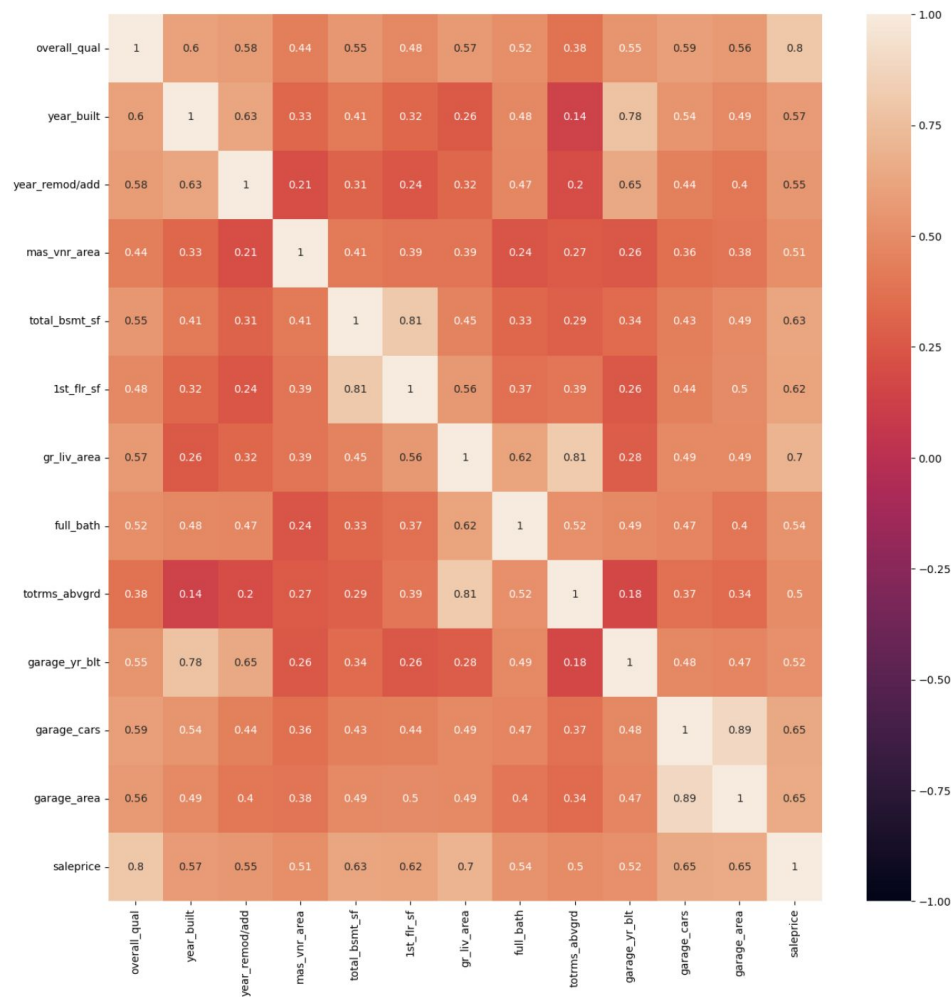# Apply Machine Learning to Predict Ames Housing
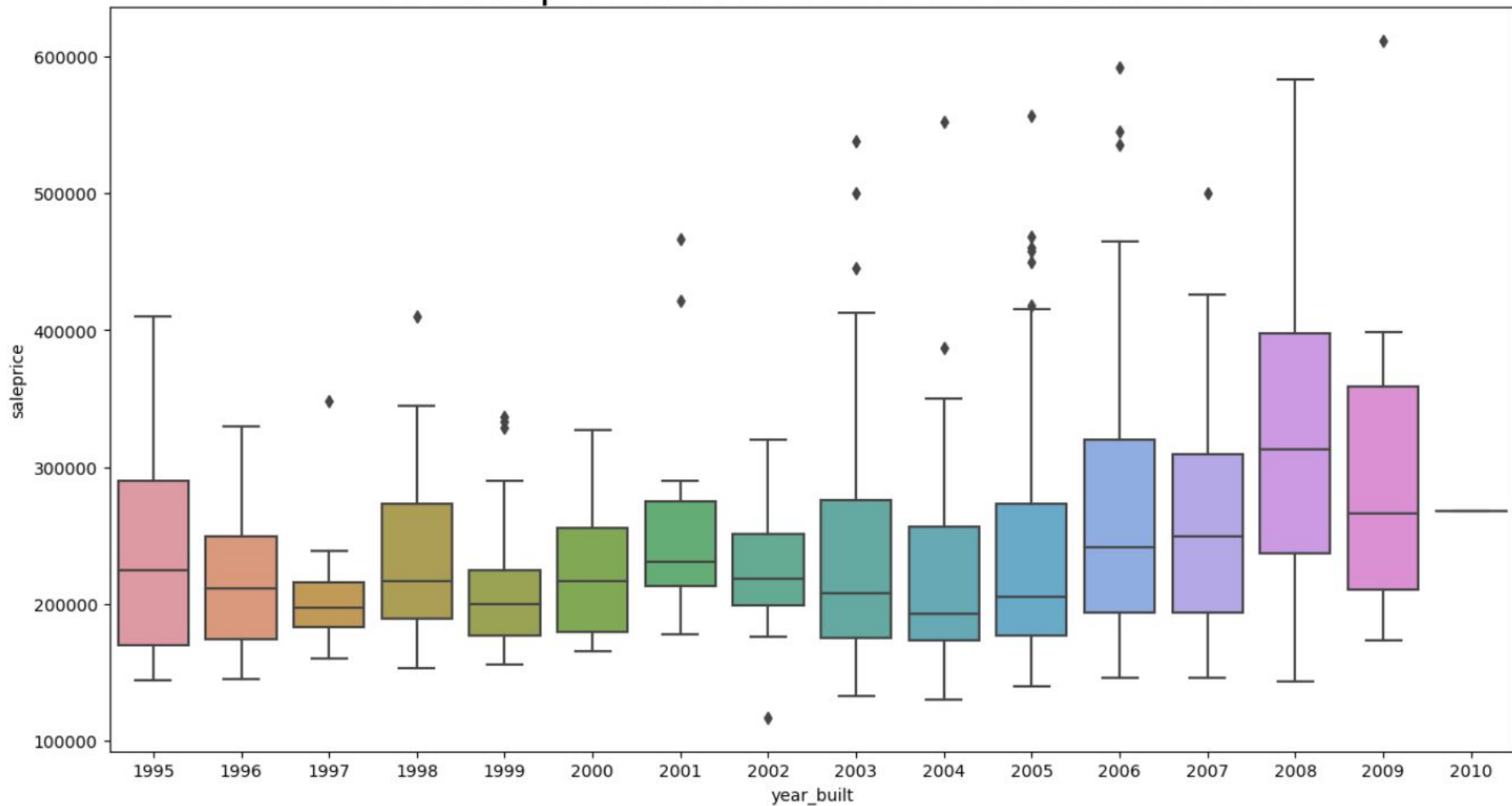
Trang Khong

