

This document contains all non-visualization-based output from Project 1 for DS-4002 Group 5. These are all the figures we will use in our presentation. Other visualizations can be found in other OUTPUT folder locations. All these figures can be obtained by running the code in:

https://github.com/aqu7he/DS-4002-Group-5/blob/main/SCRIPTS/Project_1_Master_Code.ipynb

1. Original dataset, number of observations: 1000
2. VADER sentiment counts, full dataset before cleaning (Neutral = any sentiment score between (-0.05, 0.05)): [Neutral: 757, Positive: 172, Negative: 71]
3. Number of observations in cleaned dataset: 199

The following figures are from the “full” timeframe under observation (the past year)

4. Number of observations in cleaned ‘Before’ and ‘After’ datasets: [Before: 61, After: 138]
5. Mean Compound Sentiment Scores ‘Before’ Taylor: 0.1010344262295082
6. Mean Compound Sentiment Scores ‘After’ Taylor: 0.07020434782608696
7. Results of T-test for full timeframe: (statistic=-0.5567575722172928, pvalue=0.5783250097260251, df=197.0)

The following figures are from the “narrow” timeframe under observation (9/17-9/31)

8. Number of observations in cleaned ‘Before’ and ‘After’ datasets: [Before: 3, After: 2]
9. Results of T-test for full timeframe: (statistic=1.3358756736084882, pvalue=0.2738892008255432, df=3.0)