Vivek Sakariya
B00848519

Assignment 4

Objective: The primary objective is to use concepts and tools related to Semantic Analysis, Sentiment Analysis.

There are three files which transforms the data into desired format. They are named as *Sentiment_Analysis.py*, *Semantic_Analysis.py* and *cleantext.py*.

Sentiment_Analysis.py

This file is used to obtain the sentiment of the tweet based on the list of positive and negative words. According to the match of positive and negative words the polarity is decided. For example, if in a tweet there are more positive words than negative words, the polarity of tweet becomes positive and vice-versa for negative polarity. If a tweet does not contain either positive word or negative word, or there are equal positive and negative words, then the tweet's polarity is considered as neutral. This polarity cloud is represented in tableau.

Semantic_Analysis.py

This file is used to gain the frequency count of particular very and the article which has highest F/M Ratio to be printed on the console.

The zip file also contains 4 CSV files to get the output of the frequency count.

Output of Sentiment Analysis:

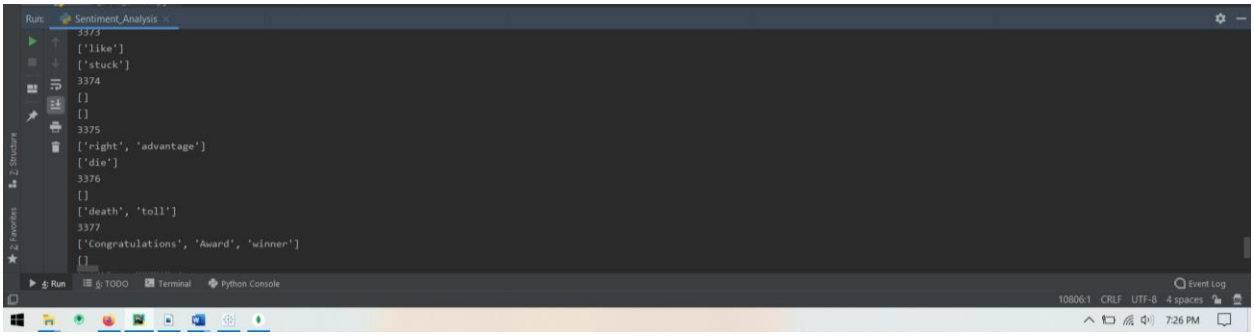

Figure 1: Sentiment Analysis

Vivek Sakariya
B00848519

Output of Tableau Visualization:

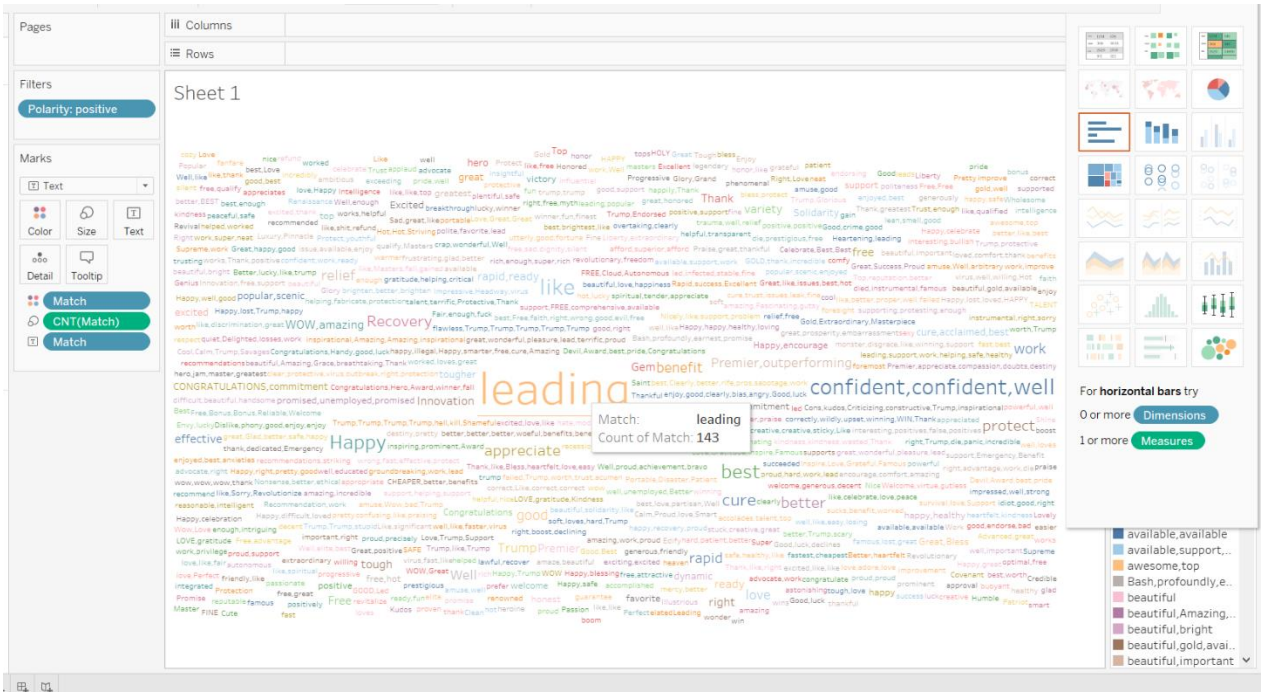

Figure 2: Positive Polarity Words

Figure 3: Negative Polarity Words

Vivek Sakariya
B00848519

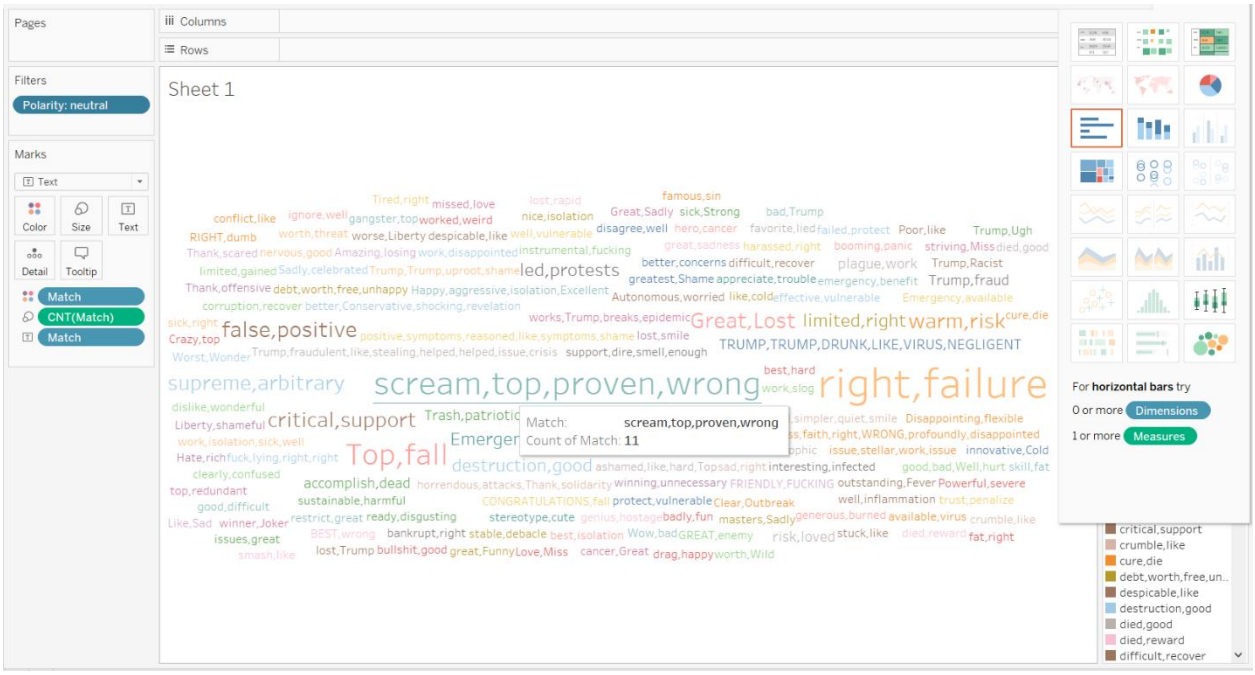

Figure 4: Neutral Polarity Words

The above figures 2,3 and 4 show the division of words based on a criterion of positive, negative and neutral polarity of a tweet.

Output of Semantic Analysis:

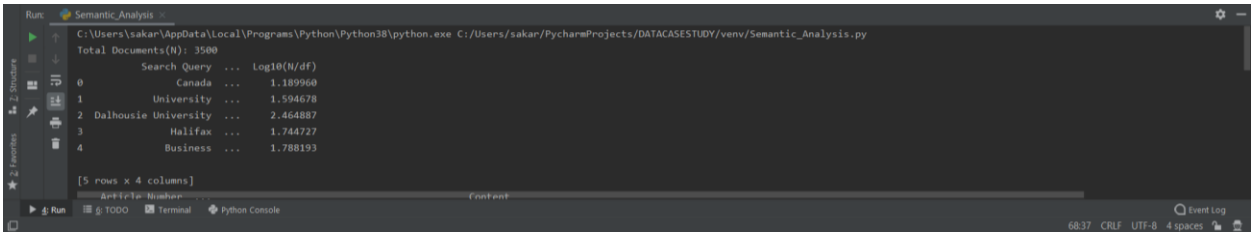

Figure 5: Dataframe of Total word and Frequency count

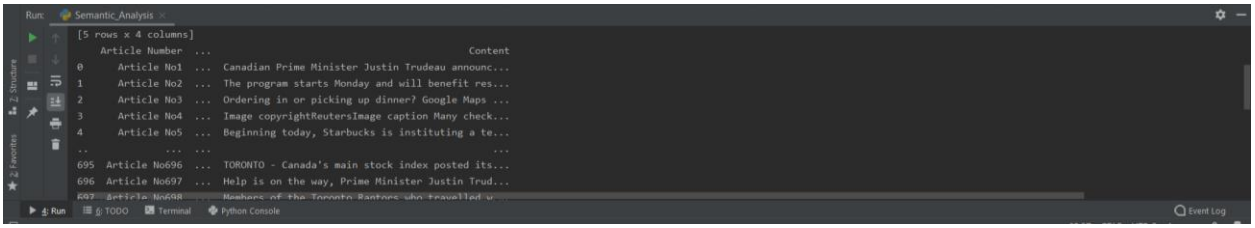

Figure 6: Dataframe showing F/M Ratio

Figure 7: Article having highest F/M ratio

The above figures 5,6 and 7 show the semantic analysis of the news articles and finding the F/M Ratio of a particular word in the news articles.

The zip file contains the screenshots of 700 News articles and 3600 tweets on which different operations are performed.

Vivek Sakariya
B00848519

References:

[1] “Add items to a list from a text file line by line in Python - CodeSpeedy”. [Online]. Available: https://www.codespeedy.com/add-items-to-a-list-from-a-text-file-in-python/. [Accessed: 13-Apr.-2020].

[2] “Extraction of Tweets using Tweepy - GeeksforGeeks”. [Online]. Available: https://www.geeksforgeeks.org/extraction-of-tweets-using-tweepy/. [Accessed: 13-Apr.-2020].

[3] “Extended Tweets — tweepy 3.8.0 documentation”. [Online]. Available: http://docs.tweepy.org/en/latest/extended_tweets.html. [Accessed: 13-Apr.-2020].

[4] “How to Create Pandas DataFrame in Python - Data to Fish”. [Online]. Available: https://datatofish.com/create-pandas-dataframe/. [Accessed: 13-Apr.-2020].

[5] “Appending to an empty DataFrame in Pandas? - Stack Overflow”. [Online]. Available: https://stackoverflow.com/questions/16597265/appending-to-an-empty-dataframe-in-pandas. [Accessed: 13-Apr.-2020].

[6] “How to Export Pandas DataFrame to a CSV File - Data to Fish”. [Online]. Available: https://datatofish.com/export-dataframe-to-csv/. [Accessed: 13-Apr.-2020].

[7] “How would you make a comma-separated string from a list of strings ...”. [Online]. Available: https://stackoverflow.com/questions/44778/how-would-you-make-a-comma-separated-string-from-a-list-of-strings. [Accessed: 13-Apr.-2020].