

```
In [1]: import numpy as np
import pandas as pd
import re
import string
import nltk
import json
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]: *#Importing the speeches from the corpus:*

```
nltk.download('inaugural')
from nltk.corpus import inaugural
inaugural.fileids()
inaugural.raw('1941-Roosevelt.txt')
inaugural.raw('1961-Kennedy.txt')
inaugural.raw('1973-Nixon.txt')
```

```
[nltk_data] Downloading package inaugural to
[nltk_data]      /Users/sabitapanchal/nltk_data...
[nltk_data]   Unzipping corpora/inaugural.zip.
```

Out[2]: 'Mr. Vice President, Mr. Speaker, Mr. Chief Justice, Senator Cook, Mrs. E isenhower, 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 spiri t, depressed by the prospect of seemingly endless war abroad and of destr uctive conflict at home.\n\nAs we meet here today, we stand on the thresh old 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 era we are abou t to enter will not be what other postwar periods have so often been: a t ime of retreat and isolation that leads to stagnation at home and invites new danger abroad.\n\nLet us resolve that this will be what it can becom e: a time of great responsibilities greatly borne, in which we renew the spirit and the promise of America as we enter our third century as a nati on.\n\nThis past year saw far-reaching results from our new policies for peace. By continuing to revitalize our traditional friendships, and by ou r missions to Peking and to Moscow, we were able to establish the base fo r a new and more durable pattern of relationships among the nations of th e world. Because of America\'s bold initiatives, 1972 will be long rememb ered as the year of the greatest progress since the end of World War II t oward 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 pea ce which can endure for generations to come.\n\nIt is important that we u nderstand both the necessity and the limitations of America\'s role in ma intaining that peace.\n\nUnless we in America work to preserve the peace, there will be no peace.\n\nUnless we in America work to preserve freedom, there will be no freedom.\n\nBut let us clearly understand the new nature of America\'s role, asa 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 i mpose 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 re duce the danger of confrontation between the great powers.\n\nWe shall do our share in defending peace and freedom in the world. But we shall expec t 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 fut ure our responsibility, or presume to tell the people of other nations ho w to manage their own affairs.\n\nJust as we respect the right of each na tion 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 indis pensable in preserving the world\'s peace, so is each nation\'s role indi spensable in preserving its own peace.\n\nTogether with the rest of the w orld, let us resolve to move forward from the beginnings we have made. Le t us continue to bring down the walls of hostility which have divided the world for too long, and to build in their place bridges of understanding

-- so that despite profound differences between systems of government, the people of the world can be friends.\n\nLet us build a structure of peace in the world in which the weak are as safe as the strong -- in which each respects the right of the other to live by a different system -- in which those who would influence others will do so by the strength of their ideas, and not by the force of their arms.\n\nLet us accept that high responsibility not as a burden, but gladly -- gladly because the chance to build such a peace is the noblest endeavor in which a nation can engage; gladly, also, because only if we act greatly in meeting our responsibilities abroad will we remain a great Nation, and only if we remain a great Nation will we act greatly in meeting our challenges at home.\n\nWe have the chance today to do more than ever before in our history to make life better in America -- to ensure better education, better health, better housing, better transportation, a cleaner environment -- to restore respect for law, to make our communities more livable -- and to insure the God-given right of every American to full and equal opportunity.\n\nBecause the range of our needs is so great -- because the reach of our opportunities is so great -- let us be bold in our determination to meet those needs in new ways.\n\nJust as building a structure of peace abroad has required turning away from old policies that failed, so building a new era of progress at home requires turning away from old policies that have failed.\n\nAbroad, the shift from old policies to new has not been a retreat from our responsibilities, but a better way to peace.\n\nAnd at home, the shift from old policies to new will not be a retreat from our responsibilities, but a better way to progress.\n\nAbroad and at home, the key to those new responsibilities lies in the placing and the division of responsibility. We have lived too long with the consequences of attempting to gather all power and responsibility in Washington.\n\nAbroad and at home, the time has come to turn away from the condescending policies of paternalism -- of "Washington knows best." A person can be expected to act responsibly only if he has responsibility. This is human nature. So let us encourage individuals at home and nations abroad to do more for themselves, to decide more for themselves. Let us locate responsibility in more places. Let us measure what we will do for others by what they will do for themselves.\n\nThat is why today I offer no promise of a purely governmental solution for every problem. We have lived too long with that false promise. In trusting too much in government, we have asked of it more than it can deliver. This leads only to inflated expectations, to reduced individual effort, and to a disappointment and frustration that erode confidence both in what government can do and in what people can do.\n\nGovernment must learn to take less from people so that people can do more for themselves.\n\nLet us remember that America was built not by government, but by people -- not by welfare, but by work -- not by shirking responsibility, but by seeking responsibility.\n\nIn our own lives, let each of us ask -- not just what will government do for me, but what can I do for myself?\n\nIn the challenges we face together, let each of us ask -- not just how can government help, but how can I help?\n\nYour National Government has a great and vital role to play. And I pledge to you that where this Government should act, we will act boldly and we will lead boldly. But just as important is the role that each and every one of us must play, as an individual and as a member of his own community.\n\nFrom this day forward, let each of us make a solemn commitment in his own heart: to bear his responsibility, to do his part, to live his ideals -- so that together, we can see the dawn of a new age of progress for America, and together, as we celebrate our 200th anniversary as a nation, we can do so proud in the fulfillment of our promise to ourselves and to the world.\n\nAs America's longest and most difficult war comes to an end, let us again learn to debate our

differences with civility and decency. And let each of us reach out for that one precious quality government cannot provide -- a new level of respect for the rights and feelings of one another, a new level of respect for the individual human dignity which is the cherished birthright of every American.

Above all else, the time has come for us to renew our faith in ourselves and in America.

In recent years, that faith has been challenged.

Our children have been taught to be ashamed of their country, ashamed of their parents, ashamed of America's record at home and of its role in the world.

At every turn, we have been beset by those who find everything wrong with America and little that is right. But I am confident that this will not be the judgment of history on these remarkable times in which we are privileged to live.

America's record in this century has been unparalleled in the world's history for its responsibility, for its generosity, for its creativity and for its progress.

Let us be proud that our system has produced and provided more freedom and more abundance, more widely shared, than any other system in the history of the world.

Let us be proud that in each of the four wars in which we have been engaged in this century, including the one we are now bringing to an end, we have fought not for our selfish advantage, but to help others resist aggression.

Let us be proud that by our bold, new initiatives, and by our steadfastness for peace with honor, we have made a break-through toward creating in the world what the world has not known before -- a structure of peace that can last, not merely for our time, but for generations to come.

We are embarking here today on an era that presents challenges great as those any nation, or any generation, has ever faced.

We shall answer to God, to history, and to our conscience for the way in which we use these years.

As I stand in this place, so hallowed by history, I think of others who have stood here before me. I think of the dreams they had for America, and I think of how each recognized that he needed help far beyond himself in order to make those dreams come true.

Today, I ask your prayers that in the years ahead I may have God's help in making decisions that are right for America, and I pray for your help so that together we may be worthy of our challenge.

Let us pledge together to make these next four years the best four years in America's history, so that on its 200th birthday America will be as young and as vital as when it began, and as bright a beacon of hope for all the world.

Let us go forward from here confident in hope, strong in our faith in one another, sustained by our faith in God who created us, and striving always to serve His purpose.

