

DSC Group 4  
Milestone 3  
2/23/22

First and foremost for our statistical analysis we wanted to centralize our focus around Goog vs Googl. As noted in milestone 2, these two ticker symbols represent shares in the same company, but the classification of the shares varies between Class C and Class A shares. The difference between Goog and Googl is that Goog does not provide shareholders a right to vote while Googl does. The expectation throughout our review was that Goog would always trade at a slightly discounted price because of that lack of votability. Using the entirety of the data's history, that conclusion did ring true:

```
> summary(googl)
      Date      Symbol      Adj.close      Close      High      Low      Open      Volume
Length:5046 Length:5046 Min. : 279.8 Min. : 279.8 Min. : 282.5 Min. : 278.5 Min. : 280.5 Min. : 465600
Class :character Class :character 1st Qu.: 551.2 1st Qu.: 551.2 1st Qu.: 556.7 1st Qu.: 546.6 1st Qu.: 551.7 1st Qu.: 1386600
Mode :character Mode :character Median : 835.1 Median : 835.1 Median : 839.0 Median : 829.0 Median : 833.0 Median : 1818700
Mean :1015.1 Mean :1015.1 Mean :1024.2 Mean :1005.2 Mean :1014.9 Mean :2442886
3rd Qu.:1211.0 3rd Qu.:1211.0 3rd Qu.:1220.8 3rd Qu.:1203.2 3rd Qu.:1211.1 3rd Qu.: 3005600
Max. :2996.8 Max. :2996.8 Max. :3030.9 Max. :2978.0 Max. :3025.0 Max. :24859915
NA's :2529 NA's :2529 NA's :2529 NA's :2529 NA's :2529 NA's :2529

S.P.Close
Length:5046
Class :character
Mode :character

> summary(goog)
      Date      Symbol      Adj.close      Close      High      Low      Open      Volume
Length:2523 Length:2523 Min. : 278.5 Min. : 278.5 Min. : 281.2 Min. : 277.2 Min. : 279.1 Min. : 7922
Class :character Class :character 1st Qu.: 539.8 1st Qu.: 539.8 1st Qu.: 543.8 1st Qu.: 535.7 1st Qu.: 539.6 1st Qu.: 1271000
Mode :character Mode :character Median : 813.7 Median : 813.7 Median : 816.7 Median : 805.1 Median : 811.7 Median : 1689526
Mean :1010.1 Mean :1010.1 Mean :1019.1 Mean :1000.4 Mean :1009.6 Mean :2310214
3rd Qu.:1206.0 3rd Qu.:1206.0 3rd Qu.:1216.2 3rd Qu.:1196.7 3rd Qu.:1205.0 3rd Qu.: 2805900
Max. :3014.2 Max. :3014.2 Max. :3042.0 Max. :2997.8 Max. :3037.3 Max. :24978074
NA's :6 NA's :6 NA's :6 NA's :6 NA's :6 NA's :6
```

As we can see from the data above, the mean of the googl adjusted close over the roughly 10 year sample was roughly \$5 or .047% higher than that of Goog. Meanwhile the median came in \$21.40 higher for googl than it did for goog which was a variance of roughly 2.5% higher for googl than goog. Based on the data one could infer that the amount of volume, daily, correlated directly to the lower price and lower median as goog had a daily volume, on average, 132,672 lower than that of googl, which equates out to a volume-by-the-day being roughly 5.4% higher for googl than goog.

```

Call:
lm(formula = DoubleClose ~ ., data = googlSelect)

Residuals:
    Min       1Q   Median       3Q      Max
-509.98  -97.70   -5.97   104.07   799.87

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.348e+03  8.570e+00  157.335 < 2e-16 ***
Adj Close    5.355e-01  4.085e-01   1.311  0.19001
Close        NA         NA         NA      NA
High         2.484e+00  4.366e-01   5.689  1.43e-08 ***
Low          -1.180e+00  4.330e-01  -2.725  0.00648 **
Open         -6.285e-01  4.048e-01  -1.553  0.12065
Volume       -4.565e-05  1.864e-06 -24.483 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 143.1 on 2511 degrees of freedom
(6 observations deleted due to missingness)
Multiple R-squared:  0.9711,    Adjusted R-squared:  0.9711
F-statistic: 1.69e+04 on 5 and 2511 DF,  p-value: < 2.2e-16

Warning: NAs introduced by coercion

Call:
lm(formula = DoubleClose ~ ., data = googlSelect)

Residuals:
    Min       1Q   Median       3Q      Max
-522.74  -95.80   -5.64   102.34   865.52

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.383e+03  8.394e+00  164.724 < 2e-16 ***
Adj Close    8.098e-01  4.172e-01   1.941  0.05235 .
Close        NA         NA         NA      NA
High         2.130e+00  4.383e-01   4.859  1.25e-06 ***
Low          -1.368e+00  4.390e-01  -3.115  0.00186 **
Open         -3.852e-01  4.089e-01  -0.942  0.34620
Volume       -4.901e-05  1.854e-06 -26.442 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 143.9 on 2511 degrees of freedom
(6 observations deleted due to missingness)
Multiple R-squared:  0.9708,    Adjusted R-squared:  0.9707
F-statistic: 1.669e+04 on 5 and 2511 DF,  p-value: < 2.2e-16

```

### ***(Top is Googl and bottom is Goog)***

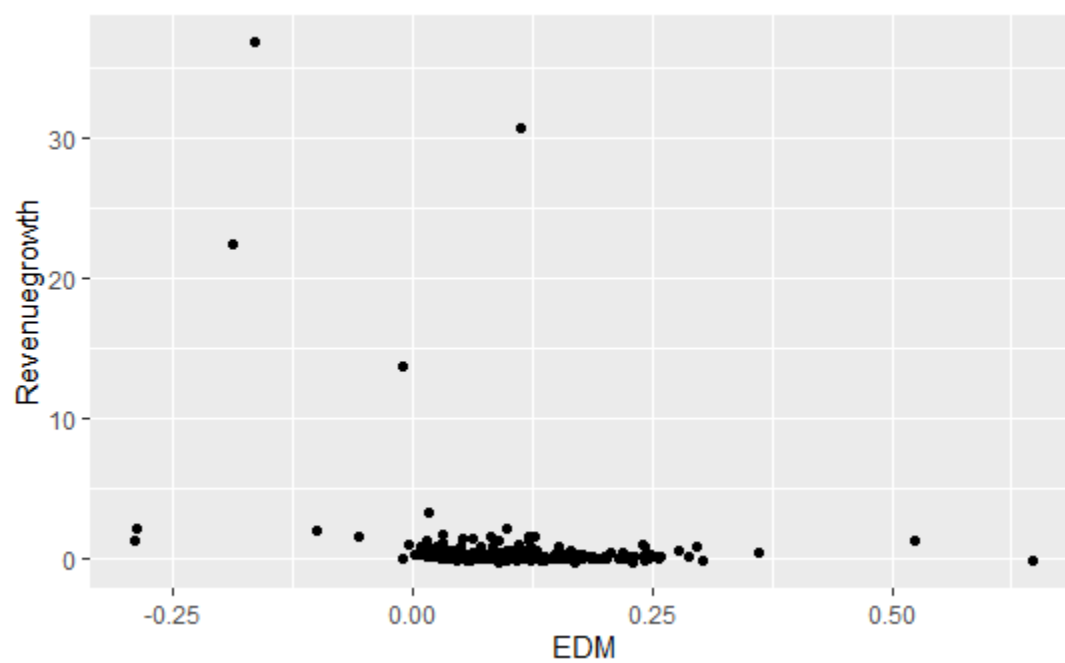
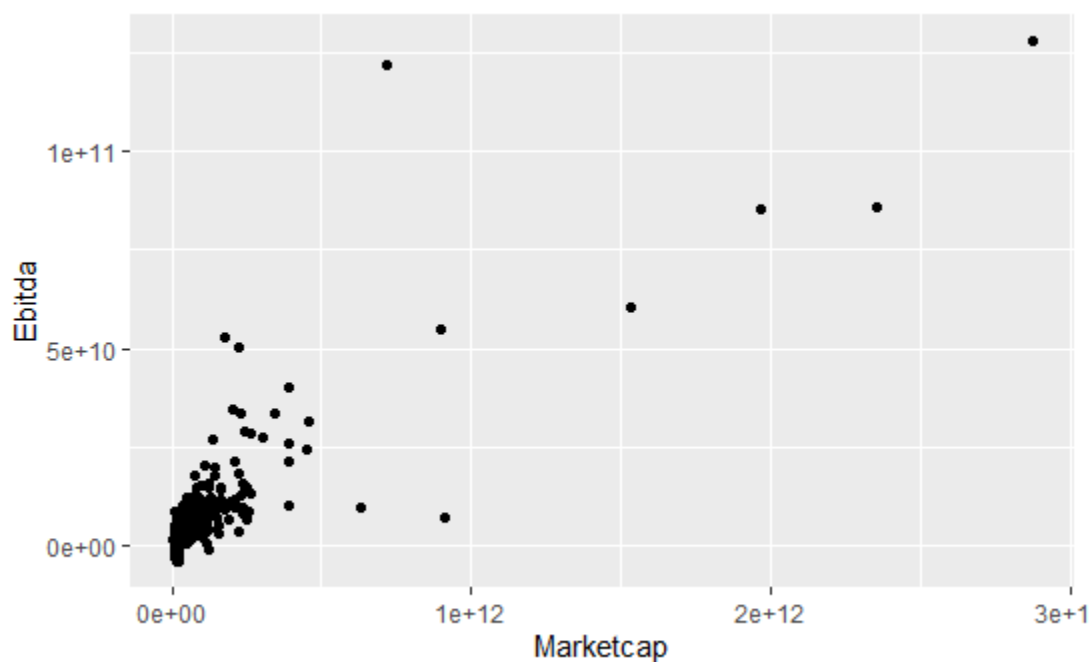
What was interesting about our findings is that although googl on average traded higher, as we can see from all our categories across the board above including the high, low and open, it wasn't true that googl closed or traded higher every single day. In fact, of the 2523 days that we analyzed, the adjusted close for Goog was higher than googl in 514 of those days (20.3%) . Of those 514 days, 450 of them have come since 1/9/2020. That means of 756 days within our data set since 1/9/2020, Goog has had a higher adj close than Googl in 59.5% of them. While some can be explained by a higher volume, as Goog has traded at a higher volume in 122 of those 450 days, the correlation isn't strong enough to conclude that. Another area of analysis is the impact full-time employees may have on the Goog stock. As the Class C shares (Goog) are the shares given as compensation to employees, it's possible that those shares are flipped with greater frequency than the Googl shares which may draw more long-term investors who desire voting rights and elongated company growth. Additionally, while this is speculative based on the

