An analysis of Trump's Tweets: What Drives Engagement?

Maximus Ong, Valentino Aceves, Joseph Crosby, Wyatt Stone
(maxong43@g.ucla.edu, josephcc713@gmail.com, oaceves@ucla.edu, wyattstone77@g.ucla.edu)

UCLA Department of Statistics

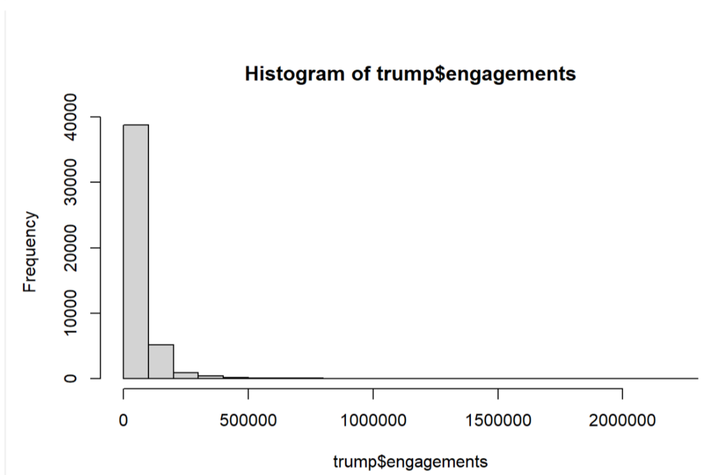
Objectives & Hypothesis

- Overall Motivation: To examine the relationship between tweet sentiment and engagement
- If there is a relationship: To examine the relationship between specific emotions and engagement
- If there is NOT a relationship: To examine if sentiment can be used to predict whether a tweet will have high/low engagement
- Hypothesis: Joy and Trust will have less correlation to Engagement than the other emotions

Exploratory Data Analysis

"Great", "Good", and "Thank" are all terms in the Joy Dictionary

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      0      72    1724   40764   67534 2278572
```

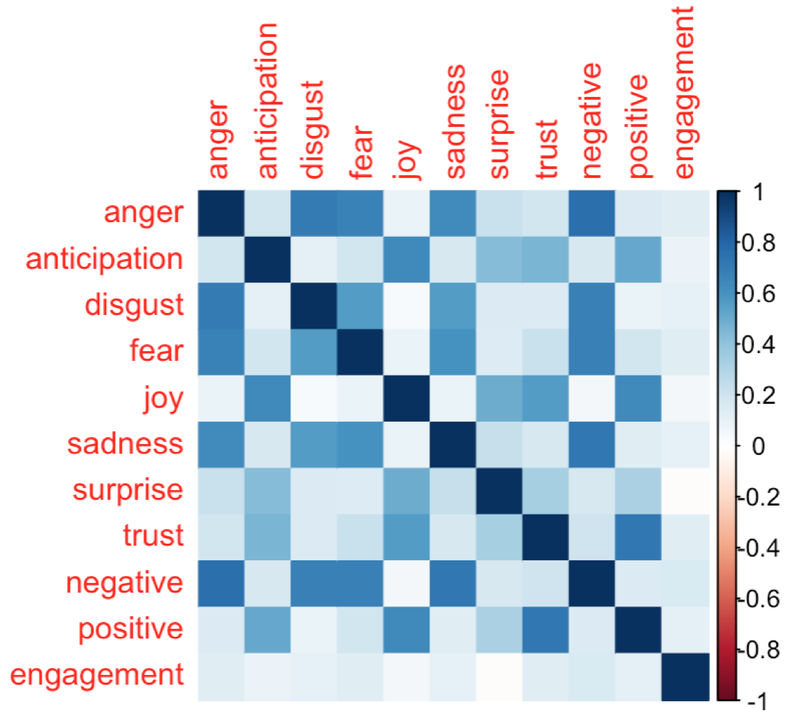


There are big outlier values in the amount of engagements. The maximum engagement value is the tweet where Trump announced his positive COVID-19 status.

Methods

- N = 45,389 tweets (after removing retweets)
- NRC Dictionary: Anger, Anticipation, Disgust, Fear, Joy, Sadness, Surprise, Trust, Positive, Negative
- Correlation Plot and Linear Model between 10 Sentiments and Engagement
- Random Forest predicting if a tweet will be High/Low Engagement based on Sentiments

Results




```
predictForest500
      above below
above  4268  3822
below  2118  6011
[1] 0.6337629
```

Discussion

- Correlation Plot: All Sentiments have Correlation < 0.2 with Engagement
- Linear Model: Joy is the only sentiment that is not significant with engagement

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    53128.8      814.0  65.265  < 2e-16 ***
anger           3139.4     1072.9   2.926 0.003435 **
anticipation    4522.0      847.0   5.339 9.44e-08 ***
disgust        -3841.9     1124.7  -3.416 0.000636 ***
fear            2579.4      913.3   2.824 0.004744 **
joy            -1858.8     1070.0  -1.737 0.082357 .
sadness        -2961.5     1033.9  -2.864 0.004182 **
surprise      -12396.3      976.9 -12.689  < 2e-16 ***
trust           8031.7      694.4  11.566  < 2e-16 ***
negative       10002.4      754.6  13.255  < 2e-16 ***
positive        1863.6      612.1   3.045 0.002331 **
```

- Random Forest: 63 Percent Success Rate in predicting whether a tweet will be High or Low Engagement based on its Sentiment

Conclusions

- Tweet Sentiment was NOT as good at predicting Engagement as we initially expected
- Joy was the emotion with the least significant relationship to Sentiment
- Trust performed better than we expected, and compared similarly to the other strong emotions
- Our Random Forest Prediction was moderately successful, as it will predict whether a tweet will be High or Low Engagement with 13% more success than a random guess
- Overall, the type of sentiment in a Trump Tweet is largely irrelevant for predicting engagement and only slightly more reliable than flipping a coin.

Future Considerations

- Our report did not consider a number of variables, so future analysis could investigate the impact of things like whether the tweet was deleted, flagged, or when it came out.
- For example, future investigation could look at the sentiments of deleted or flagged tweets, as well as the average engagement of flagged tweets.
- We also could consider if there was any relationship between the time a tweet was published and its sentiments or engagement.
- Lastly, we could have looked at specific words Trump used and see if those had any correlation with engagement.
- Our initial investigation is a good starting point that did yield interesting results, but there is always more to be analyzed.