How to Maximize Your Home Selling Price

Kings Country, WA edition

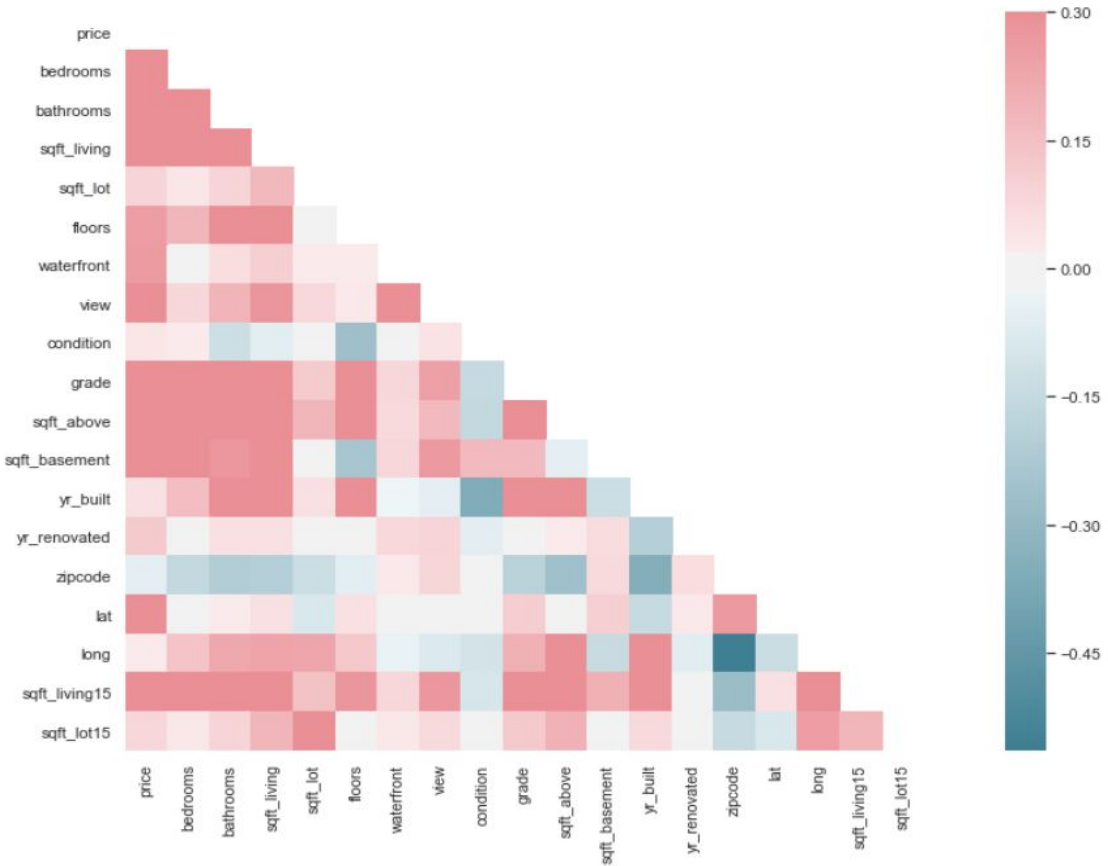
KC Home Price Potential Predictors:

- Number of Bedrooms
- Number of Bathrooms
- Square Footage of Living Area
- Average Square Footage of the Living Area of the 15 Closest Houses
- Square Footage of Total Lot of the House
- Average Square Footage of the Total Lot of the 15 Closest Houses
- Total floors of the House
- Whether or not the House is on a Waterfront
- View (Index 0 to 4 on how good the view of property is)
- Condition of House (Index 1 to 5)
- Grade (Index 1 to 13, where 1-3 falls short of building construction and design, 7 has an average level of construction and design, and 11-13 have a high quality level of construction and design)
- Square Footage of the House Above Ground Level
- Square Footage of the Basement
- The Year the House was Built
- The Year the House was Renovated (if applicable)
- Zip Code
- Latitude of the House
- Longitude of the House

Model Methodology:

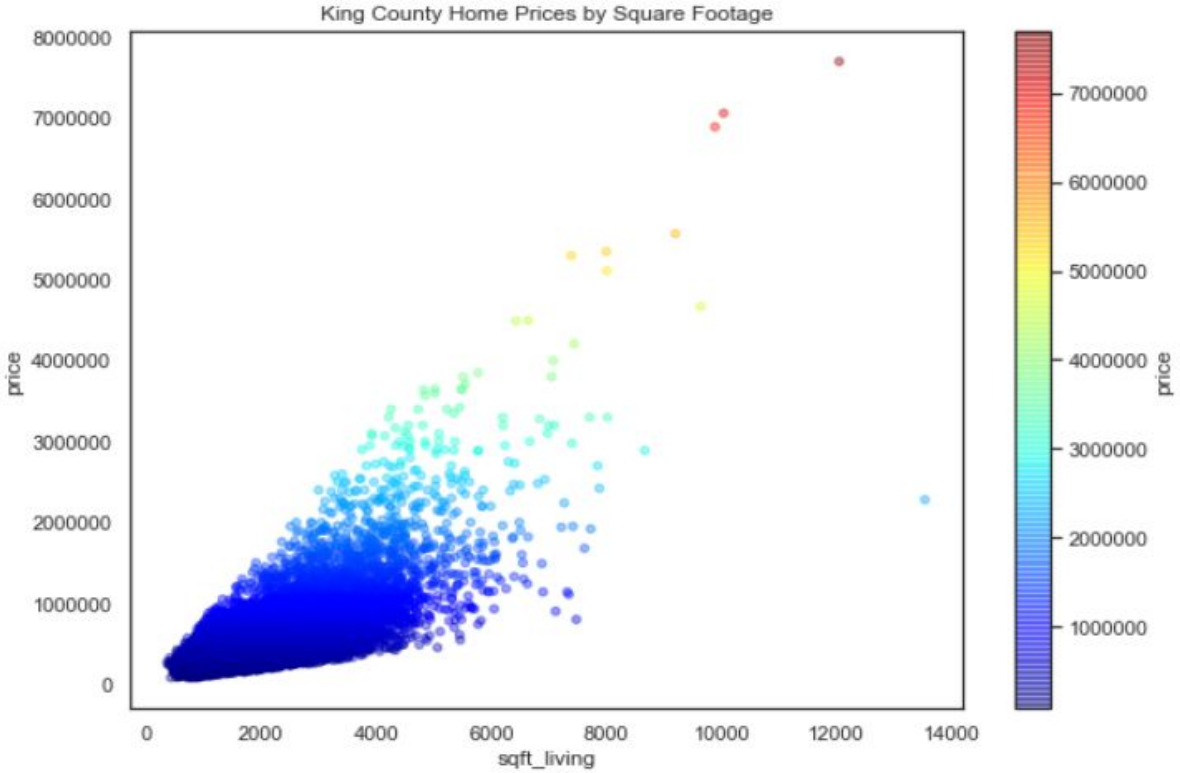
- First the data was scrubbed by removing outliers, null values, invalid entries, incorrect formatting and missing values
- Using the newly cleaned data, a regression analysis was run using the top 7 predictor variables for home prices from the previous slide
- When necessary, data was normalized for better distribution and all variables were scaled to a 0-1 scale to avoid incorrectly over-weighting variables with larger magnitudes
- Predictor variables with high correlation to other predictor variables were removed from model to prevent double counting
- After iteratively running the regression model, an optimal group of features (7 in this model) were selected to predict Kings County Home Prices
- The model was tested using k-fold cross validation method, which separates the data into test and train sets and runs k-number of tests to return the mean squared error (MSE) per iteration. The average MSE of the tests is the score of the model's accuracy
- The following slides will walk through the best predictors of KC home prices

