Import statements

In [24]:
```
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk.corpus import gutenberg
```

In [25]:
```
wrd = gutenberg.words('carroll-alice.txt')
book = ''
for word in wrd:
    book = book + word + ' '
sentences = sent_tokenize(book)
word_count = lambda sentence: len(word_tokenize(sentence))
longestsent = max(sentences, key=word_count)
```

In [26]:
```
print("\nName of the corpus is: carroll-alice.txt\n")
print('Longest sentence found using sentence tokenizer: \n\n', longestsent)
print('\nLength of the above sentence: ', len(longestsent))
sentences = book.split(". ")
longest_sent = max(sentences, key=len)
print('Longest sentence between two consecutive periods: \n\n', longest_sent)
print('\nLength of the above sentence: ', len(longest_sent))
```

```
Name of the corpus is: carroll-alice.txt

Longest sentence found using sentence tokenizer:

 Hardly knowing what she did , she picked up a little bit of stick , an
d held it out to the puppy ; whereupon the puppy jumped into the air of
f all its feet at once , with a yelp of delight , and rushed at the sti
ck , and made believe to worry it ; then Alice dodged behind a great th
istle , to keep herself from being run over ; and the moment she appear
ed on the other side , the puppy made another rush at the stick , and t
umbled head over heels in its hurry to get hold of it ; then Alice , th
inking it was very like having a game of play with a cart - horse , and
expecting every moment to be trampled under its feet , ran round the th
istle again ; then the puppy began a series of short charges at the sti
ck , running a very little way forwards each time and a long way back ,
and barking hoarsely all the while , till at last it sat down a good wa
y off , panting , with its tongue hanging out of its mouth , and its gr
eat eyes half shut .

Length of the above sentence:  944
Longest sentence between two consecutive periods:

 London is the capital of Paris , and Paris is the capital of Rome , an
d Rome -- no , THAT ' S all wrong , I ' m certain ! I must have been ch
anged for Mabel ! I ' ll try and say " How doth the little --"' and she
crossed her hands on her lap as if she were saying lessons , and began
to repeat it , but her voice sounded hoarse and strange , and the words
did not come the same as they used to do :-- ' How doth the little croc
odile Improve his shining tail , And pour the waters of the Nile On eve
ry golden scale ! ' How cheerfully he seems to grin , How neatly spread
his claws , And welcome little fishes in With gently smiling jaws !' '
I ' m sure those are not the right words ,' said poor Alice , and her e
yes filled with tears again as she went on , ' I must be Mabel after al
l , and I shall have to go and live in that poky little house , and hav
e next to no toys to play with , and oh ! ever so many lessons to learn
! No , I ' ve made up my mind about it ; if I ' m Mabel , I ' ll stay d
own here ! It ' ll be no use their putting their heads down and saying
" Come up again , dear !" I shall only look up and say " Who am I then
? Tell me that first , and then , if I like being that person , I ' ll
come up : if not , I ' ll stay down here till I ' m somebody else "-- b
ut , oh dear !' cried Alice , with a sudden burst of tears , ' I do wis
h they WOULD put their heads down ! I am so VERY tired of being all alo
ne here !' As she said this she looked down at her hands , and was surp
rised to see that she had put on one of the Rabbit ' s little white kid
gloves while she was talking

Length of the above sentence:  1595
```

In [27]:
```
wrd = gutenberg.words('bryant-stories.txt')
book = ''
for word in wrd:
    book = book + word + ' '
sentences = sent_tokenize(book)
word_count = lambda sentence: len(word_tokenize(sentence))
```

In [28]:
```
print("\nName of the corpus is: bryant-stories.txt")
longestsent = max(sentences, key=word_count)
print('\nLongest sentence found using sentence tokenizer: \n\n', longest
sent)
print('\nLength of the above sentence: ', len(longestsent))
sentences = book.split(". ")
longest_sent = max(sentences, key=len)
print('\nLongest sentence between two consecutive periods: \n\n', longes
t_sent)
print('\nLength of the above sentence: ', len(longest_sent))
```

```
Name of the corpus is: bryant-stories.txt

Longest sentence found using sentence tokenizer:

 David looked to see what the trouble was , and he saw a strange sight
: down the slope of the opposite mountain came striding a Philistine wa
rrior , calling out something in a taunting voice ; he was a gigantic m
an , the largest David had ever seen , and he was covered with armour ,
that shone in the sun : he had a helmet of brass upon his head , and he
was armed with a coat of mail , and he had greaves of brass upon his le
gs , and a target of brass between his shoulders ; his spear was so tre
mendous that the staff of it was like a weaver ' s beam , and his shiel
d so great that a man went before him , to carry it . "

Length of the above sentence:  624

Longest sentence between two consecutive periods:

 When David came before Saul , he said to the king , " Let no man ' s h
eart fail because of him ; thy servant will go and fight with this Phil
istine ." But Saul looked at David , and said , " Thou art not able to
go against this Philistine , to fight with him , for thou art but a you
th , and he has been a man of war from his youth ." Then David said to
Saul , " Once I was keeping my father ' s sheep , and there came a lion
, and a bear , and took a lamb out of the flock ; and I went out after
the lion , and struck him ; and delivered the lamb out of his mouth , a
nd when he arose against me , I caught him by the beard , and struck hi
m , and slew him ! Thy servant slew both the lion and the bear ; and th
is Philistine shall be as one of them , for he hath defied the armies o
f the living God

Length of the above sentence:  798
```

In [29]:
```
wrd = gutenberg.words('burgess-busterbrown.txt')
book = ''
for word in wrd:
    book = book + word + ' '
sentences = sent_tokenize(book)
word_count = lambda sentence: len(word_tokenize(sentence))
longestsent = max(sentences, key=word_count)
```

In [30]:
```
print("\nName of the corpus is: burgess-busterbrown.txt")
print('\nLongest sentence found using sentence tokenizer: \n\n', longest
sent)
print('\nLength of the above sentence: ', len(longestsent))
sentences = book.split(". ")
longest_sent = max(sentences, key=len)
print('\nLongest sentence between two consecutive periods: \n\n', longes
t_sent)
print('\nLength of the above sentence: ', len(longest_sent))
```

```
Name of the corpus is: burgess-busterbrown.txt

Longest sentence found using sentence tokenizer:

 Sammy Jay looked at Blacky the Crow , and Blacky looked at Chatterer ,
and Chatterer looked at Happy Jack , and Happy Jack looked at Peter Rab
bit , and Peter looked at Unc ' Billy Possum , and Unc ' Billy looked a
t Bobby Coon , and Bobby looked at Johnny Chuck , and Johnny looked at
Reddy Fox , and Reddy looked at Jimmy Skunk , and Jimmy looked at Billy
Mink , and Billy looked at Little Joe Otter , and for a minute nobody c
ould say a word .

Length of the above sentence:  444

Longest sentence between two consecutive periods:

 And so they reached the edge of the Green Forest , Buster Bear running
as hard as ever he could , Sammy Jay flying just behind him and screami
ng , " Thief , thief , thief !" at the top of his lungs , and behind hi
m Blacky the Crow , trying to catch up and yelling as loud as he could
, " Caw , caw , caw ! Come on , everybody ! Come on ! Come on !" Poor B
uster ! It was bad enough to be frightened almost to death as he had be
en up in the Old Pasture when the pail had caught over his head just as
Farmer Brown ' s boy had yelled at him

Length of the above sentence:  537
```

The above outputs show that both approaches show different outputs with different lengths and no resemblance

In [ ]:
```
```