**Name:** Arya Yemul

**Roll No:** 281015

**Batch:** A1

**Assignment-6**

**Problem Statement:**

Apply Linear Regression using a suitable library function to predict month-wise temperature and evaluate the model using performance metrics.

**Objectives:**

1. Apply **regression techniques** using Python & Scikit-learn.
2. Implement **Linear Regression** to predict temperature trends across months.
3. Evaluate model performance using **MSE, MAE, and R²**.
4. Visualize regression results using **scatter plots and trend lines**.

**Resources used:**

1. Software used: Visual Studio Code
2. Libraries used: Pandas, Matplotlib, Seaborn,numpy

**Theory:**

This project utilizes **Linear Regression**, a fundamental machine learning technique, to model and predict monthly temperature trends based on historical data. Alongside, it leverages key Python libraries for data handling, visualization, and modeling.

**Linear Regression**

**Linear Regression** is a supervised learning algorithm used for **predicting a continuous dependent variable** based on one or more independent variables. The goal is to establish a **linear relationship** between the variables in the form:

y=mx+c

Where:

* 1. y is the predicted value (temperature)
  2. xis the independent variable (month)
  3. m is the slope (rate of temperature change)
  4. ccc is the intercept

Linear regression works by minimizing the **residual sum of squares** between observed and predicted values using **least squares estimation**.

**Evaluation Metrics**

To measure model performance, the following metrics are used:

1. **Mean Squared Error (MSE):**  
   Measures the average squared difference between actual and predicted values. Lower MSE indicates better accuracy.
2. **Mean Absolute Error (MAE):**  
   The average of absolute differences between predicted and true values.
3. **R² Score (Coefficient of Determination):**  
   Indicates how well the model explains the variability of the target variable. Closer to 1 means better fit.

**Methodology:**

**1. Data Preprocessing**

• Load dataset using Pandas.  
• Handle missing values (if any).  
• Encode month names numerically.• Normalize data if needed using MinMaxScaler/StandardScaler.  
• Split data into training and testing sets (e.g., 80/20 split).

**2. Model Selection & Training**

**•** Use Linear Regression as the predictive model.  
• Train the model on the training set.  
• Fit the model to predict temperatures based on month.

**3. Evaluation Metrics**

**• Predict using the test data.  
• Evaluate model performance using:**a)MSE (Mean Squared Error)  
 b) MAE (Mean Absolute Error)  
 c) R² Score (Coefficient of Determination)

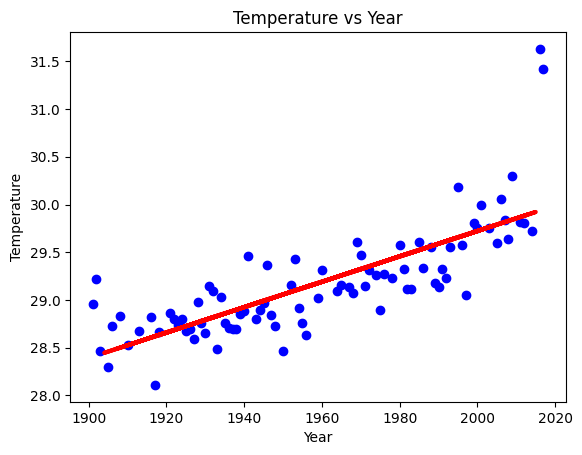
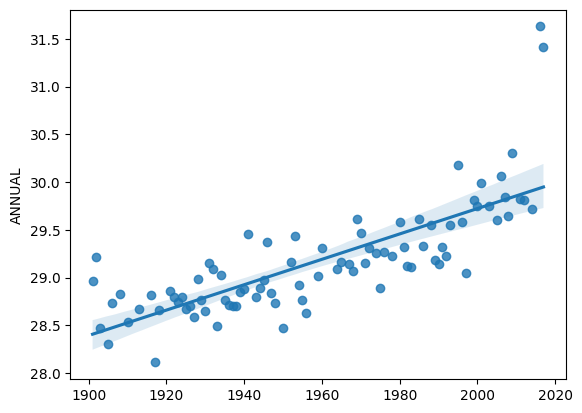
**4. Visualization**

* Plot scatter plot with regression line.
* Use line charts or residual plots for additional insights.

**Results:**

MSE: 0.051117407890753884 MAE: 0.1709352259448668 R-Sqaure : 0.7987564369156026

**RegPlot**

** **

**Conclusion:**

1. The regression model effectively predicts monthly temperatures with a reasonable error margin.
2. The evaluation metrics provide insight into model accuracy