The Correlation of THY's Stock Price & Their Supbowl Ads

Mehmet Arslan

12/6/2020

```r
library(ggplot2)
library(tidyverse)
library(ggthemes)
library(dplyr)
library(readr)
library(readxl)
library(gridExtra)
library(ggdark)
library(PerformanceAnalytics)
```

```r
project_data = read_csv("cleanThy.csv")

head(project_data, 10)
```

```
## # A tibble: 10 x 8
##    Date      Day Month  Year `Last Price` `Lowest Price` `Highest Price`  Volume
##    <chr>   <dbl> <dbl> <dbl>        <dbl>          <dbl>           <dbl>   <dbl>
##  1 18-Jul~    18     7  2019         13.2           13.0            13.3  8.52e8
##  2 17-Jul~    17     7  2019         13.0           12.6            13.0  8.25e8
##  3 14-Jul~    14     7  2019         12.7           12.6            12.9  7.43e8
##  4 13-Jul~    13     7  2019         12.6           12.6            12.9  6.50e8
##  5 12-Jul~    12     7  2019         13.0           12.9            13.2  8.20e8
##  6 11-Jul~    11     7  2019         12.9           12.7            13.3  9.51e8
##  7 10-Jul~    10     7  2019         13.3           12.9            13.4  7.97e8
##  8 7-Jul-~     7     7  2019         13.0           12.4            13.1  8.30e8
##  9 3-Jul-~     3     7  2019         12.4           12.4            12.6  1.99e8
## 10 31-May~    31     5  2019         12.5           12.4            12.6  6.86e8
```

#1.) A - There are 1,502 rows in my dataset. B - There are 5 columns in my dataset. D - There are no missing values that I can identify.

#2.) The first question I seek to answer using data visualizations: Kobe Bryant's impact on Turkish Airlines' stock price.

#2.) Did Kobe Bryant's first appearance in a commercial have any impact on Turkish Airlines' stock price?

#2.) The relevant variables will be: Date(all the dates are written in Turkish, I can translate them), lowest stock price on that day, and the highest stock price on that day.

The datasets used for this project are titled: "cleanThy", "THY 2015 February Stocks", "THY 2016 February Stocks", "THY 2017 February Stocks", "THY 2018 February Stocks", "THY 2019 February Stocks", "THY 2020 February Stocks", and, "THY Stock Price 2013 - 2020" The purpose of this project was to identify and then quantify the correlation between celebrity appearances in commercials and Turkish Airlines' stock prices from 2013 to 2020.

All stock data for Turkish Airlines was sourced from Yahoo Finance.

```r
super_cleanTHY = project_data %>%
  select(3:7)

names(super_cleanTHY) = c("Month", "Year", "Last Price",
                          "Lowest Price", "Highest Price")

head(super_cleanTHY)
```

```
## # A tibble: 6 x 5
##    Month  Year `Last Price` `Lowest Price` `Highest Price`
##    <dbl> <dbl>        <dbl>          <dbl>           <dbl>
## 1      7  2019         13.2           13.0            13.3
## 2      7  2019         13.0           12.6            13.0
## 3      7  2019         12.7           12.6            12.9
## 4      7  2019         12.6           12.6            12.9
## 5      7  2019         13.0           12.9            13.2
## 6      7  2019         12.9           12.7            13.3
```

I noticed that there were some extra variables, namely "Volume", that I wouldn't need to do anything with. I wrangled the dataset and then gave renamed the columns accordingly.

```r
monthly_low_avg = super_cleanTHY %>%
  group_by(Month) %>%
  summarize(Average = mean(`Lowest Price`, na.rm = T))

lowest_price_avg = left_join(monthly_low_avg, super_cleanTHY, by = "Month")

head(lowest_price_avg)
```

```
## # A tibble: 6 x 6
##    Month Average  Year `Last Price` `Lowest Price` `Highest Price`
##    <dbl>   <dbl> <dbl>        <dbl>          <dbl>           <dbl>
## 1      1    9.83  2019         15.4           15.2            16.2
## 2      1    9.83  2019         15.8           15.1            15.8
## 3      1    9.83  2019         15.2           15.1            15.2
## 4      1    9.83  2019         15.1           14.7            15.2
## 5      1    9.83  2019         14.7           14.6            14.9
## 6      1    9.83  2019         14.8           14.6            14.9
```

```r
monthly_high_avg = super_cleanTHY %>%
  group_by(Month) %>%
  summarize(Average = mean(`Highest Price`, na.rm = T))

highest_price_avg = left_join(monthly_high_avg, super_cleanTHY, by = "Month")

head(highest_price_avg)
```

```
## # A tibble: 6 x 6
##    Month Average  Year `Last Price` `Lowest Price` `Highest Price`
##    <dbl>   <dbl> <dbl>        <dbl>          <dbl>           <dbl>
## 1      1    10.1  2019         15.4           15.2            16.2
## 2      1    10.1  2019         15.8           15.1            15.8
## 3      1    10.1  2019         15.2           15.1            15.2
## 4      1    10.1  2019         15.1           14.7            15.2
## 5      1    10.1  2019         14.7           14.6            14.9
## 6      1    10.1  2019         14.8           14.6            14.9
```

Because I wanted to identify if there were any seasonal trends in the data, I needed to group my data by the month variable to make any seasonal trends more easily identifiable.

```
five_num_summary = super_cleanTHY %>%
  summarize(stock.med = median(as.numeric(`Last Price`)),
            stock.avg = mean(as.numeric(`Last Price`)),
            stock.sd = sd(as.numeric(`Last Price`)),
            stock.min = min(as.numeric(`Last Price`)),
            stock.max = max(as.numeric(`Last Price`)))

five_num_summary
```

```
## # A tibble: 1 x 5
##   stock.med stock.avg stock.sd stock.min stock.max
##       <dbl>     <dbl>    <dbl>     <dbl>     <dbl>
## 1      8.24      9.63     4.01      4.63      19.8
```

The average stock price, median stock price, and standard deviation were calculated directly.

```
THY_stock_trend <- read_csv("THY Stock Price 2013 - 2019.csv")

THY_stock_trend %>% ggplot(aes(x = as.factor(Month), y = Close)) +
  geom_boxplot(color = "black", fill = "blue", alpha = 0.8) +
  labs(title = "Turkish Airlines' Final Stock Price Between 2013 & 2019",
       subtitle = "Source: Yahoo Finance",
       x = "Month",
       y = "Stock Price (in Dollars)")
```

Turkish Airlines' Final Stock Price Between 2013 & 2019
Source: Yahoo Finance
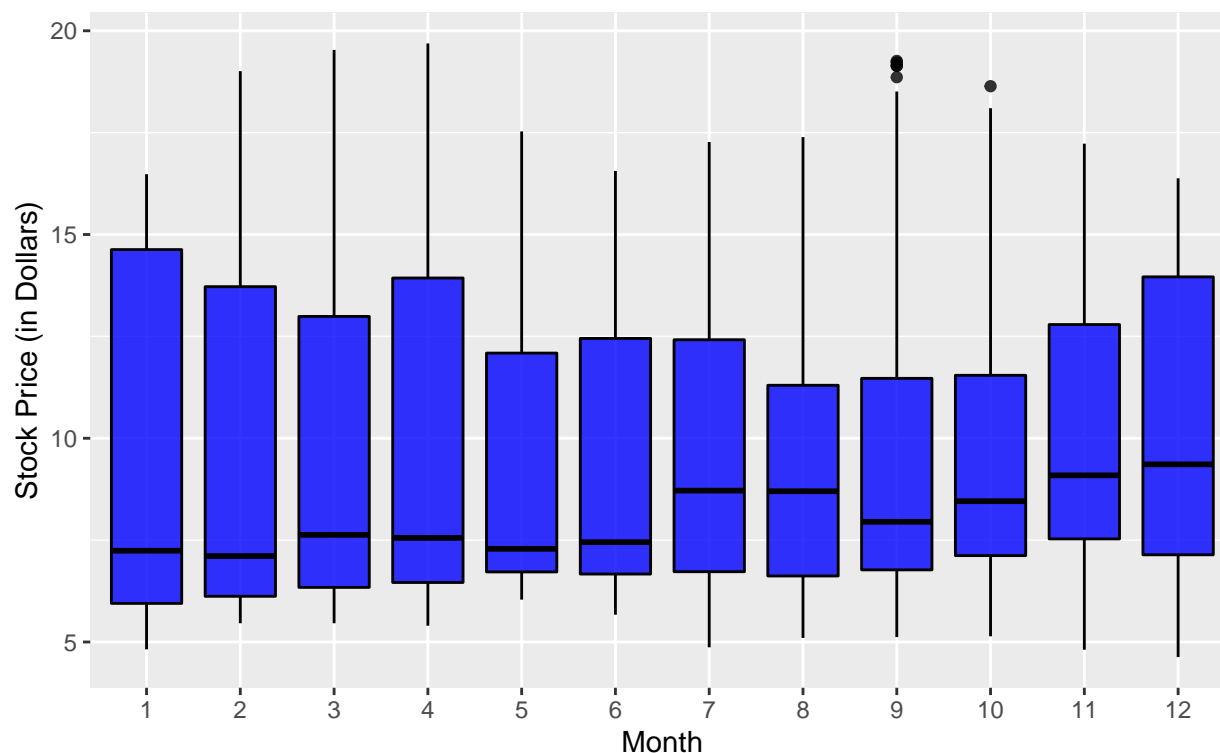
The boxplot did a decent job of displaying the variability in my data however, it doesn't exactly paint the clearest picture. I noticed that the the stock price varies quite a bit throughout the year.

```
lowest_price_avg %>% group_by(Month) %>%
  summarize(low_avg = mean(`Lowest Price`, na.rm = T),
            high_avg = mean(`Highest Price`, na.rm = T)) %>%
  ggplot(aes(Month)) +
  geom_line(aes(y = low_avg), color = "yellow", show.legend = T) +
  geom_line(aes(y = high_avg), color = "red", show.legend = T) +
  labs(title = "Average of Stock Prices Between 2013 & 2019",
       subtitle = "Souce: Yahoo Finance",
       y = "Stock Price (in Dollars)") +
  dark_theme_gray() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```
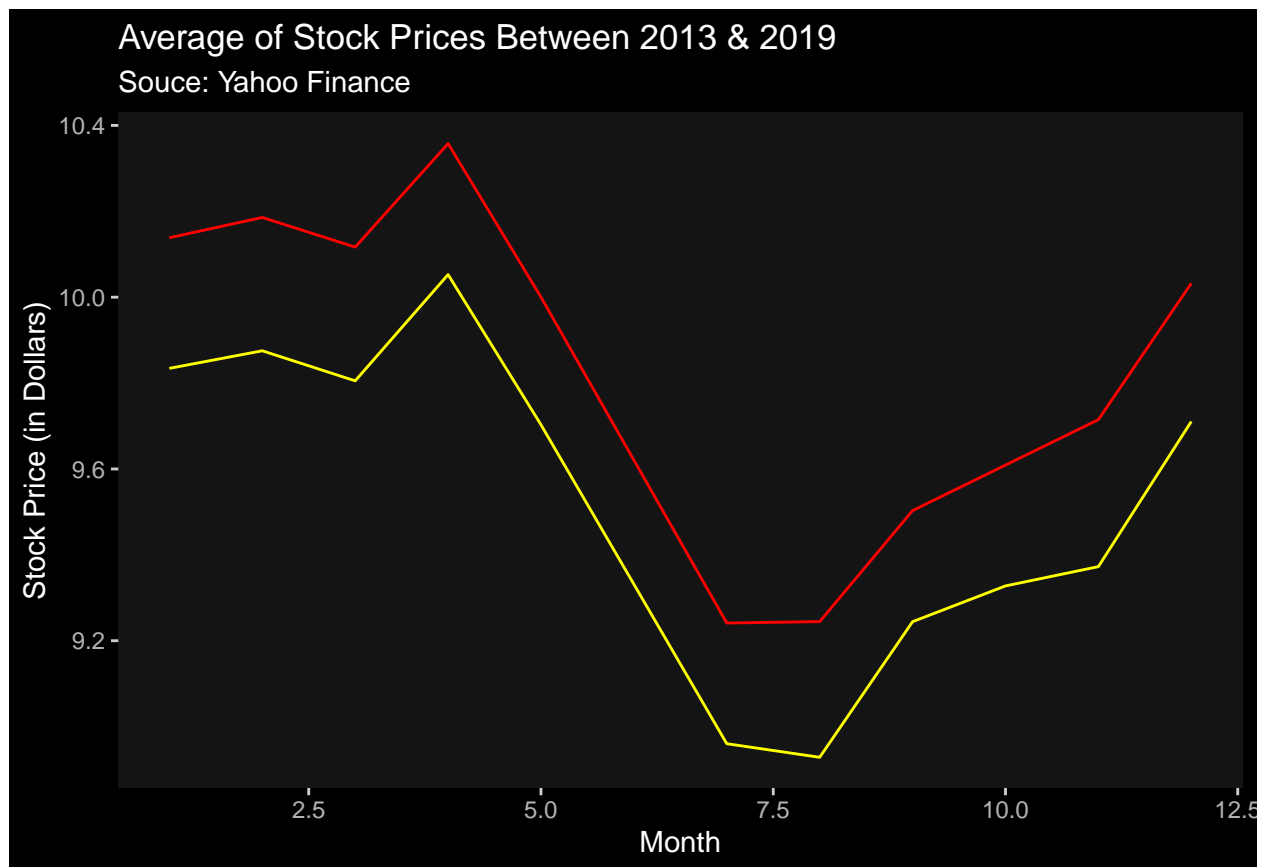

I noticed an interesting trend when plotting the stock prices in a line graph. Mainly that there is a very noticeable increase in both the highest stock price and the lowest stock price earlier in February. I did notice that there is a sharp decline in stock value starting in mid April and continuing through the summer. This is odd to say the least because more people go on vacation in the summer months, the stock should be increasing.

My original thought was that the celebrity appearances in Turkish Airlines' commercials would drive up the stock price. Considering that the largest, most recent, and most publicly available data are in regards to Turkish Airlines' Superbowl commercials, I sought to explore the correlation between Superbowl commercials and stock price.

The first Superbowl commercial that Turkish Airlines aired was during Superbowl 50 which took place on February 7, 2016. (Super Bowl History). The purpose of the ad was to generate social media buzz about

4

Turkish Airlines and potentially lead to more business. The ads in question were promoting the 2016 film "Batman V Superman." (Paul Chavez).

```r
THY_2_2016 <- read_csv("THY 2016 February Stocks.csv")

THY_2_2016 %>% ggplot(aes(Date, Close)) +
  geom_line(color = "orange") +
  geom_vline(aes(xintercept = as.Date.character("2016-02-07"))) +
  geom_label(x = as.Date.character("2016-02-07"),
             y = 6.7, label = "Superbowl 50", color = "white") +
  labs(title = "Turkish Airlines' Final Daily Stock Price for February 2016",
       subtitle = "Source: Yahoo Finance",
       y = "Final Stock Price (in Dollars)") +
  dark_theme_gray() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```
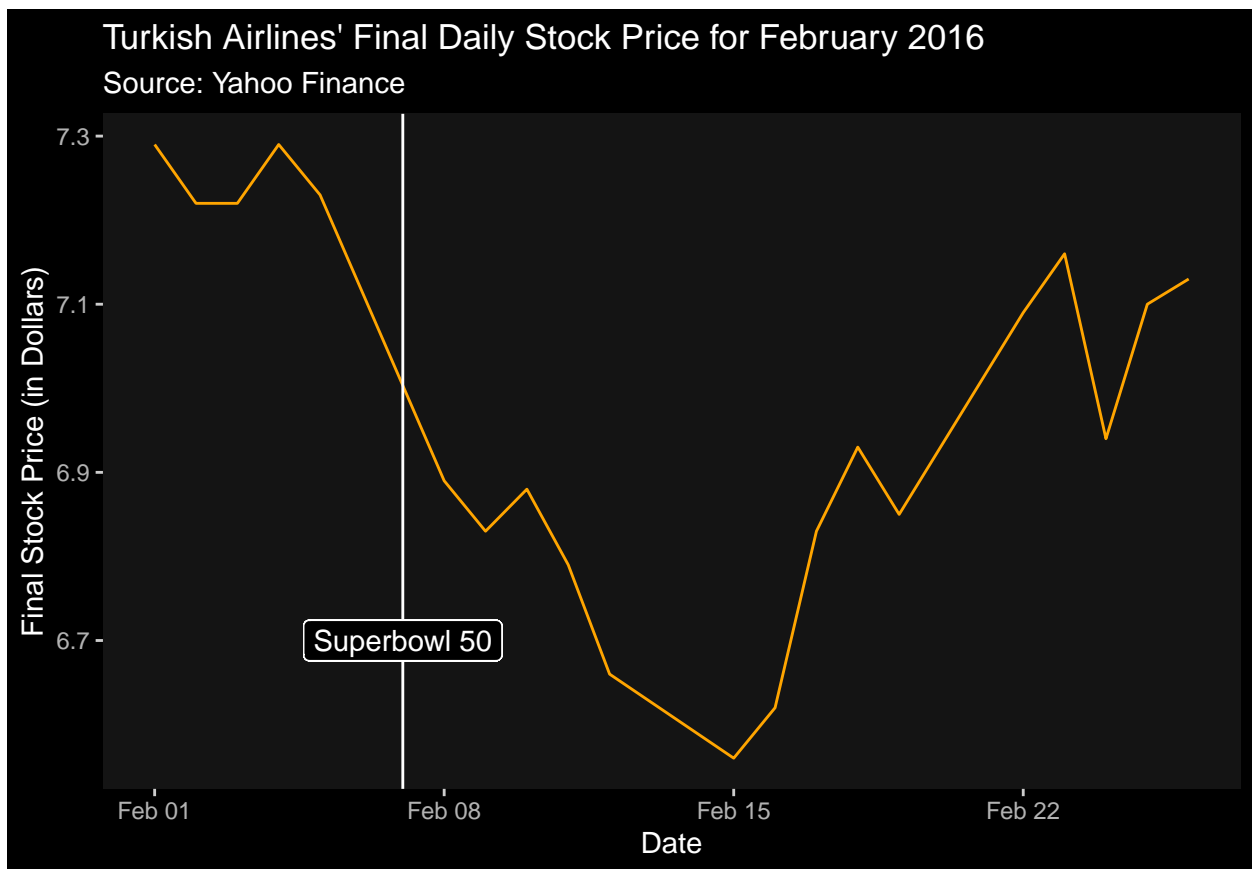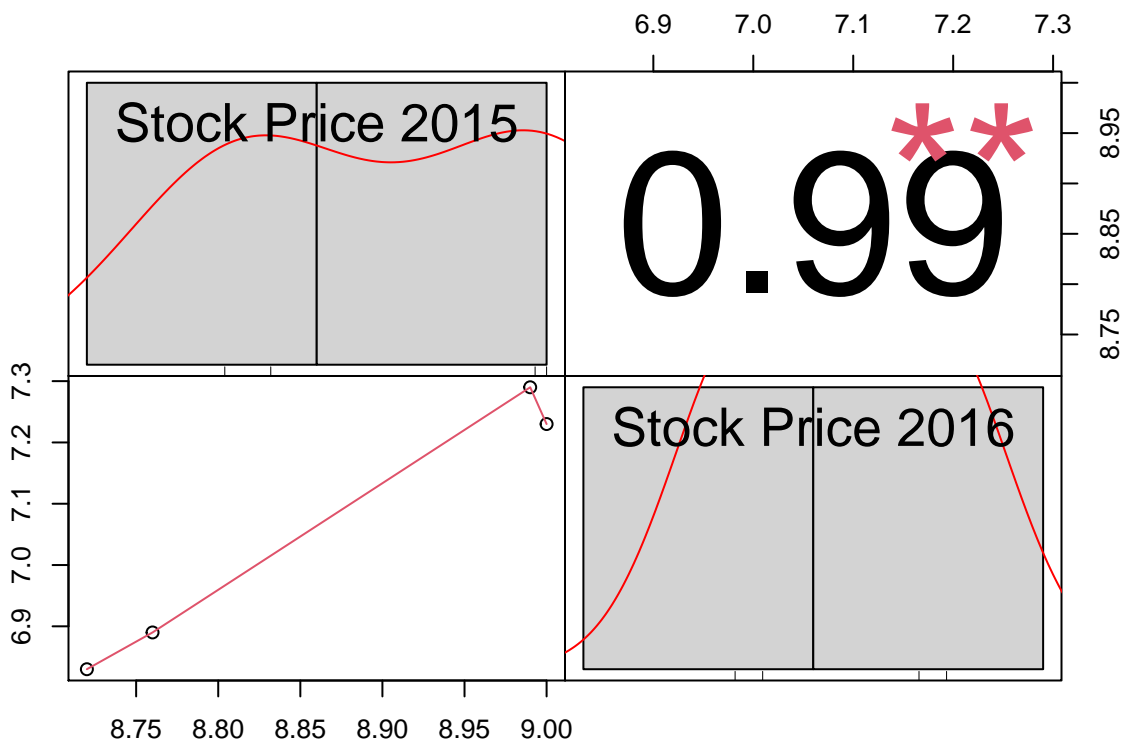


```r
Correlation_for_2016 <- read_csv("Correlation for 2016.csv")

cor(Correlation_for_2016)
```

```
##                 Stock Price 2015 Stock Price 2016
## Stock Price 2015        1.0000000        0.9912048
## Stock Price 2016        0.9912048        1.0000000
```

```r
chart.Correlation(Correlation_for_2016, histogram = T, pch = 19)
```