List of Criteria affect House Pricing

- Original construction date
- The Rates The Overall Material
- Total Square Feet of Basement Area
- Evaluates The Height of The Basement
- First Floor Square Feet
- Total Rooms Above Grade
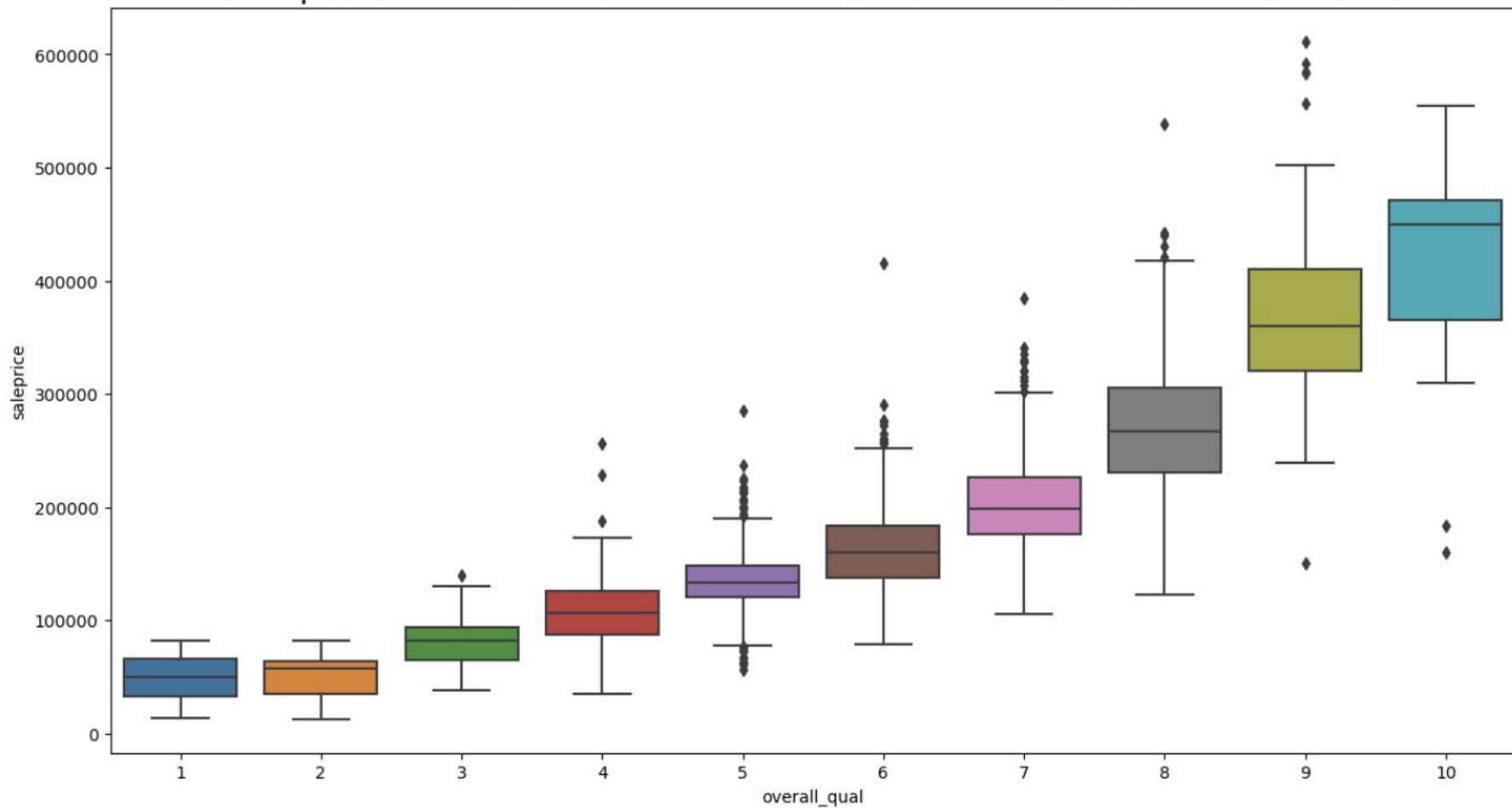- Ground Living Area
- Type of Dwelling
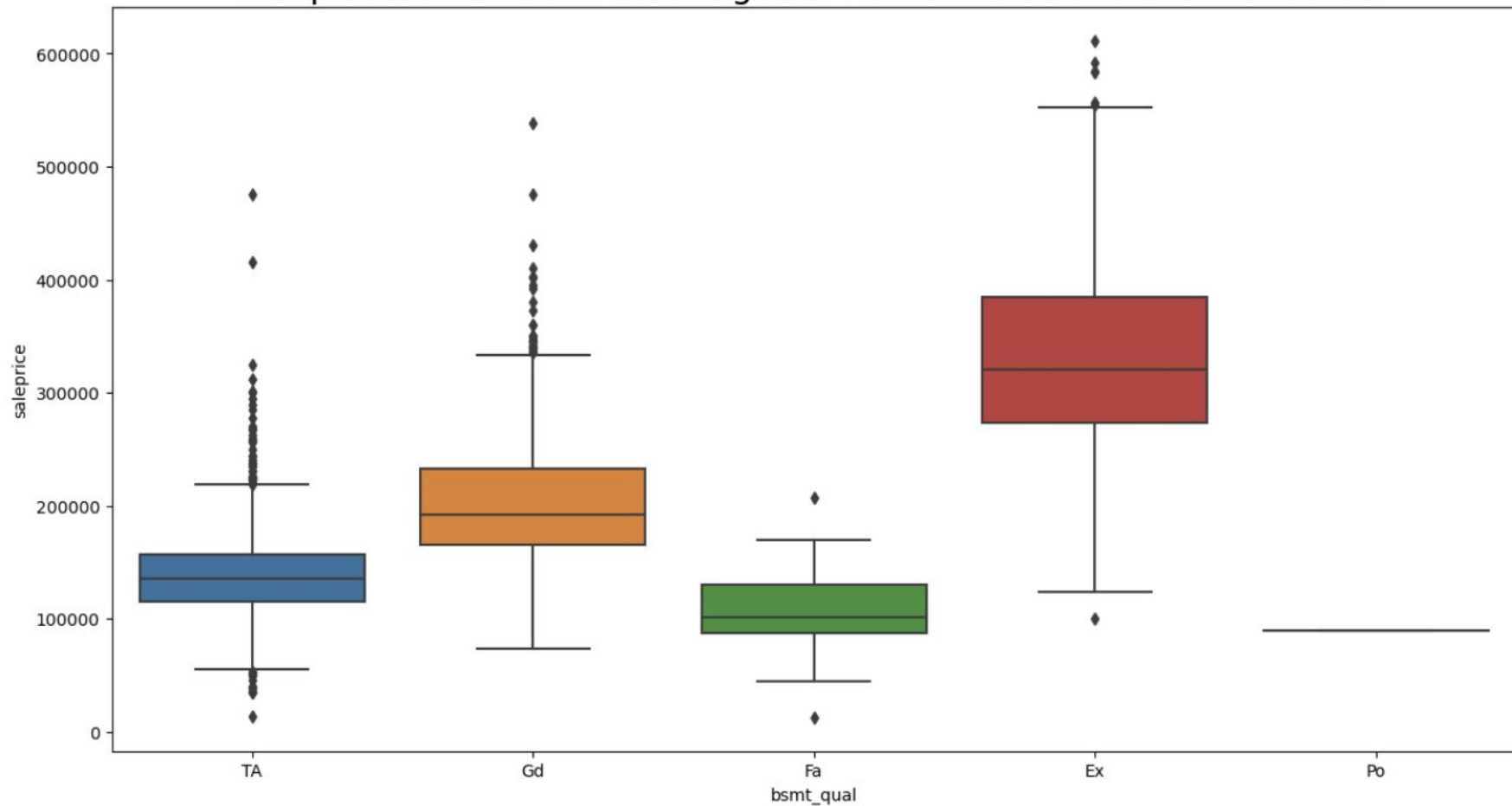
The correlation of sale price with other features

