

House Sales in King County, Washington, USA

Linear Regression Exploration

Miriam Semmar

Overview

- As we look to launch our construction business, we're gathering information on where we should build and which features to focus our efforts on in order to maximize profits.
- For this analysis, we're leveraging data from the home sales in King County
 - Includes over 20,000 homes sold in 2014 or 2015 for anywhere between \$78k and \$7.7MM.

Key Questions

- Which features of a home have the greatest impact on home price?
- Which areas of King County should we consider building in?
- How can we improve the potential sale price of a home?

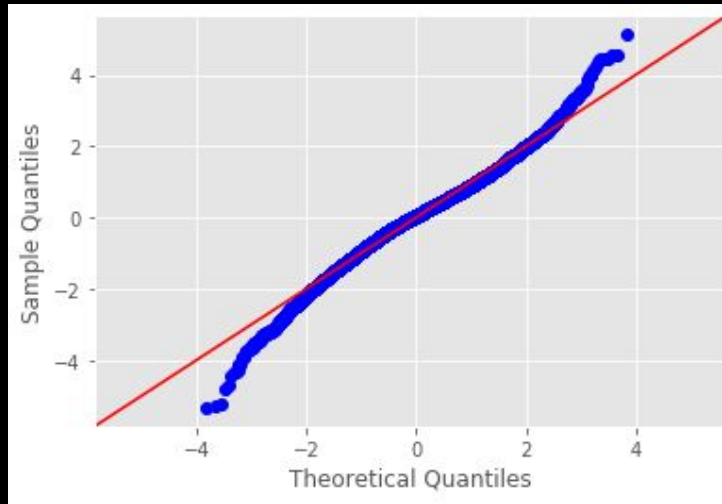
Approach

- Removal of outliers and conducting statistical transformations to normalize data
- Removal of features that did not have sufficient evidence to suggest that a correlation exists (high p-values)
- Dropping features that had a low correlation with price



Final Model

- Our model can explain 73.6% of home price variance.
- The most impactful features are location, having a basement, home grade and the living area of a home.



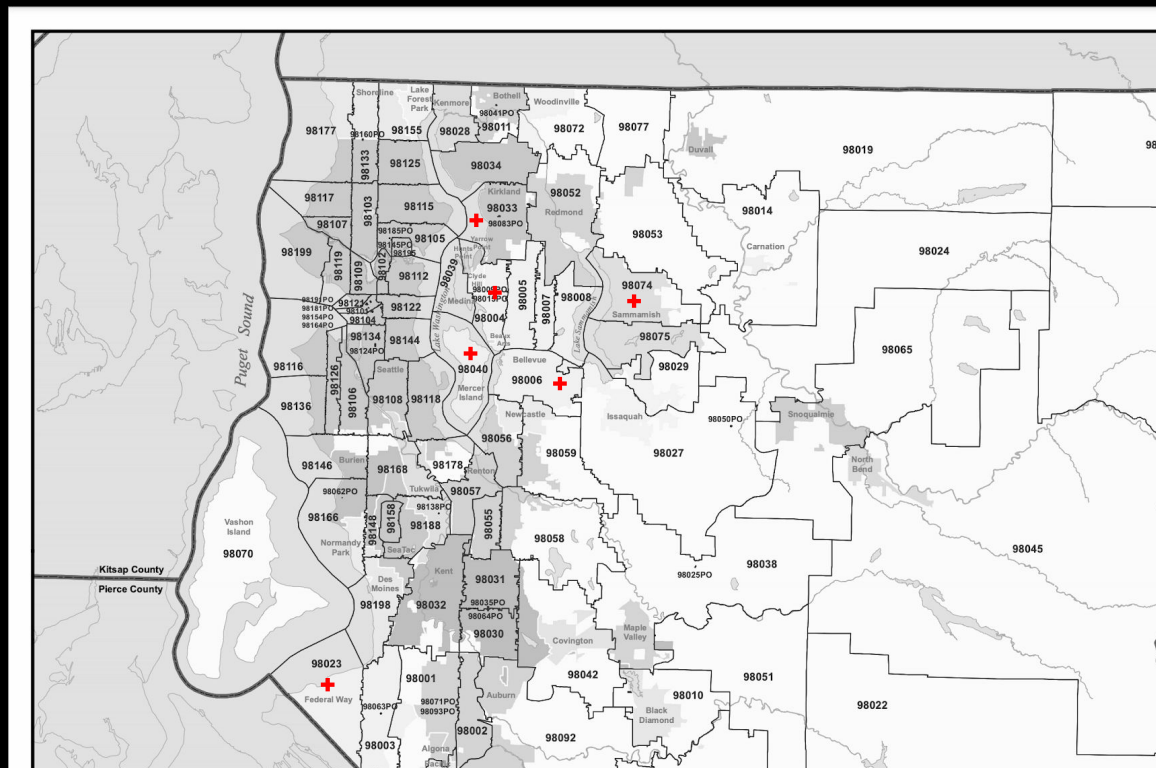
OLS Regression Results

Dep. Variable:	price	R-squared:	0.736
Model:	OLS	Adj. R-squared:	0.735
Method:	Least Squares	F-statistic:	903.1
Date:	Sat, 31 Oct 2020	Prob (F-statistic):	0.00
Time:	14:01:40	Log-Likelihood:	590.01
No. Observations:	15953	AIC:	-1080.
Df Residuals:	15903	BIC:	-696.2
Df Model:	49		
Covariance Type:	nonrobust		

Top 15 Features	coefficients
Intercept	9.807862
zipcode_98004	0.695797
zipcode_98040	0.496420
sqft_living	0.430077
zipcode_98033	0.416055
zipcode_98006	0.281822
zipcode_98075	0.252652
view_2_0	0.134445
has_basement_1_0	0.014204
grade_9	-0.085734
grade_8	-0.240607
grade_7	-0.389638
zipcode_98023	-0.391342
grade_6	-0.506110
grade_5	-0.609204

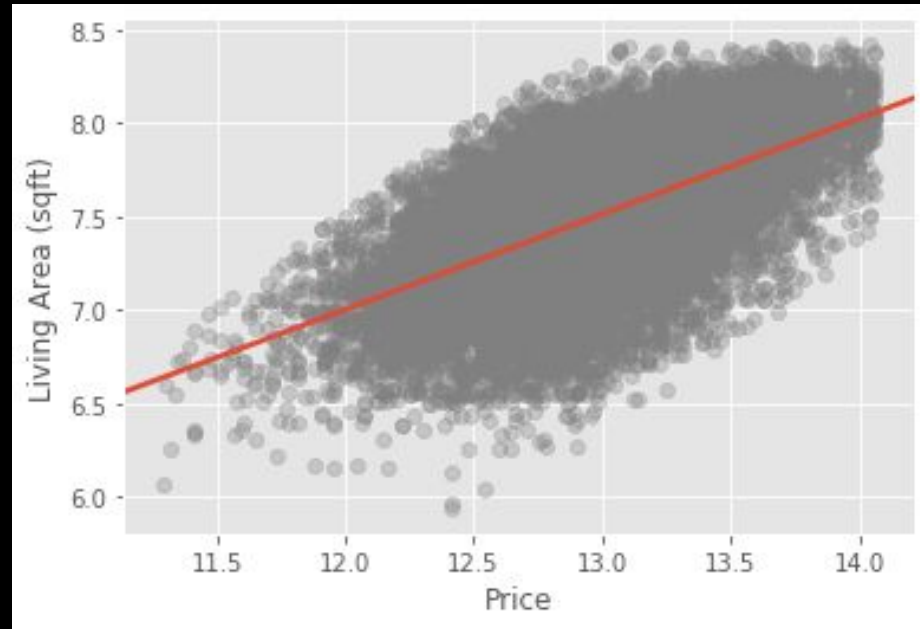
Location, location, location!

ZIP Code	Location	Coefficient	Change in Price (%)
98004	Bellevue, WA	0.70	101%
98040	Mercer Island, WA	0.50	64%
98033	Kirkland, WA	0.42	52%
98006	Bellevue, WA	0.28	33%
98075	Sammamish, WA	0.25	29%
98023	Federal Way, WA	-0.39	-32%



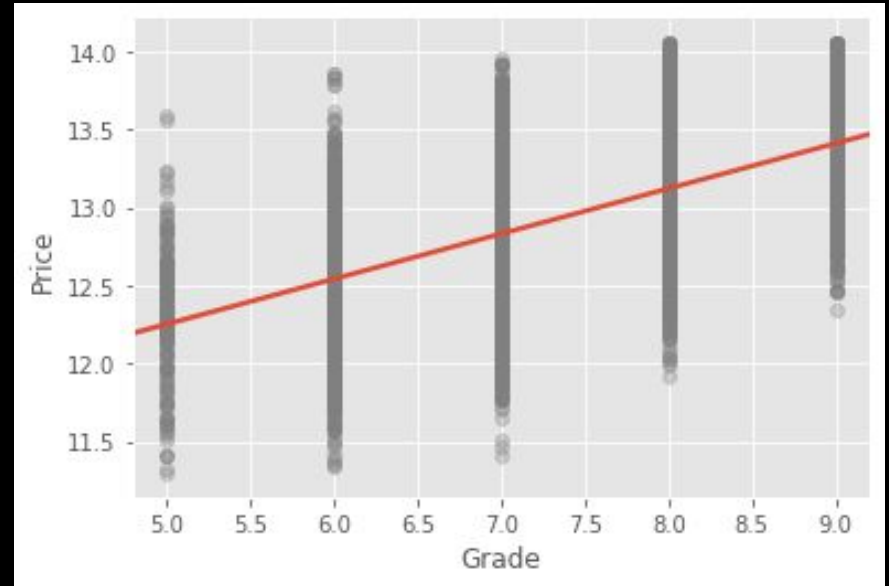
Space in the Home

- Linear relationship between price and living area.
- Model result: a 1% increase in square footage increases home price by 0.4%.
- Having a basement also increases home price by 1%.



Construction Quality (Grade)

Grade	Coefficient	Change in Price (%)
9	-0.09	-8%
8	-0.24	-21%
7	-0.39	-32%
6	-0.51	-40%
5	-0.61	-46%



Conclusions

- **Which features of a home have the greatest impact on home price?**
 - Location, size of the home and construction quality
- **Which areas of King County should we consider building in?**
 - Areas closer to Seattle: Bellevue, Mercer Island and Kirkland
- **How can we improve the potential sale price of a home?**
 - We don't only have to build homes on empty lots. We can also consider conducting remodels and focus on using higher grade materials as well as building extensions to increase the square footage of a home.

Next Steps

- Update our dataset to include more recent purchases
- Consider different methods to determine the impact of the number of bedrooms and bathrooms
- Expand our dataset to include neighboring counties to validate results as well as assess additional opportunity
- Recreate the model to include information related to:
 - School districts
 - Proximity to shopping centers
 - Proximity to transportation hubs
- Research upcoming attractions as well as local laws and regulations that may present obstacles for the business

Thank You!

Sources

- King County Dataset (Kaggle)
- US Zipcode Package (created by Sanhe Hu, <https://uszipcode.readthedocs.io>)
- Map Image:

<https://aqua.kingcounty.gov/gis/web/VMC/boundaries/zipcodes/zipcodes.pdf>