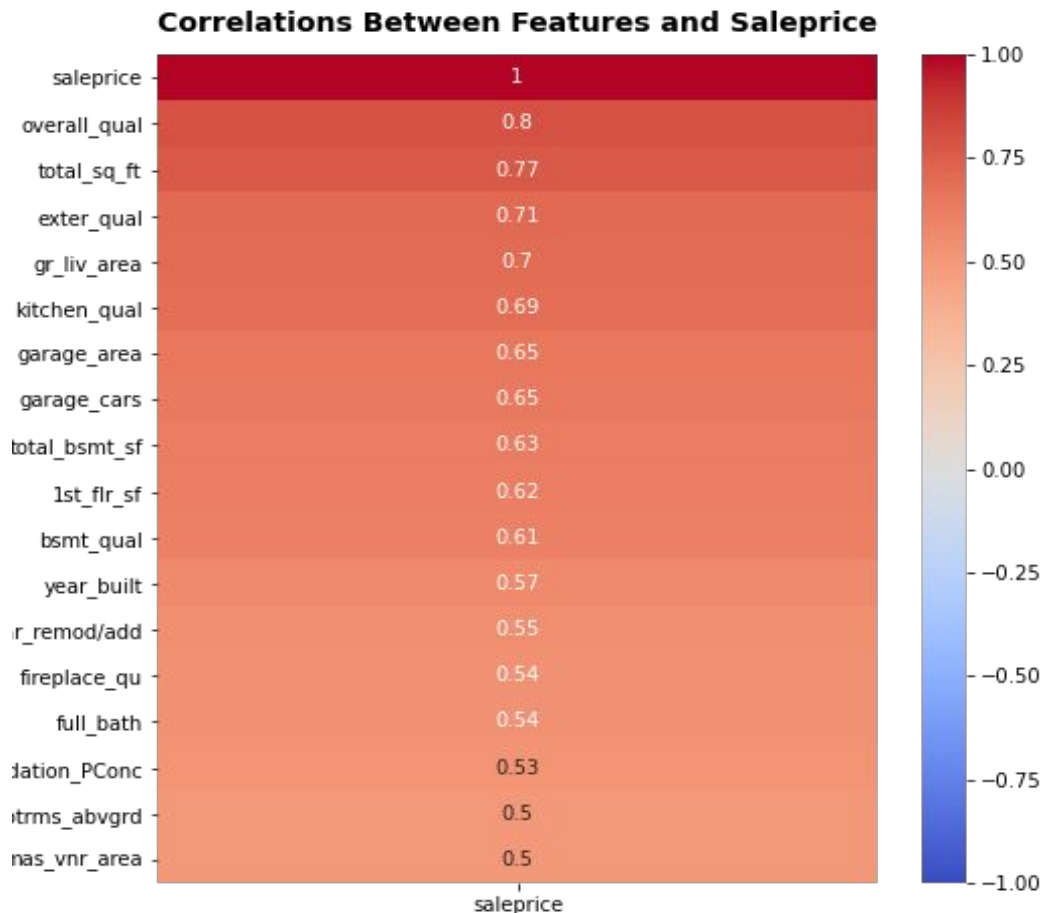# Housing Prices in Ames, Iowa
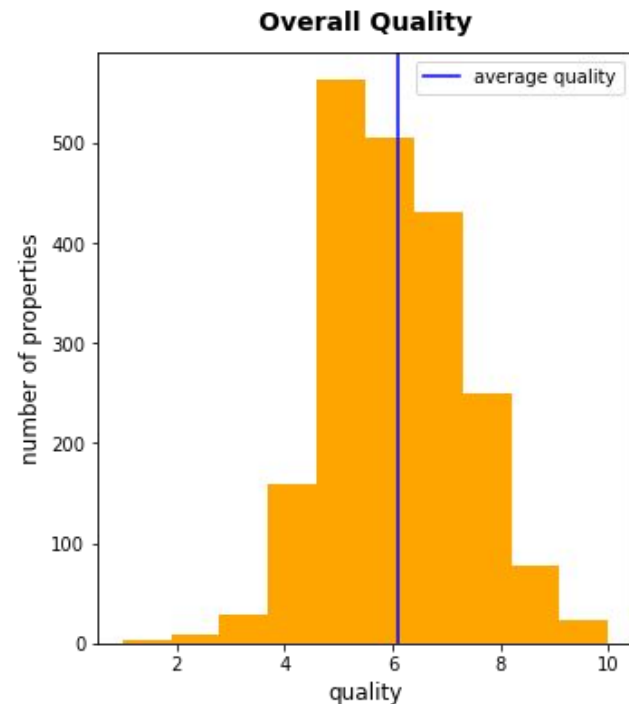