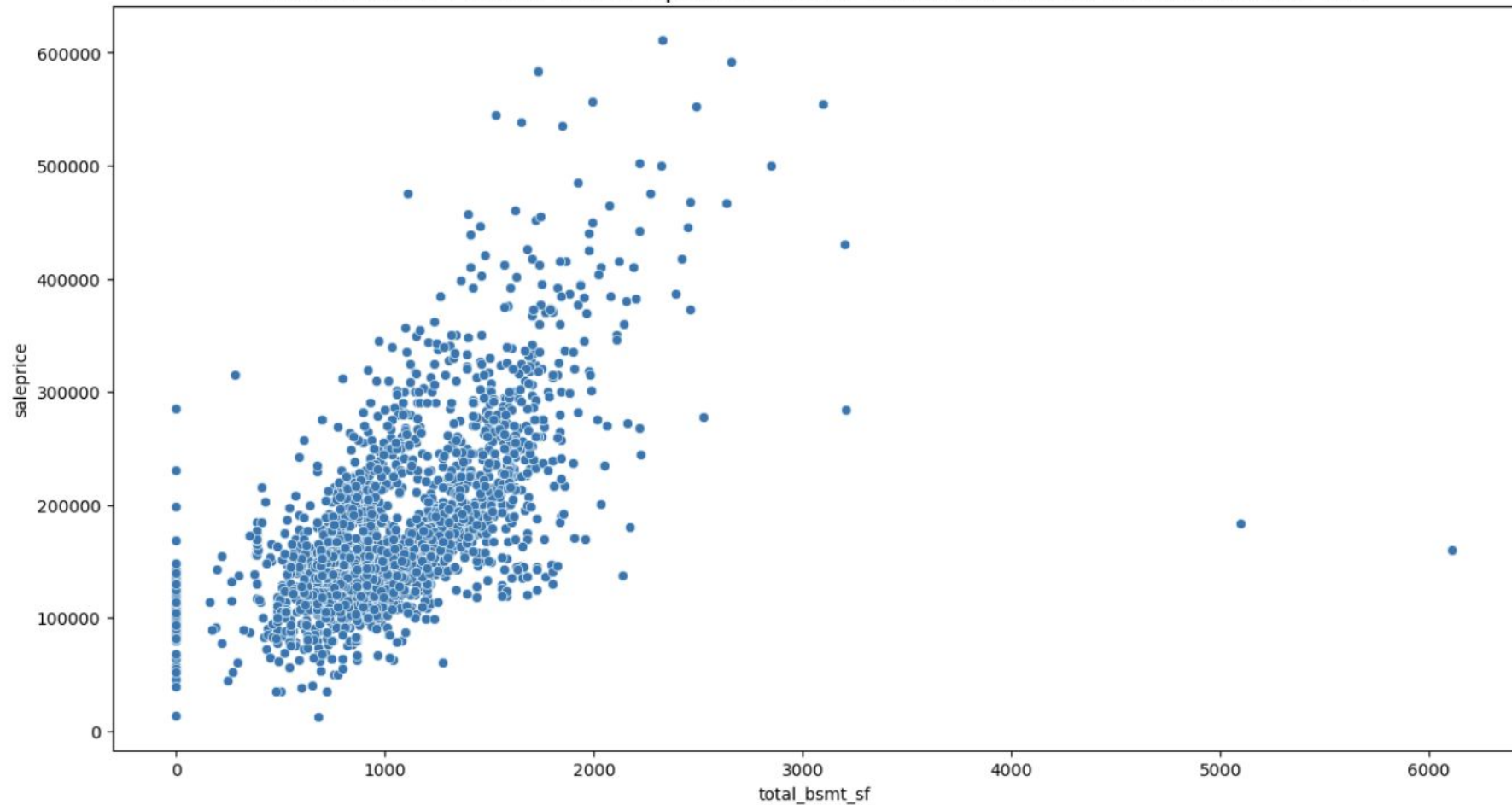The Boxplot of Year Built versus Sale Price

