

Data Science

HOUSING PRICE PREDICTION

Bao Tram Duong



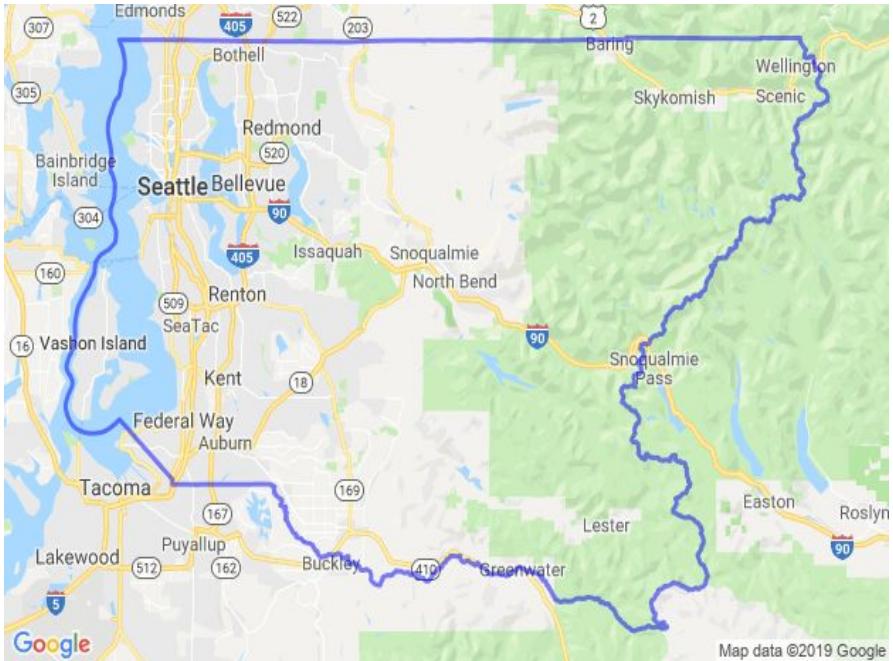
01

INTRODUCTION

- The real estate industry plays an integral part in the U.S. economy.
- It is now stands at \$33.6 trillion, which is equal to the combined GDP of the U.S. and China, the two largest economies of the world.
- Purchasing and selling a house is among the biggest commitments and a great source of income for most people. Therefore, accurate prediction of prices based on sale data can be a critical tool to assist the buyer/seller in making an informed decision.

Data Source

- King County
- Located in the U.S. state of Washington.
- The most populous county in Washington and the 12th-most populous in the U.S.
- 21,597 listings from May 2014 to May 2015



Home Features



Business Statement



01

What are the most predictive features to predict the price of a home?



02

How to increase the value of a home?



03

How does age affect the value of a home?



02

METHODOLOGY

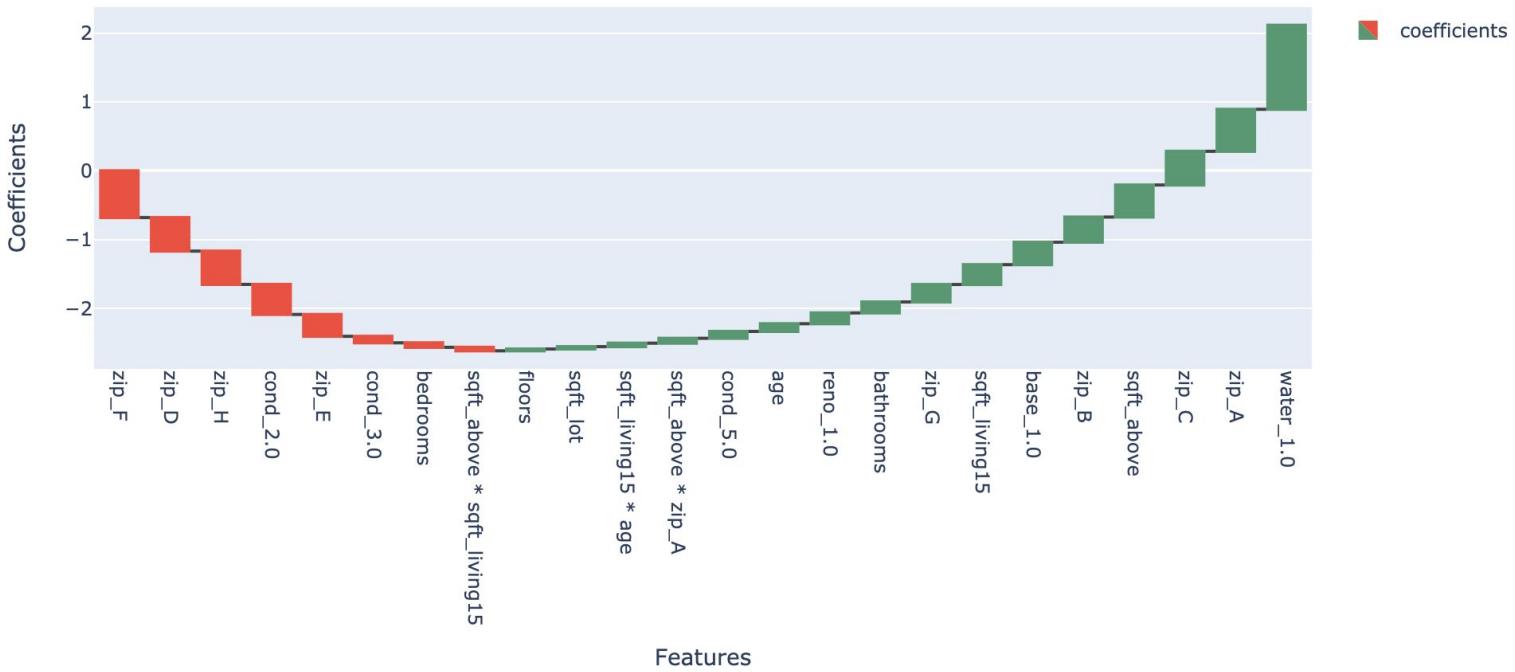
- Perform EDA to understand the insight of the data.
- Create the prediction model with the highest accuracy that is able to accurately estimate the price of the house given the features.
- **Outcome:** By cross-referencing our initial EDA and the model coefficients:
 - Predict the price of house accurately
 - Insights on what to look for when buying a new home
 - What to do to improve your current home's value