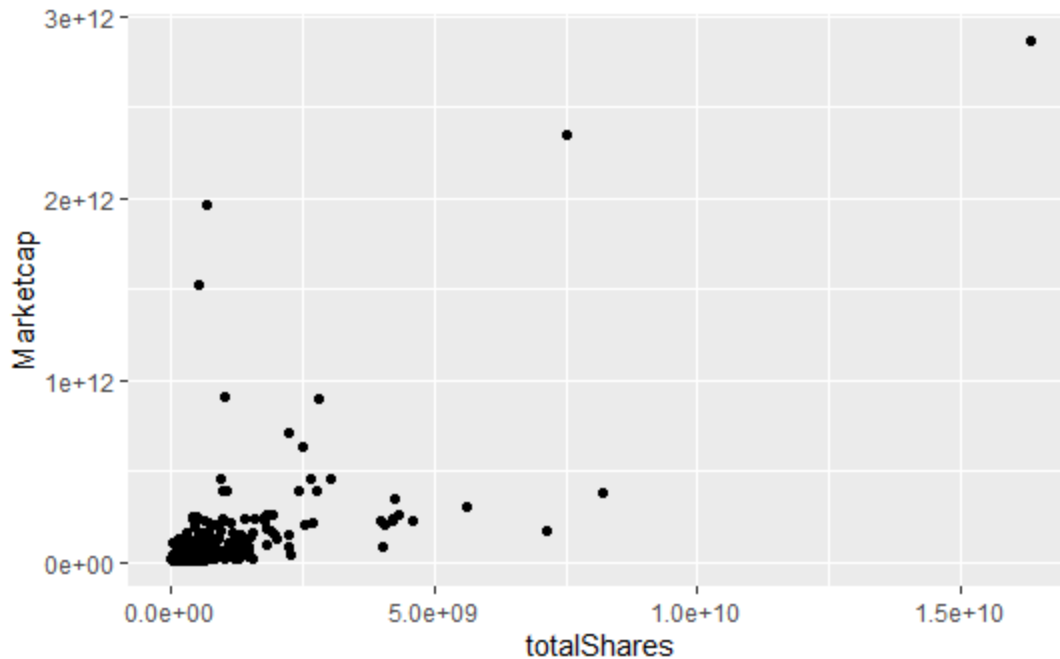
fact that employees receive Class C shares, it's possible that Google itself is buying back the Class C shares and not the Class A shares in order to escalate the Goog symbol price as an employee benefit. This is potentially supported by the data as we see our largest spikes in volume on Goog, and largest variance in volume between Goog and Googl, during the final weeks/first weeks of a new quarter. This could be anything from hiring season to buying back stock around earnings to limit the shares in circulation and inflate the share price (which may be why we see Goog higher than googl).

