**Regression Analysis Project Report: House Price Prediction**

**Introduction** :This project aims to build a Linear Regression model to predict house prices based on features such as Size, Location, and Number of Rooms. By analyzing real estate data, we explore how these attributes influence pricing. The model is trained and evaluated using performance metrics such as RMSE and R², offering valuable insights for property valuation.

**1. Dataset Overview**

* **Dataset Name:** house\_prices.csv
* **Features:**
  + Size (Numeric): Area in square feet
  + Location (Categorical): Urban, Suburban, Rural
  + Number of Rooms (Numeric)
  + Price (Numeric): Target variable

**2. Load and Explore** :We begin by loading the dataset and checking for missing values. Any incomplete rows are dropped. Exploratory Data Analysis (EDA) is used to understand the distributions and detect outliers using histograms and boxplots.

**3. Data Preprocessing**

* **Normalization**: Numerical features like Size and Number of Rooms are standardized using StandardScaler.
* **Encoding**: Location, a non-ordinal categorical variable, is transformed using One-Hot Encoding to make it suitable for the regression model.

**4. Feature Selection** :Correlation analysis is conducted on numerical columns to identify the most impactful features on Price. The heatmap visualization helps in interpreting the strength of relationships among variables.

**5. Model Training** :The data is split into training and testing sets using an 80/20 ratio. A Linear Regression model is then fitted using the training data through a pipeline that includes preprocessing steps.

**6. Model Evaluation**:The model is evaluated on test data using:

* **Root Mean Square Error (RMSE)**: Measures average prediction error.
* **R² Score**: Indicates how well the model explains variance in house prices.

**7. Results and Predictions** :The model outputs predictions for house prices which are compared against actual values. A scatterplot of actual vs. predicted values shows how well the model performs.

**8. Feature Insights:** Feature coefficients from the model indicate that:

* Size and Number of Rooms have strong positive impacts on price.
* Location also influences price, with Urban areas typically commanding higher prices.

**RESULTS**:













