GitHub Repo: https://github.com/jeremyglebe/Coursework-5290-Natural-Language-Processing

Wiki Page: https://github.com/jeremyglebe/Coursework-5290-Natural-Language-Processing/wiki/ICE-1

Code File: https://github.com/jeremyglebe/Coursework-5290-Natural-Language-Processing/blob/main/ICE-1/ice1-code.py

Questions:

1. Why we use stopwords? Why stopwords are not necessary for NLP frequency

distribution.

**We use stop words to remove words that are not very meaningful. Stop words are not necessary in our distribution because they tend to appear quite frequently, often more than actual key words, but do not mean a lot and serve mostly as grammatical structures. So we are much more interested in frequently occurring, meaningful words.**

2. Based on high frequency words what information you can extract from the graph?

**You can extract certain general descriptions of significant traits, activities, or identifying information about the subject. SpaceX had been the first to do many things, based on the occurrences of the words “original” and “first”. I would extrapolate that they perform spacecraft launches based on frequency of words such as “space” and “launch”. I would additionally think that they must have made a significant innovation in the months of May and March, as both occur frequently.**

3. Can you provide different visualization for frequency distribution? If yes, please

perform. If no, why?

**Yes, I could make a bar chart of frequently occurring words. I have done so at the end of my code.**