

Streaming Services - Stock Price Analysis

Introduction

To this day, it has been found that the television audience has been increasingly fragmenting. The main television channels share increasingly lower quotas. Among other reasons, there is the rising of different digital platforms with many original content: movies, series and documentaries that has made the viewers prefer to choose what to watch, when to watch it and where.

The options available to date are gaining more and more followers every day. The huge offer is seeking to capture the attention of any type of audience. Anyone can hire a service specialized in this type of content to enjoy it no matter when or where. In addition, the additional functions included in most streaming platforms give rise to a flexibility that ends up convincing many.

The enormous popularity of many of their own productions has become a value that they usually highlight: Game of Thrones and Stranger Things are good examples of this. Although not the only ones, since many of these options enrich their catalog with titles from other producers and even from their rivals.

Objetive

We are interested in knowing the performance of the 4 top streaming services in terms of views of their top shows and if them had impact within their stock prices.

Data

We will have access to the streaming service historical stock prices of different streaming services over the last five years. Also, we are gathering data using google trends to get to know how popular are some of this shows among the population.

```
library(quantmod)
library(dplyr)
library(tidyverse)
library(maps)
library(mxmaps)
library(gtrendsR)
```

We will start by getting the stock prices historical data of Disney, Amazon, Netflix and AT&T(HBO) from the past 10+ years using the quantmod library.

```
streaming <- c('DIS', 'AMZN', 'NFLX', 'T')
getSymbols(streaming)
```

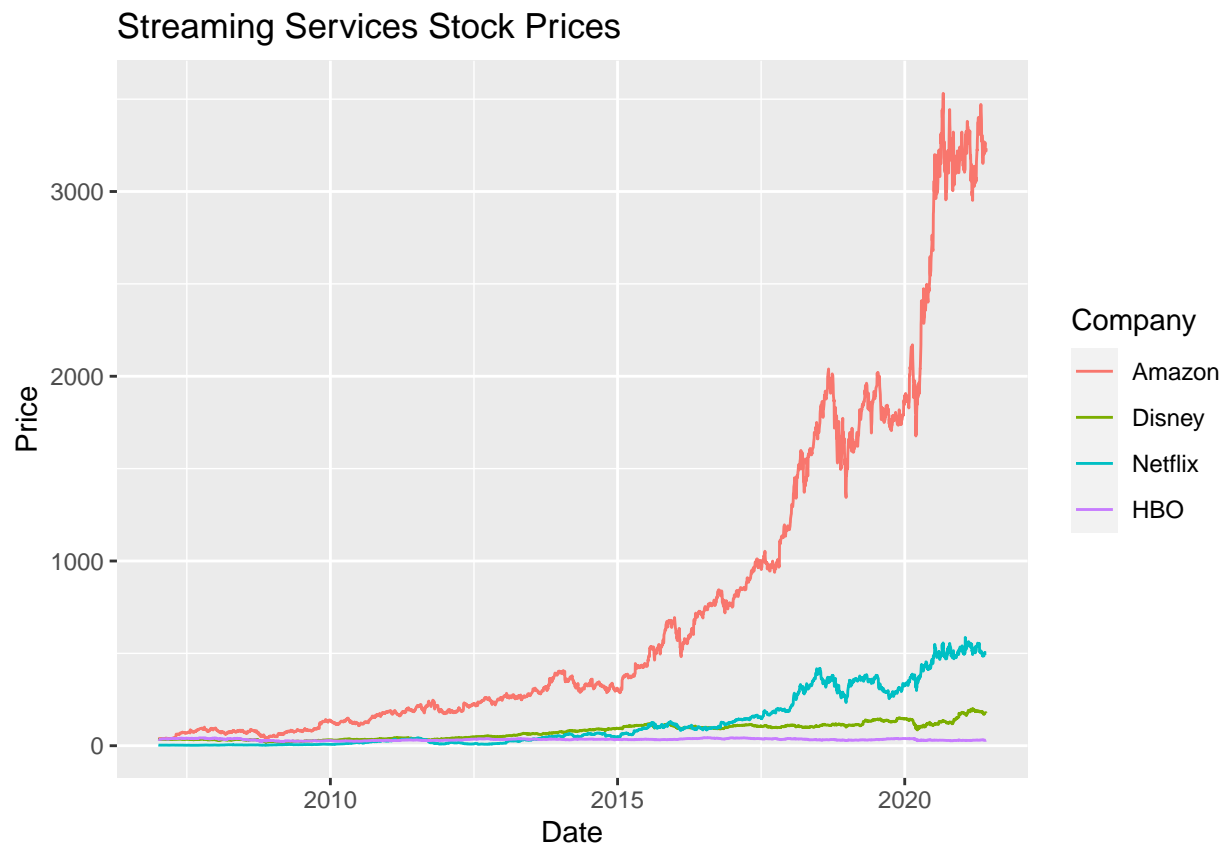
```
[1] "DIS" "AMZN" "NFLX" "T"
```

```
stocks <- data.frame('Amazon' = AMZN$AMZN.Close,
                     'Netflix' = NFLX$NFLX.Close,
                     'Disney' = DIS$DIS.Close,
                     'HBO' = T$T.Close,
                     'Date' = as.Date(row.names(as.data.frame(NFLX))))

graph <- stocks %>%
```

```
gather(value = 'value', key = 'stock', -Date)

graph %>% ggplot() +
  geom_line(aes(x = Date,
                y = value,
                color = stock)) +
  scale_color_discrete(name = 'Company',
                       labels = c('Amazon', 'Disney', 'Netflix', 'HBO')) +
  labs(title = 'Streaming Services Stock Prices',
        y = 'Price')
```



