**Exercise 1. Frequency Distribution using Histograms**

**Program 1.1: Simple Frequency Distribution**

**Aim:**

To generate a simple frequency distribution plot using NLTK.

**Procedure:**

1. Import required libraries: nltk, matplotlib.pyplot.

2. Tokenize the text into words using TreebankWordTokenizer

3. Compute the frequency distribution of the words using FreqDist.

4. Plot the frequency distribution using matplotlib.

**Code:**

import nltk

from nltk.tokenize import TreebankWordTokenizer

from nltk.probability import FreqDist

import matplotlib.pyplot as plt

*# Initialize tokenizer*

tokenizer = TreebankWordTokenizer()

*# Sample text*

text = "Natural Language Processing (NLP) with Python is fun and educational. Python

simplifies NLP tasks."

*# Tokenize text*

tokens = tokenizer.tokenize(text)

*# Print tokens*

print("Tokens:", tokens)

**Output:**

****

*# Frequency Distribution*

fdist = FreqDist(tokens)

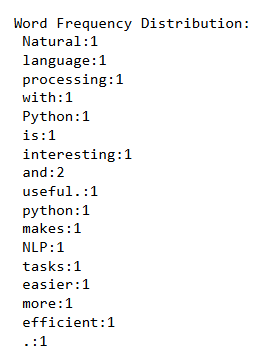
*# Print frequencies*

print("\nWord Frequency Distribution:")

for word, freq in fdist.items():

print(f"{word}: {freq}")

**Output**

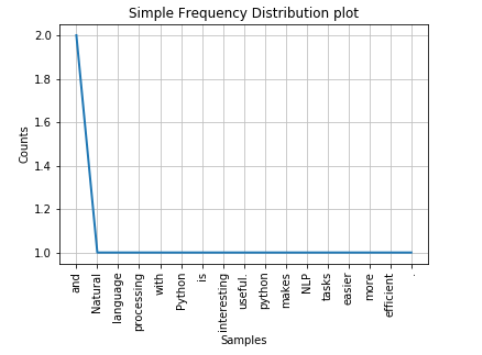
****

*# Plot frequency distribution*

fdist.plot(title='Simple Frequency Distribution Plot')

plt.show()

**Output:**

****

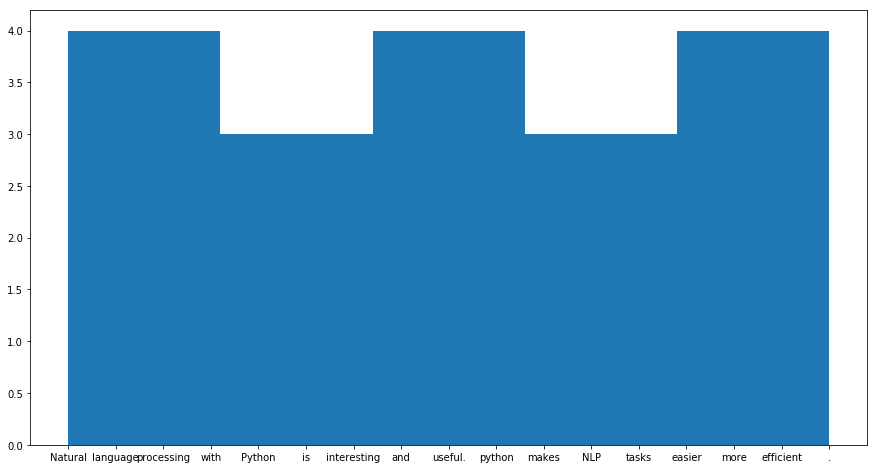
*#Histogram*

plt.figure(figsize=(15,8))

plt.hist(tokens,bins=5)

plt.show()

**Output**

****

**Result:**

Frequency distribution of words plotted successfully.