**Name:** Janhvi Katakdhond

**Roll no:** 282067

**Batch:** B3

**ASSIGNMENT 6**

**Problem Statement:**

a) Apply Linear Regression using suitable library function and predict the Month-wise temperature.  
b) Assess the performance of regression models using MSE, MAE and R-Square metrics  
c) Visualize simple regression model.

**Objective:**

The objective of this assignment is to develop a Linear Regression model to predict month-wise temperature using a suitable dataset. By applying an appropriate regression algorithm through a library function, the model will estimate temperature values based on temporal data. The performance of the regression model will be assessed using evaluation metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R²). Additionally, the relationship between the input feature(s) and the target variable will be visualized to provide a clearer understanding of the regression model’s behavior and predictive capability.

Here I used temperature dataset from Kaggle

https://www.kaggle.com/venky73/temperaturesof-india?select=temperatures.csv

**S/W Packages and H/W apparatus used:**

Software used:

1. Python 3.x

2. Google Colab

Libraries and packages used: NumPy, Pandas, Seaborn

**Theory:**

**Regression** is a fundamental technique in **supervised machine learning** used to model and analyze the relationship between a **dependent variable** (target/output) and one or more **independent variables** (features/input). The main goal of regression is to predict a **continuous** outcome based on input data.

**Types of Regression**

1. **Simple Linear Regression**

* Involves one independent variable and one dependent variable.
* Models that require relationship using a straight line
* Equation : y = β0 ​+ β1​x + ϵ

Where,

y = predicted output  
x = input feature

β0​ = intercept  
β1​ = slope of the line

ϵ = error term

1. **Multiple Linear Regression**

* Uses two or more independent variables to predict the dependent variable.
* Equation : y = β0 ​+ β1​x1 ​+ β2​x2 ​+ ⋯+ βn​xn ​+

1. **Polynomial Regression**

* A form of linear regression where the relationship between the independent and dependent variable is modeled as an nth-degree polynomial.
* **Equation**: y = β0 ​+ β1​x + β2​x2 +⋯+ βn​xn + ϵ

1. **Logistic Regression** *(technically a classification algorithm)*

* Used when the target variable is categorical (e.g., 0 or 1).
* Outputs probabilities through a sigmoid function.

**Applications of Regression**

* **Weather Forecasting**: Predicting temperature, humidity, rainfall, etc.
* **Stock Market Prediction**: Estimating future stock prices based on past trends.
* **Sales Forecasting**: Predicting future sales volumes based on historical data.
* **Real Estate**: Estimating house prices based on location, size, and other features.
* **Healthcare**: Predicting patient risk scores, disease progression, etc.
* **Business Analytics**: Modeling revenue growth, customer lifetime value, etc.

**Conclusion:**

In this assignment, Linear Regression was applied to predict month-wise temperature values based on temporal data. The model was developed using a suitable library function, and its performance was evaluated using key regression metrics such as **Mean Squared Error (MSE)**, **Mean Absolute Error (MAE)**, and **R-squared (R²)**. These metrics provided a clear understanding of how well the model was able to capture the relationship between time (months) and temperature.

The visualization of the regression line against the actual data points helped in interpreting the linear trend and validating the model's effectiveness. Overall, the assignment demonstrated how regression analysis can be a powerful tool for forecasting continuous outcomes and understanding underlying patterns in real-world datasets.