

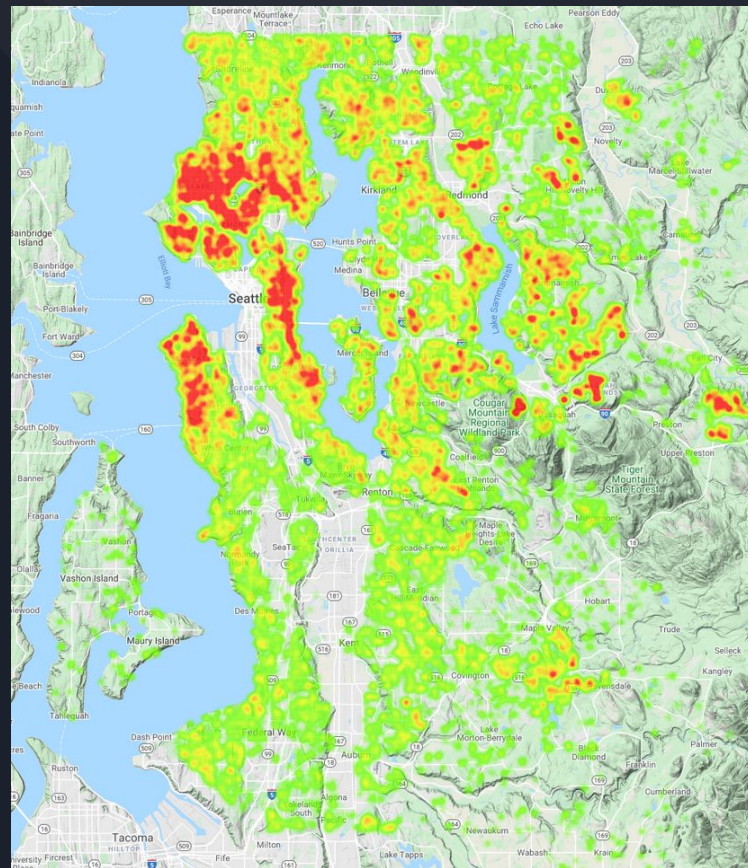


King County Real Estate Presentation

Flatiron School Module 2 Data Science Project

By: William Newton

Data & Methodology



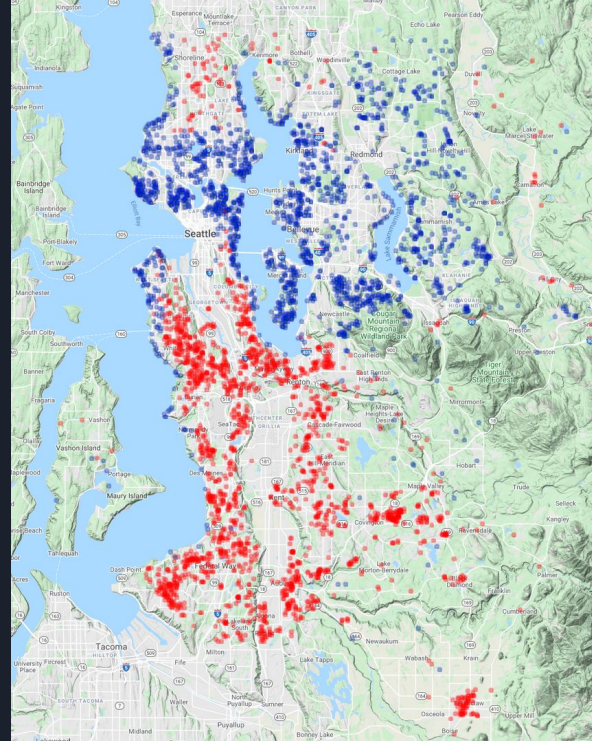


Data & Methodology

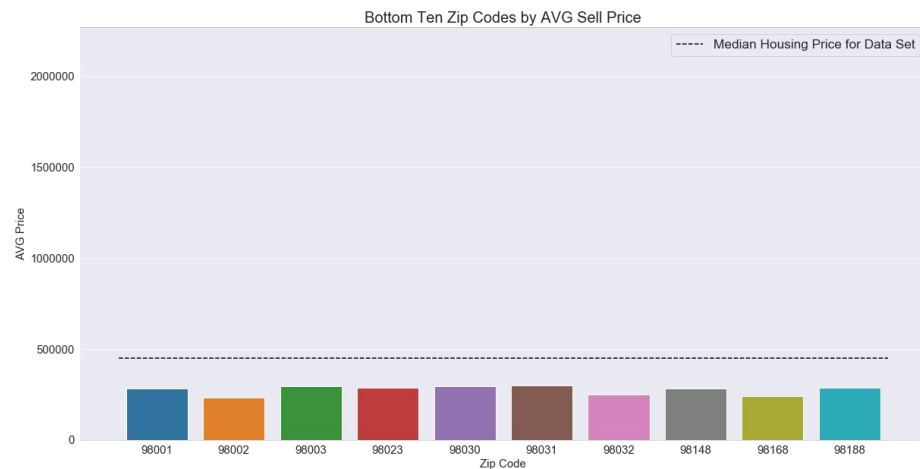
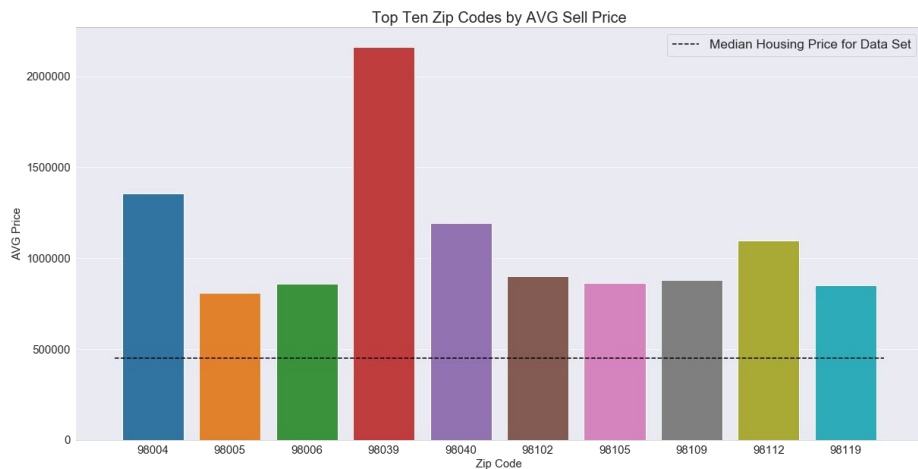
- Data used for this project is from the King County Real Estate Dataset available on Kaggle
- Data covers home sales from May 2nd, 2014 to May 24th, 2015 in the metro-Seattle area
- Focus was on building a multiple linear regression model using home data under \$1 million to predict housing prices for similar homes
- Used most effective predictors from the model for recommendations to home sellers
- Also focused on providing advice to prospective home sellers on what to do to increase their home's value and maximize selling price

Location, Location, Location

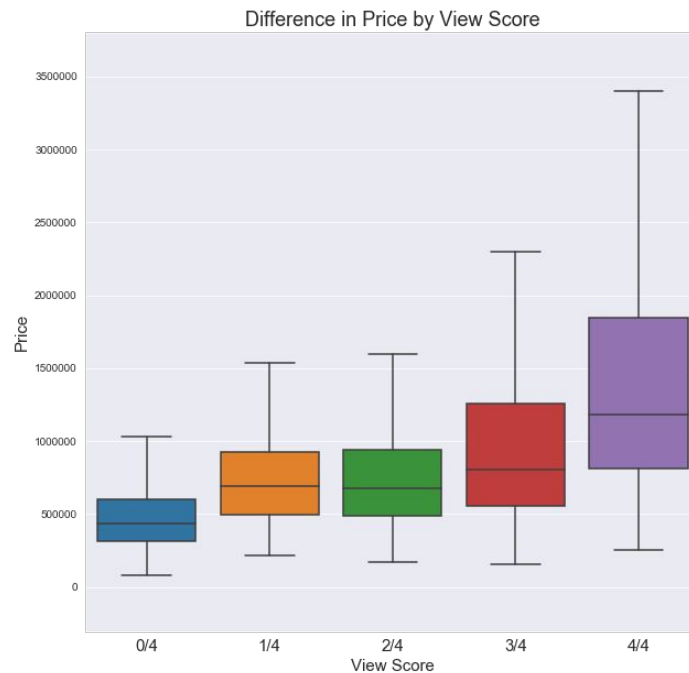
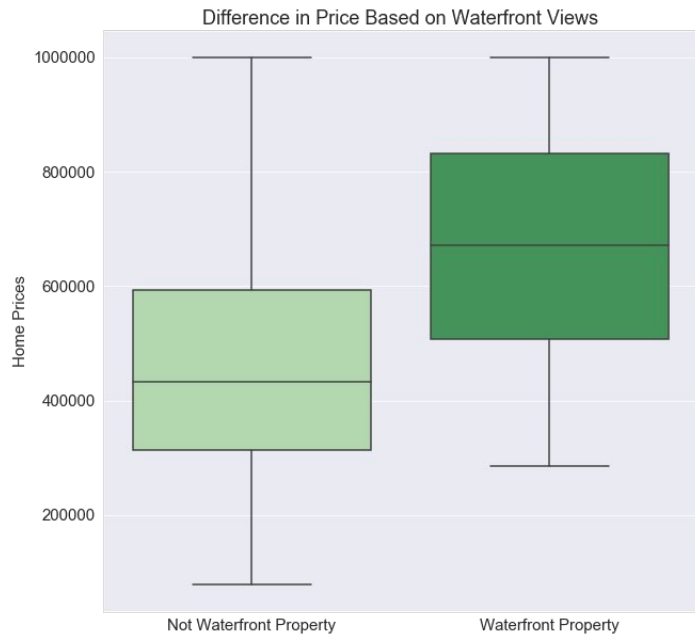
- Zip Code and other location data was highly predictive for my model and should be used to set realistic expectations for home values.
- Top 10% Sell Price in Blue
- Bottom 10% Sell Price in Red



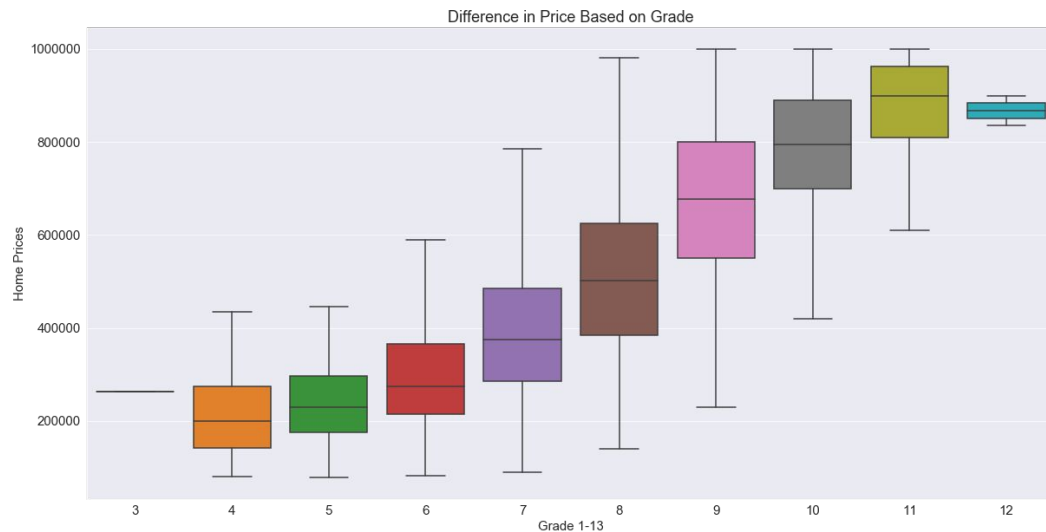
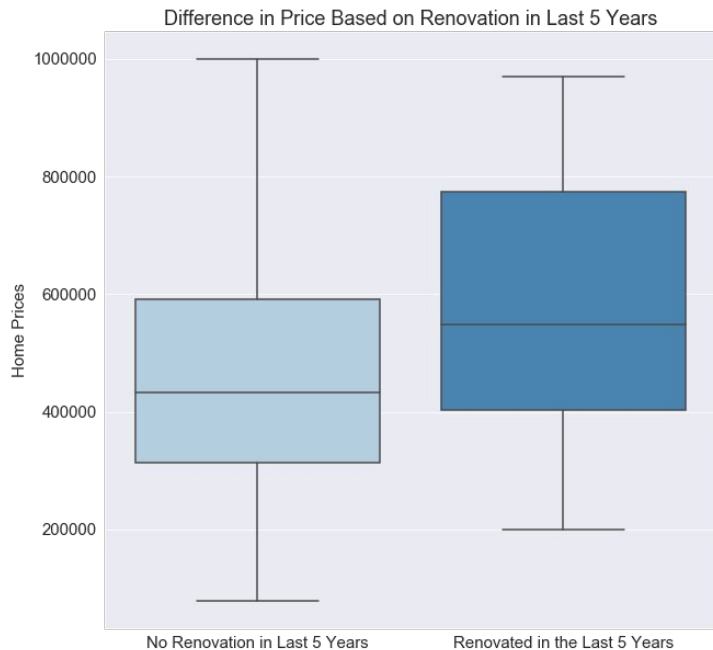
Location, Location, Location



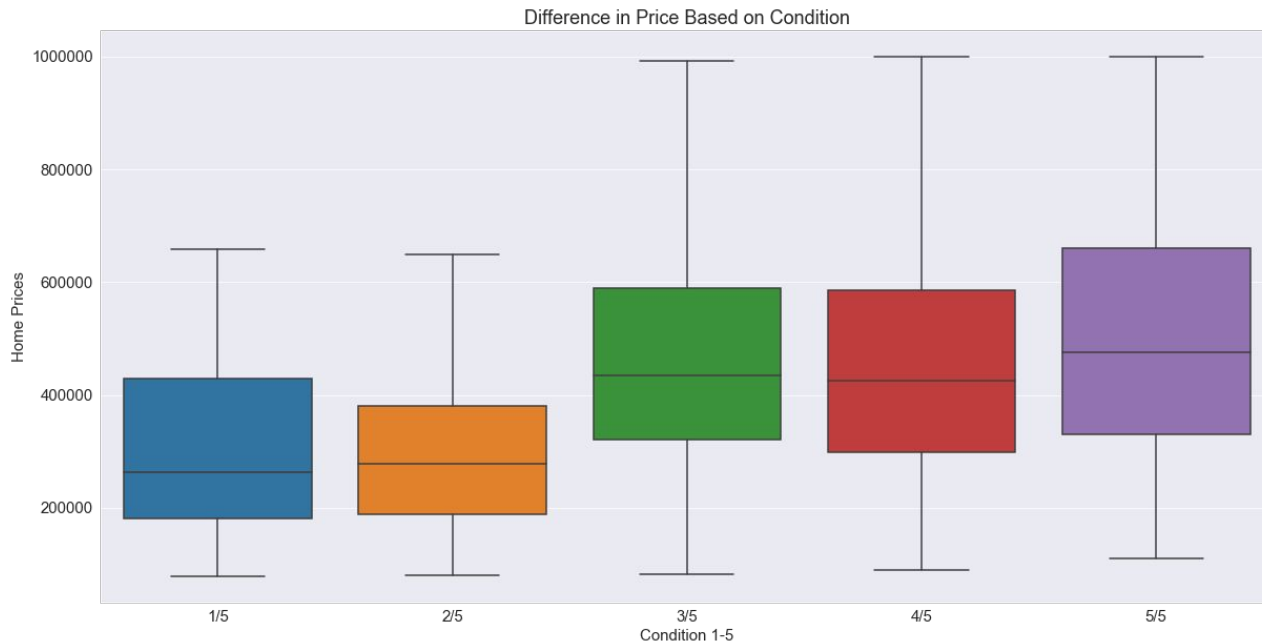
What a View!



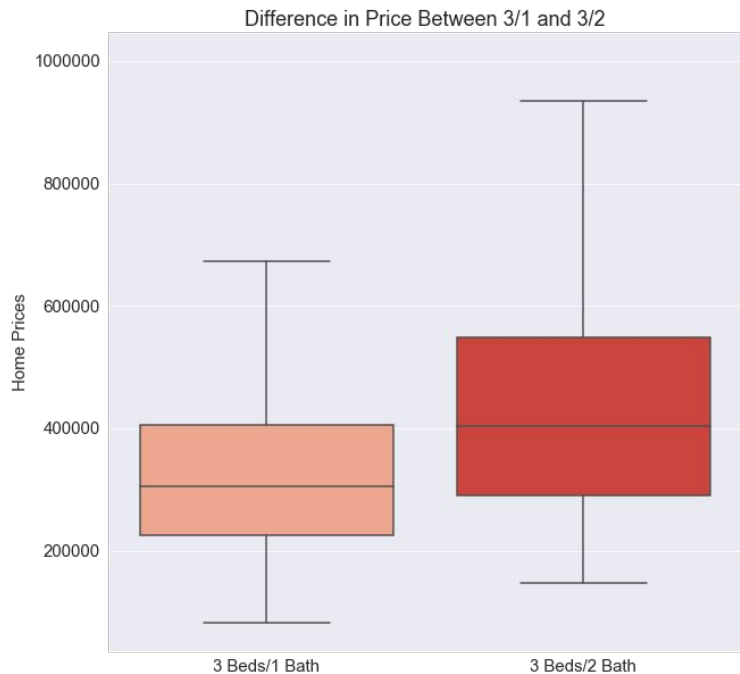
What is within your control?



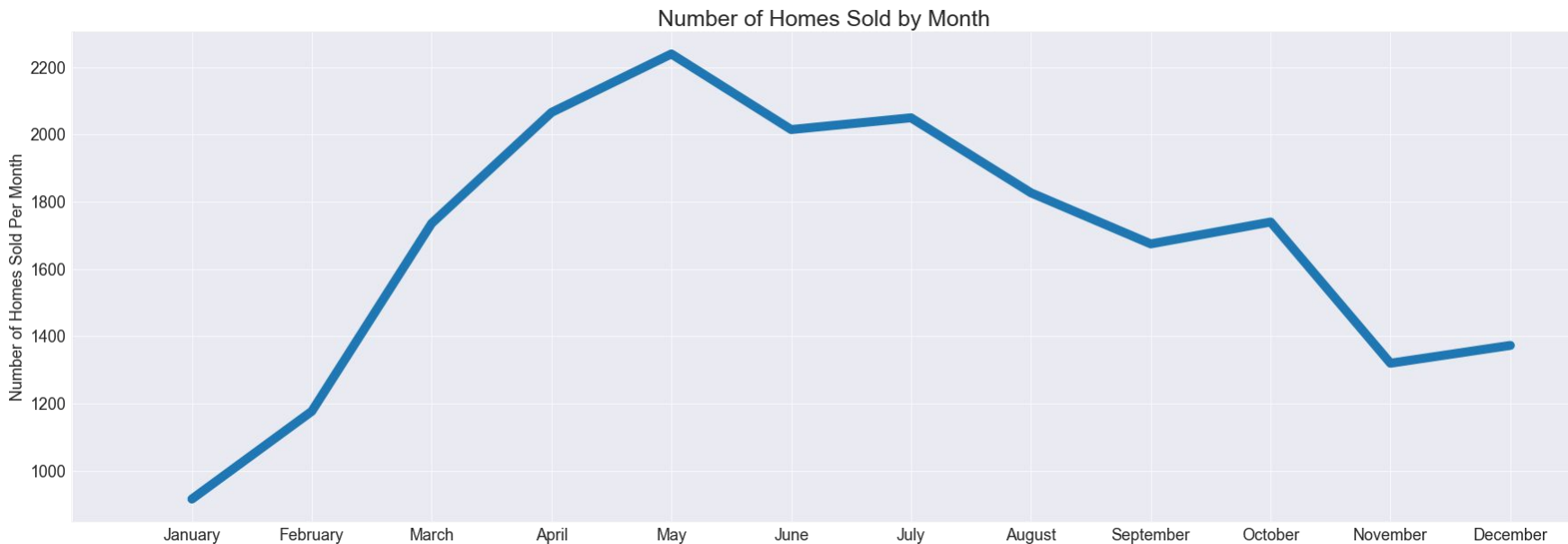
What is within your control?



Additions to your Home Can Increase Its' Value



Choose the Right Time to Put Your Home on the Market






Future Work

Given more time on this project and the data set, I would have loved to explore some additional features...

- Include demographic data to add further dimension to geographic information
- Get more data over a larger time frame to increase model fit. I have no way of knowing if this 2014 - 2015 data is an outlier if I looked at other year's data
- Further explore additional mapping libraries and functions. The gmaps library was easy to use but limited in its' functionality



Thanks and I look
forward to working
together soon!