|  |
| --- |
| Elmhurst College |
| Classifying the Price of a House in Ames, IA |
| MSD 570 Fall A 2019 |

|  |
| --- |
| Megan Cusey  10-13-2019 |

# Overview

# Business Understanding

# Data Understanding

# Data Preparation

## Data Exploration

### Feature Engineering

* Domain Knowledge
* Interactive features – combine 2 or more features (products, sums, or differences)
* Combine Sparse Classes

### Principal Component Analysis

# Modeling

## Linear Regression

# Evaluation

# Appendix

## Data Dictionary & Thoughts

|  |  |  |  |
| --- | --- | --- | --- |
| Column Name | Description | Feature Engineering Type | Initial Thoughts |
| MSSubClass | Identifies the type of dwelling involved in the sale   |  |  |  | | --- | --- | --- | | 20 | 1-STORY 1946 & NEWER ALL STYLES | | | 30 | 1-STORY 1945 & OLDER | | | 40 | 1-STORY W/FINISHED ATTIC ALL AGES | | | 45 | 1-1/2 STORY - UNFINISHED ALL AGES | | 50 | 1-1/2 STORY FINISHED ALL AGES | | 60 | 2-STORY 1946 & NEWER | | 70 | 2-STORY 1945 & OLDER | | 75 | 2-1/2 STORY ALL AGES | | 80 | SPLIT OR MULTI-LEVEL | | 85 | SPLIT FOYER | | 90 | DUPLEX - ALL STYLES AND AGES | | 120 | 1-STORY PUD (Planned Unit Development) - 1946 & NEWER | | 150 | 1-1/2 STORY PUD - ALL AGES | | 160 | 2-STORY PUD - 1946 & NEWER | | 180 | PUD - MULTILEVEL - INCL SPLIT LEV/FOYER | | 190 | 2 FAMILY CONVERSION - ALL STYLES AND AGES | |  | Maybe break out into different features?   * # of stories * Finished attic * Age? * Duplex |
| MSZoning | Identifies the general zoning classification of the sale.   |  |  | | --- | --- | | A | Agriculture | | C | Commercial | | FV | Floating Village Residential | | I | Industrial | | RH | Residential High Density | | RL | Residential Low Density | | RP | Residential Low Density Park | | RM | Residential Medium Density | |  |  |
| LotFrontage | Linear feet of street connected to property |  |  |
| LotArea | Lot size in square feet |  |  |
| Street | Type of road access to property   |  |  | | --- | --- | | Grvl | Gravel | | Pave | Paved | |  |  |
| Alley | Type of alley access to property   |  |  | | --- | --- | | Grvl | Gravel | | Pave | Paved | | NA | No alley access | |  |  |
| LotShape | General shape of property   |  |  | | --- | --- | | Reg | Regular | | IR1 | Slightly Irregular | | IR2 | Moderately Irregular | | IR3 | Irregular | |  |  |
| LandContour | Flatness of the property   |  |  | | --- | --- | | Lvl | Near Flat/Level | | Bnk | Banked – Quick and significant rise from street to building. | | HLS | Hillside – Significant slope from side to side | | Low | Depression | |  |  |
| Utilities | Type of utilities available |  |  |
| LotConfig | Lot configuration |  |  |
| LandSlope | Slope of property |  |  |
| Neighborhood | Physical locations within Ames city limits |  |  |
| Condition1 | Proximity to various conditions |  |  |
| Condition2 | Proximity to various conditions (if more than one is present) |  |  |
| BldgType | Type of dwelling |  |  |
| HouseStyle | Style of dwelling |  |  |
| OverallQual | Rates the overall material and finish of the house |  |  |
| OverallCond | Rates the overall condition of the house |  |  |
| YearBuilt | Original construction date |  |  |
| YearRemodelAdd | Remodel date (same as construction date if no remodeling or additions) |  |  |
| RoofStyle | Type of roof |  |  |
| RoofMatl | Roof material |  |  |
| Exterior1st | Exterior covering on house |  |  |
| Exterior2nd | Exterior covering on house (if more than one material) |  |  |
| MasVnrType | Masonry veneer type |  |  |
| MasVnrArea | Masonry veneer area in square feet |  |  |
| ExterQual | Evaluates the quality of the material on the exterior |  |  |
| ExterCond | Evaluates the present condition of the material on the exterior |  |  |
| Foundation | Type of foundation |  |  |
| BsmtQual | Evaluates the height of the basement |  |  |
| BsmtCond | Evaluates the general condition of the basement |  |  |
| BsmtExposure | Refers to walkout or garden level walls |  |  |
| BsmtFinType1 | Rating of the basement finished area |  |  |
| BsmtFinSF1 | Type 1 finished square feet |  |  |
| BsmtFinType2 | Rating of basement finished area (if multiple types) |  |  |
| BsmtFinSF2 | Type 2 finished square feet |  |  |
| BsmtUnfSF | Unfinished square feet of basement area |  |  |
| TotalBsmtSF | Total square feet of basement area |  |  |
| Heating | Type of hearing |  |  |
| HearingQC | Heating quality and condition |  |  |
| CentralAir | Central air conditioning |  |  |
| Electrical | Electrical system |  |  |
| 1stFlrSF | First floor square feet |  |  |
| 2ndFlrSF | Second floor square feet |  |  |
| LowQualFinSF | Low quality finished square feet (all floors) |  |  |
| GrLivArea | Above grade (ground) living area square feet |  |  |
| BsmtFullBath | Basement full bathrooms |  |  |
| BsmtHalfBath | Basement half bathrroms |  |  |
| FullBath | Bull bathrooms about grade |  |  |
| HalfBath | Half baths above grade |  |  |
| Bedroom | Bedrooms above grade (does not include basement bedrooms) |  |  |
| Kitchen | Kitchens above grade |  |  |
| KitchenQual | Kitchen Quality |  |  |
| TotRmsAbvGrd | Total rooms above grade (does not include bathrooms) |  |  |
| Functional | Home functionality ( Assume typical unless deductions are warranted) |  |  |
| Fireplaces | Number of fireplaces |  |  |
| FireplaceQu | Fireplace quality |  |  |
| GarageType | Garage location |  |  |
| GarageYrBlt | Year garage was built |  |  |
| GarageFinish | Interior finish of the garage |  |  |
| GarageCars | Size of garage in car capacity |  |  |
| GarageArea | Size of garage in square feet |  |  |
| GarageQual | Garage quality |  |  |
| GarageCond | Garage condition |  |  |
| PavedDrive | Paved driveway |  |  |
| WoodDeckSF | Wood deck are in square feet |  |  |
| OpenPorchSF | Open porch area in square feet |  |  |
| Enclosed Porch | Enclosed porch area in square feet |  |  |
| 3SsnPorch | Three season porch area in square feet |  |  |
| ScreenPorcch | Screen porch area in square feet |  |  |
| PoolArea | Pool area in square feet |  |  |
| PoolQC | Pool quality |  |  |
| Fence | Fence quality |  |  |
| MiscFeature | Miscellaneous feature not covered in other categories |  |  |
| MiscVal | Value of miscellaneous feature |  |  |
| MoSold | Month Sold (MM) |  |  |
| YrSold | Year Sold (YYYY) |  |  |
| SaleType | Type of sale |  |  |
| SaleCondition | Condition of sale |  |  |
| Sale price |  |  |  |