**ASSIGNMENT NO: 6**

**Problem Statement-**

This data consists of temperatures of INDIA averaging the temperatures of all places month-wise. Temperature values are recorded in CELSIUS.

a) Apply Linear Regression using a suitable library function and predict the Month-wise

temperature.

b) Assess the performance of regression models.

**S/W Packages and Libraries used-**

Software Package: Python

Libraries Used:

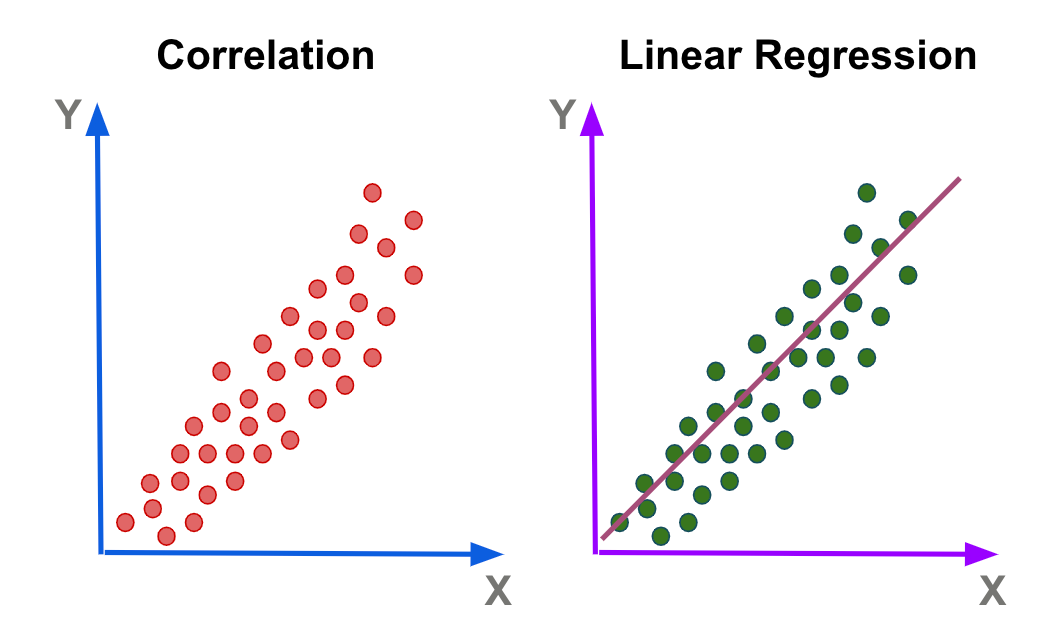
pandas: For data manipulation and analysis.

scikit-learn: For implementing machine learning algorithms, including Linear Regression.

matplotlib: For data visualization.

**Theory-**

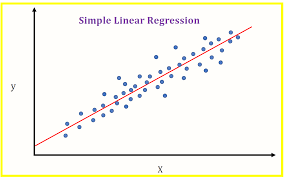
* Linear Regression:
  + It is a statistical method that is used for predictive analysis.
  + Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc.
  + Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called linear regression.
  + Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable.

****

Types of Linear Regression-

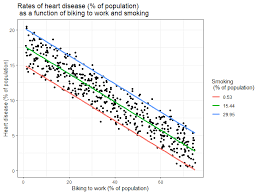
● Simple Linear Regression:

If a single independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Simple Linear Regression.



● Multiple Linear regression:

If more than one independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Multiple Linear Regression.

****

**Methodology-**

* Data Preprocessing:
  + Load the temperature data.
  + Prepare the data by separating features (months) and target variable (monthly temperatures).
  + Split the data into training and testing sets.
* Model Training:
  + Initialize a Linear Regression model.
  + Fit the model on the training data.
* Model Evaluation:
  + Predict the month-wise temperatures using the trained model.
  + Evaluate the model's performance using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-Square metrics.

**Applications:**

* Predicting the Salary of a person based on years of experience- Therefore, Experience becomes the independent variable while Salary turns into the dependent variable.
* Predicting crop yields based on the amount of rainfall- Yield is a dependent variable while the measure of precipitation is an independent variable.

**Limitations:**

* Linear Regression assumes a linear relationship between the independent and dependent variables, which may not always hold true in real-world scenarios.
* The accuracy of predictions can be affected by outliers and noise in the data.
* The model's performance heavily depends on the quality and representativeness of the training data.

**Conclusion:**

* Linear Regression can be effectively used to predict month-wise temperatures in India.
* The model's performance can be assessed using metrics like MSE, MAE, and R-Square.
* Despite its limitations, Linear Regression provides a simple and interpretable approach for temperature prediction, which can be valuable for various applications in climate science and related fields.