NLP project

Contents

Problem Statement:..... 4

1.1 Find the number of characters, words, and sentences for the mentioned documents......4

1.2 Remove all the stopwords from all three speeches. 5

1.3 Which word occurs the most number of times in his inaugural address for each president? Mention the top three words. (after removing the stop words).....6

1.4 Plot the word cloud of each of the speeches of the variable. (after removing the stopwords).....7

TABLES

Table 1: Number of words, Characters and sentences……..5
Table 2: Cleaned speech Text for each president……..…6
Table 3: Number of words, Characters and sentences after cleaning the text……..…6
Table 4: Top 3 words……..……7

IMAGES

Image 1: Roosevelt's Speech……4
Image 2: Kennedy's Speech……4
Image 3: Nixon's Speech……..5
Image 4: Dataframes of Roosevelt, Kennedy and Nixon's speech, respectively……5
Image 5: Most common words before cleaning the text for stop words……6
Image 6: Roosevelt's Speech Word Cloud. ……7
Image 7: Kennedy's Speech Word Cloud……..8
Image 8: Nixon's Speech Word Cloud……9

Problem Statement:

In this particular project, we are going to work on the inaugural corpora from the nltk in Python. We will be looking at the following speeches of the Presidents of the United States of America:

President Franklin D. Roosevelt in 1941

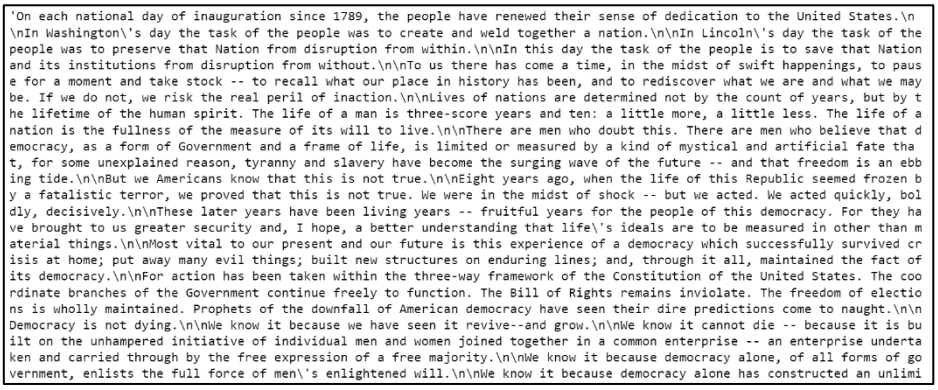
President John F. Kennedy in 1961

President Richard Nixon in 1973

(Hint: use .words(), .raw(), .sent() for extracting counts)

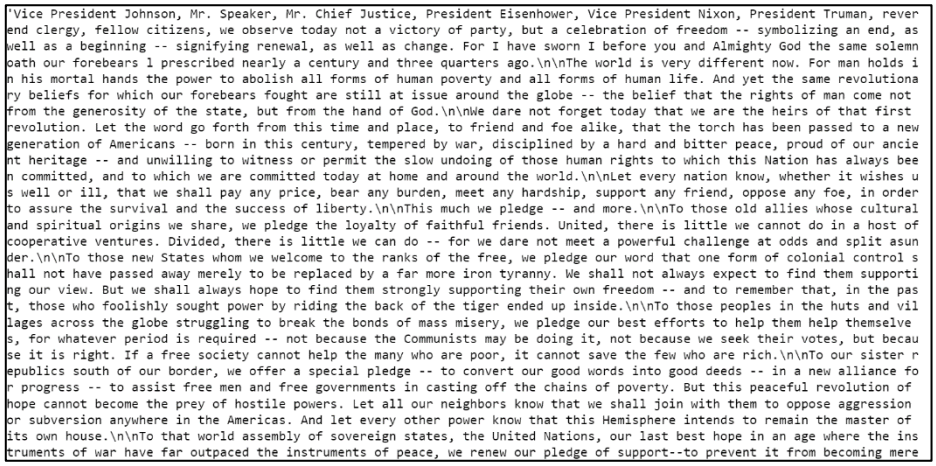
1.1 Find the number of characters, words, and sentences for the mentioned documents.

Before working on text analysis, 'nltk' library was installed in Jupyter Notebook. Then raw text file of of the speeches of the three president was downloaded from the 'inaugural' package. Following images show snippet of the three speeches.



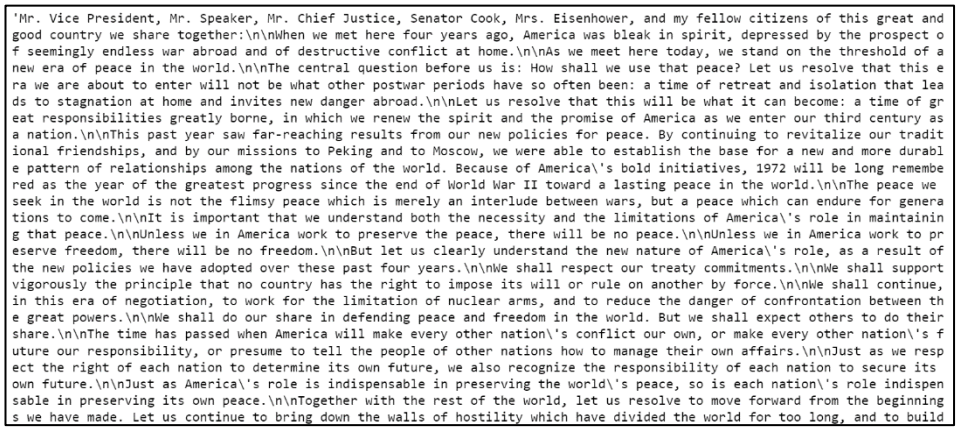
'On each national day of inauguration since 1789, the people have renewed their sense of dedication to the United States.\n
\nIn Washington\'s day the task of the people was to create and weld together a nation.\n\nIn Lincoln\'s day the task of the
people was to preserve that Nation from disruption from within.\n\nIn this day the task of the people is to save that Nation
and its institutions from disruption from without.\n\nTo us there has come a time, in the midst of swift happenings, to paus
e for a moment and take stock -- to recall what our place in history has been, and to rediscover what we are and what we may
be. If we do not, we risk the real peril of inaction.\n\nLives of nations are determined not by the count of years, but by t
he lifetime of the human spirit. The life of a man is three-score years and ten: a little more, a little less. The life of a
nation is the fullness of the measure of its will to live.\n\nThere are men who doubt this. There are men who believe that d
emocracy, as a form of Government and a frame of life, is limited or measured by a kind of mystical and artificial fate tha
t, for some unexplained reason, tyranny and slavery have become the surging wave of the future -- and that freedom is an ebb
ing tide.\n\nBut we Americans know that this is not true.\n\nEight years ago, when the life of this Republic seemed frozen b
y a fatalistic terror, we proved that this is not true. We were in the midst of shock -- but we acted. We acted quickly, bol
dly, decisively.\n\nThese later years have been living years -- fruitful years for the people of this democracy. For they ha
ve brought to us greater security and, I hope, a better understanding that life\'s ideals are to be measured in other than m
aterial things.\n\nMost vital to our present and our future is this experience of a democracy which successfully survived cr
isis at home; put away many evil things; built new structures on enduring lines; and, through it all, maintained the fact of
its democracy.\n\nFor action has been taken within the three-way framework of the Constitution of the United States. The coo
rdinate branches of the Government continue freely to function. The Bill of Rights remains inviolate. The freedom of electio
ns is wholly maintained. Prophets of the downfall of American democracy have seen their dire predictions come to naught.\n\n
Democracy is not dying.\n\nWe know it because we have seen it revive--and grow.\n\nWe know it cannot die -- because it is bu
ilt on the unhampered initiative of individual men and women joined together in a common enterprise -- an enterprise underta
ken and carried through by the free expression of a free majority.\n\nWe know it because democracy alone, of all forms of go
vernment, enlists the full force of men\'s enlightened will.\n\nWe know it because democracy alone has constructed an unlimi

Image 1: Roosevelt's Speech



'Vice President Johnson, Mr. Speaker, Mr. Chief Justice, President Eisenhower, Vice President Nixon, President Truman, rever
end clergy, fellow citizens, we observe today not a victory of party, but a celebration of freedom -- symbolizing an end, as
well as a beginning -- signifying renewal, as well as change. For I have sworn I before you and Almighty God the same solemn
oath our forebears l prescribed nearly a century and three quarters ago.\n\nThe world is very different now. For man holds i
n his mortal hands the power to abolish all forms of human poverty and all forms of human life. And yet the same revolutiona
ry beliefs for which our forebears fought are still at issue around the globe -- the belief that the rights of man come not
from the generosity of the state, but from the hand of God.\n\nWe dare not forget today that we are the heirs of that first
revolution. Let the word go forth from this time and place, to friend and foe alike, that the torch has been passed to a new
generation of Americans -- born in this century, tempered by war, disciplined by a hard and bitter peace, proud of our ancie
nt heritage -- and unwilling to witness or permit the slow undoing of those human rights to which this Nation has always bee
n committed, and to which we are committed today at home and around the world.\n\nLet every nation know, whether it wishes u
s well or ill, that we shall pay any price, bear any burden, meet any hardship, support any friend, oppose any foe, in order
to assure the survival and the success of liberty.\n\nThis much we pledge -- and more.\n\nTo those old allies whose cultural
and spiritual origins we share, we pledge the loyalty of faithful friends. United, there is little we cannot do in a host of
cooperative ventures. Divided, there is little we can do -- for we dare not meet a powerful challenge at odds and split asun
der.\n\nTo those new States whom we welcome to the ranks of the free, we pledge our word that one form of colonial control s
hall not have passed away merely to be replaced by a far more iron tyranny. We shall not always expect to find them supporti
ng our view. But we shall always hope to find them strongly supporting their own freedom -- and to remember that, in the pas
t, those who foolishly sought power by riding the back of the tiger ended up inside.\n\nTo those peoples in the huts and vil
lages across the globe struggling to break the bonds of mass misery, we pledge our best efforts to help them help themselve
s, for whatever period is required -- not because the Communists may be doing it, not because we seek their votes, but becau
se it is right. If a free society cannot help the many who are poor, it cannot save the few who are rich.\n\nTo our sister r
epublics south of our border, we offer a special pledge -- to convert our good words into good deeds -- in a new alliance fo
r progress -- to assist free men and free governments in casting off the chains of poverty. But this peaceful revolution of
hope cannot become the prey of hostile powers. Let all our neighbors know that we shall join with them to oppose aggression
or subversion anywhere in the Americas. And let every other power know that this Hemisphere intends to remain the master of
its own house.\n\nTo that world assembly of sovereign states, the United Nations, our last best hope in an age where the ins
truments of war have far outpaced the instruments of peace, we renew our pledge of support--to prevent it from becoming mere

