|  |  |  |  |
| --- | --- | --- | --- |
|  | Hope Foundation’s  Finolex Academy of Management and Technology, Ratnagiri | | |
| Department of Computer Science and Engineering (AIML) | | |
| Subject name: | Machine Learning Lab | | Subject Code: CSL604 |
| Class | TE CSE | Semester –VI (CBCGS) | Academic year: 2024-25 |
| Name of Student |  | | QUIZ Score |
| Roll No |  | Experiment No. | 01 |
| Title: Implementation of Linear Regression. | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1. Lab objectives applicable:  LOB1: To introduce platforms such as Anaconda, COLAB suitable to Machine Learning.  LOB2: To implement various Regression techniques. | | | | | |
| 2. Lab outcomes applicable:  LO1: Implement various machine learning models. | | | | | |
| 3. Learning Objectives:  1. To predict future continuous value based on past data. | | | | | |
| 4. Practical applications of the assignment/experiment:  . | | | | | |
| 5. Prerequisites:  1. Python language | | | | | |
| 1. Minimum Hardware Requirements:-   I series processor, RAM 4GB,   1. Software Requirements:-   Colab or Visual Studio or Jupyter notebook (Anaconda) | | | | | |
| 8. Quiz Questions :  https://docs.google.com/forms/d/e/1FAIpQLSdRVHVFyYcXUSTrEByBKd5cJkaboF2dWYjEAcmUPVL690ZXwg/ viewform?usp=sf\_link | | | | | |
| 9. Experiment/Assignment Evaluation: | | | | | |
| Sr. No. | Parameters | | | Marks  obtained | Out of |
| 1 | Technical Understanding (Assessment may be done based on Q & A or any other relevant method.) Teacher should mention the other method used - | | |  | 6 |
| 2 | Lab Performance | | |  | 2 |
| 3 | Punctuality | | |  | 2 |
| Date of performance (DOP) | |  | Total marks obtained |  | 10 |

Signature of Faculty

10. Theory:

Linear regression is a type of supervised machine learning algorithm that computes the linear relationship between the dependent variable and one or more independent features by fitting a linear equation to observed data.

Formula: The formula for Linear Regression represents the relationship between the dependent variable (y) and the independent variables (x) using a linear equation. Here's the general form:

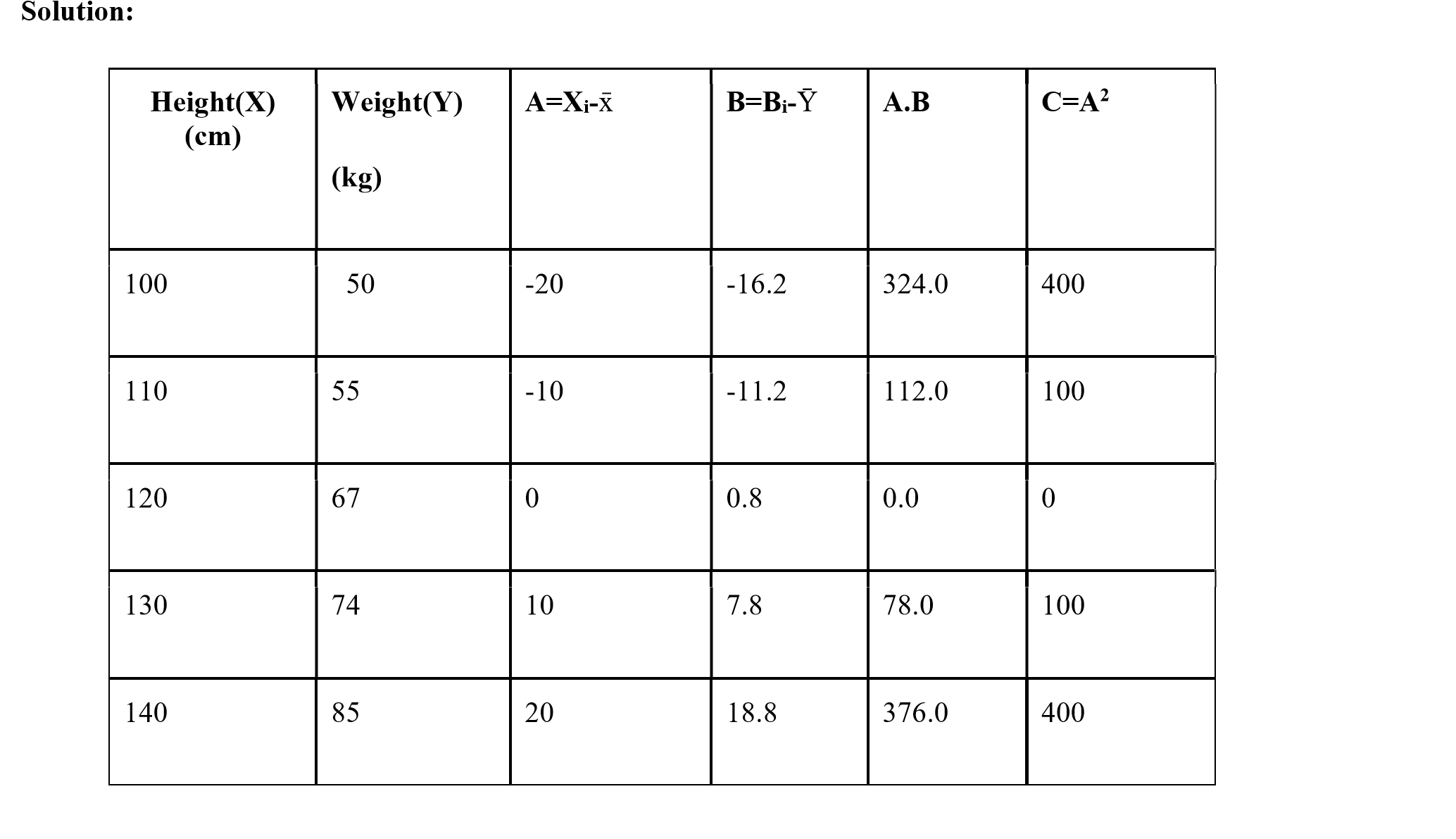
y=w0+w1x+ϵ

* y: Dependent variable (target).
* x: Independent variable (predictor).
* w0: Intercept (value of y when x=0).
* w1: Slope (rate of change of y with respect to x).
* ϵ: Error term (captures random noise or unexplained variance).

EXAMPLE:

Height(x) and weight(y) of students are given below. Predict the value of weight(y) using height(x).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Height(cm) | 100 | 110 | 120 | 130 | 140 |
| Weight(kg) | 50 | 55 | 67 | 74 | 85 |



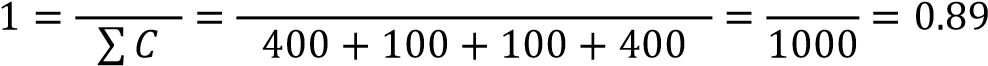
∑ 𝑋 600

𝑋 

∑ 𝑌 331

𝑌 

∑ 𝐴. 𝐵 324 + 112 + 0 + 78 + 376 890

𝑤

𝑤0 = 𝑌 − 𝑤1𝑋 = 66.2 − (0.89) ∗ 120 = −40.6

y = w0 + w1x

= (-40.6) + (0.89)x

= (-40.6) + (0.89)(150) for x = 150

=92.9

∴ For height(x) = 150 cm, predicted weight(y) = 92.9 kg

11. Installation Steps / Performance Steps and Results :

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

# Data

data = {'X': [100,110,120,130,140], 'Y': [50, 55, 67, 74,85]} df = pd.DataFrame(data) # Calculating means a = df['X'].mean() b = df['Y'].mean() print(f"Mean of X is {a}") print(f"Mean of Y is {b}") # Adding new columns to DataFrame

