**Project title: Predicting House Values in Ames, Iowa**

Project Description/Outline

With 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, our goal is to predict the final price of each home of the test data set using the training data set.

Research Question to Answer

What housing characteristics affect the price of a home most significantly?

H0: The following housing characteristics have no effect on housing prices in Ames, Iowa.

* Year remodeled
* Lot area
* Living area sq ft
* Total basement sq ft
* Overall quality

Ha: The above housing characteristics have an effect on housing prices in Ames, Iowa.

Datasets to be used

Source: <https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data>

Rough Breakdown of Tasks:

1. Preprocessing data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Numerical Columns | | Categorical Columns | | # of categories |
| LotFrontage | Feet of street connected to house | MSSubClass | Type of dwelling | 16 |
| LotArea | Lot size sqft | MSZoning | Zoning classification | 8 |
| YearBuilt | Orig. construct date | Street | Type of road access | 2 |
| YearRemodAdd | Remodel date (no remodel = same as YearBuilt) | Alley | Type of alley access | 3 |
|  |  | LotShape | Shape of property | 4 |
|  |  | LandContour | Flatness of property | 4 |
|  |  | Utilities | Utilizes available | 4 |
|  |  | LotConfig | Lot configuration | 5 |
|  |  | LandSlope | Slope of property | 3 |
|  |  | Neighborhood | Physical locations w/in Ames | 25 |
|  |  | Condition1 | Proximity to conditions | 9 |
|  |  | Condition2 | Same as above if more than 1 present | 9 |
|  |  | BldgType | Type of dwelling (simple) | 5 |
|  |  | HouseStyle | Style of dwelling | 8 |
|  |  | OverallQual | Rates material/finish of house | 1-10 |
|  |  | OverallCond | Rates condition of house | 1-10 |
|  |  | RoofStyle | Type of roof | 6 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. Adjusting categoricals/data munging
2. Data analysis
3. Regression analysis
4. Hypothesis testing
5. Creation of presentation and story