

ajlhost:8888/notebooks/AI%20AND%20ML%20Assignment%201.ipynb#

Lenovo Gmail YouTube Maps Machine Learning C... VIT Chennai - VTOP moodle

jupyter AI AND ML Assignment 1 Last Checkpoint: 18 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

In [1]: `import pandas as pd
import numpy as np`

Task 1: Create a pandas DataFrame (10 observations and 5 features)

In [3]: `data = {
 'feature1': np.random.rand(10),
 'feature2': np.random.rand(10),
 'feature3': np.random.rand(10),
 'feature4': np.random.rand(10),
 'feature5': np.random.rand(10)
}`

In [5]: `df = pd.DataFrame(data)
df`

Out[5]:

	Feature1	Feature2	Feature3	Feature4	Feature5
0	0.948090	0.506412	0.750814	0.800378	0.884926
1	0.290566	0.232384	0.774813	0.087859	0.224485
2	0.445826	0.594341	0.735345	0.165143	0.870206
3	0.132156	0.084722	0.588128	0.350249	0.410559
4	0.704462	0.383141	0.756125	0.486841	0.613644
5	0.697538	0.509447	0.058927	0.585474	0.746566
6	0.360148	0.792889	0.187848	0.846338	0.687813
7	0.646738	0.174250	0.328208	0.306903	0.678944
8	0.896629	0.513190	0.313654	0.765850	0.928075
9	0.015918	0.544771	0.438575	0.642417	0.911739

5	0.697538	0.509447	0.058927	0.585474	0.746566
6	0.360148	0.792889	0.187848	0.846338	0.687813
7	0.646738	0.174250	0.328208	0.306903	0.678944
8	0.896629	0.513190	0.313654	0.765850	0.928075
9	0.015918	0.544771	0.438575	0.642417	0.911739

Task 2: Check the info of 'df'

In [7]: `print("Task 2: Info of 'df'")
print(df.info())`

```
Task 2: Info of 'df'  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 10 entries, 0 to 9  
Data columns (total 5 columns):  
#   Column      Non-Null Count  Dtype  
---  -  
0   Feature1    10 non-null     float64  
1   Feature2    10 non-null     float64  
2   Feature3    10 non-null     float64  
3   Feature4    10 non-null     float64  
4   Feature5    10 non-null     float64  
dtypes: float64(5)  
memory usage: 528.0 bytes  
None
```

Task 3: Check the descriptive statistics of 'df'

Task 3: Check the descriptive statistics of 'df'

```
In [8]: print("\nTask 3: Descriptive statistics of 'df'")
print(df.describe())
```

```
Task 3: Descriptive statistics of 'df'
      Feature1  Feature2  Feature3  Feature4  Feature5
count  10.000000  10.000000  10.000000  10.000000  10.000000
mean    0.513007    0.433555    0.493244    0.503745    0.695696
std     0.315269    0.215171    0.263886    0.268778    0.230926
min     0.015918    0.084722    0.058927    0.087859    0.224485
25%    0.307962    0.270073    0.317293    0.317739    0.629969
50%    0.546282    0.507929    0.513351    0.536158    0.717189
75%    0.702731    0.536875    0.746947    0.734992    0.881246
max     0.948090    0.792889    0.774813    0.846338    0.928075
```

Task 4: Check the 4th index observation with 'loc' slicing operator

```
In [9]: print("\nTask 4: 4th index observation with 'loc'")
print(df.loc[3])
```

```
Task 4: 4th index observation with 'loc'
Feature1    0.132156
Feature2    0.084722
Feature3    0.588128
Feature4    0.350249
Feature5    0.410559
Name: 3, dtype: float64
```

Task 4: Check the 4th index observation with 'loc' slicing operator

```
In [9]: print("\nTask 4: 4th index observation with 'loc'")
print(df.loc[3])
```

```
Task 4: 4th index observation with 'loc'
Feature1    0.132156
Feature2    0.084722
Feature3    0.588128
Feature4    0.350249
Feature5    0.410559
Name: 3, dtype: float64
```

Task 5: Check the null values in your 'df'

```
In [10]: print("\nTask 5: Null values in 'df'")
print(df.isnull().sum())
```

```
Task 5: Null values in 'df'
Feature1    0
Feature2    0
Feature3    0
Feature4    0
Feature5    0
dtype: int64
```

```
In [ ]:
```