In the above chart, we see that there was a noticeable uptick just before February 8, 2016. (Superbowl History). Afterwards, it looks to be more normalized but there was in fact a sudden and very sharp increase following the Superbowl. To prove any correlation with the data, I used the February 2015 dataset as my benchmark. This was for two reasons: February 2015 was the final year before Turkish Airlines aired any Superbowl commercials, and they have aired one in every Superbowl from 2016 to 2020.

```
THY_2_2017 <- read_csv("THY 2017 February Stocks.csv")

THY_2_2017 %>% ggplot(aes(Date, Close)) +
  geom_line(color = "deepskyblue1") +
  geom_vline(xintercept = as.Date.character("2017-02-05")) +
  geom_label(x = as.Date.character("2017-02-05"), y = 5.70,
             label = "Superbowl 51") +
  labs(title = "Turkish Airlines' Final Daily Stock Price for February 2017",
       subtitle = "Source: Yahoo Finance",
       y = "Final Stock Price (in Dollars)") +
  dark_theme_gray()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```
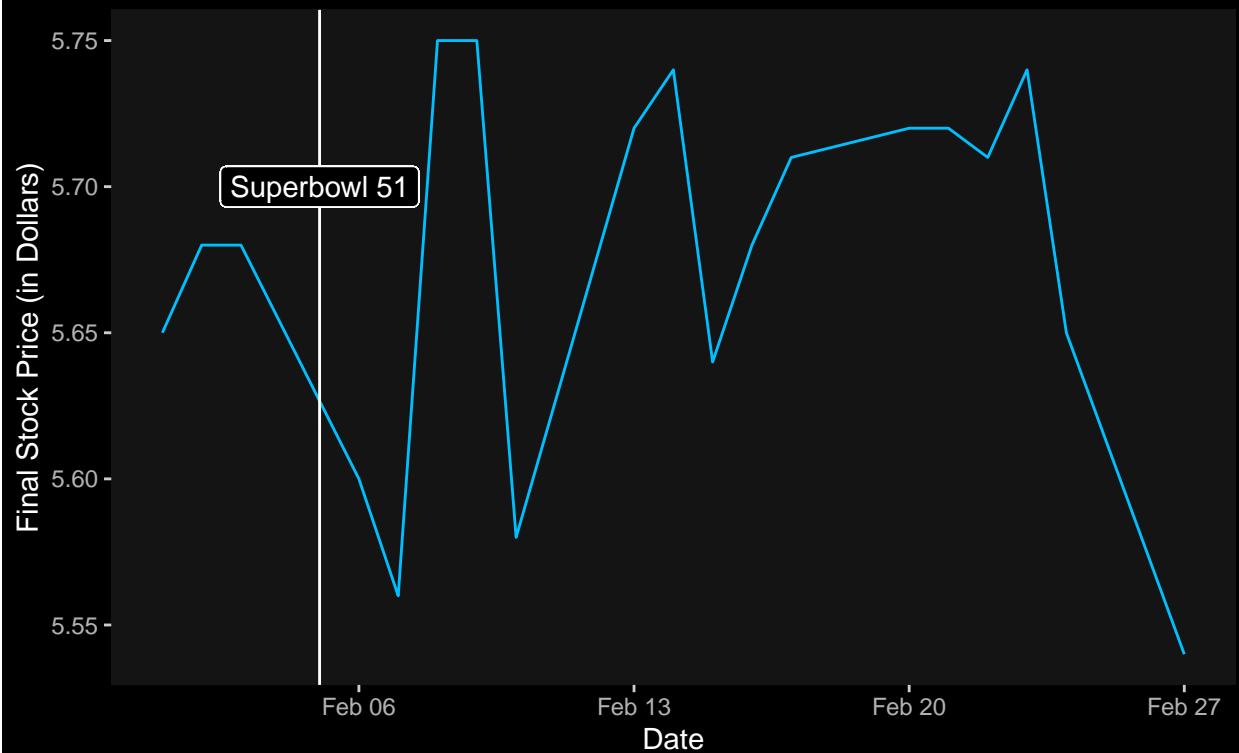
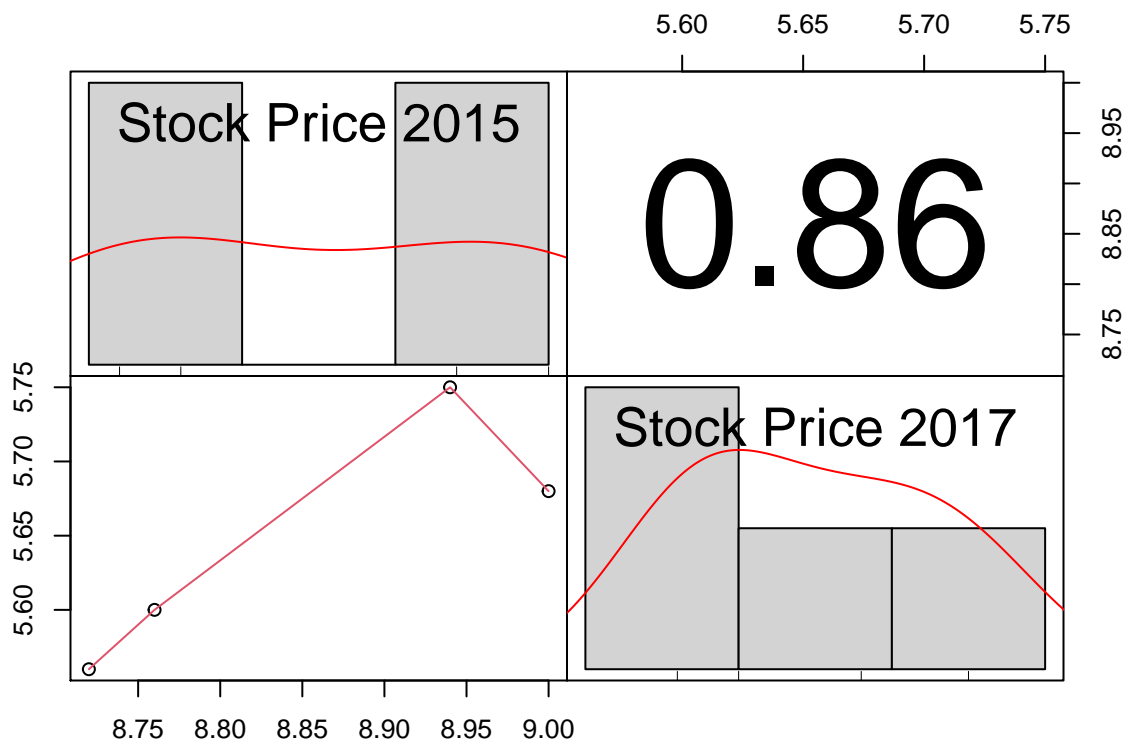Turkish Airlines' Final Daily Stock Price for February 2017
Source: Yahoo Finance

```
Correlation_for_2017 <- read_csv("Correlation for 2017.csv")

cor(Correlation_for_2017)
```

```
##                 Stock Price 2015 Stock Price 2017
## Stock Price 2015        1.0000000        0.8617353
## Stock Price 2017        0.8617353        1.0000000
```

```
chart.Correlation(Correlation_for_2017, histogram = T, pch = 19)
```

Superbowl 51 took place on February 5, 2017. (Super Bowl History). When trading resumed on the following Monday, we can see a very sharp increase in stock value. The increase could be attributed to the fact that Turkish Airlines was the chief sponsor of the event. (THY Super Bowl Sponsor). I theorize that this is attributed to Morgan Freeman's appearance as well as THY's "Enhancing Your World Through Wonderment" ad campaign. Such an ad is a powerful one, and that alone could contribute to the near impossible rise in stock when trading resumed.

