In [1]:

```
pip install nltk
```

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: nltk in c:\programdata\anaconda3\lib\site-packages (3.7)
Requirement already satisfied: regex>=2021.8.3 in c:\programdata\anaconda3\lib\site-packages
(from nltk) (2022.7.9)
Requirement already satisfied: click in c:\programdata\anaconda3\lib\site-packages (from nltk)
(8.0.4)
Requirement already satisfied: joblib in c:\programdata\anaconda3\lib\site-packages (from nlt
k) (1.1.0)
Requirement already satisfied: tqdm in c:\programdata\anaconda3\lib\site-packages (from nltk)
(4.64.1)
Requirement already satisfied: colorama in c:\programdata\anaconda3\lib\site-packages (from cl
ick->nltk) (0.4.5)
Note: you may need to restart the kernel to use updated packages.
```

In [1]:

```
import nltk
nltk.download()
```

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

Out[1]:

```
True
```

In [3]:

```
from nltk.tokenize import sent_tokenize
nltk.download('punkt')
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\ADMIN\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

Out[3]:

```
True
```

In [6]:

```
txt="Find someting you love to do and you'll never have to work a day in your life. Don't be afrid to learn.
ts=sent_tokenize(txt)
ts
```

Out[6]:

```
["Find someting you love to do and you'll never have to work a day in your life.",
 "Don't be afrid to learn.",
 'Knowledge is wieghtless, a treasure you can always carry easily.']
```

In [8]:

```
from nltk.tokenize import word_tokenize
tw=word_tokenize(txt)
tw
```

Out[8]:

```
['Find',
 'someting',
 'you',
 'love',
 'to',
 'do',
 'and',
 'you',
 "'ll",
 'never',
 'have',
 'to',
 'work',
 'a',
 'day',
 'in',
 'your',
 'life',
 '.',
 'Do',
 "n't",
 'be',
 'afrid',
 'to',
 'learn',
 '.',
 'Knowledge',
 'is',
 'wieghtless',
 ',',
 'a',
 'treasure',
 'you',
 'can',
 'always',
 'carry',
 'easily',
 '.']
```

In [9]:

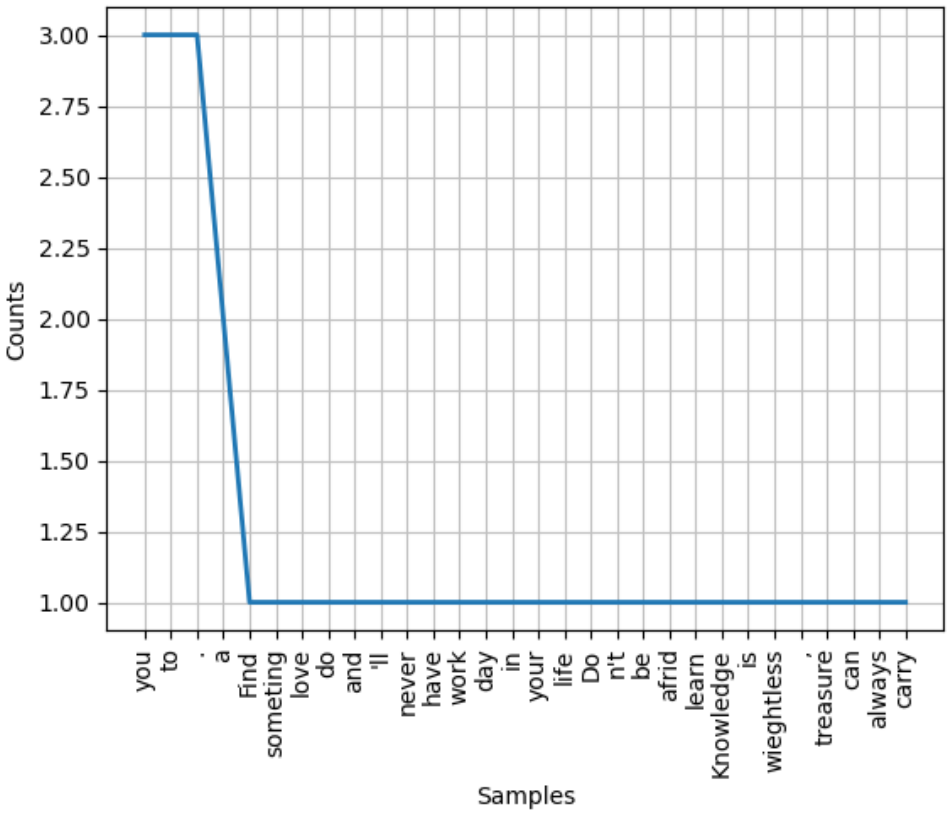
```
from nltk.probability import FreqDist
fdist=FreqDist(tw)
fdist
```

Out[9]:

```
FreqDist({'you': 3, 'to': 3, '.': 3, 'a': 2, 'Find': 1, 'someting': 1, 'love': 1, 'do': 1, 'an
d': 1, "'ll": 1, ...})
```

