In [1]:
```
pip install nltk
```

```
Requirement already satisfied: nltk in c:\users\shree\anaconda3\lib\site-packages (3.6.1)
Requirement already satisfied: regex in c:\users\shree\anaconda3\lib\site-packages (from nltk) (2021.4.4)
Requirement already satisfied: click in c:\users\shree\anaconda3\lib\site-packages (from nltk) (7.1.2)
Requirement already satisfied: tqdm in c:\users\shree\anaconda3\lib\site-packages (from nltk) (4.59.0)
Note: you may need to restart the kernel to use updated packages.
Requirement already satisfied: joblib in c:\users\shree\anaconda3\lib\site-packages (from nltk) (1.0.1)
```

In [6]:
```
import nltk
nltk.download()
```

```
showing info https://raw.githubusercontent.com/nltk/nltk_data/gh-pages/index.xml
```

Out[6]:
```
True
```

In [ ]:
```
pip install gensim
```

In [ ]:
```
pip install pattern
```

In [7]:
```
#Tokenizing text data
from nltk.tokenize import sent_tokenize, \
        word_tokenize, WordPunctTokenizer
```

In [12]:
```
# Define input text
input_text = "Do you know how tokenization works? It's actually  quite interesting! Let's analyze a couple of sentences and  figure it out."
```

In [13]:
```
#Divide the input text into sentence tokens:
# Sentence tokenizer
print("\nSentence tokenizer:")
print(sent_tokenize(input_text))
```

```
Sentence tokenizer:
['Do you know how tokenization works?', "It's actually  quite interesting!", "Let's analyze a couple of sentences and  figure it out."]
```

In [14]:
```
#Divide the input text into word tokens:
# Word tokenizer
print("\nWord tokenizer:")
print(word_tokenize(input_text))
```

```
Word tokenizer:
['Do', 'you', 'know', 'how', 'tokenization', 'works', '?', 'It', "'s", 'actually', 'quite', 'interesting', '!', 'Let', "'s", 'analyze', 'a', 'couple', 'of',
'sentences', 'and', 'figure', 'it', 'out', '.']
```

In [17]:
```
#Divide the input text into word tokens using the WordPunct tokenizer:
# WordPunct tokenizer
print("\nWord punct tokenizer:")
print(WordPunctTokenizer().tokenize(input_text))
```

```
Word punct tokenizer:
['Do', 'you', 'know', 'how', 'tokenization', 'works', '?', 'It', "'", 's', 'actually', 'quite', 'interesting', '!', 'Let', "'", 's', 'analyze', 'a', 'couple',
'of', 'sentences', 'and', 'figure', 'it', 'out', '.']
```

In [ ]:
```
```