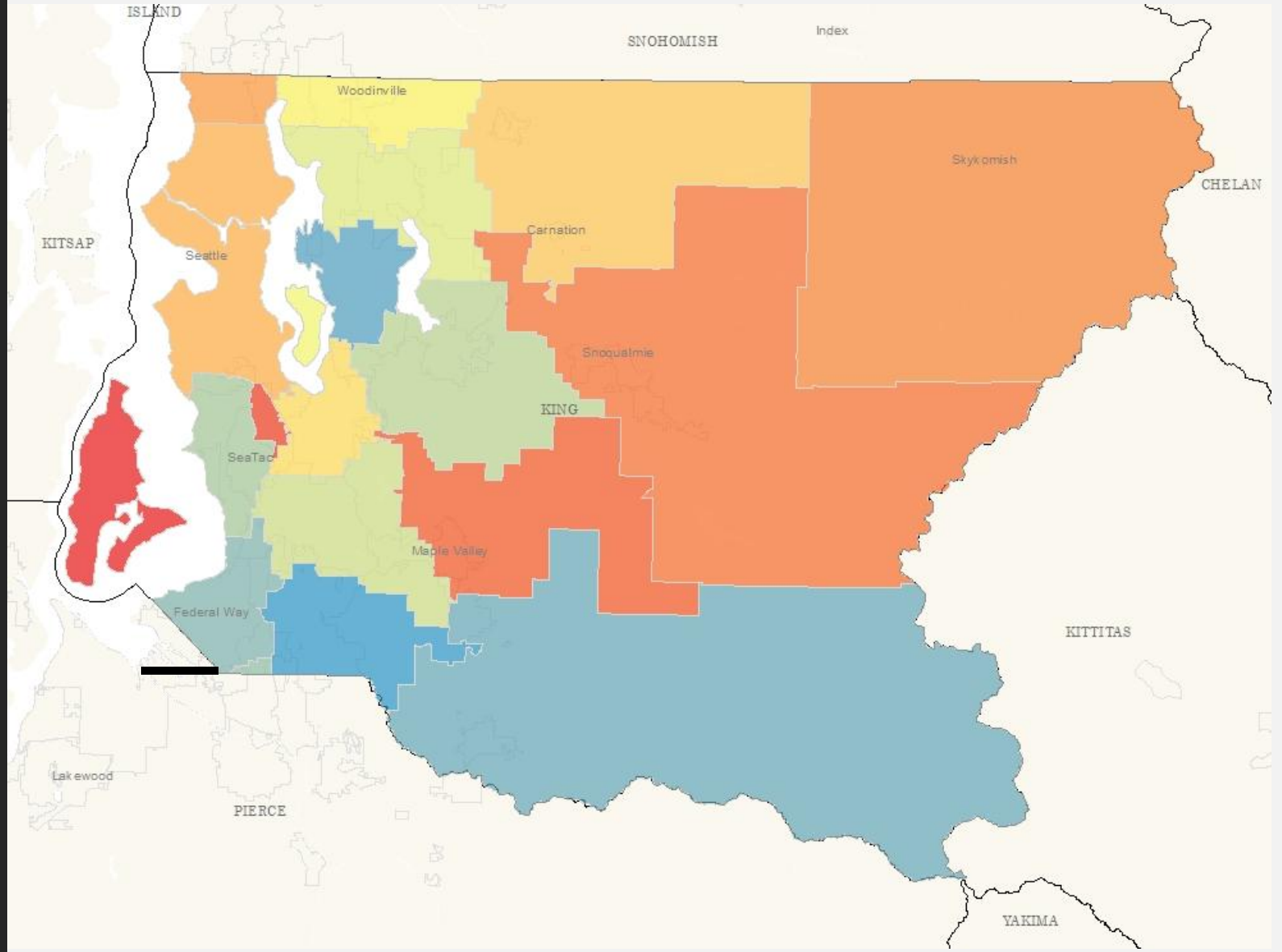
Regression Analysis on Homes in King County

Author: Anastasiia Leskiv

Project overview

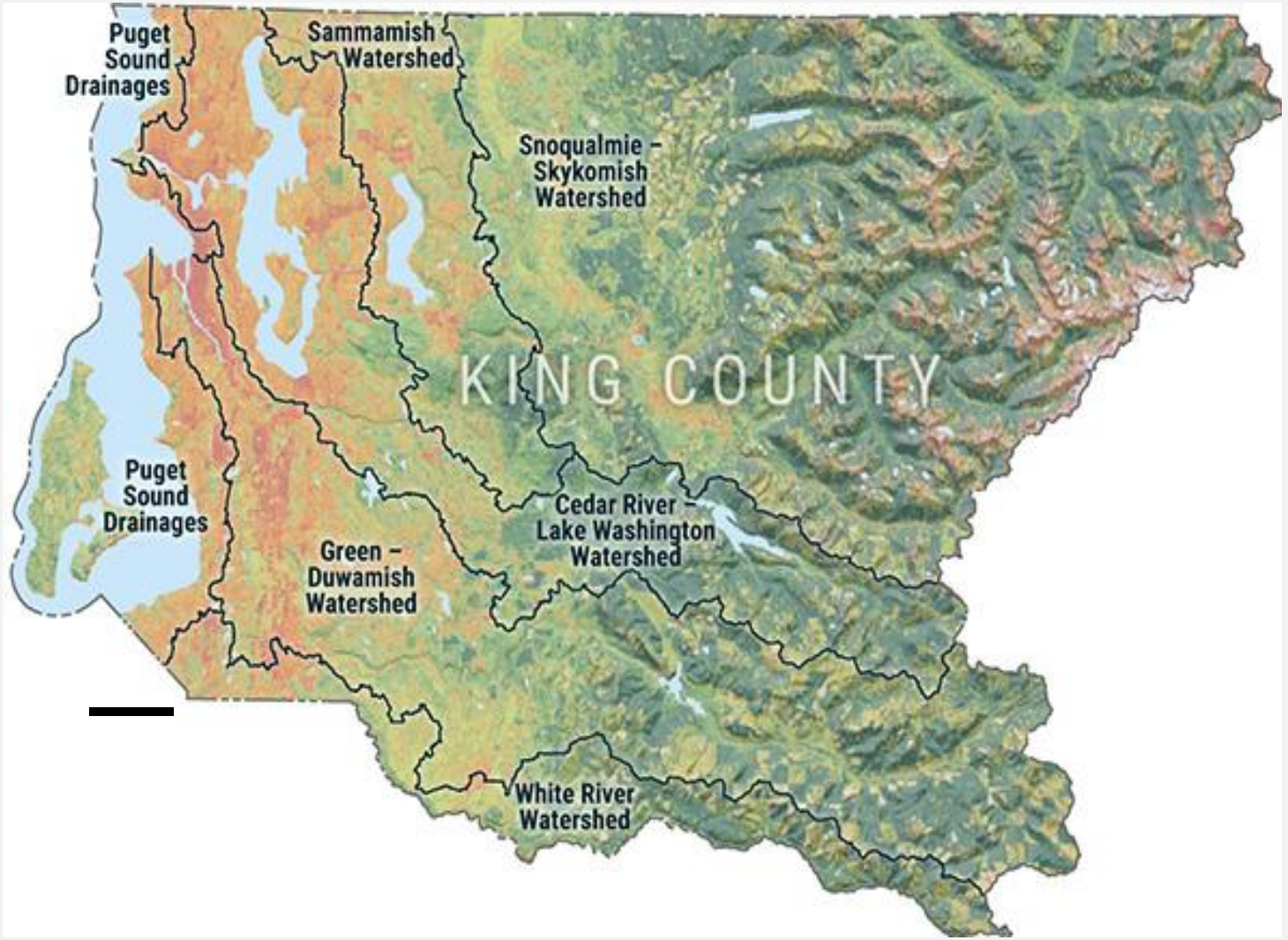
A client wants to buy single family home. I observed properties of King County for various home features and analyzed it.
Build a predictive regression model to predict the price of a home.



Business and Data Understanding

For the best prediction I used
-Data collecting
-Data cleaning
-Modeling
-Exploratory data analysis
-Visualization

For my client I checked
-The chipset house
-The most expensive
-How the footage of the home (sqft_living) affect the price?
-What features effect price



Tools

-CSV
-Pandas
-Numpy
-Matplotlib
-Seaborn
-Statsmodel
-SciPy Stats
-LinearRegression
-Lasso Regration

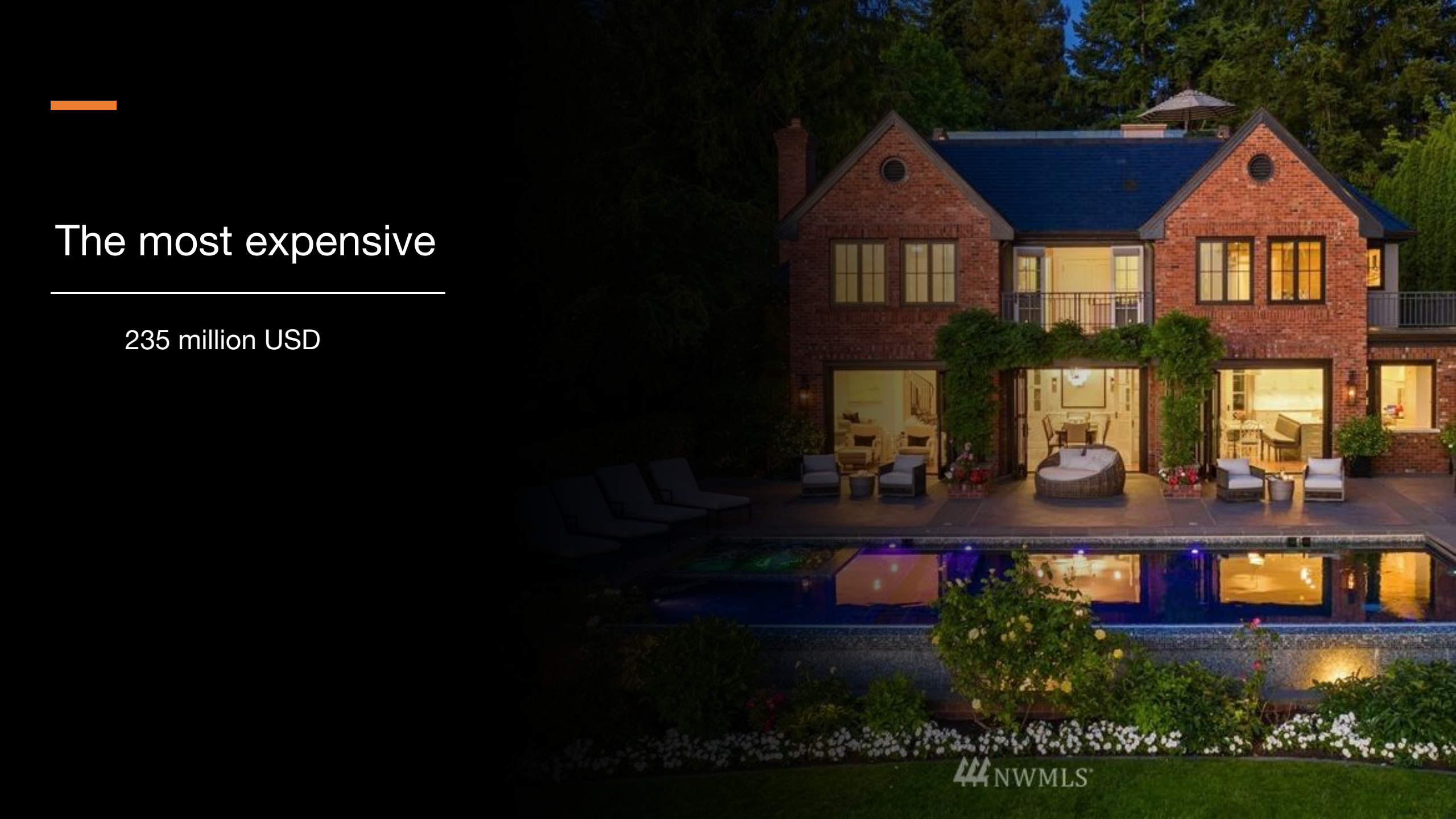
Data

Home Features I used for this project

-Price of each home sold
-Number of bedrooms
-Number of bathrooms
-Square footage of the interior living space
-Waterfront
-Greenbelt
-Condition
-Heat source
-Year Build
-The year of the house’s last renovation
-Address

The most expensive

235 million USD



The cheapest

40 thousand USD

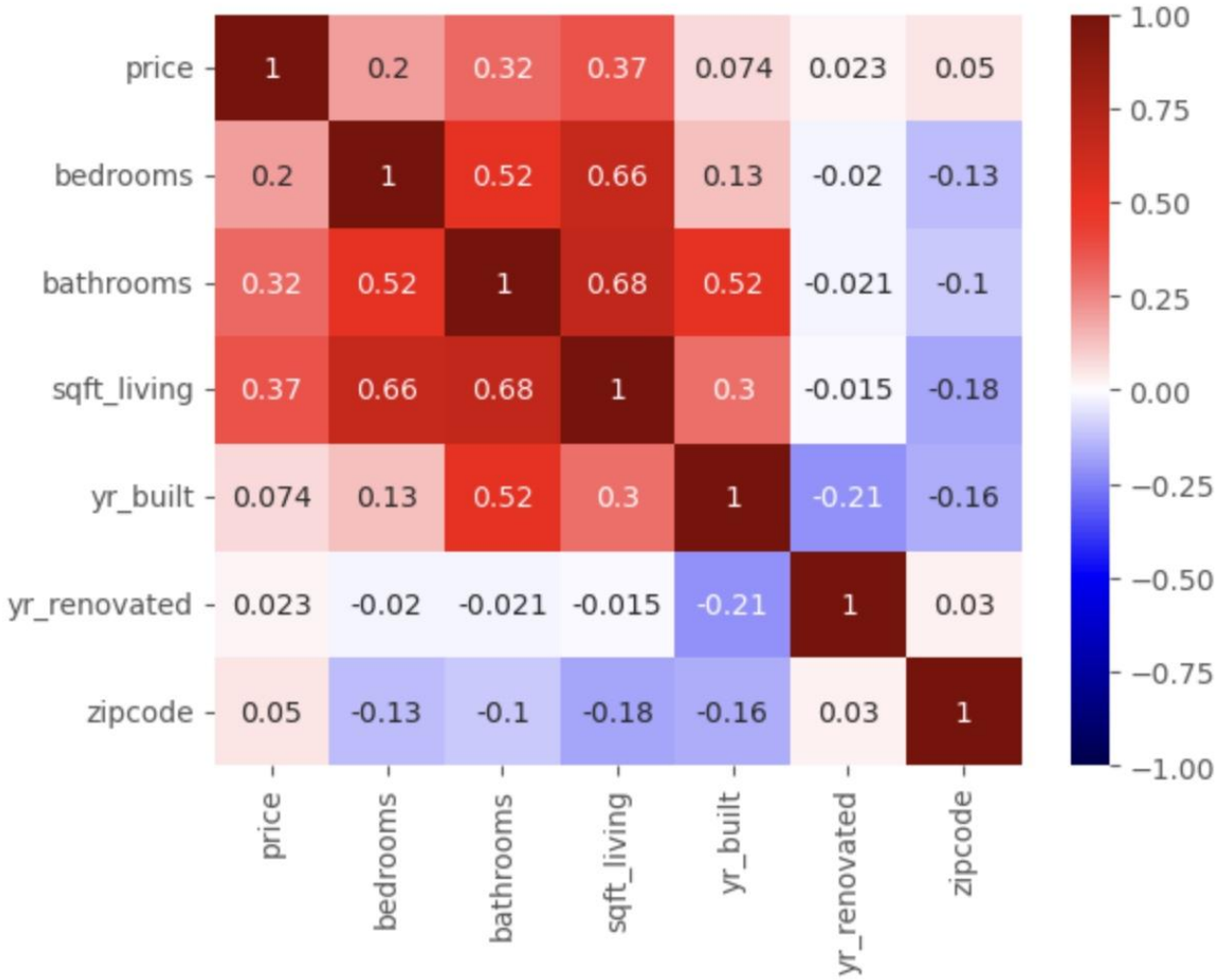
The cheapest houses
for
3 bedrooms 2 bathrooms

Federal Way
Renton

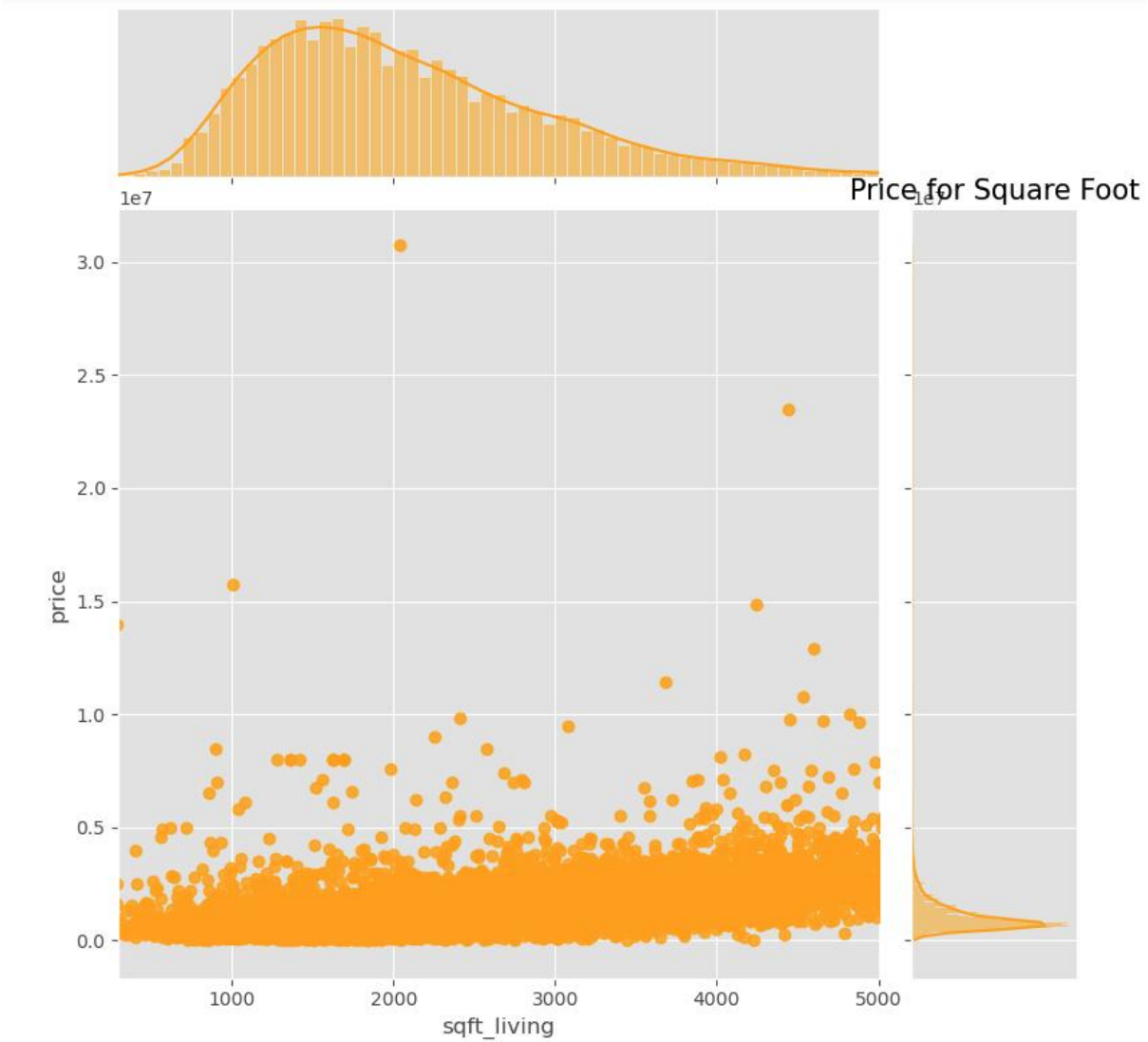
Expensive Houses
for
3 bedrooms 2 bathrooms

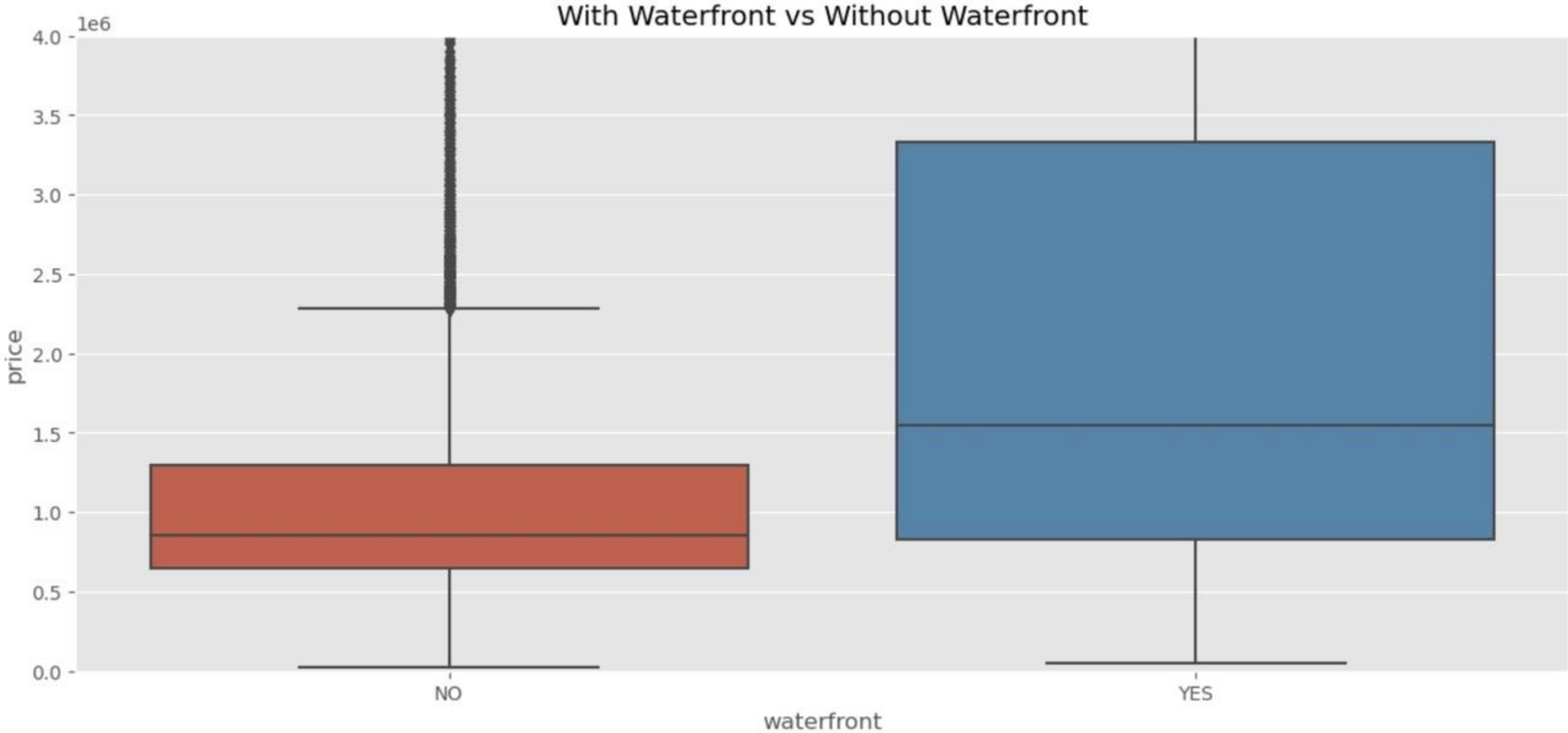
-Seattle

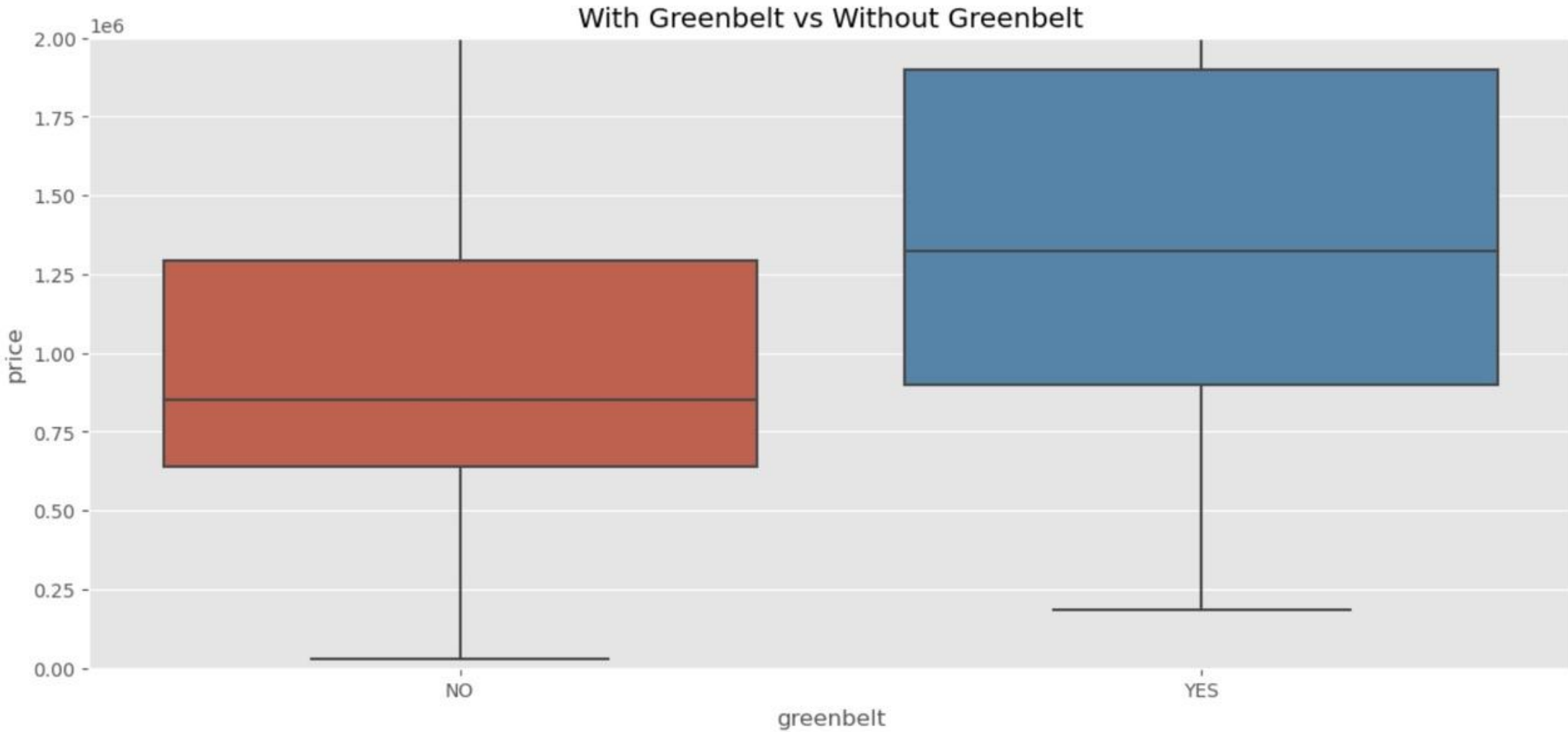
During the exploratory analysis I saw a good correlation between square footage and price

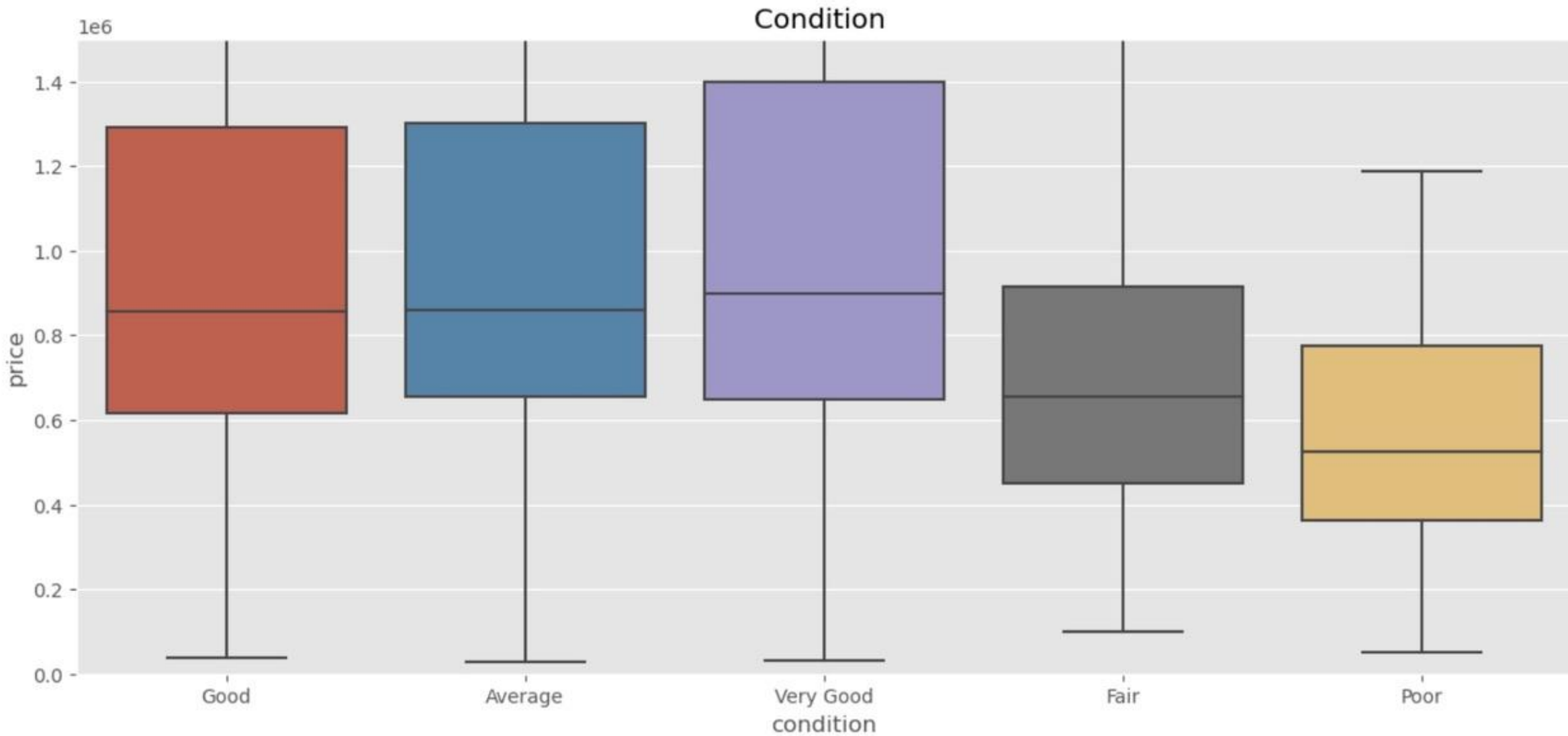


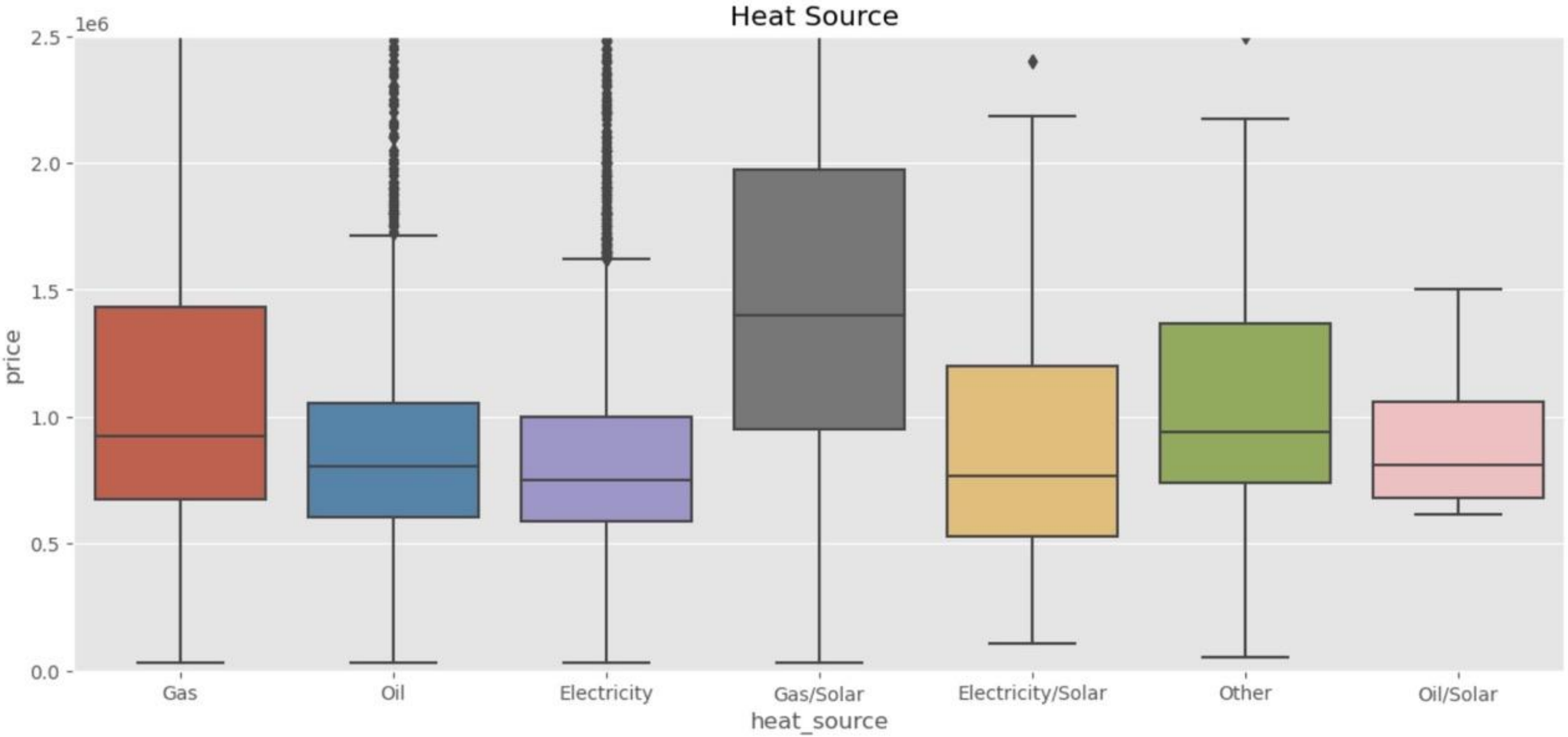
The estimated change in price for every sqft and came up with the price of 560 USD.



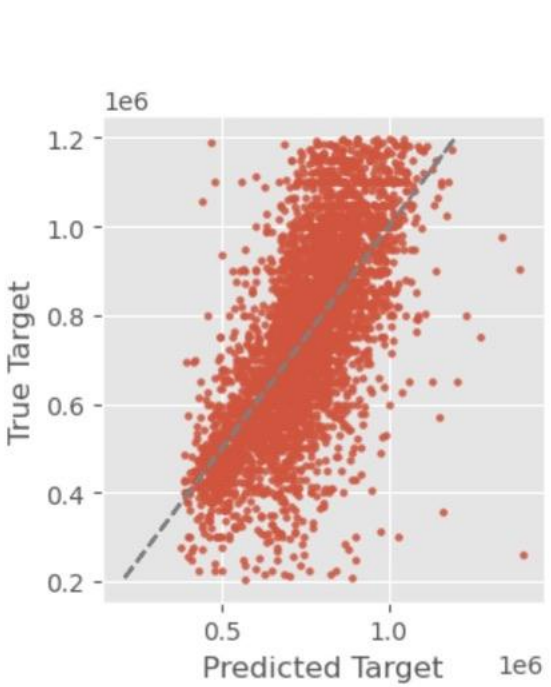

With Waterfront vs Without Waterfront
1e6
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
price
NO
YES
waterfront

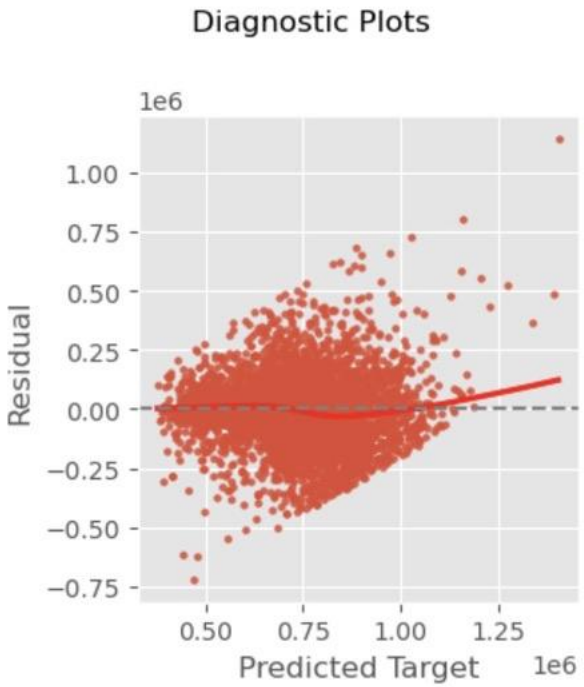

With Greenbelt vs Without Greenbelt
1e6
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
price
NO
YES
greenbelt

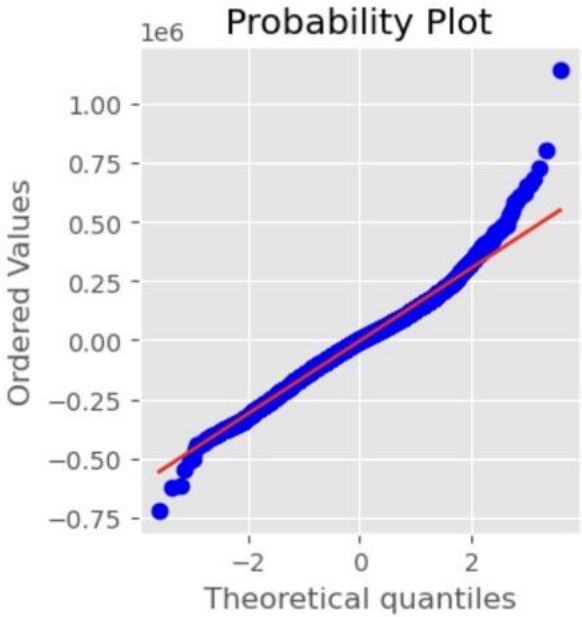

Condition
1e6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
price
Good
Average
Very Good
Fair
Poor
condition


Heat Source
1e6
2.5
2.0
1.5
1.0
0.5
0.0
price
Gas
Oil
Electricity
Gas/Solar
Electricity/Solar
Other
Oil/Solar
heat_source

Linear Regressions



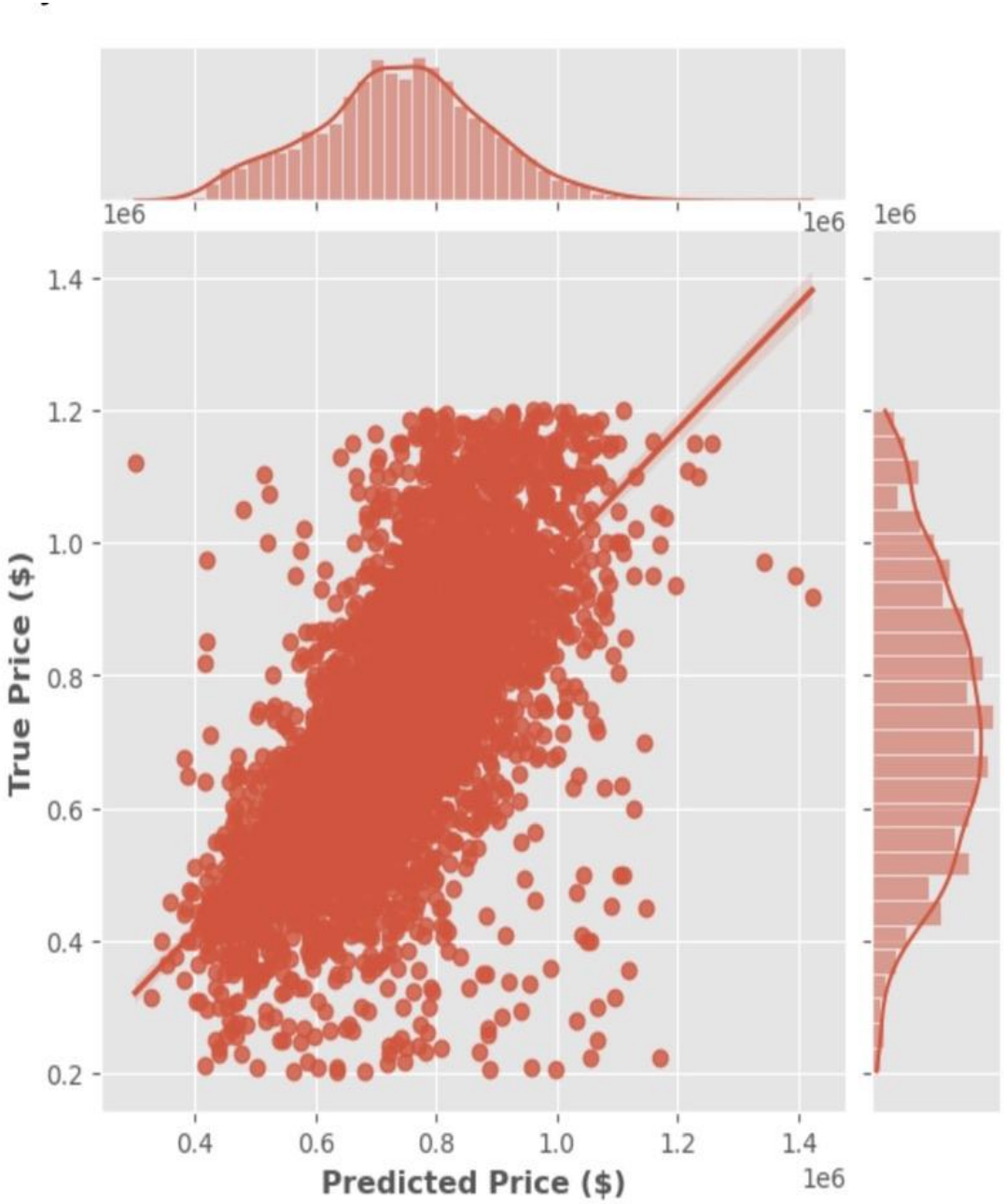




Ridge Model - Reduces effect of colinear features

R2 Adjusted = 0.46

MAE(mean absolute error) – $116,866



City	Mean variance
Osceola	1360.2
Renton	1878.3
Longmont	-1276.1

Recommendation

I would recommend to my client to buy a house with 3 bedrooms 2 bathrooms in Renton city.

1350 square feet

243thousand USD

Next Steps

Schools in the area

Property tax

Email prokopivanastasiia@gmail.com
Linkedin https://www.linkedin.com/in/anastasiia-leskiv-684b4b250/

Thank You