```
In [3]: import os
os.chdir("/Users/sabitapanchal/nltk_data/corpora/inaugural")
```

Merging the 3 text files into one text file:

```
In [4]: filenames=['1941-Roosevelt.txt', '1961-Kennedy.txt', '1973-Nixon.txt']
```

```
In [5]: #Combining the 3 text files into one:

with open('newfile.txt','w') as outfile:
    for names in filenames:
        with open(names) as infile:
            outfile.write(infile.read())
            outfile.write("\n")
```

```
In [6]: #Reading the new file:

speech=open('newfile.txt','r').read()
```

Total number of characters in the documents:

```
In [7]: print('Total number of characters in the documents:',len(speech))
```

Total number of characters in the documents: 25183

Total number of words in the documents:

```
In [8]: print('Total number of words in the documents:',len(speech.split()))
```

Total number of words in the documents: 4569

Total number of sentences in the documents:

```
In [9]: nltk.download('punkt')
from nltk.tokenize import sent_tokenize
```

```
[nltk_data] Downloading package punkt to
[nltk_data]      /Users/sabitapanchal/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

```
In [10]: sentences=sent_tokenize(speech)
print('Total number of sentences in the documents:',len(sentences))
```

Total number of sentences in the documents: 188

```
In [11]: print(speech)
```

y which successfully survived crisis at home; put away many evil things; built new structures on enduring lines; and, through it all, maintained the fact of its democracy.

