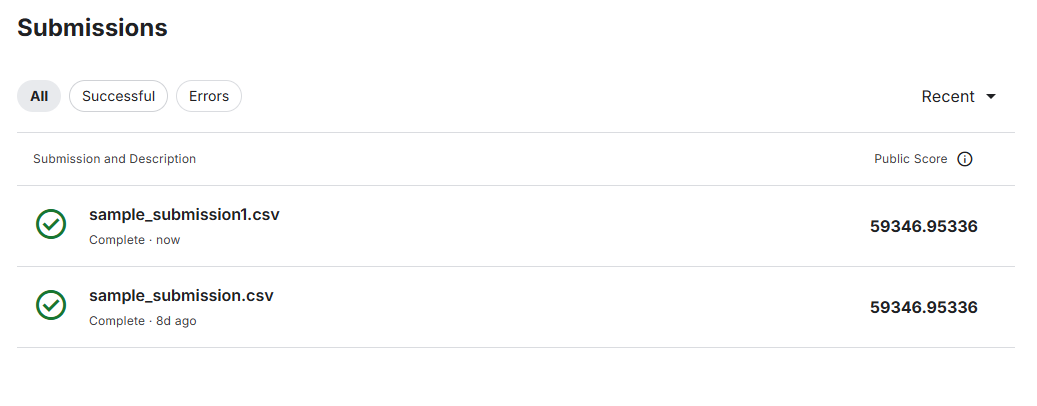
# Name : Hashim Nadeem

**Roll No: SU92-BDSFM-F25-001**

****

# House Prices Kaggle Competition - Step by Step Guide

This document explains what was done with your Kaggle Housing Prices competition files and how to create a submission.

## 1. Sample Submission

- You provided a file called `sample\_submission[1].csv`.  
- It contains columns: Id, SalePrice.  
- This is the format Kaggle expects for submissions.

## 2. Baseline Submission

- We saved your sample submission directly as `submission.csv`.  
- This can be uploaded to Kaggle as a baseline submission (just to test the pipeline).

## 3. Model Training Plan

To generate real predictions, you need `train.csv` and `test.csv` from the competition.  
Steps:  
1. Load `train.csv` (contains features + SalePrice).  
2. Load `test.csv` (contains features only).  
3. Preprocess data:  
 - Fill missing values.  
 - Encode categorical features.  
4. Train a regression model (RandomForestRegressor used here).  
5. Predict SalePrice for the test set.  
6. Save predictions in the required CSV format.

## 4. Python Script

I created a script `train\_and\_submit.py` that:  
- Trains the model.  
- Prints validation score.  
- Saves a trained model and features.  
- Outputs `submission\_model.csv`.

## 5. Notebook (Easy Mode)

I also created a Jupyter Notebook `HousePrices\_Easy\_Notebook.ipynb`.  
It explains step by step:  
1. Load data  
2. Train/validate  
3. Predict  
4. Save submission CSV  
This is beginner-friendly and easier to follow.

## 6. Uploading to Kaggle

- Go to the Kaggle competition submission page.  
- Drag and drop `submission\_model.csv` (or `submission.csv`).  
- Kaggle will evaluate and give you a score on the leaderboard.

## 7. Next Steps

- Improve preprocessing (log transform SalePrice, feature engineering).  
- Try advanced models like XGBoost or LightGBM.  
- Use cross-validation for better evaluation.  
- Tune hyperparameters.