

Siddhant Patil

211112238 CSE-2
ML Assignment 1

Ans.1.

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

df = pd.read_csv("/content/Student_Performance.csv")

df.head()

df.shape

# Column type info and data type
df.info()

# Number of Null values in each column
df.isna().sum()

# Mean, median and mode of columns of data
df.describe()
```

ML_theory_assignment.ipynb has been attached for all the output .
df.shape (important for next answer)

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 6 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   Hours Studied                        10000 non-null  int64
 1   Previous Scores                     10000 non-null  int64
 2   Extracurricular Activities          10000 non-null  object
 3   Sleep Hours                         10000 non-null  int64
 4   Sample Question Papers Practiced    10000 non-null  int64
 5   Performance Index                   10000 non-null  float64
dtypes: float64(1), int64(4), object(1)
memory usage: 468.9+ KB
```

Ans.2.



Maulana Azad National Institute of Technology, Bhopal (M.P.)

Roll No. 211112238

Scholar No. 211112238

Date

SIDDHANT PATIL

211112238

CSE-2

ML Assignment

SESSIONAL PAPER

② Let x_1 = Hours studied (in hrs)

x_2 = Previous scores (scores obtained in previous tests)

x_3 = Extracurricular Activities (Boolean, participates)

x_4 = Sleep hours (in hrs)

x_5 = Sample papers practised.

& x_6 = Performance Index.

Clearly from o/p of notebook datatype of x_i are int or float & continuous.

$\forall x_i : i \in [1, 5] \Rightarrow x_i$ is feature of dataset
and x_6 is output

Clearly,

Multiple Regression can be used.

GD Multiple Regression

$$\text{Eqn: } w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + w_5 x_5 + w_6 x_6 + w_0 = 0$$

$\{ \text{eqn of hyperplane of best fitting above points} \}$

$$\Rightarrow \sum_{i=1}^6 w_i x_i + w_0 = 0$$

$\{ w_0 \text{ is intercept} \}$

Considering, hyperplane passing through origin $\therefore w_0 = 0$
 $\Rightarrow \sum_{i=1}^6 w_i x_i = 0$

$$\text{Let } w = [w_1, w_2, w_3, \dots, w_6] \quad \{ \text{vector} \}$$

$$x = [x_1, x_2, x_3, \dots, x_6] \quad \{ \text{vector} \}$$

$$\therefore w^T x = 0 \quad \{ \text{in matrix form} \}$$

Now use multiple regression to get "w" vector
& then too further predictions