

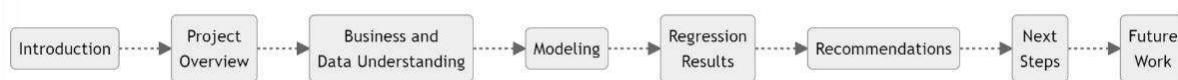


BUSINESS ANALYSIS OF HOUSE SALES IN KING COUNTY, WASHINGTON

Phase 2 Group 24 Project Presentation
September 8, 2023

Group 24 Members

1. *Samuel Gichanga*
2. *Leonard Gachimu*
3. *Moses Thiong'o*
4. *Evangeline Ngunjiri*



Introduction

King County highlights:

- The most populous county in Washington.
- Population estimates as of July 1, 2022 was **2,266,789** while the Housing Units were **1,004,796**.
- The County is the 13th most populous in the USA.



Project Overview

To build a data analytics model for Team Howlett Realtors, a real estate agency that helps homeowners buy and/or sell homes in King County, Washington.

Five major methods were utilized in this project to drive our analysis



Methods used in the analysis

Business Problem

Our team seeks to help Team Howlett Real Estate Agents to make data-driven pricing strategies for homes in King County.

Business Understanding

Major Factors that Influence the Price of a Home in King County:

1. Location of the house
2. Age of the House
3. Condition

4. The number of bedrooms and bathrooms etc.

Business Objectives

1. To build a data analytics model for Team Howlett Real Estate Agents
2. To help realize the best value to a home buyer and optimal returns to the home seller.
3. To provide insights on how weather seasons, the grade of the house, and the number of bedrooms may affect sales performance.
4. To find out whether there is a variation in sales volume and mean sale prices across different locations in King County.

Study Questions

What is the:

1. Relationship between weather season and sales performance?
2. Relationship between the number of bedrooms and sales performance?
3. Best weather season to buy a house?
4. Best performing house grades?
5. Optimal range of a house grade for different budgets?
6. Optimal range of bedrooms for different budgets?
7. What are the variables with the highest influence on the sale price of a house?
8. How are the house prices and house sale volumes distributed around the county?

Data Overview

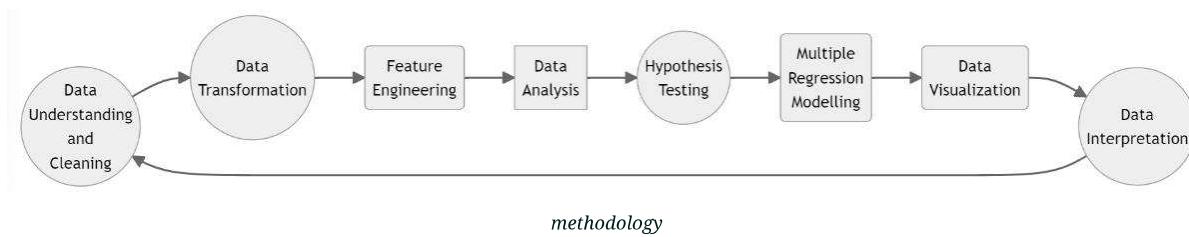
King County Housing Dataset contains information about house prices between May 2014 to May 2015.



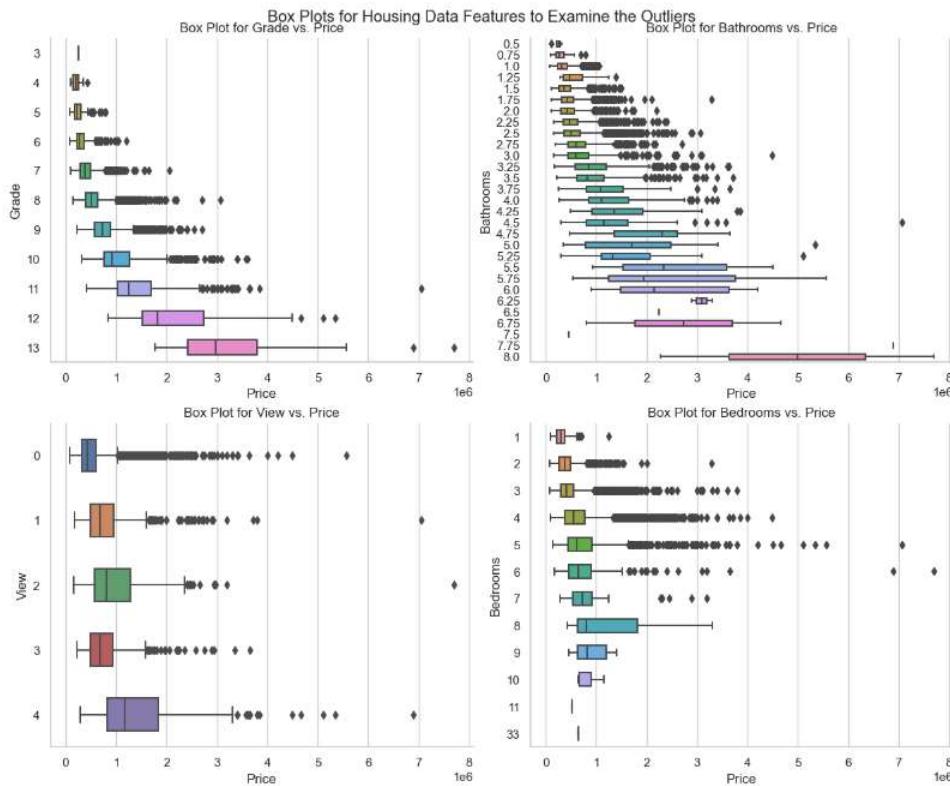
The Dataset has 21,597 rows and 21 columns.

Some of the columns like the date were converted to relevant formats to allow analysis

Methodology



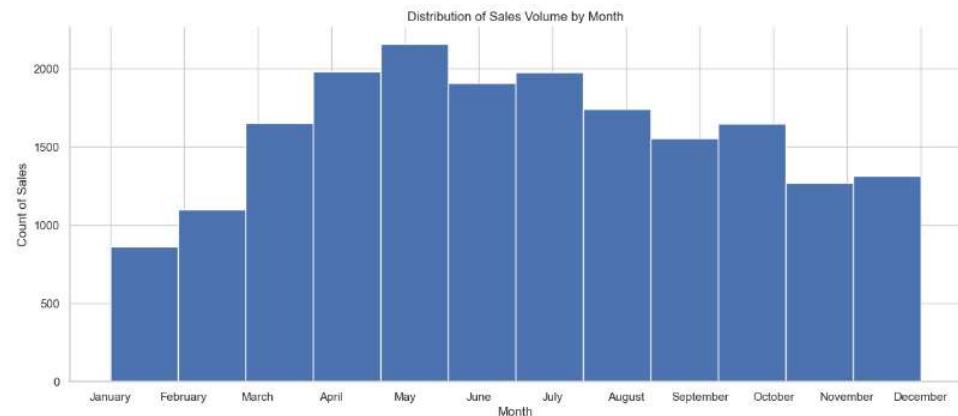
The boxplot for Bedrooms vs. Price shows a house with 33 bedrooms. That was dropped in our analysis.



Results

1. The relationship between weather season and sales performance

There is a significant variation in the volume of sales across the months in year.



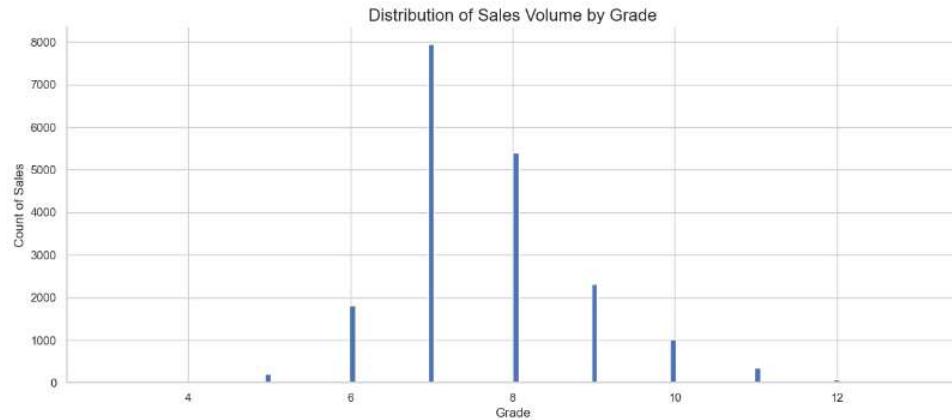
2. Distribution of mean sale prices over the months

It shows that there are significant differences in mean sale prices over the months.



3. Relationship between house grade and sales performance

- Grade 7 has a peak sales volume of about 8,000 houses.
- The lower grades of 3, and 4, and the higher grades of 11 and 12 each have sales volumes around 100 or fewer.