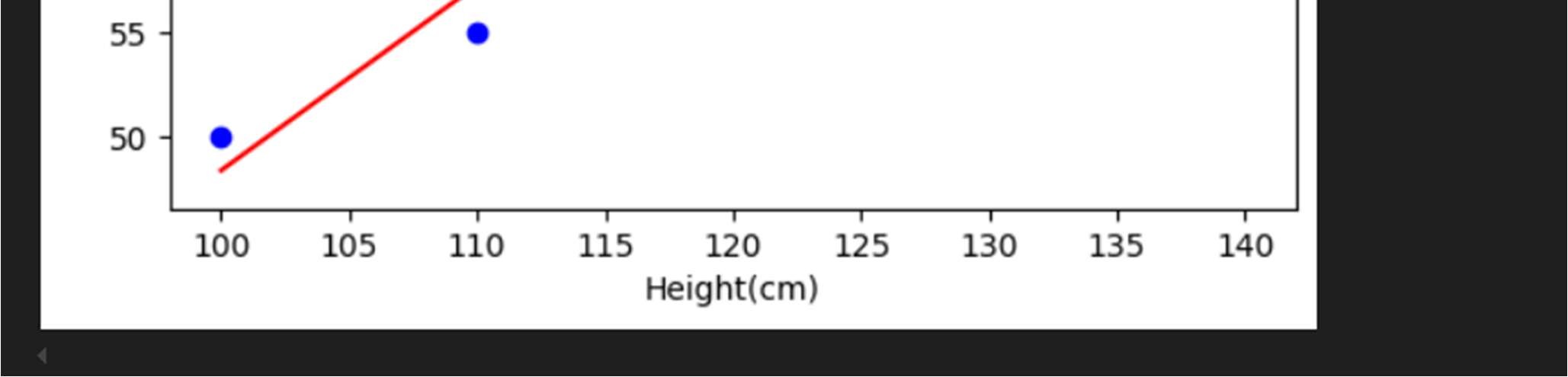
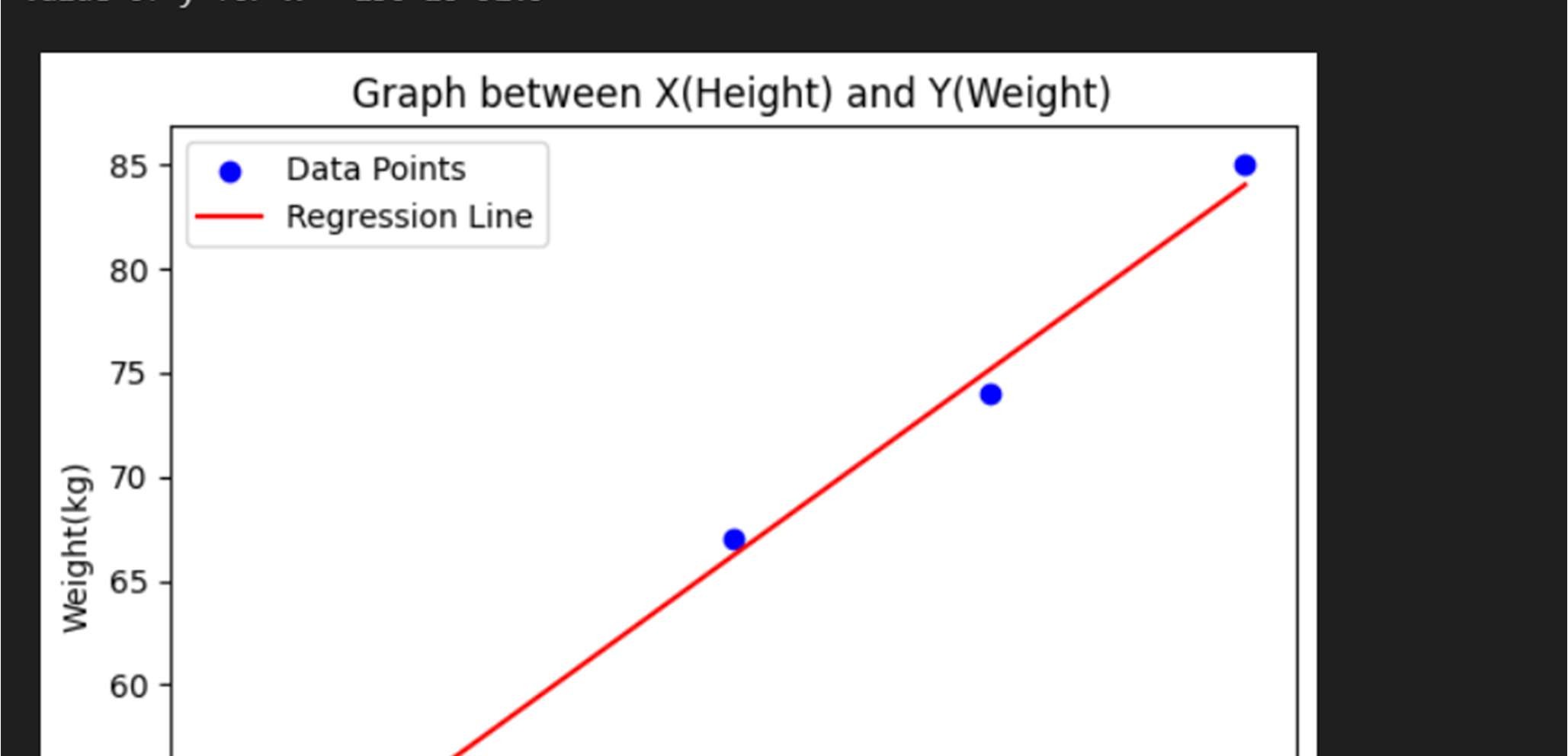
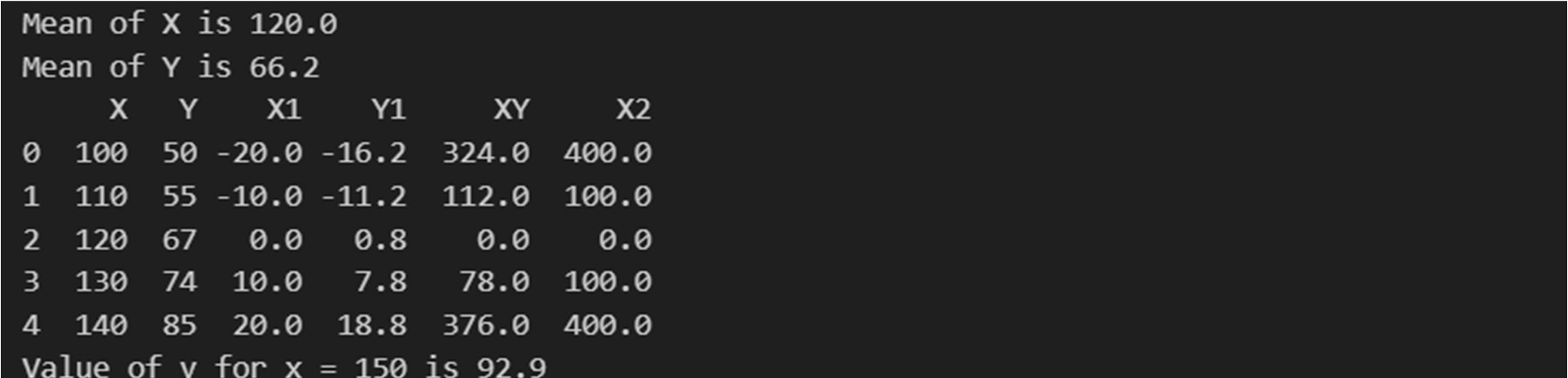
df['X1'] = df['X'] - a df['Y1'] = df['Y'] - b df['XY'] = df['X1'] \* df['Y1'] df['X2'] = df['X1'] \* df['X1'] print(df)

# Calculating regression coefficients w1 = df['XY'].sum() / df['X2'].sum() w0 = b - (df['XY'].sum() / df['X2'].sum()) \* a # Predicting value of y for a given x x = int(input("Enter the value of x: ")) y = w0 + w1\*x print(f"Value of y for x = {x} is {y}")

# Plotting the graph

plt.scatter(df['X'], df['Y'], color='blue', label='Data Points') plt.plot(df['X'], w0 + w1 \* df['X'], color='red', label='Regression Line') plt.title('Graph between X(Height) and Y(Weight)') plt.xlabel('Height(cm)') plt.ylabel('Weight(kg)') plt.legend() plt.show()

OUTPUT:



1. Learning Outcomes Achieved
   1. Students are able toimplement regression to predict future continuos value.
2. Conclusion:

* 1. Applications of the Studied Technique in Industry

Linear regression is widely used in various industries due to its simplicity and interpretability. Some notable applications include:

* + - Finance: Predicting stock prices, evaluating risk factors, or modeling consumer credit risk.
    - Marketing: Estimating the impact of advertising spend on sales or customer acquisition metrics.
  1. Engineering Relevance

Linear regression holds significant relevance in engineering disciplines:

* + Data-Driven Design: Engineers use regression to model relationships between design parameters and performance metrics.
  + Process Optimization: Helps in identifying key variables that influence efficiency and quality in manufacturing or chemical processes.
  1. Skills Developed

Programming Proficiency: Gaining experience in Python libraries like NumPy, pandas, scikit-learn, and matplotlib.

Critical Thinking: Evaluating the limitations and assumptions of linear regression, such as linearity independence, and normality of residuals.

1. References:

* 1. Nathalie Japkowicz & Mohak Shah, ―Evaluating Learning Algorithms: A Classification Perspective , Cambridge.
  2. Marc Peter Deisenroth, Aldo Faisal, Cheng Soon Ong, ―Mathematics for machine learning
  3. Samir Roy and Chakraborty, ―Introduction to soft computing , Pearson Edition.
  4. Ethem Alpaydın, ―Introduction to Machine Learning , MIT Press McGraw-Hill Higher

Education

* 1. Peter Flach, ―Machine Learning , Cambridge University Press
  2. Tom M. Mitchell, ―Machine Learning , McGraw Hill
  3. Kevin P. Murphy, ―Machine Learning ― A Probabilistic Perspective , MIT Press