King County Houses Multiple Regression Analysis

Phase 2 Project: April 2023

Business Overview

- The housing demand in Washington State is high, with Seattle being the leading city in terms of population.
- As such, Tella Real Estate Agency, based in King County, has undertaken a research to find out the best performing metrics when it comes to house sale prices determination.
- By identifying the key factors that affect the sale prices of houses, the agency can develop more effective marketing strategies, help sellers target the right buyers, and make better investment decisions.

Business Problem

- The agency has a dataset of home features collected from the various house sales in King County.

- The agency wants to understand the relationship between these features and the sale prices of houses in King County, and particularly, which features most affect the prices.

Objectives

- To understand the relationships between the various features and the house sale prices.
- To determine the peak house sale season.
- To build a multiple linear regression model that identifies the most important factors that influence the sale price of a house.
- To use the model to gain insights into the home sales market and improve decision-making processes.
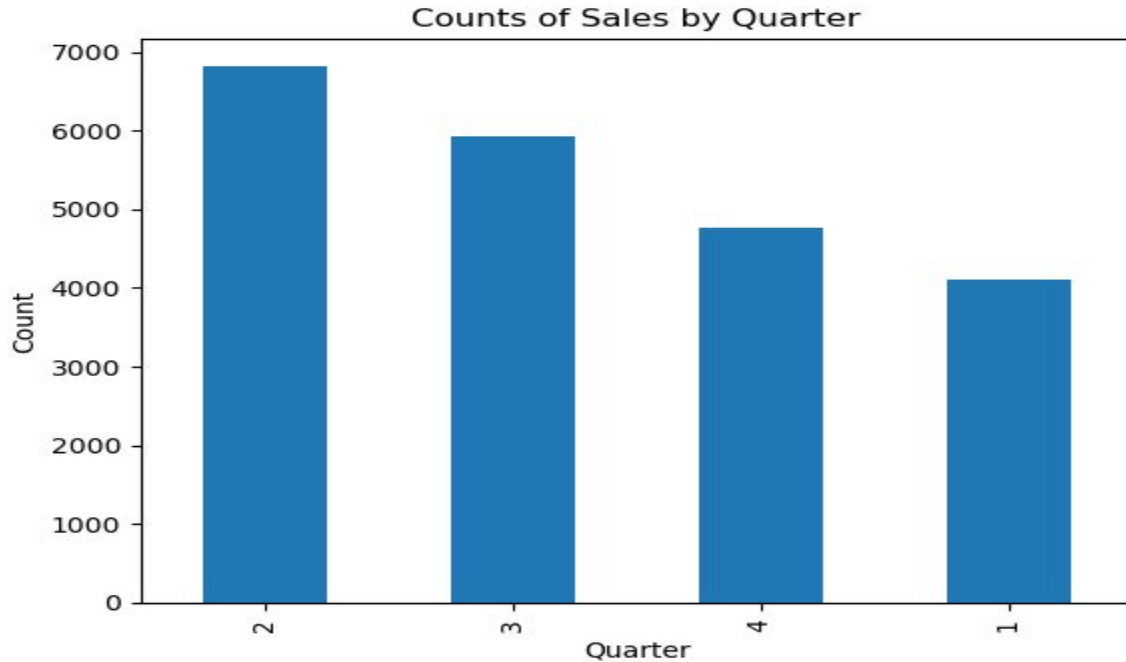
Data Understanding

- We used King County Houses data for our analysis.

