SmartBridge Externship

Artificial Intelligence

Assignment-1

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

In [1]:
```
import pandas as pd
import numpy as np

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

# Create a 4x4 array of random values
data = np.random.rand(4, 4)

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

# Print the DataFrame
print(df)
```

```
   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'

In [3]:
```
# Rename the columns
new_columns = {
    'Feature1': 'Random value 1',
    'Feature2': 'Random value 2',
    'Feature3': 'Random value 3',
    'Feature4': 'Random value 4'
}

df = df.rename(columns=new_columns)
print(df)
```

```
   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.

```
In [4]:
# Get descriptive statistics
statistics = df.describe()

# Print the descriptive statistics
print(statistics)
```

```
       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.

```
In [5]:
# Check for null values
null_values = df.isnull().sum()

# Get the data types of columns
column_types = df.dtypes

# Print the null values and data types
print("Null values:\n", null_values)
print("\nData types:\n", column_types)
```

```
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.

In [6]:

```
# Display columns using loc method
loc_columns = df.loc[:, ['Random value 2', 'Random value 3']]
print("Columns using loc method:\n", loc_columns)

# Display columns using iloc method
iloc_columns = df.iloc[:, [1, 2]]
print("\nColumns using iloc method:\n", iloc_columns)
```

```
Columns using loc method:
    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

Columns using iloc method:
    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
```