Modeling Assigment #3

Brandon Moretz

### Introduction

To accurately forecast the value of a home, we must find a relevant dataset that contains accurate information of comparable inventory so that we can explore the significant variables of a home which ultimately determine the sale price of the residence. Once we have explored the data set and selected an appropriate sample from the population, our task will be to create both single and multivariate regression models that leverages these key indicators in the data to predict the value of a home given based upon its features. Once we have constructed the models, we will form hypothesis tests at our stated confidence intervals and conduct statistical significance tests upon these models.

In this report, we will use the Ames dataset which is an alternative to the famous Boston housing data to perform exploratory data analysis through variable derivation, validation, selection and visualization to measure the relevance of these indicators as they pertain to the value of the home in terms of a dollar estimate.

### preparing the categorical variables

For this part of the lab we will take a systematic approach to examining the relationships between the categorical variables in the data set in relation to the desired response variable. We will look at the subset of 43 columns that contain categorical information and extract the R2, residual standard error from the model fitted to predict the sale price, as well as the mean difference between levels, the number of levels, and the percent of the data that is populated with this attribute.

The reason we chose these metrics is due to the variance explained by each category is an indicator of the relative “goodness-of-fit”, and the RSE and mean difference give us a sense of the variance found in each of the levels, where the lower the variance and higher the R2 will give us a good idea of how useful this metric will be in predictive modeling. The full results of this exercise can be found in the [appendix](#_Appendix).

### Conclusion

### Appendix