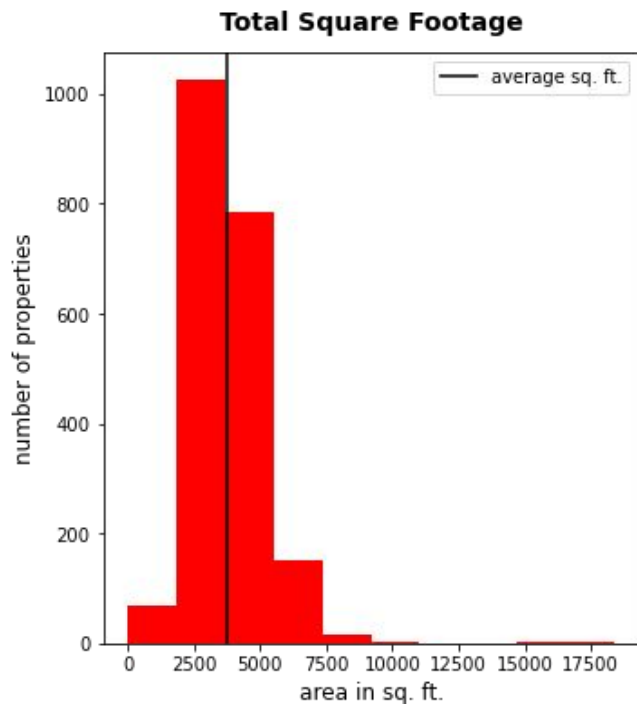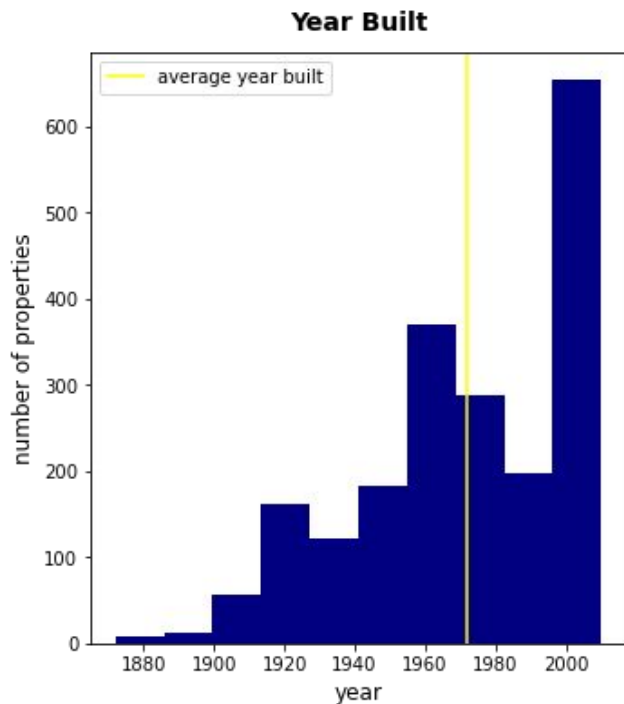
Anthony Zurke