Our plan for further analysis on this topic is to further analyze an OLS model utilizing adjusted close of goog against googl while factoring in S&P and overall share volume by day with MLR. In order to assist with this process, and improve on its potential ease of analysis, we have created a calculated column within our data sets that takes the variance between the two adjusted close prices by subtracting the Googl\$AdjClose from the Goog\$AdjClose. This will help us look key in on a single variable. Additionally as a part of our data cleanup process we have separated the Goog and Googl symbol from the stocks dataset creating two distinct datasets that contain data only related to those two symbols. In addition, we paired the S&P close from the index data set to the Goog and Googl new-data-sets by their corresponding dates.

Lastly, the intent is to further analyze the correlation between the adjusted close differential and possibly even the day of the week and volume. Initial analysis and statistical looks on the day of the week telling a story shows us that since 2020, of the 122 days where the volume and close price has been higher for Goog, 32 of those days have fallen on a Friday, or roughly 26.2%. When we dig a bit deeper though we show no significant correlation between day of the week and adj close being higher for Goog as the distribution of the 514 days is nearly evenly distributed amongst each day of the week. When we dug deeper on volume in general for Google, we realized that the volume being higher on Friday wasn't driven by Goog closing higher, but was a generalized finding for both shares. Of the top 100 volume days for Goog and Googl since 2012, 38 of them fell on a Friday and higher volume directly correlated to Friday as

volume was higher on average on Fridays than any other day of the week. In general, the volume on Friday for both symbols was 13% lower on Thursday, 18% lower on Wednesday, 16% lower on Tuesday, and 16% lower on Monday.





One interesting finding we will be investigating further with given statistical techniques is shown below. We show four significant outliers pertaining revenue growth and total shares; of which a company who has fewer shares in circulation is experiencing substantial revenue growth. Additionally full-time employees as a stagnant variable to overall company growth shows an interesting pattern with 5 companies following outside the statistical pattern exhibited by the other companies. Looking at these two variables and components via OLS and additional techniques has led us to further investigate this relationship to determine any predictive qualities that may exist.

One additional relationship and regression we have taken a deep dive into pertains to the EBITDA in relation to a company's market cap. One thing that has stood out is that while there are companies that exude similar ratios, being 1% or less EBITDA in relation to said market cap, no other companies in the top 20 of total market cap have an EBITDA ratio as low as Tesla has. In order to really analyze this data via regression techniques we were forced to remove banks from the equation as they did not have a reported EBITDA which was skewing the data downward given the zero dollar values in relation to their market caps. There were

other companies that had similar ratios, and some were even negative such as the airplane industries who had negative cashflows as shown by their EBITDA, but for a company with the growth rate and the market cap of Tesla, ranking in the top 10 of the S&P companies, Tesla was a significant outlier when it came to listed earnings before interest, taxes and depreciation. While looking into this regression another question arose that requires some potential further investigation and that is does a lower share price promote faster growth due to a more accessible price for the casual investor? We noticed that some of the higher growth stocks represented stocks that came in at a more reasonable price in relation to their total market cap.