```
THY_2_2018 <- read_csv("THY 2018 February Stocks.csv")

THY_2_2018 %>% ggplot(aes(Date, y = Close)) +
  geom_line(color = "green") +
  geom_vline(xintercept = as.Date.character("2018-02-04")) +
  geom_label(x = as.Date.character("2018-02-04"), y = 18,
             label = "Superbowl 52", color = "white") +
  labs(title = "Turkish Airlines' Final Daily Stock Price for February 2018",
       subtitle = "Source: Yahoo Finance",
       y = "Final Stock Price (in Dollars)") +
  dark_theme_gray() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```

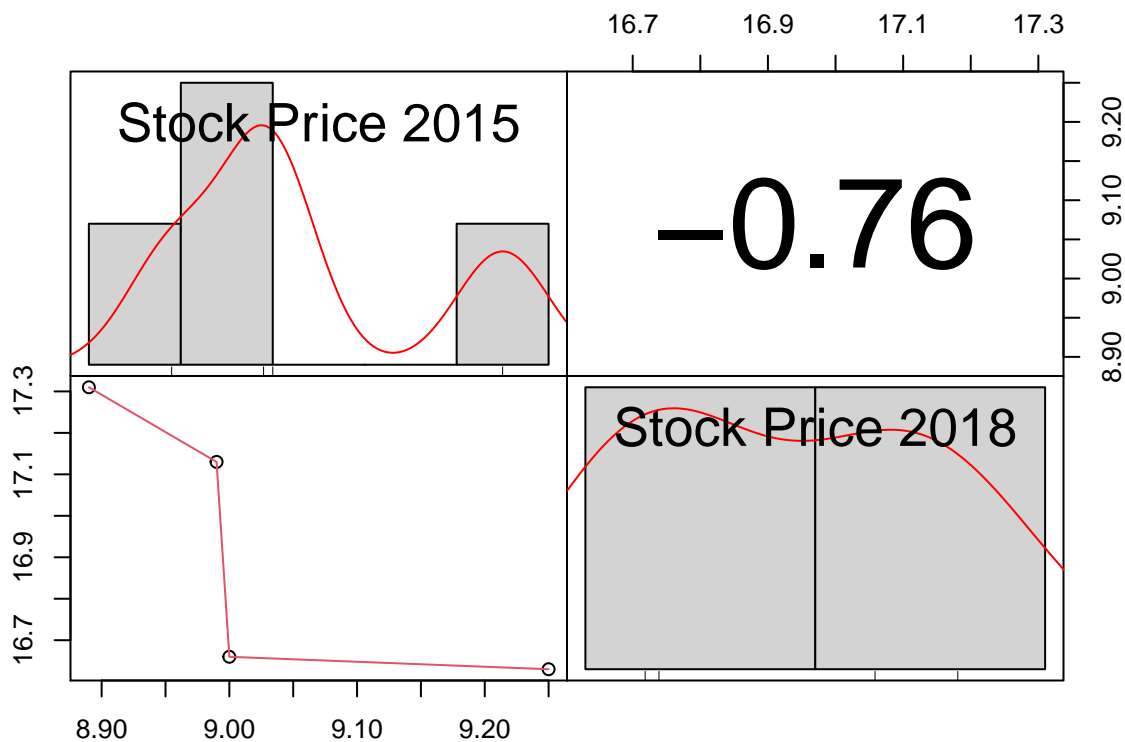Turkish Airlines' Final Daily Stock Price for February 2018
Source: Yahoo Finance
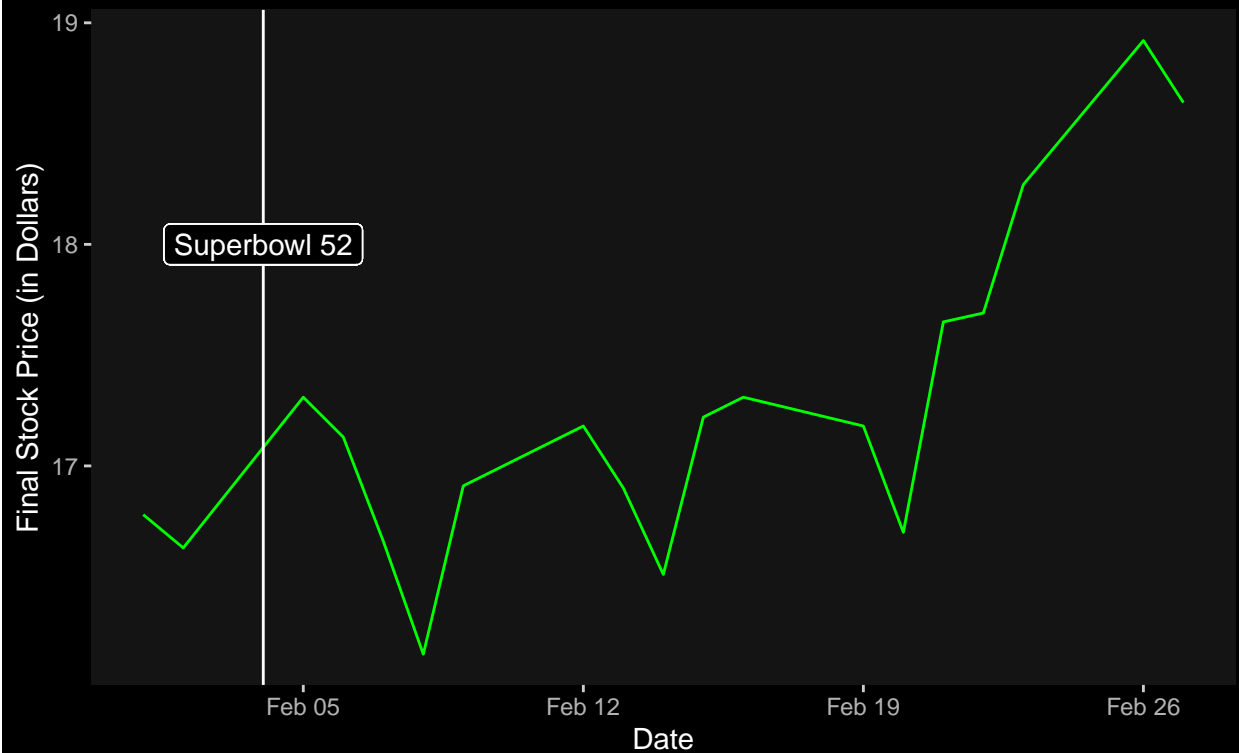
```r
Correlation_for_2018 <- read_csv("Correlation for 2018.csv")

cor(Correlation_for_2018)
```

```
##                 Stock Price 2015 Stock Price 2018
## Stock Price 2015        1.0000000       -0.7614635
## Stock Price 2018       -0.7614635        1.0000000
```

```r
chart.Correlation(Correlation_for_2018, histogram = T, pch = 19)
```

