Selling Sunset Real Estate

Kings CountY Home Sales

This analysis examines homes in Kings County, Washington. The business case was to determine the home attributes that will affect the price and in turn guarantee sales.

Outline

Business Understanding

Data Understanding

Data Analysis

Model Results and Findings

Conclusion, Limitations and Recommendations

Business Underdanding

Business Problem

The real estate company is trying to determine what types of houses to focus on selling. Statistical analysis will be done, specifically Linear Regression modeling, as this helps gain an in-depth understanding of a huge dataset by splitting the data into smaller samples in order to make sense of it.

Data Understanding

The dataset contained information about
different home attributes like:

- The number of bedrooms and bathrooms
- Square footage of the living space and number of floors
- The condition of the house
- Whether the house has a waterfront and a view
- The grade of the house which is the quality of construction
- How old the house is and the year of renovation
- Location

Data Understanding

The dataset contained information about 21597 homes. These homes have different home attributes like:

- The number of bedrooms and bathrooms
- Square footage of the living space and number of floors
- The condition of the house
- Whether the house has a waterfront and a view
- The grade of the house which is the quality of construction
- How old the house is and the year of renovation
- Location

Data Analysis

The dataset contained information about 21597 homes. These homes have different home attributes like:

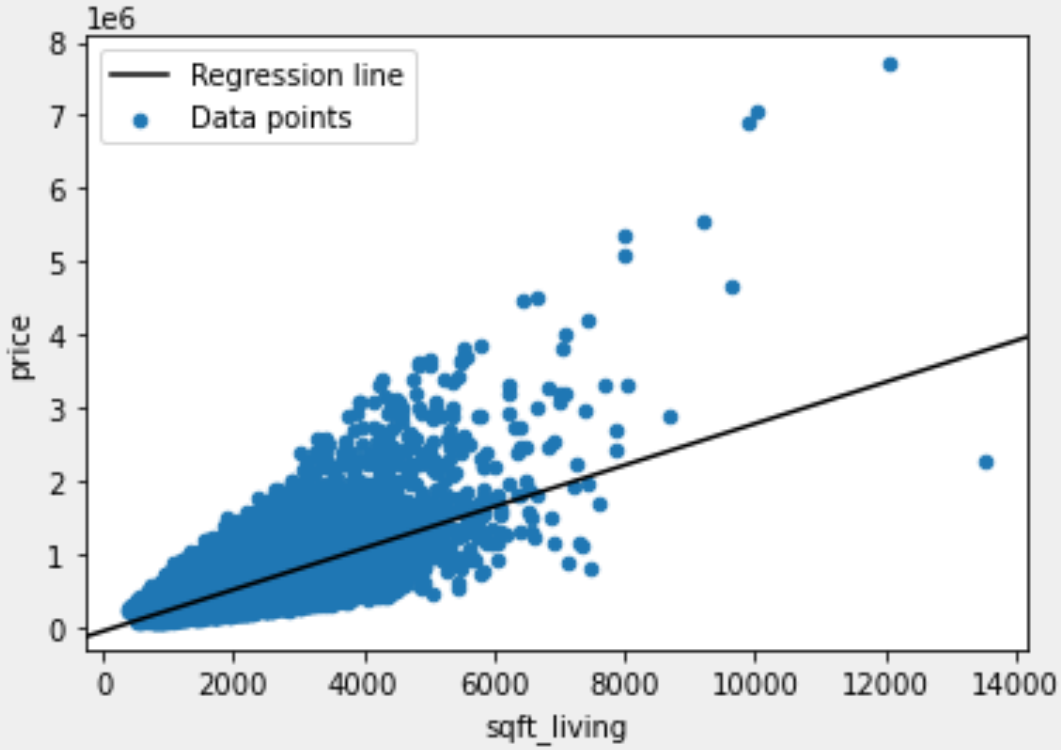
- The number of bedrooms and bathrooms
- Square footage of the living space and number of floors
- The condition of the house
- Whether the house has a waterfront and a view
- The grade of the house which is the quality of construction
- How old the house is and the year of renovation
- Location

Data Analysis

Modeling

The use of Ordinary Least Regression to do analysis.
For the first model, the simple linear regression model, only one variable was used, the one the highest correlation to price which was Square footage

Baseline model



Data Analysis

Modeling, cont

Baseline model

The model explains 49.6% proportion of the variation which was low and why we need a multiple linear regression adding more variables that will affect the price.

Model Results

- A low R-squared insists that more models should be done with more independent variables.
- The final model included all the variables except the date and location and had an accuracy of 68%, and lesser RMSE than the baseline model. This means that the model had better accuracy and less off than the baseline model.

Conclusion

- : Increasing the square footage is associated with a higher sale price.
- Generally a positive improvement of any of these attributes an increase in sales.

Recommendations

- An increase in the square footage of living space, the house is expected to sell for more.
- Maintenance to improve the condition, the expected sale price should go up.
- The agency should invest in a quality design of the house by the quality of contractors during renovation.

Limitations

- The model 's accuracy is still not stellar of 0.680
- A variation of 32% still cannot be explained.
- The square footage of living space had to be augmented to give a linear relationship with price.