For action has been taken within the three-way framework of the Constitution of the United States. The coordinate branches of the Government continue freely to function. The Bill of Rights remains inviolate. The freedom of elections is wholly maintained. Prophets of the downfall of American democracy have seen their dire predictions come to naught.

Democracy is not dying.

We know it because we have seen it revive--and grow.

We know it cannot die -- because it is built on the unhampered initiative of individual men and women joined together in a common enterprise -- an enterprise undertaken and carried through by the free expression of a free majority.

We know it because democracy alone of all forms of government enlists +

```
In [ ]:
```

```
In [ ]:
```

Removing stopwords:

```
In [12]: from nltk.corpus import stopwords
stop = set(stopwords.words('english') + list(string.punctuation))

from nltk.tokenize import word_tokenize
tokens=word_tokenize(speech)

filtered_speech=[w for w in tokens if not w in stop]
print(filtered_speech)

know , democracy , alone , constructed , unlimited , civilization ,
'capable', 'infinite', 'progress', 'improvement', 'human', 'life', 'We',
'know', 'look', 'surface', 'sense', 'still', 'spreading', 'every', 'conti
nent', '--', 'humane', 'advanced', 'end', 'unconquerable', 'forms', 'huma
n', 'society', 'A', 'nation', 'like', 'person', 'body', '--', 'body', 'mu
st', 'fed', 'clothed', 'housed', 'invigorated', 'rested', 'manner', 'meas
ures', 'objectives', 'time', 'A', 'nation', 'like', 'person', 'mind', '--
', 'mind', 'must', 'kept', 'informed', 'alert', 'must', 'know', 'understa
nds', 'hopes', 'needs', 'neighbors', '--', 'nations', 'live', 'within',
'narrowing', 'circle', 'world', 'And', 'nation', 'like', 'person', 'somet
hing', 'deeper', 'something', 'permanent', 'something', 'larger', 'sum',
'parts', 'It', 'something', 'matters', 'future', '--', 'calls', 'forth',
'sacred', 'guarding', 'present', 'It', 'thing', 'find', 'difficult', '--
', 'even', 'impossible', '--', 'hit', 'upon', 'single', 'simple', 'word',
'And', 'yet', 'understand', '--', 'spirit', '--', 'faith', 'America', 'I
t', 'product', 'centuries', 'It', 'born', 'multitudes', 'came', 'many',
'lands', '--', 'high', 'degree', 'mostly', 'plain', 'people', 'sought',
'early', 'late', 'find', 'freedom', 'freely', 'The', 'democratic', 'aspir
ation', 'mere', 'recent', 'phase', 'human', 'history', 'It', 'human', 'hi
story', 'It', 'permeated', 'ancient', 'life', 'early', 'peoples', 'It',
```

Text preprocessing:

In [13]: *#Tokenizing the 3 text files separately to enable word frequency count:*

```
speech1=open('1941-Roosevelt.txt','r').read()
speech2=open('1961-Kennedy.txt','r').read()
speech3=open('1973-Nixon.txt','r').read()
```

```
token1=word_tokenize(speech1)
token2=word_tokenize(speech2)
token3=word_tokenize(speech3)
token1
```

```
' ',
'the',
'people',
'have',
'renewed',
'their',
'sense',
'of',
'dedication',
'to',
'the',
'United',
'States',
'. ',
'In',
'Washington',
"s",
'day',
'the',
'task'.
```


In [14]: *#Converting to dataframe for better processing:*

```
data1=pd.DataFrame(token1,columns=['Words'],index=None)
data2=pd.DataFrame(token2,columns=['Words'],index=None)
data3=pd.DataFrame(token3,columns=['Words'],index=None)
data1
```

Out[14]:

	Words
0	On
1	each
2	national
3	day
4	of
...	...
1521	the
1522	will
1523	of
1524	God
1525	.

