Name: Rishit Maheshwari

Registration No.: 21BCE0789

Campus: VIT Vellore

Task 1 - Create a pandas dataframe (DataFrame name as 'df') with numpy random values (4 features and 4 observation)

Code –

```
import pandas as pd

import numpy as np

# Set the random seed for reproducibility

np.random.seed(42)

# Create random data

data = np.random.rand(4, 4)

# Create a pandas DataFrame

df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

print(df)
```

Output –

```
In [1]: runfile('C:/Users/DELL/OneDrive/Desktop/python_exe/untitled15.py', wdir='C:/Users/DELL/OneDrive/
Desktop/python_exe')
   Feature1  Feature2  Feature3  Feature4
0  0.374540  0.950714  0.731994  0.598658
1  0.156019  0.155995  0.058084  0.866176
2  0.601115  0.708073  0.020584  0.969910
3  0.832443  0.212339  0.181825  0.183405
```

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

Code –

```
import pandas as pd

import numpy as np

# Set the random seed for reproducibility

np.random.seed(42)

# Create random data

data = np.random.rand(4, 4)

# Create a pandas DataFrame

df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3',
'Feature4'])

# Rename the columns

df.columns = ['Random value 1', 'Random value 2', 'Random value 3',
'Random value 4']

print(df)
```

```python
import pandas as pd
import numpy as np

# Set the random seed for reproducibility
np.random.seed(42)

# Create random data
data = np.random.rand(4, 4)

# Create a pandas DataFrame
df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

# Rename the columns
df.columns = ['Random value 1', 'Random value 2', 'Random value 3', 'Random value 4']

print(df)
```

Output –

```
In [2]: runfile('C:/Users/DELL/OneDrive/Desktop/python_exe/untitled17.py', wdir='C:/Users/DELL/OneDrive/
Desktop/python_exe')
   Random value 1  Random value 2  Random value 3  Random value 4
0        0.374540        0.950714        0.731994        0.598658
1        0.156019        0.155995        0.058084        0.866176
2        0.601115        0.708073        0.020584        0.969910
3        0.832443        0.212339        0.181825        0.183405
```

Task 3- Find the descriptive statistics of the 'df' dataframe.

Code –

import pandas as pd

import numpy as np

Set the random seed for reproducibility

np.random.seed(42)

Create random data

data = np.random.rand(4, 4)

Create a pandas DataFrame

df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

Rename the columns

df.columns = ['Random value 1', 'Random value 2', 'Random value 3', 'Random value 4']

Get descriptive statistics

descriptive_stats = df.describe()

print(descriptive_stats)

```
import pandas as pd
import numpy as np

# Set the random seed for reproducibility
np.random.seed(42)

# Create random data
data = np.random.rand(4, 4)

# Create a pandas DataFrame
df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

# Rename the columns
df.columns = ['Random value 1', 'Random value 2', 'Random value 3', 'Random value 4']

# Get descriptive statistics
descriptive_stats = df.describe()

print(descriptive_stats)
```

Output –

```
In [3]: runfile('C:/Users/DELL/OneDrive/Desktop/python_exe/untitled17.py', wdir='C:/Users/DELL/OneDrive/
Desktop/python_exe')
       Random value 1  Random value 2  Random value 3  Random value 4
count        4.000000        4.000000        4.000000        4.000000
mean         0.491029        0.506780        0.248122        0.654537
std          0.291252        0.386153        0.329856        0.350875
min          0.156019        0.155995        0.020584        0.183405
25%          0.319910        0.198253        0.048709        0.494845
50%          0.487828        0.460206        0.119954        0.732417
75%          0.658947        0.768733        0.319367        0.892110
max          0.832443        0.950714        0.731994        0.969910
```

Task 4 - Check for the null values in 'df' and find the data type of the columns.

Code –

```
import pandas as pd

import numpy as np

# Set the random seed for reproducibility

np.random.seed(42)

# Create random data

data = np.random.rand(4, 4)

# Create a pandas DataFrame

df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3',
'Feature4'])

# Rename the columns

df.columns = ['Random value 1', 'Random value 2', 'Random value 3',
'Random value 4']

# Check for null values

null_values = df.isnull().sum()

# Find data types of columns

data_types = df.dtypes

print("Null Values:")

print(null_values)

print("\nData Types:")

print(data_types)
```

```python
import pandas as pd
import numpy as np

# Set the random seed for reproducibility
np.random.seed(42)

# Create random data
data = np.random.rand(4, 4)

# Create a pandas DataFrame
df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

# Rename the columns
df.columns = ['Random value 1', 'Random value 2', 'Random value 3', 'Random value 4']

# Check for null values
null_values = df.isnull().sum()

# Find data types of columns
data_types = df.dtypes

print("Null Values:")
print(null_values)

print("\nData Types:")
print(data_types)
```

Output –

```
In [4]: runfile('C:/Users/DELL/OneDrive/Desktop/python_exe/untitled18.py', wdir='C:/Users/DELL/OneDrive/
Desktop/python_exe')
Null Values:
Random value 1    0
Random value 2    0
Random value 3    0
Random value 4    0
dtype: int64

Data Types:
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.

Code –

```
import pandas as pd

import numpy as np

# Set the random seed for reproducibility

np.random.seed(42)

# Create random data

data = np.random.rand(4, 4)

# Create a pandas DataFrame

df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3',
'Feature4'])

# Rename the columns

df.columns = ['Random value 1', 'Random value 2', 'Random value 3',
'Random value 4']

# Display 'Random value 2' and 'Random value 3' columns using loc[]

loc_result = df.loc[:, ['Random value 2', 'Random value 3']]

# Display 'Random value 2' and 'Random value 3' columns using iloc[]

iloc_result = df.iloc[:, [1, 2]]

print("Using loc[]:")

print(loc_result)

print("\nUsing iloc[]:")

print(iloc_result)
```

```python
import pandas as pd
import numpy as np

# Set the random seed for reproducibility
np.random.seed(42)

# Create random data
data = np.random.rand(4, 4)

# Create a pandas DataFrame
df = pd.DataFrame(data, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

# Rename the columns
df.columns = ['Random value 1', 'Random value 2', 'Random value 3', 'Random value 4']

# Display 'Random value 2' and 'Random value 3' columns using loc[]
loc_result = df.loc[:, ['Random value 2', 'Random value 3']]

# Display 'Random value 2' and 'Random value 3' columns using iloc[]
iloc_result = df.iloc[:, [1, 2]]

print("Using loc[]:")
print(loc_result)

print("\nUsing iloc[]:")
print(iloc_result)
```

Output –

```
In [5]: runfile('C:/Users/DELL/OneDrive/Desktop/python_exe/untitled19.py', wdir='C:/Users/DELL/OneDrive/
Desktop/python_exe')
Using loc[]:
   Random value 2  Random value 3
0        0.950714        0.731994
1        0.155995        0.058084
2        0.708073        0.020584
3        0.212339        0.181825

Using iloc[]:
   Random value 2  Random value 3
0        0.950714        0.731994
1        0.155995        0.058084
2        0.708073        0.020584
3        0.212339        0.181825
```