# HOUSE PRICES

## Description

Ask a home buyer to describe their dream house, and they probably won't begin with the height of the basement ceiling or the proximity to an east-west railroad. But this playground competition's dataset proves that much more influences price negotiations than the number of bedrooms or a white-picket fence.

With 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, this competition challenges you to predict the final price of each home.

## Goal

It is your job to predict the sales price for each house. For each Id in the test set, you must predict the value of the SalePrice variable.

## Metric

Submissions are evaluated on Root-Mean-Squared-Error (RMSE) between the logarithm of the predicted value and the logarithm of the observed sales price. (Taking logs means that errors in predicting expensive houses and cheap houses will affect the result equally.)

## Install

This project requires Python 2.7 and the following Python libraries installed:

* [NumPy](http://www.numpy.org/)
* [matplotlib](http://matplotlib.org/)
* [seaborn](https://stanford.edu/~mwaskom/software/seaborn/)
* [scikit-learn](http://scikit-learn.org/stable/)
* [XGBoost](https://xgboost.readthedocs.io/en/latest/)
* [LightGBM](https://lightgbm.readthedocs.io/en/latest/)

You will also need to have [jupyter notebook](https://jupyter.org/install.html) installed to run and execute the script.

## Code

Script is in code folder. All input data are in input folder and the detailed description of the data can be found in [Kaggle](https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data). The submission file is in Output folder.