SpringBoard - Data Science Program

Predicting Housing Prices in Ames, Iowa

Capstone Project 2 - Proposal

Lisa Lorentzon
November 2020

Problem Definition and Client

When buying a house, home buyers often are not immediately imagining the height of the basement ceiling or the exact size of a property, but about the number of bedrooms or bathrooms. This project aims to find out what features of a house influences price negotiations to help home buyers/sellers and/or real estate agents calculate what the final selling price of a property in Ames, Iowa will be based on its features.

The goal of this project is to build models to predict the final price of homes in Ames, Iowa based on 79 different property features.

Datasets

The datasets used for this project were compiled by Dean De Cock and have been retrieved from https://www.kaggle.com/c/house-prices-advanced-regression-techniques/overview. The dataset consists of 1460 properties with 79 features listed for each property.

Anticipated Data Science Approach

I anticipate to build regression models which will be evaluated according to the following metrics: R^2 Coefficient, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), upper and lower bound worst-case scenarios for the test set, and feature importance/impact.

Deliverables

At the conclusion of this project I will provide a GitHub repo containing my code through Jupyter Notebooks as well as my findings and conclusions via a written report and a presentation slide deck.