Kings County Housing Analysis

Presented By:
Hoang
Madoria
Marcelo

Outline

- **Overview**
- **Business Problem**
- **Data & Methods**
- **Models**
- **Conclusion**

Overview

Zillow is looking for a model that can help predict house prices. They asked us to come up with a plan to buy a large amount of houses.

Business Problem

Analysis of the King's County house features show what characteristics are important to determine the fair price of houses.

- Predict the fair value of houses.

- Which features are more important in determining house prices?

- Which city has less expensive houses: Seattle, Redmond or Kent?

Data Dive and Methods

- **The Kings County Housing dataset:**

  - Over 21 thousand houses
  - 21 different features

- Regression models with key predictive features were trained and tested as tools capable of predicting prices of houses in Kings County.
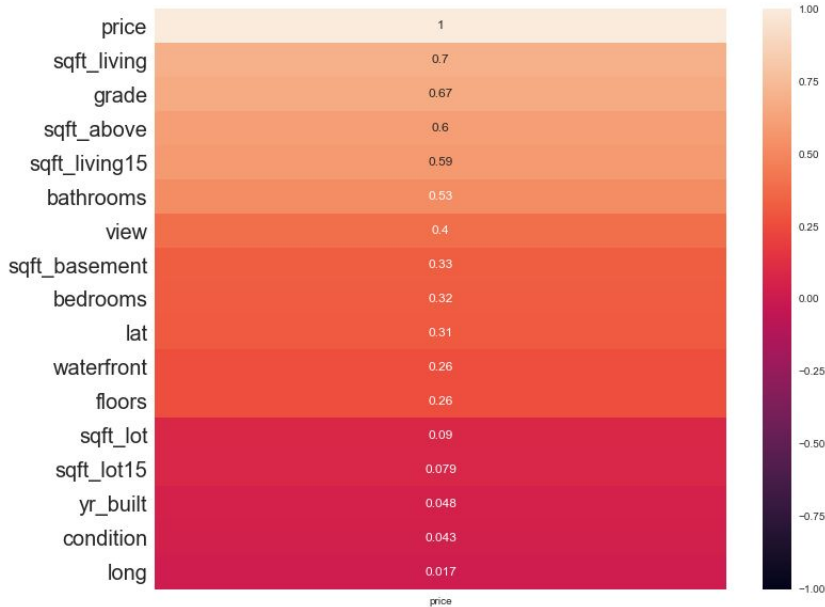
Models and Regression Results

Models

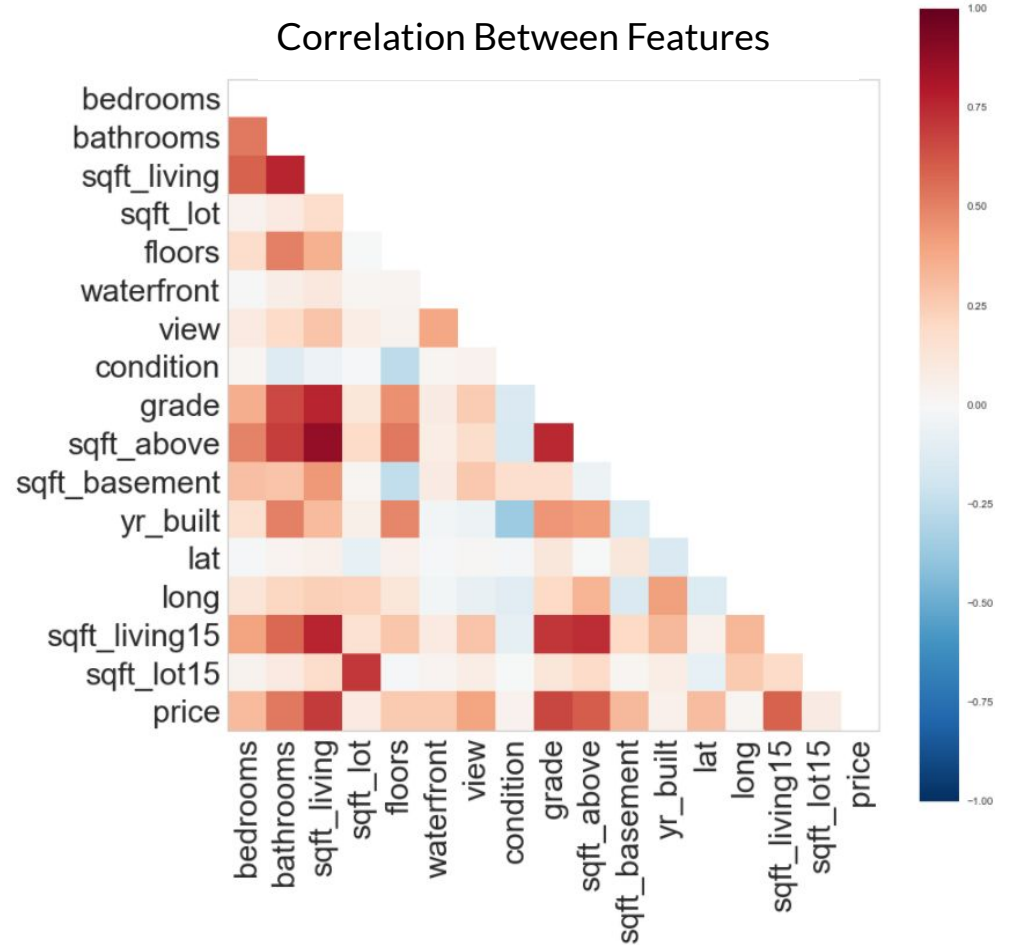- Several models were created in the analysis process

