

Malachi Stoll
Andrew Piispanen
Professor Li
MIS3640-01

Assignment 3: Twitter Trends in The World, D.C. and #Trump

1. Project Overview

For the text analysis project, our group utilized Twitter's API (through Jupyter Notebook) to analyze tweets based on trends by locations, common words and hashtags. We also looked to see if American political tweets influenced global trends in society. To find this, we looked to see if global tweets coincided with the tweets made in the Washington D.C. area. Then, we also looked at the popular #Trump hashtag, and broke down the text of individual words and other hashtags associated with the #trump hashtag. With this project, we were utilizing our group's previous experiences working with API's as well as Malachi's skills in Jupyter Notebooks and a variety of other Python libraries not taught in class in order to maximize the efficiency in our code and in our analysis.

2. Implementation

To start, the code had to first access Twitter's API. We utilized the libraries of "Pickle" (to conceal API keys for privacy when sharing), "os" (for operating system searches, only for the "path.exists" function), "json" (to break down and to better work with the JSON data that the API returns), and "Twitter" (providing us easy functions to break down Twitter's API) in order to load the data.

The first step of the system is to provide the user with local, U.S.A, and global trends, and then analyze common overlaps between D.C. and global trends. Twitter's Python library provides the "trends()" function that receives information on trending topics based on the application's own analysis on the data, and returns the top trends of any given location. To perform this task, we had to find both the U.S and the local I.D. for Washington D.C. Although not provided in the code, the ID was found by querying the locations and finding the ID for the local area of "Washington D.C". Then, it was put into the query as a constant for searching. The code first starts by returning the trends in the world, in the US, and then in Washington DC. Then, to simplify information, we combined the three queries into one dictionary of three separate sets (world, us and dc), and return the intersection of the US and DC's trends.

Lastly, the code analyzed the hashtag "#trump" in two separate ways—by looking through the text of each tweet first associated with the hashtag and then looking at the top hastags that are also associated with the "#trump" hashtag. To do this, we pulled a collection of tweets containing the hashtag "#trump". This was done via a "search.tweets" function that Twitter provides, allowing us to automatically search the term and receive a series of results. The format of the result pulled each tweet's text, screen names and hashtags, so that we could further query the information. Then, the queried tweets were counted by frequency of words and hashtags, so that we could return the top individual words and individual hashtags that were found in these tweets.

3. Results

One of the most interesting trends that we found in our data was in our data was in the text analysis of the #trump tweets. Below is the list of individual words that were most common in the tweets:

[('#Trump', 35), ('RT', 27), ('to', 13), ('is', 11), ('the', 10), ('of', 8), ('and', 8), ('a', 7), ('&', 5), ('#MAGA', 5)]

And then the top 10 trending hashtags associated with #trump:

[('Trump', 38), ('MAGA', 5), ('DonaldTrump', 4), ('trump', 4), ('POTUS', 2), ('FLOTUS', 2), ('NorthKorea', 2), ('worldwar3', 2), ('WorldWarIII', 2), ('OSU', 2)]

Besides the #trump hashtag, there is another common similarity. The hashtag "MAGA" appears as both a common word and a common hashtag. Upon further research, "MAGA" could mean two different things, depending on who is making the tweet—"Make America Great Again", or "Morons Are Governing America". The hashtag, although predominantly used to boast Trump and his presidential motto—could also be used as political satire to bash his presidency. This demonstrates an instance where looking at top tweets may not always help to gain further insight on what the general public thinks, because it may not always have the connotations the tweeter intends on it for having.

Another interesting topic that came up in our text analysis was fairly concerning from a societal standpoint. The idea of "WorldWarIII" being thrown around in multiple text formats can show the general concern that the world currently has for an impending

all-out war. Although it is not a trend that is always stated in the media, the general concern for WW3 was clearly present. This helped show our group the type of insight that the Twitter API can give, for it can give us trends in society that may not be easily found through the news or through old methods of receiving public viewpoints (surveys, opinions, polls).

4. Reflection

In regards to the process, our approach was more exploratory than structured. The basic way we came across making the code was by putting together the API retriever and then toying with new and interesting ways to query the data. During the project, we discovered that transferring the codes from Jupyter Notebooks to Visual Basic Studio was the biggest problem that created significant malfunctions. However, the code worked well once we were both utilizing only the Python editor Jupyter. For testing and trials, we compiled the code and ran debugging altogether. Since the project did not have a set outcome, we started by making queries and eventually ended with the queries based on intersection of areas and then based on the topic of Trump. In the end, the majority of the code was made by Malachi, and the write up was made by Andrew. The project was a great opportunity for both of us to continue to work with Python and API's, as we both had experience and were interested in learning more about their usages.