Superbowl 52 took place on February 4th, 2018 (Superbowl History.) The ad featured well known physician, Dr. Mehmet Oz. (Five Senses). From the stock prices, it is now evident that Dr. Mehmet Oz had a strong negative impact on THY's stock price.

```
THY_2_2019 <- read_csv("THY 2019 February Stocks.csv")

THY_2_2019 %>% ggplot(aes(Date, Close)) +
  geom_line(color = "cyan") +
  geom_vline(xintercept = as.Date.character("2019-02-03")) +
  geom_label(x = as.Date.factor("2019-02-03"), y = 14.0,
             label = "Superbowl 53", color = "white") +
  labs(title = "Turkish Airlines' Final Stock Price for February 2019",
       subtitle = "Source: Yahoo Finance",
       y = "Final Stock Price (in Dollars)") +
  dark_theme_gray() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```
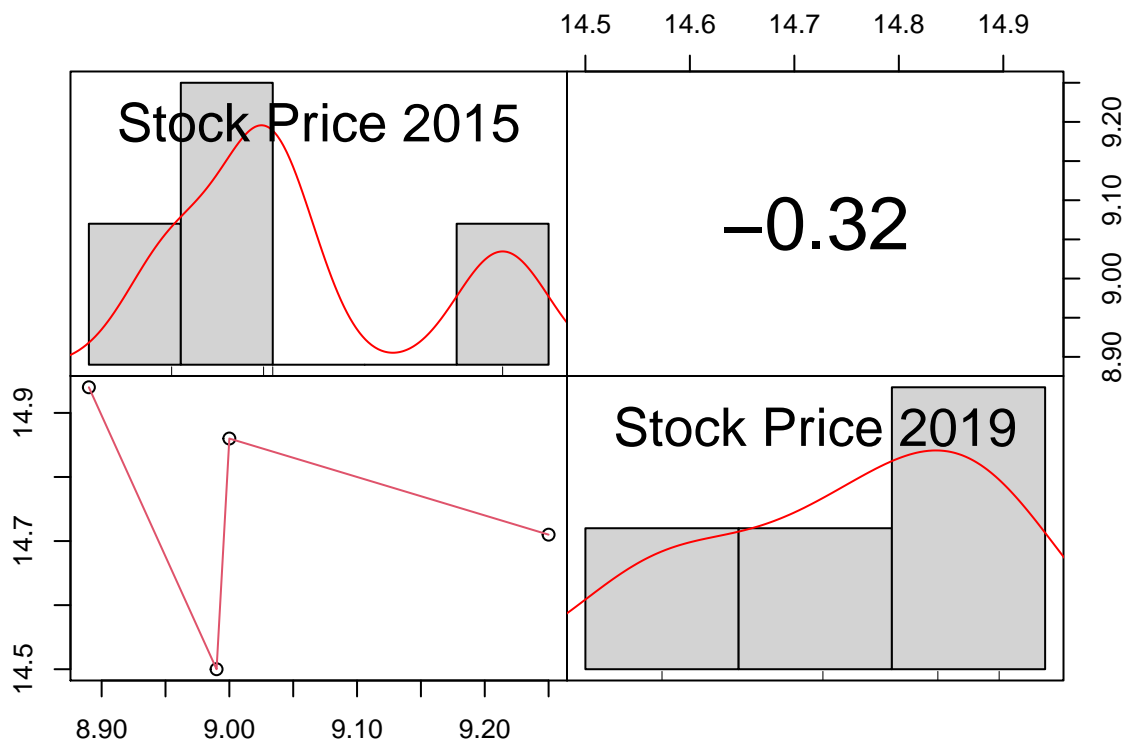
Turkish Airlines' Final Stock Price for February 2019
Source: Yahoo Finance

```
Correlation_for_2019 <- read_csv("Correlation for 2019.csv")

cor(Correlation_for_2019)
```

```
##                 Stock Price 2015 Stock Price 2019
## Stock Price 2015     1.0000000       -0.3229245
## Stock Price 2019    -0.3229245        1.0000000
```

```
chart.Correlation(Correlation_for_2019, histogram = T, pch = 19)
```

Superbowl 53 took place on February 3, 2019. (Super Bowl History). This ad featured a short-film directed by Ridley Scott. The short-film featured actresses Sylvia Hoakes and Aure Atika. The short film's plot was a proverbial game of cat-and-mouse between spies set in Istanbul, Turkey. (The Journey). This advertisement sought to convey the newly rebuilt Istanbul International Airport following the bombing that occurred in 2016. (Istanbul Airport Attack). The other incentive was to show the scenic parts of the city of Istanbul. The short film was received well, but negatively impacted the stock price when trading resumed on February 4, 2019.

Notice how strangely Morgan Freeman's appearance impacted the stock price as compared to everyone else?

```
THY_2_2020 <- read_csv("THY 2020 February Stocks.csv")

THY_2_2020 %>% ggplot(aes(Date, Close)) +
  geom_line(color = "white") +
  geom_vline(xintercept = as.Date.character("2020-02-02")) +
  geom_label(x = as.Date.character("2020-02-02"), y = 13,
             label = "Superbowl 54", color = "white") +
  labs(title = "Turkish Airlines' Final Daily Stock Price for February 2020",
       subtitle = "Source: Yahoo Finance",
       y = "Final Stock Price (in Dollars)" ) +
  dark_theme_gray() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```

Turkish Airlines' Final Daily Stock Price for February 2020
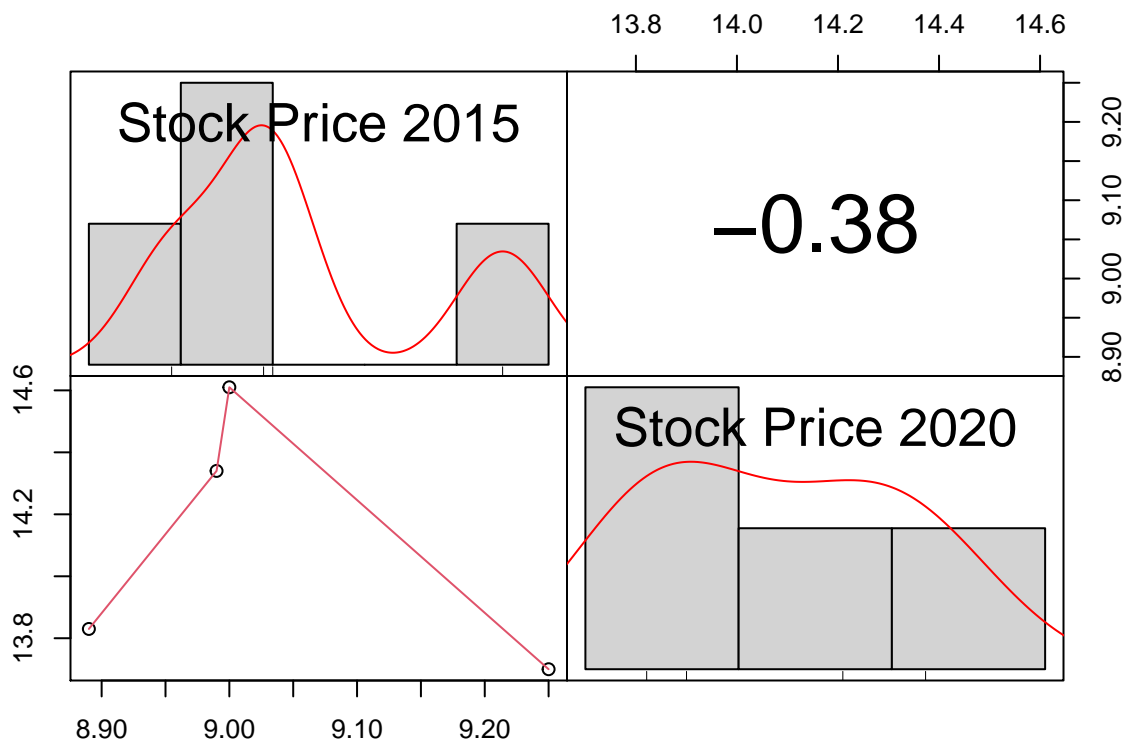Source: Yahoo Finance
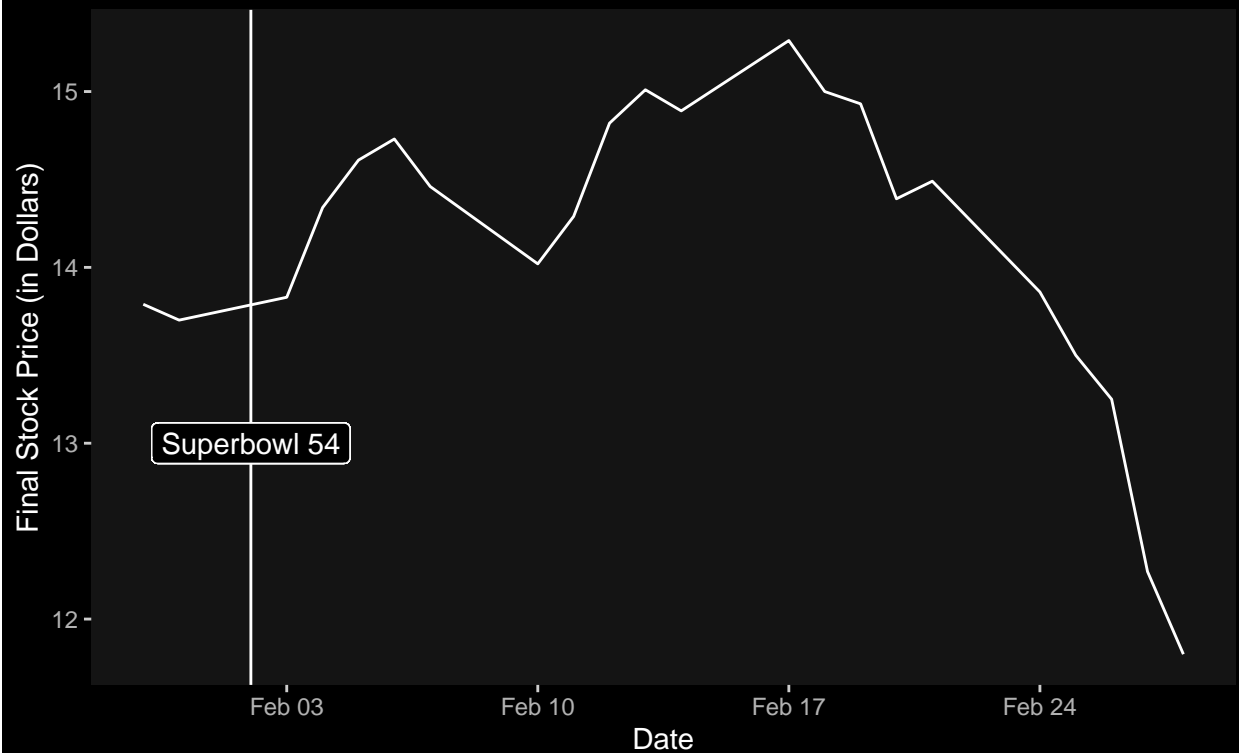
```r
Correlation_for_2020 <- read_csv("Correlation for 2020.csv")

cor(Correlation_for_2020)
```

```
##                 Stock Price 2015 Stock Price 2020
## Stock Price 2015         1.000000        -0.382817
## Stock Price 2020        -0.382817         1.000000
```

```r
chart.Correlation(Correlation_for_2020, histogram = T, pch = 19)
```

Superbowl 54 took place on February 2, 2020. (Superbowl History). When trading resumed on February 3, 2020, Turkish Airlines' stock increased significantly and was more sustainable. Incidentally, this was the first year that Turkish Airlines didn't feature a celebrity endorsement of any kind. In fact, the ad they aired in Superbowl 54 was very aggressive and directed by David Madrid. (Walsh. S.M.). The purpose of the ad was to illustrate that Turkish Airlines had more destinations than almost any other airline and was proposing the very powerful message of "Step On Earth." A simple, yet powerful message that was conveyed clearly had a significantly more noticeable impact on stock price when trading resumed. In fact, the stock price increased when trading resumed and stayed elevated for quite some time. This was a feat that hadn't been accomplished in any of their previous commercials. I think that this can be attributed to how powerful their ad campaign was for this year.

When I first started the project, I stumbled upon a study of consumer habits in different continents. The study clearly showed that American Consumers respond very well to aggressive and powerful ads. I think I have submitted enough to prove that this is the case.

```
correlation_THY <- read_excel("Book2.xlsx")

cor(correlation_THY)

##            Close 2016 Close 2017 Close 2018 Close 2019 Close 2020
## Close 2016  1.0000000 -0.3520738  0.1016274  0.4726536 -0.6084129
## Close 2017 -0.3520738  1.0000000 -0.2764093 -0.4136929  0.4606287
## Close 2018  0.1016274 -0.2764093  1.0000000 -0.2450932 -0.3404772
## Close 2019  0.4726536 -0.4136929 -0.2450932  1.0000000 -0.6532644
## Close 2020 -0.6084129  0.4606287 -0.3404772 -0.6532644  1.0000000

chart.Correlation(correlation_THY, histogram = T, pch = 19)
```
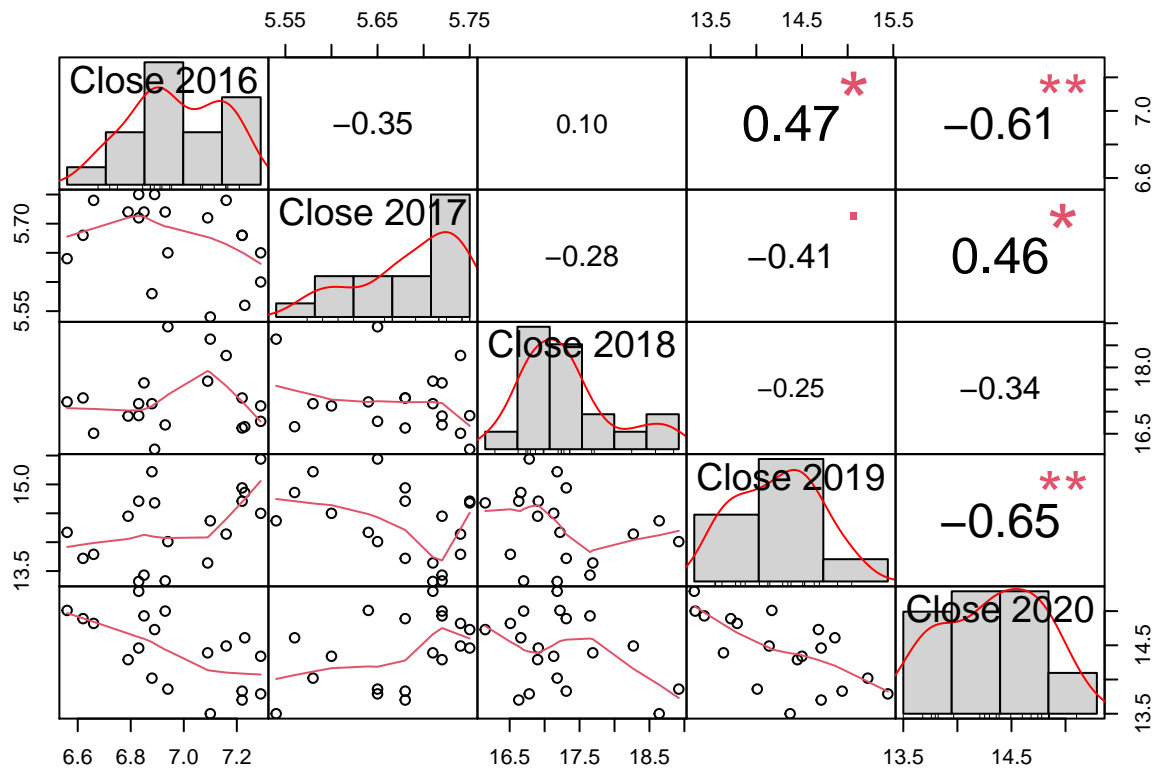
In summation, from what I have learned in my process of basic analysis, I can definitively state that celebrity endorsements and appearances had an overall negative impact on Turkish Airlines' stock price. I can also conclude that the most successful ad campaign, by the numbers at least, was David Madrid's bold advertisement. I have learned a lot and this whole process has been eye opening and I was tremendously excited to learn a new skill

```
Stock_predict_2020 <- read_csv("THY 2020 December Stocks.csv")


Stock_predict_2020 %>% ggplot(aes(Date, Close), na.rm = T) +
  geom_line(color = "salmon") +
  geom_abline(intercept = -917.99, slope = 0.05) +
  labs(title = "Turkish Airlines' Daily Stock Price for November - December 2020",
       subtitle = "Source: Yahoo Finance",
       y = "Final Stock Price (in Dollars)") +
  dark_theme_gray() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())
```

