**Housing Prices Prediction (Kaggle Competition)**

**📌 Project Overview**

This repository contains my solution for the Housing Prices Competition for Kaggle Learn Users.

The goal of this competition is to predict house sale prices based on various property features. It is a classic **regression problem**, widely used as an introductory project for machine learning practice.

**📂 Project Structure**

├── data/ # Dataset folder (not uploaded to GitHub)

├── notebooks/ # Jupyter Notebooks

├── src/ # Python scripts (data cleaning, feature engineering, modeling)

├── submissions/ # Generated submission files

└── README.md # Project documentation

**📊 Dataset**

The dataset is provided by Kaggle. **Note:** Due to Kaggle’s rules, the dataset files (train.csv, test.csv) are **not included in this repository**.

You can download the dataset directly from the competition page:  
👉 House Prices Dataset

Main files:

* train.csv: Training data with features and sale prices
* test.csv: Test data with features (target values not provided)
* sample\_submission.csv: Example submission file

**🔧 How to Run**

**1. Clone the repository**

git clone https://github.com/<your-username>/housing-prices-kaggle.git

cd housing-prices-kaggle

**2. Install dependencies**

pip install -r requirements.txt

**3. Download the dataset**

Place the train.csv and test.csv files inside the data/ folder.

**4. Run notebooks**

jupyter lab

Open the notebooks under notebooks/ to explore the code and reproduce results.

**🧠 Methods & Models**

* **Data Cleaning:** handling missing values, outliers
* **Feature Engineering:** categorical encoding, scaling, feature selection
* **Modeling Approaches:**
  + Linear Regression
  + Random Forest Regressor
  + XGBoost / LightGBM
* **Evaluation Metric:** Root Mean Squared Error (RMSE)

**📈 Results**

Best submission on Kaggle achieved:

* **RMSE:** *<your score here>*

**📜 License**

This project is for learning and research purposes only. Please follow the Kaggle Terms of Service.