1526 rows × 1 columns

In [15]: *#Converting each speech into lowercase for further processing:*

```
data1['Words'] = data1['Words'].apply(lambda x: " ".join(x.lower() for x in x.split()))
data2['Words'] = data2['Words'].apply(lambda x: " ".join(x.lower() for x in x.split()))
data3['Words'] = data3['Words'].apply(lambda x: " ".join(x.lower() for x in x.split()))
```

In [16]: *#For each speech, we will remove the stopwords and punctuation before finding*

```
stop = set(stopwords.words('english') + list(string.punctuation))

data1['Words'] = data1['Words'].apply(lambda x: " ".join(x for x in x.split() if x not in stop))
data2['Words'] = data2['Words'].apply(lambda x: " ".join(x for x in x.split() if x not in stop))
data3['Words'] = data3['Words'].apply(lambda x: " ".join(x for x in x.split() if x not in stop))
data1['Words'].head(10)
```

Out[16]:

```
0
1
2      national
3         day
4
5  inauguration
6         since
7         1789
8
9
Name: Words, dtype: object
```

In [17]: *#Removing '--' before finding word frequencies:*

```
char=['--']
data1['Words'] = data1['Words'].apply(lambda x: " ".join(x for x in x.split() if x != '--'))
data2['Words'] = data2['Words'].apply(lambda x: " ".join(x for x in x.split() if x != '--'))
data3['Words'] = data3['Words'].apply(lambda x: " ".join(x for x in x.split() if x != '--'))
```

In [18]: *#Removing the blank spaces:*

```
data1['Words'].replace('', np.nan, inplace=True)
data2['Words'].replace('', np.nan, inplace=True)
data3['Words'].replace('', np.nan, inplace=True)

data1.dropna(subset=['Words'], inplace=True)
data2.dropna(subset=['Words'], inplace=True)
data3.dropna(subset=['Words'], inplace=True)

data1.head(10)
```

Out[18]:

	Words
2	national
3	day
5	inauguration
6	since
7	1789
10	people
12	renewed
14	sense
16	dedication
19	united

Finding most frequent words in each speech:

In [19]: `from nltk.probability import FreqDist`

In [20]: `print("The 3 most frequently used words in President Roosevelt's speech are")
freq1 = pd.Series(' '.join(data1['Words']).split()).value_counts()[:3]
freq1`

The 3 most frequently used words in President Roosevelt's speech are:

Out[20]:

nation	12
know	10
spirit	9
dtype:	int64

```
In [21]: print("The 3 most frequently used words in President Kennedy's speech are:")
freq2 = pd.Series(' '.join(data2['Words']).split()).value_counts()[:3]
freq2
```

The 3 most frequently used words in President Kennedy's speech are:

```
Out[21]: let      16
us         12
world      8
dtype: int64
```

```
In [22]: print("The 3 most frequently used words in President Nixon's speech are:")
freq3 = pd.Series(' '.join(data3['Words']).split()).value_counts()[:3]
freq3
```

The 3 most frequently used words in President Nixon's speech are:

```
Out[22]: us         26
let        22
america    21
dtype: int64
```

Creating word cloud for each speech:

```
In [23]: #Filtering out more words that don't add meaning:
```

```
stop = list(stopwords.words('english'))
stop.extend(["let", "us", "new", "s", "mr."])
```

```
In [24]: corpus1 = data1['Words'].apply(lambda x: ' '.join([z for z in x.split() if
corpus1.head(10)
```

```
Out[24]: 2          national
3             day
5    inauguration
6             since
7             1789
10         people
12        renewed
14             sense
16    dedication
19         united
Name: Words, dtype: object
```

```
In [25]: corpus2 = data2['Words'].apply(lambda x: ' '.join([z for z in x.split() if
corpus2.head(10)
```

```
Out[25]: 0         vice
1      president
2        johnson
4
5        speaker
7
8         chief
9        justice
11     president
12    eisenhower
Name: Words, dtype: object
```

```
In [26]: corpus3 = data3['Words'].apply(lambda x: ' '.join([z for z in x.split() if
corpus3.head(10)
```

```
Out[26]: 0
1         vice
2     president
4
5        speaker
7
8         chief
9        justice
11     senator
12         cook
Name: Words, dtype: object
```

```
In [27]: data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 632 entries, 2 to 1524
Data columns (total 1 columns):
#   Column  Non-Null Count  Dtype
---  -
0    Words    632 non-null    object
dtypes: object(1)
memory usage: 9.9+ KB
```

```
In [28]: data=[data1[ 'Words' ],data2[ 'Words' ],data3[ 'Words' ]]  
headers=[ "Roosevelt", "Kennedy", "Nixon" ]  
df=pd.concat(data,axis=1,keys=headers)  
df.head(10)
```

