

```
In [ ]: pip install nltk
```

```
In [1]: from urllib import request  
# Little women in plain text  
url = "https://www.gutenberg.org/cache/epub/514/pg514.txt"  
response = request.urlopen(url)  
raw = response.read().decode('utf8')  
type(raw)
```

```
Out[1]: str
```

```
In [3]: len(raw)
```

```
Out[3]: 1047545
```

```
In [5]: raw[:100]
```

```
Out[5]: '\uffeffThe Project Gutenberg eBook of Little Women\r\n    \r\nThis ebook is for the use of anyone anywhere in '
```

```
In [6]: from nltk import word_tokenize  
tokens = word_tokenize(raw)  
type(tokens)
```

```
Out[6]: list
```

```
In [7]: len(tokens)
```

```
Out[7]: 233062
```

```
In [8]: tokens[5:10]
```

```
Out[8]: ['Little', 'Women', 'This', 'ebook', 'is']
```

```
In [9]: import nltk  
text = nltk.Text(tokens)  
type(text)
```

```
Out[9]: nltk.text.Text
```

```
In [10]: text[1024:1062]
```

```
Out[10]: ['any',  
          'of',  
          'you',  
          'suffer',  
          'as',  
          'I',  
          'do',  
          ',',  
          ',',  
          ',',  
          'cried',  
          'Amy',  
          ',',  
          ',',  
          ',',  
          'for',  
          'you',  
          'don',  
          ',',  
          ',',  
          't',  
          'have',  
          'to',  
          'go',  
          'to',  
          'school',  
          'with',  
          'impertinent',  
          'girls',  
          ',',  
          'who',  
          'plague',  
          'you',  
          'if',  
          'you',  
          'don',  
          ',',  
          ',',  
          't',  
          'know',  
          'your',  
          'lessons']
```

```
In [11]: text.collocations()
```

```
Mrs. March; Aunt March; Project Gutenberg™; Mr. Laurence; Mr. Brooke;  
Mr. Bhaer; old gentleman; Project Gutenberg; old lady; said Meg;  
United States; young lady; dare say; great deal; young ladies;  
Literary Archive; every day; Gutenberg™ electronic; one another; said  
Amy
```

```
In [19]: raw.find("PART 1")
```

```
Out[19]: 802
```

```
In [20]: raw.rfind("END OF THE PROJECT GUTENBERG EBOOK LITTLE WOMEN")
```

```
Out[20]: 1028719
```

```
In [21]: raw = raw[802:1028719]
```

```
In [22]: text.concordance("Mrs.")
```

Displaying 25 of 235 matches:

hile making these maternal inquiries Mrs. March got her wet things off , her w
. As they gathered about the table , Mrs. March said , with a particularly hap
pecial message to you girls , ” said Mrs. March , patting her pocket as if she
rought round the happy coming home . Mrs. March broke the silence that followe
will find your guidebook , ” replied Mrs. March . They talked over the new pla
. “ I thought you ’ d do it , ” said Mrs. March , smiling as if satisfied . “
es with old hats and her own cloak . Mrs. March gave the mother tea and gruel
enacted escort with great dignity . Mrs. March was both surprised and touched
Old Mr. Laurence sent it , ” replied Mrs. March . “ The Laurence boy ’ s grand
’ s roses are sweeter to me , ” said Mrs. March , smelling the half-dead posy
! A regular note of invitation from Mrs. Gardiner for tomorrow night ! ” crie
Have a good time , dearies ! ” said Mrs. March , as the sisters went daintily
eg , as she turned from the glass in Mrs. Gardiner ’ s dressing room after a p
ring was , it was an event to them . Mrs. Gardiner , a stately old lady , gree
make a great racket getting ready . Mrs. March was very busy trying to finish
distracted with your worry , ” cried Mrs. March , crossing out the third spoil
and Papa had sent him away . I heard Mrs. King crying and Mr. King talking ver
id Jo , after a minute ’ s silence . Mrs. March smiled and began at once , for
and flew about to get ready , for as Mrs. March said , he was ‘ a little gentl
use on the other side of the hedge . Mrs. March wanted to talk of her father w
ough to show how grateful he was for Mrs. March ’ s motherly welcome , their c
that little nunnery over there , and Mrs. March is doing more for him than we
ad just occurred to him , he said to Mrs. March ... “ The boy neglects his mus
quite took her breath away . Before Mrs. March could reply , Mr. Laurence wen
ll pay for the making up , ” replied Mrs. March , who took peculiar pleasure i

```
In [ ]:
```