```
!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
                     /root/nltk data...
     [nltk_data]
     [nltk_data]
                   Unzipping taggers/averaged_perceptron_tagger.zip.
     [nltk data] Downloading package conll2000 to /root/nltk data...
                   Unzipping corpora/conll2000.zip.
     [nltk_data]
     [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...
                   Unzipping corpora/stopwords.zip.
     [nltk data]
     [nltk data] Downloading package wordnet to /root/nltk data...
                   Package wordnet is already up-to-date!
     [nltk data]
     [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: <a href="https://pypi.org/simple">https://us-python.pkg.dev/colab-wheels/pub</a>.
     Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (2.23
     Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packa
```

https://colab.research.google.com/drive/1Ng-tRzJEZkXPaErTN1KX0uTSHHuey1Ga#scrollTo=w-undunKuCKB&printMode=true

```
blob=TextBlob(data)

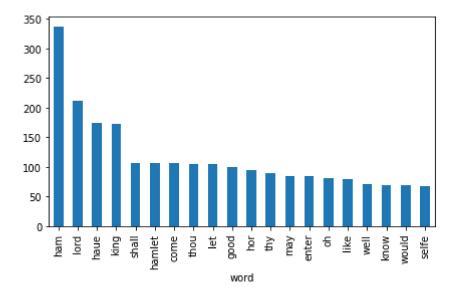
items = blob.word_counts.items()
stop_words = stopwords.words('english')
items = [item for item in items if item[0] not in stop_words]

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

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

	word	count	1
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	

axes = df.plot.bar(x='word', y='count', legend=False)
plt.gcf().tight_layout()



import imageio
image_file="https://media.cheggcdn.com/media/216/21621ee5-e80f-47f3-9145-513f2229b390/phploeB
mask image = imageio.imread(image file)

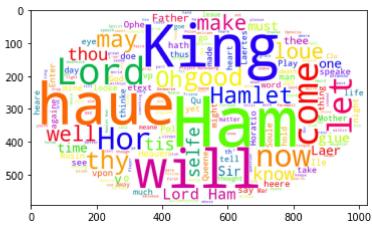
from wordcloud import WordCloud

wordcloud = WordCloud(colormap='prism', mask=mask_image, background_color='white')

wordcloud = wordcloud.generate(data)

plt.imshow(wordcloud)

<matplotlib.image.AxesImage at 0x7f546be469d0>



Colab paid products - Cancel contracts here

✓ 0s completed at 8:35 PM

X