

si-assignment-1

September 5, 2023

```
[1]: #Task 1 - Create a pandas dataframe (DataFrame name as 'df') with numpy random
      ↪ values (4 features and 4 observation)
```

```
[2]: import pandas as pd
```

```
[20]: data = {
      'Feature1' : [1.2, 2.3, 3.4, 4.5],
      'Feature2' : [5.6, 6.7, 7.8, 8.9],
      'Feature3' : [9.0, 10.1, 11.2, 12.3],
      'Feature4' : [13.4, 14.5, 15.6, 16.7]
    }
```

```
[21]: df = pd.DataFrame(data)
      df
```

```
[21]:
```

	Feature1	Feature2	Feature3	Feature4
0	1.2	5.6	9.0	13.4
1	2.3	6.7	10.1	14.5
2	3.4	7.8	11.2	15.6
3	4.5	8.9	12.3	16.7

```
[22]: #Task 2 - Rename the task - 1 'df dataframe column names to 'Random value 1',
      ↪ 'Random value 2', 'Random value 3' & 'Random value 4'
```

```
[23]: df.rename(columns = { "Feature1" : "Random value 1", "Feature2" : "Random value
      ↪ 2", "Feature3" : "Random value 3", "Feature4" : "Random value 4"},inplace =
      ↪ True)
      df
```

```
[23]:
```

	Random value 1	Random value 2	Random value 3	Random value 4
0	1.2	5.6	9.0	13.4
1	2.3	6.7	10.1	14.5
2	3.4	7.8	11.2	15.6
3	4.5	8.9	12.3	16.7

```
[24]: #Task 3 - Find the descriptive statistics of the 'df dataframe.
```

```
[25]: print(df.describe())
```

	Random value 1	Random value 2	Random value 3	Random value 4
count	4.000000	4.000000	4.000000	4.000000
mean	2.850000	7.250000	10.650000	15.050000
std	1.420094	1.420094	1.420094	1.420094
min	1.200000	5.600000	9.000000	13.400000
25%	2.025000	6.425000	9.825000	14.225000
50%	2.850000	7.250000	10.650000	15.050000
75%	3.675000	8.075000	11.475000	15.875000
max	4.500000	8.900000	12.300000	16.700000

```
[26]: #Task 4 - Check for the null values in 'df' and find the data type of the
      ↪ columns.
```

```
[27]: df.isnull()
```

```
[27]:
```

	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

```
[28]: print(df.dtypes)
```

```
Random value 1    float64
Random value 2    float64
Random value 3    float64
Random value 4    float64
dtype: object
```

```
[29]: #Task 5 - Display the 'Random value 2' & 'Random value 3' columns with location
      ↪ method and index location method.
```

```
[30]: #Location Based
      print(df.loc[:, ['Random value 2', 'Random value 3']])
```

	Random value 2	Random value 3
0	5.6	9.0
1	6.7	10.1
2	7.8	11.2
3	8.9	12.3

```
[31]: #Index Location Based (Zero Index)
      print(df.iloc[:, [1,2]])
```

	Random value 2	Random value 3
0	5.6	9.0

1	6.7	10.1
2	7.8	11.2
3	8.9	12.3

[]: