Data Science Project 3: Chicago CTA Tweets

Team: Kelly, Garima, Gaby, Joe, Amy  
  
**Chosen topic and rational:**  
We want to create a report of public CTA Sentiment

* CTA has made a number of improvements over the past decade
* We want to track sentiment as the improvements are made.
* The Twitter API will allows us to look at tweets that have occurred over the last 7 days

Link to datasets

* Twitter API
* Airline Twitter sentiment <https://www.figure-eight.com/data-for-everyone/>

Visuals

1. Bar chart or pie chart of total number of positive/negative/neutral statements
2. Word cloud
3. Plot of model and score
4. Horizontal bar chart of top 10 words and their frequency

Presentation Highlights

1. Goal and reasoning
2. Where the data came from
3. Models used, best fit
4. Discussion of use of what we produce
5. Discussion of what could explain results
   1. Construction
   2. Shootings/suicide
   3. Robbery

Tools:

* NLP - Scikit learn
* Textblob – may use for cleaing
* Pandas
* HTML
* Flask
* Tweepy
* Javascript
* Plotly.js
* vaderSentiment