Shri Ramdeobaba College of Engineering and Management, Nagpur

Department of Computer Science and Engineering

Natural Language Processing Lab

Name : Shantanu Mane
Branch : CSE - AIML (VI^{th} SEM)
Roll Num : E-63

AIM :

1. Write a Python NLTK program to find the number of male and female names in the "names" corpus.Print the first 10 male and female names. Note: The names corpus contains a total of around 2943 male (male.txt) and 5001 female (female.txt) names.
2. Write a Python NLTK program to print the first 15 random combine labeled male and labeled female names from names corpus.

1. Importing the Dependencies

```
import random

from nltk.corpus import names
```

Part A:

A.1 Analyzing the Names Corpus

```
names.fileids()
```

```
['female.txt', 'male.txt']
```

```
male_names = names.words('male.txt')
female_names = names.words('female.txt')
```

A.1.1 Number of Male and Female Names

```
print(f"total male names : {len(male_names)}\n"
      f"\ntotal female names : {len(female_names)}")
```

```
total male names : 2943

total female names : 5001
```

A.1.2 First 10 Male and Female Names

```
print(
    f"first 10 male names in the names corpus : {male_names[:10]}\n"
    f"\nfirst 10 female names in the names corpus : {female_names[:10]}"
)
```

```
first 10 male names in the names corpus : ['Aamir', 'Aaron', 'Abbey', 'Abbie', 'Abbot', 'Abbott', 'Abby', 'Abdel', 'Abdul',
'Abdulkarim']

first 10 female names in the names corpus : ['Abagael', 'Abagail', 'Abbe', 'Abbey', 'Abbi', 'Abbie', 'Abby', 'Abigael',
'Abigail', 'Abigale']
```

Part B

B.1 Getting 15 Random Male and Female Names

```
print(f"15 random male names : {random.sample(male_names, 15)}")
```

```
15 random male names : ['Foster', 'Thibaud', 'Dabney', 'Edward', 'Lay', 'Griffith', 'Levin', 'Grover', 'Philip', 'Jamey', 'Douglass', 'Jay', 'Sanderson', 'Flynn', 'Nathanael']
```

```
print(f"15 random female names : {random.sample(female_names, 15)}")
```

```
15 random female names : ['Vi', 'Donetta', 'Lauren', 'Laverne', 'Traci', 'Christa', 'Delinda', 'Mirella', 'Ruthann', 'Dori', 'Hildy', 'Julee', 'Jany', 'Evette', 'Karrah']
```

```

```