The King of Home Renovation

By: Grant Edwards

Summary

- King Renovators is looking for a way to find homes that are potential investment opportunities.
- King county Washington, is a large and competitive housing market.
- Use historical data to help predict the homes values by building a linear regression model.

Why Build a Model?

Benefits:

- Identify what features in a home have the largest impact on the price of a home.
- Quickly identify potential properties that are below their predicted price and could be flipped or remodeled for a profit from the known variables.

Business Problem

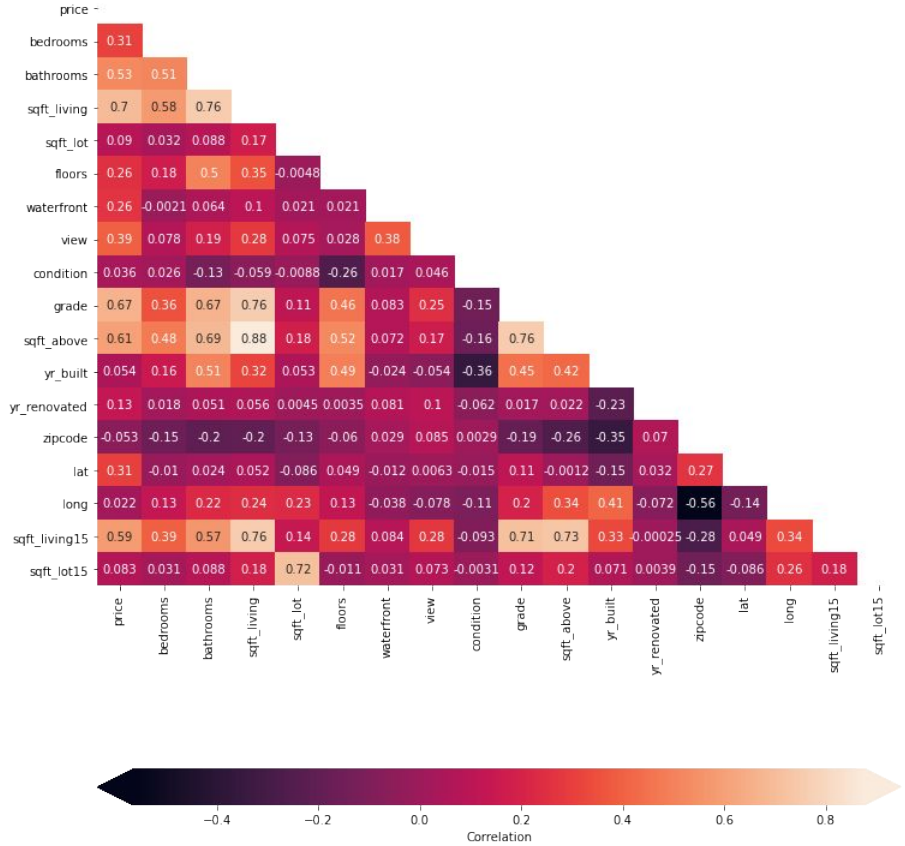
Goal: Using historical data to identify opportunity to renovate and flip homes

- Can we accurately predict the price of a home based on the known variables?
- What features affect a home's value the most?
- Can we use our model to identify homes that are under their predicted market value?

Data

- The data is historical home sales in King County, WA
- Contains information on various features of a home as well as the price.

Initial Look at Correlation of Data



- Focusing on price column we see the correlation to other features of the home.
- Square footage of living space (sqft_living) has highest correlation, with grade having a similar value.

Higher correlated values are lighter, or darker if there is a negative correlation. The closer a value is to 1 or -1, the more significant the variable while being closer to 0 indicates little to no correlation between the dependant and independent variable.

Building a Linear Regression Model

For our first model we ran a simple linear regression without modifying our variables.

- 70% of homes values were based on known variables.

Refining the Model

For our next model we dropped some of the less significant variables, and removed outliers.

- 68% of homes values were based on known variables.

Log of the Price - The Final Model

For our final model we took the log of the home price.

- 69.5% of homes values were based on known variables.
- Key Features
  - Latitude (The more north, the better)
  - Grade (design and quality of the home)
  - Square Footage of Living Space
  - Year Built (Inverse, older homes had higher values)

The Most Influential Variables

- Latitude - Strongest indicator, good way to locate homes, cannot be changed.
- Grade - Measure of design and quality of construction, can be improved.
- Square Footage of Home - Strong indicator, can be increased.
- Year Built - Older homes have higher values, remodel can help with value.

Finding Potential Homes

	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	grade	yr_built	zipcode	lat	resids	pred_price
12539	90000.0	2	1.00	790	2640	1.0	0	0	3	7.0	1973	98034	47.7351	-1.419347	372098.0
2587	134000.0	2	1.50	980	5000	2.0	0	0	3	7.0	1922	98014	47.7076	-1.302061	492700.0
21034	900000.0	5	6.00	7120	40806	2.0	0	4	3	12.0	2007	98006	47.5451	-1.265836	3191450.0
4020	800000.0	7	6.75	7480	41664	2.0	0	2	3	11.0	1953	98166	47.4643	-1.165535	2566112.0
326	274975.0	3	2.50	3030	45004	2.0	0	0	3	9.0	1987	98077	47.7721	-1.159401	876626.0
1220	130000.0	3	1.00	1110	7520	1.0	0	0	4	7.0	1960	98033	47.6830	-1.123715	399914.0
12332	160000.0	2	1.00	1140	23030	1.0	0	0	3	8.0	1980	98028	47.7637	-1.093386	477498.0
14255	130000.0	2	1.00	840	6654	1.0	0	0	3	7.0	1951	98133	47.7319	-1.081848	383517.0
16828	170000.0	1	0.75	850	5600	1.0	0	2	3	6.0	1903	98019	47.7654	-1.045246	483497.0
18318	130000.0	3	1.00	1200	7000	2.0	0	0	1	7.0	1908	98116	47.5883	-1.043729	369172.0

Looking at the homes with the highest negative residual values, we can find some homes that would have potential for a remodel or resale for a profit.

Conclusions

- Through our linear regression model we could predict ~70% of the home price.
- We found that the latitude, grade, square footage, and year built were the strongest predictors for the homes price.
- We identified homes that are listed below their predicted value, shortlisting potential business opportunities.

Thank You!

Email: grantedwards11@gmail.com
GitHub: @gzedwards
LinkedIn: www.linkedin.com/in/grant-edwards-25206914a