Smart Bridge Assignment - 1

Name: Kuchibhatla Mohan Datta

Reg_No : 20MID0012

Q1. Create a pandas dataframe (DataFrame name as 'df') with numpy random values (4 features and 4 observation)

```
import numpy as np
import pandas as pd

# Set random 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=['Feature 1', 'Feature 2', 'Feature 3', 'Feature 4'])

# Display the DataFrame
print(df)
```

```
   Feature 1  Feature 2  Feature 3  Feature 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
```

Q2. Rename the task - 1 'df' dataframe column names to 'Random value 1', 'Random value 2', 'Random value 3' & 'Random value 4'

```
df.rename(columns = {'Feature 1':'Random value 1','Feature 2':'Random value 2','Feature 3':'Random value 3','Feature 4':'Random value 4'}

print("\nAfter modifying first column:\n", df.columns)
print(df)
```

```
After modifying first column:
 Index(['Random value 1', 'Random value 2', 'Random value 3', 'Random value 4'], dtype='object')
   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
```

Q3.Find the descriptive statistics of the 'df' dataframe.

```
# Descriptive stat

df.describe(include='all')
```

	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

Q4. Check for the null values in 'df' and find the data type of the columns.

```
# Checking null values
print(df.isnull())
#print(df.isnull().all())
#print(df.isnull().sum())

# For knowing datatype of each column

print(df.dtypes)

print(df.info())
```

```
   Random value 1  Random value 2  Random value 3  Random value 4
0           False           False           False           False
1           False           False           False           False
2           False           False           False           False
3           False           False           False           False
Random value 1    float64
Random value 2    float64
Random value 3    float64
Random value 4    float64
dtype: object
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4 entries, 0 to 3
Data columns (total 4 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Random value 1  4 non-null      float64
 1   Random value 2  4 non-null      float64
 2   Random value 3  4 non-null      float64
 3   Random value 4  4 non-null      float64
dtypes: float64(4)
memory usage: 256.0 bytes
None
```

Q5. Display the 'Random value 2' & 'Random value 3' columns with location method and index location method.

```
# Location (loc) & Index location (iloc) ---> Syntax = loc[:,:], iloc[:,:]

# With Location Method

# give the actual row and column name as a parameter
print(df)

print(df.loc[:,'Random value 2':'Random value 3'])

# With Iloc Method

# give the actual row and column index value as a parameter

df.iloc[:,1:3]
```

```
   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
   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
```

	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