Percentage of the Variation of our Data from the Mean
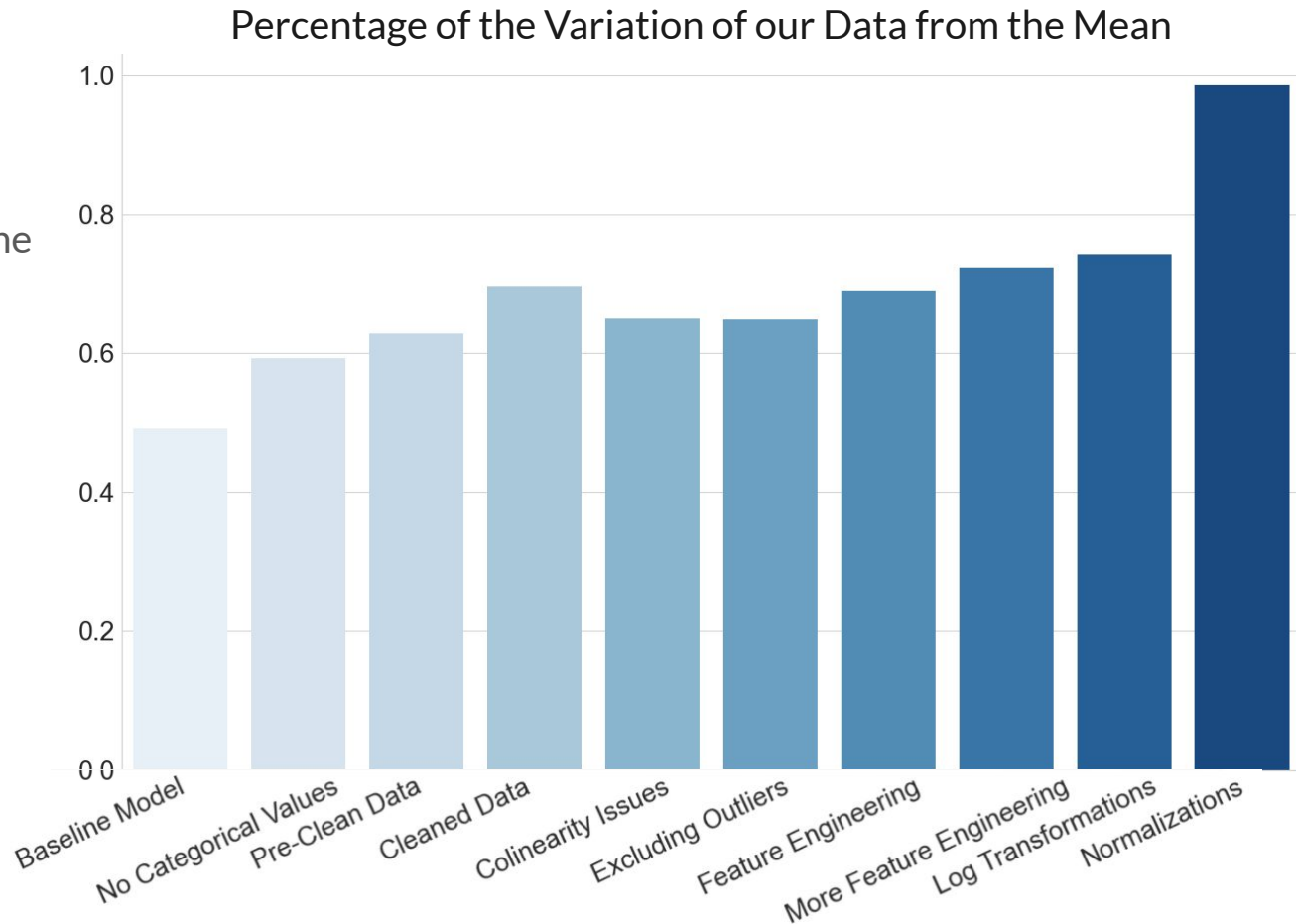
- Multiple highly correlated features are not good for regression models

Correlation of Features with Price

Correlation Between Features

- More appropriate correlation matrix
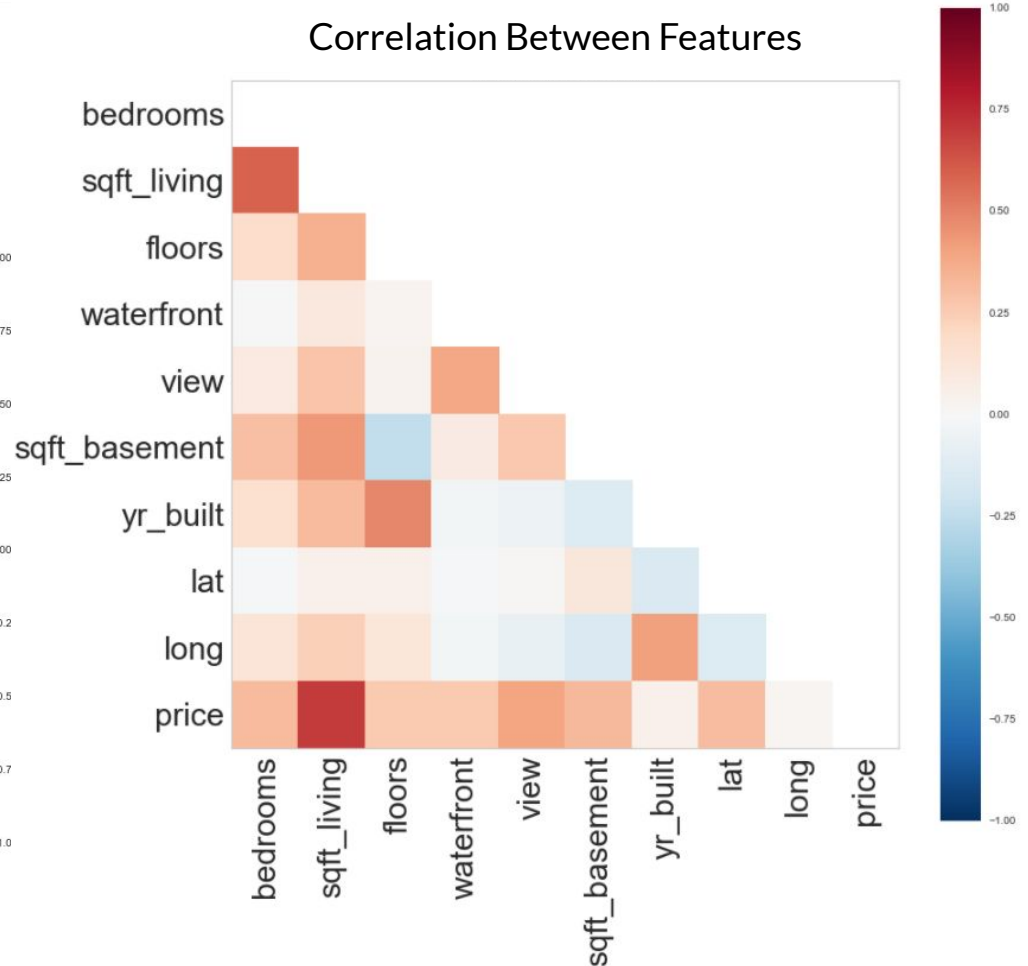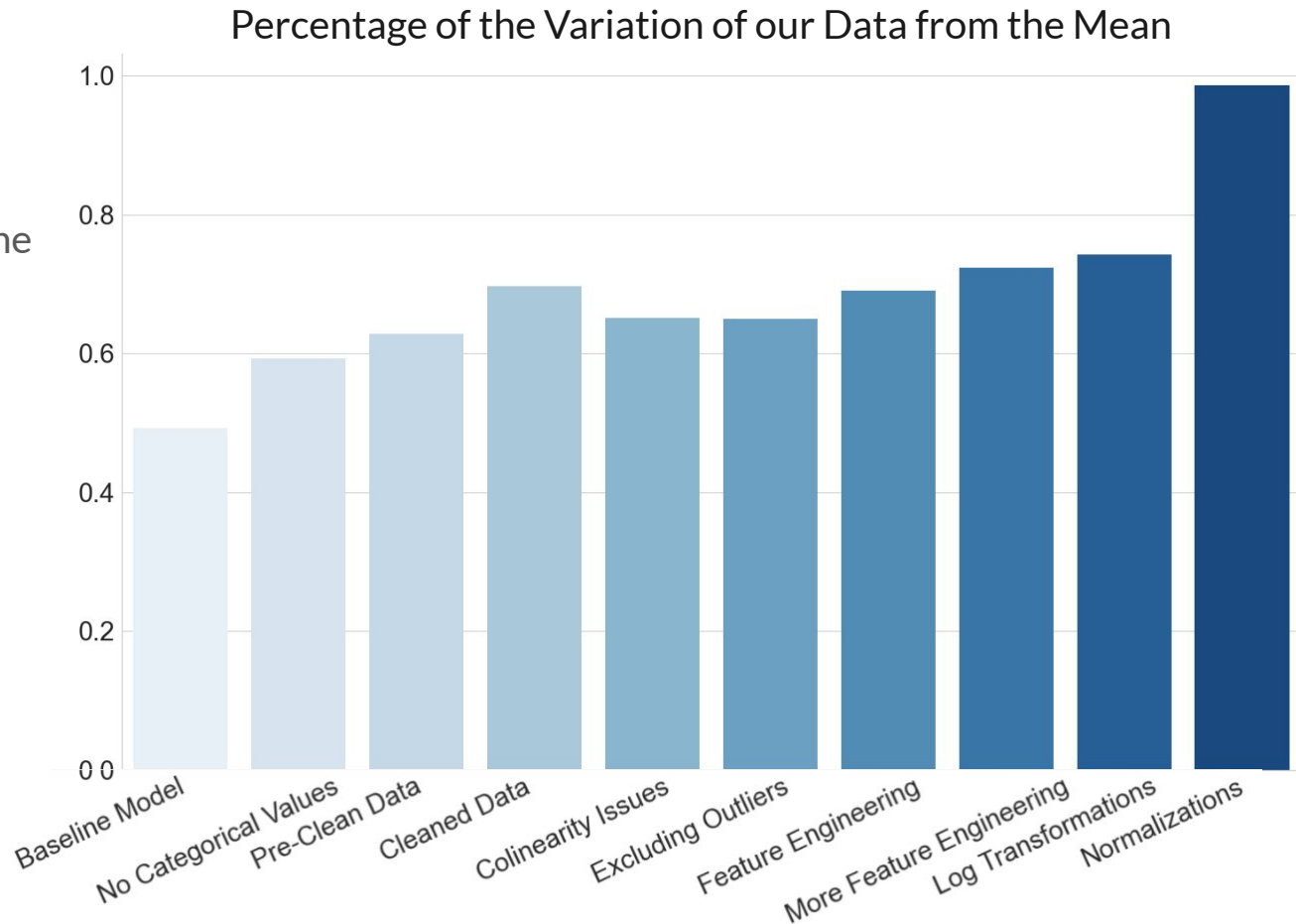
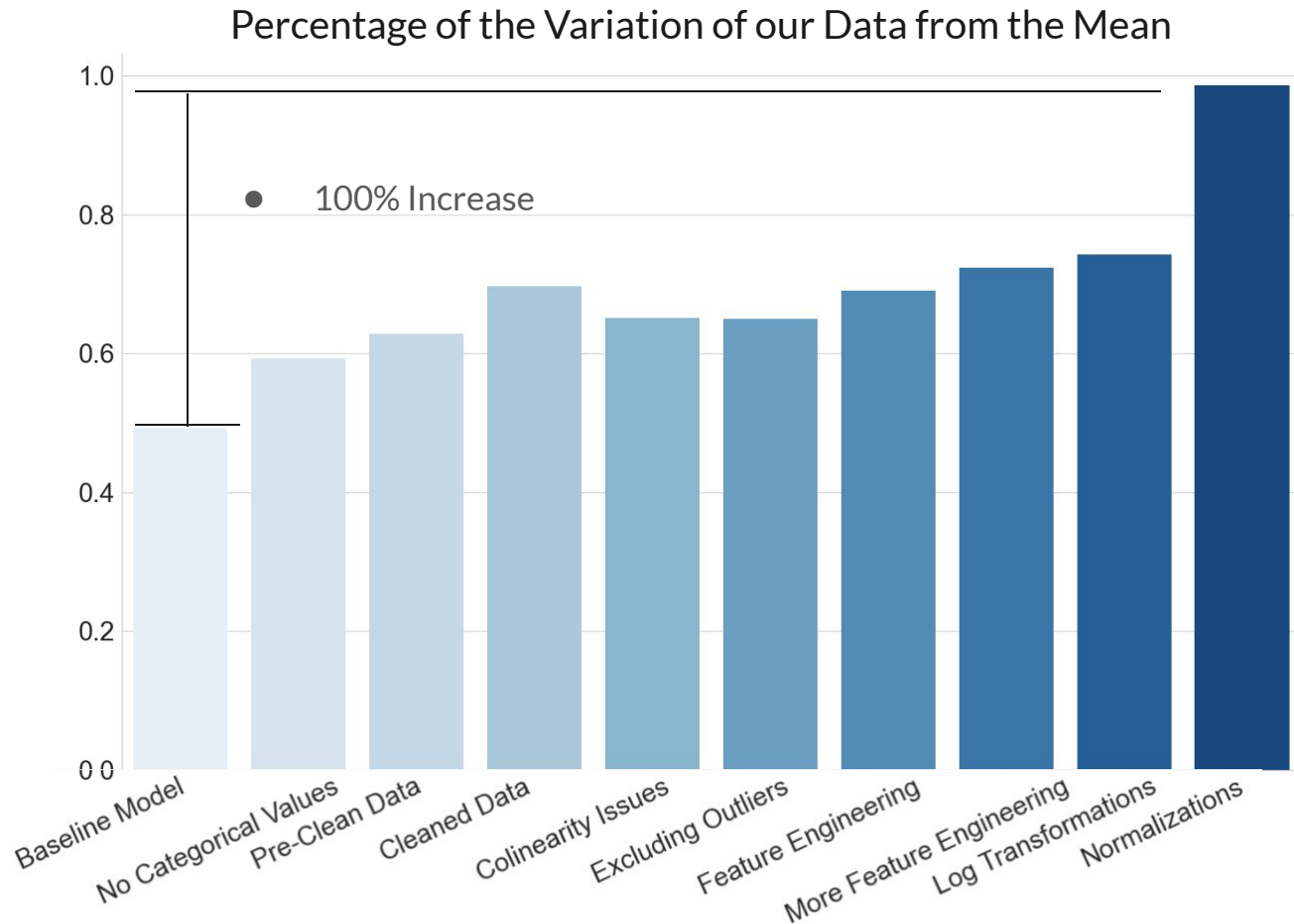## Correlation of Features with Price
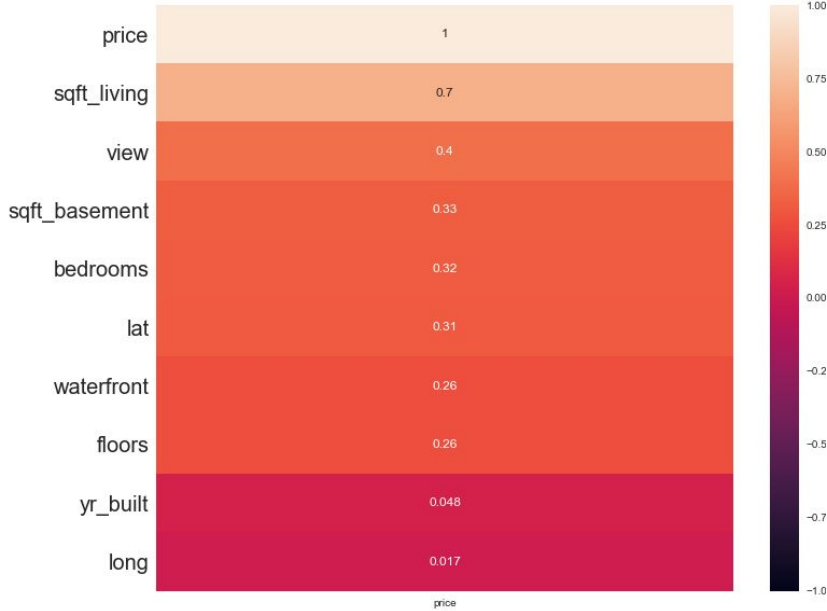
Correlation Between Features

Models

● Several models were created in the analysis process

Percentage of the Variation of our Data from the Mean

Models

Percentage of the Variation of our Data from the Mean

● 100% Increase

Baseline Model, No Categorical Values, Pre-Clean Data, Cleaned Data, Colinearity Issues, Excluding Outliers, Feature Engineering, More Feature Engineering, Log Transformations, Normalizations

Models

- How much money a better model can save

Model Error

- 30% Decrease

Y-axis: Thousands of Dollars ($)

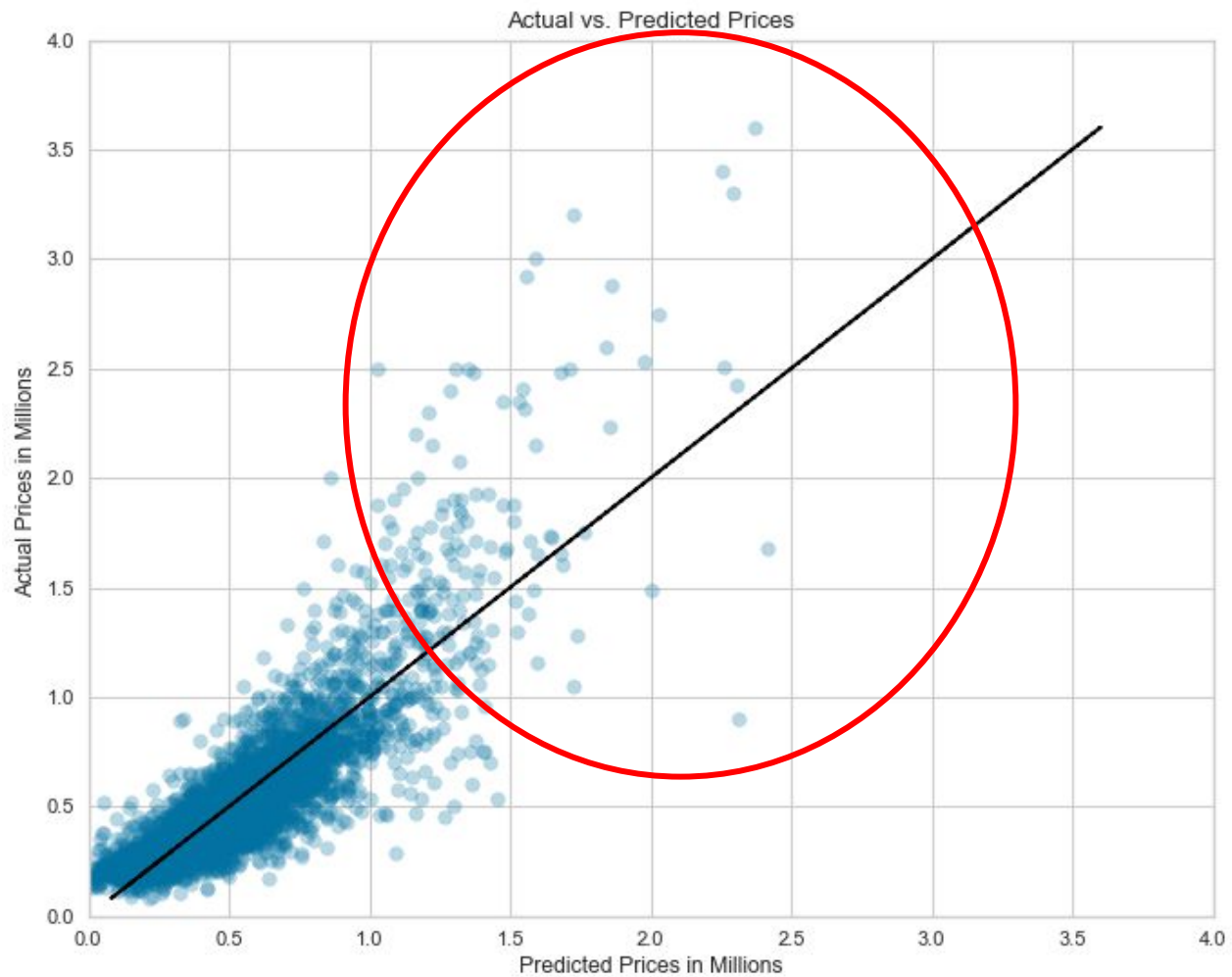Categories: Baseline Model, No Categorical Values, Pre-Clean Data, Cleaned Data, Colinearity Issues, Excluding Outliers, Feature Engineering, More Feature Engineering