Problem Statement

- There are many things to consider when a person is looking for a home or a homeowner is looking to sell. In this analysis we must find the features that most impact a home's sale price.

Data Correlation

- Numerical features were measured against the sale price.
- Features with correlation greater than 0.5

Correlations Between Features and Saleprice

	saleprice
saleprice	1
overall_qual	0.8
total_sq_ft	0.77
exter_qual	0.71
gr_liv_area	0.7
kitchen_qual	0.69
garage_area	0.65
garage_cars	0.65
total_bsmt_sf	0.63
1st_flr_sf	0.62
bsmt_qual	0.61
year_built	0.57
r_remod/add	0.55
fireplace_qu	0.54
full_bath	0.54
dation_PConc	0.53
trms_abvgrd	0.5
nas_vnr_area	0.5

Data Visualizations

Model Performance

Model	R2 Training Score	R2 Testing Score	Avg. Cross Val Score
MLR	0.928	0.902	0.832
Lasso	0.925	0.905	0.873
Ridge	0.928	0.903	0.851

Data Visualizations

Lasso - Distribution of Residuals
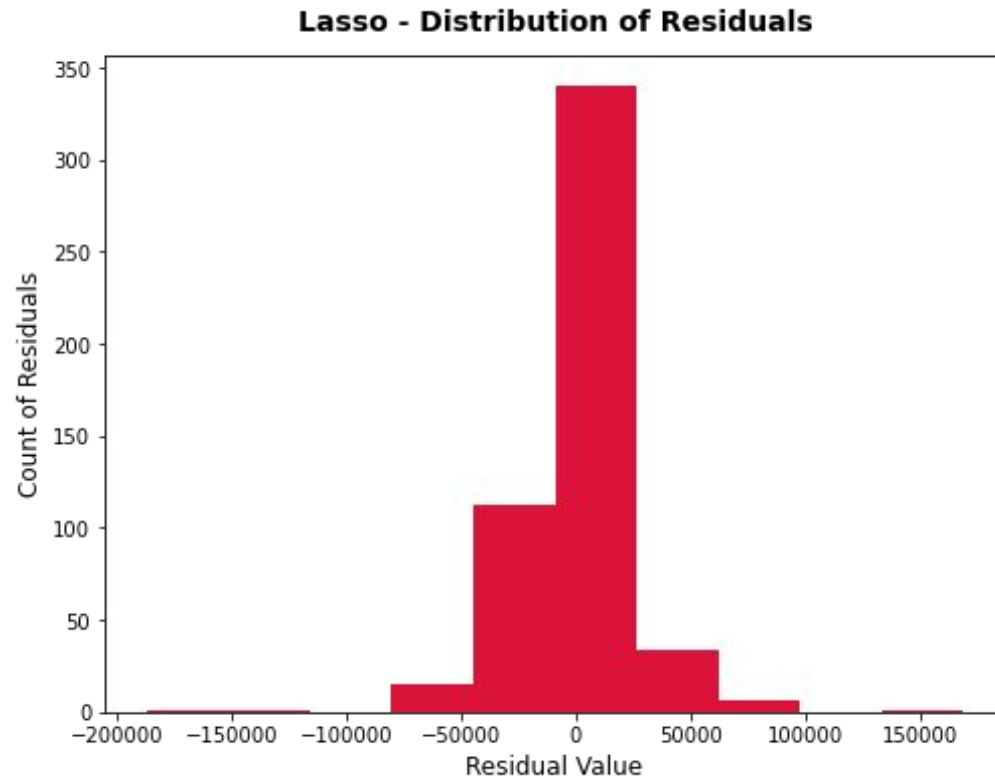
Conclusion

- The Lasso model performed the best with a testing score of .91, and cross val score of .87. The ultimate sale price of a home is highly dependent upon the neighborhood, total square footage, the number of basement bathrooms, the number of main floor bathrooms, the overall quality of the property, the number of fireplaces, and having a paved driveway, all of which are positively correlated with the sale price. The Lasso model is able to predict most sale prices plus or minus $24699.36.