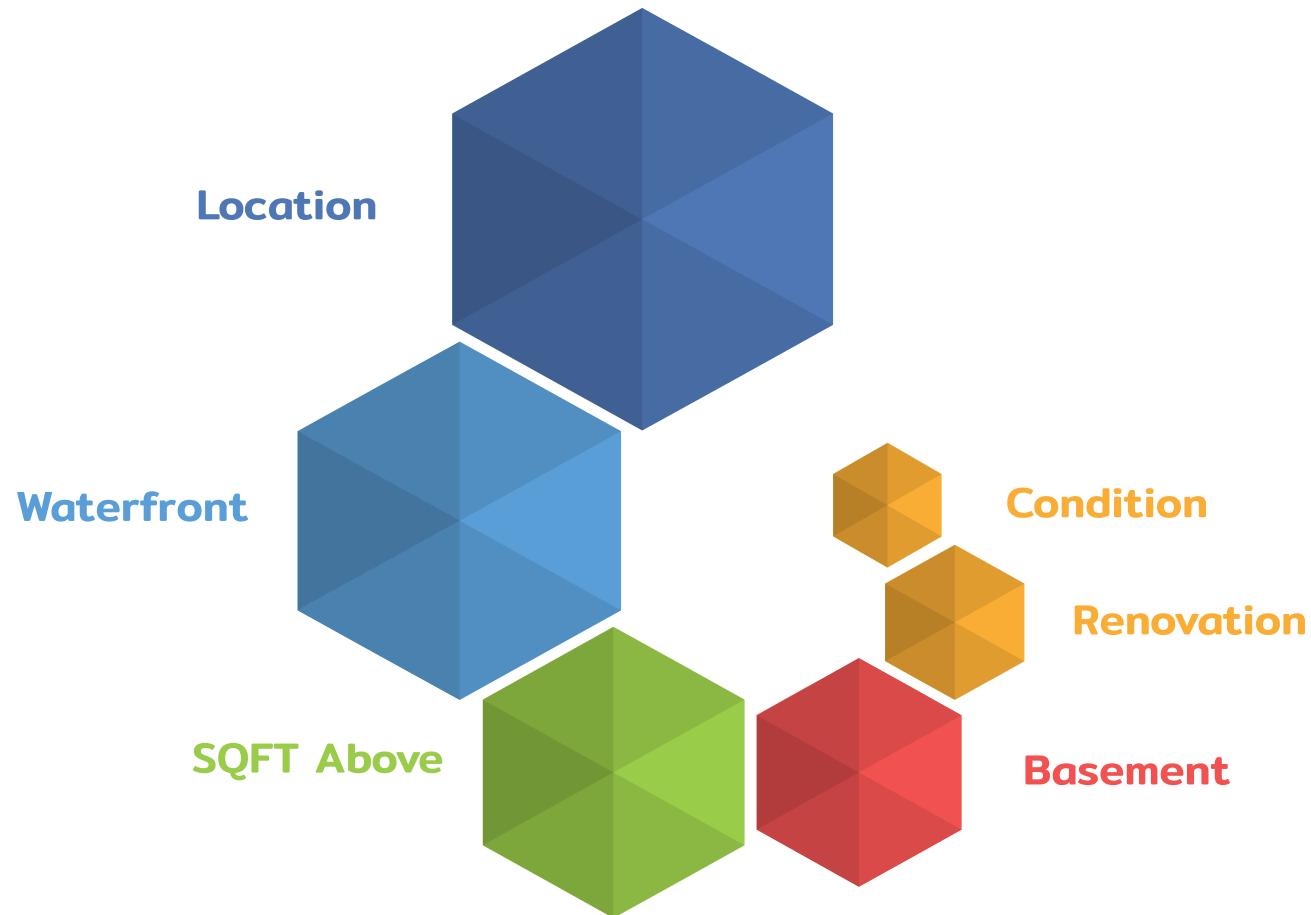


**What are the most
predictive features to
predict the price of a home?**

Accuracy
78%

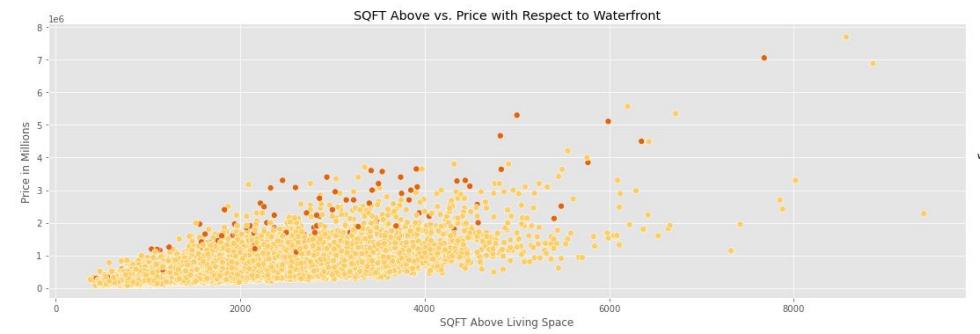
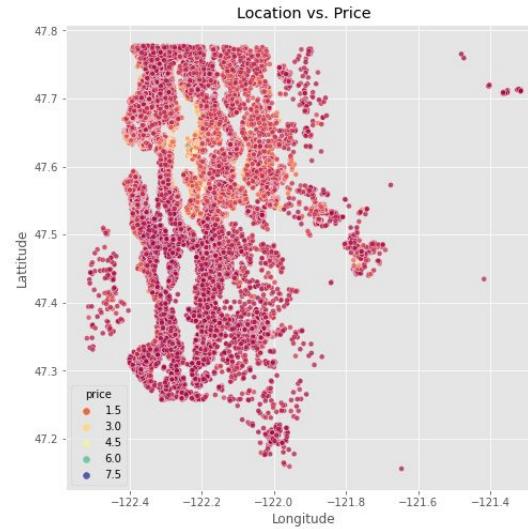
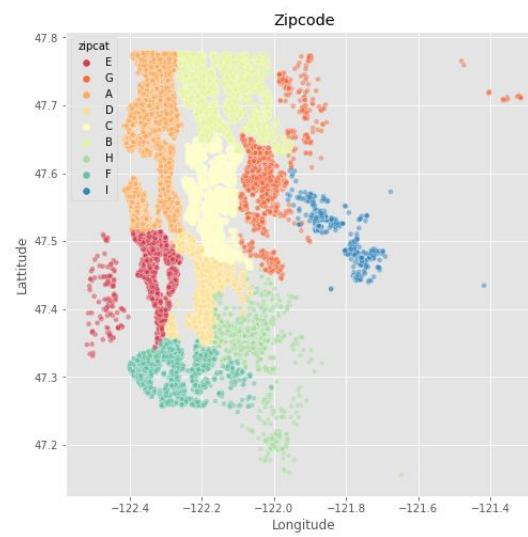
Model Fa Coefficients



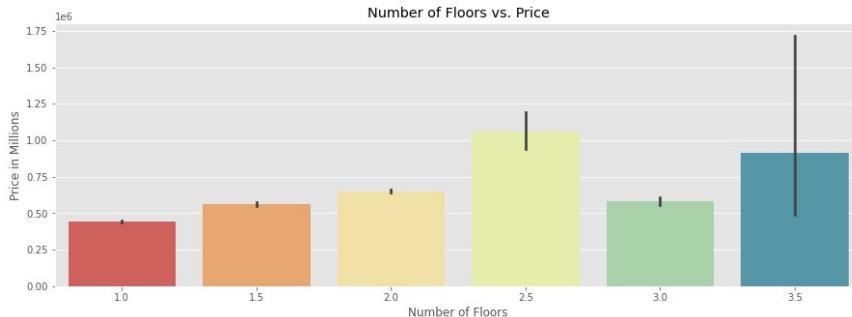
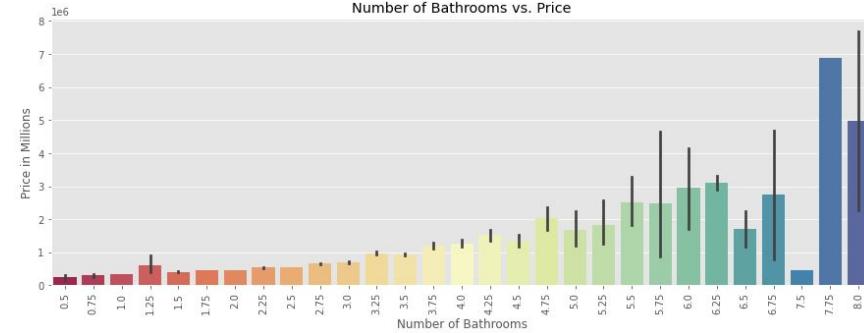
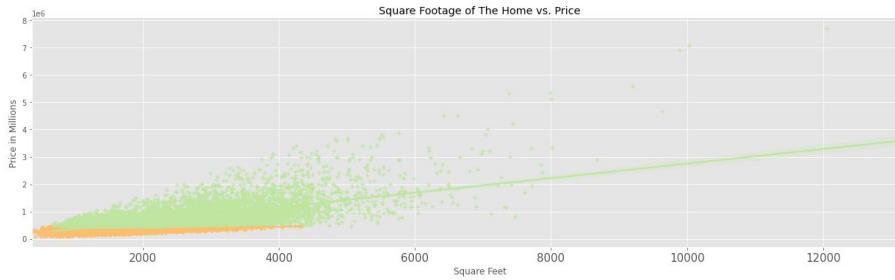




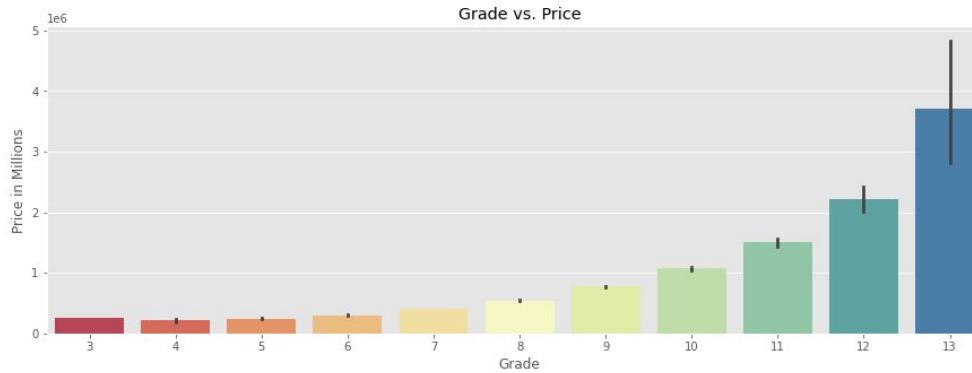
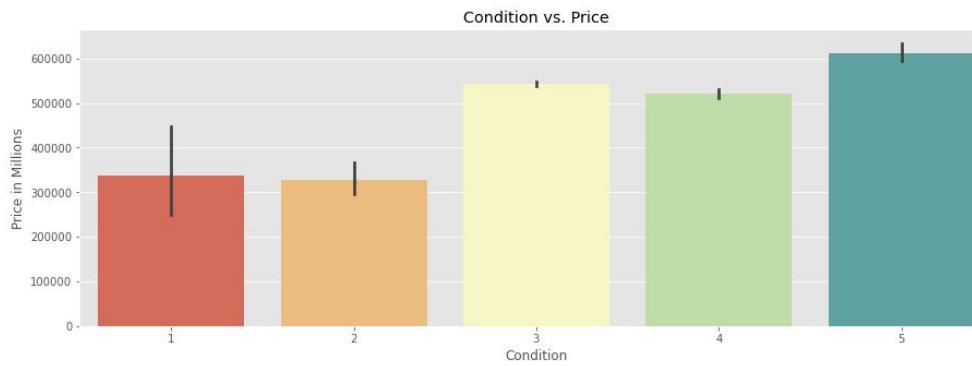
How to increase the value of a home?



SQFT Livable Space



Condition Grade





**How does age affect the
value of a home?**

Renovation





03

ACTIONABLE INSIGHTS



Location

- Waterfront
- Neighbors



Square Footage of Livable Space

- Bedrooms
- Bathrooms
- Floors



Renovation

- Basement



Condition: 3 and above



Grade: 8 and above



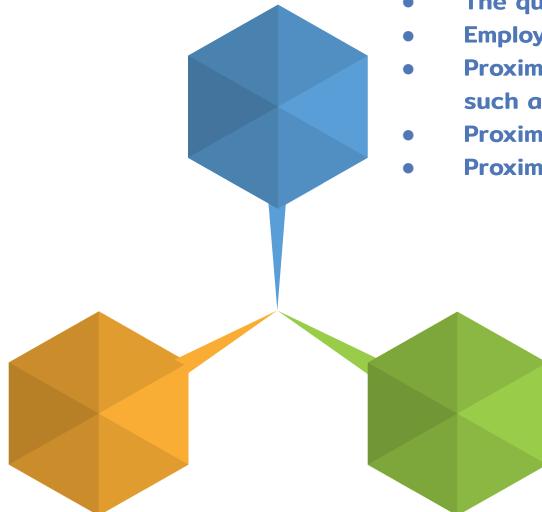
Age

- Renovation
- Improve condition & grade

04

OTHER FACTORS

When is the best time to buy or sell?



How hot (or cold) is the area's real estate market?

- The quality of local schools
- Employment opportunities
- Proximity to shopping, entertainment, and common services such as hospitals
- Proximity to highways, utility lines, and public transit
- Proximity to the nearest major city



