

```
In [1]: import urllib.request

url = "https://www.gutenberg.org/cache/epub/10/pg10.txt"

with urllib.request.urlopen(url) as response:
    pg10 = response.read().decode('utf-8')
```

```
In [14]: lines = pg10.splitlines()
num_lines = len(lines)
print(f"Number of lines: {num_lines}")
```

Number of lines: 99968

```
In [18]: pg10_lower = pg10.lower()

words = []
current_word = ''
for char in pg10_lower:
    if 'a' <= char <= 'z':
        current_word += char
    else:
        if current_word:
            words.append(current_word)
            current_word = ''
if current_word:
    words.append(current_word)

word_count = len(words)

print(f"Word Count: {word_count}")
```

Word Count: 795227

```
In [21]: apostle_occurences = words.count("apostle")
print(f"The word 'apostle' appears {apostle_occurences} times in the text")
```

The word 'apostle' appears 47 times in the text

```
In [22]: frequencies = {}
for word in words:
    frequencies[word] = frequencies.get(word, 0) + 1

most_common = None
max_count = 0
for word in frequencies:
    if frequencies[word] > max_count:
        most_common = word
        max_count = frequencies[word]
```

```
print(f" The word '{most_common}' is the most common word.")
```

The word 'the' is the most common word.