

SmartInternz (Evening Batch)

Assignment-1

Name: Prakhar Agarwal

Reg No.: 21BIT0034

Email ID: prakhar.agarwal2021@vitstudent.ac.in

```
[1]: import pandas as pd
import numpy as np
# Creating random values using numpy
np.random.seed(42) # For reproducibility
random_values = np.random.rand(4, 4) # 4 features and 4 observations
# Creating a DataFrame using pandas
df = pd.DataFrame(random_values, columns=["Height", "Weight", "Age", "Score"])
print(df)
```

	Height	Weight	Age	Score
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

```
[2]: import pandas as pd
import numpy as np
# Creating random values using numpy
np.random.seed(42) # For reproducibility
random_values = np.random.rand(4, 4) # 4 features and 4 observations
# Creating a DataFrame using pandas
df = pd.DataFrame(random_values, columns=["Height", "Weight", "Age", "Score"])
# Renaming the columns
new_column_names = {
    "Height": "Random value 1",
    "Weight": "Random value 2",
    "Age": "Random value 3",
    "Score": "Random value 4"
}
df.rename(columns=new_column_names, inplace=True)
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

```
[3]: import pandas as pd
import numpy as np
# Creating random values using numpy
np.random.seed(42) # For reproducibility
random_values = np.random.rand(4, 4) # 4 features and 4 observations
# Creating a DataFrame using pandas
df = pd.DataFrame(random_values, columns=["Random value 1", "Random value 2",
s"Random value 3", "Random value 4"])
# Descriptive statistics
descriptive_stats = df.describe()
print(descriptive_stats)
```

	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

```
[4]: import pandas as pd
import numpy as np
# Creating random values using numpy
np.random.seed(42) # For reproducibility
random_values = np.random.rand(4, 4) # 4 features and 4 observations
# Creating a DataFrame using pandas
df = pd.DataFrame(random_values, columns=["Random value 1", "Random value 2",
s"Random value 3", "Random value 4"])
# Checking for null values
null_values = df.isnull().sum()
# Finding data types of columns
column_data_types = df.dtypes
print("Null Values:")
print(null_values)
print("\nData Types of Columns:")
print(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 of Columns:

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

```
[5]: import pandas as pd
import numpy as np
# Creating random values using numpy
np.random.seed(42) # For reproducibility
random_values = np.random.rand(4, 4) # 4 features and 4 observations
# Creating a DataFrame using pandas
df = pd.DataFrame(random_values, columns=["Random value 1", "Random value 2",
    "Random value 3", "Random value 4"])
# Using .loc[] to access columns by label-based location
columns_label = df.loc[:, ["Random value 2", "Random value 3"]]
print("Using .loc[]:")
print(columns_label)
# Using .iloc[] to access columns by index-based location
columns_index = df.iloc[:, [1, 2]]
print("\nUsing .iloc[]:")
print(columns_index)
```

Using .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 .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