**Name = Nihar**

**Sap ID = 500091867**

**Roll No = R2142210525**

**Experiment No -3**

**Title:- Implementation of Linear and Multiple Linear Regression**

**Linear regression is a statistical method used to analyze the relationship between two variables by fitting a linear equation to the observed data. The goal of linear regression is to find the best fit line, which describes the relationship between a dependent variable (also called the response variable) and one or more independent variables (also called predictor variables or features).**

**In simple linear regression, there is only one independent variable, and the relationship between the independent variable and the dependent variable is represented by a straight line. The line is determined by finding the values of the slope and intercept that minimize the difference between the predicted values of the dependent variable and the actual values.**

**In multiple linear regression, there are more than one independent variables, and the relationship between the independent variables and the dependent variable is represented by a hyperplane. The hyperplane is determined by finding the values of the coefficients that minimize the difference between the predicted values of the dependent variable and the actual values.**

**Code:-**

**import statsmodels.api as sm**

**import pandas as pd**

**import numpy as np**

**import matplotlib.pyplot as plt**

**data = pd.read\_csv("diabetes.csv")**

**print(data.head())**

**x = data['Age'].tolist()**

**y = data['BMI'].tolist()**

**result = sm.OLS(y,x).fit()**

**print(result.summary())**

**y\_pred = result.predict(x)**

**plt.scatter(x,y, color = 'r')**

**plt.scatter(x,y\_pred, color= 'g')**

**plt.plot(x,y\_pred, color = 'b')**

**plt.show()**

**Output:-**



