

515 HOMEWORK 3:

CUSTOMER SERVICE ANALYTICS

Your assignment is to create a tool that allows the user to assess the linguistic properties of major companies' customer service tweets. Your program should first ask the user to specify the type of analysis they would like to perform, and it should then ask the user which company they would like to analyze. Your program should allow the user to perform an unlimited number of analyses.

The dataset is available at https://dgoldberg.sdsu.edu/515/customer_service_tweets_full.json and contains approximately 1,000 tweets for each of twelve major companies' customer service Twitter handles, listed alphabetically below:

@amazonhelp	@AppleSupport	@Ask_Spectrum	@AskPlayStation
@comcastcares	@hulu_support	@SpotifyCares	@sprintcare
@TMobileHelp	@Uber_Support	@UPSHelp	@XboxSupport

The data is formatted as a JSON array. An example of the formatting is below:

```
[  {"Company": "@sprintcare", "Text": "I understand. I would  
like to assist you. We would need to get you into a private  
secured link to further assist."},  
  ...  
  {"Company": "@UPSHelp", "Text": "Hello, please click the  
link to let us know how we can assist you. Click the link  
to DM us with your tracking and phone number. ^E.W  
https://t.co/wKJHDXWGRQ"} ]
```

Your tool should support the following types of analyses:

- **Polarity:** your tool should calculate the average sentiment polarity across that company's customer service tweets. Ensure that you calculate sentiment scores per-tweet rather than per-sentence.
- **Subjectivity:** your tool should calculate the average sentiment subjectivity across that company's customer service tweets. Ensure that you calculate sentiment scores per-tweet rather than per-sentence.
- **Formality index:** this metric is an attempt to estimate how formally or informally text is written based upon its parts of speech ([find the part of speech tagging scheme here](#)). Scores close to 0 reflect very informal text, and scores close to 100 reflect very formal

text. Your tool should calculate the average formality across that company's customer service tweets. Formality is computed as follows:

f = number of nouns, adjectives, prepositions, and determiners
 c = number of pronouns, verbs, adverbs, and interjections

$$50 \left(\frac{f - c}{f + c} + 1 \right)$$

For the purposes of this assignment, treat any tag that contains "NN" as a noun; any tag that contains "JJ" as an adjective; any tag that contains "IN" as a proposition; any tag that contains "DT" as a determiner; any tag that contains "PR" as a pronoun; any tag that contains "VB" as a verb; any tag that contains "RB" as an adverb; and any tag that contains "UH" as an interjection.

Write your code such that the entirety of your program is case insensitive (for example, the program should behave equivalently if the user enters "yes", "Yes", or "YES" or if they enter "@sprintcare" or "@SprintCare"). However, do not perform any spellchecking on this assignment (unfortunately, the spellchecker we have used corrects "polarity" to "popularity").

Some considerations as you write your code:

- Consider the possibility that, when loading the dataset, some connection issue occurs (that is, a status code other than 200). Ensure that your code handles this case and provides the user with a helpful printout if it does occur.
- Consider the possibility that the user enters an unsupported mode (something other than polarity, subjectivity, or formality). In this case, print out a short warning message and allow the user to continue analyzing data. It is also possible that the user will enter an unsupported Twitter handle, but you do not need to handle this case for this assignment.
- Ensure that your prompts and output are crisp, professional, and well-formatted. For example, ensure that you have used spaces appropriately and checked your spelling. Ensure that graphs are appropriately titled and that axes are appropriately labeled.
- Adding comments in your code is encouraged. You may decide how best to comment your code. At minimum, please use a comment at the start of your code to describe its basic functionality.

Please use the following as a template for the tool's expected functionality:

Welcome to the customer service analyzer!

Which analysis would you like to perform
(polarity/subjectivity/formality)? POLarity

Which Twitter handle would you like to analyze? @upshelp

@upshelp: -0.020923390993824426

Would you like to run another analysis (yes/no)? yes

Which analysis would you like to perform
(polarity/subjectivity/formality)? Subjectivity

Which Twitter handle would you like to analyze? @SPOTIFYCARES

@SPOTIFYCARES: 0.29245743199261076

Would you like to run another analysis (yes/no)? yes

Which analysis would you like to perform
(polarity/subjectivity/formality)? Formality

Which Twitter handle would you like to analyze? @xboxsupport
@xboxsupport: 59.37499613549362

Which analysis would you like to perform
(polarity/subjectivity/formality)? finances

Which Twitter handle would you like to analyze? @amazonhelp

Sorry, that type of analysis is not supported. Please try again.

Would you like to run another analysis (yes/no)? no