

# ICE-1

## Questions

### Coding Exercise:

Write a code to tokenize the text and grab contents from a webpage where you can find information about SpaceX. Use stopwords and use the above strategy to calculate the frequency. The graph should display the first 10 high distribution words on the webpage while ignoring others. If the frequency of the word is less the 5 times ignore those words as well.

## Questions

1. Why do we use stopwords? Why stopwords are not necessary for NLP frequency distribution.
2. Based on high-frequency words what information you can extract from the graph?
3. Can you provide different visualization for frequency distribution? If yes, please perform. If no, why?