

# assignment-1-slack

October 19, 2023

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

# Defining the number of observations and features
num_observations = 4
num_features = 4

# Creating a NumPy array with random values
data = np.random.rand(num_observations, num_features)

# Create the DataFrame
df = pd.DataFrame(data, columns=['Feature 1', 'Feature 2', 'Feature 3',
↪ 'Feature 4'])

print(df)
```

	Feature 1	Feature 2	Feature 3	Feature 4
0	0.052624	0.471770	0.626420	0.670821
1	0.275578	0.819893	0.966660	0.240851
2	0.434525	0.552991	0.330094	0.415837
3	0.472862	0.989545	0.890138	0.384700

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

```
[ ]: # Rename the columns
df.rename(columns={'Feature 1': 'Random value 1',
                  'Feature 2': 'Random value 2',
                  'Feature 3': 'Random value 3',
                  'Feature 4': 'Random value 4'}, inplace=True)

# Display the updated DataFrame
print(df)
```

	Random value 1	Random value 2	Random value 3	Random value 4
0	0.052624	0.471770	0.626420	0.670821
1	0.275578	0.819893	0.966660	0.240851

2	0.434525	0.552991	0.330094	0.415837
3	0.472862	0.989545	0.890138	0.384700

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

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

# Display 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	0.308897	0.708550	0.703328	0.428052
std	0.191009	0.239178	0.288364	0.178895
min	0.052624	0.471770	0.330094	0.240851
25%	0.219840	0.532686	0.552338	0.348737
50%	0.355052	0.686442	0.758279	0.400268
75%	0.444109	0.862306	0.909269	0.479583
max	0.472862	0.989545	0.966660	0.670821

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

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

# Find the data types of the columns
data_types = df.dtypes

# Display the results
print("Null Values:")
print(null_values)
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

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.

```
[ ]: # Using label-based location method
columns_by_label = df.loc[:, ['Random value 2', 'Random value 3']]

# Display the selected columns
print(columns_by_label)
```

	Random value 2	Random value 3
0	0.471770	0.626420
1	0.819893	0.966660
2	0.552991	0.330094
3	0.989545	0.890138

```
[ ]: # Using index-based location method
columns_by_index = df.iloc[:, [1, 2]] # Index 1 corresponds to 'Random value_
↪2', and Index 2 corresponds to 'Random value 3'

# Display the selected columns
print(columns_by_index)
```

	Random value 2	Random value 3
0	0.471770	0.626420
1	0.819893	0.966660
2	0.552991	0.330094
3	0.989545	0.890138

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

```
[ ]:
```