Image 2: Kennedy's Speech



'Mr. Vice President, Mr. Speaker, Mr. Chief Justice, Senator Cook, Mrs. Eisenhower, and my fellow citizens of this great and good country we share together:\n\nWhen we met here four years ago, America was bleak in spirit, depressed by the prospect o f seemingly endless war abroad and of destructive conflict at home.\n\nAs we meet here today, we stand on the threshold of a new era of peace in the world.\n\nThe central question before us is: How shall we use that peace? Let us resolve that this e ra we are about to enter will not be what other postwar periods have so often been: a time of retreat and isolation that lea ds to stagnation at home and invites new danger abroad.\n\nLet us resolve that this will be what it can become: a time of gr eat responsibilities greatly borne, in which we renew the spirit and the promise of America as we enter our third century as a nation.\n\nThis past year saw far-reaching results from our new policies for peace. By continuing to revitalize our tradit ional friendships, and by our missions to Peking and to Moscow, we were able to establish the base for a new and more durabl e pattern of relationships among the nations of the world. Because of America\'s bold initiatives, 1972 will be long remembe red as the year of the greatest progress since the end of World War II toward a lasting peace in the world.\n\nThe peace we seek in the world is not the flimsy peace which is merely an interlude between wars, but a peace which can endure for genera tions to come.\n\nIt is important that we understand both the necessity and the limitations of America\'s role in maintainin g that peace.\n\nUnless we in America work to preserve the peace, there will be no peace.\n\nUnless we in America work to pr eserve freedom, there will be no freedom.\n\nBut let us clearly understand the new nature of America\'s role, as a result of the new policies we have adopted over these past four years.\n\nWe shall respect our treaty commitments.\n\nWe shall support vigorously the principle that no country has the right to impose its will or rule on another by force.\n\nWe shall continue, in this era of negotiation, to work for the limitation of nuclear arms, and to reduce the danger of confrontation between th e great powers.\n\nWe shall do our share in defending peace and freedom in the world. But we shall expect others to do their share.\n\nThe time has passed when America will make every other nation\'s conflict our own, or make every other nation\'s f uture our responsibility, or presume to tell the people of other nations how to manage their own affairs.\n\nJust as we resp ect the right of each nation to determine its own future, we also recognize the responsibility of each nation to secure its own future.\n\nJust as America\'s role is indispensable in preserving the world\'s peace, so is each nation\'s role indispen sable in preserving its own peace.\n\nTogether with the rest of the world, let us resolve to move forward from the beginning s we have made. Let us continue to bring down the walls of hostility which have divided the world for too long, and to build

Image 3: Nixon's Speech

To get the number of characters, words and sentences, firstly all characters, words and sentences were extracted, using *.raw, .words* and *.sents* function and then they the length of those character, words and sentences were calculated. Te results are as follows:

President's Name	Number of characters	Number of words	Number of sentences
Roosevelt	7571	1536	68
Kennedy	7618	1546	52
Nixon	9991	2048	69

Table 1: Number of words, Characters and sentences

1.2 Remove all the stopwords from all three speeches.

Before removing the stop words and cleaning the text, we will convert the whole text into dataframe since it will be easier to implement python functions to clean data. To convert test into dataframes:

1. The text is split line by line to convert it into a list
2. Then it is stored in a new dataframe using pd.Dataframe function. Following are the dataframes that we get.

	speech
0	On each national day of inauguration since 178...
1	In Washington's day the task of the people was...
2	In Lincoln's day the task of the people was to...
3	In this day the task of the people is to save ...
4	To us there has come a time, in the midst of s...

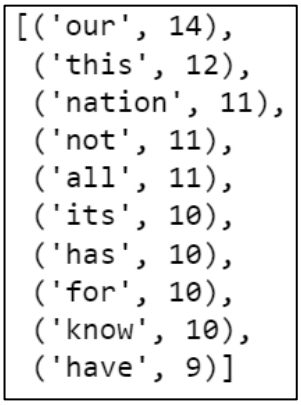
	speech
0	Vice President Johnson, Mr. Speaker, Mr. Chief...
1	The world is very different now. For man holds...
2	We dare not forget today that we are the heirs...
3	Let every nation know, whether it wishes us we...
4	This much we pledge -- and more.

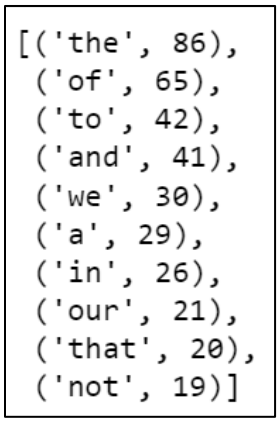
	speech
0	Mr. Vice President, Mr. Speaker, Mr. Chief Jus...
1	When we met here four years ago, America was b...
2	As we meet here today, we stand on the thresho...
3	The central question before us is: How shall w...
4	Let us resolve that this will be what it can b...