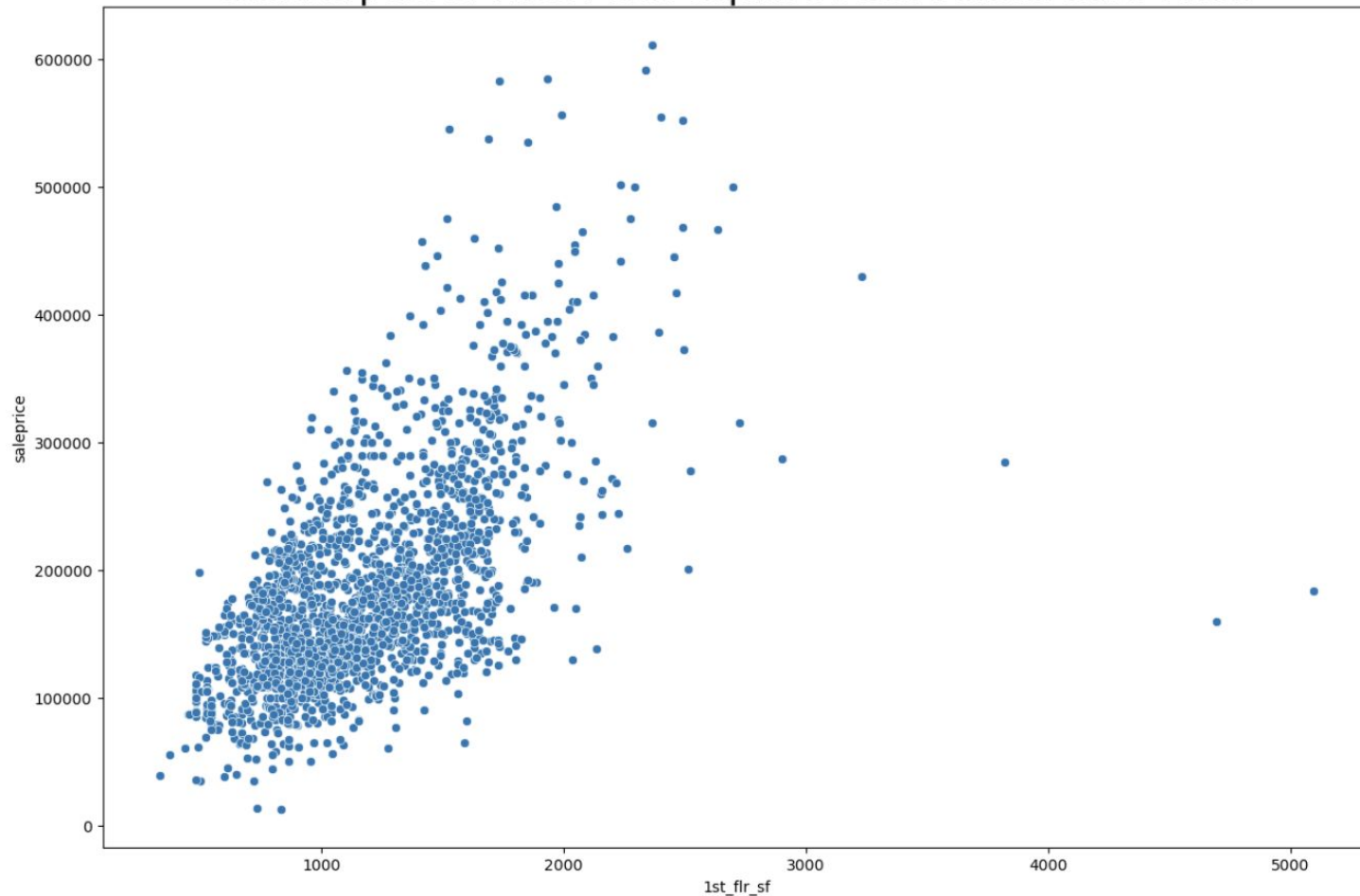The Boxplot of Rates The Overall Material and Finish of The House versus Sale Price

