Housing Price Prediction

Brian Nam

INTRODUCTION

Motivation: As a person actively looking for a house was interested in housing prices in San Diego

Objectives:

- Web Scrape sold listings
- Clean the Data
- Create a regression model

Goal: Be able to approximately predict housing price in San Diego

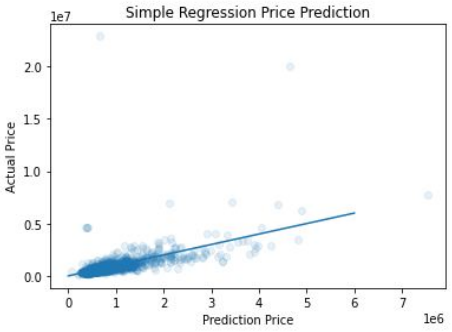
WorkFlow

Data: 8455 sold housing listings between (4/7/21 to 10/12/20) with price as dependent variable and bed, bath, sqft, address, built year, renovated year, lot size, story, garage, parking as independent variables(features)

Outliers

3 houses in La Jolla area with the view of the ocean were sold at more than 20 million dollars.

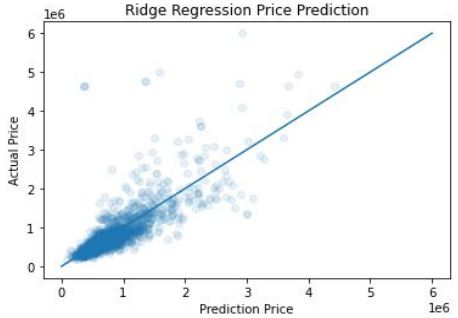
Regression Model



Simple Regression
R^2 = 0.344
mae = 254,243

Feature engineering:

- Neighborhood
- Property Type
- Region



Ridge Regression
R^2 = 0.598
mae = 206,732

High	Mid	Low
P> 75%	25%<P<75%	P<25%

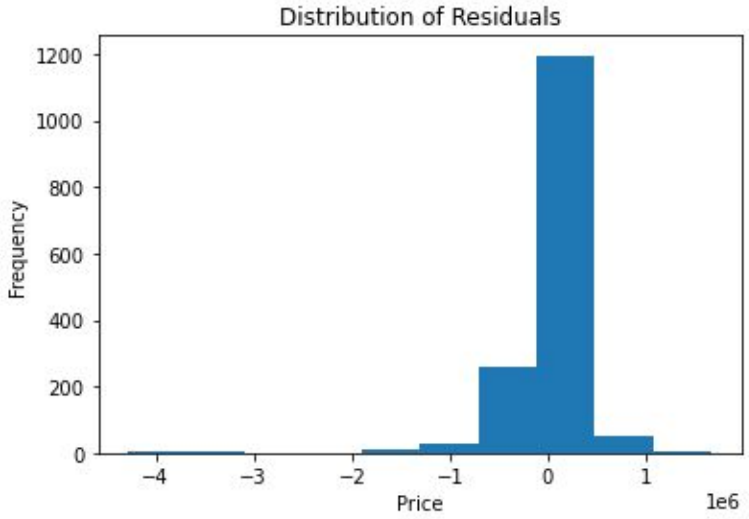
Coefficients

Feature	Scaled Coefficients
Sqft	522605.14
Renovated Year	62722.51
Property Type (House vs Condo/Apt/Townhouse)	41861.66
Region(Central vs Other)	33104.92

Predictions VS Reality

Estimate = 618,494
Real = 618,500
Residual = -6

Estimate = 1,570,339
Real = 5,000,000
Residual = -3,429,661



Conclusion

- Through ridge regression modeling, I was able to account for 60% of the variance in housing price.
- Gained an insight in what factors have the most impact in housing prices.
- According to the value of the mean absolute, the predictions were off by $206,732 which is a third of the standard deviation.
- The model shows weakness in predicting houses with very high values but mostly due to insufficient data.

Fun Fact

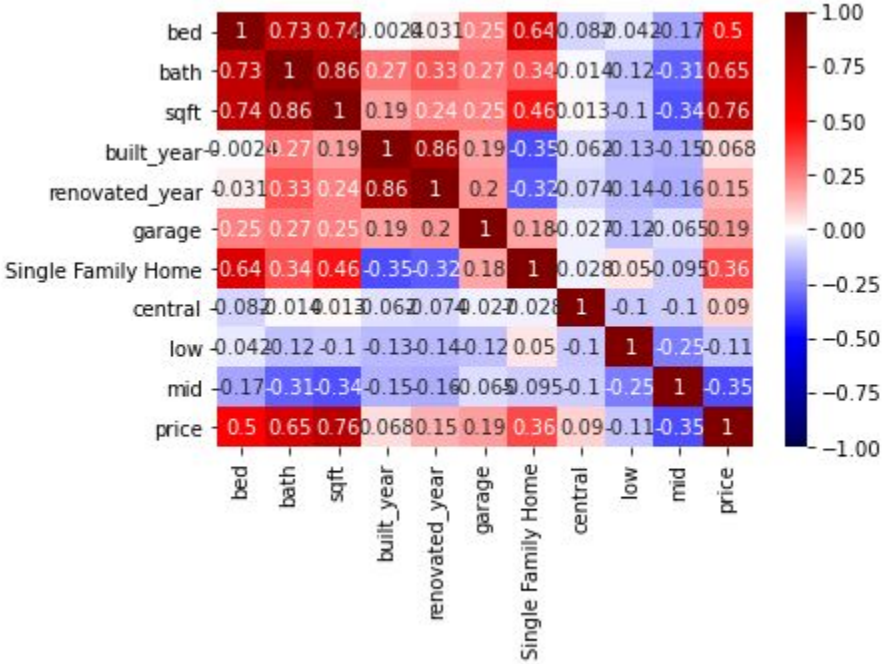
- Currently I am looking for apartments that are 2 bed, 2bath, in the price range of 450,000 to 550,000.
- Filtering through my test set, I found 6 matches.

	property_type	city	price	bed	bath
2655	Condo/Townhome	SD, MISSION VALLEY	457900	2	2.0
3790	Condo/Townhome	SD, MISSION VALLEY	538000	2	2.0
4384	Condo/Townhome	SD, MISSION VALLEY	535000	2	2.5
1051	Condo/Townhome	SD, MISSION VALLEY	488750	2	2.0
1378	Condo/Townhome	SD, MISSION VALLEY	466000	2	2.0
4374	Condo/Townhome	SD, MISSION VALLEY	505000	2	2.0

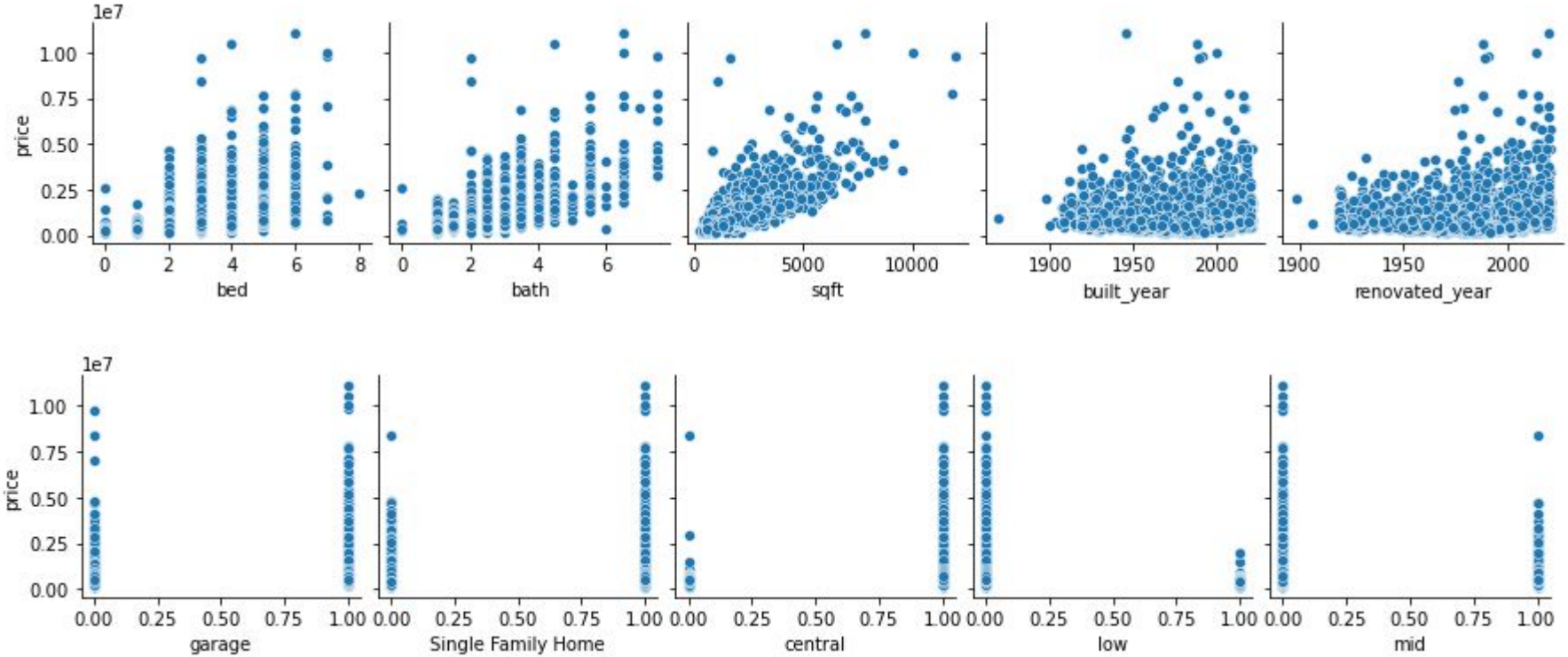
	Real_Values	Predicted_Values	residual
2655	457900.0	515501.896704	57601.896704
3790	538000.0	758898.877126	220898.877126
4384	535000.0	669271.607360	134271.607360
1051	488750.0	515501.896704	26751.896704
1378	466000.0	510098.131042	44098.131042
4374	505000.0	500524.972346	4475.027654

- Mean residual for this dataframe is $81,349 which is significantly lower than the mae for the entire test set.

Appendix


bed
bath
sqft
built_year
renovated_year
garage
Single Family Home
central
low
mid
price
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Appendix


1e7
price
bed
bath
sqft
built_year
renovated_year
garage
Single Family Home
central
low
mid