Actual vs. Predicted Prices
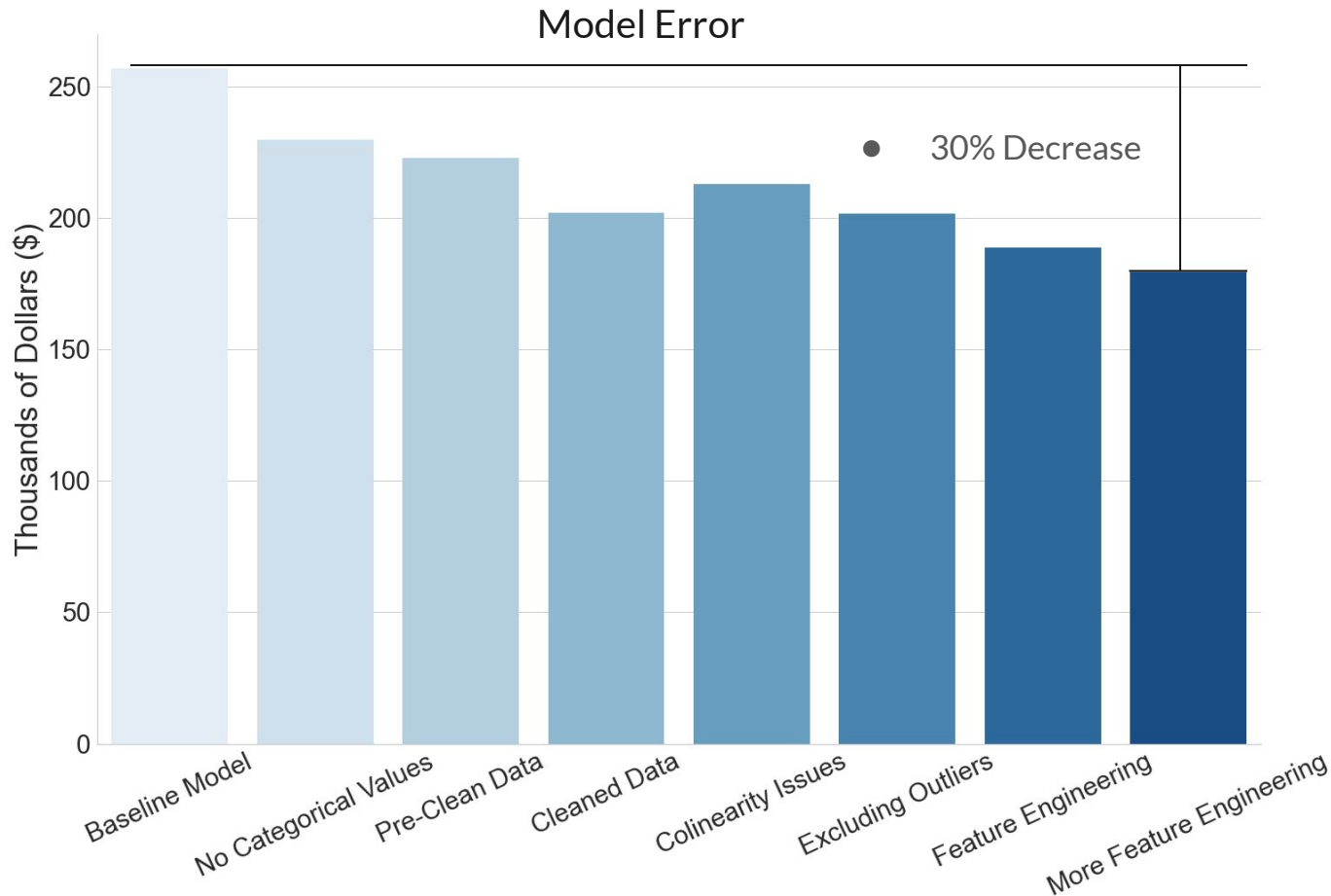
R-squared=
72.5%

RMSE =
$178,958

Actual vs. Predicted Prices

Actual vs. Predicted Prices
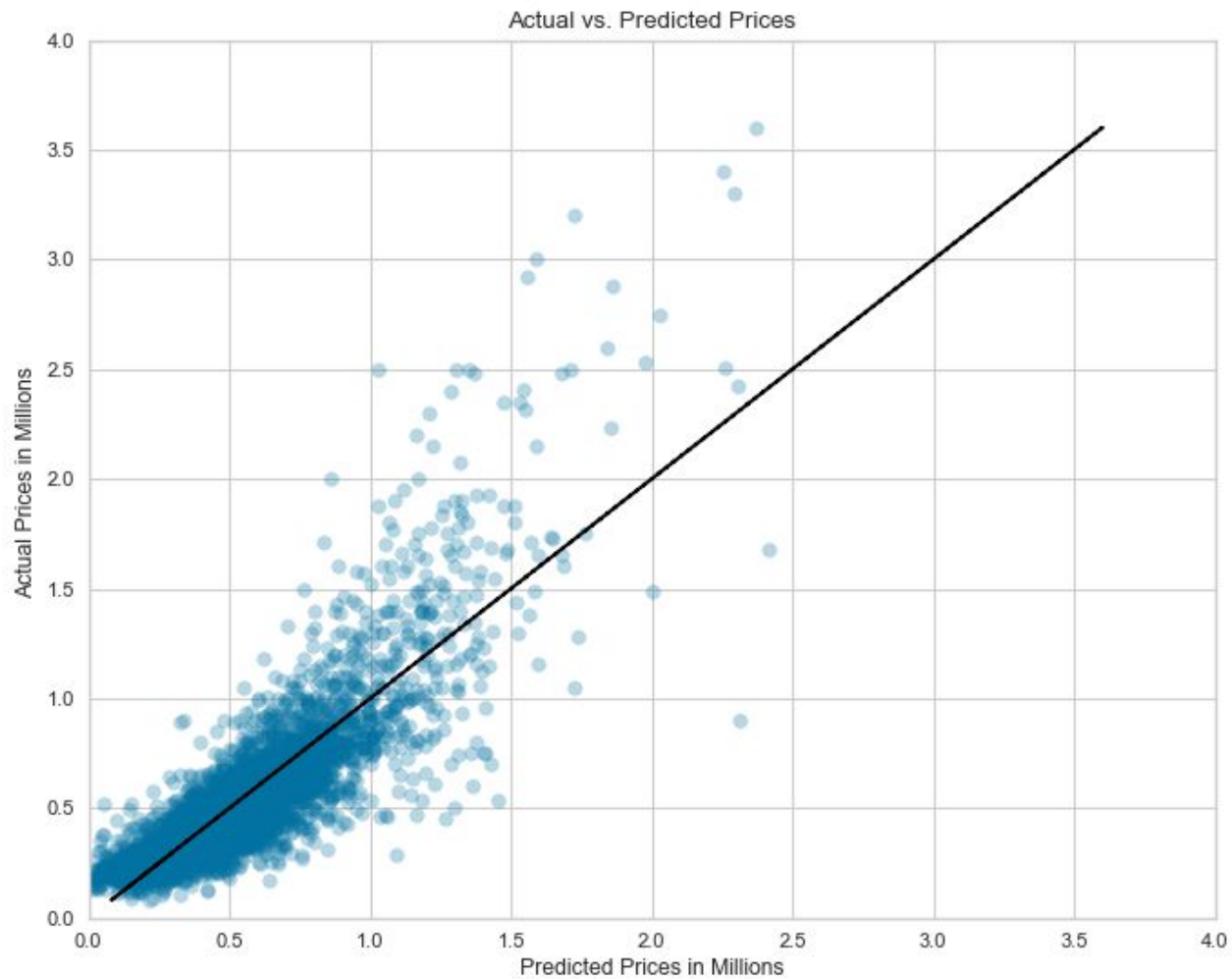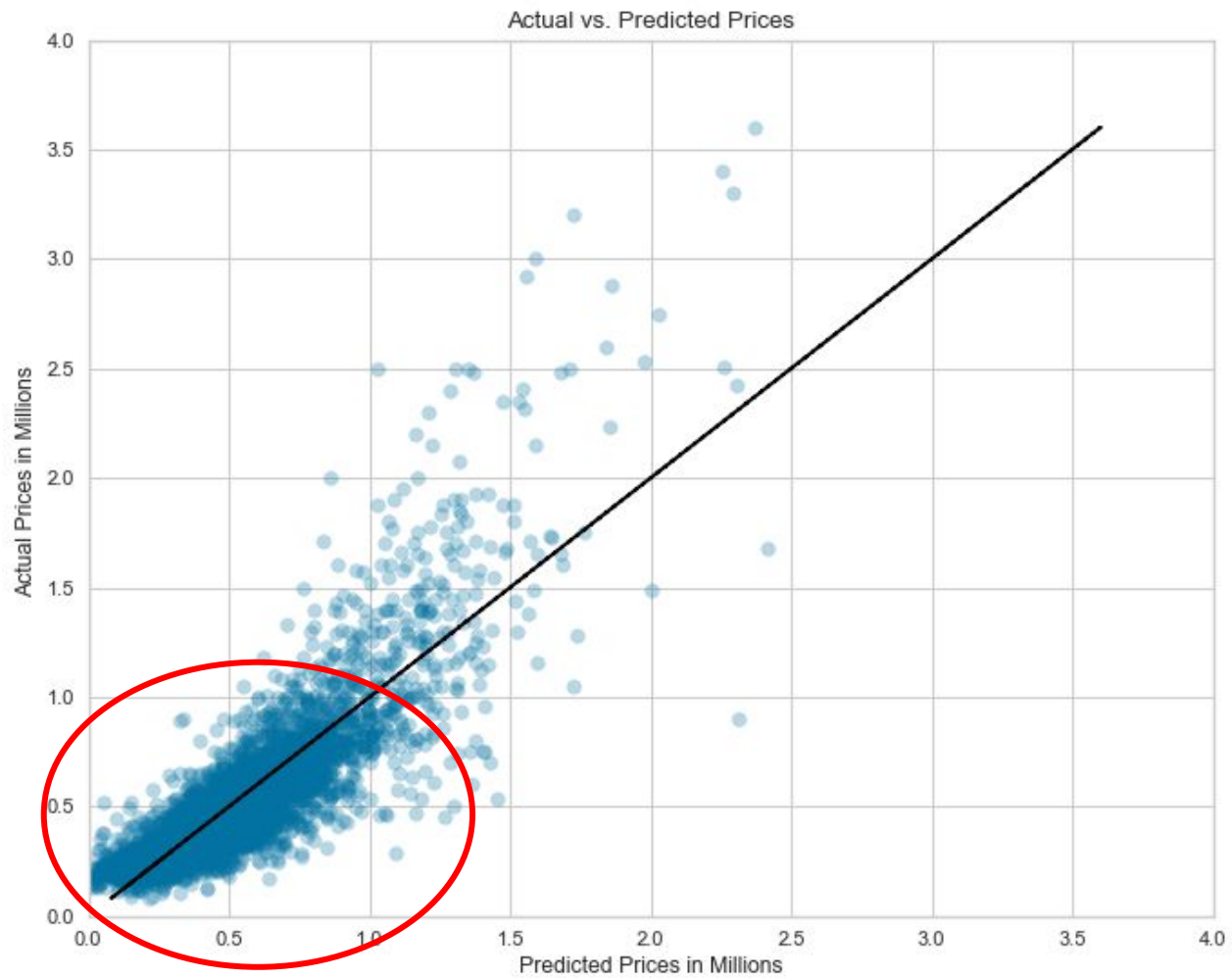
Coefficients

