**NLP Assignment 1: Due 21st January 2023**

1. **Use requests and beautiful soup to download a page from the internet** 
   1. Follow the example under <https://www.digitalocean.com/community/tutorials/how-to-work-with-web-data-using-requests-and-beautiful-soup-with-python-3>
2. **Once you have the texts from the webpage, perform some basic text analytics**
3. Split the text into sentences
4. Tokenize
5. Person POS tagging
6. Remove the stop words
7. Perform lemmatization
8. Plot a frequency distribution of the most important words in the document before and after lemmatization
   1. What are the top 5 most important words before and after lemmatization?
   2. What are the bottom 5 words before and after lemmatization?

You are free to use NLTK or Spacy.