HOUSE PRICE PREDICTION

**EXPLAINATION**

**house price prediction** problem using machine learning. Below is a step-by-step breakdown of the code:

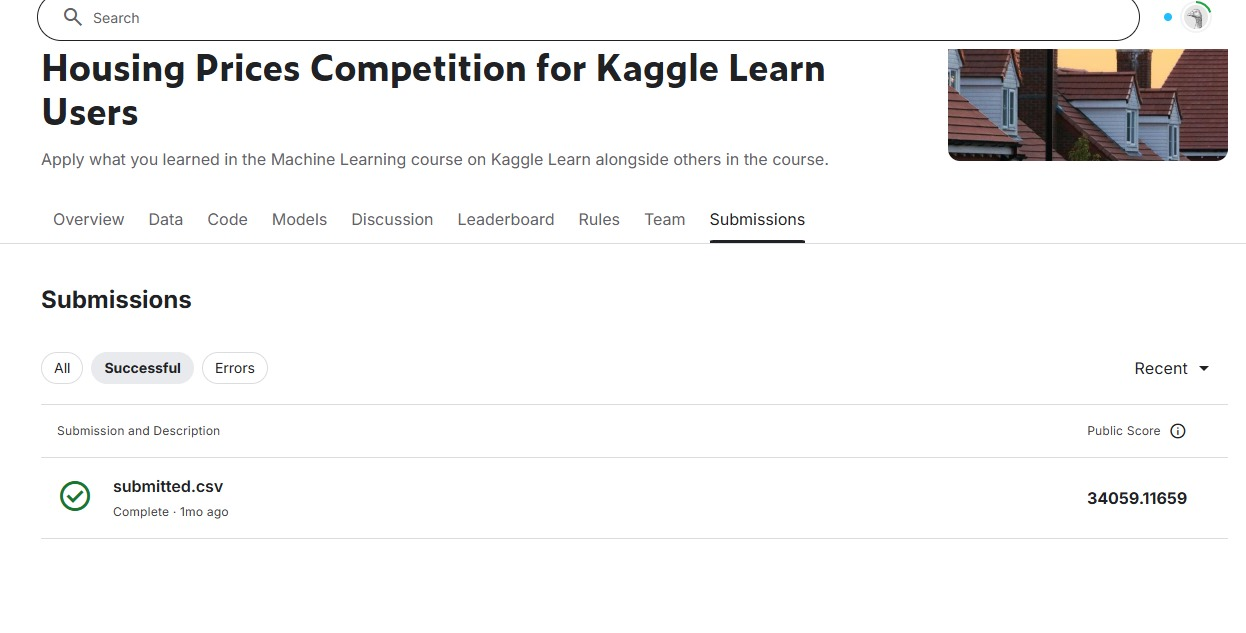
The provided Jupyter Notebook is designed for house price prediction using machine learning. It starts by importing necessary libraries like pandas for data manipulation and numpy for numerical operations. The dataset is divided into a training dataset (df\_train) and a test dataset (df\_test), both read from CSV files. The training dataset is then examined using .info() to understand the data types and check for missing values.

Next, several irrelevant or redundant columns are removed to simplify the dataset. Missing values in specific categorical columns, such as BsmtQual and Electrical, are filled with the most frequent value (mode) to maintain data consistency. Since machine learning models require numerical input, categorical features like Street, LotShape, and Neighborhood are converted into numerical values using Label Encoding.

Once the data is cleaned and preprocessed, the independent variables (features) are separated from the dependent variable (SalePrice), which represents house prices. A Random Forest Classifier is then used to train the model on the preprocessed data. However, since this is a regression problem (predicting continuous values), a Random Forest Regressor would be more appropriate. After training, the same preprocessing steps are applied to the test dataset to ensure consistency before making predictions.

Overall, the notebook follows a structured approach of data cleaning, feature engineering, model training, and prediction to estimate house prices based on various features.

**Output**

****