	coef	varname
0	600251.734733	waterfront
1	65448.436009	view
2	2717.348880	distance_kent
3	284.755704	sqft_living
4	-180.599350	sqft_basement
5	-23405.346026	floors
6	-29661.915030	distance_redmond
7	-40054.625132	distance_seattle
8	-40911.789568	bedrooms
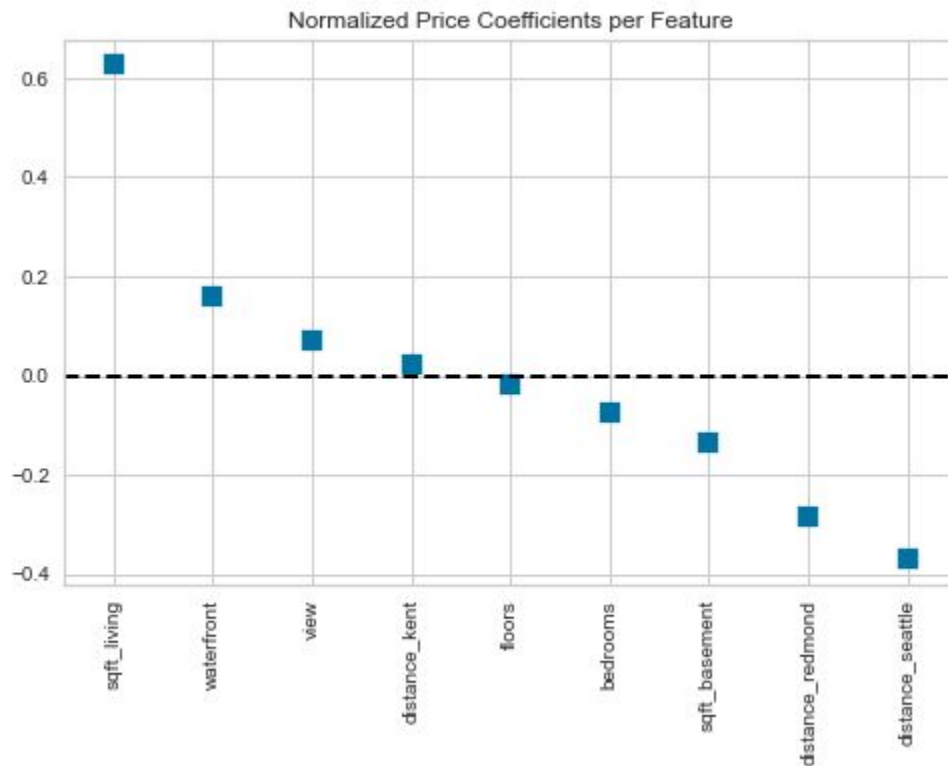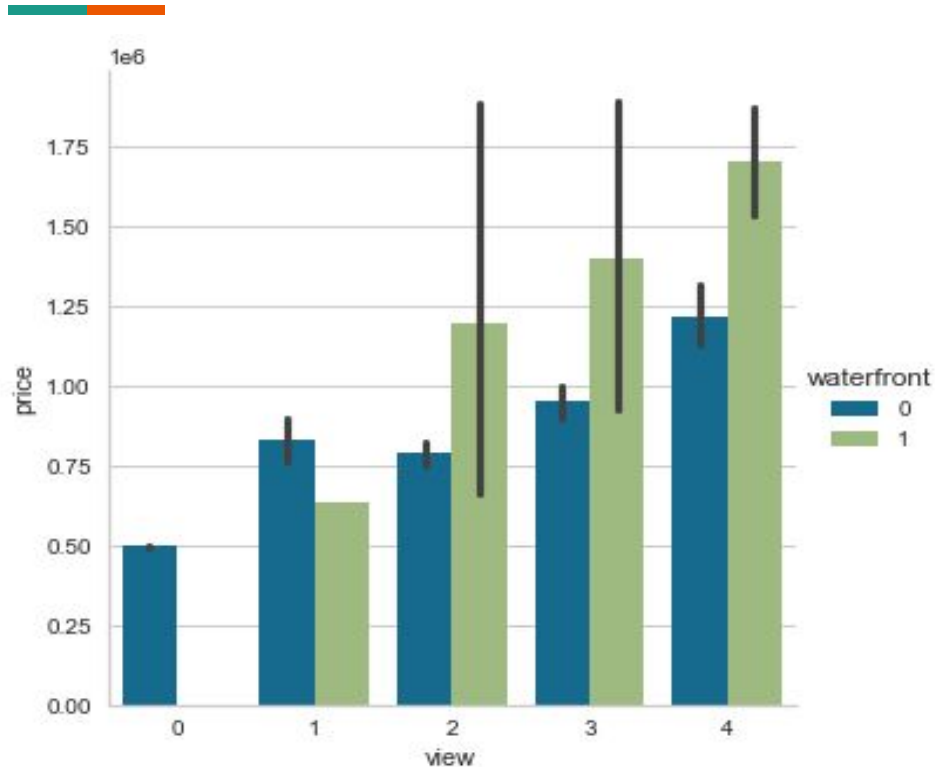
Normalized Price Coefficients per Feature

