visualize changes in stock prices every single day. Besides that, because we want to see the date when sectors (tech, medical, groceries) swap their position in stock price rankings as well, we can easily point out those special days by intersections among lines.

Data pre-processing:

- Because we only crawled the data set of the most dramatic period of COVID-19 (1/22/2020 - 7/27/2020), we need to create subset of stock prices in that time.
- As I mentioned, bar plots are great to illustrate the distribution, however, it increase the density of the plot, makes it harder to observe differences in bars in the case where **there are too many bars**. Therefore, with the range of 6 months (nearly 200 days), I recalculate and **only leverage the average stock price for each week** to enhance the clarity of the plot while maintain the general distribution.
- Besides that, because **number of cases exponentially increased**, initial data points would be meaningless if we directly plot the raw data. Therefore, we try to consider the growth speed of the number of global COVID-19 cases, and represent this metric with a new variable: **percentage** ($percentage = number\ of\ new\ case / total\ of\ case * 100$).
- If there had been no COVID-19, the stock price of companies would have increased or at least been unchanged, therefore, the decrease in stock price created a loss value for the companies. The more stock transactions, the higher loss they had to suffer.
- Therefore, we come up with a metrics to evaluate their losses: **$loss = (expected\ price - actual\ price) * volume$**
- Because each company have different stock price, and we want to focus on the growth of them in the pandemic period, it would be **unfair if we directly plot the raw price**. Therefore, **percentage compared to the initial price would be our choice**.

Other techniques:

- Colors are used to show the comparison of price in that week with the based price (in the week before COVID-19), so that we can see **how hard or how long COVID-19 badly affected the stock prices of each companies**.
- We added some **annotations for maxima points** to see whether the most dramatic COVID-19 expansion match with the lowest stock price in the time series.

Theme definition:

```

theme_covid <- function(){
  theme_minimal(base_family = "Montserrat") +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold", color = "#c83538"),
    plot.subtitle = element_text(hjust = 0.5, face = "bold"), # This will center the subtitle
    legend.position = "bottom",
    axis.text.x = element_text(angle = 45,color = "#35426e", face = "bold", size = 12 ),
    axis.text.y = element_text(color = "#35426e", face = "bold", size = 12),
    panel.grid.major.x = element_blank(), # Remove major x grid lines
    panel.grid.minor.x = element_blank(),
    panel.grid.major.y = element_line(color = "grey", size = 0.2),
    panel.grid.minor.y = element_blank())
}

theme_covid2 <- function(){
  theme_minimal(base_family = "Montserrat") +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold", color = "#c83538"),
    plot.subtitle = element_text(hjust = 0.5, face = "bold"), # This will center the subtitle
    axis.text.x = element_text(angle = 45,color = "#35426e", face = "bold", size = 12 ),
    axis.text.y = element_text(color = "#35426e", face = "bold", size = 12),
    panel.grid.major.x = element_blank(), # Remove major x grid lines
    panel.grid.minor.x = element_blank(),
    panel.grid.major.y = element_line(color = "grey", size = 0.2),
    panel.grid.minor.y = element_blank())
}

```

2.3 Analysis

2.3.1. Verify the assumption “Did COVID-19 actually affect Tech Giants’ stock prices?”

```

create_covid_plot("AAPL")

## `summarise()` has grouped output by 'Year'. You can override using the
## `.` argument.

## Warning: The `trans` argument of `sec_axis()` is deprecated as of ggplot2 3.5.0.
## i Please use the `transform` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.

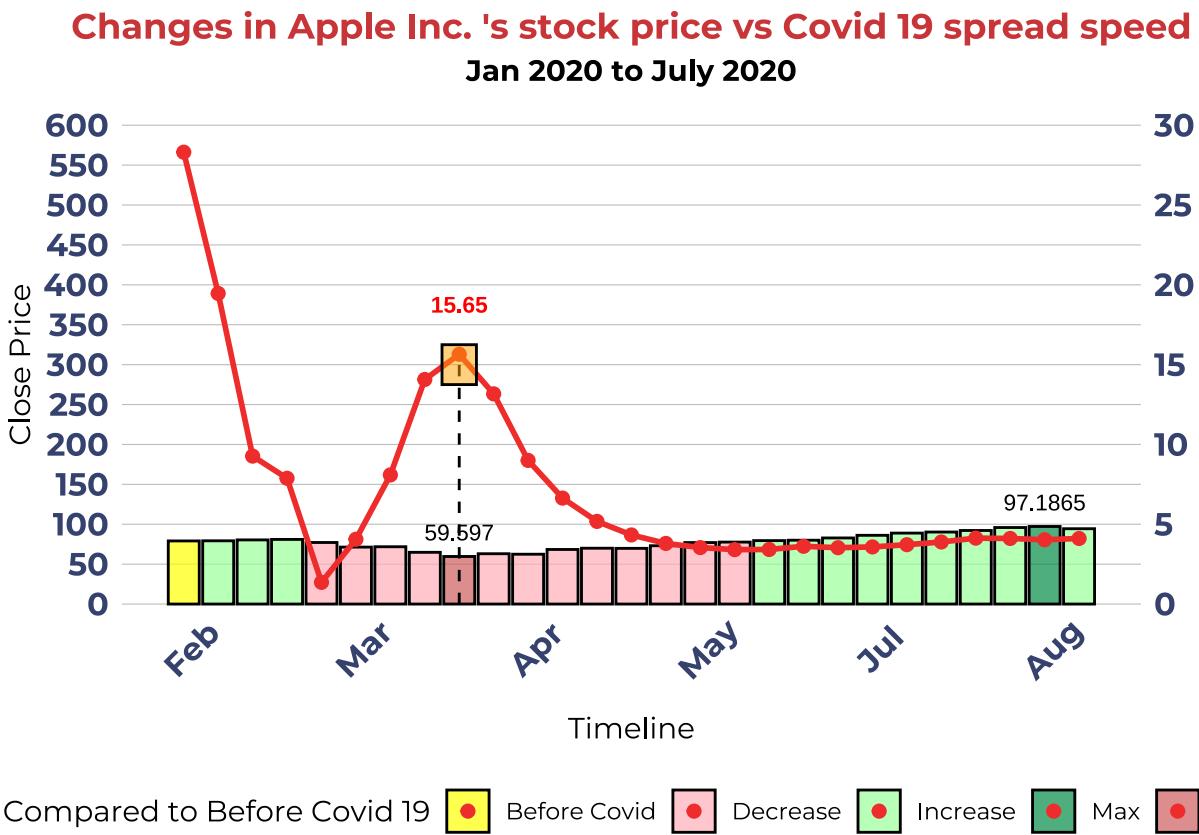
```

```

## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## Warning: The `size` argument of `element_line()` is deprecated as of ggplot2 3.4.0.
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

