DA 420: Project 7
Amy Hong

Part 1: Project Plan

For my final project, I would like to use text/sentiment analysis. Since presidential election is coming near, I would like to do text/sentiment analysis on twitter comments about top three or four candidates. This sentiment analysis will reveal the overall polarity of each candidate, such in way whether people have negative, neutral or positive sentiment toward each candidate. Also, I will find top 5 (or top 10) keywords, and if time permits, I would like to create words cloud for each candidate.

I will use methods shown in this website (https://hackernoon.com/text-processing-and-sentiment-analysis-of-twitter-data-22ff5e51e14c) to extract twitter comments. First, I have to create a twitter account and apply for data usage (I created my account and am waiting for twitter to grant my permission to use twitter data). Once I successfully get permission from Twitter for data usage, I will follow steps described in the website I provided above to extract and process text data. Here is my brief outline for my project:

1. Set topic for the project: Text/sentiment analysis on twitter comments of top three or four presidential candidates
2. Extracting twitter comments: I still need to decide how many comments I would like to collect and which candidates as keywords I would like to use. Method-wise, I created a twitter account and applied for the data usage. Upon permission, I would extract comments using method in the website I mentioned, which most likely using twitter's APIs and R packages
3. Processing twitter comments: I will use the textbook as one source and will do research on R blog postings for hints and tips. I also need to do research on which R packages are required for processing text.
4. Analysis: again, I will use the textbook as one main source and will do research on R blog postings for hits and tips. I also need to do research on which R packages are required for sentiment analysis.
5. Final step: put everything together, including introduction, methods, codes, outputs, explanation/conclusion, and reference

Here are some references I will use for creating my sentiment analysis model:
https://dataaspirant.com/2018/03/22/twitter-sentiment-analysis-using-r/
https://towardsdatascience.com/twitter-sentiment-analysis-and-visualization-using-r-22e1f70f6967
https://cran.r-project.org/web/packages/saotd/vignettes/saotd.html
https://github.com/Twitter-Sentiment-Analysis/R
https://www.earthdatascience.org/courses/earth-analytics/get-data-using-apis/sentiment-analysis-of-twitter-data-r/
https://www.r-bloggers.com/twitter-data-analysis-in-r/
https://www.r-bloggers.com/twitter-sentiment-analysis-with-r/
https://scholar.afit.edu/cgi/viewcontent.cgi?article=2853&context=etd (Scholarly source)