View and Waterfront vs Price

Homes with a higher view rating are also more likely to be a waterfront property.

Waterfront increase: $564,000-637,000

View per rating increase: $61,300-69,600

View and Waterfront vs Price

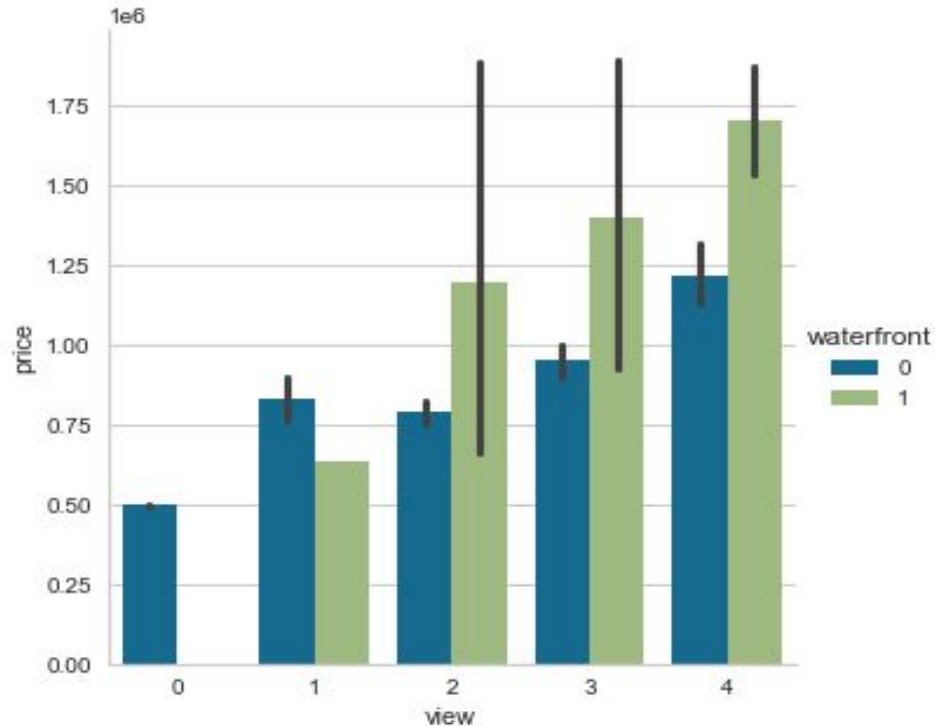
Homes with a higher view rating are also more likely to be a waterfront property.

Waterfront increase: $564,000-637,000

View per rating increase: $61,300-69,600

View and Waterfront vs Price
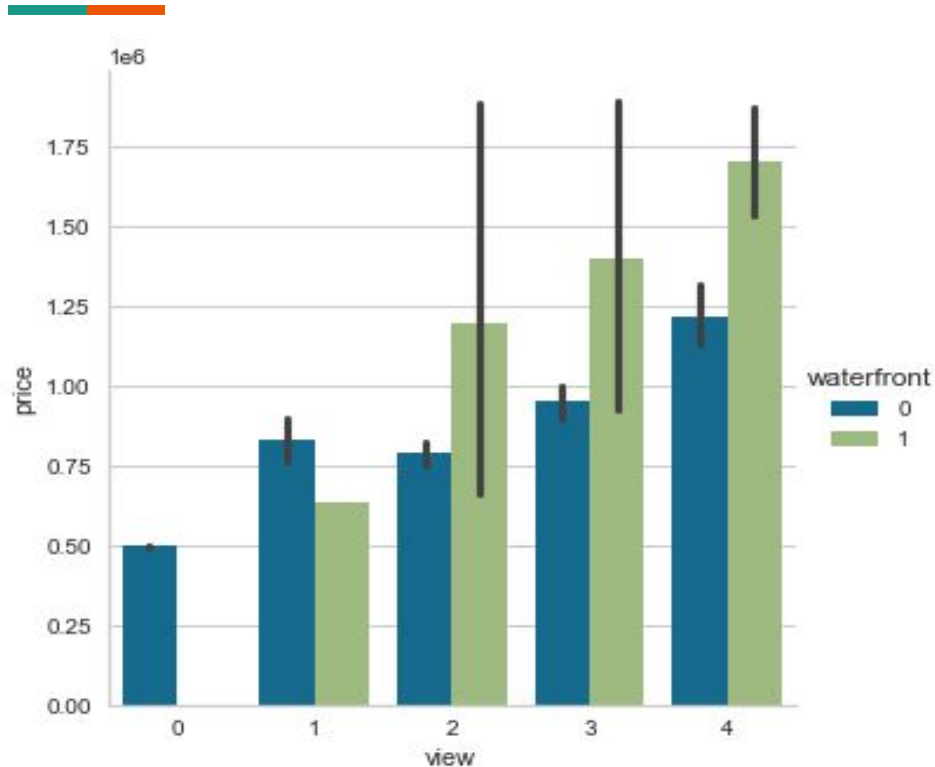

Homes with a higher view rating are also more likely to be a waterfront property.

Waterfront increase:
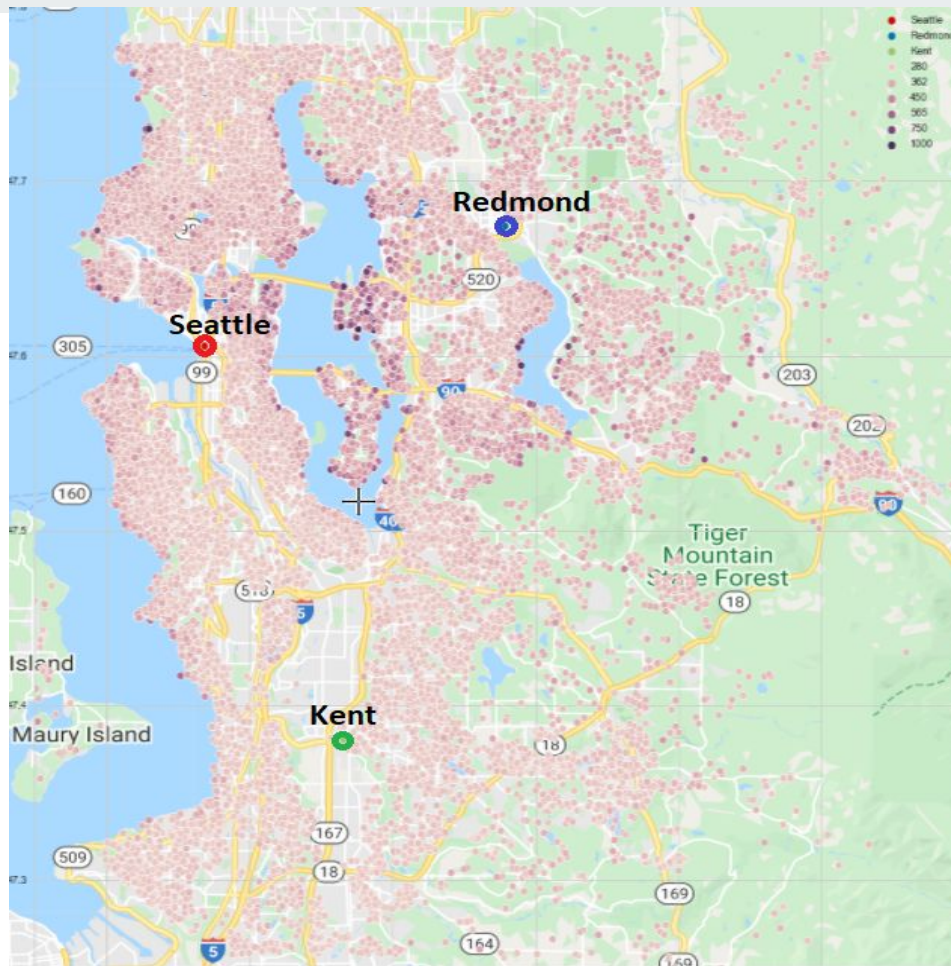$564,000-637,000

View per rating increase:
$61,300-69,600

Location v Price

Higher priced houses around Lake Washington and Lake Sammamish

Distance from Kent increases:
+$2,161 - 3,274

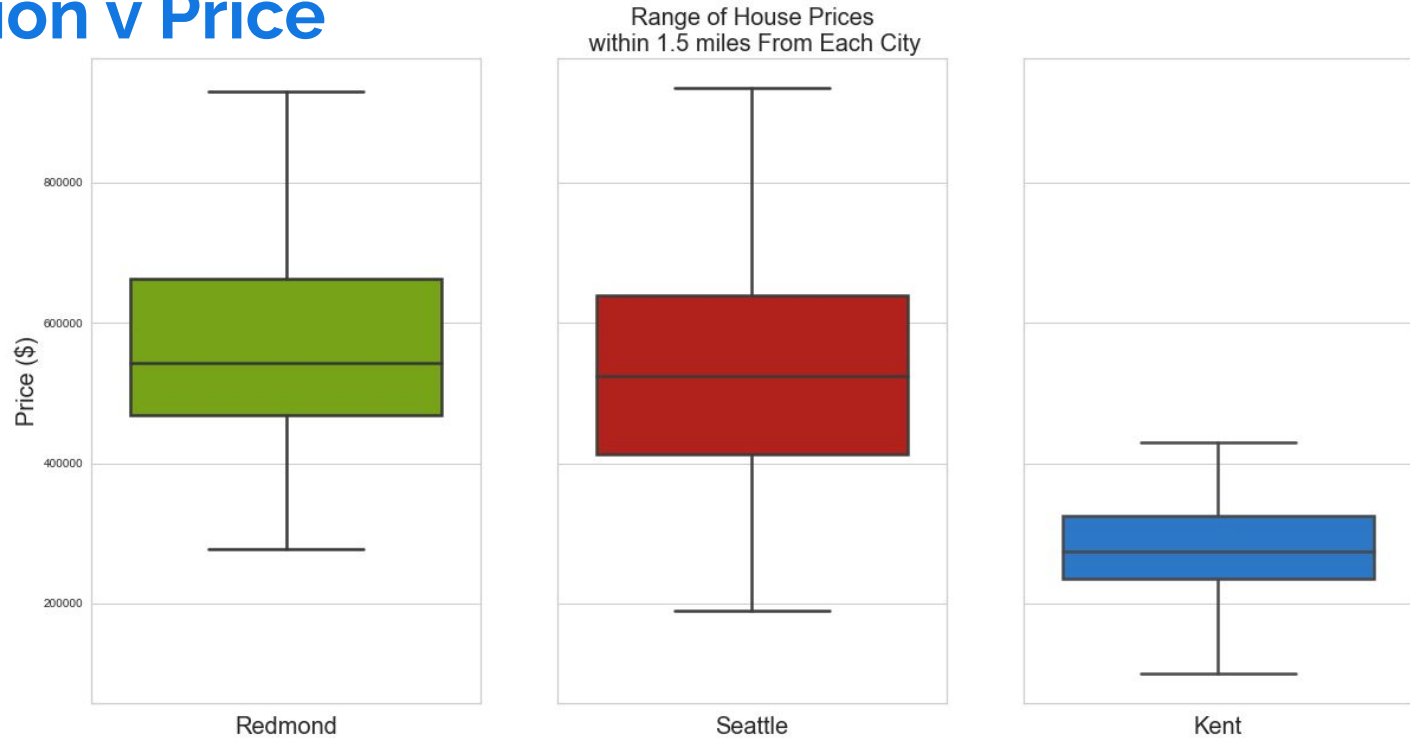Distance from Seattle decreases:
-$38,800 - 41,300

Distance from Redmond decreases:
-$28,200 - 31,100

Location v Price

- Range of prices close to city center

Range of House Prices
within 1.5 miles From Each City

Redmond Seattle Kent

Conclusions

Recommendations

- Use our predictive model to help predict house prices

- Avoid looking at expensive houses close to bodies of water and with great view

- Look around Kent to find a higher number of cheaper houses

Next Steps

- Find more features that home buyers often value highly to add to the model.

- Correlate the information of this model with ones for other states.

Thanks!

Any questions?

Find us at:

Hoang Nguyen
GitHub @DataDoggo215
hvnguyen90@gmail.com

Marcelo Scatena
GitHub @moscatena
moscatena@gmail.com

Madoria Thomas
GitHub @madoriathomas
deaudrey011@hotmail.com

Credits

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by SlidesCarnival and Google
- Photographs by SeattleMet and King County

More info on how to use this template at slidescarnival.com/help-use-presentation-template
This template is free to use under Creative Commons Attribution license. You can keep the Credits slide or mention SlidesCarnival and other resources used in a slide footer.