Linear Regression using Statistics

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

plt.rcParams["figure.figsize"] = (20.0,10.0)

data= pd.read\_csv(r"C:\Users\kumars262\Desktop\DS\DataSets\headbrain.csv")

X=data["Head Size(cm^3)"].values

Y=data["Brain Weight(grams)"].values

X\_mean = np.mean(X)

Y\_mean = np.mean(Y)

num = 0

denom = 0

for i in range(len(X)):

num += (X[i]-X\_mean)\*(Y[i]-Y\_mean)

denom += (X[i]-X\_mean)\*\*2

m = num/denom

c = Y\_mean - (m\*(X\_mean))

Y\_pred = m\*X+c

plt.scatter(X,Y\_pred)

plt.scatter(X,Y)

plt.plot(X,Y\_pred, color='green')

#accuracy Check

a=0.0

b=0.0

for i in range(len(Y)):

a+=(Y[i]-Y\_pred[i])\*\*2

b+=(Y[i]-Y\_mean)\*\*2

R2 = (1-(a/b))\*100

R2

#2.RMSE method

SSE = 0

TSS = 0

for i in range(len(Y)):

SSE += (Y[i] - Y\_pred[i])\*\*2

TSS += (Y[i] - Y\_mean)\*\*2

MSE = SSE/(len(Y))

RMSE = np.sqrt(MSE)

RMSE

R22 = (1-(SSE/TSS))\*100

#3. Adjusted R2

#R\_Adj = 1-(1-R^2)((n-1)/(n-p-1))

R\_adjusted = 1-((1-R2)\*((len(Y)-1)/(len(Y)-1-1)))

R\_adjusted

