Real Estate Analysis

Business Problem

How should a real estate firm price a house based on property features such as lot, footage, seasonality, and renovations?

Analysis: What did we examine?

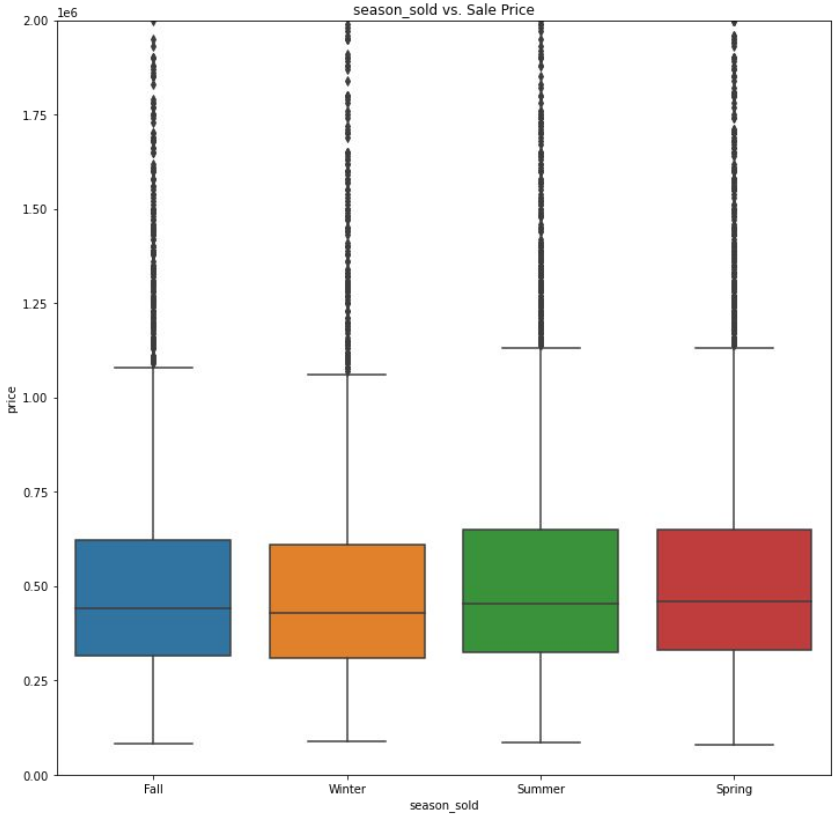
Target - Price is prediction target

Features - Various Property features and zipcode in the King County housing market of Seattle WA

Conclusion: What we found

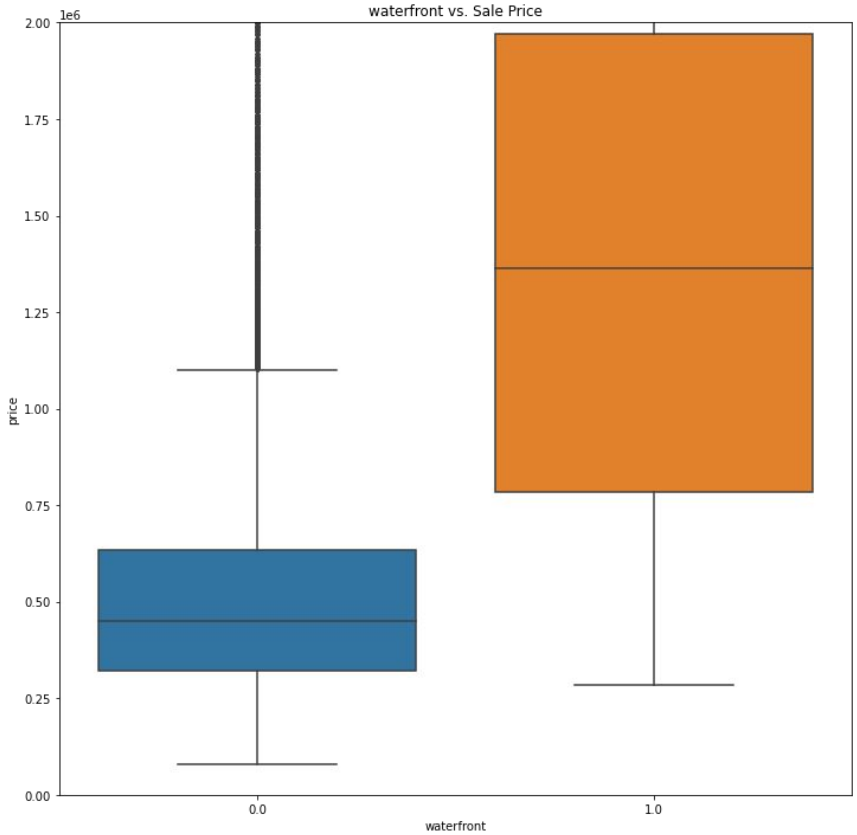
- Various business insights by feature
- A model which can price mid priced houses

