**Multiple Linear Regression**

**Problem Statement** – Built this model for a housing dataset and predict the price of a house using the various potential predictor variables provided.

**Explanation –**

* Identify the variables affecting house prices, e.g., area, number of rooms, bathrooms, etc.
* Build model that relates house prices with variables, such as the number of rooms, area, number of bathrooms, etc.
* Calculate the accuracy of the model to understand how well these variables predict the house prices

**Building Model steps** -

1. Reading and Understanding the Data
2. Visualize the Data to find out multicollinearity between independent variables. During this step we will also identify if some predictors directly have a strong association with the outcome variable.
3. Data Preparation
   1. Dealing with categorical variables (create Dummy variables)
4. Splitting the Data into Training and Testing Sets
   1. Feature scaling using Standardization and normalization
   2. Dividing into X and Y sets for the model building
5. Building a linear model
   1. Built the model using combined approach of coarse tuning and fine tuning for feature selection
   2. Calculated VIF to check the multicollinearity
   3. Rebuild the model until VIFs and p-values both are within an acceptable range
6. Residual Analysis of the train data
   1. Check if the error terms are normally distributed
7. Model assessment
   1. Assess the model using adjusted R-squared score
8. Making Predictions Using the Final Model
   1. Scaling on test set
   2. Predict using the model built on test set
9. Model Evaluation
   1. Plot the graph for actual versus predicted values