```



```
create_covid_plot("AMZN")
```

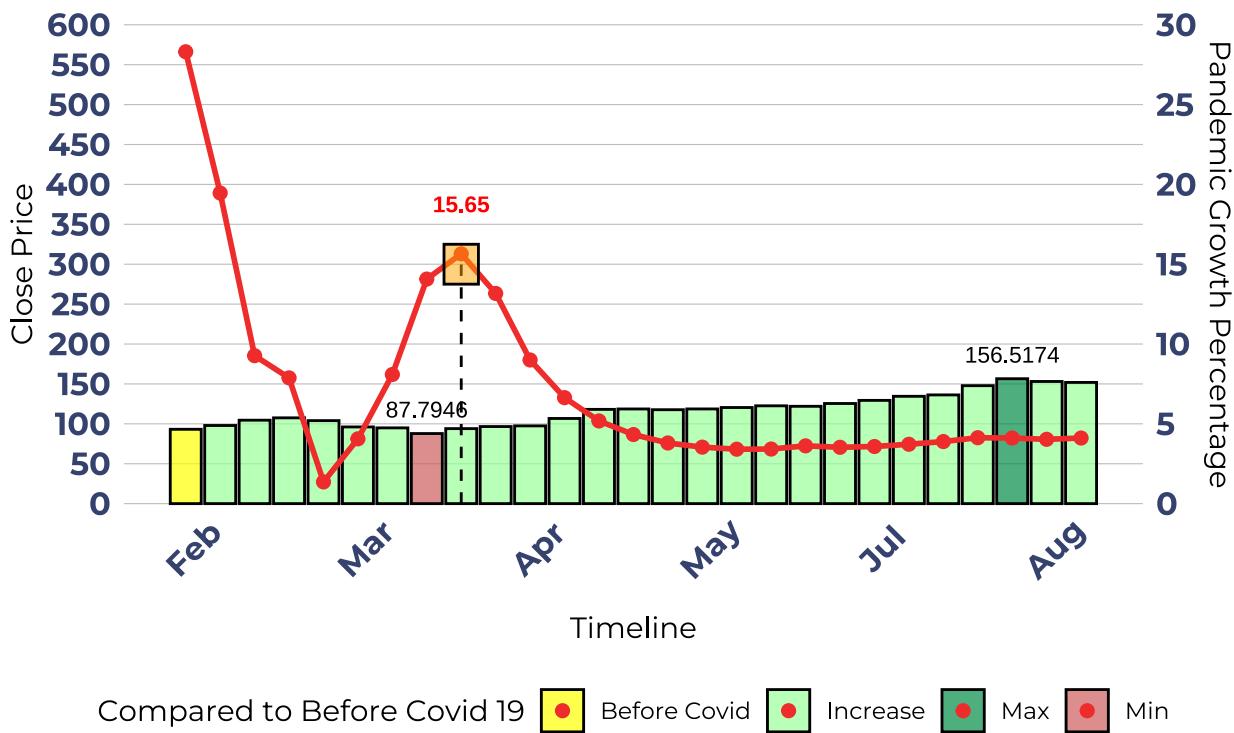
```

## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.

```

Changes in Amazon.com, Inc. 's stock price vs Covid 19 spread speed

Jan 2020 to July 2020

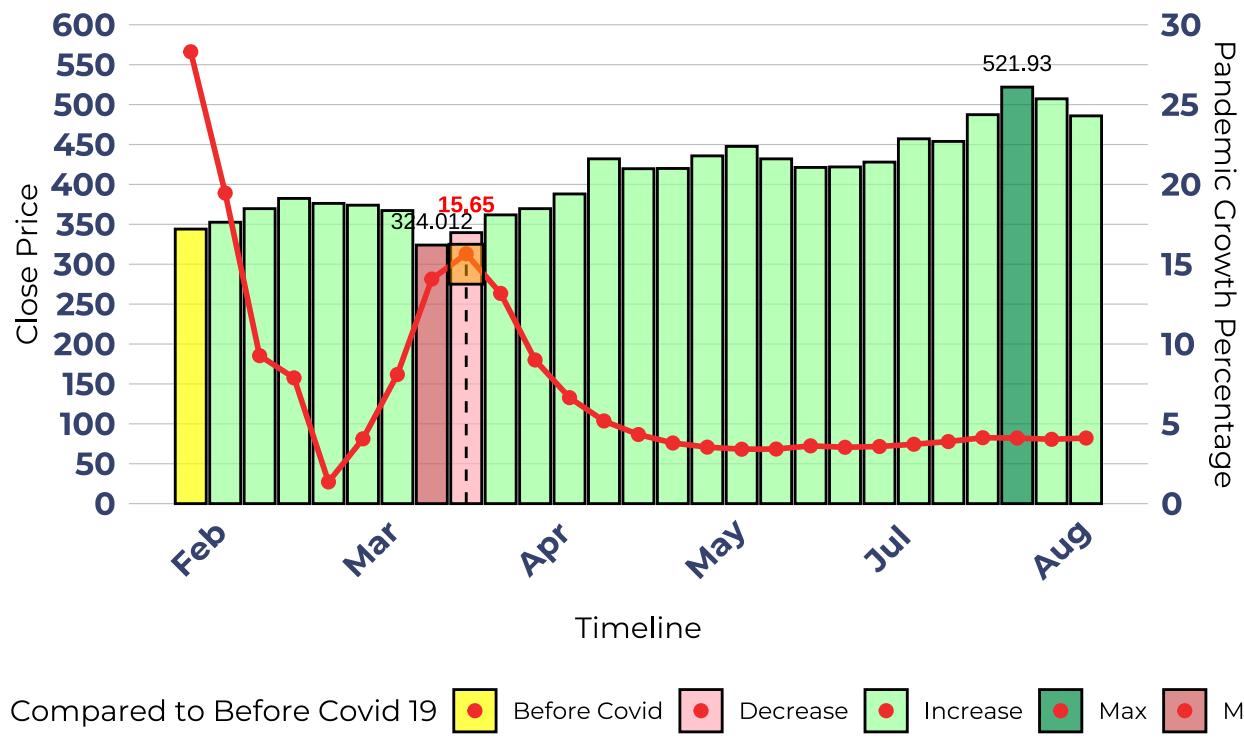


```
create_covid_plot("NFLX")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Netflix, Inc. 's stock price vs Covid 19 spread speed

Jan 2020 to July 2020

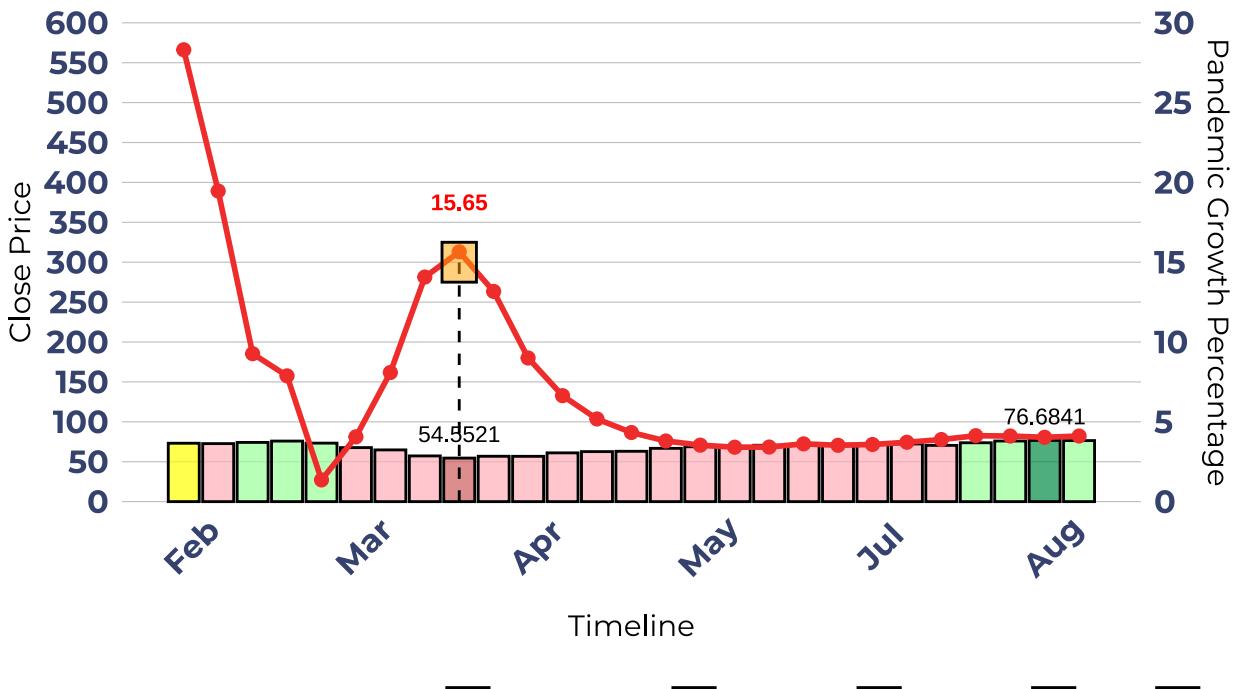


```
create_covid_plot("GOOGL")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Alphabet Inc. 's stock price vs Covid 19 spread speed

Jan 2020 to July 2020

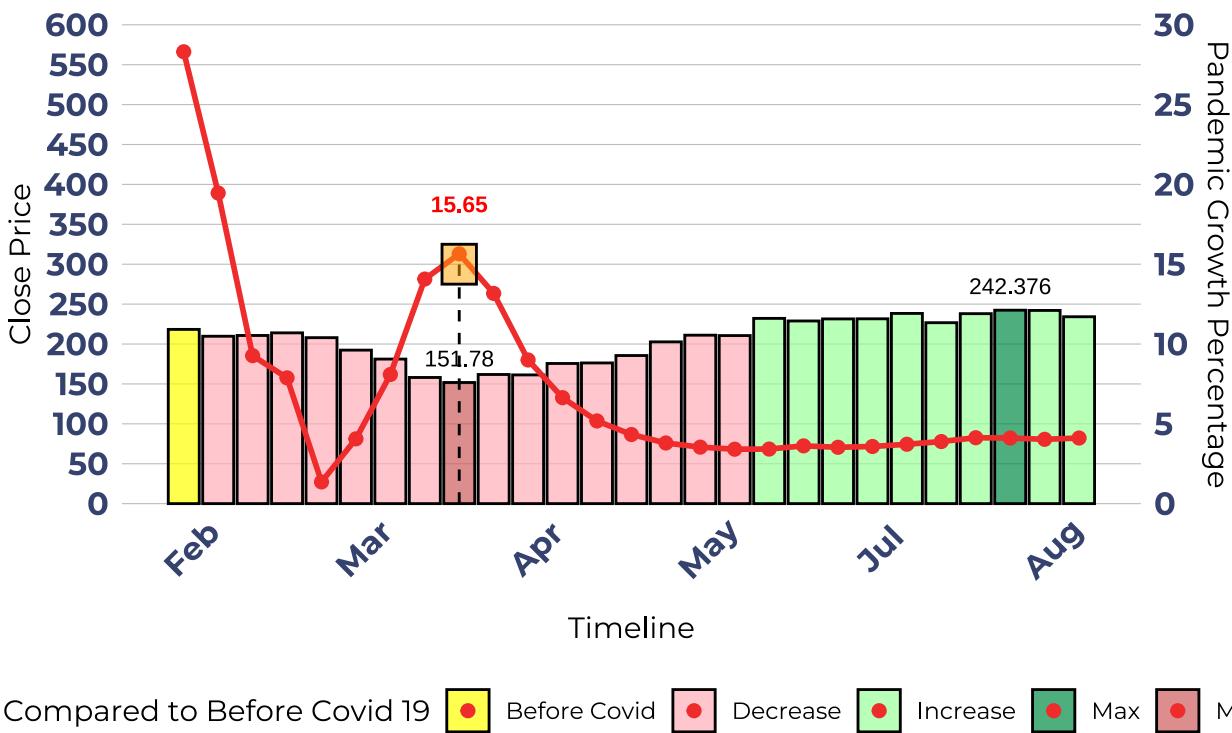


```
create_covid_plot("META")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Meta Platforms, Inc. 's stock price vs Covid 19 spread specific

Jan 2020 to July 2020

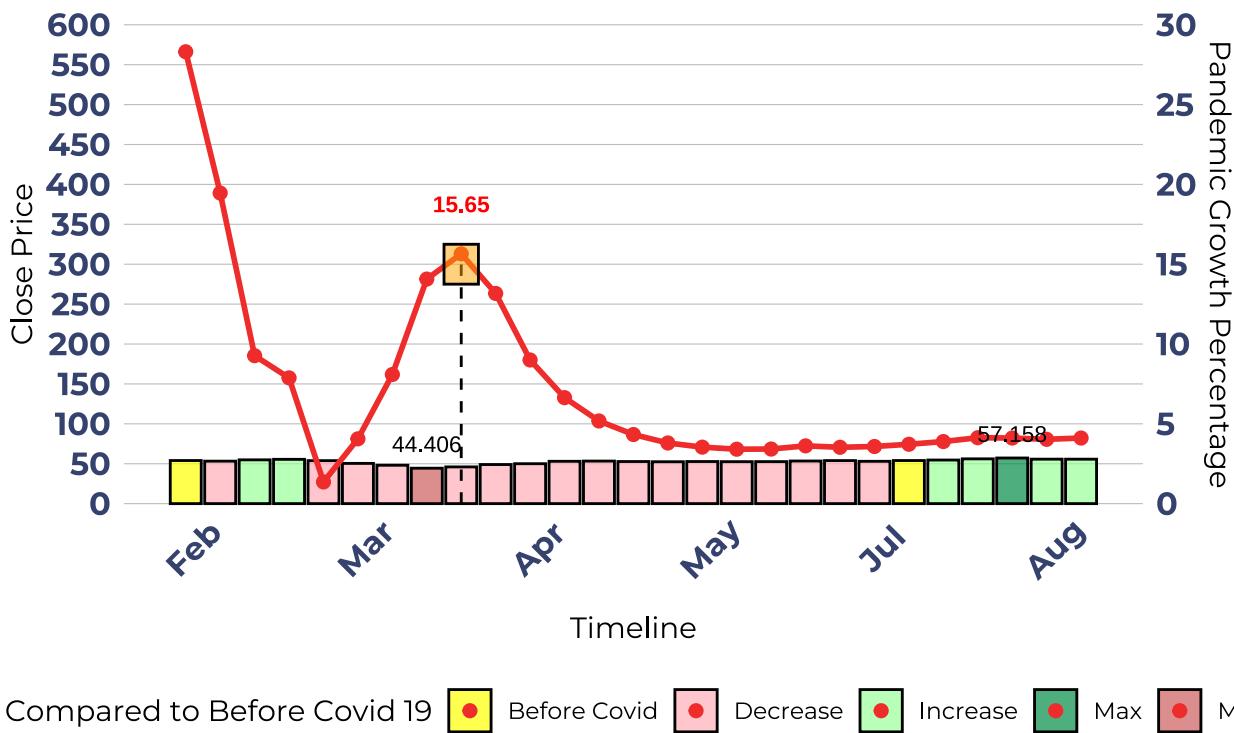


```
create_covid_plot("ORCL")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Oracle Corporation's stock price vs Covid 19 spread specific

Jan 2020 to July 2020

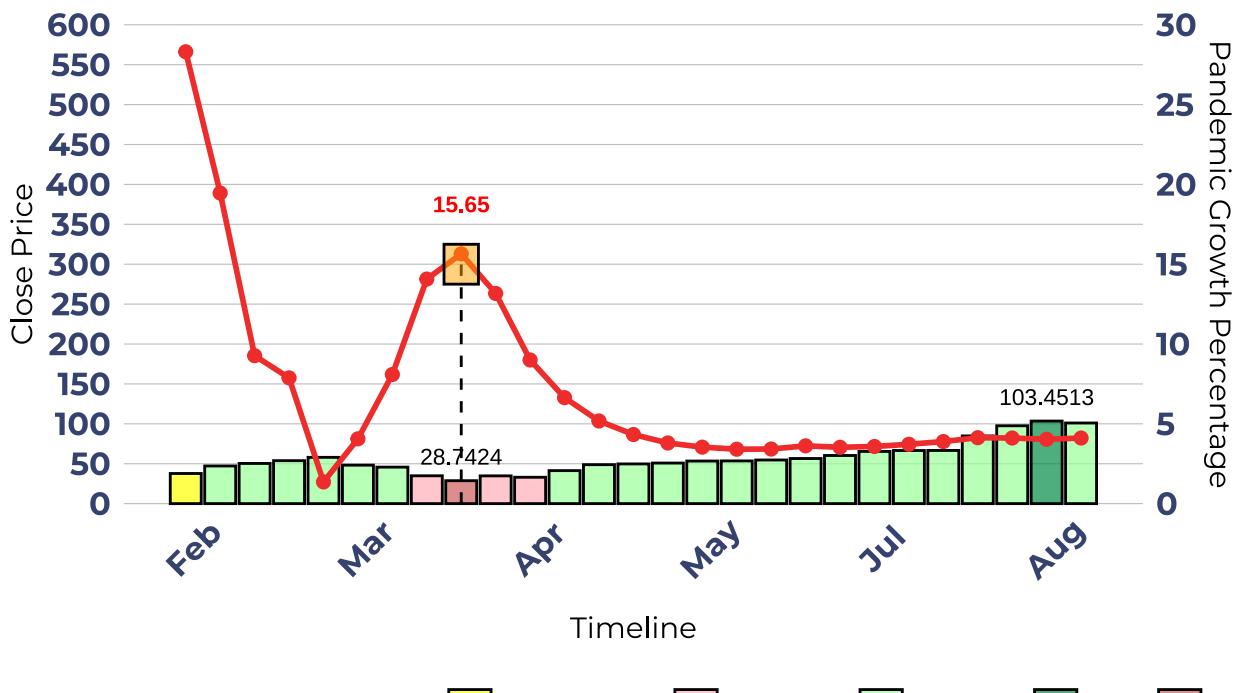


```
create_covid_plot("TSLA")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Tesla, Inc. 's stock price vs Covid 19 spread speed

Jan 2020 to July 2020



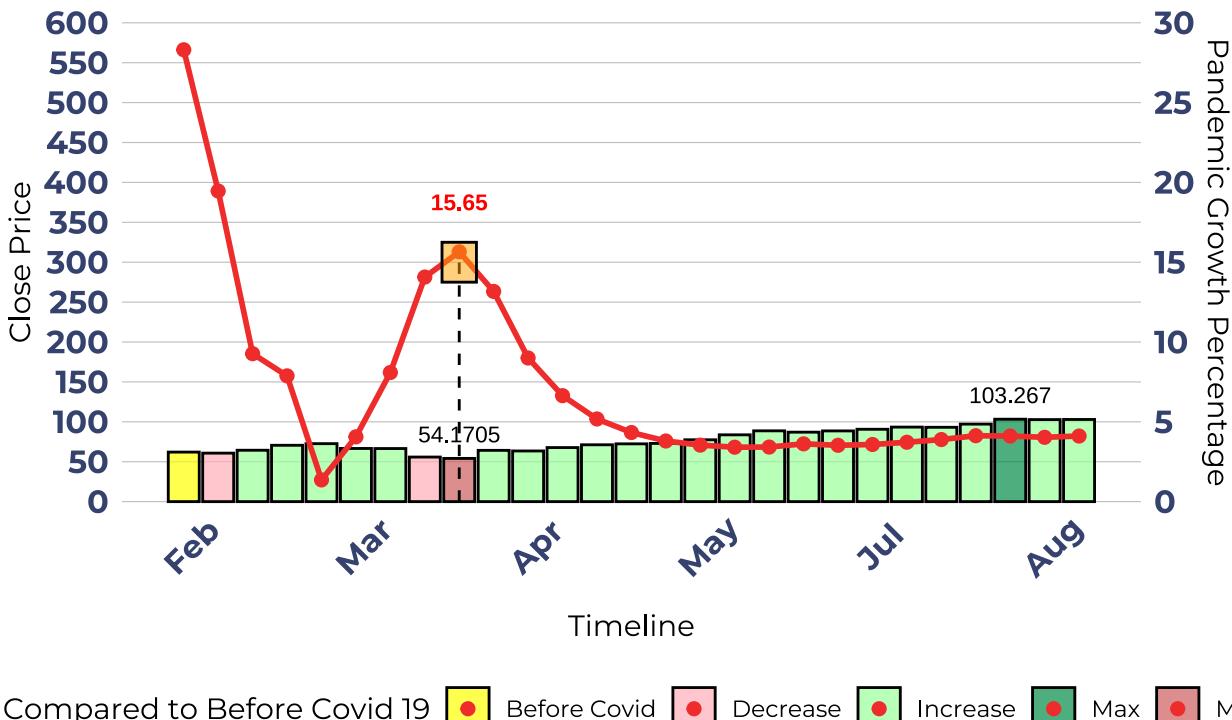
Compared to Before Covid 19 ● Before Covid ● Decrease ● Increase ● Max ● Min

```
create_covid_plot("NVDA")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in NVIDIA Corporation's stock price vs Covid 19 spread specific