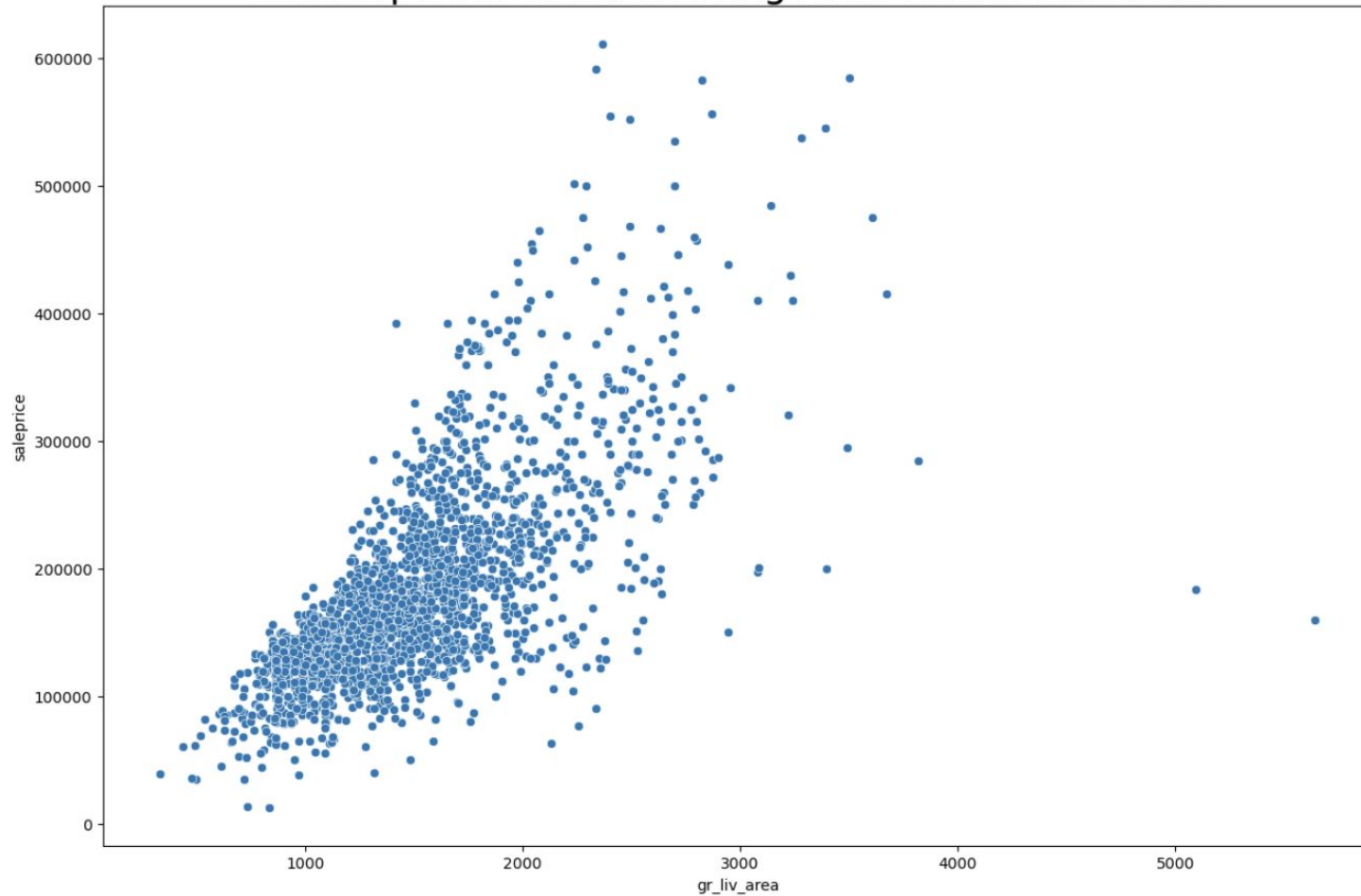Boxplot of Evaluates The Height of The Basement versus Sale Price

The Scatter Plot of Total Square Feet of Basement Area versus Sale Price
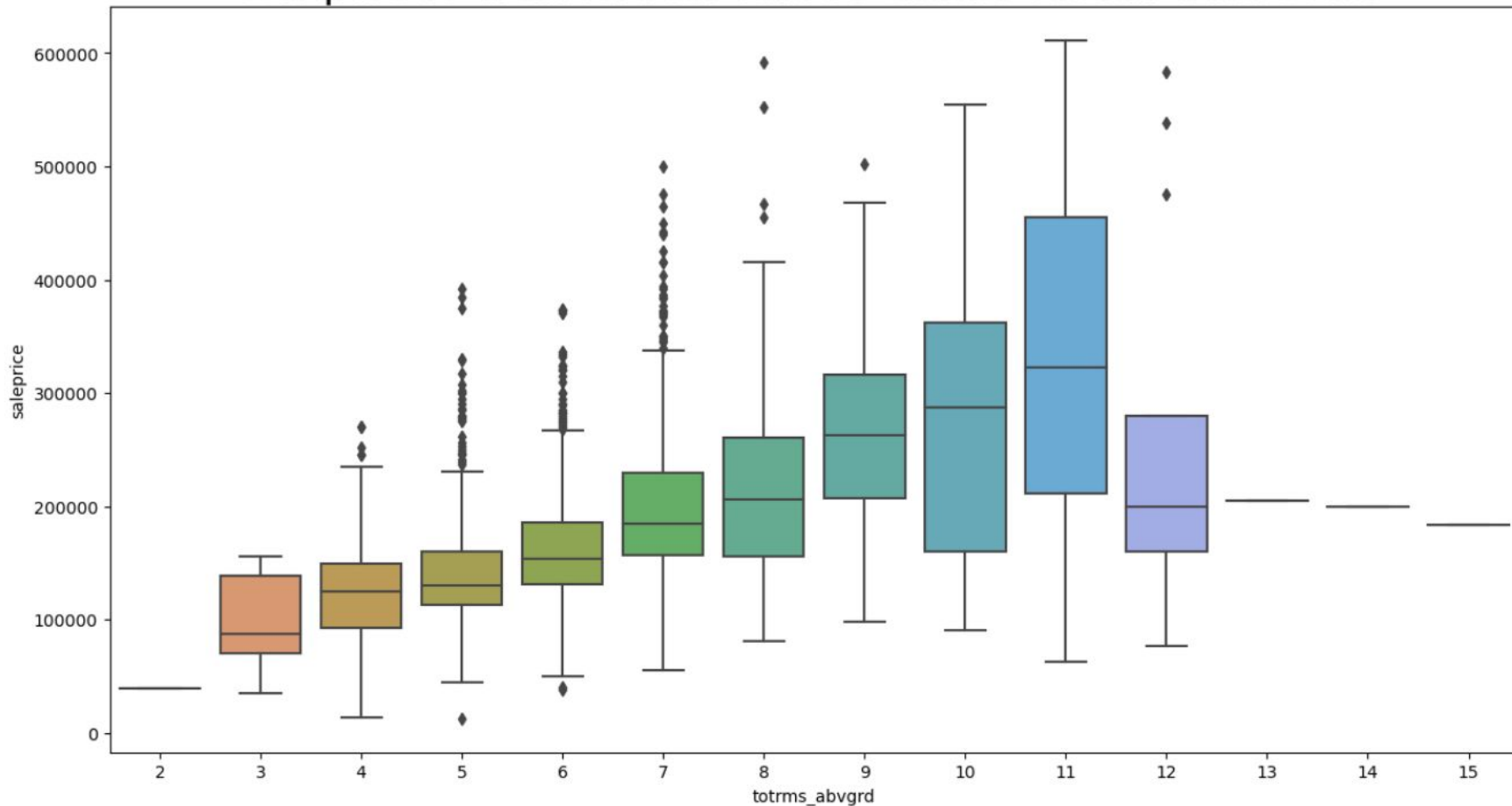
Scatterplot of First Floor Square Feet versus Sale Price
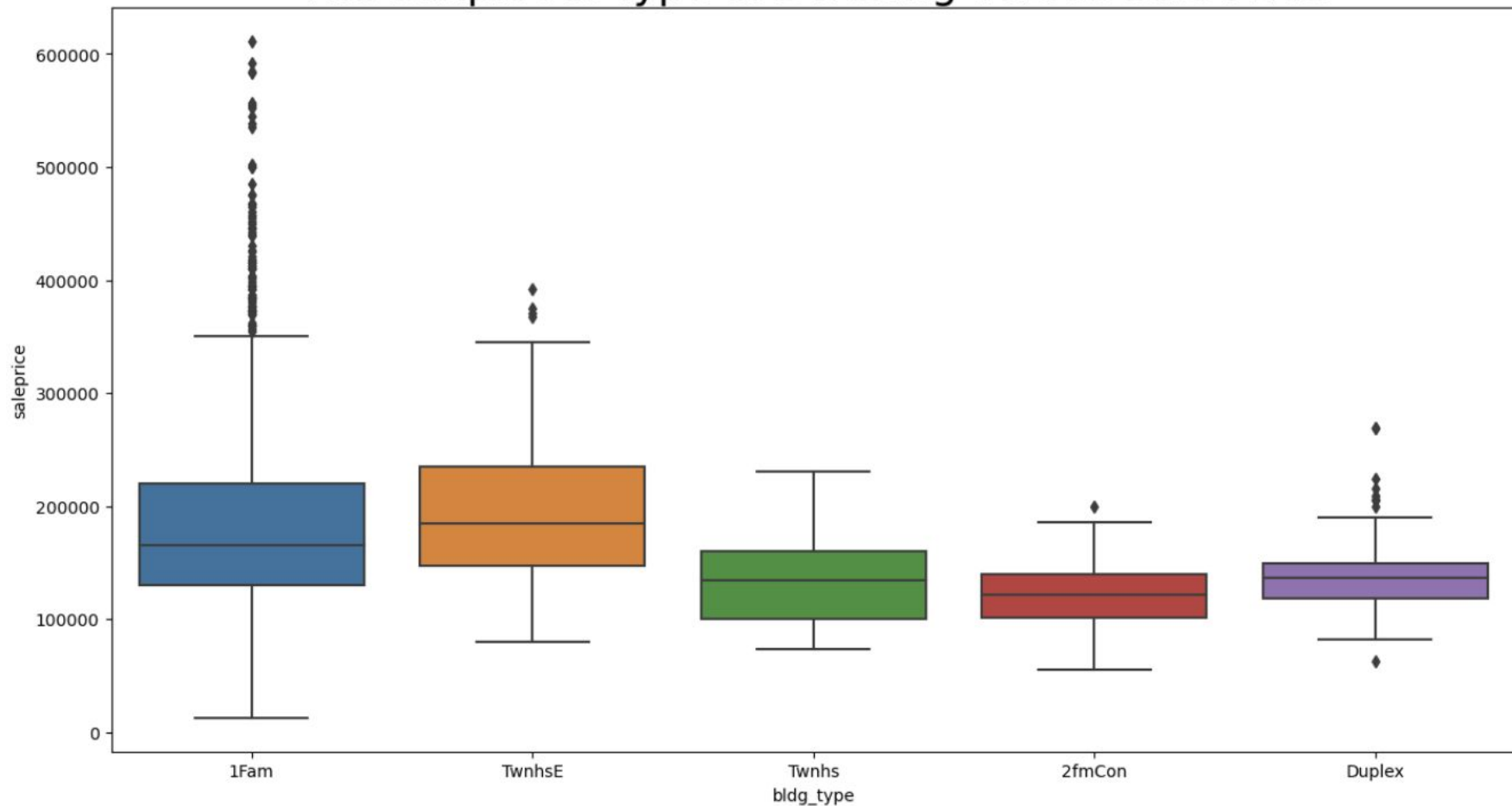
Scatterplot of Ground Living Area versus Sale Price

Boxplot of Total Rooms Above Grade versus Sale Price

The Boxplot of Type of Dwelling versus Sale Price

Imputing and Modeling

❖ Impute missing values with mode for categoricals feature and mean for continuous features.

❖ Model selected to predict ames housing price
  ➢ Linear Regression
  ➢ Lasso Regression
  ➢ Ridge Regression

Discussion

- Models are predicted base on 87 features including some objects features after OneHotEncorder by applying Linear Regression model, Ridge model, and Lasso model.
- The train score of all the models are lower than test score, those models are underfit.
- To improve the underfit of both models, add more features and try some different range of alpha on ridge and lasso model.

Conclusion

- Those models was tested did not have a good performance.
- Apply different model.
- Add more features

Reference

(https://www.forbes.com/advisor/mortgages/real-estate/what-to-look-for-when-buying-a-house/)

(https://www.whitehouse.gov/cea/written-materials/2021/09/09/housing-prices-and-inflation/))

(https://lovehomedesigns.com/where-are-basements-common/#:~:text=On%20average%2C%20in%20the%20United,%2C%20Oklahoma%2C%20Iowa%2C%20Missouri.)

(http://jse.amstat.org/v19n3/decock/DataDocumentation.txt))