ARYAN SHARMA 21BBS0106

Vellore Institute of Technology (VIT), Vellore

Evening Slot (6P.M to 8P.M)

Assignment - 1

Task 1 :

Create a pandas dataframe (DataFrame name as 'df') with numpy random values (4 features and 4 observation)

```
import pandas as pd
import numpy as np

data = np.random.randn(4, 4)
df = pd.DataFrame(data)
print("Task 1: DataFrame:")
print(df)
```

```
Task 1: DataFrame:
          0         1         2         3
0 -0.425566  2.362216  0.375277  0.342605
1  0.409795 -0.133370  0.206561  0.660088
2  0.400739 -1.137500  1.654505 -1.044264
3  0.828268  1.362428 -0.716839  0.362412
```

Task - 2:

Rename the task - 1 'df dataframe column names to 'Random value 1', 'Random value 2', 'Random value 3' & 'Random value 4'

```
df.columns = ['Random value 1', 'Random value 2', 'Random value 3',
'Random value 4']
print("Task 2: Renamed columns:")
print(df)
```

```
Task 2: Renamed columns:
   Random value 1  Random value 2  Random value 3  Random value 4
0       -0.425566        2.362216        0.375277        0.342605
1        0.409795       -0.133370        0.206561        0.660088
2        0.400739       -1.137500        1.654505       -1.044264
3        0.828268        1.362428       -0.716839        0.362412
```

Task 3

Find the descriptive statistics of the 'df' dataframe.

```
descriptive_stats = df.describe()
print("Task 3: Descriptive Statistics:")
print(descriptive_stats)
```

```
Task 3: Descriptive Statistics:
       Random value 1  Random value 2  Random value 3  Random value 4
count        4.000000        4.000000        4.000000        4.000000
mean         0.303309        0.613443        0.379876        0.080210
std          0.525253        1.553782        0.975965        0.763586
min         -0.425566       -1.137500       -0.716839       -1.044264
25%          0.194163       -0.384403       -0.024289       -0.004112
50%          0.405267        0.614529        0.290919        0.352508
75%          0.514413        1.612375        0.695084        0.436831
max          0.828268        2.362216        1.654505        0.660088
```

Task 4

Check for the null values in 'df' and find the data type of the columns.

```
null_values = df.isnull().values.any()
data_types = df.dtypes
print("Task 4: Null Values and Data Types:")
print(null_values)
print(data_types)
```

```
Task 4: Null Values and Data Types:
False
Random value 1    float64
Random value 2    float64
Random value 3    float64
Random value 4    float64
dtype: object
```

Task - 5

Display the 'Random value 2' & 'Random value 3' columns with location method and index location method.

```
random_value_loc = df.loc[:, 'Random value 2':'Random value 3']
random_value_idx = df.iloc[:, 1:3]
print("Task 5: 'Random value 2' & 'Random value 3' Columns:")
print("Using loc:")
print(random_value_loc)
print("\nUsing iloc:")
print(random_value_idx)
```

```
Task 5: 'Random value 2' & 'Random value 3' Columns:
Using loc:
```

```
   Random value 2  Random value 3
0        2.362216        0.375277
1       -0.133370        0.206561
2       -1.137500        1.654505
3        1.362428       -0.716839

Using iloc:
   Random value 2  Random value 3
0        2.362216        0.375277
1       -0.133370        0.206561
2       -1.137500        1.654505
3        1.362428       -0.716839
```