

Home Sale Price Prediction

Ames, Iowa housing dataset modeling,
Advanced regression machine learning
algorithms

Capstone Project by Pavithra Rajkumar
For Springboard Data Science Intensive
certificate

Goals

- Home sale price prediction with a high level accuracy promotes faster closing of transactions
- Feature engineering - Develop features to predict home prices
- Build models to predict home prices for the Ames, Iowa dataset with a high level of accuracy
- Modeling using machine learning algorithms



Dataset

All these variables are features for our consideration in modeling

➤ Discrete variables

- (# of kitchen, # bedrooms, # bathrooms, Month sold etc.)

➤ Categorical variables

▶ Nominal variables

- Sale Type, Sale Condition, Neighborhood, Building Type

▶ Ordinal variables

- Lot Shape, Utilities, Land Slope

Data Wrangling

➤ Missing Values removed

- ▶ PoolQC
- ▶ MiscFeatures
- ▶ Basement
- ▶ Alley
- ▶ FireplaceQC

➤ Missing Features, value changed to None

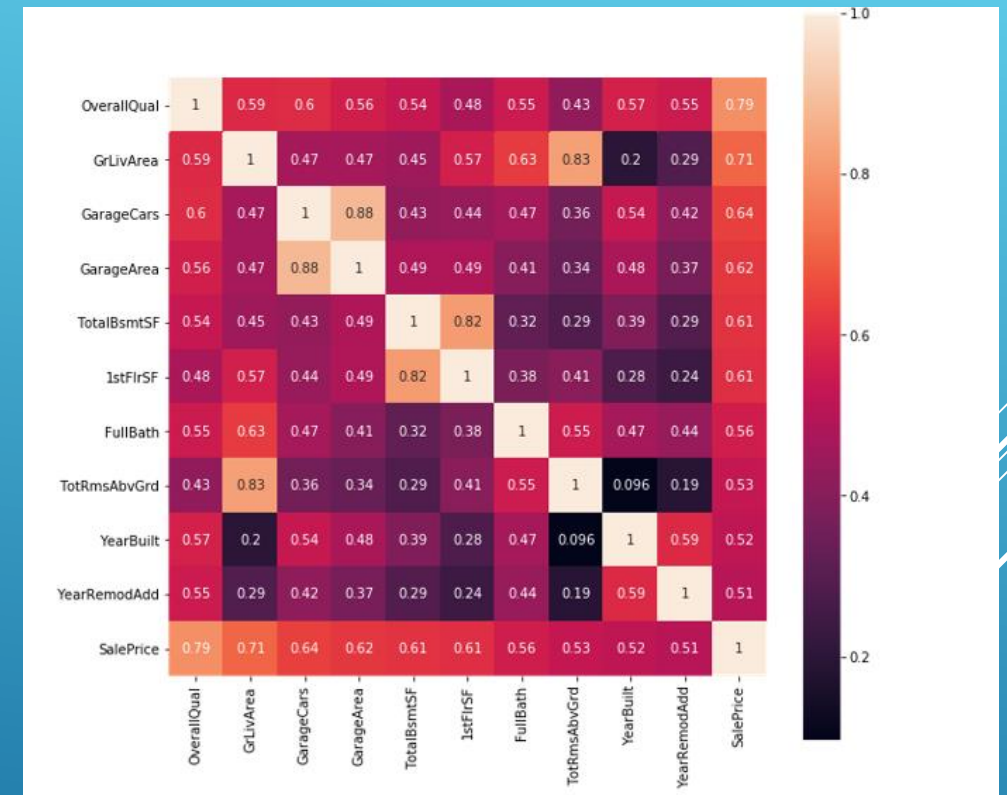
- ▶ Garage
- ▶ Basement

➤ Missing Features, value with related values and grouped

- ▶ PoolQC & Pool Area = 0
- ▶ Garage, NaN and related variable
- ▶ Electrical removed since all homes should have it

