Where should you invest your money?

Javier Herbas, 2020

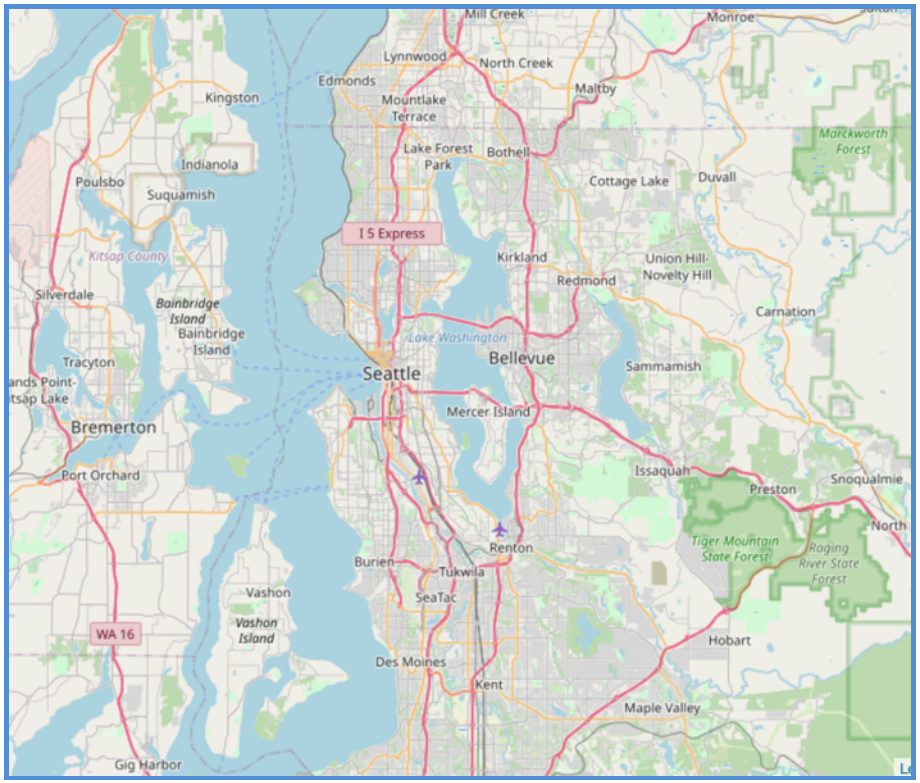
Project Details

Stakeholders: Investors trying to build a real state based portfolio in King County

Threshold: 1 million US$ / property

Property specifications: None. Should be provided by us.

Database: King County dataset (years: 2014-2015)



Objectives

1. Find where would the best area to invest in King County
2. Compare house prices based on different characteristics
3. Create a Price Prediction Model
4. Analyze Model

Key Features

- Sectors

4 different sectors were created based on lat and long coord

- Square Footage

Analysis of the square footage price per sector

- Waterfront

Waterfront Avg. Price vs Inland Avg. Price

- Number of Bedrooms

Analysis of average price variation by number of bedroom per sector

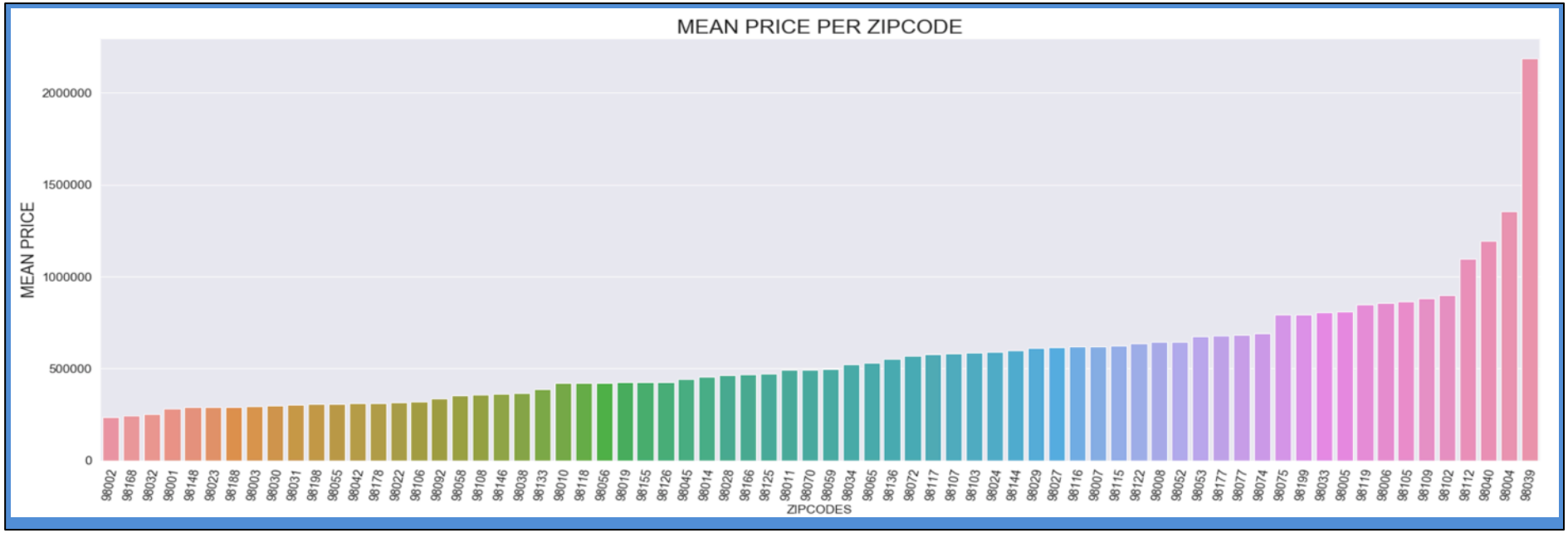
- Bathrooms

Analysis of average price variation by number of bathrooms per sector

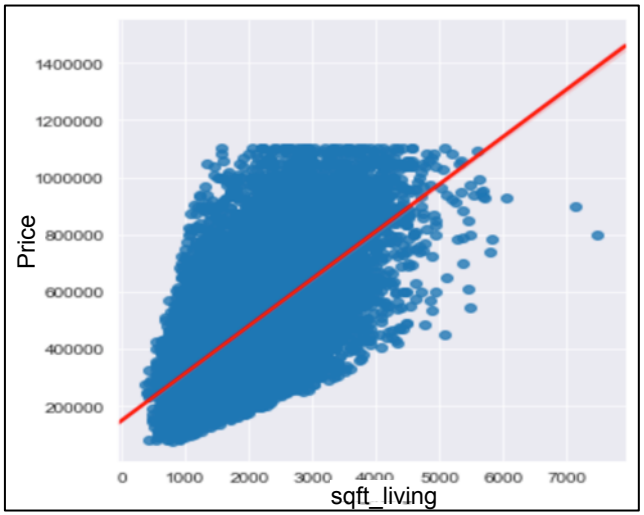
Sector Creation

- Dealing with 70 zip codes might be complicated, but ultimately it might result in a more accurate recommendation since the average price is clearly control by them
- Sectors managed to separate price ranges up to a certain degree and can be considered a good starting point

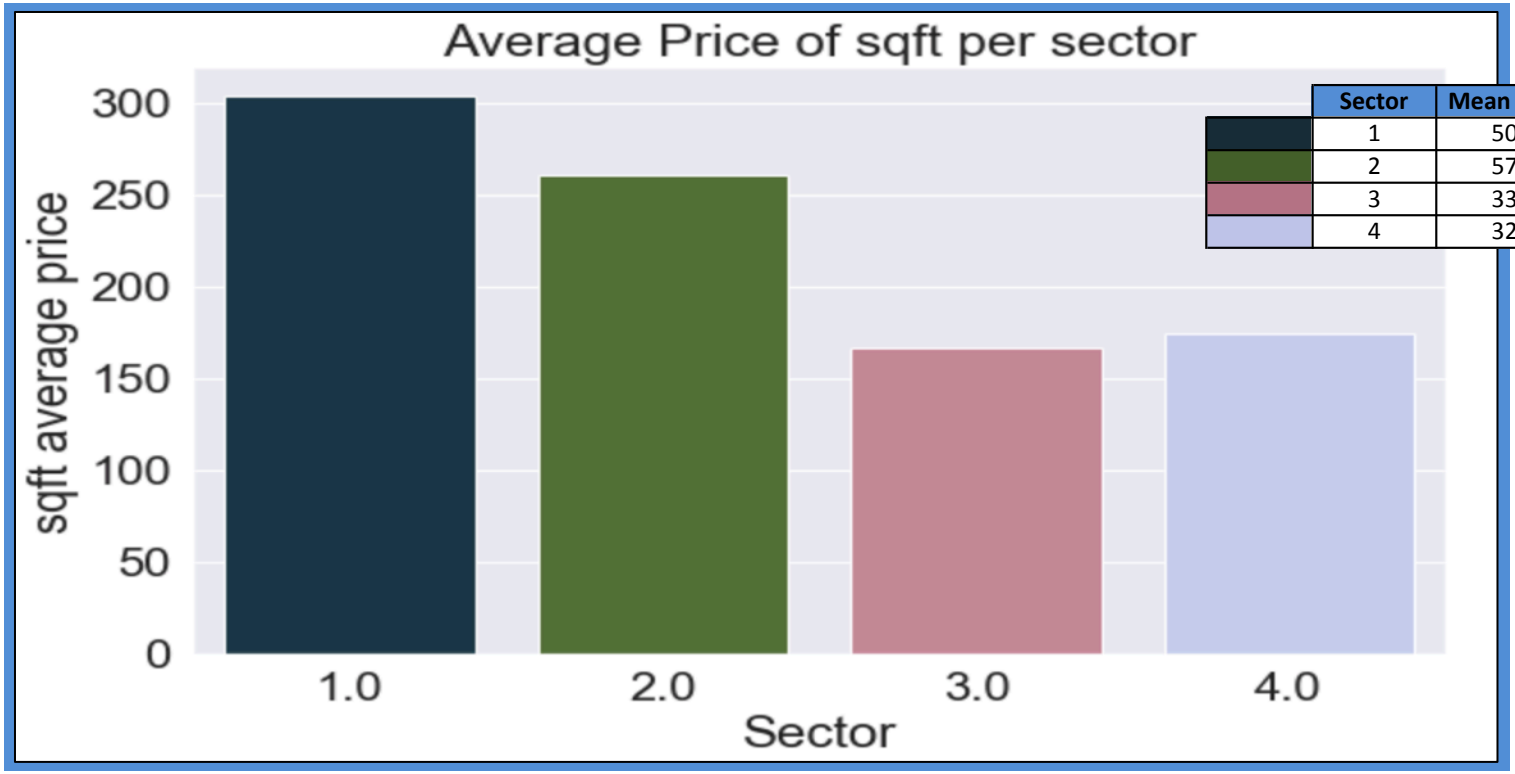
Sector definition based on lat & long and house price



Which sector should you be looking at? and, what price per square foot will you get?



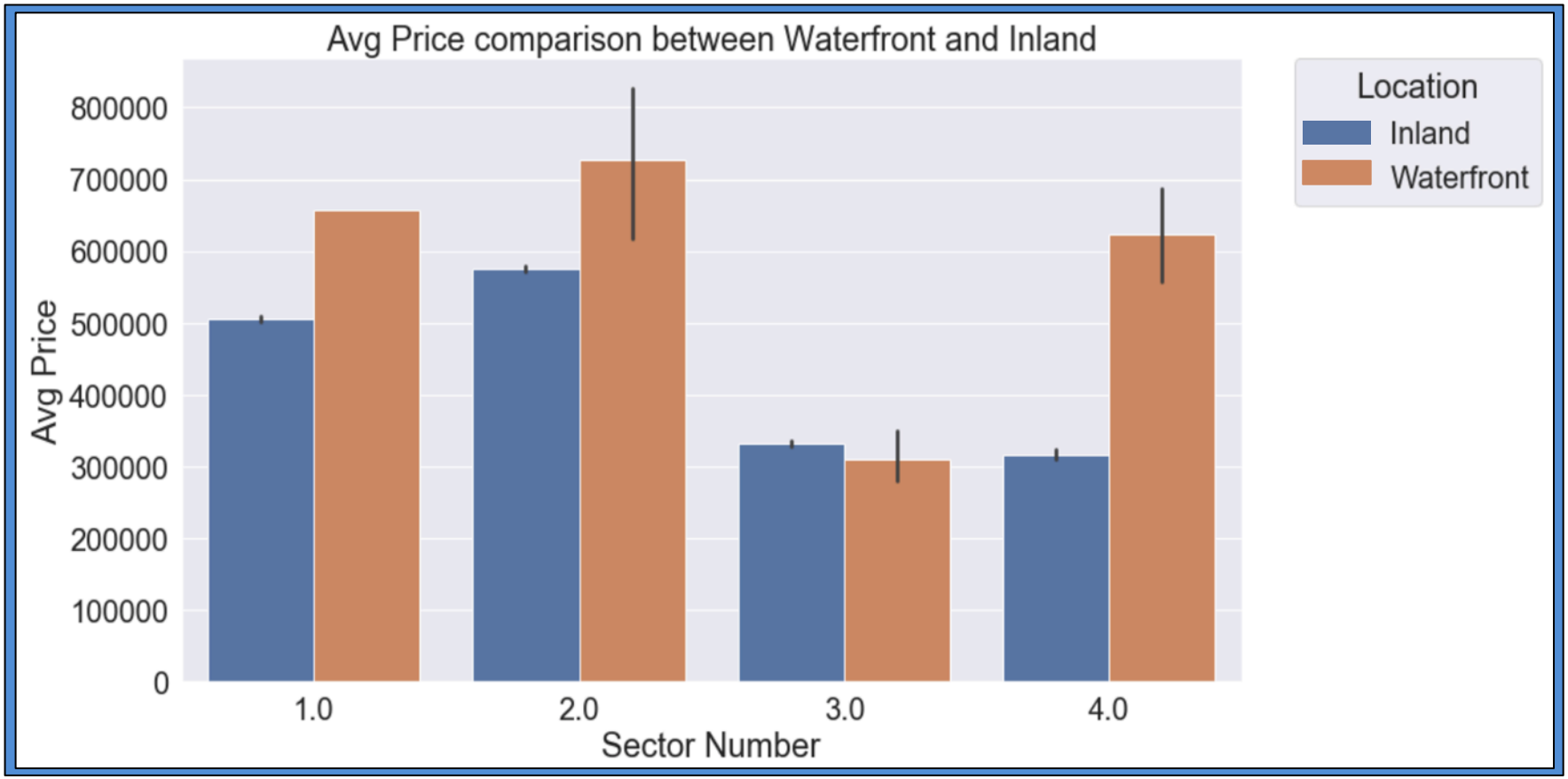
- Sector 1 contains the highest price per square foot
- Sector 2 covers the majority of the houses located in the waterfront
- Bigger houses within Sector 2 didn't passed the threshold/filter value of 1 million $



