Lab Task 3

Download the dataset from the following link:

https://www.kaggle.com/c/house-prices-advanced-regression-techniques/overview

The Ames Housing dataset was compiled by Dean De Cock for use in data science education. It's an incredible alternative for data scientists looking for a modernized and expanded version of the often cited Boston Housing dataset.

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, you have to predict the final price of each home.

Goal

Predict the sales price for each house. For each Id in the test set, you have to 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.)

Submit the thoroughly commented code along with snapshots of RMSE taken in Spyder environment.

Participation to kaggle website competition is optional.