Next, we will use `gtrendsR` to gather data about the popularity of google search terms over time to get to know the impact each show has had among the population and its released date. A visual of the data can help us understand easily the impact magnitude of each show.

```
hits <- gtrends(keyword=c('game of thrones',
                          'the mandalorian',
                          'stranger things',
                          'the marvelous mrs. maisel'))
trend <- hits$interest_over_time
trend <- trend %>% mutate(hits = as.numeric(hits))
```

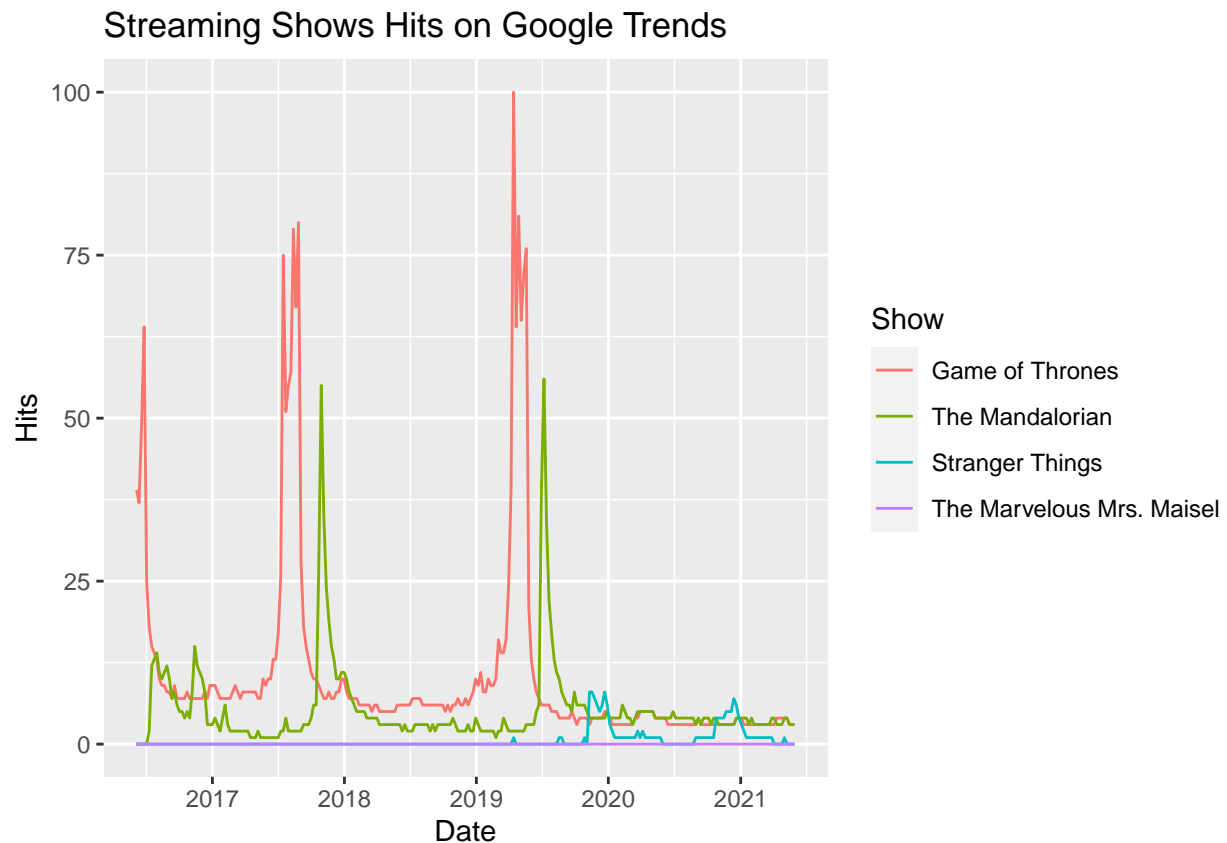
```
## Warning in mask$eval_all_mutate(quo): NAs introducidos por coerción
```

```

trend <- trend %>%
  replace_na(list(hits = 0))

trend %>%
  ggplot() + geom_line(aes(date,
                            hits,
                            color = keyword)) +
  scale_color_discrete(name = 'Show',
                      labels = c('Game of Thrones',
                                  'The Mandalorian',
                                  'Stranger Things',
                                  'The Marvelous Mrs. Maisel' )) +
  labs(title = 'Streaming Shows Hits on Google Trends',
       y = 'Hits',
       x = 'Date')

