

## NLP Homework 2

Due Monday 10/12, 11:59 PM

In this assignment, you will continue to work on the texts that you've extracted in the first homework (in the field "text"). To gain some understanding of the sentiment in these texts, you will extract those sentences that contain adjective phrases or exclamation marks.

Once you extract these sentences, you will do a descriptive statistics about them, e.g., the average length of a sentence, top 50 adjective phrases (by frequency), top 50 adjective words, top 50 nouns or verbs, etc. You are encouraged to combine these findings with the other fields of the text (e.g., country, replies\_count, Facebook activities, etc.). You can use tables or figures to show these summaries.

Write a report to explain your processing and analysis process. Use Python screenshots to show your programming.

As always, please explain your process – both for cleaning and for analysis.

And, after you present the results (from the descriptive statistics and from the frequency analysis), please provide your interpretation of the results. And, what do you learn from these results?

The last thing to do is to envision how you would conduct a sentiment analysis on the review texts.

### **What to submit for Homework:**

Each student should submit a report that includes the grammars, the Python screenshots, and the written explanations.

### **How to Submit Homework:**

Go to the Blackboard system and the Assignment for Homework 2 and attach your report. Your submission should include:

- 1). Description of the cleaning and analysis processes (in report)
- 2). Results of the analysis and interpretation of the results, including tables/figures if any (in report)
- 3). Thoughts on how to conduct sentiment analysis (in report)
- 3). Python processing screenshot (in report)
- 4). Your Python code (submit in one separate folder zipped)