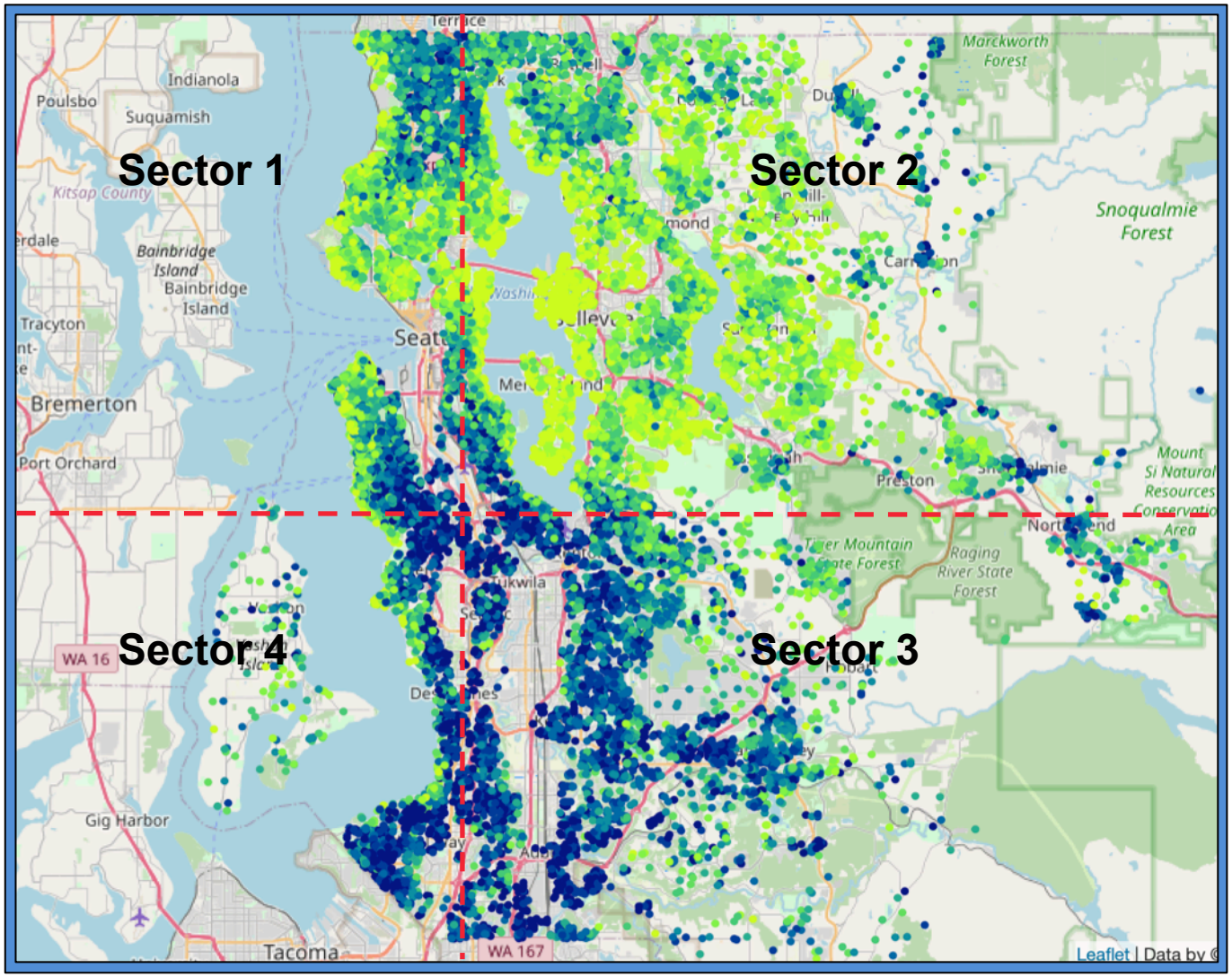
Sector	Mean Price/House	Mean Sqft Living
1	505438.50	1661.59
2	575834.19	2208.68
3	331316.47	1981.53
4	324299.67	1861.76

Should you be looking at the waterfront or in-land view?

- Buying in the waterfront is always more expensive with the exception of Sector 3, which has the lowest average price
- With the exception of Sector 3, the price uplift from in-land to waterfront property is rather significant