THANK YOU

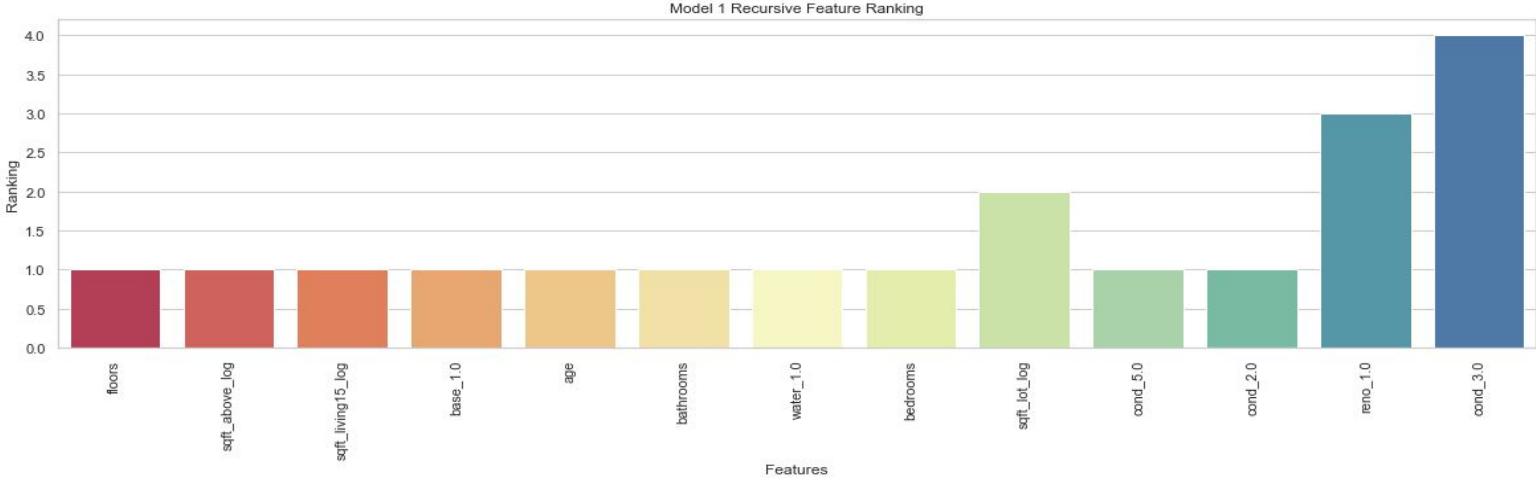


Q & A

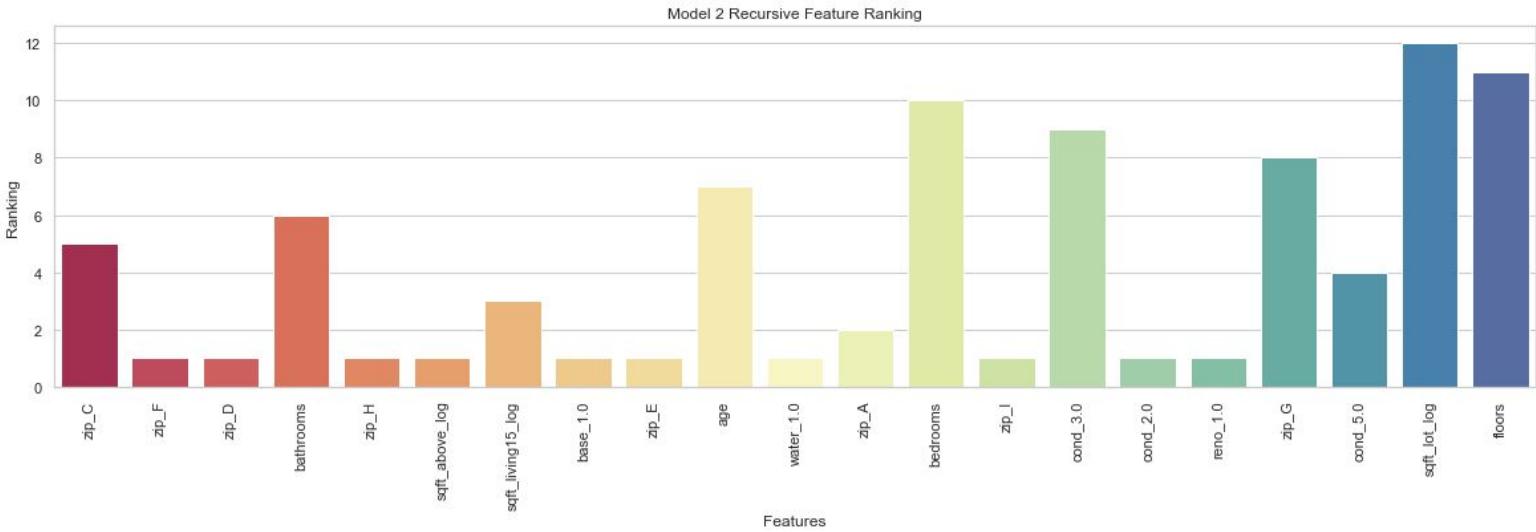


APPENDIX

Accuracy
59%

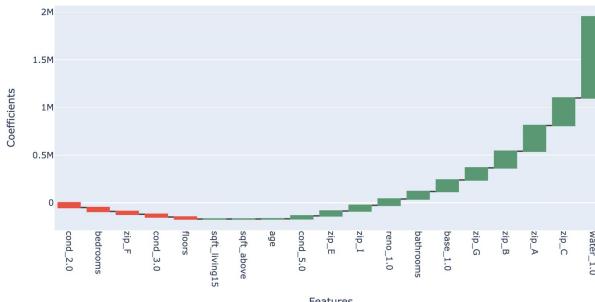


Accuracy
73%

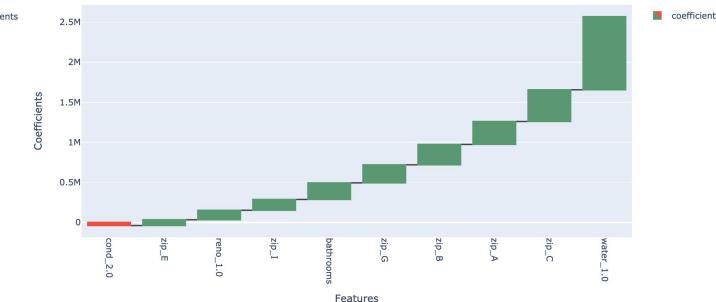


Coef

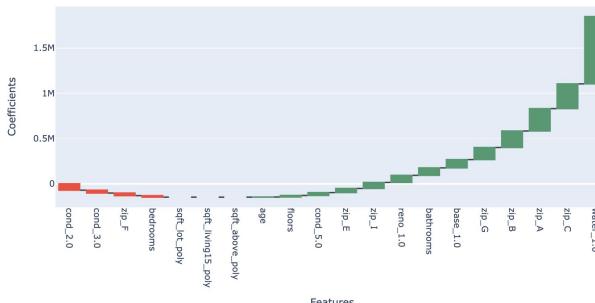
Model A Coefficients



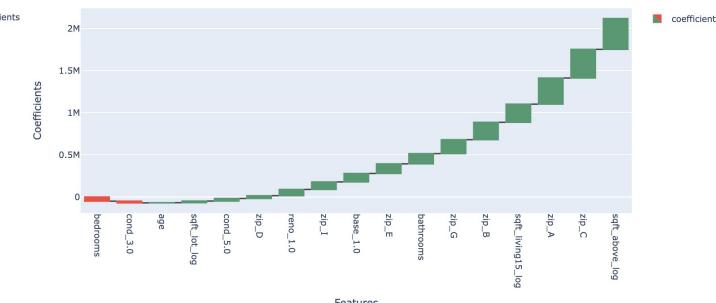
Model B Coefficients



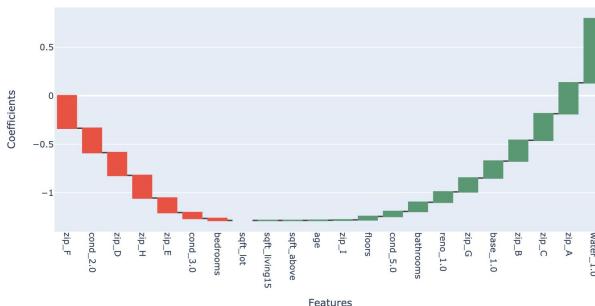
Model C Coefficients



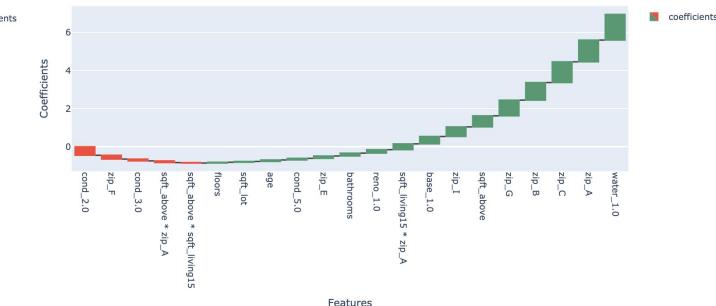
Model D Coefficients



Model E Coefficients

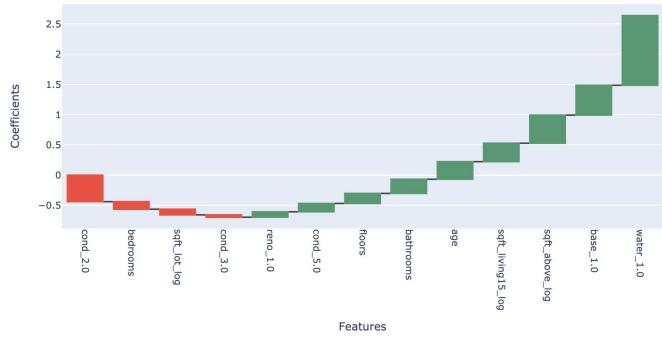


Model F Coefficients

