```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import nltk
4 from pathlib import Path
5 from textblob import TextBlob
6 from nltk.corpus import stopwords
7 from operator import itemgetter
```

```
1 nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
True
```

```
1 nltk.download('punkt')
```

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

```
1 text = Path('romeoandjuliet.txt').read_text()
```

```
1 blob = TextBlob(text)
```

```
1 stop_words = stopwords.words('english')
2
```

```
1 items = blob.word_counts.items()
```

```
1 items = [item for item in items if item[0] not in stop_words]
```

```
1 sorted_items = sorted(items, key=itemgetter(1), reverse=True)
```

```
1 top20 = sorted_items[1:21]
```

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

	word	count
0	romeo	440
1	1	435
2	etc	428
3	juliet	325
4	thou	324
5	2	317
6	iii	307
7	v	288
8	3	280
9	ii	269
10	see	255
11	love	238
12	iv	230
13	thy	229
14	capulet	195
15	one	191
16	4	191
17	thee	180
18	nurse	179
19	may	165

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

1