Out[28]:

	Roosevelt	Kennedy	Nixon
0	NaN	vice	mr.
1	NaN	president	vice
2	national	johnson	president
3	day	NaN	NaN
4	NaN	mr.	mr.
5	inauguration	speaker	speaker
6	since	NaN	NaN
7	1789	mr.	mr.
8	NaN	chief	chief
9	NaN	justice	justice

```
In [29]: pip install wordcloud
```

```
Requirement already satisfied: wordcloud in /opt/anaconda3/anaconda3/lib/  
python3.8/site-packages (1.8.1)  
Requirement already satisfied: numpy>=1.6.1 in /opt/anaconda3/anaconda3/  
lib/python3.8/site-packages (from wordcloud) (1.19.2)  
Requirement already satisfied: pillow in /opt/anaconda3/anaconda3/lib/pyt  
hon3.8/site-packages (from wordcloud) (7.2.0)  
Requirement already satisfied: matplotlib in /opt/anaconda3/anaconda3/li  
b/python3.8/site-packages (from wordcloud) (3.2.2)  
Requirement already satisfied: cycler>=0.10 in /opt/anaconda3/anaconda3/l  
ib/python3.8/site-packages (from matplotlib->wordcloud) (0.10.0)  
Requirement already satisfied: python-dateutil>=2.1 in /opt/anaconda3/ana  
conda3/lib/python3.8/site-packages (from matplotlib->wordcloud) (2.8.1)  
Requirement already satisfied: kiwisolver>=1.0.1 in /opt/anaconda3/anacon  
da3/lib/python3.8/site-packages (from matplotlib->wordcloud) (1.2.0)  
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 i  
n /opt/anaconda3/anaconda3/lib/python3.8/site-packages (from matplotlib->  
wordcloud) (2.4.7)  
Requirement already satisfied: six in /opt/anaconda3/anaconda3/lib/python  
3.8/site-packages (from cycler>=0.10->matplotlib->wordcloud) (1.15.0)  
Note: you may need to restart the kernel to use updated packages.
```

```
In [46]: from wordcloud import WordCloud

wc_a= ' '.join(corpus2)

wordcloud = WordCloud(width = 2000, height = 2000,
                        background_color = 'black',
                        min_font_size = 10, random_state=100).generate(wc_a)

# plot the WordCloud image
plt.figure(figsize = (10, 10), facecolor = None)
plt.show(wc_a)
plt.axis("off")
plt.xlabel('Word Cloud')
plt.tight_layout(pad = 0)

print("Word Cloud for President Kennedy's speech")
plt.show()
```

<Figure size 720x720 with 0 Axes>

Word Cloud for President Kennedy's speech

```
In [43]: from wordcloud import WordCloud

wc_a= ' '.join(corpus1)

wordcloud = WordCloud(width = 2000, height = 2000,
                        background_color = 'black',
                        min_font_size = 10, random_state=100).generate(wc_a)

# plot the WordCloud image
plt.figure(figsize = (10, 10), facecolor = None)
plt.imshow(wc_a)
plt.axis("off")
plt.xlabel('Word Cloud')
plt.tight_layout(pad = 0)

print("Word Cloud for President Roosevelt's speech")
plt.show()
```

<Figure size 720x720 with 0 Axes>

Word Cloud for President Roosevelt's speech

```
In [44]: from wordcloud import WordCloud

wc_a= ' '.join(corpus3)

wordcloud = WordCloud(width = 2000, height = 2000,
                        background_color = 'black',
                        min_font_size = 10, random_state=100).generate(wc_a)

# plot the WordCloud image
plt.figure(figsize = (10, 10), facecolor = None)
plt.imshow(wc_a)
plt.axis("off")
plt.xlabel('Word Cloud')
plt.tight_layout(pad = 0)

print("Word Cloud for President Nixon's speech")
plt.show()
```

<Figure size 720x720 with 0 Axes>

Word Cloud for President Nixon's speech

In []: