```python
import pandas as pd
import numpy as np
import string
import spacy
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
from sklearn.feature_extraction.text import CountVectorizer
import wordcloud
import matplotlib.pyplot as plt
from wordcloud import WordCloud,STOPWORDS
```

```python
data=pd.read_csv('Downloads/stop.txt',error_bad_lines=False)
data
```

	a
0	a's
1	able
2	about
3	above
4	according
...	...
565	yours
566	yourself
567	yourselves
568	z
569	zero

570 rows × 1 columns

```python
data=[a.strip() for a in data.a]
data=[a for a in data if a]
data
```

```
["a's",
 'able',
 'about',
 'above',
 'according',
 'accordingly',
 'across',
 'actually',
```

```
join=''.join(data)
join
```

"a'sableaboutaboveaccordingaccordinglyacrossactuallyafterafterwardsagainagainstain'tallallowallowsalmostalonealongalreadyalsoalthoughalwaysamamongamongstanandanotheranyanybodyanyhowanyoneanythinganywayanywaysanywhereapartappearappreciateappropriateareAren'taroundasasideaskaskingassociatedatavailableawayawfullybbbecamebecausebecomebecomesbecomingbeenbeforebeforehandbehindbeingbelievebelowbesidebesidesbestbetterbetweenbeyondbothbriefbutbycc'monc'scamecancan'tcannotcantcausecausescertaincertainlychangesclearlycocomcomecomesconcerningconsequentlyconsiderconsideringcontaincontainingcontainscorrespondingcouldcouldn'tcoursecurrentlyddefinitelydescribeddespitediddidn'tdifferentdodoesdoesn'tdoingdon'tdonedowndownwardsduringeeacheduegeighteitherelseelsewhereenoughentirelyespeciallyetetceveneneverEveryeverybodyeveryoneeverythingeverywhereeexactlyexampleexceptffarfewfifthfirstfivefollowedfollowingfollowsforformerformerlyforthfourfromfurtherfurthermoreggetgetsgettinggivengivesgogoesgoinggonegotgottengreetingshhadhadn'thappenshardlyhashasn'thavehaven'thavinghehe'shellohelphenceherehereherebyhereinhereuponhersherselfhihimhimselfhishitherhopefullyhowhowbeithoweverii'di'lli'mi'veieifignoredimmediateininasmuchincindeedindicateindicatedindicatesinnerinsofarinsteadintoinwardisisn'tititit'dit'llit'sitsitselfjjustkkeepkeepskeptknowknowsknownllastlatelylaterlatterlatterlyleastlesslestletlet'slikelikedlikelylittlelooklookinglooksltdmmainlymanymaymaybememeanmeanwhilemerelymightmoremoreovermostmostlymuchmustmymyselfnnamenamelyndnearnearlynecessaryneedneedsneithernevernheverthelessnewnextninenonobodynonnonenoonenornormallynotnothingnovelnownowhereoobviouslyoffoftenohokokayoldononceoneonesonlyontoorotherothersotherwiseoughtouroursourselvesoutoutsideoveroverallownpparticularparticularlyperperhapsplacedpleaseplus possiblepresumablyprobablyprovidesqquequiteqvrratherrdrereallyreasonablyregardingregardlessregardsrelativelyrespectivelyrightssaidsamesawsaysayingsayssecondsecondlyseeseeingseemseemedseemingseemsseenselfselvessensiblesentseriousseriouslysevenseveralshallsheshouldshouldn'tsincesixsosomesomebodysomehowsomeonesomethingsometimesometimessomewhatsomewheresoonsorryspecifiedspecifyspecifyingstillsubsuchsupsurett'staketakentelltendsththanthankthanksthanxthatthat'sthatsthetheirtheirsthemthemselvesthenthencetherehereethere'sthereafterthereby therefore thereintheresthereuponthesetheythey'dthey'llthey'rethey'vethinkthirdthisthoroughthoroughlythosethoughthreethroughthroughoutthruthustotogethertootooktowardtowardstriedtriestrulytrytryingtwicetwouununderunfortunatelyunlessunlikelyuntiluntoupuponususeusedusefulusesusingusuallyuucpvvaluevariousveryviavizvswwantwantswaswasn'twaywewe'dwe'llwe'rewe'vewelcomewellwentwerewerwen'twhatwhat'swhateverwhenwhencewheneverwherewhere'swhereafterwhereaswherebywhereinwhereuponwhereverwhetherwhichwhilewhitherwhowho'swhoeverwholewhomwhosewhywillwillingwishwithwithinwithoutwon'twonderwouldwouldwouldn'txxyyesyetyouyou'dyou'llyou'reyou'veyouryoursyourselfyourselveszzero"

```
trans=join.translate(str.maketrans(",",string.punctuation))
trans
```

```
word=word_tokenize(trans)
word
```

```python
stop=stopwords.words('english')
stop.append('the')
s=[word for word in word if not word in stop]
s
```

```
"a'sableaboutaboveaccordingaccordinglyacrossactuallyafterafterwardsagainagainstain'tallallowallowsalmostalonealongalreadyalsoal
thoughalwaysamamongamongstanandanotheranyanybodyanyhowanyoneanythinganywayanywaysanywhereapartappearappreciateappropriateareare
n'taroundasasideaskaskingassociatedatavailableawayawfullybbebecamebecausebecomebecomesbecomingbeenbeforebeforehandbehindbeingbe
lievebelowbesidebesidesbestbetterbetweenbeyondbothbriefbutbycc'monc'scamecancan'tcannotcantcausecausescertaincertainlychangescl
earlycocomcomecomesconcerningconsequentlyconsiderconsideringcontaincontainingcontainscorrespondingcouldcouldn'tcoursecurrentlyd
definitelydescribeddespitediddidn'tdifferentdodoesdoesn'tdoingdon'tdonedowndownwardsduringeeacheduegeeighteitherelseelsewhereeno
ughentirelyespeciallyetetceneverevereveryeveryeverybodyeveryoneeverythingeverywhereexexactlyexampleexceptffarfewfifthfirstfivefollowed
followingfollowsforformerformerlyforthfourfromfurtherfurthermoreggetgetsgettinggivengivesgogoesgoinggonegotgottengreetingshhadh
adn'thappenshardlyhashasn'thavehaven'thavinghehe'shellohelphenceherherehere'shereafterherebyhereinhereuponhersherselfhhimhimhimse
lfhishitherhopefullyhowhowbeithoweverii'di'lli'mi'veieifignoredimmediateininasmuchincindeedindicateindicatedindicatesinnerinsof
arinsteadintoinwardsisisn'titit'dit'llit'sitsitselfjjustkkeepkeepskeptknowknowsknownllastlatelylaterlatterlatterlyleastlesslestl
etlet'slikelikedlikelylittlelooklookinglooksltdmmainlymanymaymaybememeanmeanwhilemerelymightmoremoreovermostmostlymuchmustmymys
elfnnamenamelyndnearnearlynecessaryneedneedsneithernevernevertheless newnextninenonobodynonnonenoonenornormallynotnothingnovelno
wnowhereoobviouslyofoffoftenohokokayoldononceoneonesonlyontoorotherothersotherwiseoughtouroursourselvesoutoutsideoveroverallown
pparticularparticularlyperperhapsplacedpleasepluspossiblepresumablyprobablyprovidesqquequiteqvrratherrdrereallyreasonablyregard
ingregardlessregardsrelativelyrespectivelyrightssaidsamesawsaysayingsayssecondsecondlyseeseeingseemseemedseemingseemsseenselfse
lvessensiblesentseriousseriouslysevenseveralshallsheshouldshouldn'tsincesixsosomesomebodysomehowsomeonesomethingsometimesometim
essomewhatsomewheresoonsorryspecifiedspecifyspecifyingstillsubsuchsupsurett'staketakentelltendsththanthankthanksthanxthatthat's
thatsthetheirtheirsthemthemselvesthenthencetherethere'sthereafterherebythereforethereintherestthereuponthesetheythey'dthey'llth
ey'rethey'vethinkthirdthisthoroughthoroughlythosethoughthreethroughthroughoutthruthustotogethertootooktowardtowardstriedtriestr
ulytrytryingtwicetwouununderunfortunatelyunlessunlikelyuntiluntoupuponususeusedusefulusesusingusuallyuucpvvaluevariousveryviavi
zvswwantwantswaswasn'twaywewe'dwe'llwe'rewe'vewelcomewellwentwerewereen'twhatwhat'swhateverwhenwhencewheneverwherewhere'swhereaf
terwhereaswherebywhereinwhereuponwhereverwhetherwhichwhilewhitherwhowho'swhoeverwholewhomwhosewhywillwillingwishwithwithinwitho
utwon'twonderwouldwouldwouldn'txxyyesyetyouyou'dyou'llyou'reyou'veyouryoursyourselfyourselveszzero"
```

```python
low=[a.lower() for a in s]
low
```

```python
ps=PorterStemmer()
p=[ps.stem(word) for word in low]
p
```

```python
nlp=spacy.blank('en')
nlp
```

```python
cv=CountVectorizer()
x=cv.fit_transform(p)
```

```python
print(cv.vocabulary_)
```

"a'sableaboutaboveaccordingaccordinglyacrossactuallyafterafterwardsagainagainstain'tallallowallowsalmostalonealongalreadyalsoal
thoughalwaysamamongamongstanandanotheranyanybodyanyhowanyoneanythinganywayanywaysanywhereapartappearappreciateappropriateareare
n'taroundasasideaskaskingassociatedataavailableawayawfullybbebecamebecausebecomebecomesbecomingbeenbeforebeforehandbehindbeingbe
lievebelowbesidebesidesbestbetterbetweenbeyondbothbriefbutbycc'monc'scamecancan'tcannotcantcausecausescertaincertainlychangescl
earlycocomcomecomesconcerningconsequentlyconsiderconsideringcontaincontainingcontainscorrespondingcouldcouldn'tcoursecurrentlyd
definitelydescribeddespitediddidn'tdifferentdodoesdoesn'tdoingdon'tdonedowndownwardsduringeeacheduegeighteitherelseelsewhereeno
ughentirelyespeciallyetetcevenevereveryeverybodyeveryoneeverythingeverywhereexexactlyexampleexceptffarfewfifthfirstfivefollowed
followingfollowsforformerformerlyforthfourfromfurtherfurthermoreggetgetsgettinggivengivesgogoesgoinggonegotgottengreetingshhadh
adn'thappenshardlyhashasn'thavehaven'thavinghehe'shellohelphenceherhereherehere'shereafterherebyhereinhereuponhersherselfhihimhimse
lfhishitherhopefullyhowhowbeithoweverii'di'lli'mi'veieifignoredimmediateininasmuchincindeedindicateindicatedindicatesinnerinsof
arinsteadintoinwardisisn'titit'dit'llit'sitsitselfjjustkkeepkeepskeptknowknowsknownllastlatelylaterlatterlatterlyleastlesslestl
etlet'slikelikedlikelylittlelooklookinglooksltdmmainlymanymaymaybememeanmeanwhilemerelymightmoremoreovermostmostlymuchmustmymys
elfnnamenamelyndnearnearlynecessaryneedneedsneithernevernevertheelessnewnextnineononobodynonnonenoonenornormallynotnothingnovelno
wnowhereoobviouslyofoffoftenohokokayoldononceoneonesonlyontoorotherothersotherwiseoughtouroursourselvesoutoutsideoveroverallown
pparticularparticularlyperperhapsplacedpleaseplspossiblepresumablyprobablyprovidesqquequiteqvrratherrdrereallyreasonablyregard
ingregardlessregardsrelativelyrespectivelyrightssaidsamesawsaysayingsayssecondsecondlyseeseeingseemseemedseemingseemsseenselfse
lvessensiblesentseriousseriouslysevenseveralshallsheshouldshouldn'tsincesixsosomesomebodysomehowsomeonesomethingsometimesometim
essomewhatsomewheresoonsorryspecifiedspecifyspecifyingstillsubsuchsupsurett'staketakentelltendsththanthankthanksthanxthatthat's
thatsthetheirtheirsthemthemselvesthenthencetherethere'sthereafterherebythereforethereinthereserethereuponthesetheytheythey'dthey'llth
ey'rethey'vethinkthirdthisthoroughthoroughlythosethoughthreethroughthroughoutthruthustotogethertootooktowardtowardstriedtriestr
ulytrytryingtwicetwouununderunfortunatelyunlessunlikelyuntiluntoupupononususeusedusefulusesusingusuallyuucpvvaluevariousveryviavi
zvswwantwaswasn'twaywewe'dwe'llwe'rewe'vewelcomewellwentwereweeren'twhatwhat'swhateverwhenwhencewheneverwherewhere'swhereaf
terwhereaswherebywhereinwhereuponwhereverwhetherwhichwhilewhitherwhowho'swhoeverwholewhomwhosewhywillwillingwishwithwithinwitho
utwon'twonderwouldwouldwouldn'txxyyesyetyouyou'dyou'llyou'reyou'veyouryoursyourselfyourselveszzero"

```python
print(cv.get_feature_names()[0:16])
print(x.toarray()[0:16])
```

```python
cv=CountVectorizer(ngram_range=(1,30),analyzer='word',max_features=30)
x=cv.fit_transform(p)
```

```python
print(cv.vocabulary_)
```

```python
print(cv.get_feature_names()[0:16])
print(x.toarray()[0:16])
```

```
"a'sableaboutaboveaccordingaccordinglyacrossactuallyafterafterwardsagainagainstain'tallallowallowsalmostalonealongalreadyalsoal
thoughalwaysamamongamongstanandanotheranyanybodyanyhowanyoneanythinganywayanywaysanywhereapartappearappreciateappropriateareare
n'taroundasasideaskaskingassociatedatavailableawayawfullybbebecamebecausebecomebecomesbecomingbeenbeforebeforehandbehindbeingbe
lievebelowbesidebesidesbestbetterbetweenbeyondbothbriefbutbycc'monc'scamecancan'tcannotcantcausecausescertaincertainlychangescl
earlycocomcomecomesconcerningconsequentlyconsiderconsideringcontaincontainingcontainscorrespondingcouldcouldn'tcoursecurrentlyd
definitelydescribeddespitediddidn'tdifferentdodoesdoesn'tdoingdon'tdonedowndownwardsduringeeacheduegeighteitherelseelsewhereeno
ughentirelyespeciallyetetcevenevereveryeverybodyeveryoneeverythingeverywhereexactlyexampleexceptffarfewfifthfirstfivefollowed
followingfollowsforformerformerlyforthfourfromfurtherfurthermoreggetgetsgettinggivengivesgogoesgoinggonegotgottengreetingshhadh
adn'thappenshardlyhashasn'thavehaven'thavinghehe'shellohelphenceherherehere'shereafterherebyhereinhereuponhersherselfhihimhimse
lfhishitherhopefullyhowhowbeithoweverii'di'lli'mi'veieifignoredimmediateininasmuchincindeedindicateindicatedindicatesinnerinsof
arinsteadintoinwardisisn'titit'dit'llit'sitsitselfjjustkkeepkeepskeptknowknowsknownllastlatelylaterlatterlatterlyleastlesslestl
etlet'slikelikedlikelylittlelooklookinglooksltdmmainlymanymaymaybememeanmeanwhilemerelymightmoremoreovermostmostlymuchmustmymys
elfnnamenamelyndnearnearlynecessaryneedneedsneitherneverneverthelessnewnextninenonobodynonnonenoonenornormallynotnothingnovelno
wnowhereoobviouslyofofffoftenohokokayoldononceoneoneonesonlyontoorotherothersotherwiseoughtouroursourselvesoutoutsideoveroverallown
pparticularparticularlyperperhapsplacedpleasepluspossiblepresumablyprobablyprovidesqquequiteqvrratherrdrereallyreasonablyregard
ingregardlessregardsrelativelyrespectivelyrightssaidsamesawsaysayingsayssecondsecondlyseeseeingseemseemedseemingseemsseenselfse
lvessensiblesentseriousseriouslysevenseveralshallsheshouldshouldn'tsincesixsosomesomebodysomehowsomeonesomethingsometimesometim
essomewhatsomewheresoonsorryspecifiedspecifyspecifyingstillsubsuchsupsurett'staketakentelltendsththanthankthanksthanxthatthat's
thatsthetheirtheirsthemthemselvesthenthencetherethere'sthereaftertherebythereforethereintheresthereuponthesetheythey'dthey'llth
ey'rethey'vethinkthirdthisthoroughthoroughlythosethoughthreethroughthroughoutthruthustotogethertootooktowardtowardstriedtriestr
ulytrytryingtwicetwouununderunfortunatelyunlessunlikelyuntiluntoupuponususeuseduseful usesusingusuallyuucpvvaluevariousveryviavi
zvswwantwantswaswasn'twaywewe'dwe'llwe'rewe'vewelcomewellwentwereweren'twhatwhat'swhateverwhenwhencewheneverwherewhere'swhereaf
terwhereaswherebywhereinwhereuponwhereverwhetherwhichwhilewhitherwhowho'swhoeverwholewhomwhosewhywillwillingwishwithwithinwitho
utwon'twonderwouldwouldwouldn'txyyesyetyouyou'dyou'llyou'reyou'veyouryoursyourselfyourselveszzero"
```

```python
def plot_cloud(wordcloud):
    plt.figure(figsize=(16,8))
    plt.imshow(wordcloud)


stopwords=STOPWORDS
stopwords.add('the')
wordcloud=WordCloud(height=3000,width=1000,background_color='black',max_words=30,color
map='Set2',stopwords=stopwords).generate(join)
plot_cloud(wordcloud)
```