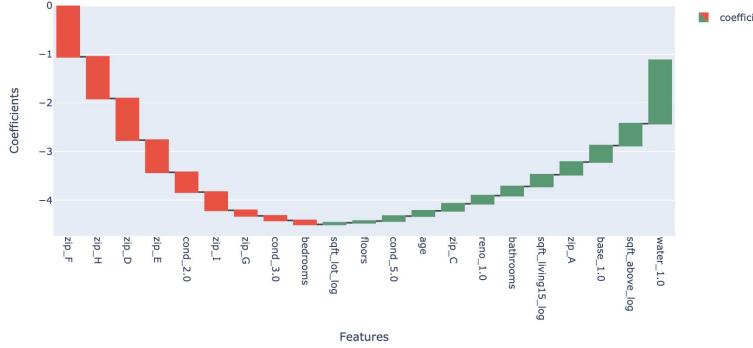


Coef

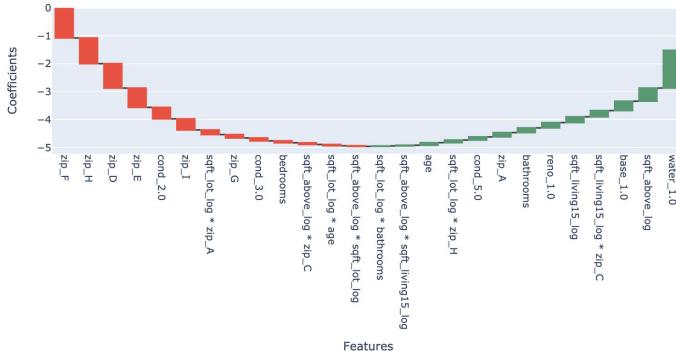
Model 1 Coefficients



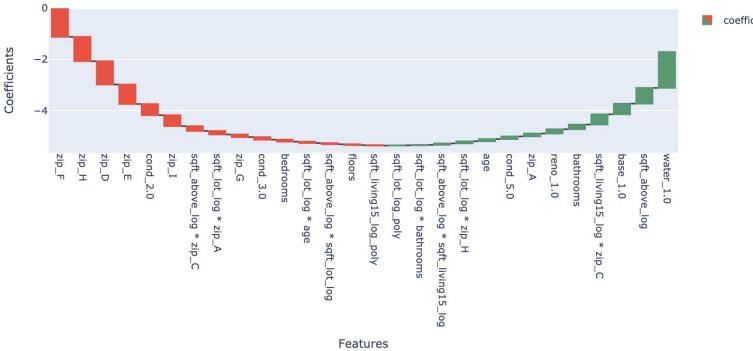
Model 2 Coefficients



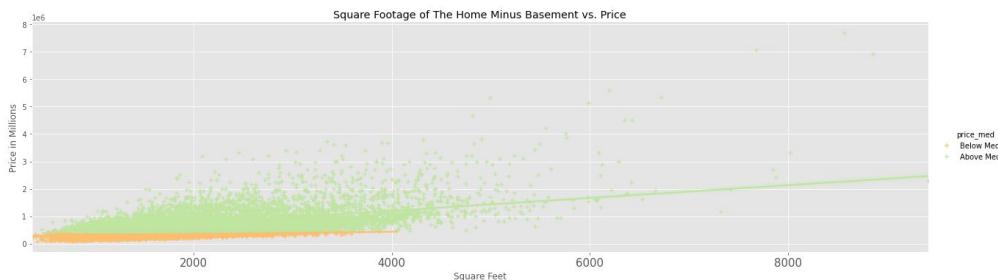
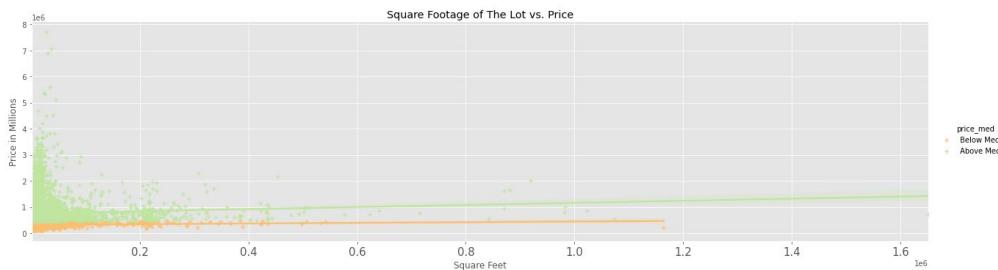
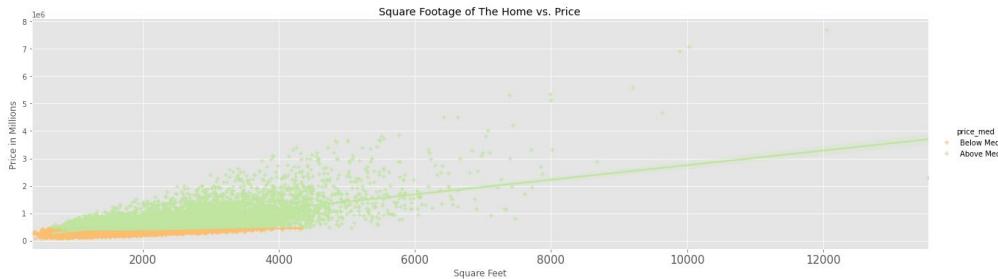
Model 4 Coefficients



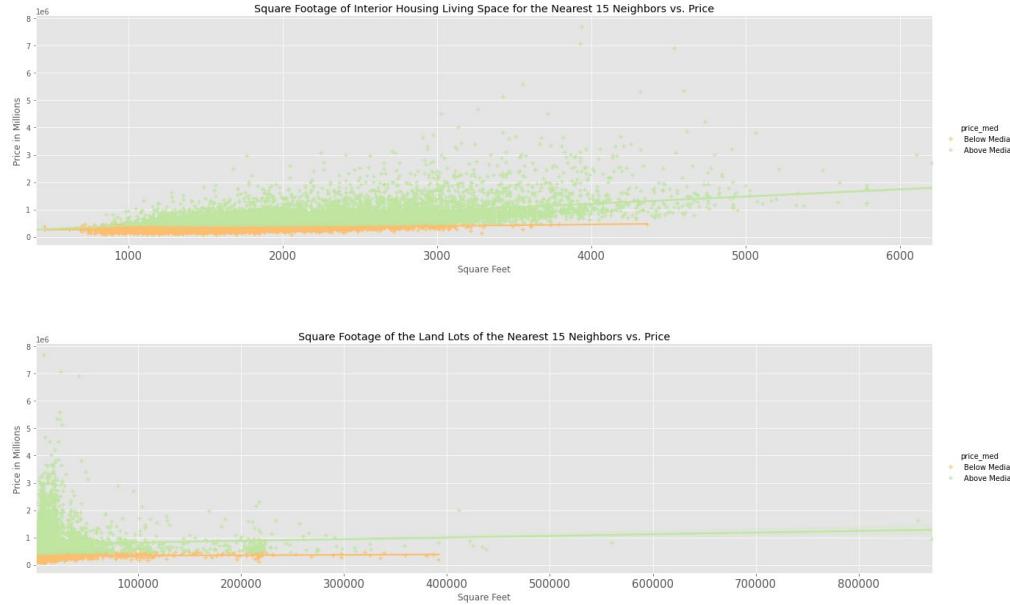
Model 5 Coefficients



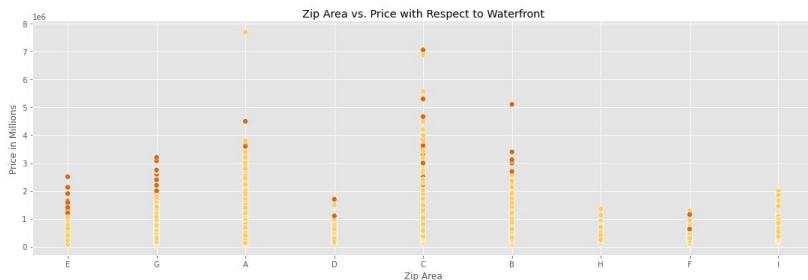
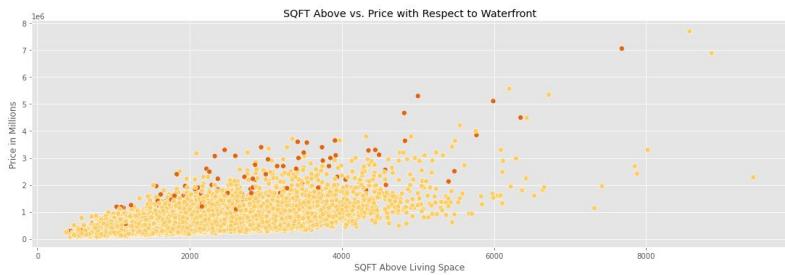
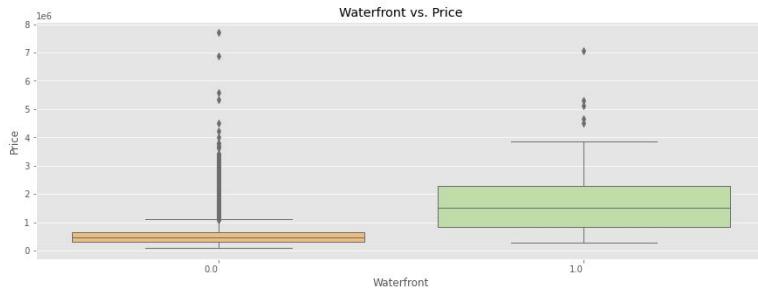
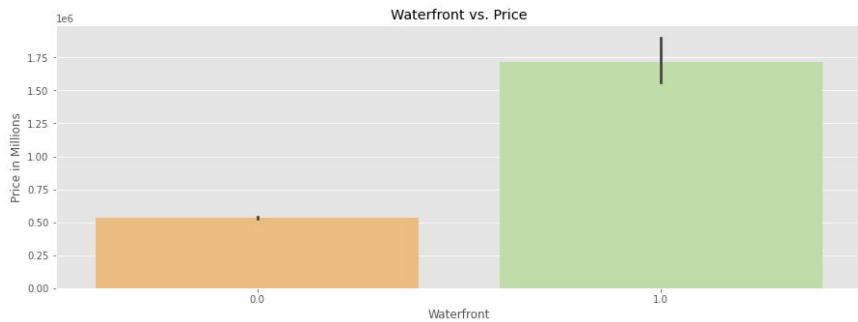
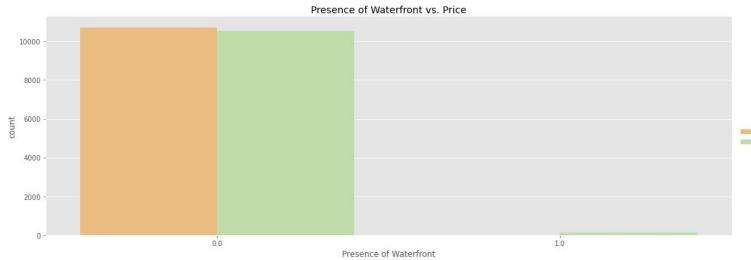
SQFT



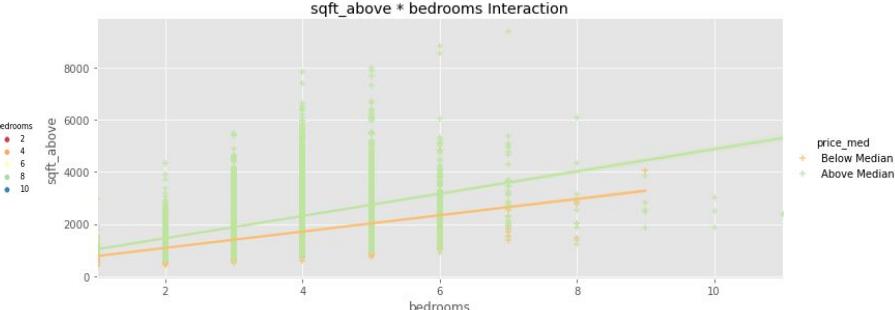
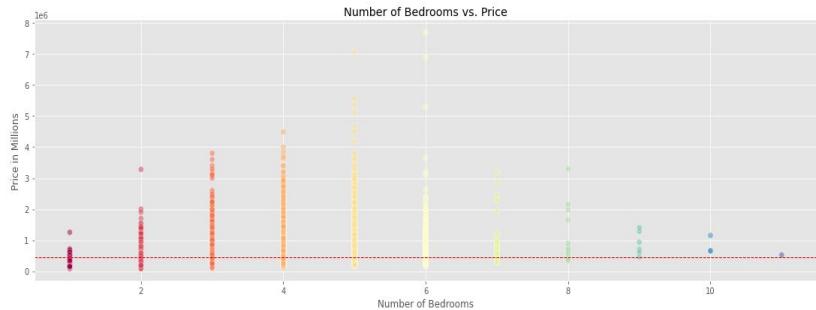
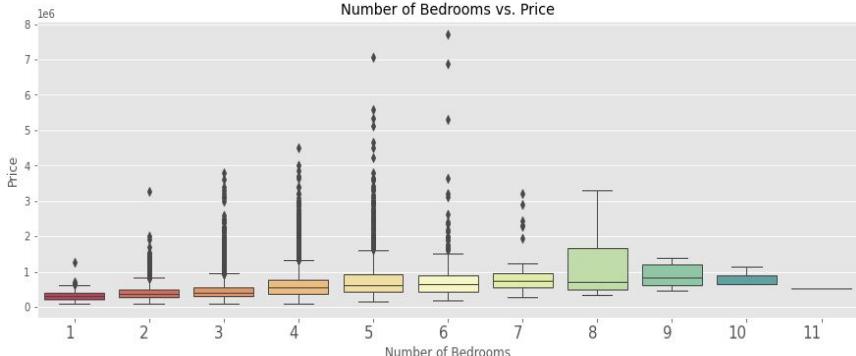
15 Neighbors



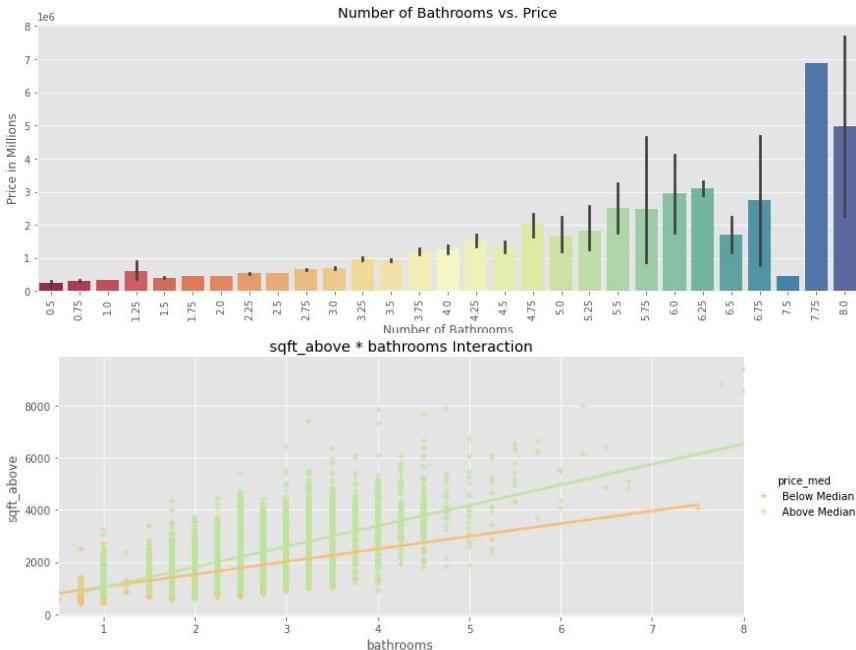
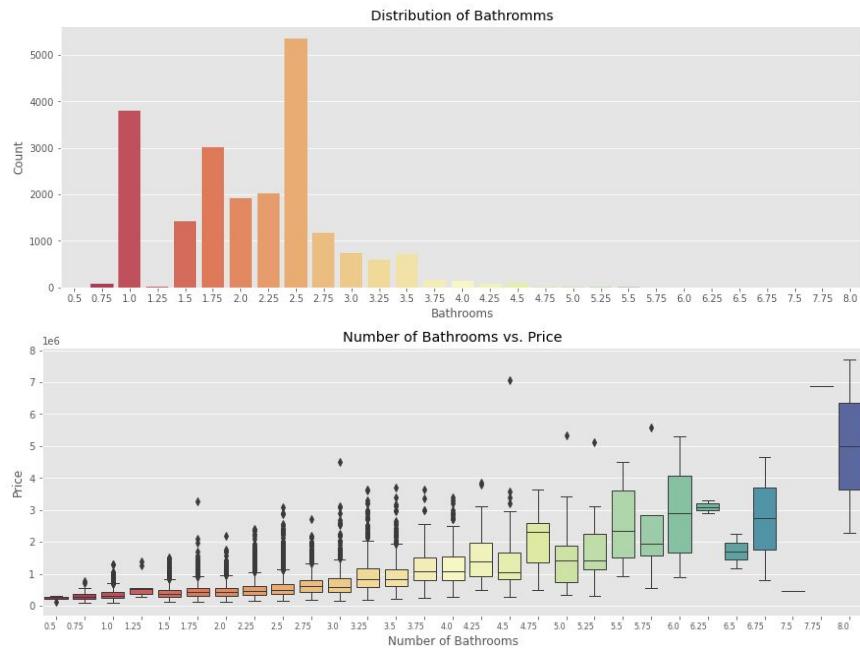
Waterfront



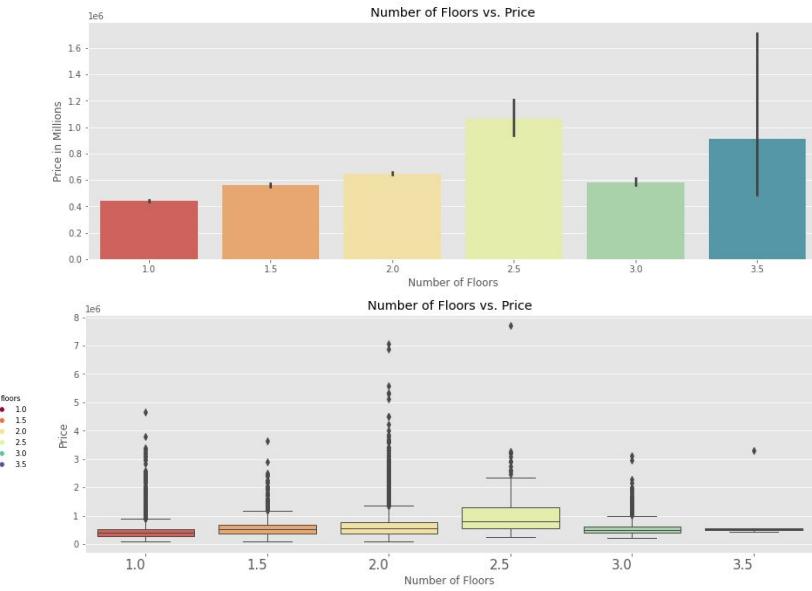
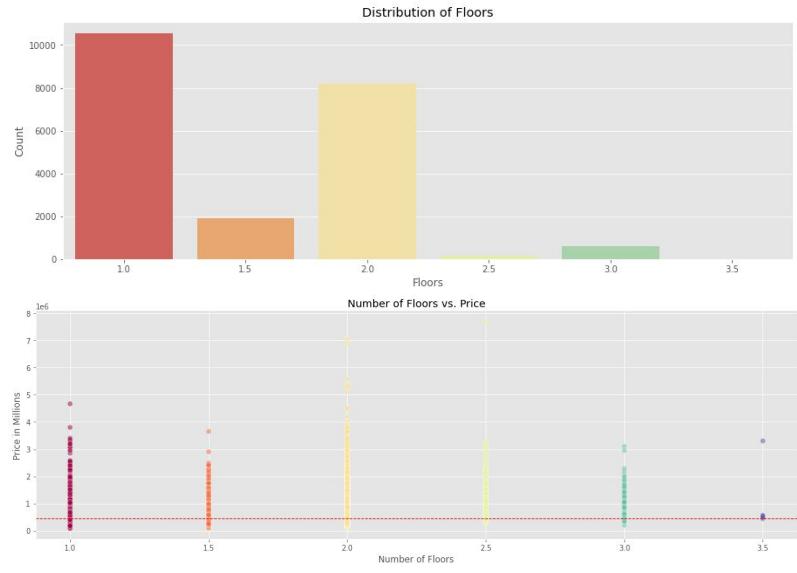
Bedrooms



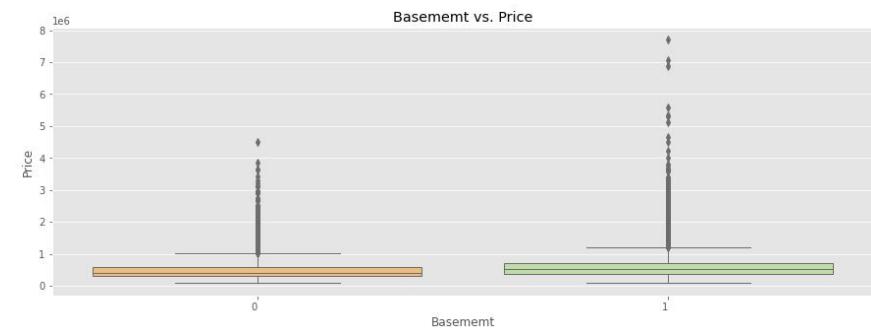
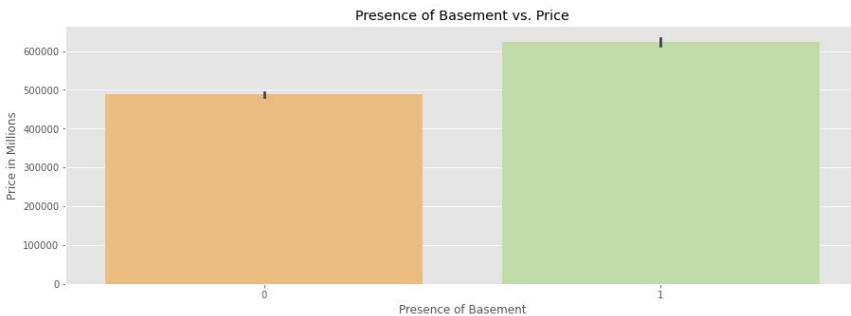
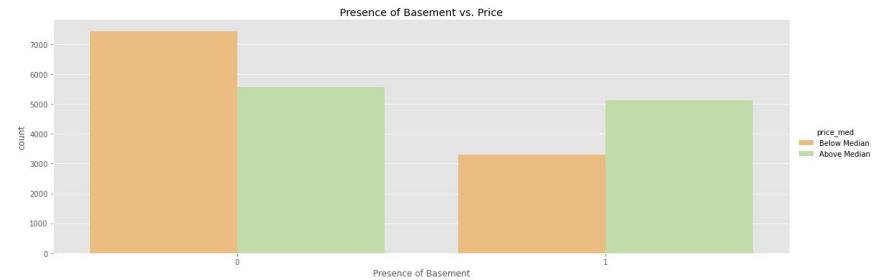
Bathrooms



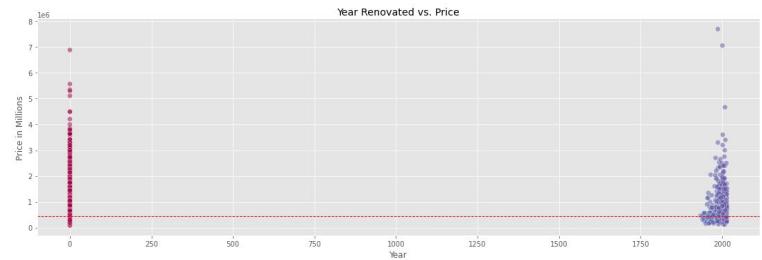
Floors



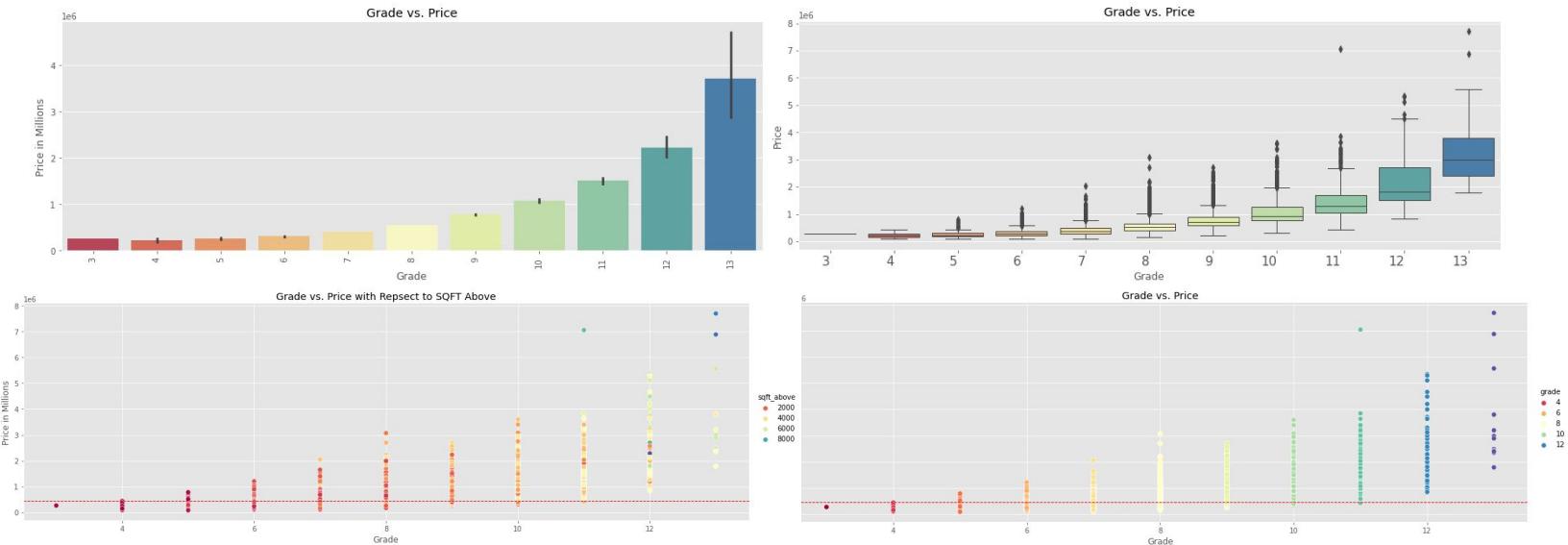
Basement



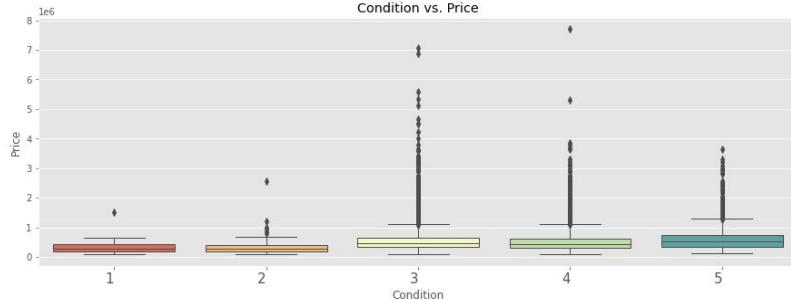
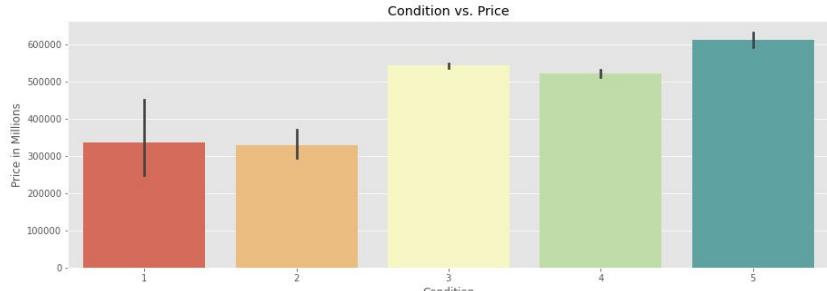
Renovation