Jan 2020 to July 2020

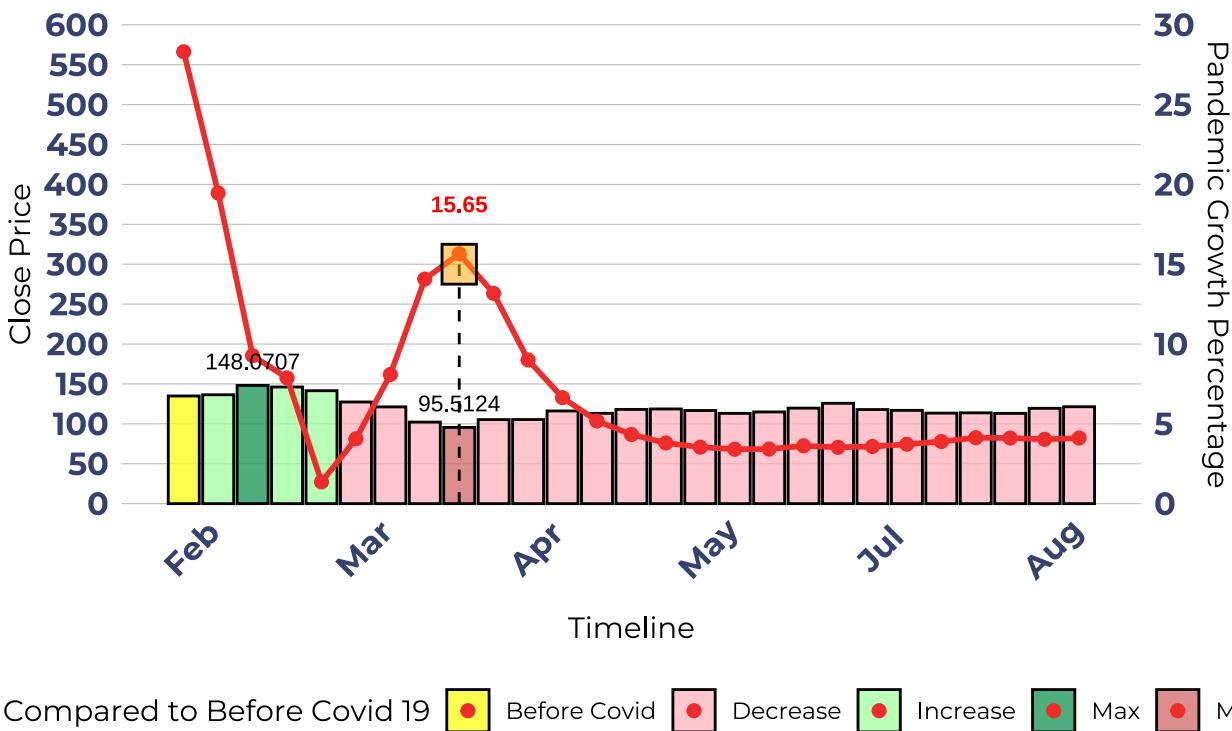


```
create_covid_plot("IBM")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

International Business Machines Corporation's stock price vs Covid 19

Jan 2020 to July 2020

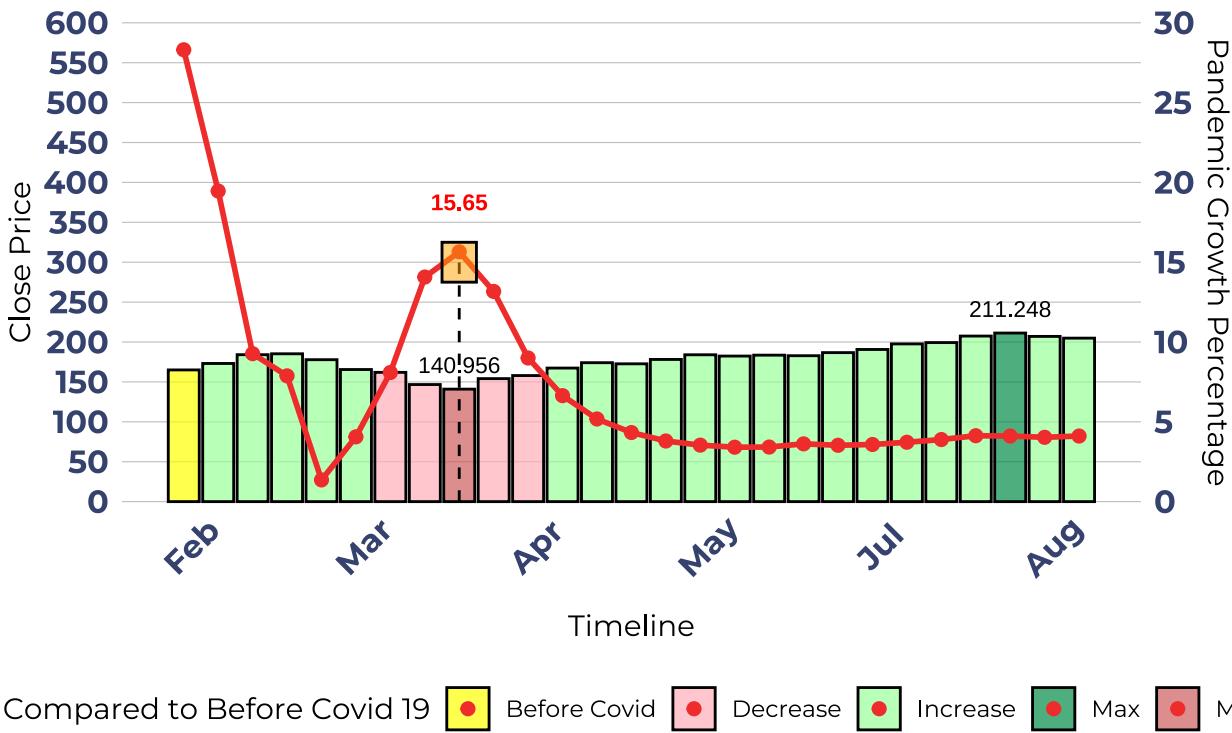


```
create_covid_plot("MSFT")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Microsoft Corporation's stock price vs Covid 19 spread specific

