

Twitter has become a prominent tool for political communication. During the 2016 elections twitter was used extensively by politicians to speak directly to their political opponents and the American people. All of the 100 US senators have an official verified twitter account

(<https://www.birdsonganalytics.com/blog/2017/02/21/full-list-of-us-senators-on-twitter-2017/>).

According to George Washington University all but 7 congresspersons have twitter accounts

(<https://gwu-libraries.github.io/sfm-ui/posts/2017-05-23-congress-seed-list>).

Many news companies and websites have been dedicated to analyzing political tweets. The BBC published 7 articles and 5 TV segments involving Donald Trump's tweets between June and September 2017. USA Today published an opinion article on July 10th, 2017 suggesting President Trump's twitter account may be against the law. TweetCongress.org is a website that provides information about all the tweets issued by members of congress. They also offer statistics about most followed and new followers for each congressperson. However, even with all of this monitoring and analysis of political tweets it is unclear if statements made *via* twitter are then translated into political action.

This capstone project will analyze President Trump's tweets. Some of the questions I will try to answer.

1. What political issues does he tweet about?
2. Do his tweets translate into official actions of any kind?

The Data:

The twitter data I will be using is from the website Trump Twitter Archive

(<http://www.trumptwitterarchive.com/about>). This website monitors President Trump's twitter feed in real time, since 1/27/17. It includes all of the metadata available, including things like time of tweet, retweets, and comments. I will also be using data provided by the Federal Register (federalregister.gov). The office of the federal register publishes and archives Presidential proclamations, executive orders, and other documents that the President determines has general applicability and legal. Using this website's API I will obtain all Determinations, Executive Orders, Memorandums, Notices, and Proclamations made by the current Executive Office of the President.

Methodology:

1. Clean the data
 - a. Twitter Data
 - i. Import the twitter data into a Pandas' dataframe
 - ii. Convert the time of the tweet to a datetime object
 - iii. Using python's NLP libraries, remove stop words and punctuation from the tweet's text
 - b. Federal Register data
 - i. Import Federal Register data into a Pandas' dataframe
 1. Columns are document type
 2. Index is datetime action/document occurred
 3. Cells are body of document

- ii. Use Python's NLP libraries to remove stop words and punctuation

2. EDA

- a. Tokenize twitter and document data with NLP libraries
- b. Determine 1-gram, 2-gram, and 3-gram phrases
- c. Create graphs of most used 1-gram, 2-gram, 3-gram phrases for both twitter and document data

3. Analysis

- a. Do topic analysis for each tweet and document
- b. Create dictionaries of time binned n-grams, and topics for both tweets and documents
- c. Compare dictionaries for similar values
 - i. Same words used
 - ii. Same topics
- d. Determine % of time that political topic tweets, had a corresponding topic in the document dataset
- e. Determine average time difference in tweeting and publish political action

4. Deliverables

- a. Description of how cleaning, EDA, and analysis was conducted along with the corresponding code
- b. Graphs, figures for the results and an explanation of what the results imply
- c. A conclusion that summarizes the analysis, results and explains the limitations of the analysis.