- The dataset had 21 columns translating to 21 observed features, and 21,597 rows representing the number of house sales between May 2014 and May 2015.
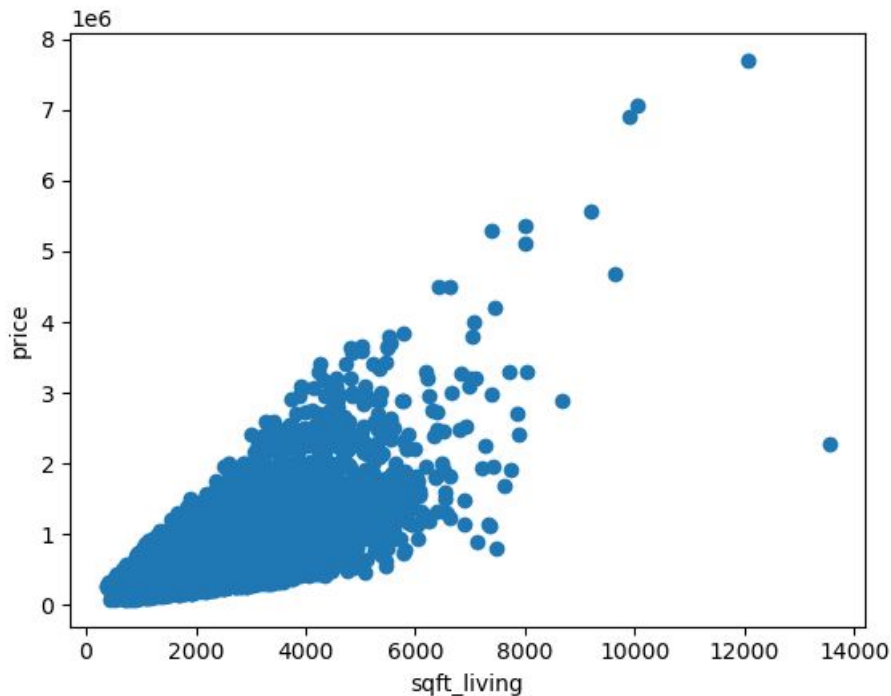
Price Distribution
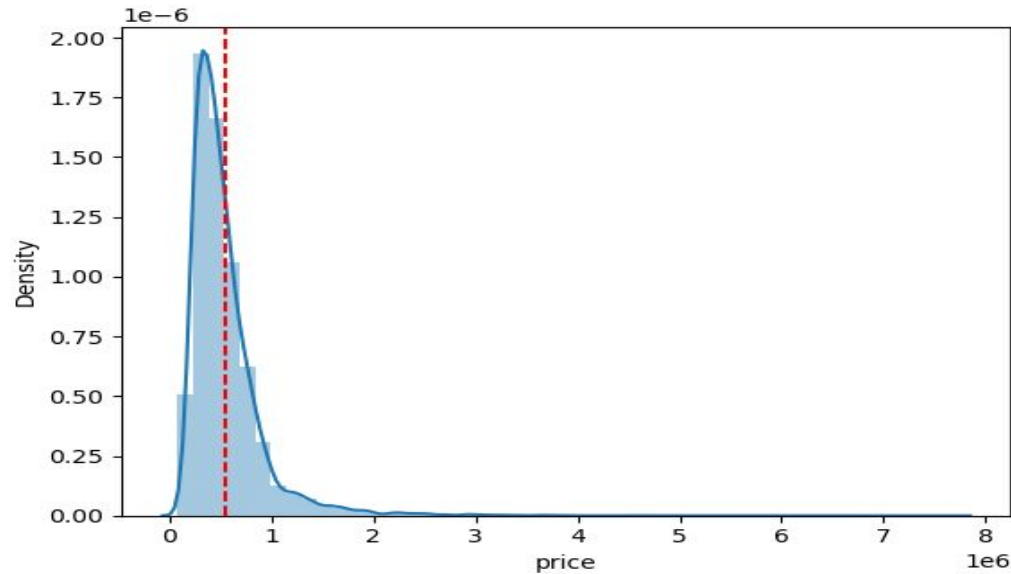
- This indicates that most houses are lowly priced

Peak House Sale Season

Counts of Sales by Quarter

● Most house sales are made in the spring season.

Baseline Model

- Our baseline model was a linear regression of price against square foot living.
- The model could predict 49.3% of the prices, so we went ahead to examine other models that would better predict price.

Best-performing model

- The best model explained 61.7% of house prices.

- The predictor variables in the model were:

  - Bedrooms, sqft_living, floors, grade, sqft_basement, yr_renovated, age, waterfront, view, condition, sqft_lot, sqft_lot15

Conclusion

- The variables that have a major influence on the price of the house are; square foot living, age of the house,good condition of the house,if the house is on a waterfront and has an excellent view.
- For those looking for economical housing options, it might be wise to consider sacrificing spacious living quarters or a scenic waterfront view.
- The variables that has the least influence on the price of the house are; grade, number of bedrooms, sqft lot, sqft basement and sqft lot 15.

Recommendation

- Renovate their house since this increases the value of the house.
- Ensure that the houses are in good condition before putting it into the market for sale.
- Increase square footage of living space.
- Put up their houses for sale in peak season-Spring.

Thank You

Contributors:

- Billy Mwangi
- Lynne Mutwiri
- Sharon Kimani
- Susan Kanyora
- Kellen Kinya
- Derrick Wekesa

Github: https://github.com/billymwangi/project-phase-2