

# **Assignment - 1**

Name: Gautam Prakash

Reg no: 20BCE2598

Email: gautam.prakash2020a@vitstudent.ac.in

AI Assignment 1	
Task - 1	Create a pandas dataframe (DataFrame name as 'df') with numpy random values (4 features and 4 observation)
Task - 2	Rename the task - 1 'df' dataframe column names to 'Random value 1', 'Random value 2', 'Random value 3' & 'Random value 4'
Task - 3	Find the descriptive statistics of the 'df' dataframe.
Task - 4	Check for the null values in 'df' and find the data type of the columns.
Task - 5	Display the 'Random value 2' & 'Random value 3' columns with location method and index location method.

## Question 1:

```
import pandas as pd

import numpy as np

# Create a NumPy array with random values
np.random.seed(42) # Setting seed
data = np.random.rand(4, 4) # 4 features and 4 observations

# Create the DataFrame
df = pd.DataFrame(data, columns=['F1', 'F2', 'F3', 'F4'])

print(df)
```

Question 1

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

# Create a NumPy array with random values
np.random.seed(42) # Setting seed
data = np.random.rand(4, 4) # 4 features and 4 observations

# Create the DataFrame
df = pd.DataFrame(data, columns=['F1', 'F2', 'F3', 'F4'])

print(df)
```

	F1	F2	F3	F4
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

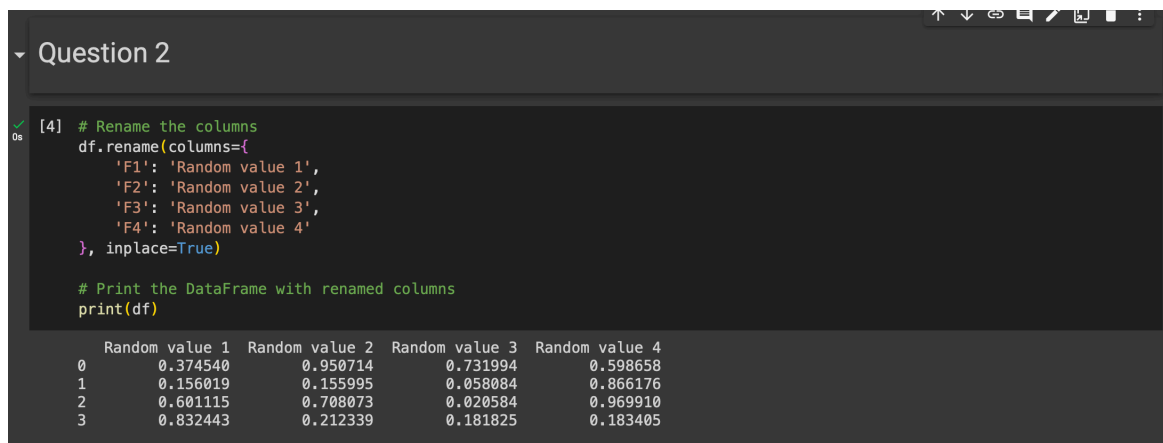
## Question 2:

```
# Rename the columns
```

```
df.rename(columns={  
    'F1': 'Random value 1',  
    'F2': 'Random value 2',  
    'F3': 'Random value 3',  
    'F4': 'Random value 4'  
}, inplace=True)
```

```
# Print the DataFrame with renamed columns
```

```
print(df)
```



```
[4] # Rename the columns  
df.rename(columns={  
    'F1': 'Random value 1',  
    'F2': 'Random value 2',  
    'F3': 'Random value 3',  
    'F4': 'Random value 4'  
}, inplace=True)  
  
# Print the DataFrame with renamed 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

### Question 3:

```
# Descriptive statistics of the DataFrame
```

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

```
# Print the descriptive statistics
```

```
print(statistics)
```

#### Question 3

```
✓ [5] # Descriptive statistics of the DataFrame  
1s 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

## Question 4:

```
# Check for null values and column types  
df.info()
```

### Question 4

```
0s  # Check for null values and column types  
df.info()  
  
<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
```

## Question 5:

```
# Display columns using label-based location method (.loc)
columns_loc = df.loc[:, ['Random value 2', 'Random value 3']]
print("Using label-based location method (.loc):")
print(columns_loc)
```

```
# Display columns using index-based location method (.iloc)
columns_iloc = df.iloc[:, [1, 2]]
print("\nUsing index-based location method (.iloc):")
print(columns_iloc)
```

### Question 5

```
✓ [7] # Display columns using label-based location method (.loc)
0s columns_loc = df.loc[:, ['Random value 2', 'Random value 3']]
print("Using label-based location method (.loc):")
print(columns_loc)

# Display columns using index-based location method (.iloc)
columns_iloc = df.iloc[:, [1, 2]]
print("\nUsing index-based location method (.iloc):")
print(columns_iloc)
```

Using label-based location method (.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 index-based location method (.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