Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Probability and Statistics (01CT1401)	**Aim:** Shoe Size Prediction Using Regression (Case Study)	
Assigment :- 1	**Date:-** 15-02-2024	**Enrollment No:-** 92200133030

Problem Statement:-

A shoe size is an alphanumerical indication of the fitting size of a shoe for a person. Often it just consists of a number indicating the length because many shoemakers only provide a standard width for economic reasons. Several different shoe-size systems are used worldwide. These systems define in what they measure, what unit of measurement they use, and where the size 0 (or 1) is positioned. Only a few systems also take the width of the feet into account. Some regions use different shoe-size systems for different types of shoes (e.g., men's, women's, children's, sport, or safety shoes). Perform a survey to collect data regarding the Height of a person and the Shoe size (at least 100 persons), anonymously. Perform linear regression over the data collected using Least Square Estimation.

Calculation:-

The dataset and the Table Are written in Excel Sheet Named **'Dataset.xlsx** Sheet Name Is **'Table'.**

For US Size Prediction:-

No.of Observations (n) = 109
$\sum$**xi(**Sum of Height) = 18901
$\sum$**yi(**Sum of Shoe Sizes) = 866.5
$\bar{x}$ = 173.4036697
$\bar{y}$ = 7.949541284
$\sum xi^2$ = 3297423
$\sum yi^2$ = 7126.75
$\sum$**xi * yi** = 151112.5
$\sum ei^2$ = 201.498

S_{xy} = ΣXiYi- [(ΣXi)(ΣYi) / n]
= 151112.5 – [(173.4036697)(7.949541284) / 109]
= 858.2201835

S_{xx} = ΣXi 2 - (ΣXi) 2 / n
= 3297423 – [30068.83267 / 109]
= 19920.23853

B1 = S_{xy} / S_{xx}
= 858.2201835 / 19920.23853
= 0.043082827

Marwadi University Marwadi Chandarana Group NAAC A+	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Probability and Statistics (01CT1401)	**Aim:** Shoe Size Prediction Using Regression (Case Study)	
Assigment :- 1	**Date:-** 15-02-2024	**Enrollment No:-** 92200133030

B0 = $1 / n (\sum y_i - B1 \sum x_i)$

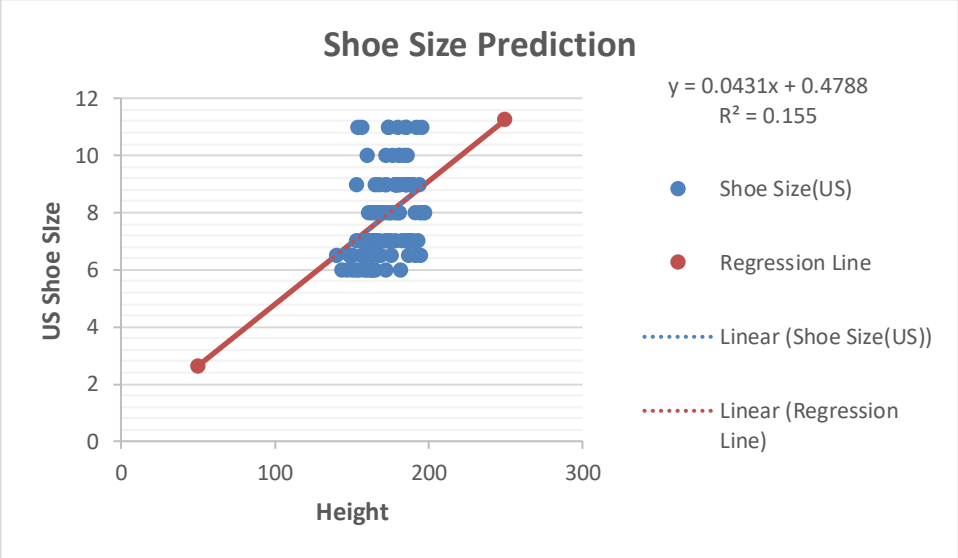
$= [1 / 109] (866.5 - (0.043082827) (866.5)$

$= 0.478821041$

Regression Line :- $\hat{y} = B0 + B1X$

$\hat{y} = 0.478821041 + 0.043082827(X)$

Graph:-



Code :-

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
from sklearn import linear_model

Dataset = pd.read_excel("./Dataset.xlsx", "Python_Data")

regression = linear_model.LinearRegression()
regression.fit(Dataset[["Hight"]], Dataset["Shoe Size(US)"])

print(f"B0 = {regression.intercept_} \nB1 = {regression.coef_}")
```

```
plt.figure(figsize=(18, 9))
plt.scatter(x=Dataset["Hight"], y=Dataset["Shoe Size(US)"], color="red", marker="*")
plt.plot(Dataset["Hight"], regression.predict(Dataset[["Hight"]]), color="blue")
plt.xlabel("Hight")
plt.ylabel("Shoe Size")
plt.title("Shoe Size Prediction")
plt.show()
```

Output :-

```
PS D:\Aryan Data\Usefull Data\Semester - 4\Probability and Statistics\TASK 1 - Regression Case study assignment> py
thon -u "d:\Aryan Data\Usefull Data\Semester - 4\Probability and Statistics\TASK 1 - Regression Case study assignme
nt\Code.py"
B0 = 0.4788210413456211
B1 = [0.04308283]
PS D:\Aryan Data\Usefull Data\Semester - 4\Probability and Statistics\TASK 1 - Regression Case study assignment>
```