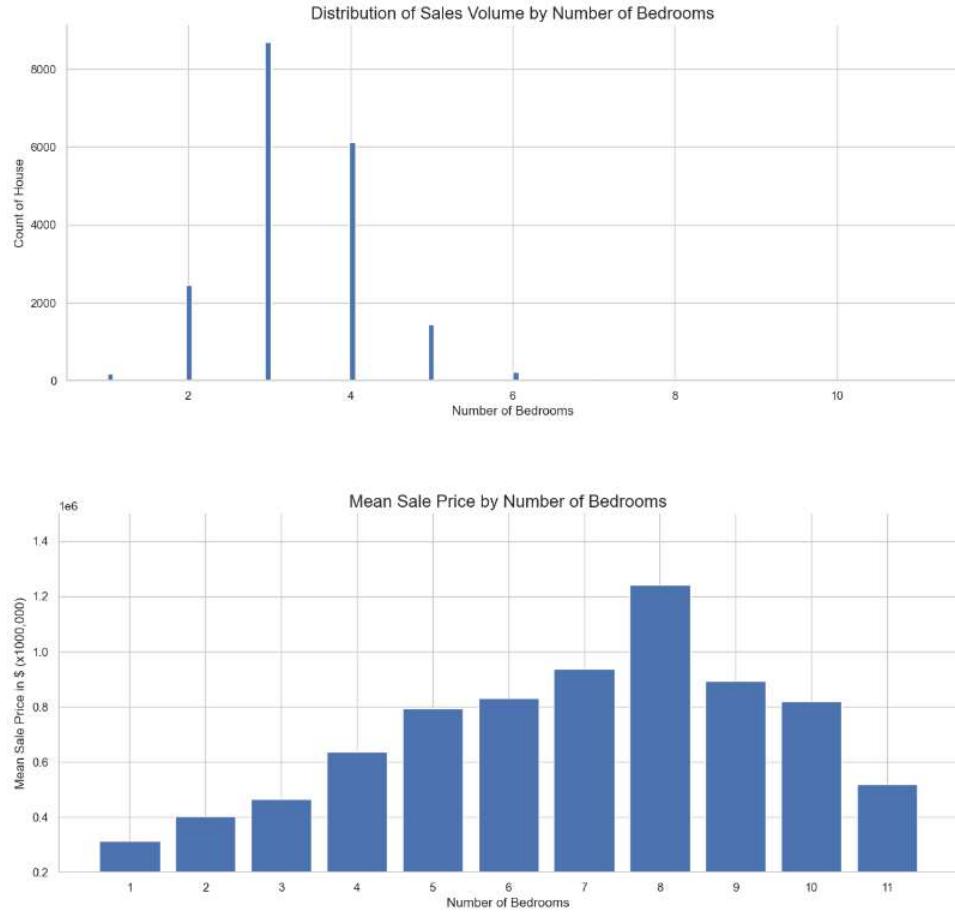
- Between grades 3 and 8, the mean price is between **\$250,000 to \$500,000**.
- From grade 9 to 13, the mean sale price rises from around **\$600,000** to over **\$3,500,000**.



4. The relationship between number of bedrooms and sales performance

Houses with 3 bedrooms have a sales peak of above **8,000 houses**.

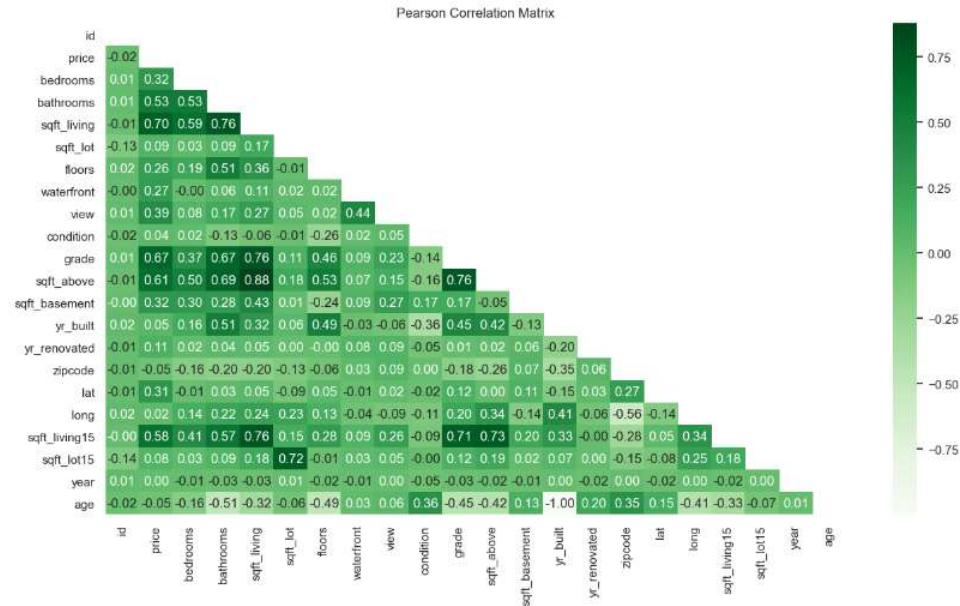
BUSINESS ANALYSIS OF HOUSE SALES IN KING COUNTY, WASHINGTON

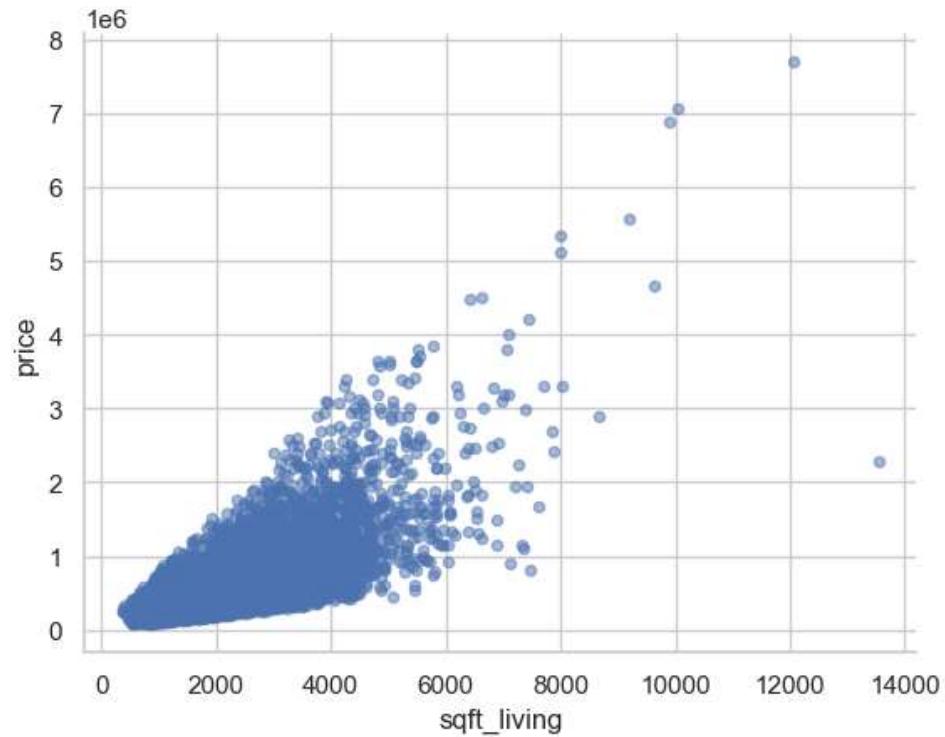


The distribution of mean sale price by number of bedrooms

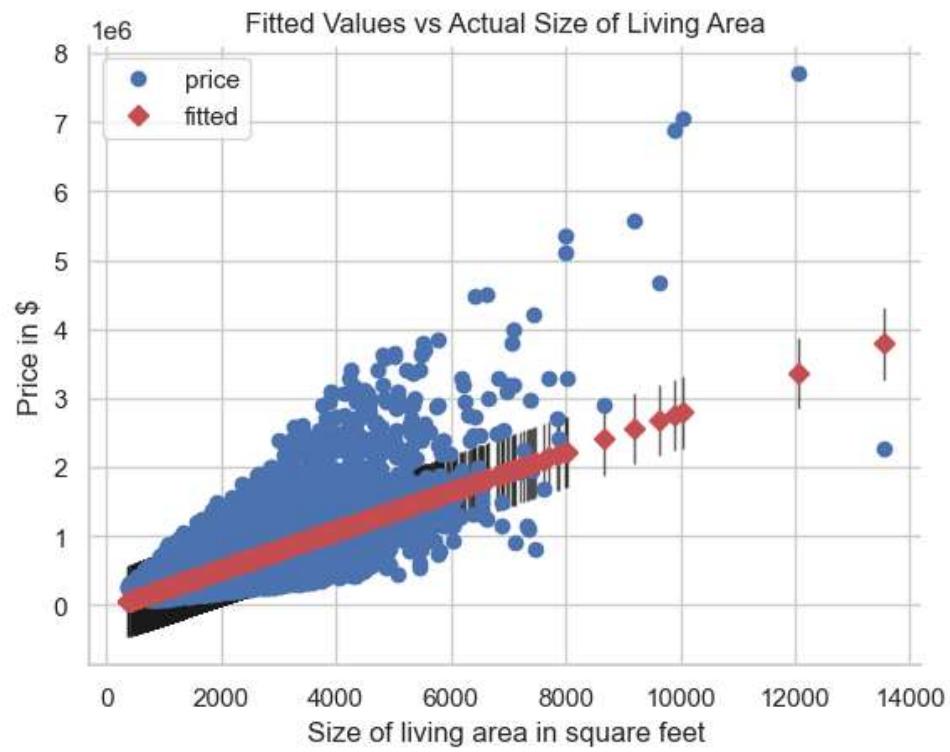
5. Pearson Correlation Matrix

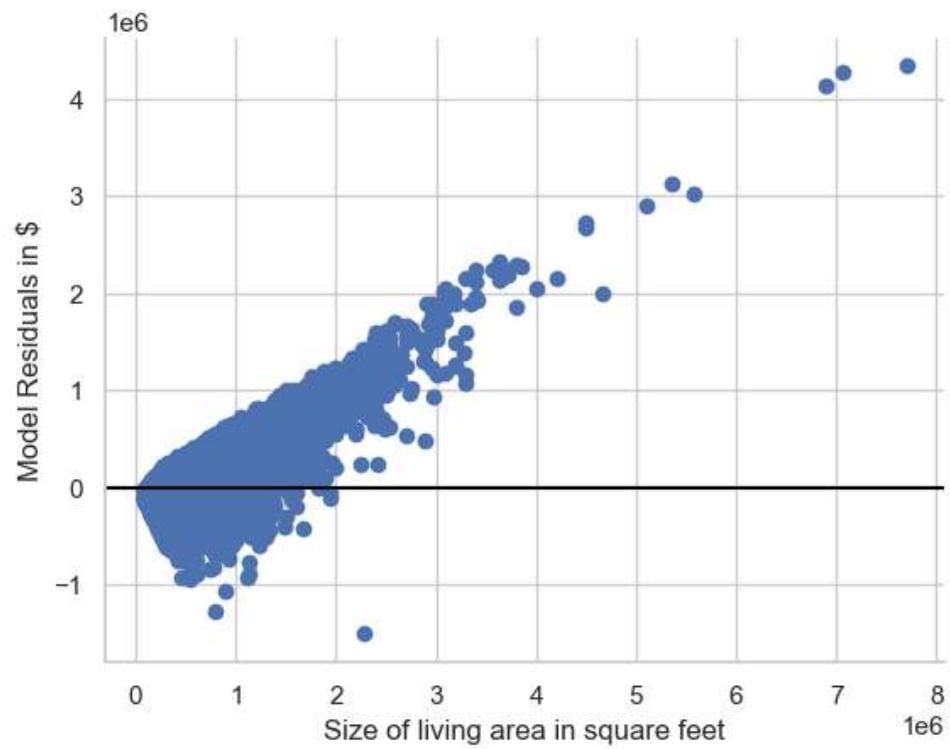
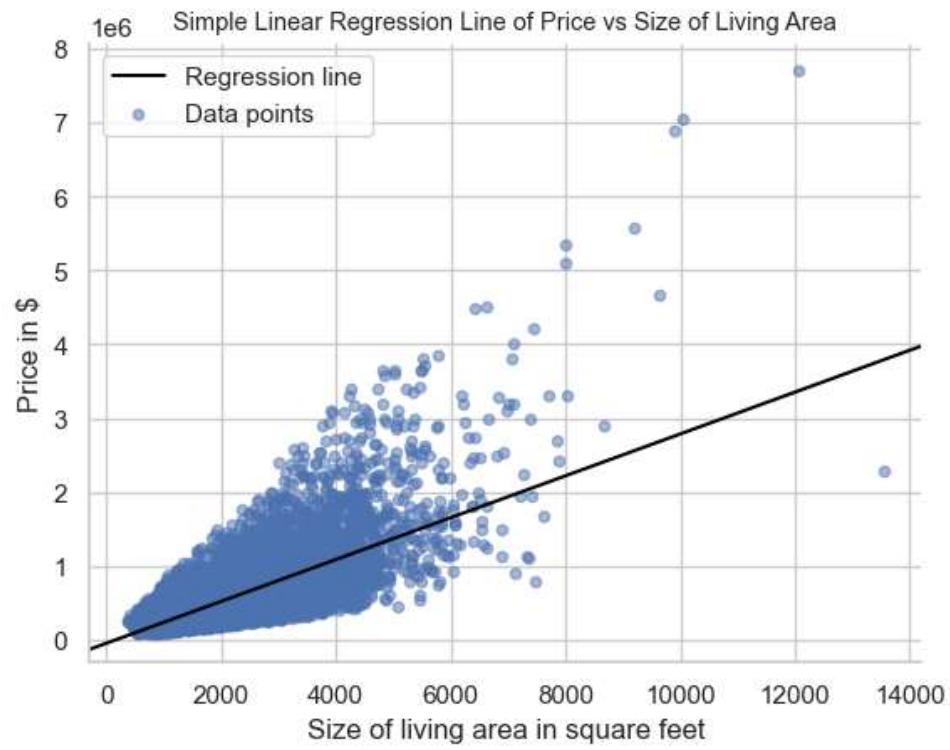
The correlation coefficients show that the size of the living area in square feet ('sqft_living' column) has the highest correlation with price.

