Springboard--Data Science Career Program
Capstone Project #1 - Real Estate Price Predictor
Project Proposal
By Kevin Cole
January 2020

Problem: As an individual, or as a real estate investor, finding properties that are undervalued and avoiding properties that are overvalued can help the long term prospects of the investment. Therefore the goal of this project is to build models that can estimate the price of properties as a function of the features of such properties.

Clients: The clients would be anybody interested in buying/selling a house. They would be interested in using predictive models to give them a sense of what the house should cost, so they can make informed decisions to buy or not buy, sell or not sell.

Data: House_prices_advanced_techniques (Kaggle[1])

This Dataset contains 1459 data points over 80 predictor variables. This dataset also has an additional 1460 data points with the same predictor variables, absent a target variable. This second set of data points will be useful for analysis between predictor variables, but not so much for measuring accuracy of the model since target variable truth data isn't present.

Approach: The approach would be to build and evaluate supervised regression learning models. There would be some analysis as to the best predictor variables to use and feed into the training model. For instance, such variables would have strong correlation with the sale price.

Deliverables: The project deliverables will include Jupyter notebooks containing methods used and code to support the analysis. In addition, a final report will be delivered as well as a presentation slide deck.

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