Jan 2020 to July 2020



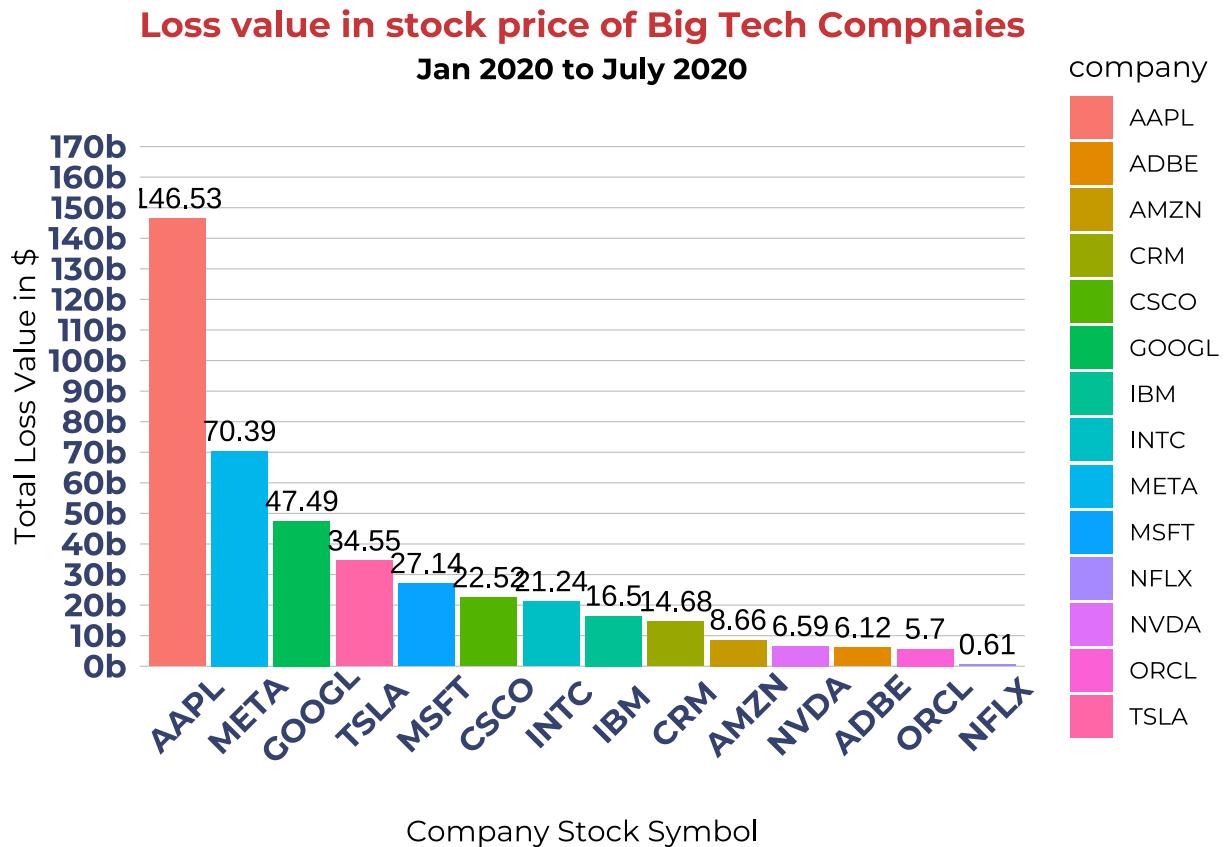
2.3.2. Which company is most affected (metric: expected loss value) ?

```
print(sorted_df)
```

```
##      company  loss_value    volume
## 2      AAPL 146532797810 121214000
## 10     META  70387759593 13163100
## 7      GOOGL 47491417059 27124000
## 15     TSLA  34548093748 240730500
## 11     MSFT 27139145392 30160900
## 6      CSC0 22515816955 15576000
## 9      INTC 21240160026 107526500
## 8      IBM 16499190016 3905136
## 5      CRM 14677418881 3015700
## 4      AMZN 8662239393 83410000
## 13     NVDA 6590426189 29213200
## 3      ADBE 6116644253 1622000
## 14     ORCL 5696346615 10319600
## 12     NFLX 611774097 7863100
## 1      test        0       0
```

```
for (symbol in loss_company$stock_symbol) {
  loss_company <- calculate_loss(loss_company, symbol)
}
```

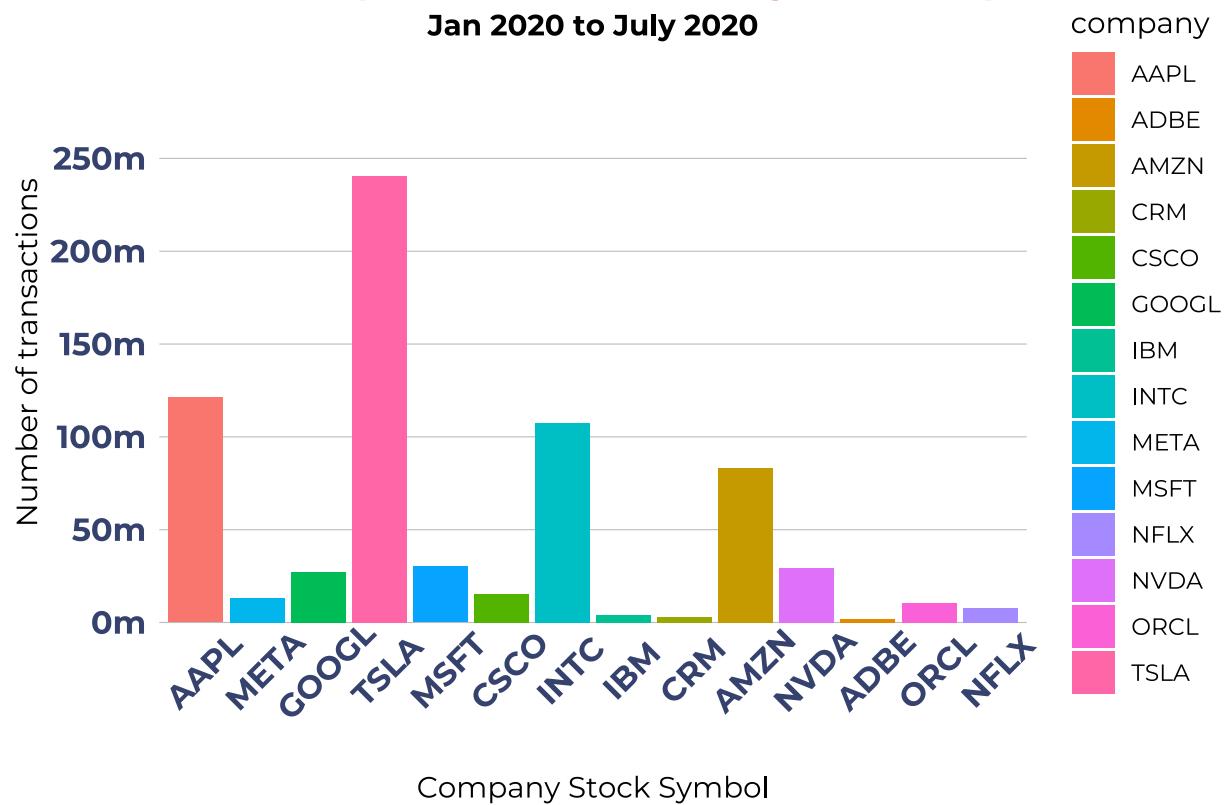
```
loss_plot+theme_covid2()
```



```
volume_plot +theme_covid2()
```

Volume of low price transaction in Big Tech Companies

Jan 2020 to July 2020



2.3.3. Compare tech giants stock price to giants in other fields (Walmart: sale, Pfizer: medical)

```
compare_to_other("AAPL")
```

Comparision of stock price between Apple Inc. and other fields

Jan 2020 to July 2020



```
compare_to_other("AMZN")
```

Comparision of stock price between Amazon.com, Inc. and other f Jan 2020 to July 2020



```
compare_to_other("NFLX")
```

