In [1]:
```
!pip install NLTK
```

Requirement already satisfied: NLTK in c:\intel\python310\lib\site-packages (3.
7)
Requirement already satisfied: click in c:\intel\python310\lib\site-packages (f
rom NLTK) (8.1.3)
Requirement already satisfied: tqdm in c:\intel\python310\lib\site-packages (fr
om NLTK) (4.64.1)
Requirement already satisfied: joblib in c:\intel\python310\lib\site-packages
(from NLTK) (1.1.0)
Requirement already satisfied: regex>=2021.8.3 in c:\intel\python310\lib\site-p
ackages (from NLTK) (2022.9.13)
Requirement already satisfied: colorama in c:\intel\python310\lib\site-packages
(from click->NLTK) (0.4.5)

WARNING: You are using pip version 22.0.4; however, version 22.3 is available.
You should consider upgrading via the 'C:\Intel\python310\python.exe -m pip ins
tall --upgrade pip' command.

In [2]:
```
import nltk
```

In [3]:
```
nltk.download()
```

showing info https://raw.githubusercontent.com/nltk/nltk_data/gh-pages/index.xm
l (https://raw.githubusercontent.com/nltk/nltk_data/gh-pages/index.xml)

Out[3]: True

In [5]:
```
from urllib import request
url="http://www.gutenberg.org/files/1342/1342-0.txt"
response=request.urlopen(url)
```

In [6]:
```
raw=response.read().decode('utf8')
```

```
In [7]: print(raw[0:75])
        print(raw[-1000:])
```

The Project Gutenberg eBook of Pride and prejudice, by Jane Austen

This
e

Section 5. General Information About Project Gutenberg-tm electronic works

Professor Michael S. Hart was the originator of the Project
Gutenberg-tm concept of a library of electronic works that could be
freely shared with anyone. For forty years, he produced and
distributed Project Gutenberg-tm eBooks with only a loose network of
volunteer support.

Project Gutenberg-tm eBooks are often created from several printed
editions, all of which are confirmed as not protected by copyright in
the U.S. unless a copyright notice is included. Thus, we do not
necessarily keep eBooks in compliance with any particular paper
edition.

Most people start at our website which has the main PG search
facility: www.gutenberg.org

This website includes information about Project Gutenberg-tm,
including how to make donations to the Project Gutenberg Literary
Archive Foundation, how to help produce our new eBooks, and how to
subscribe to our email newsletter to hear about new eBooks.

```
In [8]: start=raw.find("It is a truth")
        end=raw.find("wife.")
        print(raw[start:end+5])
```

It is a truth universally acknowledged, that a single man in possession
of a good fortune must be in want of a wife.

```
In [9]: end=raw.find("End the Projects")
        print(raw[end-150:end])
```

t Gutenberg Literary
Archive Foundation, how to help produce our new eBooks, and how to
subscribe to our email newsletter to hear about new eBooks.

```
In [10]: start=raw.find("It is truth")
         end=raw.find("End the project")
         body=raw[start:end]
         body=body.strip()
         len(body)
         print(body[:80],"[ -------- ],body[-80:]")
```

 [-------],body[-80:]

In [11]:
```python
words=body.split()
print(len(words))
print(words[0:10])
```

```
0
[]
```

In [ ]: