



King County Housing Data Analysis

Nicole Michaud
June, 2023

Project Goals

- What pre-existing features of houses are predictive of higher sale prices so that house flippers can buy houses to remodel that will have the best chances of returning higher profits?





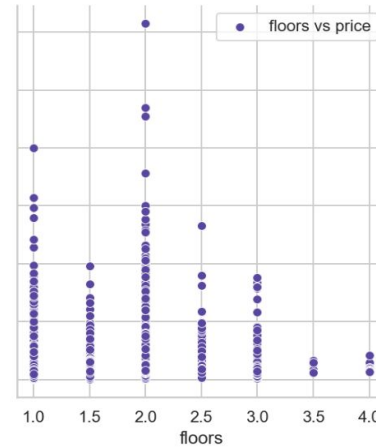
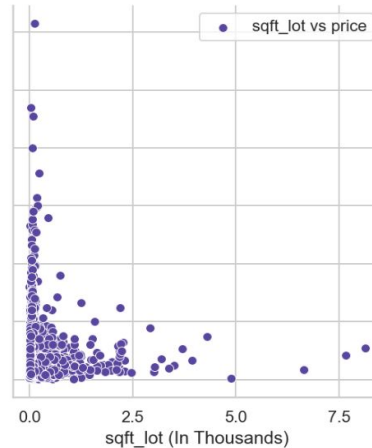
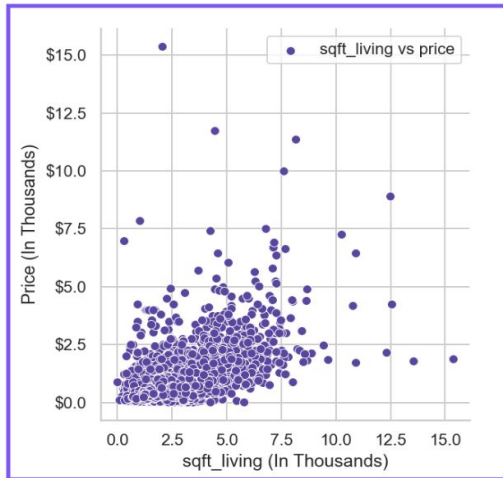
Data Understanding

The features from the Kings County housing dataset that I took into consideration include:

- Square footage of the living area
- Square footage of the lot
- Number of floors
- If house is on a waterfront
- If house is adjacent to a green belt
- If house has traffic noise or other nuisances
- The quality of the view of the house

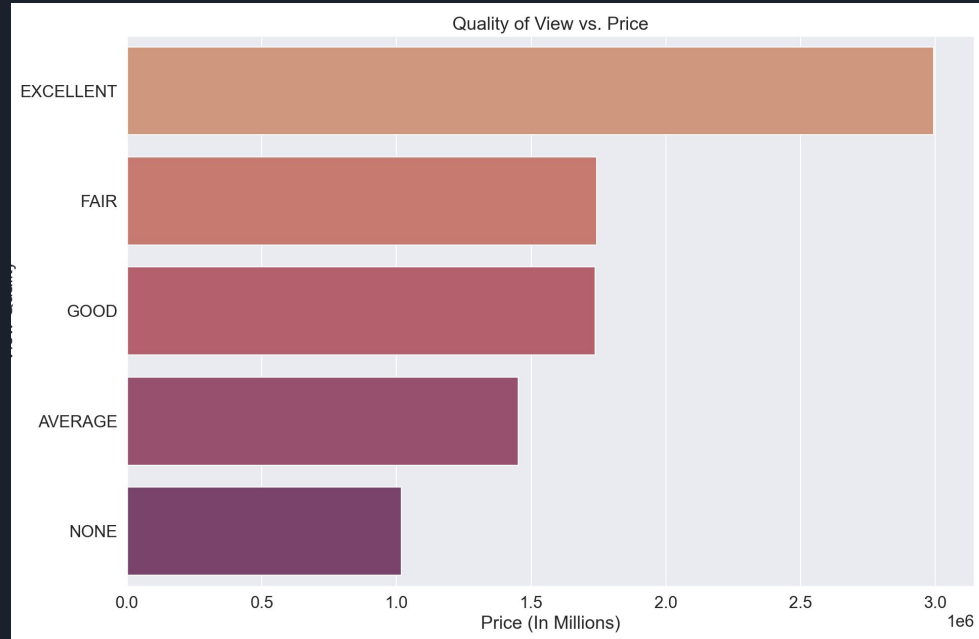
Methods & Data Analysis

- By looking at which variables within the data were most highly correlated with price, I determined that the variable that had the strongest relationship with price was the square feet of living area in the home



Methods & Data Analysis

- I created models for each of the four relevant categorical variables– waterfront, greenway, nuisance, and view– and of those found that view was the best predictor of price.





Results

- The best predictor of price was found to be a combination of the house's square feet of living space and the quality of the view the house has.
- The model with these two variables tells us:
 - For each **increase by 1 square foot** of living space, we expect to see an increase in price of about **\$521**.
 - Compared to a house with an **average** view, for a house with an **excellent** view we see an associated increase in price of about of about **\$1,196,00**.
 - For a house with **no view**, we see an associated decrease in price of about **\$182,400**.
- This model is **estimated to be off by** about \$388,647 on a given prediction.

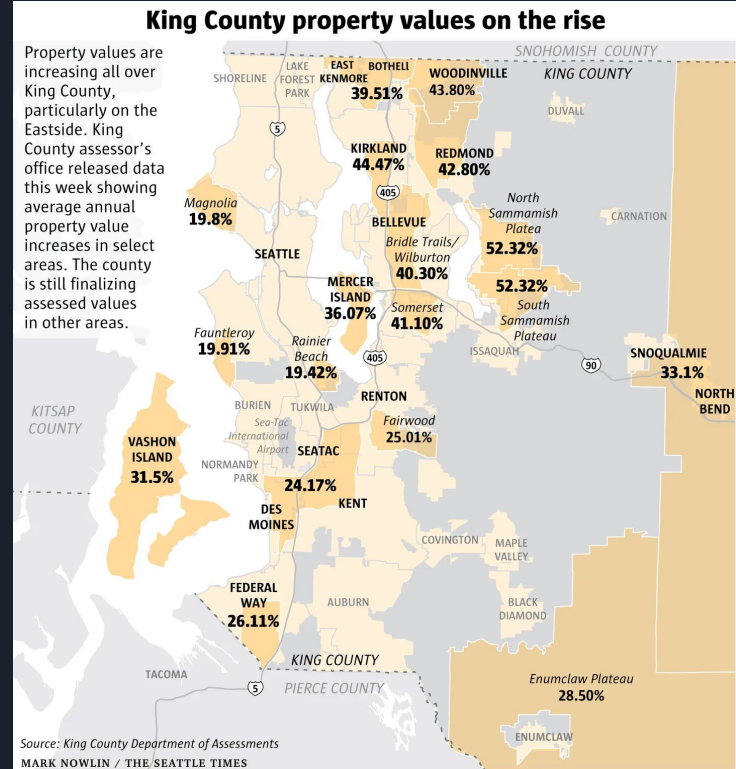


Conclusion

- Two factors that to look for when deciding which houses to buy in order to flip and make the highest possible profits are square feet of living space and quality of view.
- However, this model is **unreliable** for houses that are not within the price range of \$684,205.75 to \$2,901,227.43.

Next Steps

- Further analysis should be conducted on **more datasets** that include **more features** that are likely to be related to sale price, such as housing subdivision.





Thank You!

- Any Questions?
- Contact Info:
 - Nicole Michaud
 - LinkedIn: www.linkedin.com/in/nicole-michaud2

