Objective: Estimate individual medical insurance expenses based on factors such as age, BMI, smoking status, and other health-related variables.

Dataset Overview:

- **age:** Age of the primary policyholder.
- **sex:** Gender of the insurance policyholder (female or male).
- **bmi:** Body Mass Index, a measure that uses the ratio of weight to height (kg/m^2) to categorize body weight, ideally ranging from 18.5 to 24.9.
- **children:** Number of dependents covered by the health insurance.
- **smoker:** Indicates if the person smokes.
- **region:** The geographical area in the U.S. where the policyholder resides (northeast, southeast, southwest, northwest).
- **charges:** Medical expenses billed to the insurance for the individual.

Steps:

Import Libraries:

- Gather all necessary libraries for data manipulation and linear regression modeling.

Load the Dataset:

- Read the dataset into a data frame to start working with it.

Explore the Data:

- Get an overview of the dataset by looking at the first few rows, summary statistics, and data types.

Handle Missing Data:

- Check for any missing values and address them appropriately, such as by filling them in or dropping rows/columns.

Convert Categorical Variables:

- Convert categorical variables like ‘sex’, ‘smoker’, and ‘region into numerical values using encoding techniques.

Define Features and Target:

- Separate the dataset into input features (independent variables) and the target variable (dependent variable, "charges").

- **Split the Dataset:**

- Divide the data into training and testing sets to evaluate the model's performance on unseen data.

- **Train the Linear Regression Model:**

- Fit a linear regression model using the training dataset.

- **Make Predictions:**

- Use the model to predict "charges" on the test dataset.

Evaluate the Model:

- Assess the model's accuracy using metrics like Mean Squared Error (MSE) and R-squared (R^2).

Analyze Model Coefficients:

- Examine the model coefficients to understand the impact of each feature on the predicted charges.

Model Refinement (Optional):

- Based on the model's performance, consider refining the model by adjusting features, transformations, or adding interaction terms.