```



Google trends data was normalized from 0 to 1 to get a proportional sight of the shows. As you may see, Game of thrones seems the most poplar show to this day with an increasing popularity each season. The mandalorian occupies the second spot with a constant trend and Stranger Things falls behind followed by The marvelous Mrs. Maisel.

Us Trends Geographical Distribution

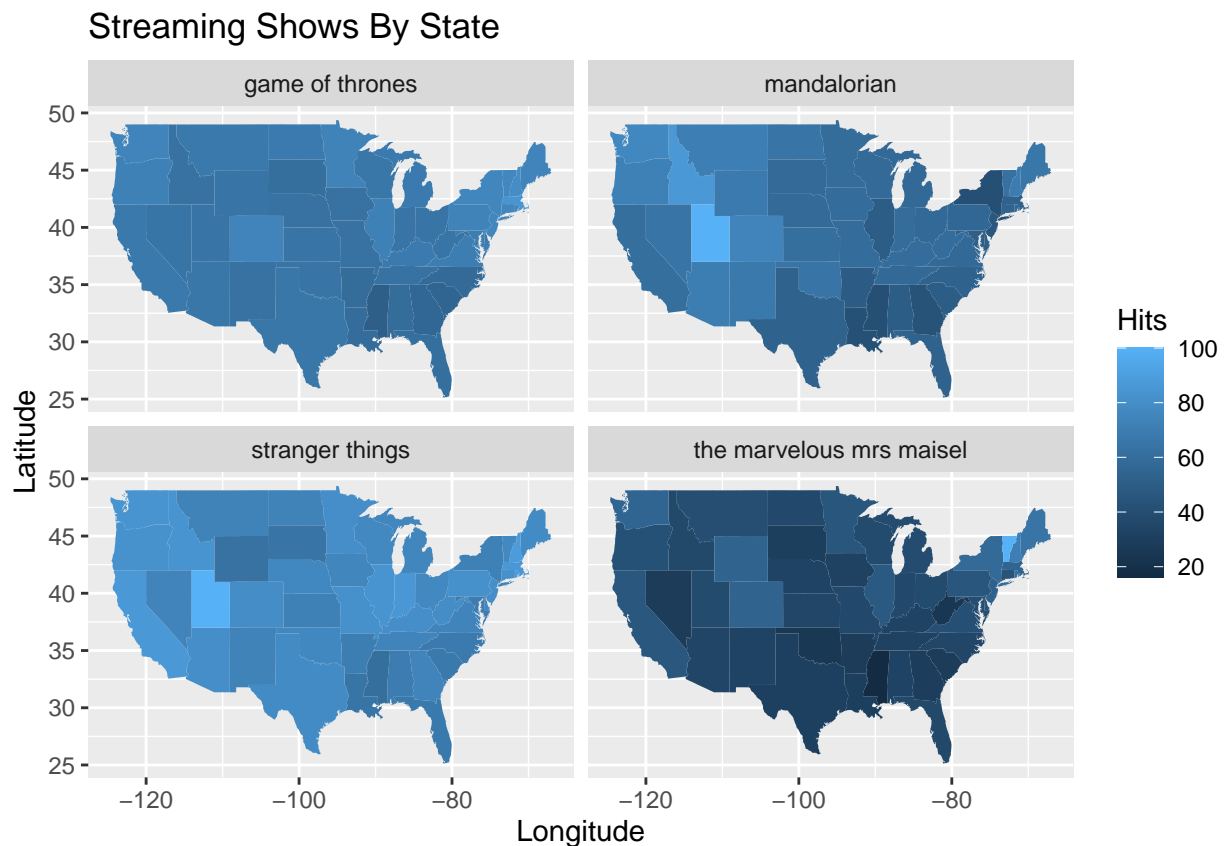
An interesting google trends factor to get to know the viewers market is to get a sight of the fandom distribution, where are the show more popular? This way companies can know where to focus their efforts

to increase sells revenue or where do they have to invest in advertising. This type of insights may help companies get a better understanding of who is watching their shows.

```
us_trends <- gtrends(keyword=c('mandalorian', 'stranger things', 'game of thrones', 'the marvelous mrs
us_states <- us_trends$interest_by_region
us_states <- us_states %>%
  mutate(location = tolower(location))

us_map <- map_data('state')

us_states %>% ggplot(aes(map_id = location)) +
  geom_map(aes(fill = hits),
    map = us_map) +
  expand_limits(x = us_map$long,
    y = us_map$lat) +
  facet_wrap(~ keyword,
    nrow = 2) +
  labs(title = 'Streaming Shows By State',
    y = 'Latitude',
    x = 'Longitude') +
  scale_fill_continuous(name = 'Hits')
```



Game of Thrones and Stranger Things popularity seems to have a uniform popularity among the United States slightly increasing at xx. The Mandalorian appears to be more popular in the west. On the other

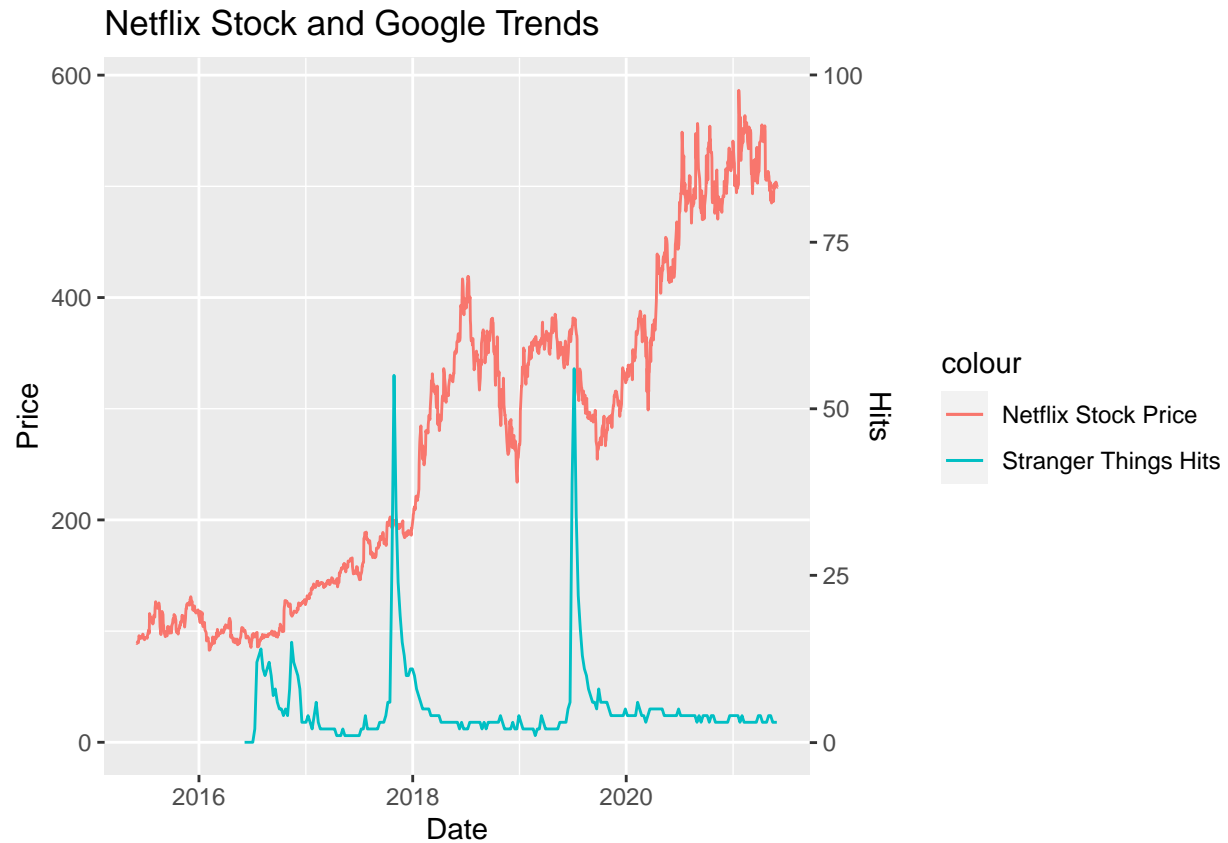
hand , since The Marvelous Mrs. Maisel is set in New York City,it seems to be most popular in the north east.

Companies Stock Prices vs Show Popularity

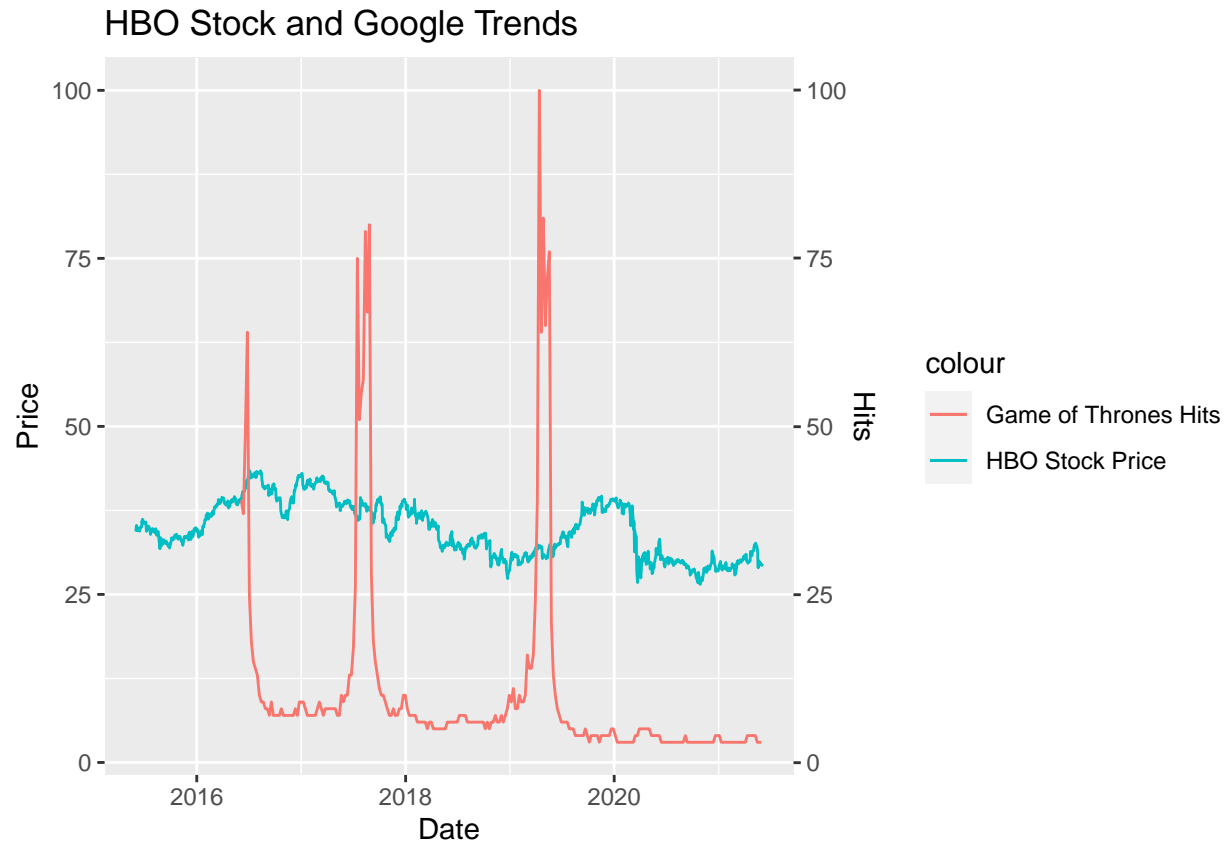
Lets gather some insights of each company using the financial data by plotting stock prices vs google trend popularity of their respective shows. Even though we have stock data since 2010 and before we have to adjust the dates since many of these shows didn't exist at that time.

```
stocks_2015 <- stocks %>%
  filter(Date > as.Date('2015-06-01'))

netflix <- ggplot() +
  geom_line(stocks_2015,
    mapping = aes(x = Date,
                  y= NFLX.Close,
                  color = 'Netflix Stock Price')) +
  geom_line(filter(trend,
    keyword == 'stranger things'),
    mapping = aes(as.Date(date),
                  hits*6,
                  color = 'Stranger Things Hits')) +
  scale_y_continuous(name = 'Price',
    sec.axis = sec_axis(~./6,
                        name='Hits')) +
  labs(title = 'Netflix Stock and Google Trends')
netflix
```



```
HBO <- ggplot() +
  geom_line(stocks_2015,
    mapping = aes(x = Date,
      y= T.Close,
      color = 'HBO Stock Price')) +
  geom_line(filter(trend,
    keyword == 'game of thrones'),
    mapping = aes(as.Date(date),
      hits,
      color = 'Game of Thrones Hits')) +
  scale_y_continuous(name = 'Price',
    sec.axis = sec_axis(~.,
      name='Hits')) +
  labs(title = 'HBO Stock and Google Trends')
HBO
```



The release of *Stranger Things* Season 2 in 2018 may have had an important impact on Netflix stocks since they went up shortly after. On the other hand, *Game of Thrones* last seasons didn't had an important impact on HBO stock prices, this may be due to HBO is part of a larger company (AT&T) and HBO's success doesn't has a huge impact in a company with different bussiness going on.

