TSLA Stock Prediction Model

Braden Hanna

March 2021

1 Introduction

Tesla is a new Goliath in the stock market. According to Google Finance, Tesla stock has gone from 72.24 dollars per share in March 2020, to 883.09 dollars per share at it's high in January 2021 [1]. This project will attempt to predict the TSLA stock's high price based on TSLA stock's opening price and NDAQ stock opening price(NASDAQ stock exchange where TSLA is sold), using a multiple linear regression model.

2 R Script

```
#Braden Hanna
#3/14/21
#Computation Statistics CS 467

library(ggplot2)

# Load the data from csv files to dataframe
tsla <- read.csv("TSLA.csv")
ndaq <- read.csv("NDAQ.csv")

# Remove no trading data i.e volume == 0
tsla <- subset(tsla, Volume > 0)
ndaq <- subset(ndaq, Volume > 0)

# Remove NAN
tsla <- na.omit(tsla)
ndaq <- na.omit(ndaq)

# Rename columns before joining the data
names(tsla)[names(tsla)=="High"] <- "TSLAHigh"
names(tsla)[names(tsla)=="Open"] <- "TSLAOpen"
names(ndaq)[names(ndaq)=="Open"] <- "NDAQOpen"
```

```
# Select the columns to join
tsla <- subset(tsla, select=c("Date", "TSLAOpen", "TSLAHigh"))
ndaq <- subset(ndaq, select=c("Date", "NDAQOpen"))

# Join the data
data <- merge(tsla, ndaq, by="Date")

# Multiple linear regression
model <- lm(TSLAHigh~TSLAOpen+NDAQOpen, data=data)
print(summary(model))
print(confint(model))

# Plot the regression model with smooth function
data$TSLAHighPredicted<-predict(model)
scatter <- ggplot(data, aes(TSLAHighPredicted, TSLAHigh))
scatter + geom_point() + geom_smooth(method = "lm", colour = "Red")
                + labs(x = "TSLA High Predicted", y = "TSLA High")
```

3 Graphs and Analysis

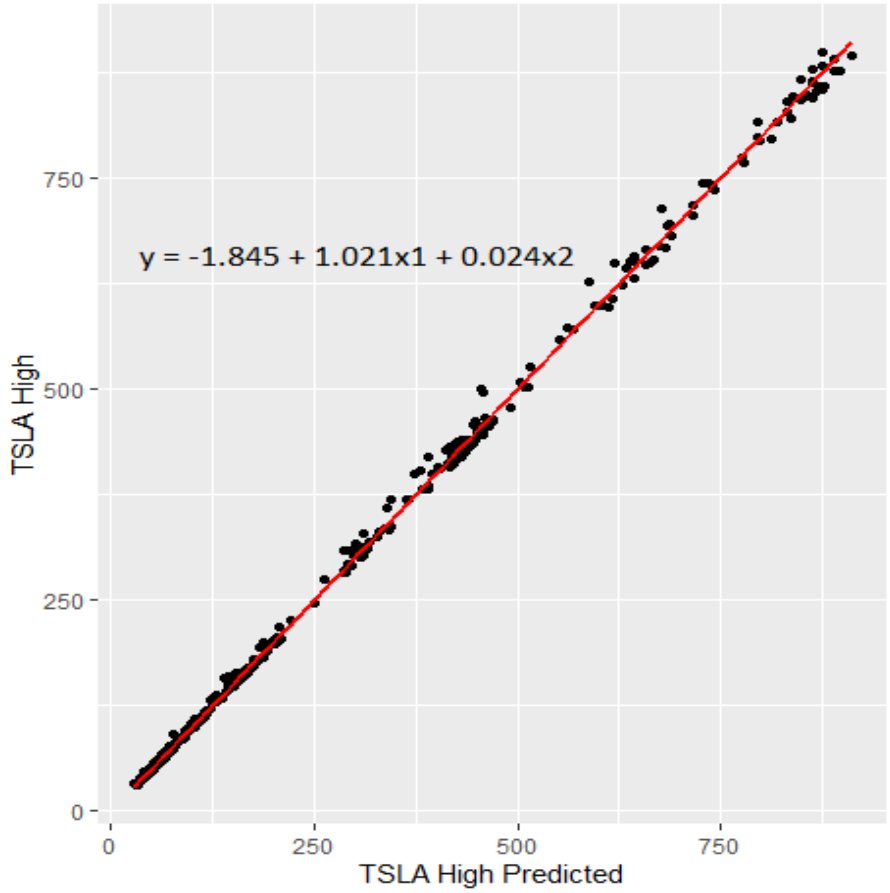

Figure 1: TSLA stock's high price from 2015-2021 vs TSLA stock's high price predicted with multiple linear regression using TSLA open price and NDAQ open price as predictor variables.

The linear regression equation for this model was found to be y = -1.845 + 1.021x1 + 0.024x2, where x1 = TSLA opening price and x2 = NDAQ opening price. The 2.5 percent confidence interval for the intercept was -2.781 and the 97.5 percent confidence interval for the intercept was -0.908. The 2.5 percent confidence interval for the TSLA opening price was 1.019 and the 97.5 percent confidence interval for the TSLA opening price was 1.023. The 2.5 percent confidence interval for the NDAQ opening price was 0.011 and the 97.5 percent confidence interval for the NDAQ opening price was 0.037. The R2 value was 0.999 and the adjusted R2 value was 0.999. The t-value of the TSLA opening price was 1026.834 and the p-value was 2e-16. The t-value of the NDAQ opening price was 3.719 and the p-value of the NDAQ opening price was 2.07e-4.

4 Conclusion

Because of the low p-values(almost 0) of the predictor variables, it can be concluded that TSLA opening price and NDAQ opening price are significant predictors of TSLA high price. The adjusted R2 value was 0.999, suggesting that this is an excellent model for predicting TSLA stock's high price.

References

[1] Google FInance. Tesla inc on nasdaq, March 2021.