Q2.1

Find the number of characters, words, and sentences for the mentioned documents.

Character counts for the mentioned documents

Roosevelt: 7571 Kennedy: 7618 Nixon: 9991

Words counts for the mentioned documents

Roosevelt: 1536 Kennedy: 1546 Nixon: 2028

Sentences counts for the mentioned documents

Roosevelt: 68 Kennedy: 52 Nixon: 69

Q2.2

Using the set of "English" stop words available in NLTK library along with the set of punctuation available in string library, the stop words were removed from all the mentioned documents.

Q2.3

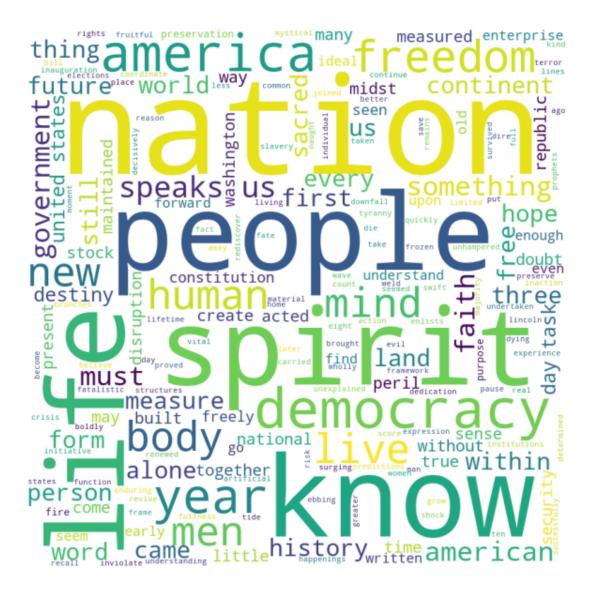
After removing the stop words, word occurs the most number of times in the inaugural address for each president are as follows (along with the number of times they appeared):

```
Top 3 words for Roosevelt: [('nation', 12), ('know', 10), ('spirit', 9)] Top 3 words for Kennedy: [('let', 16), ('us', 12), ('world', 8)] Top 3 words for Nixon: [('us', 26), ('let', 22), ('america', 21)]
```

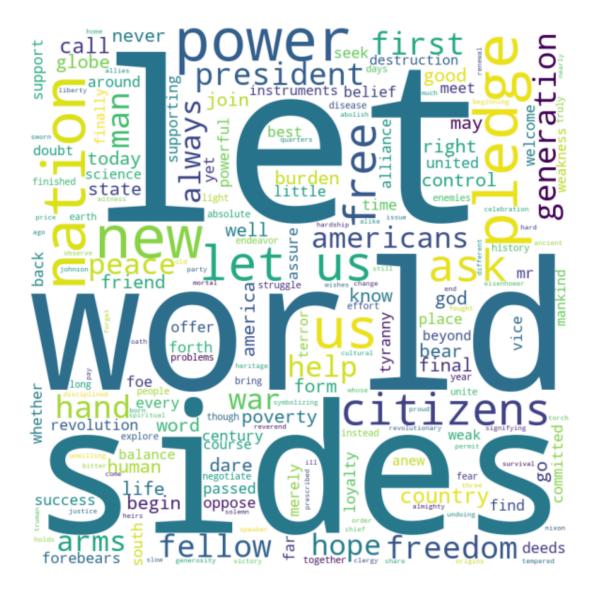
Q2.4

Plot the word cloud of each of the speeches of the variable. (after removing the stopwords)

For Roosevelt



For Kennedy



For Nixon

