SMART – BRIDGE

ASSIGNMENT – 1

NAME – AYUSHI JAIN

REGISTRATION NO – 21BPS1275

CAMPUS – VIT CHENNAI

File Edit Format View Help

```
Task - 1 Create a pandas dataframe (DataFrame name as 'df') (10 observation and 5 features)

Task- 2 Check the info of 'df'

Task 3- Check the descriptive statistics of 'df'

Task 4- check the 4th index observation with 'loc' slicing operator.

Task 5 - Check the null values in your 'df'
```

Step -1) installing jupyter notebook and anaconda navigator

Step -2) opening command prompt and typing jupyter notebook

```
C:\Users\Ayushi Jain>jupyter notebook
[I 2023-08-30 19:55:36.756 ServerApp] Package notebook took 0.0000s to import
[I 2023-08-30 19:55:36.819 ServerApp] Package jupyter_lsp took 0.0622s to import
[W 2023-08-30 19:55:36.819 ServerApp] A `_jupyter_server_extension_points` function was not found in jupyter_lsp. In
d, a `_jupyter_server_extension_paths` function was found and will be used for now. This function name will be depre
d in future releases of Jupyter Server.
[I 2023-08-30 19:55:36.864 ServerApp] Package jupyter_server_terminals took 0.0440s to import
[I 2023-08-30 19:55:36.865 ServerApp] Package jupyterlab took 0.0000s to import
[I 2023-08-30 19:55:38.147 ServerApp] Package notebook_shim took 0.0000s to import
[W 2023-08-30 19:55:38.148 ServerApp] A `_jupyter_server_extension_points` function was not found in notebook_shim.
ead, a `_jupyter_server_extension_paths` function was found and will be used for now. This function name will be dep
ted in future releases of Jupyter Server.
[I 2023-08-30 19:55:38.149 ServerApp] jupyter_lsp | extension was successfully linked.
[I 2023-08-30 19:55:38.156 ServerApp] jupyter_server_terminals | extension was successfully linked.
[I 2023-08-30 19:55:38.166 ServerApp] jupyterlab | extension was successfully linked.
[I 2023-08-30 19:55:38.174 ServerApp] notebook | extension was successfully linked.
[I 2023-08-30 19:55:38.808 ServerApp] notebook_shim | extension was successfully linked.
[I 2023-08-30 19:55:38.892 ServerApp] notebook_shim | extension was successfully loaded.
[I 2023-08-30 19:55:38.893 ServerApp] jupyter_lsp | extension was successfully loaded.
[I 2023-08-30 19:55:38.894 ServerApp] jupyter_server_terminals | extension was successfully loaded.
[I 2023-08-30 19:55:38.905 LabApp] JupyterLab extension loaded from C:\Users\Ayushi Jain\AppData\Local\Programs\Pyth
ython311\Lib\site-packages\jupyterlab
```

Step -3) copying the link and pasting it in the browser

```
2023-08-30 19:55:38.987 ServerApp]

    To access the server, open this file in a browser:
        file:///C:/Users/Ayushi%20Jain/AppData/Roaming/jupyter/runtime/jpserver-9156-open.html
    Or copy and paste one of these URLs:
        http://localhost:8888/tree?token=ff6a14bcd3323021c023f622564ec1e40b02d7ddb325f7c9
        http://127.0.0.1:8888/tree?token=ff6a14bcd3323021c023f622564ec1e40b02d7ddb325f7c9
W 2023-08-30 19:55:39.018 ServerApp] Could not determine npm prefix: [WinError 2] The system cannot find the file spec
ed
[I 2023-08-30 19:55:39.060 ServerApp] Skipped non-installed server(s): bash-language-server, dockerfile-language-server
odejs, javascript-typescript-langserver, jedi-language-server, julia-language-server, pyright, python-language-server,
ython-lsp-server, r-languageserver, sql-language-server, texlab, typescript-language-server, unified-language-server, v
```

Step -4) open jupyter notebook and run the following commands

#task1

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

[2]: data = {
         'Feature1': np.random.rand(10),
         'Feature2': np.random.randint(1, 100, 10),
         'Feature3': np.random.choice(['A', 'B', 'C'], 10),
         'Feature4': np.random.randn(10),
         'Feature5': np.random.choice([True, False], 10)
     }

     df = pd.DataFrame(data)
```

Code –

```
data = {

    'Feature1': np.random.rand(10),

    'Feature2': np.random.randint(1, 100, 10),

    'Feature3': np.random.choice(['A', 'B', 'C'], 10),

    'Feature4': np.random.randn(10),

    'Feature5': np.random.choice([True, False], 10)

}


df = pd.DataFrame(data)
```

#task2

Code –

```
print("Task 2:")

print(df.info())

print("\n")
```

output -

```
print("Task 2:")
print(df.info())
print("\n")
```

```
Task 2:
<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     int32
 2   Feature3  10 non-null     object
 3   Feature4  10 non-null     float64
 4   Feature5  10 non-null     bool
dtypes: bool(1), float64(2), int32(1), object(1)
memory usage: 422.0+ bytes
None
```

#task3

Code –

print("Task 3:")

print(df.describe())

print("\n")

output –

```
[4]:  print("Task 3:")
      print(df.describe())
      print("\n")

      Task 3:
              Feature1   Feature2   Feature4
      count  10.000000  10.000000  10.000000
      mean    0.458031  53.700000  -0.584314
      std     0.389249  33.206592   1.088595
      min     0.023106   3.000000  -3.209356
      25%     0.133895  24.750000  -0.860326
      50%     0.384976  58.500000  -0.286491
      75%     0.836360  79.250000   0.012554
      max     0.960585  98.000000   0.470227
```

#task4

Code –

```
print("Task 4:")

print(df.loc[4])

print("\n")
```

output –

```
5]:  print("Task 4:")
     print(df.loc[4])
     print("\n")

     Task 4:
     Feature1      0.960585
     Feature2            65
     Feature3             C
     Feature4     -0.173789
     Feature5          True
     Name: 4, dtype: object
```

#task5

Code –

```
print("Task 5:")

print(df.isnull().sum())
```

output

```
[6]:  print("Task 5:")
      print(df.isnull().sum())

      Task 5:
      Feature1    0
      Feature2    0
      Feature3    0
      Feature4    0
      Feature5    0
      dtype: int64

[ ]:
```