Checking For Linear Regression

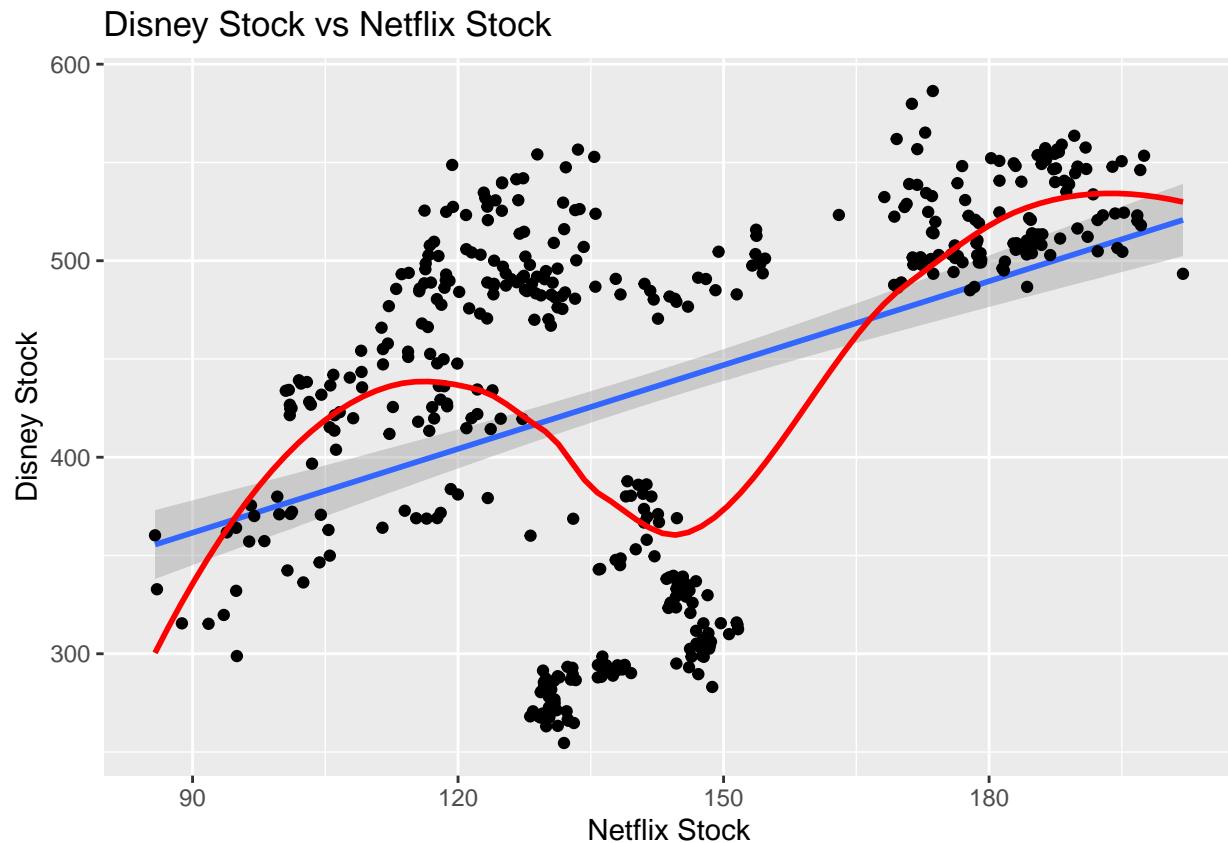
Finally, we wanted to see if there was a correlation between any of these services. Perhaps if Netflix's stock rises, HBO's stock will fall. We created a plot with Netflix's stock on the x axis and Disney's stock on the y axis. Note that we only used data points starting from September 2019 since that was the month of Disney Plus' launch.

We also added a LOESS smoother to visualize the regression line.

```
disney <- stocks %>% filter(Date > as.Date('2019-09-01'))

disney %>% ggplot(aes(x = DIS.Close,
                     y = NFLX.Close)) +
  geom_smooth(method = 'lm') +
  geom_point() +
  geom_smooth(se = F,
             color = 'red') +
  labs(title = 'Disney Stock vs Netflix Stock',
       y = 'Disney Stock',
       x = 'Netflix Stock')
```

```
## 'geom_smooth()' using formula 'y ~ x'
## 'geom_smooth()' using method = 'loess' and formula 'y ~ x'
```



As you can see, this relationship is not linear. Actually unsurprising, it would have been extraordinary if Disney and Netflix stock prices were so closely correlated.

Conclusions and Recommendations

The streaming industry has been rising since 2010, every year shows get more popular, every year stocks prices rise. I particular 2020 was a good year for them, since the covid 19 pandemic hit and people had to stay at home, the necessity of entertainment has become a demand in the market. The next few years will become critical to the entertainment industry, since streaming services has rise over Hollywood and big blockbusters. With more competition, new problems will come, but we will have data science to help us find new solutions.

Back in the day, Netflix used to rule the streaming services, since it was the pioneer or at least the one that had success. It provided 90% of what you actually wanted to see, no commercials or interruptions, all what you wanted to see, whenever you wanted to see. Nowadays we're seeing Netflix lose content to Disney+, Amazon and others, and eventually you will require a lot more than just one or two streaming services too see all your favorite shows and movies.

For this project we compared different methods to evaluate the success not just of the top 4 streaming services but the top ways they offer. Quantmod helped us to gather stock prices, which are a good indicator of the path companies have had. We also used Google Trends to get to know the popularity of the shows and its distribution among the US population. These two data sets gave a general overview of the streaming companies.