Correlation Matrix



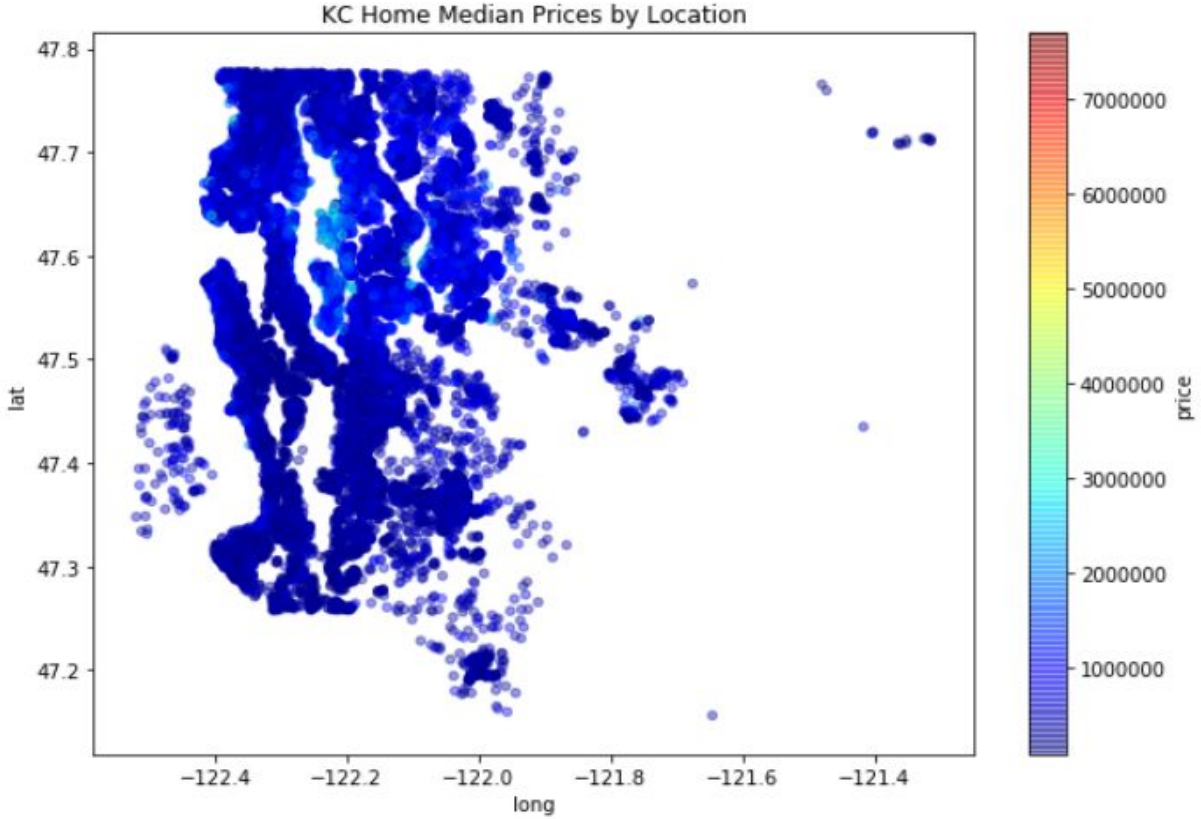
- The correlation matrix to the left shows how each predictor correlated with the target variable, price
- In other words, how well does each independent variable predict the outcome of home prices in KC
- From looking at the visualization alone, we can see that bedrooms, bathrooms, sqft_living, grade, sqft_above, latitude and sqft_living15 are among the best predictors of home prices

