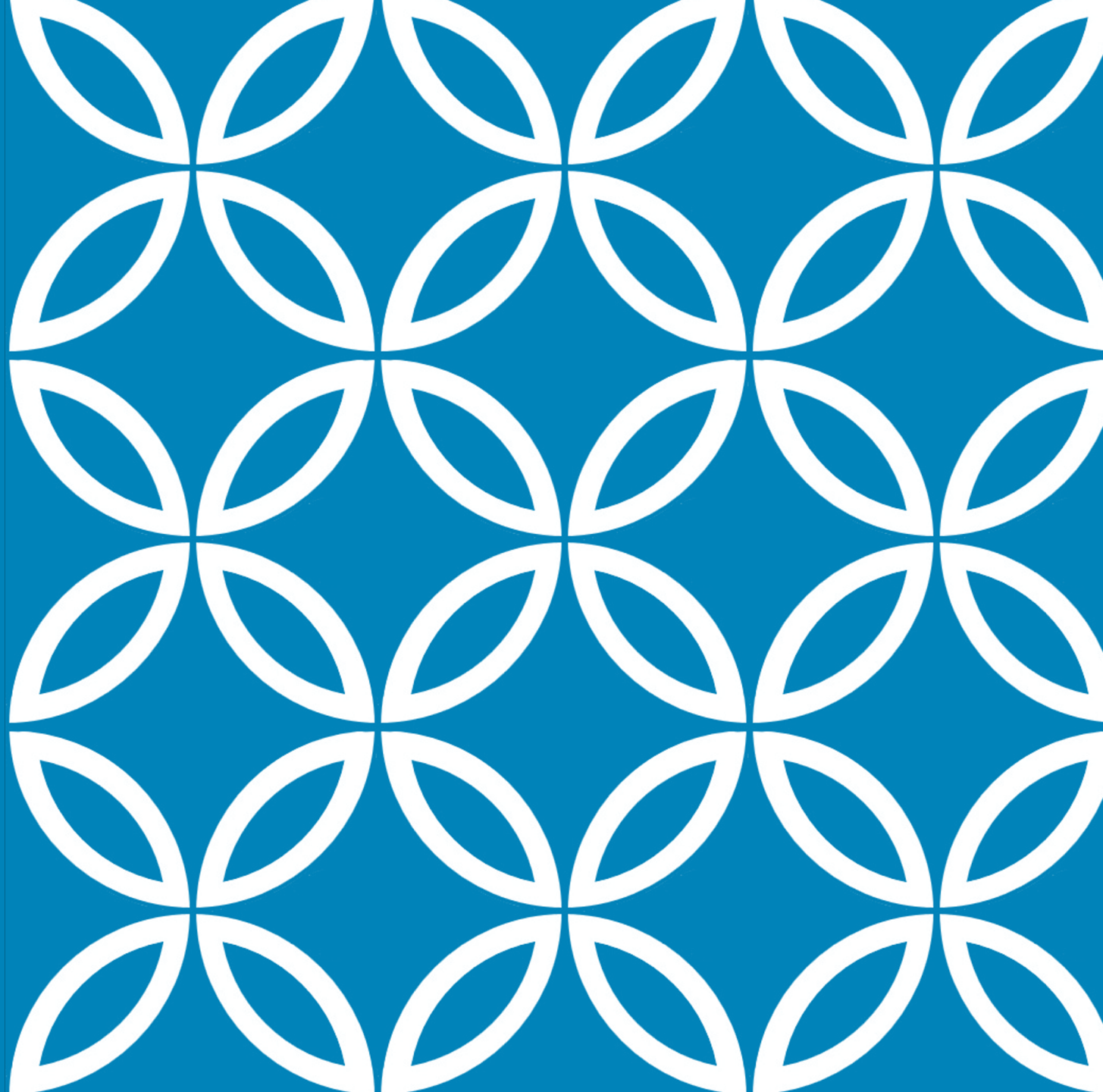# HOUSE PRICES ADVANCED REGRESSION ANALYSIS

Aakash S.