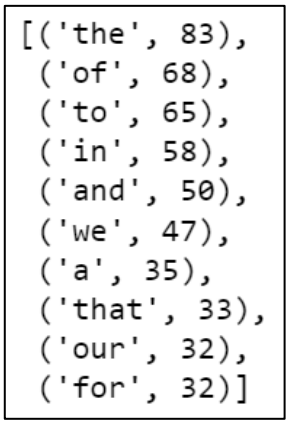
Image 4: Dataframes of Roosevelt, Kennedy and Nixon's speech, respectively

After converting the speech into dataframes, following steps are taken to clean the text data-

1. All the punctuation and exceptional strings or symbols are removed using 'replace' function.
2. Then all the words were converted into lower case. This will help in uniform recognition of same words.
3. The words are stemmed using Porter Stemmer. Stemming helps to convert all words into root words, helping to group the similar words to its origin, the root word.
4. Then we extract each word and make a series list out of it by first splitting the rows into words and then joining them with a space between them. This is stored in a new dataframe.
5. From here we can count to see the most common words.







Roosevelt's speech **Kennedy's Speech** **Nixon's Speech**

Image 5: Most common words before cleaning the text for stop words

6. To remove the stop words, 'stop word' corpus for English Language and punctuations are removed
7. Then frequency of cleaned text is counted again. The useless words, which do not contribute to the meaning of speech is removed again from the speech text and the words are recounted. The cleaned speech text looks as follows:

President	Cleaned Speech text
Roosevelt	'on each national day inauguration since 1789 people have renewed th se dedication united states'
Kennedy	'vice president johnson mr speaker mr chief justice president eisenhower vi ce president nixon president truman reverend clergy fellow citizens we obse rve today not a victory of party but a celebration of freedom symbolizing a n end as well as a beginning signifying renewal as well as change for i hav e sworn i before you and almighty god the same solemn oath our forebears l prescribed nearly a century and three quarters ago'
Nixon	'mr vice president mr speaker mr chief justice senator cook mrs eisenhower and my fellow citizens of this great and good country we share together'

Table 2: Cleaned speech Text for each president

President	Number of words BEFORE cleaning for stop words	Number of words AFTER cleaning for stop words
Roosevelt	1536	627
Kennedy	1546	677
Nixon	2028	811

Table 3: Number of words, Characters and sentences after cleaning the text

1.3 Which word occurs the most number of times in his inaugural address for each president? Mention the top three words. (after removing the stop words)

The top three words of each president after removing the stopwords are as follows:

President	Top 3 words
Roosevelt	nation:11 know: 10 spirit: 9
Kennedy	us: 12 world: 8 sides:8
Nixon	us: 26 peace: 19 world: 16

Table 4: Top 3 words

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

Word Cloud is a data visualisation technique used to represent test data, where the size of each words depends on its frequency, emphasizing its importance. To make a word cloud of cleaned test, 'wordcloud' is first installed using *!pip install wordcloud*. Following are the word cloud of each president's speech:

Image 6: Roosevelt's Speech Word Cloud

From the word cloud it seems that words in Roosevelt's speech are, 'Nation', 'Spirit', 'people', 'life' and 'democracy'.