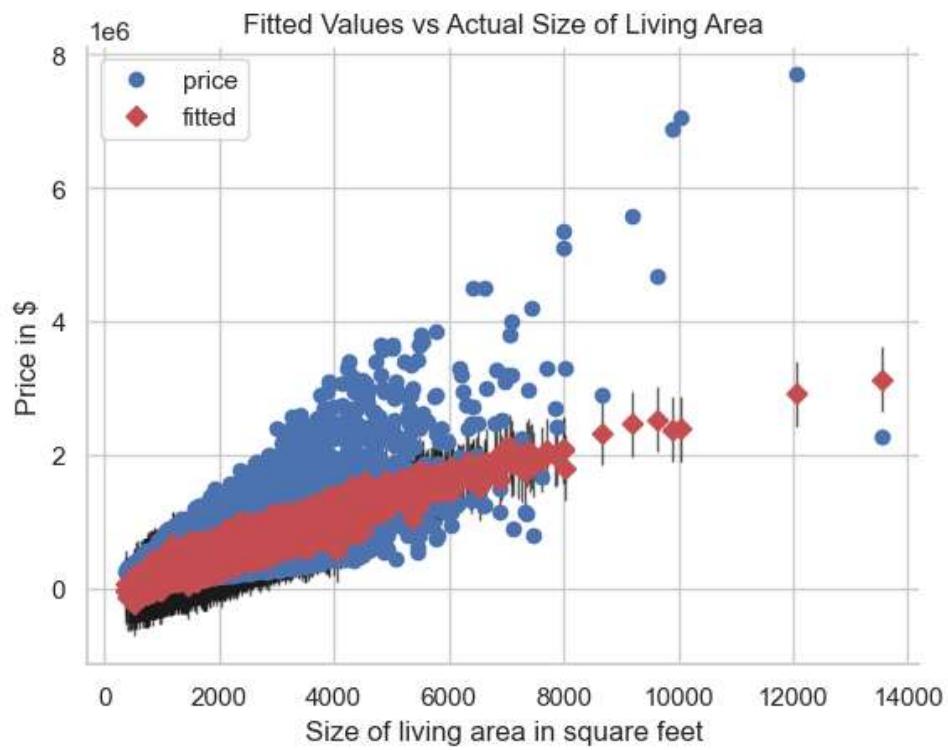
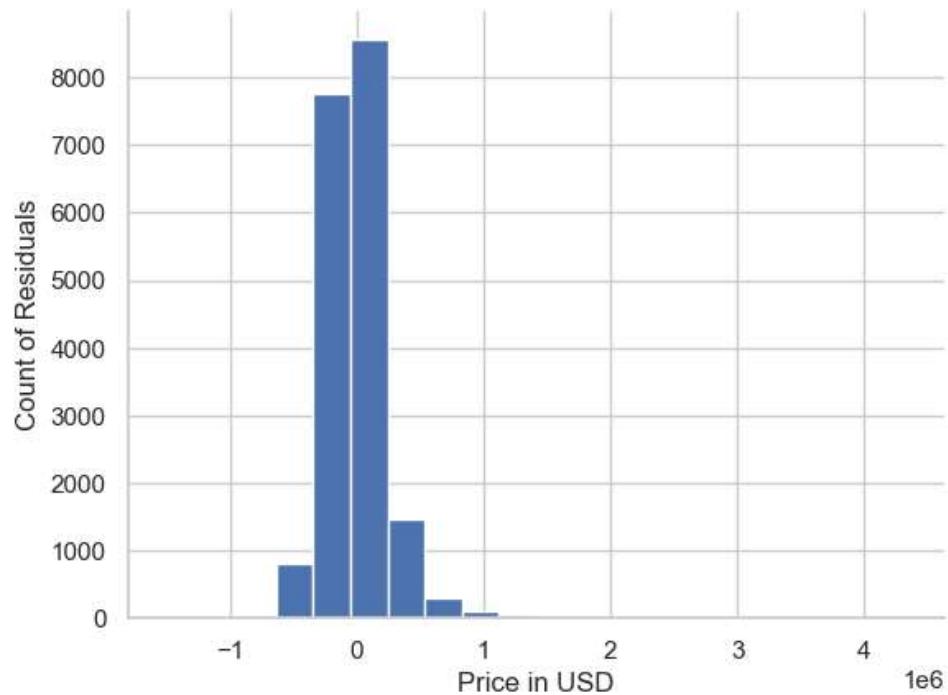




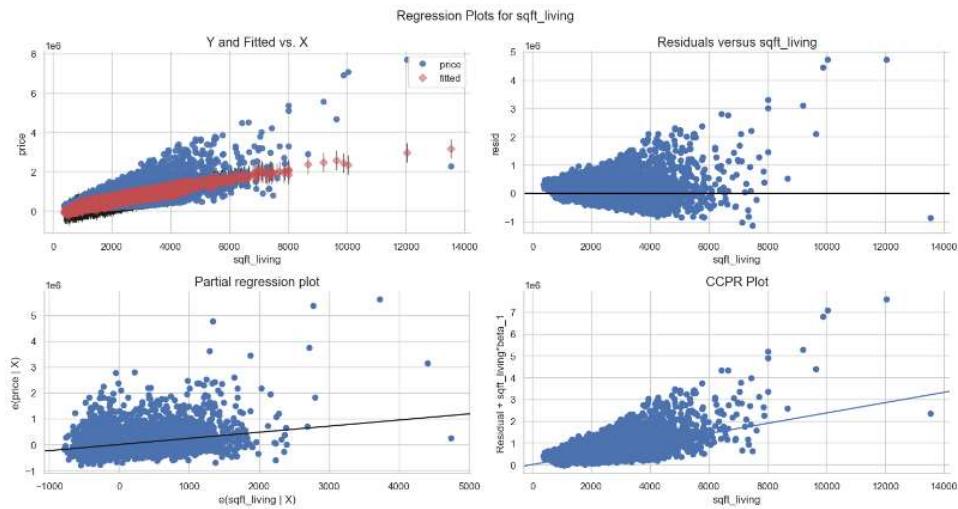
A scatter plot of living area size vs. price