# PROJECT GOAL

Evaluate which of the six models created would give us the best score to predict house sale prices in Ames, Iowa

# STAKEHOLDERS

Home Buyer's

Landlords

Ames, Iowa Village Officials

Realtors

Surrounding Residents

# EXPLORATORY DATA ANALYSIS

On the right, I took the the SalePrice ( target) against the TotalBsmtSF ( total basement square feet). I took this scatter plot to understand the variability for the stake holders. Seen in the plot, as the square foot increased, the SalePrice of the house also increased.
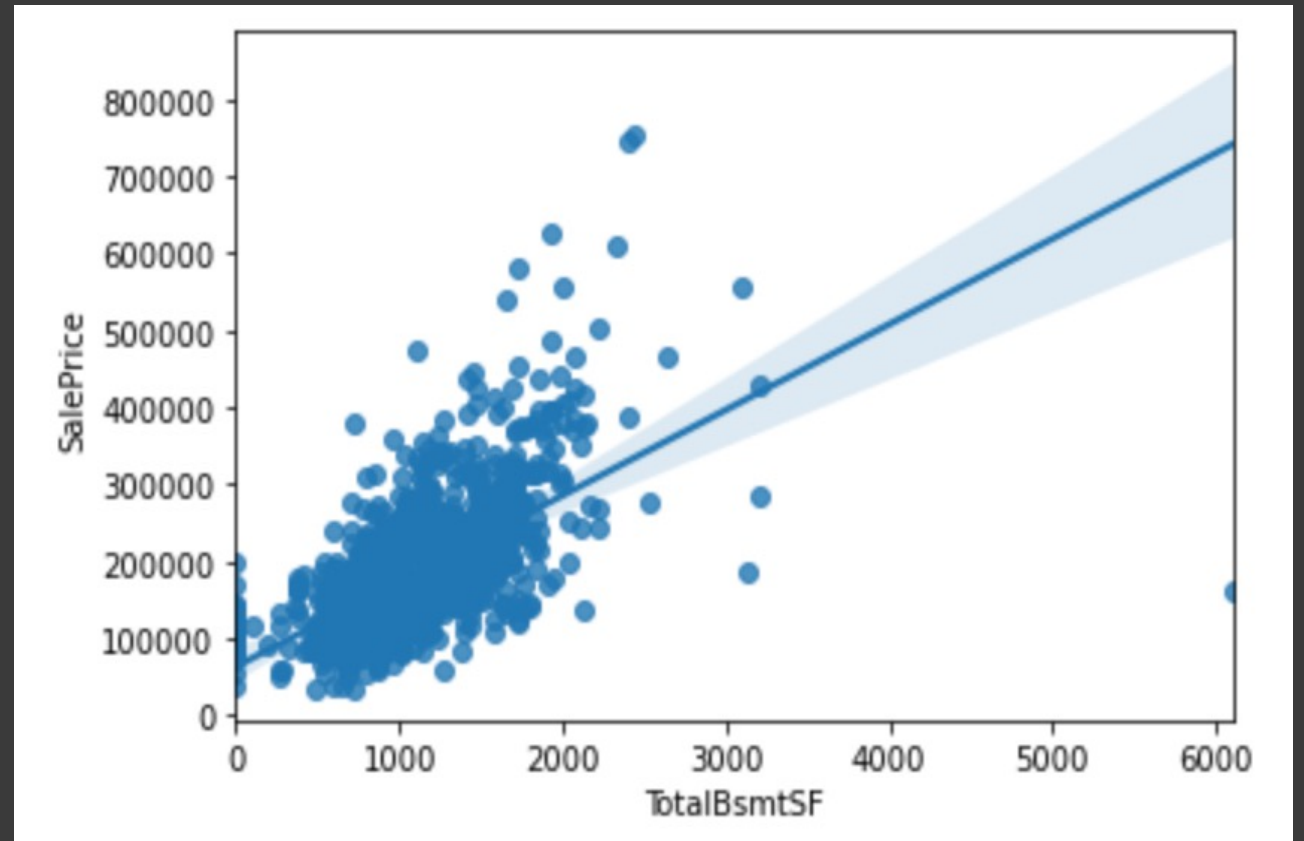
In the scatter plot to the right, we see the positive correlation in the increase of living area and sale price. This helps us understand how the sale price of houses can fluctuate, especially with outliers in the dataset.
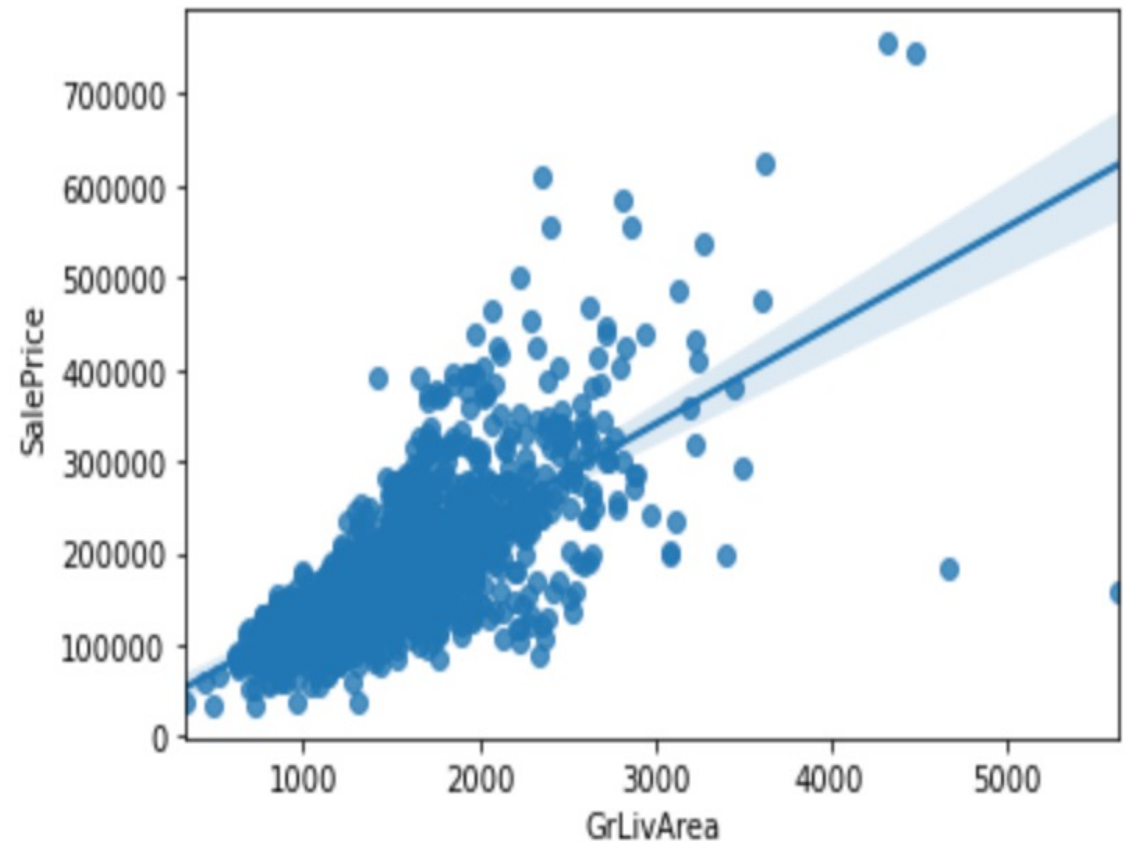
Another plot, demonstrating the positive relationship in the overall Quality. The higher quality, the higher the price.
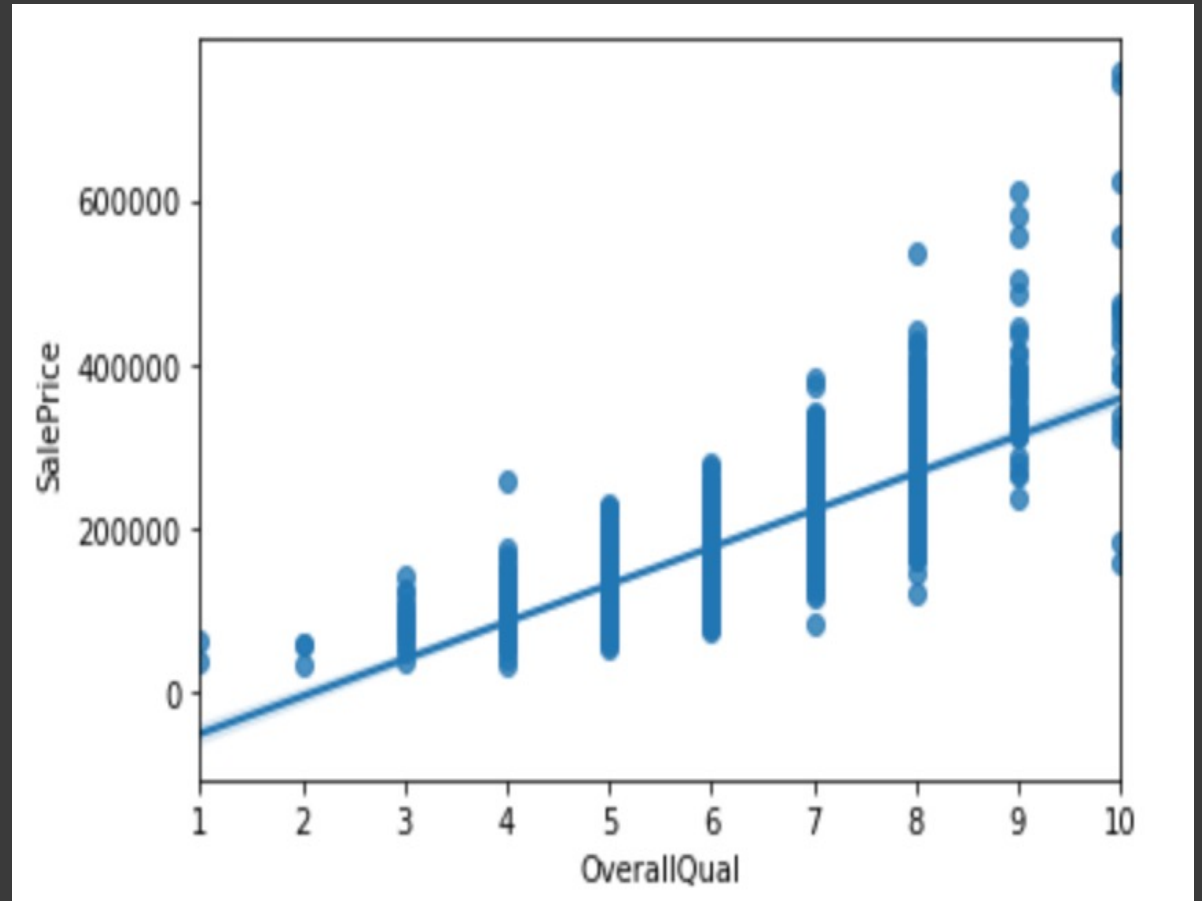
