

# Proposal

---

The example from class where we plotted the sentiment of Donald trump's tweets on different platforms really interested me. Not because of the politics, but for the idea of being able to trace tweet sentiment in the first place.

I am proposing to create a shiny app that, given a twitter username, plots their tweet sentiment over time. This would require pulling data from the twitter website or API. I so far have been successful with scraping as the API approval process is extremely difficult, and I generated a full dataset for Donald trump along with the tweet VADER sentiment score using some python scripting. All visualization will be done with shiny and R. It just so happens that both the scraping library and sentiment analysis tools I found were written in python, so it made the most sense.

## Problem

Evaluating the state of the world and it's impact on people is incredibly difficult. One group of people may be affected differently than others, different people have different viewpoints, and sometimes there are changes in population that can be noticed but often not measured.

Because people today spend so much time and share so much information publically on the internet, tracing their general state of mind on a positive or negative scale is very feasible. This information can be used to track general sentiment for all kinds of things using keyword detection or geolocation. Even more interesting is the idea that someone could see their general state of mind through the sentiment of their social media usage and presence. What if the general sentiment of someone's social media feed could be graphed?

If that kind of information could be graphed for a specific person, mental health issues, life pressures, or difficult circumstances can be detected in a person. Maybe certain circumstances can be correlated to someone's mental state. Maybe people can track their opinions around certain keywords over time. This kind of information can be very useful information for a lot of differences.

## Data

The datasets are generated using a combination of the above libraries, described as below:

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is *specifically attuned to sentiments expressed in social media*. It is fully open-sourced under the [\[MIT License\]](#) (we sincerely appreciate all attributions and readily accept most contributions, but please don't hold us liable).

---

TwitterScraper

A simple script to scrape for Tweets using the Python package requests to retrieve the content and BeautifulSoup4 to parse the retrieved content.

One of the bigger disadvantages of the Search API is that you can only access Tweets written in the **past 7 days**. This is a major bottleneck for anyone looking for older past data to make a model from. With TwitterScraper there is no such limitation.

So far I've generated a dataset for trump without too much issue.

```
,Unnamed:
0,has_media,hashtags,img_urls,is_replied,is_reply_to,likes,links,parent_tweet_id,
replies,reply_to_users,retweets,screen_name,text,text_html,timestamp,timestamp_ep
ochs,tweet_id,tweet_url,user_id,username,video_url,sentiment
0,0,False,[],[],True,False,83329,[],,13045,[],13029,realDonaldTrump,Will be
starting The White House news conference at 5:30 P.M. Eastern., "<p
class=""TweetTextSize js-tweet-text tweet-text"" data-aria-label-part=""0""
lang=""en"">Will be starting The White House news conference at 5:30 P.M.
Eastern.</p>",2020-03-29 20:57:55,2020-03-29
20:57:55,1244368213133144064,/realDonaldTrump/status/1244368213133144065,25073877
,Donald J. Trump,,{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}"
1,1,False,[],[],True,False,67633,[],,5550,[],15692,realDonaldTrump,"So proud of
the @USACEHQ, @FEMA, and the Federal Government for the 2,900 bed hospital they
built in 4 days (way ahead of schedule) in the Javits Center for NY. We are now
moving the teams to join others so they can continue to build more
hospitals/beds. Keep up the GREAT WORK!", "<p class=""TweetTextSize js-tweet-text
tweet-text"" data-aria-label-part=""0"" lang=""en"">So proud of the <a
class=""twitter-atreply pretty-link js-nav"" data-mentioned-user-id=""161329769""
dir=""ltr"" href=""/USACEHQ""><s>@</s><b>USACEHQ</b></a>, <a class=""twitter-
atreply pretty-link js-nav"" data-mentioned-user-id=""16669075"" dir=""ltr""
href=""/fema""><s>@</s><b>FEMA</b></a>, and the Federal Government for the 2,900
bed hospital they built in 4 days (way ahead of schedule) in the Javits Center
for NY. We are now moving the teams to join others so they can continue to build
more hospitals/beds. Keep up the GREAT WORK!</p>",2020-03-29 20:06:35,2020-03-29
20:06:35,1244355295033331712,/realDonaldTrump/status/1244355295033331718,25073877
,Donald J. Trump,,{'neg': 0.0, 'neu': 0.821, 'pos': 0.179, 'compound': 0.8938}"
2,2,False,[],[],True,False,52207,[],,2802,[],9308,realDonaldTrump,Thank you very
much to Ken Langone for being a great American and for your wonderful comments on
@TeamCavuto., "<p class=""TweetTextSize js-tweet-text tweet-text"" data-aria-
label-part=""0"" lang=""en"">Thank you very much to Ken Langone for being a great
American and for your wonderful comments on <a class=""twitter-atreply pretty-
link js-nav"" data-mentioned-user-id=""223970563"" dir=""ltr""
href=""/TeamCavuto""><s>@</s><b>TeamCavuto</b></a>.</p>",2020-03-29
20:03:36,2020-03-29
20:03:36,1244354540486512640,/realDonaldTrump/status/1244354540486512640,25073877
,Donald J. Trump,,{'neg': 0.0, 'neu': 0.608, 'pos': 0.392, 'compound': 0.8834}"
```

```
3,3,False,[],[],True,False,505503,[],,87053,[],96483,realDonaldTrump,"I am a
great friend and admirer of the Queen & the United Kingdom. It was reported that
Harry and Meghan, who left the Kingdom, would reside permanently in Canada. Now
they have left Canada for the U.S. however, the U.S. will not pay for their
security protection. They must pay!", "<p class=\"\"TweetTextSize js-tweet-text
tweet-text\"\" data-aria-label-part=\"\"0\"\" lang=\"\"en\"\">I am a great friend and
admirer of the Queen & the United Kingdom. It was reported that Harry and
Meghan, who left the Kingdom, would reside permanently in Canada. Now they have
left Canada for the U.S. however, the U.S. will not pay for their security
protection. They must pay!</p>",2020-03-29 19:00:26,2020-03-29
19:00:26,1244338645198352384,/realDonaldTrump/status/1244338645198352386,25073877
,Donald J. Trump,,{"neg': 0.022, 'neu': 0.706, 'pos': 0.271, 'compound':
0.9381}"
```

## Outputs

### 1. General sentiment over time

1. Generate graphs of average sentiment of an account's tweets using the scraped tweet information and the text VADER sentiment. yaxis of 1 to -1, positive to negative sentiment
2. Set start and end date for the display
3. Optionally overlay confidence intervals

### 2. Specific sentiment

1. Limit tweets to certain keywords of times or day
2. Find keywords with high or low sentiment (without stopwords)