Assignment 1

Name: Dereddy Venkata Yogitha

Registration Number: 21BAI1645

Email ID: dereddyvenkata.yogitha2021@vitstudent.ac.in

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

```
import pandas as pd
import numpy as np

data = {
    'Age': np.random.randint(18, 65, 4),
    'Income': np.random.randint(30000, 80000, 4),
    'Experience': np.random.uniform(0.5, 20, 4),
    'CustomerSatisfaction': np.random.randint(1, 6, 4)
}

df = pd.DataFrame(data)
print(df)
```

```
   Age  Income  Experience  CustomerSatisfaction
0   58   46929    1.413835                     4
1   36   45694    8.143124                     3
2   29   42323   12.288618                     4
3   47   56209   18.285079                     5
```

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

```
df.rename(columns={'Age':'Random value 1','Income':'Random value 2','Experience':'Random value 3','CustomerSatisfaction':'Random value 4'},in
df
```

	Random value 1	Random value 2	Random value 3	Random value 4
0	58	46929	1.413835	4
1	36	45694	8.143124	3
2	29	42323	12.288618	4
3	47	56209	18.285079	5

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

```
statistics = df.describe()
print(statistics)
```

```
       Random value 1  Random value 2  Random value 3  Random value 4
count        4.000000        4.000000        4.000000        4.000000
mean        42.500000    47788.750000       10.032664        4.000000
std         12.714821     5941.441485        7.095685        0.816497
min         29.000000    42323.000000        1.413835        3.000000
25%         34.250000    44851.250000        6.460802        3.750000
50%         41.500000    46311.500000       10.215871        4.000000
75%         49.750000    49249.000000       13.787733        4.250000
max         58.000000    56209.000000       18.285079        5.000000
```

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

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

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

null_values = df.isnull().any()
print("\nNull Values:")
print(null_values)

# Finding the data types of the columns
data_types = df.dtypes
print("\nData Types:")
print(data_types)
```

```
Null Values:
Random value 1    0
Random value 2    0
Random value 3    0
Random value 4    0
dtype: int64

Null Values:
Random value 1    False
Random value 2    False
Random value 3    False
Random value 4    False
dtype: bool

Data Types:
Random value 1      int64
Random value 2      int64
Random value 3    float64
Random value 4      int64
dtype: object
```

> Task 5 - Display the 'Random value 2' and 'Random value 3' columns with location method and index location method

```
# Displaying columns with location method
columns_loc = df.loc[:, ['Random value 2', 'Random value 3']]
print("Columns using .loc[]:")
print(columns_loc)

# Displaying columns using index location method
columns_iloc = df.iloc[:, [1, 2]]
print("\nColumns using .iloc[]:")
print(columns_iloc)
```

```
Columns using .loc[]:
   Random value 2  Random value 3
0           46929        1.413835
1           45694        8.143124
2           42323       12.288618
3           56209       18.285079

Columns using .iloc[]:
   Random value 2  Random value 3
0           46929        1.413835
1           45694        8.143124
2           42323       12.288618
3           56209       18.285079
```

✓ 0s completed at 12:47 AM