Comparision of stock price between Netflix, Inc. and other field

Jan 2020 to July 2020



```
compare_to_other("GOOGL")
```

Comparision of stock price between Alphabet Inc. and other field

Jan 2020 to July 2020



```
compare_to_other("META")
```

Comparision of stock price between Meta Platforms, Inc. and other

Jan 2020 to July 2020



```
compare_to_other("ORCL")
```

Comparision of stock price between Oracle Corporation and other firms

Jan 2020 to July 2020



```
compare_to_other("NVDA")
```

Comparision of stock price between NVIDIA Corporation and other

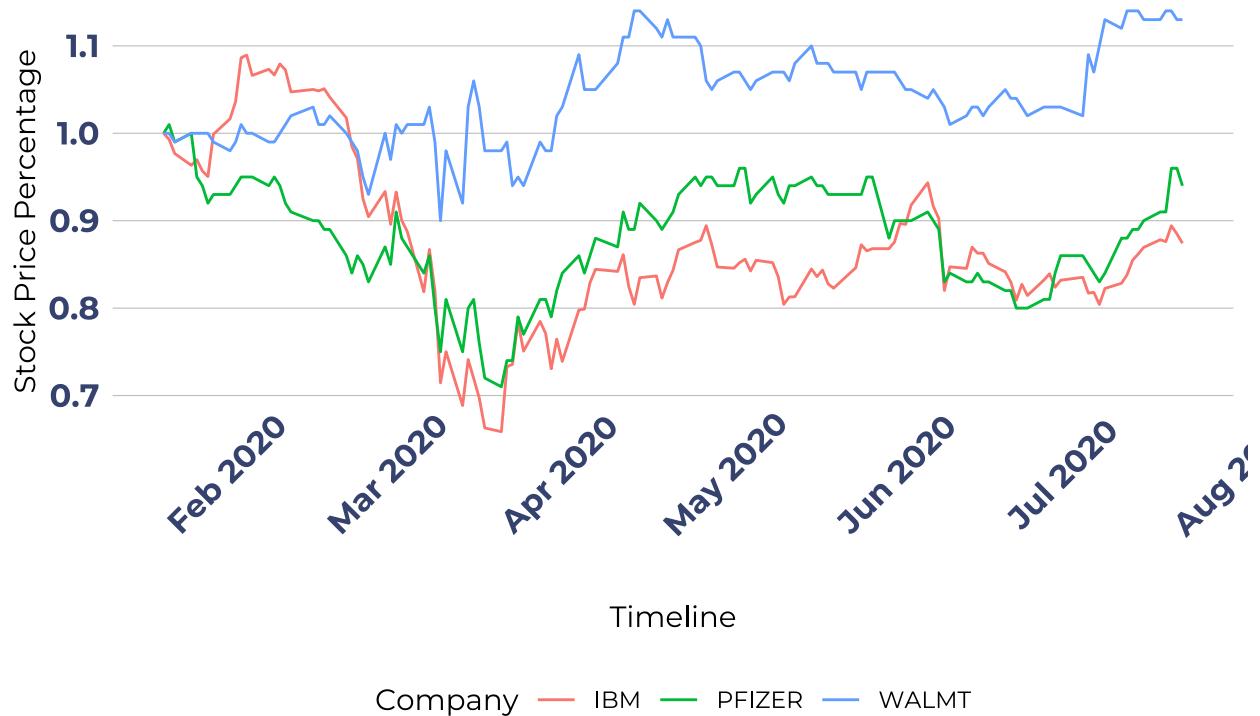
Jan 2020 to July 2020



```
compare_to_other("IBM")
```

of stock price between International Business Machines Corporation

Jan 2020 to July 2020



```
compare_to_other("MSFT")
```

Comparision of stock price between Microsoft Corporation and other

Jan 2020 to July 2020



We want to visualize Average Stock Prices of every tech companies, therefore we can observes the influence of COVID-19 on different sectors.

```
sec_plot <- ggplot(combine_df, aes(x = date, y = percentage, color = stock_symbol)) +  
  geom_line() +  
  scale_x_date(date_breaks = "1 month", date_labels = "%b %Y") +  
  labs(x = "Timeline", y = "Stock Price Percentage", color = "Company", title = paste("Comparision be  
  subtitle = "Jan 2020 to July 2020")  
sec_plot + theme_covid()
```

Comparision between Tech Companies and others in stock price

Jan 2020 to July 2020



2.4 Discussion

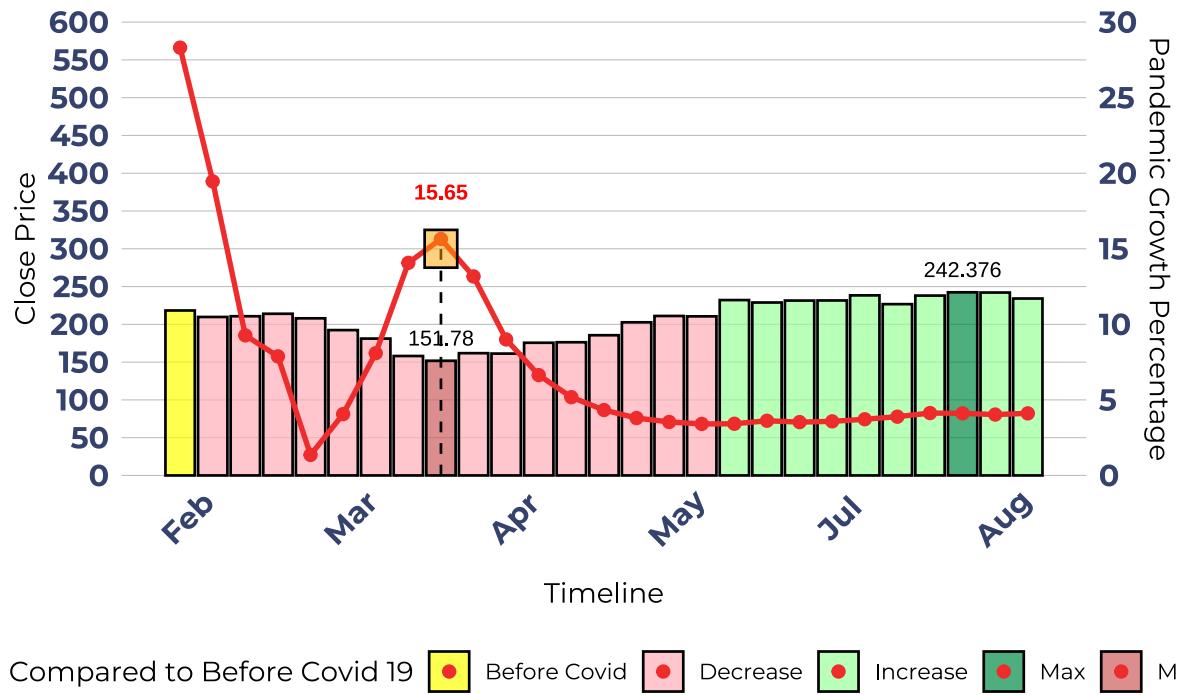
- From the set of plots that illustrates the change in stock price and COVID-19 spread speed over time, we can see that the stock price seems to be in inverse ratio to the growth of COVID-19 which the faster number of COVID-19 cases increases, the faster stock prices decrease.
- Especially, as we can see from the vertical lines starting from the peak of COVID-19 growing speed, this point corresponds to the lowest stock price week in most companies. On the other hand, the highest prices appeared when COVID-19 spread more slowly in late July.
- This phenomenon is easily explained by the positive correlation between the pandemic and the global economic crisis when everyone tends to store food, medicines, etc. rather than investing in technology companies. They were willing to sell stocks in lower prices.
- Moreover, based on the number of red and green bars for each company, we observed that entertainment (Netflix), young (Tesla), or online shopping (Amazon) companies suffer less damage than others. This could be thanks to adaptive and flexible policies from those companies or the high demand for online activities from customers in lockdown situations. Besides that, those bars also show that there are some companies (NVIDIA, Tesla, Netflix, Microsoft, Amazon) that recover sooner and better compared to the rest, this suggests that we should research those companies' strategies dealing with COVID-19.

```
create_covid_plot("META")
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `groups` argument.
```

