In [1]:
```
import nltk
from nltk.tokenize import word_tokenize
```

In [2]:
```
#1
text = input("Enter text : ")
tokens = word_tokenize(text)
print("Total words : ", len(tokens))
```

```
Enter text : Hi, I love chocolate ice cream.
Total words :  8
```

In [3]:
```
#2
text = input("Enter text : ")
tokens = set(word_tokenize(text))
tokens
tag = nltk.pos_tag(tokens)
print("Different words : ", len(tokens))
tag
```

```
Enter text : Hi, I love chocolate ice cream.
Different words :  8
```

Out[3]:
```
[('ice', 'NN'),
 ('I', 'PRP'),
 ('Hi', 'NNP'),
 (',', ','),
 ('love', 'VB'),
 ('chocolate', 'NN'),
 ('.', '.'),
 ('cream', 'NN')]
```

In [4]:
```
#3
text = input("Enter text : ")
words = nltk.word_tokenize(text)
fdist = nltk.FreqDist(words)
fdist
```

```
Enter text : Hi, I love chocolate ice cream.
```

Out[4]:
```
FreqDist({'Hi': 1, ',': 1, 'I': 1, 'love': 1, 'chocolate': 1, 'ice': 1, 'cream': 1, '.': 1})
```

In [5]:
```
#4
text = input("Enter text : ")
words = nltk.word_tokenize(text)
fdist = nltk.FreqDist(words)
fdist
for i in fdist:
    print(i + ' :', (fdist[i] / len(words)) * 100)
```

```
Enter text : Hi, I love chocolate ice cream.
Hi : 12.5
, : 12.5
I : 12.5
love : 12.5
chocolate : 12.5
ice : 12.5
cream : 12.5
. : 12.5
```