

```
!python -m textblob.download_corpora
```

```
[nltk_data] Downloading package brown to /root/nltk_data...
[nltk_data]   Unzipping corpora/brown.zip.
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]   /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger.zip.
[nltk_data] Downloading package conll2000 to /root/nltk_data...
[nltk_data]   Unzipping corpora/conll2000.zip.
[nltk_data] Downloading package movie_reviews to /root/nltk_data...
[nltk_data]   Unzipping corpora/movie_reviews.zip.
Finished.
```

```
import pandas as pd
import matplotlib.pyplot as plt
from textblob import TextBlob
from nltk.corpus import stopwords
from operator import itemgetter
import nltk
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('punkt')
```

```
↳ [nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
True
```

```
!pip install requests
```

```
import requests
```

```
target_url = 'http://www.gutenberg.org/files/2265/2265.txt'
```

```
response = requests.get(target_url)
```

```
data = response.text
```


```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (2.23.0)
Requirement already satisfied: certifi<=2017.4.17 in /usr/local/lib/python3.7/dist-packages (2017.4.17)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (2.5)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (1.25.0)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (3.0.2)
```

```
blob=TextBlob(data)
```

```
items = blob.word_counts.items()
stop_words = stopwords.words('english')
items = [item for item in items if item[0] not in stop_words]
```

```
sorted_items = sorted(items, key=itemgetter(1), reverse=True)
top20 = sorted_items[0:20]
```

```
df = pd.DataFrame(top20, columns=['word', 'count'])
df
```

	word	count	
0	ham	337	
1	lord	211	
2	haue	175	
3	king	173	
4	shall	107	
5	hamlet	106	
6	come	106	
7	thou	105	
8	let	104	
9	good	99	
10	hor	95	
11	thy	90	
12	may	85	
13	enter	85	
14	oh	81	
15	like	79	
16	well	71	
17	know	70	
18	would	69	
19	selfe	68	

[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 8:35 PM

