**Project Title: House Price Prediction**

**Project Description:**

I undertook a machine learning project focused on predicting house prices using Python and popular data science libraries. The project involved several key steps, including data exploration, preprocessing, feature engineering, and the development of predictive models.

**Project Documentation:**

1. Loading the Data Set:

I began the project by loading the housing dataset using Pandas. This dataset served as the foundation for the entire project.

2. Data Exploration:

To gain insights into the data, I performed exploratory data analysis (EDA). This involved using Pandas and visualization libraries like Matplotlib and Seaborn to understand the dataset's structure, summary statistics, and the distribution of key variables.

3. Data Preprocessing:

To prepare the data for modeling, I addressed missing values by dropping rows with missing data. This ensured that the dataset was clean and ready for further analysis.

4. Feature Engineering:

I conducted feature engineering to create new variables that could improve the model's performance. This included:

Log-transforming certain variables to address skewness.

Encoding categorical features using one-hot encoding to make them suitable for machine learning models.

Creating new features like bedroom\_ratio and household\_rooms to capture additional information.

5. Linear Regression Model:

I built a Linear Regression model using the Scikit-Learn library. This model aimed to predict house prices based on the selected features.

The dataset was split into training and testing sets to evaluate the model's performance.

After training the model on the training set, I calculated the R-squared score (0.4062) to assess its predictive capability on the test set.

6. Random Forest Model:

In addition to Linear Regression, I implemented a Random Forest Regressor using Scikit-Learn.

This ensemble model was trained on the same training set and evaluated with an R-squared score (0.2149) on the test set.

This project showcased my proficiency in data preprocessing, feature engineering, and machine learning modeling techniques. It also demonstrated my ability to use data visualization to gain insights from the dataset. The project's documentation highlights my skills in Python programming, data manipulation, and model evaluation.