House Pricing and Seasonality



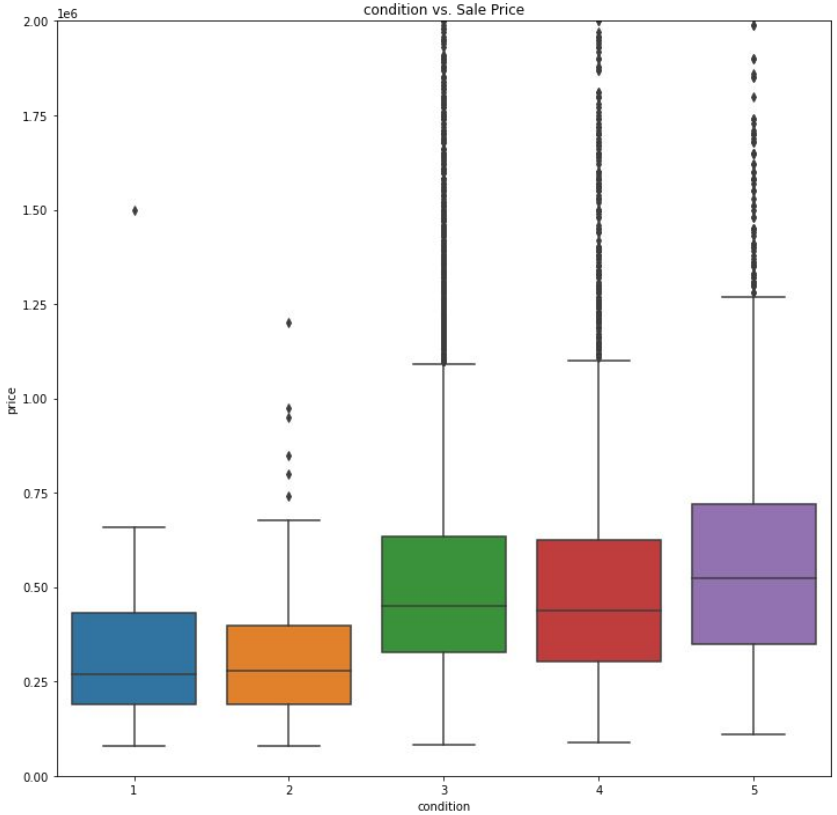
More houses purchased during Spring and Summer tend to sell for higher prices.

House Pricing and Waterfront Properties



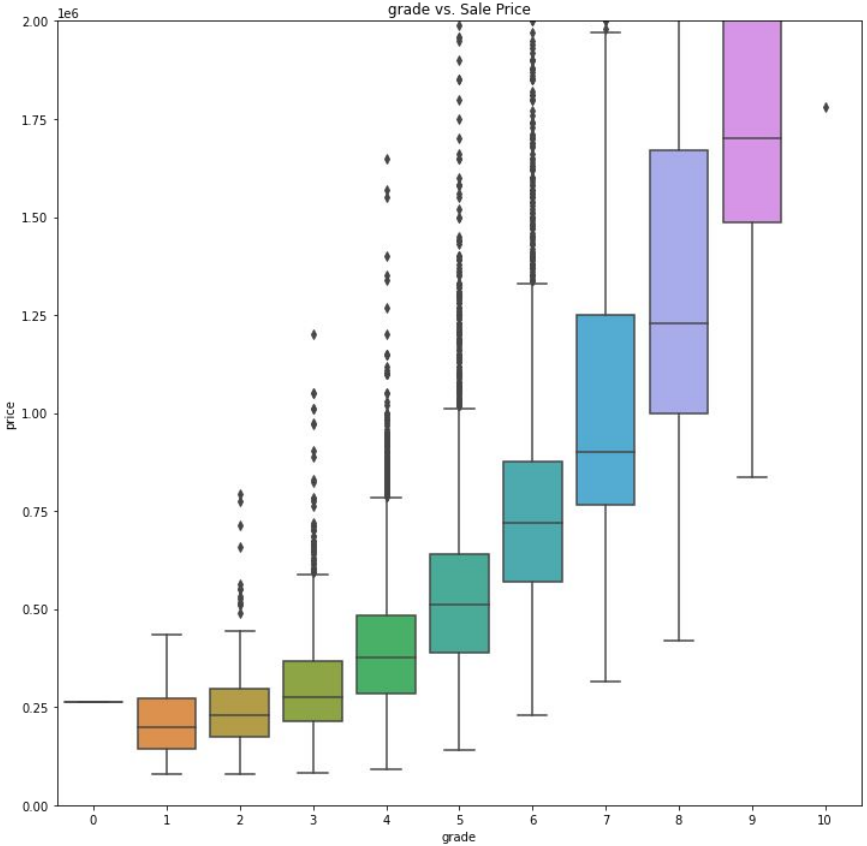
Here we see confirmed what we would intuitively assume; waterfront properties result in an increase in price.

House Pricing and Condition



For condition of the house a higher condition indicates a house which is well maintained and overall in good condition. Looking at house condition vs price we can see that houses above a grade 2 tend to be in a similar pricing group and houses grades 2 and under tend to be in a lower pricing group.

House Pricing and King County Grading



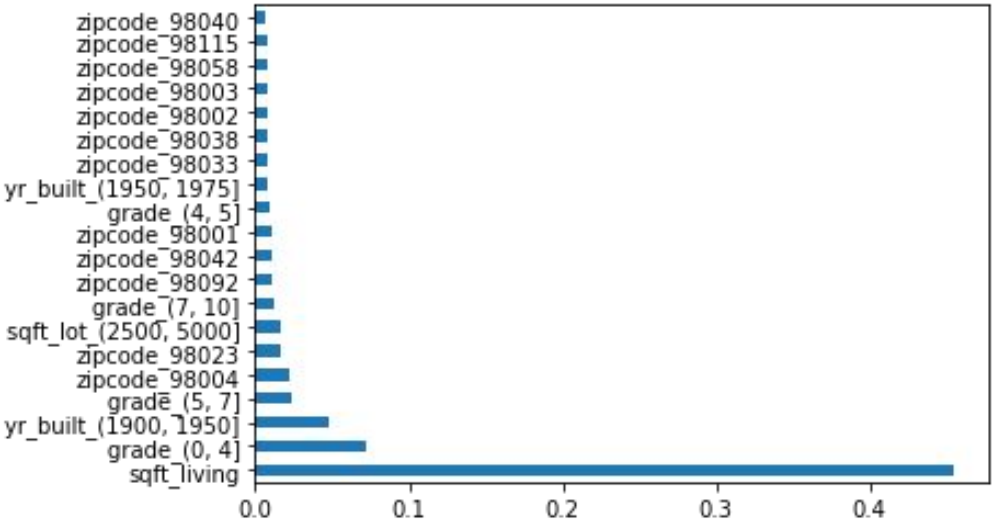
King County has a process for grading houses. A higher grade indicating a better house a lower grade indicating a less desirable house. Looking at house grades vs price we can see that houses graded 4 and below tend to be more tightly clustered and in a more similar pricing group. As the grade increases we tend to see a greater spread in pricing groups.

House Pricing and King County Grading

- When using the grade in the model we found that the King County process for giving house grades is only a good price indicator for houses with lower grades. When modeling we found that grades 0-4 were a good indicator of a decrease in value but that grades above 4 were not a good price indicator. This likely due to the greater variability in prices at higher grades and could indicate that houses with grade 5+ are generally acceptable to buyers while there is a more dramatic drop in price for less desirable houses.

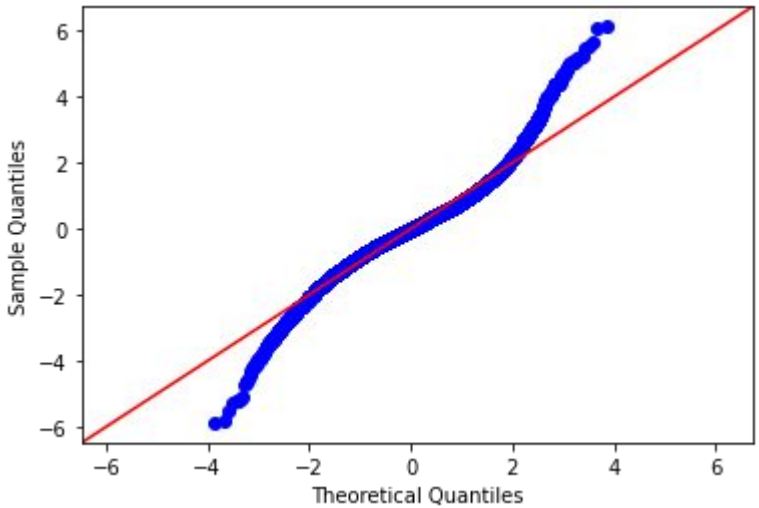
House Pricing and Square Footage

The 20 most important features for price modeling:



Square Footage is by far the strongest indicator of how a house should be priced. This field also has low error unlike some of the other fields listed in the top 20 for feature importance.

OLS Model



This plot shows how close our predictions were to actual prices in our data set. As we can see our predictions are close in the mid range house prices but it does not perform as well on the extremes aka the low end and high end of the housing market.

Recommendations

- Sell houses in the Spring and Summer when the market is strong.
- When pricing a house the overall square footage should be the most important factor.
- Properties with a grade of 4 and under will have a lower price to comparable properties.
- Our pricing model is most effective at estimating mid range housing prices. It is not advisable to use this model for the low end or high end of the King County housing market.

Further Analysis

- Expand data set beyond houses sold in 2014 and 2015. Having more current data would make the model more accurate.
- Adjust modeling to better estimate low and high priced houses.
  - Investigate what features would help us to estimate extremes