Changes in Meta Platforms, Inc.'s stock price vs Covid 19 spread speed

Jan 2020 to July 2020



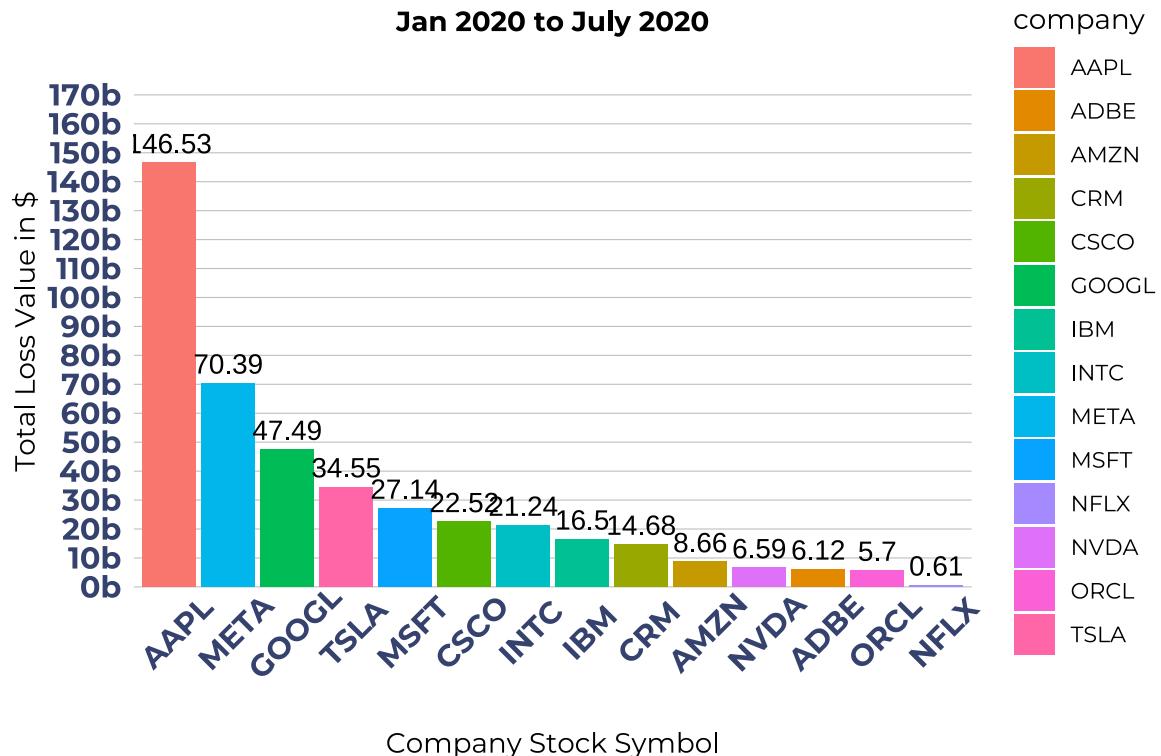
Compared to Before Covid 19 [Yellow Square] Before Covid [Red Square] Decrease [Green Square] Increase [Dark Green Square] Max [Red Circle] Min

- From the second plot, the bar chart clearly shows that Netflix (0.61 billion \$) is the most successful company in the pandemic situation while Apple lost the most money (146.53 billion \$). Looking back to the previous visualization, we can see that the decrease in the price of Apple's stock is not significant, however, the thing that made them lose much money is the volume of transactions. Although they lost only tens of bucks for each sell/buy transaction, their huge number of "low-price" transactions (~121 mil, less than only Tesla) caused their significant loss.

```
loss_plot+theme_covid2()
```

Loss value in stock price of Big Tech Companies

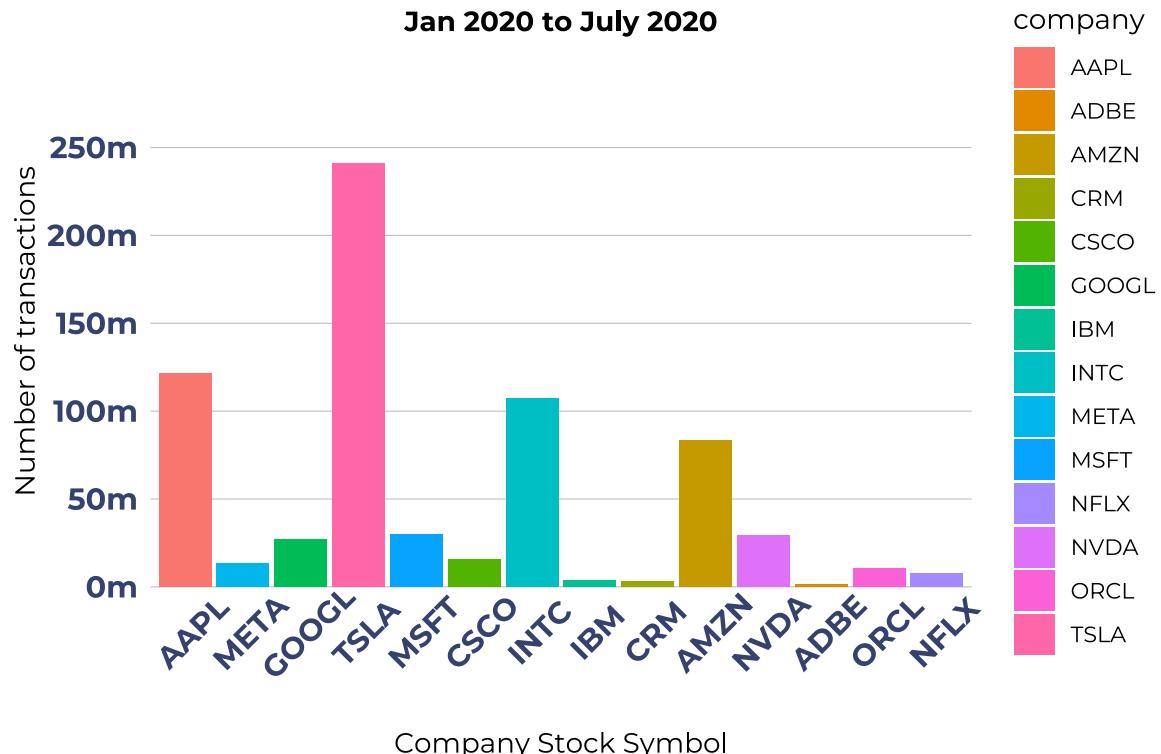
Jan 2020 to July 2020



```
volume_plot + theme_covid2()
```

Volume of low price transaction in Big Tech Companies

Jan 2020 to July 2020



- For the final chart, we can see that except for Netflix and Amazon, other tech companies tended to perform worse than Walmart and equal compared to Pfizer. Further more, in each chart, tech companies often reach the lowest points which show that those companies were damaged more badly compared to the medical or grocery sectors. This can be explained by the demand of people in the pandemic situation, they prefer food, drink, and necessities which can be found at Walmart rather than technological products. Therefore, the sale of this company can be better, so that it can handle the damage better as well.
- However, the greater slope of the going up part of tech companies' lines indicates that they recovered faster compared to Walmart and Pfizer after being hit by COVID-19. Some of them even reach a better stock price at the end. The average line plot shows the same story.

```
compare_to_other("NVDA")
```

Comparision of stock price between NVIDIA Corporation and other Jan 2020 to July 2020



```
compare_to_other("AMZN")
```

Comparision of stock price between Amazon.com, Inc. and other f Jan 2020 to July 2020



```
compare_to_other("NFLX")
```

Comparision of stock price between Netflix, Inc. and other field

Jan 2020 to July 2020