Turkish Airlines' Daily Stock Price for November – December 2020
Source: Yahoo Finance
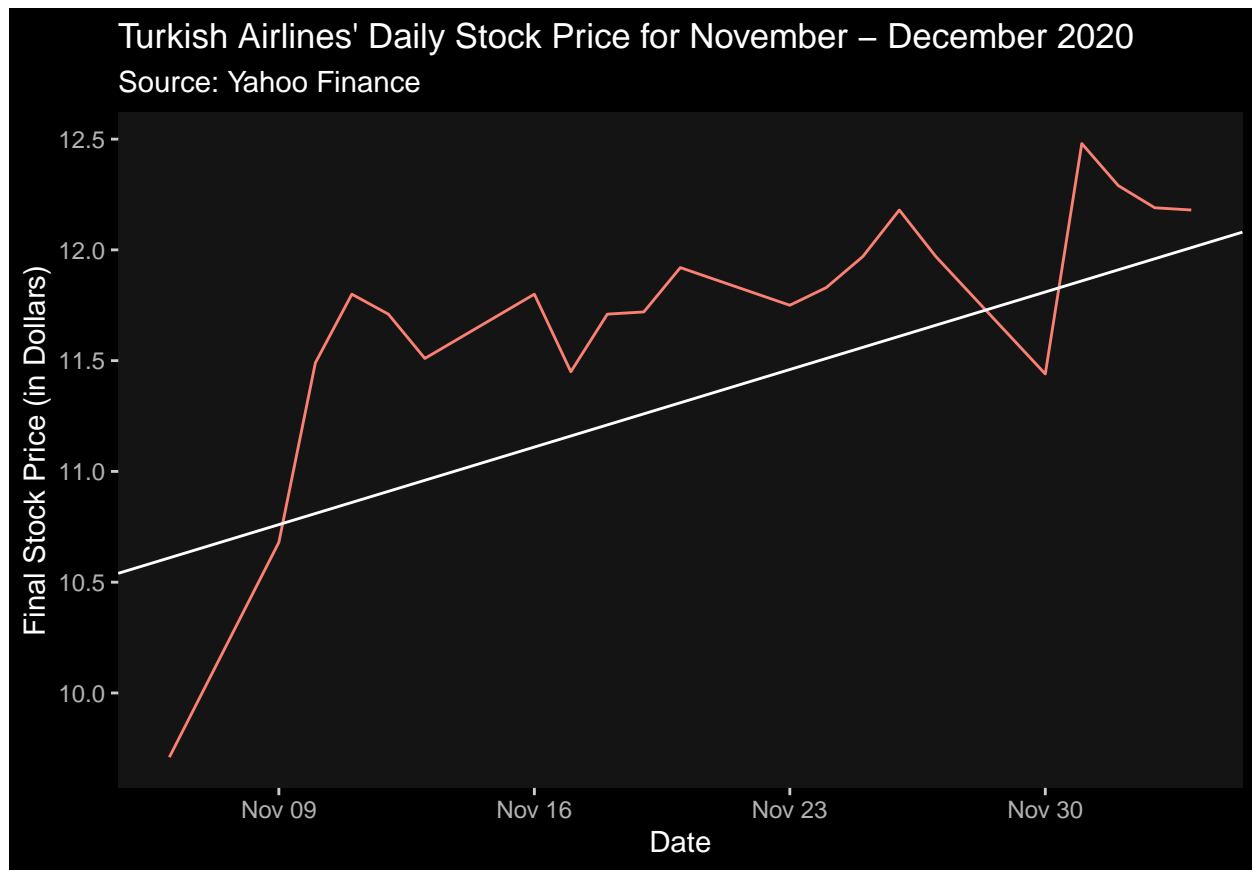
```
linearmodel = lm(Close~Date, data = Stock_predict_2020)

linearmodel

##
## Call:
## lm(formula = Close ~ Date, data = Stock_predict_2020)
##
## Coefficients:
## (Intercept)         Date
##  -917.98979      0.05002
```

```
(price_thy_end = (0.05*27 + 12.18))
```

```
## [1] 13.53
```

Having pulled the growth rate of the stock price, I computed the potential stock price for the end of 2020 directly. I used the slope and the new intercept to obtain that the final stock for THY as of December 31, 2020 will be \$13.53. This is an example of an extremely simple linear model and I don't expect that I will be right.

EDIT as of 5/13/2021: the actual closing price of THY for December 31, 2020 was \$13.03. I was a little too high, but December 2020 saw less growth than I was expecting at the time.

Works Cited:

Dailymail.com, Paul Chavez For. "Ben Affleck and Jesse Eisenberg Star in Turkish Airlines Ad Promoting Flights to Gotham." Daily Mail Online, Associated Newspapers, 8 Feb. 2016, www.dailymail.co.uk/tvshowbiz/article-3436824/Ben-Affleck-Jesse-Eisenberg-star-Batman-Lex-Luthor-Turkish-Airlines-ad-promoting-flights-Gotham.html.

"Istanbul Ataturk Airport Attack: 41 Dead and More than 230 Hurt." BBC News, BBC, 29 June 2016, www.bbc.com/news/world-europe-36658187.

"Super Bowl History." Pro, www.pro-football-reference.com/super-bowl/.

"A Thrilling Journey Awaits You." THE JOURNEY | Turkish Airlines ®, www.turkishairlines.com/en-int/thejourney/.

"TURK HAVA YOLLARI (THYAO.IS) Stock Price, News, Quote & History." Yahoo! Finance, Yahoo!, 26 Nov. 2020, finance.yahoo.com/quote/THYAO.IS?p=THYAO.IS.

"Turkish Airlines Sponsored the Super Bowl & Not Everyone Was Happy about It." KARRYON, 6 Feb. 2017, karryon.com.au/industry-news/airline/turkish-airlines-sponsored-the-us-super-bowl-not-everyone-was-happy-about/.

"Universal and Turkish Airlines Are Betting Big on Super Bowl Ads." Skift, 1 Feb. 2018, skift.com/2018/02/01/universal-and-turkish-airlines-are-betting-big-on-super-bowl-ads/.

Walsh, S.M. "WATCH: Turkish Airlines' Super Bowl 2020 Commercial." Heavy.com, 2 Feb. 2020, heavy.com/entertainment/2020/02/turkish-airlines-super-bowl-2020-commercial/.