In [10]:

```
import matplotlib.pylab as plt
fdist.plot(30,cumulative=False)
plt.show
```



Out[10]:

```
<function matplotlib.pyplot.show(close=None, block=None)>
```

In [11]:

```
from nltk.corpus import stopwords
stw=set(stopwords.words("english"))
stw
```

```
 'just',
 'll',
 'm',
 'ma',
 'me',
 'mightn',
 "mightn't",
 'more',
 'most',
 'mustn',
 "mustn't",
 'my',
 'myself',
 'needn',
 "needn't",
 'no',
 'nor',
 'not',
 'now',
 'o',
```

In [14]:

```
filtsent=[]
for q in tw:
    if q not in stw:
        filtsent.append(q)
print('tokenized sentence ',tw)
print('filtered sentence ', filtsent)
```

```
tokenized sentence  ['Find', 'someting', 'you', 'love', 'to', 'do', 'and', 'you', "'ll", 'neve
r', 'have', 'to', 'work', 'a', 'day', 'in', 'your', 'life', '.', 'Do', "n't", 'be', 'afrid',
'to', 'learn', '.', 'Knowledge', 'is', 'wieghtless', ',', 'a', 'treasure', 'you', 'can', 'alwa
ys', 'carry', 'easily', '.']
filtered sentence  ['Find', 'someting', 'love', "'ll", 'never', 'work', 'day', 'life', '.', 'D
o', "n't", 'afrid', 'learn', '.', 'Knowledge', 'wieghtless', ',', 'treasure', 'always', 'carr
y', 'easily', '.']
```

In [15]:

```
from nltk.stem import PorterStemmer
from nltk.tokenize import sent_tokenize,word_tokenize
ps=PorterStemmer()
stmw=[]
for w in filtsent:
    stmw.append(ps.stem(w))
print("Filtered Sentence:",filtsent)
print("Stemmed Sentence:",stmw)
```

```
Filtered Sentence: ['Find', 'someting', 'love', "'ll", 'never', 'work', 'day', 'life', '.', 'D
o', "n't", 'afrid', 'learn', '.', 'Knowledge', 'wieghtless', ',', 'treasure', 'always', 'carr
y', 'easily', '.']
Stemmed Sentence: ['find', 'somet', 'love', "'ll", 'never', 'work', 'day', 'life', '.', 'do',
"n't", 'afrid', 'learn', '.', 'knowledg', 'wieghtless', ',', 'treasur', 'alway', 'carri', 'eas
ili', '.']
```

In [22]:

```
from nltk.stem.wordnet import WordNetLemmatizer
lem=WordNetLemmatizer()
for y in filtsent:
    print("Lemmatized Word:",lem.lemmatize(y))
    print("Stemmed Word:",ps.stem(y))
```

```
Lemmatized Word: Find
Stemmed Word: find
Lemmatized Word: someting
Stemmed Word: somet
Lemmatized Word: love
Stemmed Word: love
Lemmatized Word: 'll
Stemmed Word: 'll
Lemmatized Word: never
Stemmed Word: never
Lemmatized Word: work
Stemmed Word: work
Lemmatized Word: day
Stemmed Word: day
Lemmatized Word: life
Stemmed Word: life
Lemmatized Word: .
Stemmed Word: .
Lemmatized Word: Do
Stemmed Word: do
Lemmatized Word: n't
Stemmed Word: n't
Lemmatized Word: afrid
Stemmed Word: afrid
Lemmatized Word: learn
Stemmed Word: learn
Lemmatized Word: .
Stemmed Word: .
Lemmatized Word: Knowledge
Stemmed Word: knowledg
Lemmatized Word: wieghtless
Stemmed Word: wieghtless
Lemmatized Word: ,
Stemmed Word: ,
Lemmatized Word: treasure
Stemmed Word: treasur
Lemmatized Word: always
Stemmed Word: alway
Lemmatized Word: carry
Stemmed Word: carri
Lemmatized Word: easily
Stemmed Word: easili
Lemmatized Word: .
Stemmed Word: .
```

In [26]:

```
nltk.pos_tag(filtsent)
```

Out[26]:

```
[('Find', 'IN'),
 ('someting', 'VBG'),
 ('love', 'NN'),
 ("'ll", 'MD'),
 ('never', 'RB'),
 ('work', 'VB'),
 ('day', 'NN'),
 ('life', 'NN'),
 ('.', '.'),
 ('Do', 'VBP'),
 ("n't", 'RB'),
 ('afrid', 'VB'),
 ('learn', 'VB'),
 ('.', '.'),
 ('Knowledge', 'NNP'),
 ('wieghtless', 'NN'),
 (',', ','),
 ('treasure', 'NN'),
 ('always', 'RB'),
 ('carry', 'VBP'),
 ('easily', 'RB'),
 ('.', '.')]
```

In [ ]:

```

```