**COMPLEX DATA ANALYTICS**

**Making Word Cloud Using Python**

Collect textual data on a subject of your choice. You can collect Tweets by hashtags or keywords (e.g., “#Kavanaugh” or “Kavanaugh” or “midterm election”), Yelp reviews of prisons (or the book “The Little Prince”, or any product or service you are interested in), webpages or RSS feeds from your favorite news websites.

Make a word cloud visualization of the frequent words (or bigrams, or hashtags), and briefly discuss implications you draw from the output, as well as any other interesting analyses you wish to perform on the data in the future.

* You are not required to collect the data by Python, but you are strongly encouraged to do so to practice your Python skill.
* Minimum data entry requirement:

Tweets: at least 100 Tweets

Reviews: at least 100 textual reviews

Web pages: at least 5 webpages on the same topic (or similar topics)

* You need to read and process your data in Python, and make sure you remove stop words.
* Optional: color your words based on their sentiment

Deliverables:

1. A one-page word document including the following:
   1. Describe your data sources
   2. Present the visualization
   3. Briefly discuss any implications you draw from the output, as well as any other interesting analyses you wish to perform on the data in the future
2. Your Python script
3. Optional: Any output files you generated from the text data (word frequencies, bigram frequencies, etc.)

Have fun!