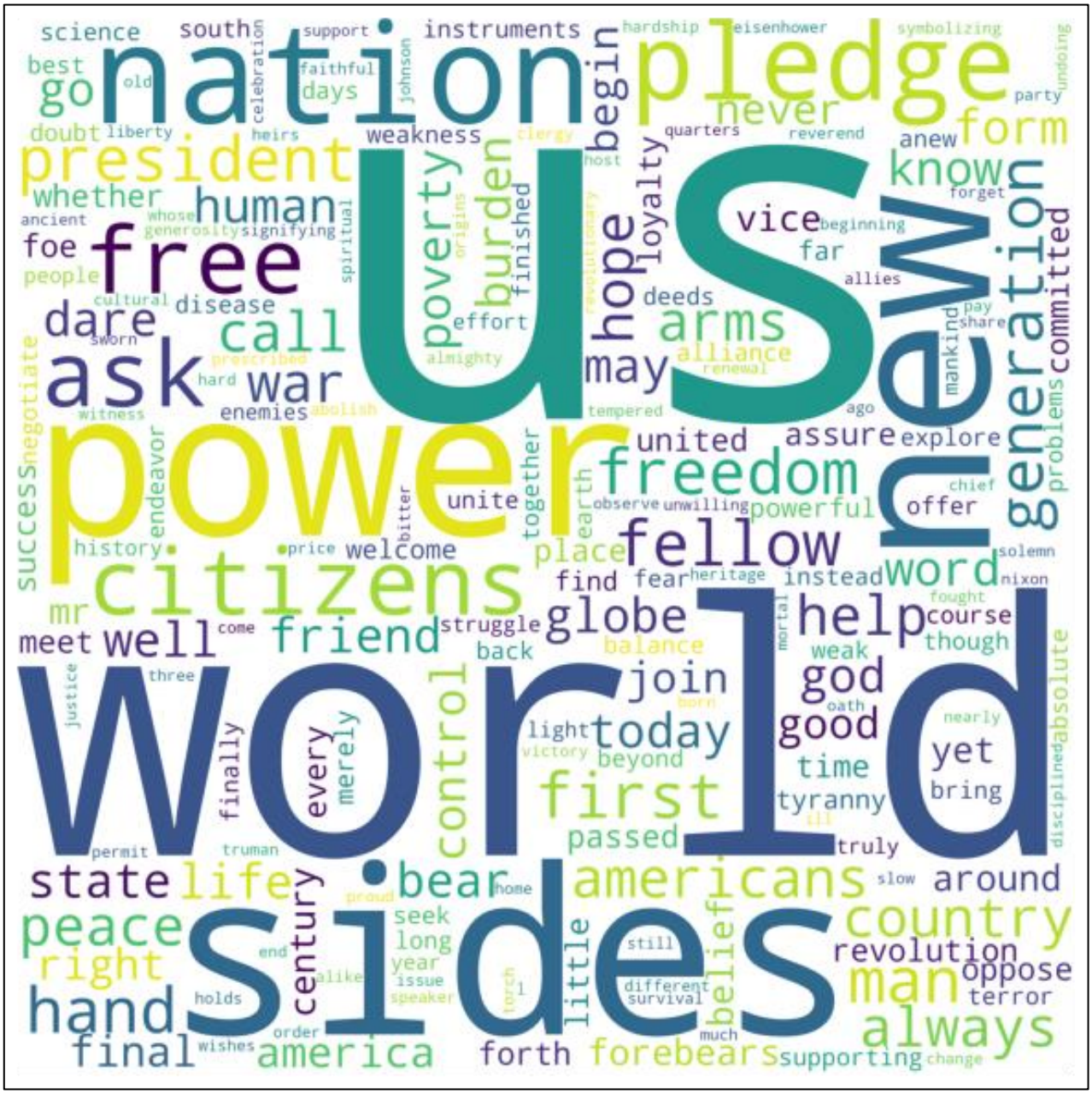


Image 7: Kennedy's Speech Word Cloud

From the word cloud it seems that words in Kennedy's speech are, 'us', 'world', 'new', 'power' and 'sides'.

Image 8: Nixon's Speech Word Cloud

From the word cloud it seems that words in Nixon's speech are, 'us', 'world', 'peace', 'nation' and 'new'.

Insights:

1. 'nation' seems to be the common popular word in all three presidents' speech
2. Freedom, power and peace seems to be one of the driving themes in the respective speeches of each president
3. Roosevelt's speech revolves around spirit of people of America, Kennedy's speech is about stand of America in the World and Nixon's speech is about responsibility of peace that America can bring to the world.