On the right, we have the Sale Type feature and on the left we have the house style feature. Both boxplots are correlated with what a buyer might purchase. In the Sale type box plot, we see a higher mean for the New house type, which can connect to a purchase of the 2Story house. These help us understand how the averages on the x feature can have a price difference on Sale Price.
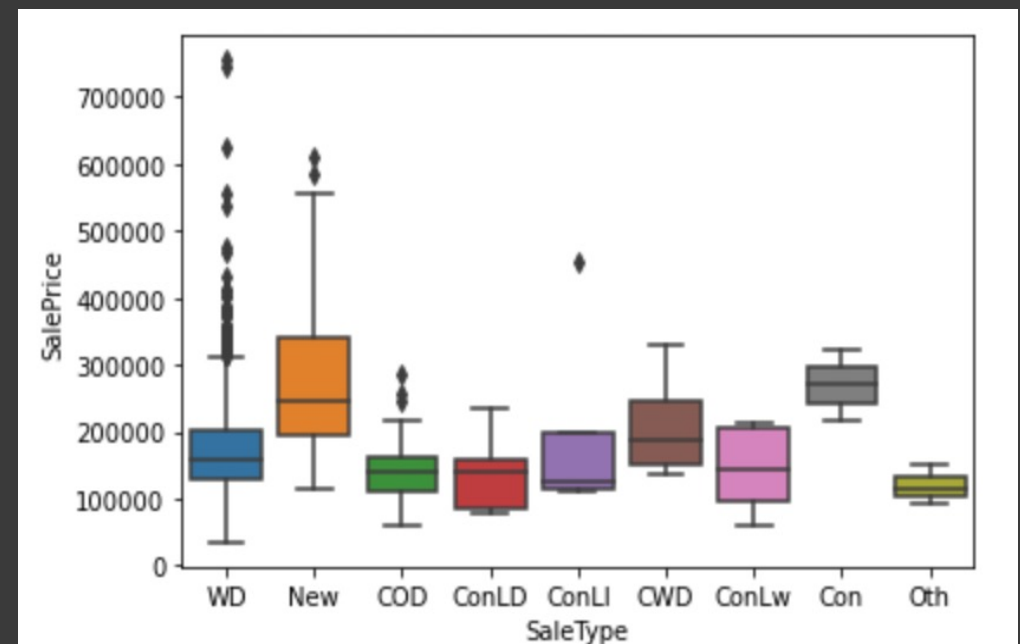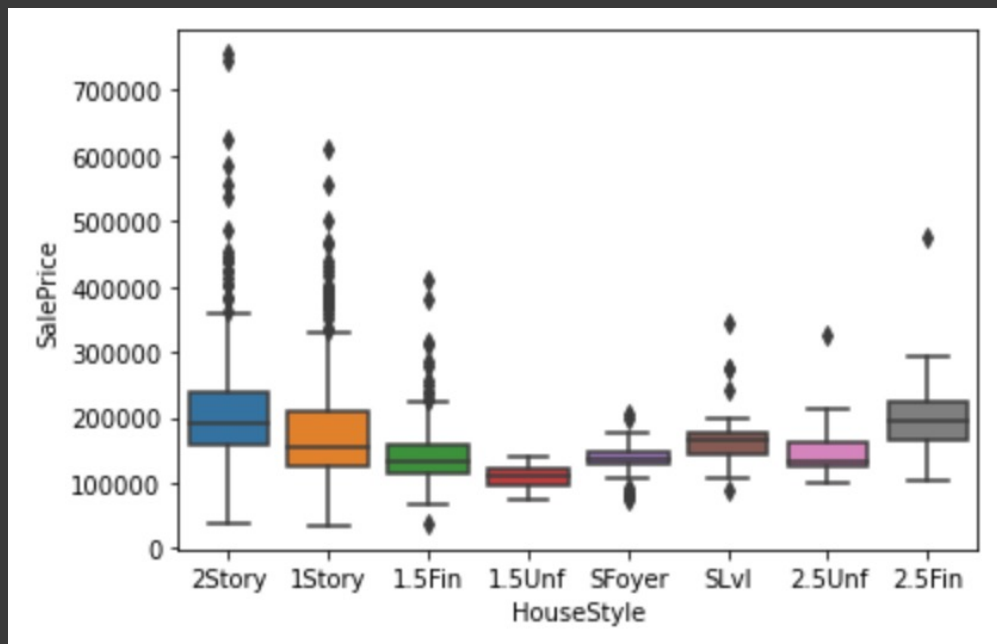
# DATA WRANGLING & PREPROCESSING

Pandas HTML Profiling

Dropped duplicates

Looked for numeric data

Created dummy variable for categorical variables

# MODELING

Linear Regression

L1: Lasso Model

L2: Ridge Model

Random Forest Regressor

Decision Tree Regressor

**Gradient Boosting Regressor**

# CONCLUSION | RECOMMENDATION FOR STAKEHOLDERS

Stakeholders can take away the sale prices and how it may increase or value another home-owner in the area, be able to create new features base on the housing market, and help everyone understand the impact of buying powers.