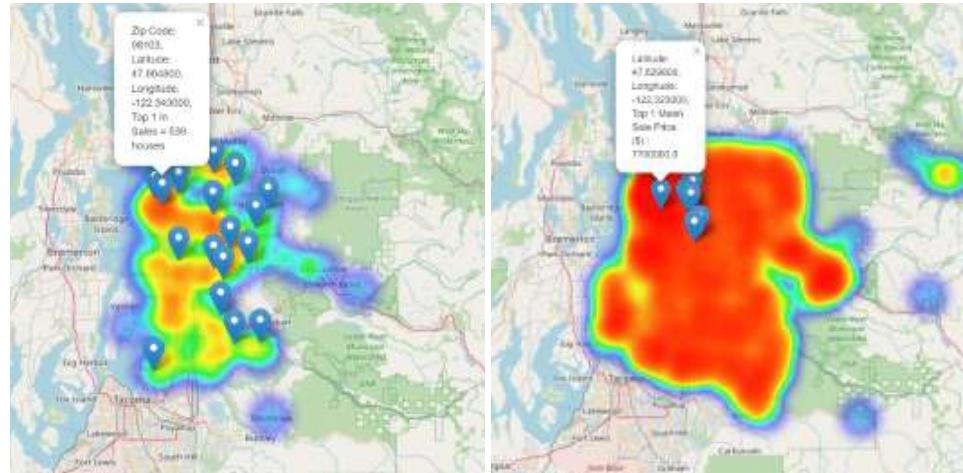


6. Regression modelling of the set of variables with the highest influence on the sale price of a house.



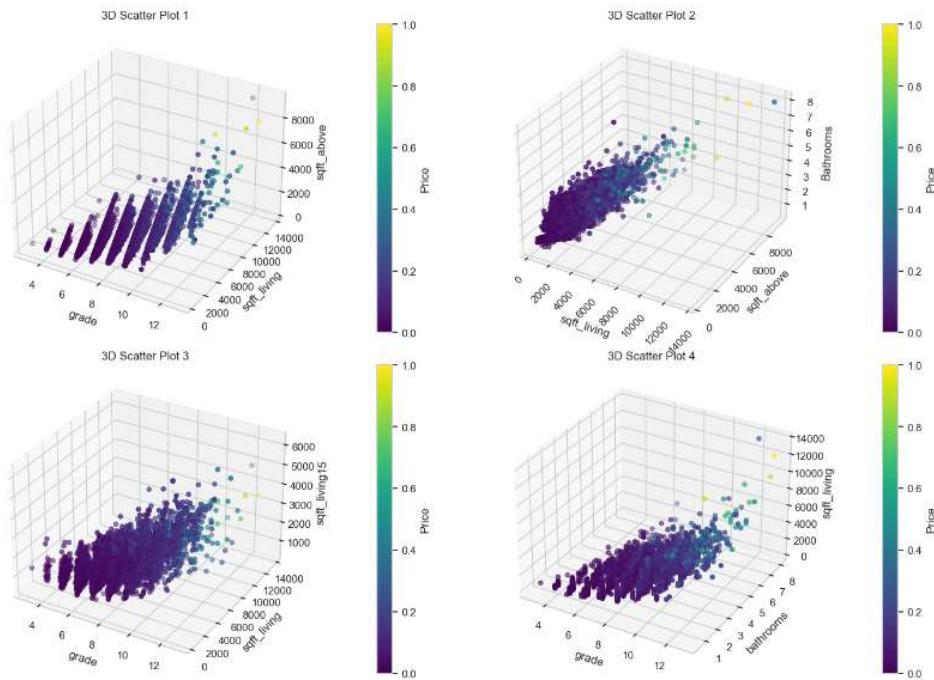
7. Distribution of house prices and house sale volumes around the county

- Most of the top 20 locations in the count of sales are in the northwestern region of Seattle.
- Most of the top 10 mean sale prices came from the northwestern region of Seattle.



8. 3D Scatter Plots

- Different combinations of the most influential predictor variables show a high correlation among them.



Conclusion

- i.) The size of the interior living area, the grade of the house, the size of the house excluding the basement, the average size of interior living space for the closest 15 houses, and the number of bathrooms, are the 5 most influential factors affecting price.
- ii.) Winter and Fall are off-peak seasons while Spring and Summer are peak seasons. May has the highest volume at **above 2,000 houses**. April has the highest mean sale price of **above \$550,000**.
- iii.) Grade 7 houses have a peak sales volume of about **8,000 houses**.
- iv.) The median number of 3 bedrooms has the highest sales volume of 8,000 houses. However, the peak mean price is **\$1,200,000** for 8 bedrooms. For houses with 3 bedrooms, the mean price is only **\$465,000**.
- v.) Most of the best-performing locations in terms of number of houses sold and the mean sale price are in the Seattle

region.

Recommendations for Team Howlett Realtors

- i.) Buyers and sellers should give higher consideration to the size of the interior living area, the grade of the house, the size of the house excluding the basement, the average size of interior living space for the closest 15 houses, and the number of bathrooms.
- ii.) A buyer seeking a significant discount should buy in Winter or Fall. A seller seeking the highest margin should sell in Summer.
- iii.) Buyers and sellers looking for or selling low-budget houses of around \$500,000 or less should consider houses between grades 3 and 8. Those seeking to spend or earn more than \$500,000 should go for grades 9 to 13.
- iv.) Buyers and sellers looking for low-budget houses should also consider between 1 and 4 bedrooms. Those seeking to spend or earn more than \$800,000 should go for houses with between 5 and 10 bedrooms.
- v.) If a buyer wants a wider variety of high-grade homes on sale, they should search in Seattle region. However, this is also the region with the highest mean sale price.