Leading Predictors: Square Footage of Living Area



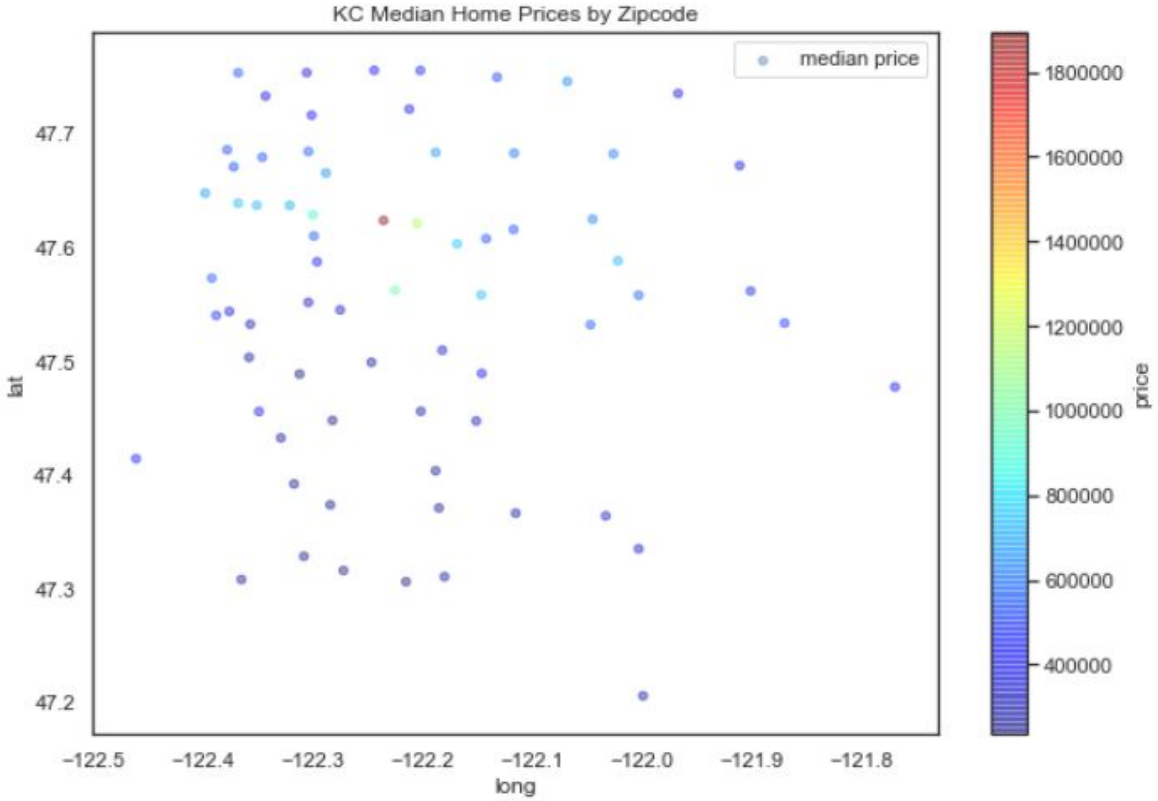
- The Square footage of the living space had one of the highest correlations to home prices
- As shown on the figure to the left, there is a linear relationship of home prices to square footage
- Most houses in Kings County (KC) fall within a $200-$600k price range
- A majority of homes in KC have a sqft_living of 500 to 3,000 sqft
- As shown in the next couple slides, predictability of pricing increases as we hone in on location

Leading Predictors: Location

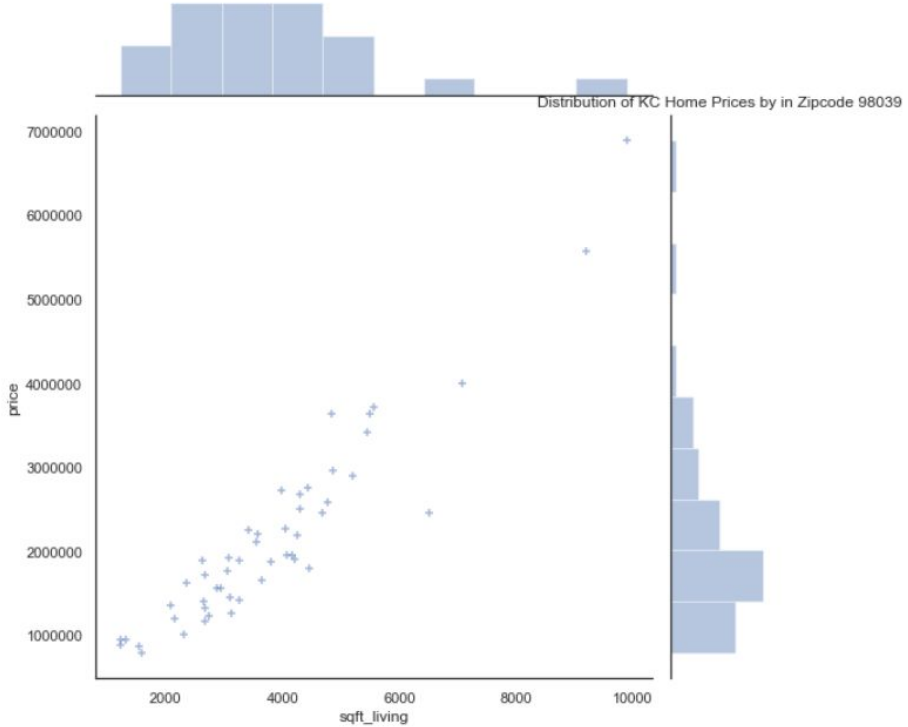


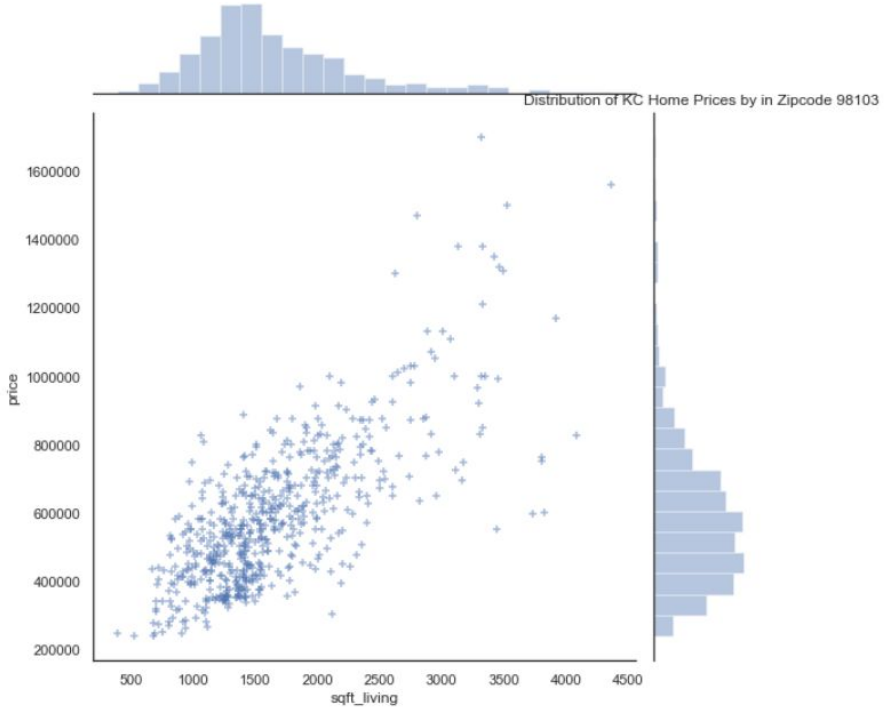
- The image to the left is a heated scatter plot of KC Home Prices by latitude and longitude
- Based on the colors, a cluster of homes in the Northwestern region (latitude range: 47.5 to 47.7, longitude range: -122.4 to -122) have the highest home prices
- The farther away homes are from this cluster, the lower the home prices
- From our analysis, prices were more sensitive to location based on North-South vs. East-West

Leading Predictors: Location (cont'd)



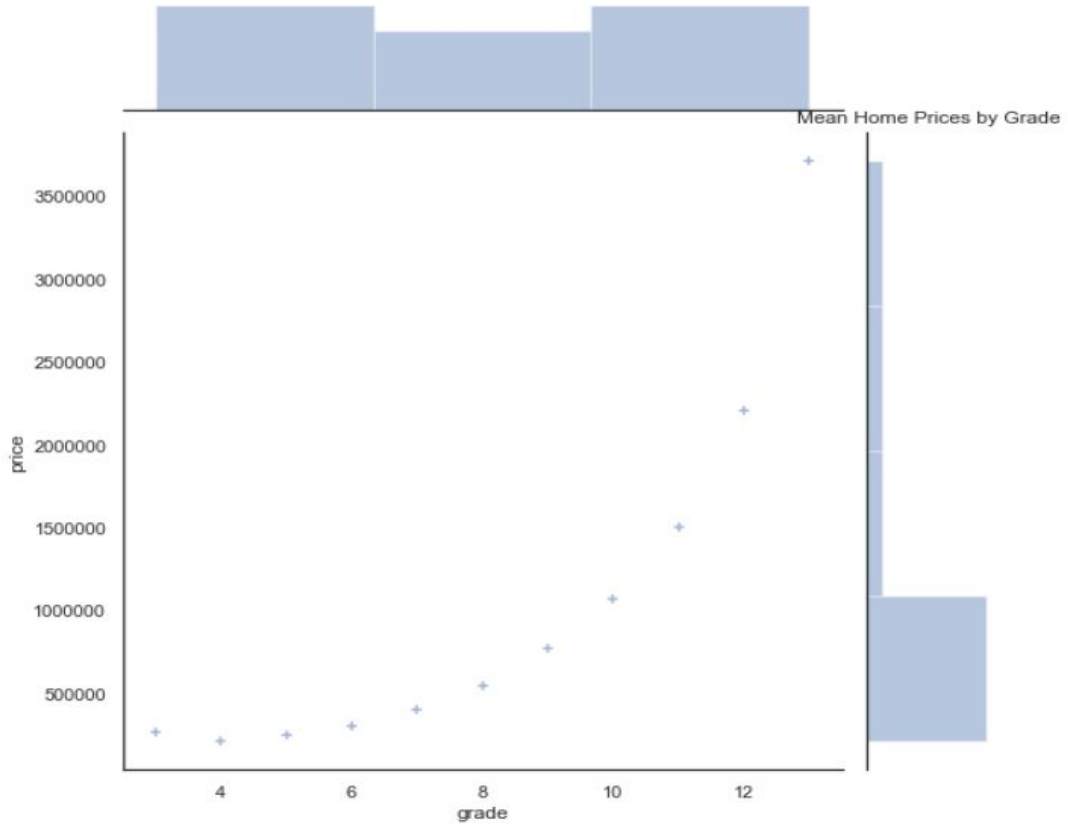
- To hone in on the importance of location on housing prices, we grouped the data by zip code
- The figure to the left shows a clearer picture of distribution of home prices in Kings County by location
- The red dot shows the most expensive zip code, 98039, in Kings County
- The further away you get from this neighborhood, the lower the median home prices





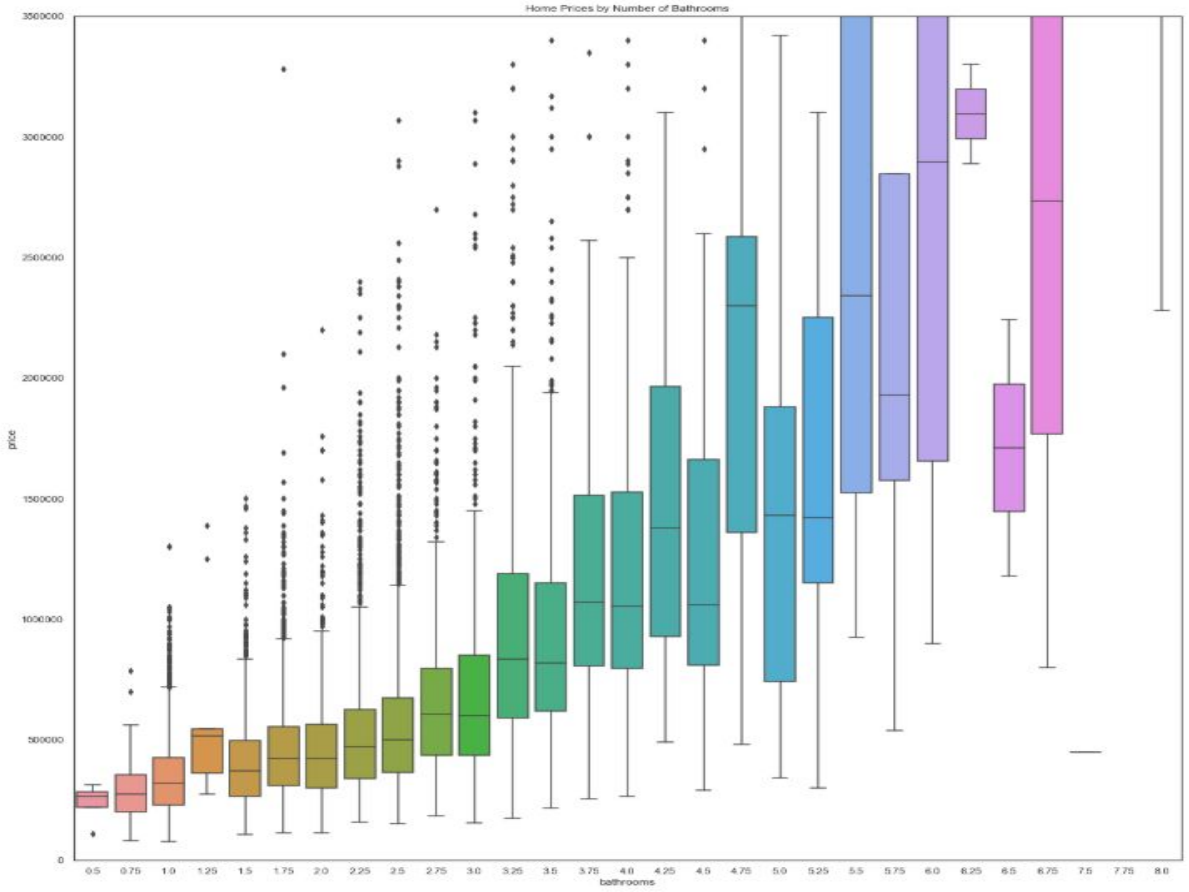
- Taking two sample zip codes (98039 and 98103), we can see that when split by location there is a clear linear relationship between prices and sqft of the living space as well as between prices and location
- Zip code 98039 (left) represents the most expensive zip code in KC, with square footage of 2,000+ and home prices ranging from just under $1mn to $7mn
- Zip code (right) represents the most densely populated zip code in KC, with square footage of homes clustered in the 500 to 3,000 range and prices clustered in the $200k to $1mn range

Leading Predictors: Grade (Index 1-13)



- The price distribution of KC Homes by grade has a clear exponential relationship. This suggests that home prices increase more rapidly for each incremental grade level.
- According to HomeAdvisor, "a typical residential excavation job runs between $1433 and $5057 with an average of $3069."
- For homeowners looking to maximize the selling price of their homes, improving their homes' grading can be a more cost-effective means of increasing sales price vs. expensive renovations to increase sqft

Leading Predictors: Bathrooms



- Bathrooms are also one of the top predictors of KC home prices
- Renovations, however, such as bathrooms can get expensive
- According to Than Merrill, Founder of FortuneBuilders.com, the following are typical costs of home renovations:
  - Low ($25,000 to $45,000): Interior and exterior painting, small repairs (like refinishing cabinets) and new landscaping.
  - Medium ($46,000 to $75,000): The low-cost upgrades above, plus a total kitchen renovation and minor bathroom upgrade.
  - High ($76,000 and up): Low- and medium-cost upgrades, plus fixing any foundation issues, roof and sewer line problems.*

*prices vary depending on size of home, location and size of the project.

Conclusion

OLS Regression Results

Dep. Variable:	log_price	**R-squared:**	0.747
Model:	OLS	**Adj. R-squared:**	0.747
Method:	Least Squares	**F-statistic:**	9097.
Date:	Fri, 26 Apr 2019	**Prob (F-statistic):**	0.00
Time:	03:13:53	**Log-Likelihood:**	30962.
No. Observations:	21596	**AIC:**	-6.191e+04
Df Residuals:	21588	**BIC:**	-6.184e+04
Df Model:	7		
Covariance Type:	nonrobust		

```
5-Fold Cross Validation MSE:  -0.003053667590188688
10-Fold Cross Validation MSE:  -0.003049395477199787
20-Fold Cross Validation MSE:  -0.0030345829122914872
```

```
grade          0.703747
sqft_living    0.695209
bathrooms      0.551270
log_lat        0.449120
bedrooms       0.350871
waterfront     0.170724
yr_built       0.080608
```

- The seven features included in the final model are: grade, sqft_living, bathrooms, log_lat, bedrooms, waterfront and yr_built
- Overall the adjusted R-Squared is 0.747, meaning that this model explains 74.7% of the variability in our response data, price
- Grade was most positively correlated with housing prices, meaning that a 1 unit increase in grade leads to a change in price by 70%
- Similarly for sqft_living, a 1 unit increase in sqft_living would result in a 70% change in price
- For each bathroom added, price will increased 55%
- Looking at the log_lat variable, we can interpret this as a 1% change in log_lat will lead to a 0.45% change in price.