**HOUSE PREDICTION USING LINEAR REGRESSION**

**Title**: House Price Prediction Using Linear Regression

**Introduction:**

This project aims to build a machine learning model that predicts house prices based on various features. Linear regression, a fundamental supervised learning algorithm, is employed to create this predictive model. The dataset used for this project is 'USA\_Housing.csv,' which presumably contains information about houses in the United States.

**Project Steps:**

Importing Libraries: The project begins by importing essential Python libraries such as NumPy, pandas, matplotlib, and seaborn for data manipulation, visualization, and modeling.

**Loading the Dataset:**

The dataset ('USA\_Housing.csv') is loaded into a pandas DataFrame named 'HouseDF.' This dataset likely contains features such as average area income, house age, number of rooms, number of bedrooms, and population in a given area, along with the corresponding house prices.

**Data Exploration:**

HouseDF.head(): Displays the first few rows of the dataset to get a glimpse of its structure.

HouseDF.info(): Provides information about the dataset, including data types and missing values.

HouseDF.describe(): Generates summary statistics for the numerical columns.

HouseDF.columns: Lists the columns or features in the dataset.

**Data Visualization:**

sns.pairplot(HouseDF): Creates a pairplot to visualize relationships between pairs of numerical features.

sns.heatmap(HouseDF.corr(), annot=True): Generates a heatmap to visualize the correlation between different features, which can help in feature selection.

**Data Preprocessing:**

Selecting Features and Target Variable: Features (X) and the target variable (y) are selected from the dataset.

train\_test\_split: The data is split into training and testing sets for model evaluation.

**Linear Regression Model:**

LinearRegression(): Initializes a linear regression model.

lm.fit(X\_train, y\_train): Trains the model using the training data.

**Model Evaluation:**

pd.DataFrame(lm.coef\_, X.columns, columns=['Coefficient']): Displays the coefficients of the linear regression model, indicating the importance of each feature in predicting house prices.

predictions = lm.predict(X\_test): Uses the trained model to make predictions on the test data.

**Visualization of Results:**

plt.scatter(y\_test, predictions): Creates a scatter plot to visualize how well the predicted house prices align with the actual house prices.

sns.distplot((y\_test-predictions), bins=50): Generates a distribution plot to examine the residuals (the differences between predicted and actual values) to assess the model's performance.

**Conclusion:**

This project demonstrates the application of linear regression for predicting house prices based on various features. By visualizing the data, training the model, and evaluating its performance, you can gain insights into the factors that influence house prices and create a predictive tool for estimating house prices in the given dataset. The coefficients of the linear regression model can also provide valuable information about the relative importance of each feature in the prediction.