Student No: 190050834

Daniel K

Coding Project Proposal

Project Title: Most Frequently Used Adjective per Word List

Description:

The idea for this project is to create a program that analyses a text for nouns and checks to see if an adjective is used. They will be recorded, and the function will return either: a list of all modified nouns and the adjectives used to modify them in order of most frequent use along with the exact number if possible afterward, or a set of charts for each noun that shows its most frequent adjectives in bars, to instead illustrate the data, for example.

This will allow the user to analyse how topics are described or conveyed in particular contexts, such as news articles on certain topics, or even an excerpt of a story for more general descriptions of everyday things. Any text that is of interest to the user can be used.

Resources and Approach to the Project:

To incorporate any type of data visualisations and charts, I will need to make use of the numpy and matplotlib modules. Text data can be taken from whatever articles or books that may be desired. I will take text data from various news articles to test the program since they are the most convenient source of digital text to use.

As of now, I expect I will make use of separate functions to copy and store the text, analyse it for occurrences of nouns and any adjectives that precede it while creating a collection referring to each noun that will contain the adjectives used, and possibly another function to order the collection if needed, as well as a function to create charts. I could use dictionaries to store the data, each one referring to the noun, containing a key for each adjective used with the value being the number of times it has been used with that noun. Another approach could be to store the adjectives used in general as a set, and using another function to search the text again for these specific instances of modified nouns and counting them here instead of using the previous function.

(5) Maybe even a dictionary of dictionaries!

(3)

While developing the program, I will test each function by returning debug values to make sure variables are tracking and updating as expected. This will help me to figure out at which point of the code an error may be found. Using 'assert' lines also helps to meet this end.

(4) debug mode in IDLE?

Initial Plan

I plan to first note down pseudocode and the overall plan on paper to determine the variables I will need to use as well as the order I will need to complete certain functions to work together. Following this I will begin coding on Python IDLE. As of now I don't expect to need the use of other tools such as Markdown or Jupyter. I will make sure the first function works as intended – with all variables how I expect them to be by the end of execution as I planned in my pseudocode notes. I will repeat this development and testing process until the last function has been written before finalising by using various different texts to make sure the program accounts for all types of situations in the text.

Yes, this is a good proposal. A few q suggestions / queries

1) How will you recognise a word as a noun or an adjective? Are you proging to have a ready-made list you will supply to your app?

② Store your text in files (Seriously)
Then if you make changes to anything you can test/re-run with same data easily.

③ Your decisions about data structures (set vs dict) will have a big impact on your code and how re-usable/extensible it will be. My intuition is that a set of words (noun/adjective) is OK for simple things ~~but~~ like testing is one is present but associations usually need dictionaries or something of equivalent complexity at least.

④ If not already, become familiar with debug mode in IDLE.

⑤ Jupyter notebooks + Anaconda do make chatting simple. With IDLE you will have to install (PIP) the libraries yourself