

AI ML Assignment 1

Name – Kajal

Reg no. – 21BAC10039

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

```
import numpy as np
import pandas as pd

# Define the dimensions of the DataFrame
num_observation = 4
num_features = 4
random_values = np.random.rand(num_observation, num_features)
df = pd.DataFrame(random_values, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])
print(df)
```

Output:

	Feature1	Feature2	Feature3	Feature4
0	0.685148	0.664873	0.153992	0.925722
1	0.911381	0.615026	0.370979	0.910019
2	0.437747	0.817209	0.683688	0.045233
3	0.509270	0.930961	0.051106	0.261148

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

```
import numpy as np
import pandas as pd

num_observation = 4
num_features = 4
```

```

random_values = np.random.rand(num_observation, num_features)

df = pd.DataFrame(random_values, columns=['Feature1', 'Feature2', 'Feature3', 'Feature4'])

new_column_names = {
    'Feature1': 'Random value 1',
    'Feature2': 'Random value 2',
    'Feature3': 'Random value 3',
    'Feature4': 'Random value 4'
}

df = df.rename(columns=new_column_names)

print(df)

```

Output:

	Random value 1	Random value 2	Random value 3	Random value 4
0	0.942576	0.822405	0.219070	0.233292
1	0.292157	0.567574	0.607525	0.693132
2	0.621904	0.529611	0.780811	0.976691
3	0.032625	0.346951	0.167922	0.861397

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

```

import numpy as np
import pandas as pd

num_observation = 4
num_features = 4

random_values = np.random.rand(num_observation, num_features)

df = pd.DataFrame(random_values, columns=['Random value 1', 'Random value 2', 'Random
value 3', 'Random value 4'])

```

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

```
print(statistics)
```

Output:

	Random value 1	Random value 2	Random value 3	Random value 4
count	4.000000	4.000000	4.000000	4.000000
mean	0.486688	0.529157	0.663541	0.343777
std	0.312270	0.138133	0.218250	0.122834
min	0.104452	0.403839	0.441120	0.235029
25%	0.350245	0.451232	0.494181	0.239306
50%	0.491089	0.494817	0.673721	0.338197
75%	0.627532	0.572743	0.843081	0.442668
max	0.860122	0.723155	0.865600	0.463683

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

```
import numpy as np
```

```
import pandas as pd
```

```
num_observation = 4
```

```
num_features = 4
```

```
random_values = np.random.rand(num_observation, num_features)
```

```
df = pd.DataFrame(random_values, columns=['Random value 1', 'Random value 2', 'Random  
value 3', 'Random value 4'])
```

```
null_values = df.isnull().sum()
```

```
column_data_types = df.dtypes
```

```
print("Null Values:\n", null_values)
```

```
print("\nData Types:\n", column_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.

```
import numpy as np
import pandas as pd
```

```
num_observation = 4
num_features = 4
```

```
random_values = np.random.rand(num_observation, num_features)
```

```
df = pd.DataFrame(random_values, columns=['Random value 1', 'Random value 2', 'Random
value 3', 'Random value 4'])
```

```
random_value_2_loc = df.loc[:, 'Random value 2']
print("\nUsing loc:\n", random_value_2_loc)
```

```
random_value_3_iloc = df.iloc[:, 2]
print("\nUsing iloc:\n", random_value_3_iloc)
```

Output:

Using loc:

0 0.767460

1 0.218724

2 0.324387

3 0.150744

Name: Random value 2, dtype: float64

Using iloc:

0 0.956199

1 0.688478

2 0.103176

3 0.838887

Name: Random value 3, dtype: float64