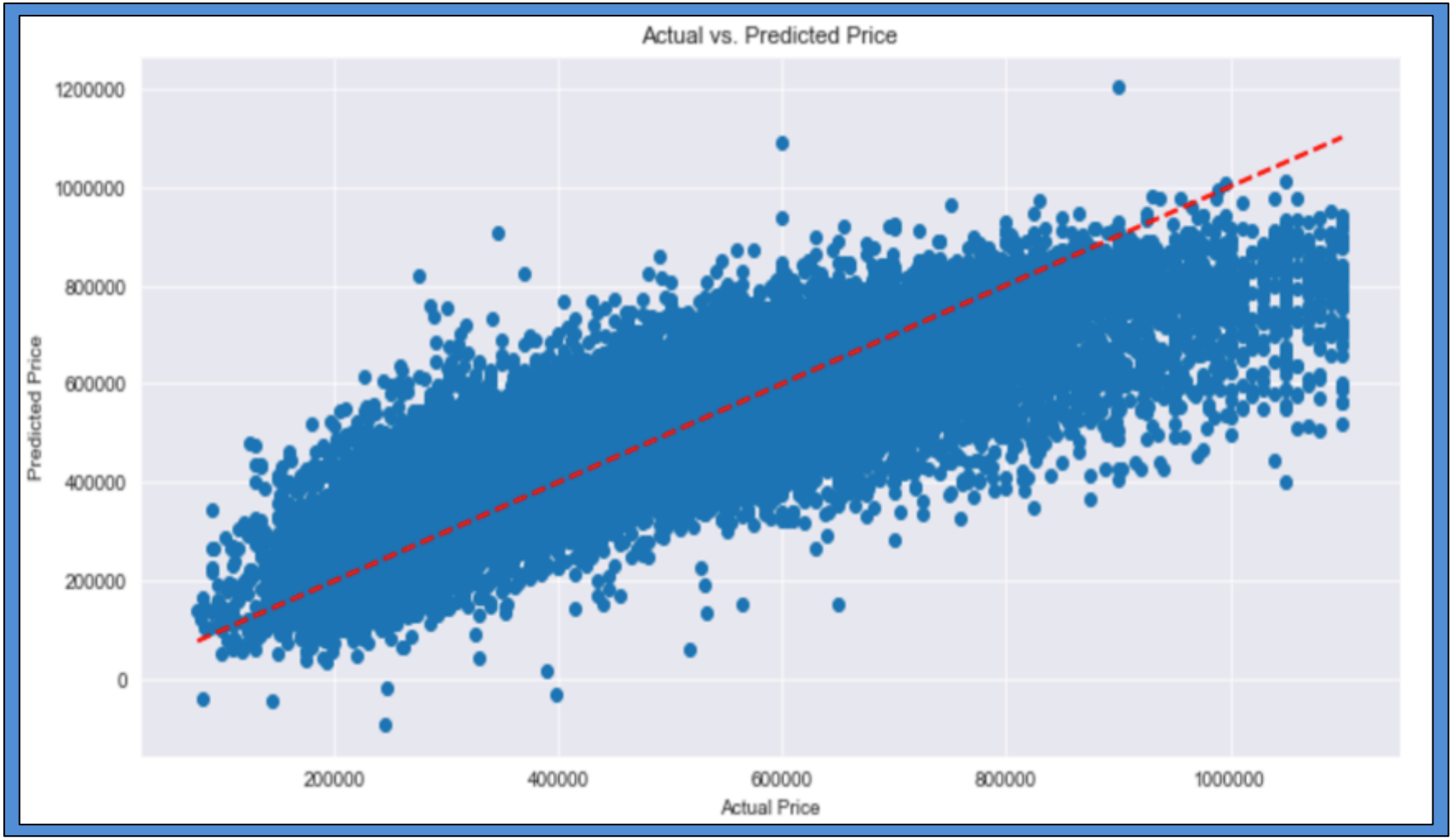
Should you be looking at the waterfront or in-land?



Color variation correlation with the water is quite clear

As expected, being in the waterfront will demand a higher price!

Actual vs. Predicted Price



Model Result Details

Input Data: King County dataset from 2014 - 2015

Metric: Mean Absolute Error (mse) & R^2
R2 : 0.693
Predictive Error (mse): 111,359% (11.3% on price range)
Price Range: 1,000,000$ (~95% of the data)
Average Price: 468,468.61$

SUMMARY:

- According to my model the sector (location), square foot of living, and the waterfront view are the variables that will contribute the most to predicting price.

- Other good predictors worth looking at are the renovations and the grade given to the property (Grade_C)

Recommendations

- Look for the best square footage within sector 1 and sector 2.
- Preferably look for houses with waterfront view
- Within the given budget, it is advisable to stay between 600 to 700,000$ to avoid entering the zone where the model under-predicts the prices by over 100,000$
- If you want to minimize the error stay below 600,000$

Way Forward

- Extending the study into individual zip-codes could not only refine the model but allow more accurate recommendations on specific locations of where to invest

- Get mode data :

  - The one available only covers 2014-2015 and as an investor it would be interesting to analyze price variations with time
  - School and public transportation distance to properties would be interesting to add to the model as well

Thank you

Javier Herbas, 2020