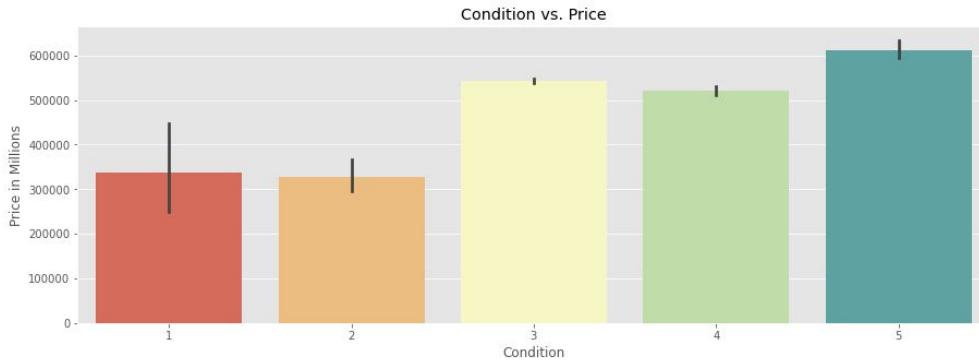
Grade



Condition



Condition



1 = Poor - Worn out.

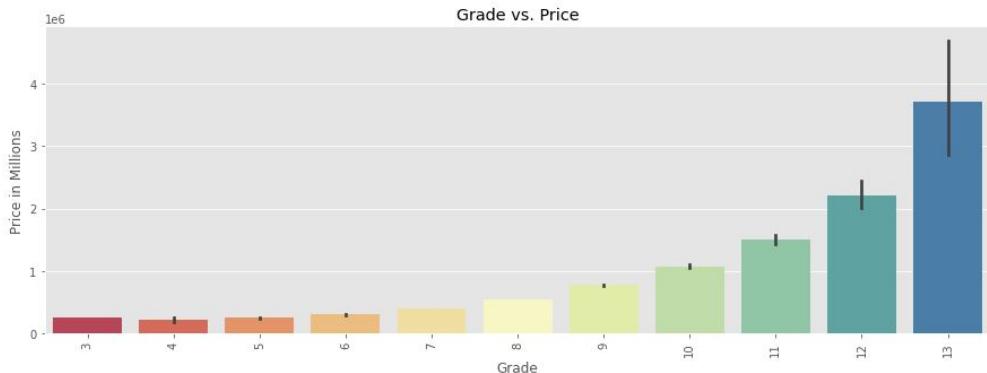
2 = Fair - Badly worn.

3 = Average

4 = Good

5 = Very Good

Grade



1-3 = Falls short of minimum building standards.

4 = Generally older, low quality construction.
Does not meet code.

5 = Low construction costs and workmanship.
Small, simple design.

6 = Lowest grade currently meeting building
code. Low quality materials and simple designs.

7 = Average grade of construction and design.

8 = Just above average in construction and
design.

9 = Better architectural design with extra
interior and exterior design and quality.

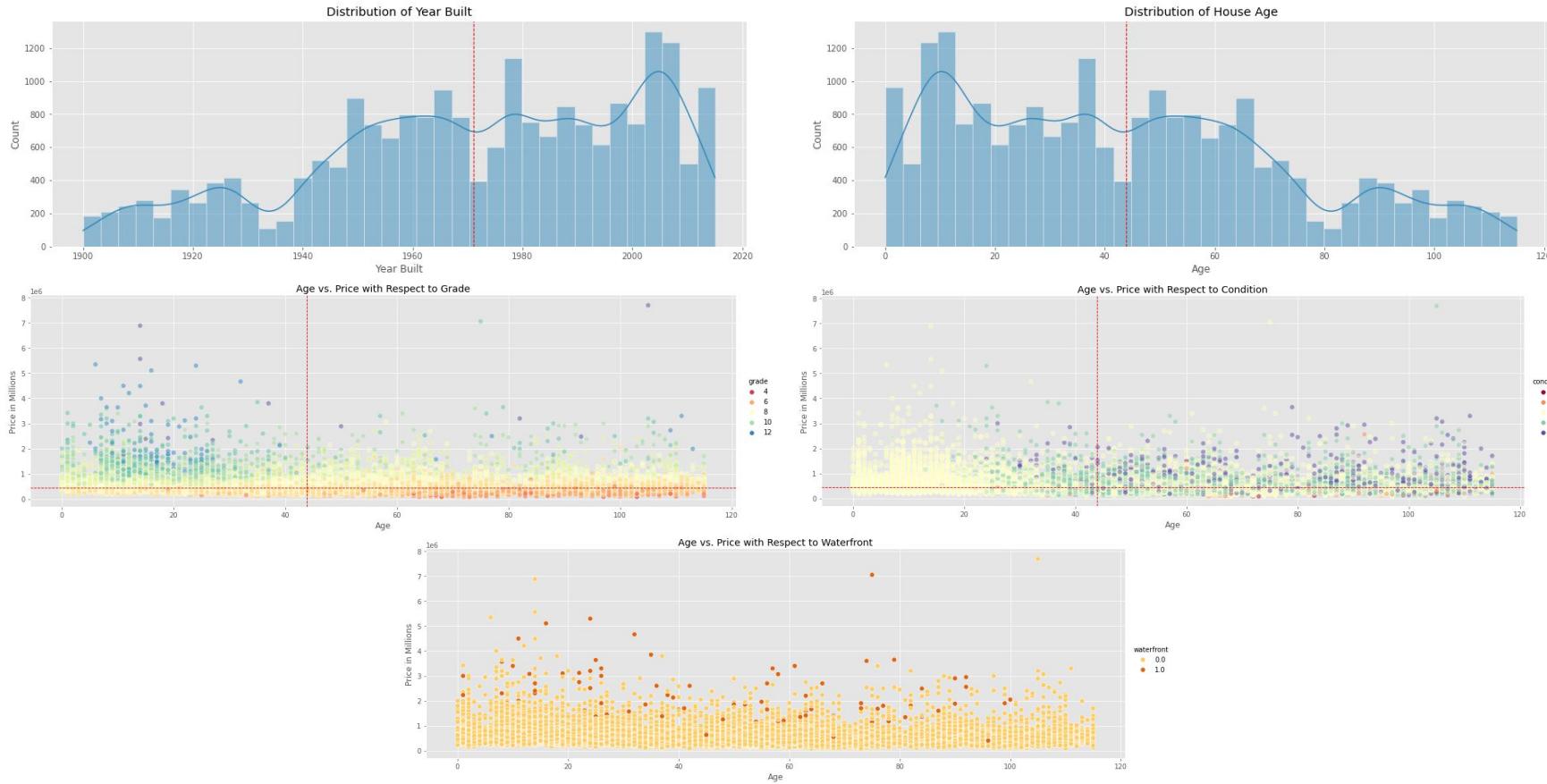
10 = High quality features.

11 = Custom design and higher quality finish
work with added amenities.

12 = Custom design and excellent builders.

13 = Generally custom designed and built.
Mansion level.

Age



Location

