House Prices Competition Summary

Competition Overview

- **Ideal Participants**: Suitable for data science students with some R or Python experience and basics of machine learning, looking to expand their skills.
- **Competition Focus**: Participants will use advanced regression techniques to predict house prices based on 79 explanatory variables.
- **Dataset**: The Ames Housing dataset, compiled by Dean De Cock, is used, offering a detailed set of variables to model beyond the basic features like bedroom count or fence type.

Starting Point

- **Getting Started Notebook**: A starter notebook is provided to help participants begin analyzing the data quickly.

Key Skills to Practice

- Creative feature engineering.
- Advanced regression techniques, including random forest and gradient boosting.

Evaluation Criteria

- **Goal**: Predict the sales price for each house based on its features.
- **Metric**: Performance is measured using Root-Mean-Squared-Error (RMSE) on the logarithmic scale of prices.

Submission Guidelines

- Format: The submission file should include two columns: Id and SalePrice.