Data Visualization

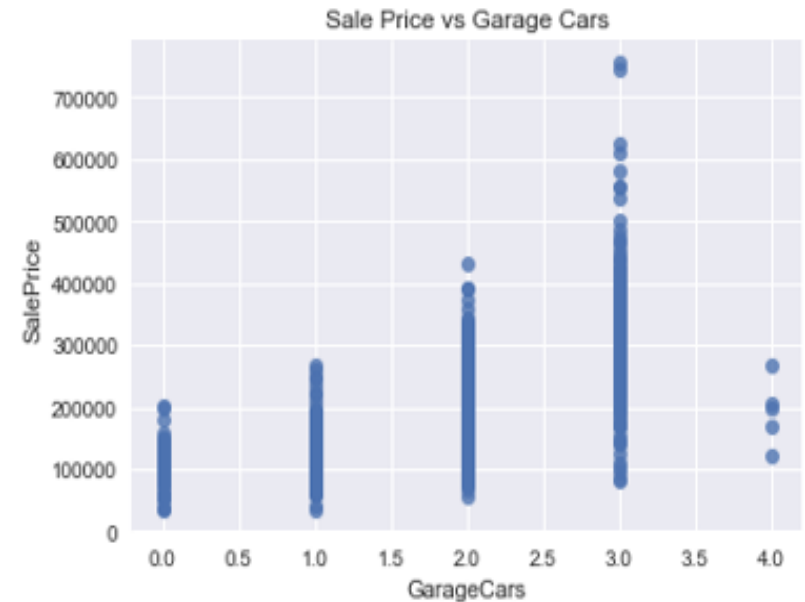
- Correlation of features using correlation matrix
- Overall Quality and Ground Floor living area have a strong correlation to Sales Prices



Exploratory Data Analysis

Outliers of Sales Prices when plotted with the following variables

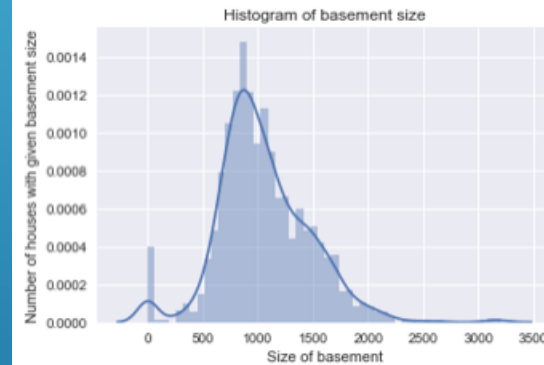
- Ground Living area
- Year Built
- Size of Basement
- Garage cars



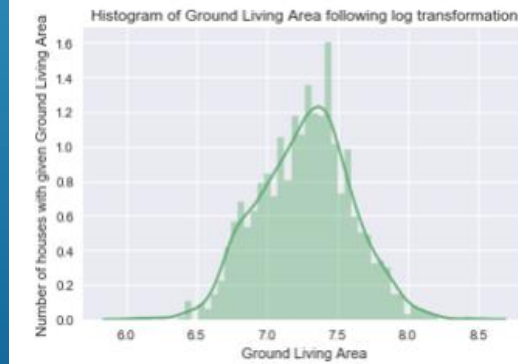
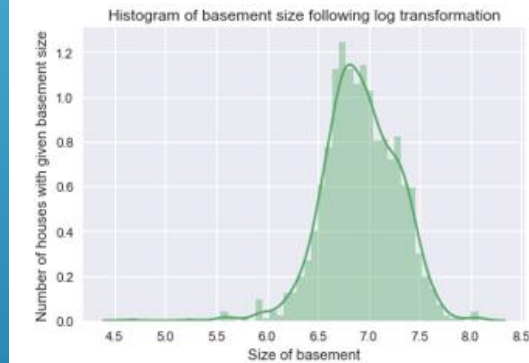
Exploratory Data Analysis

- Sales price
- Size of Basement
- Ground Living Area

Skewed & outliers



After Log transformation



Regression Models

Result Summary for R2 Score			
Ridge Regression	Lasso Regression	Support Vector Machine	Gradient Boosting Regression
0.825067	0.826482	0.826482	0.834250

- R Square has been used to evaluate the models
- Gradient Boosting Regression has the best accuracy and was used to predict the home prices

Thank you

