King County Real Estate Services Housing Analysis

The industry at a glance

Summary

The goal of the project was to find a model of the data that would predict the price of the houses in King County as accurately as possible.

Ultimately, we achieved a model that explained 85% of the variations in price using the features in our data like Square feet of the house, number of bedrooms and so on.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

King County Real estate services wants to contract a construction company to build some houses in King County and the area of Seattle. It wants to know what features to focus on to build a variety of houses at different prices. We shall focus on:

- Generating a model that accurately predicts the House prices

Data

The data I used came from [Kaggle](). It contains data on House sales in King County.

Some of the features included Square feet of the house, view, grade, bathrooms, floors and so on. We used the features in our model to predict the price of houses in the area.
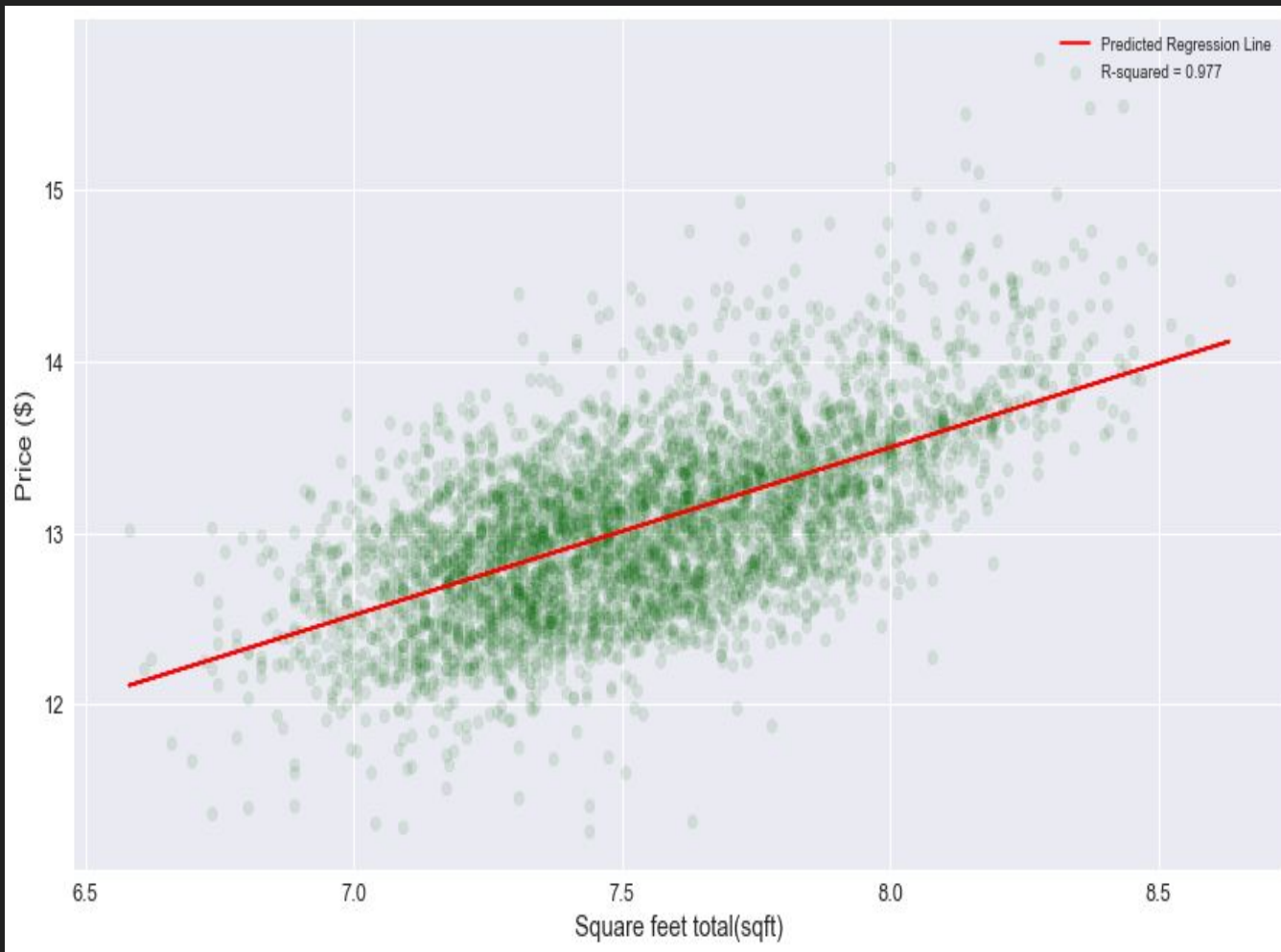
Methods

Before I could utilize the variables aforementioned, the data had to undergo a process of preparation to ensure that the insights I would generate would be as accurate as possible. This included:

- Dropping rows with empty values
- Making sure the data is in the correct format
- Creating variables necessary for analysis
- Normalizing the data and scaling it.
- Using linear regression to model the data
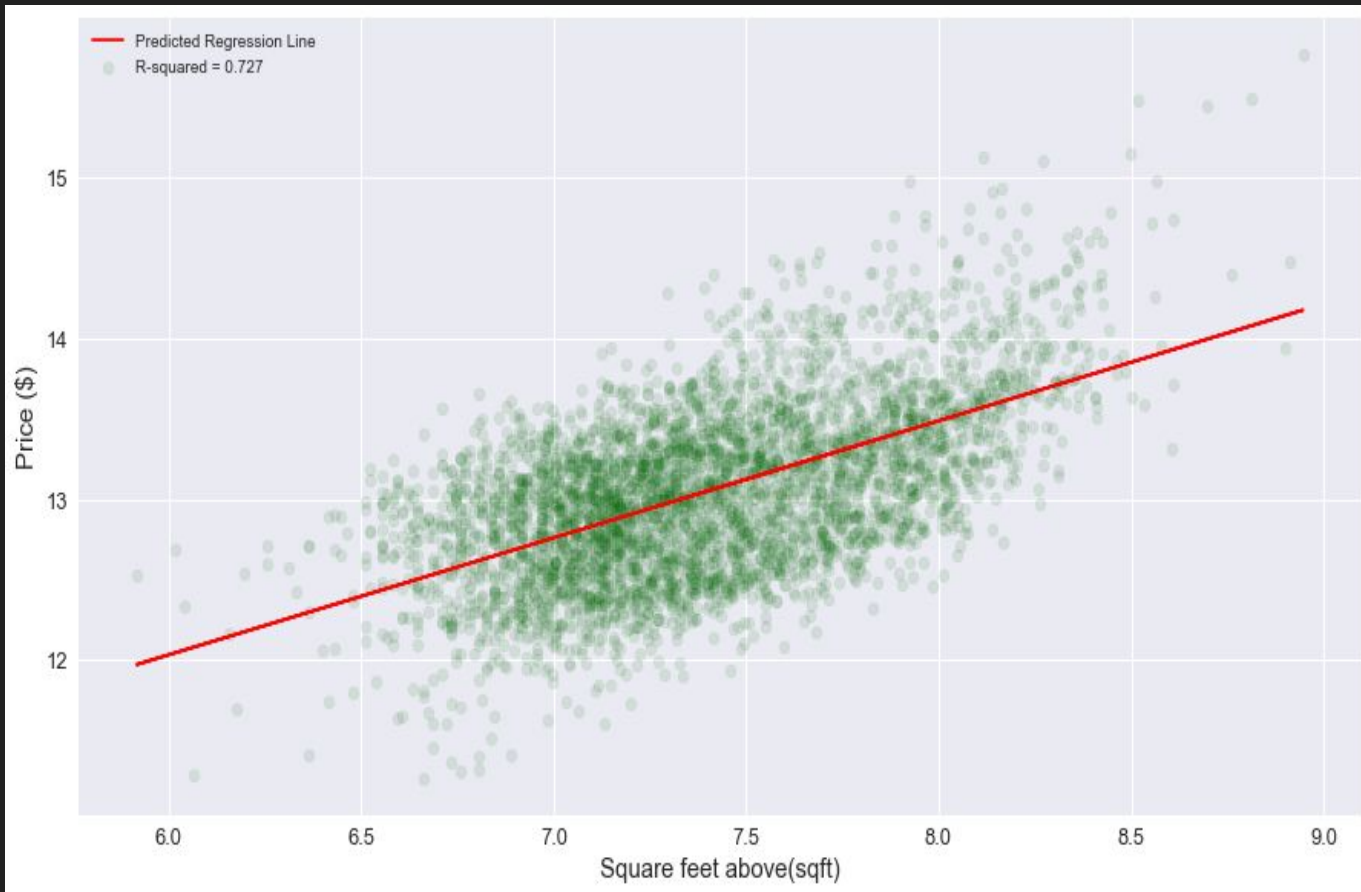
Results

Residuals for Ridge Model

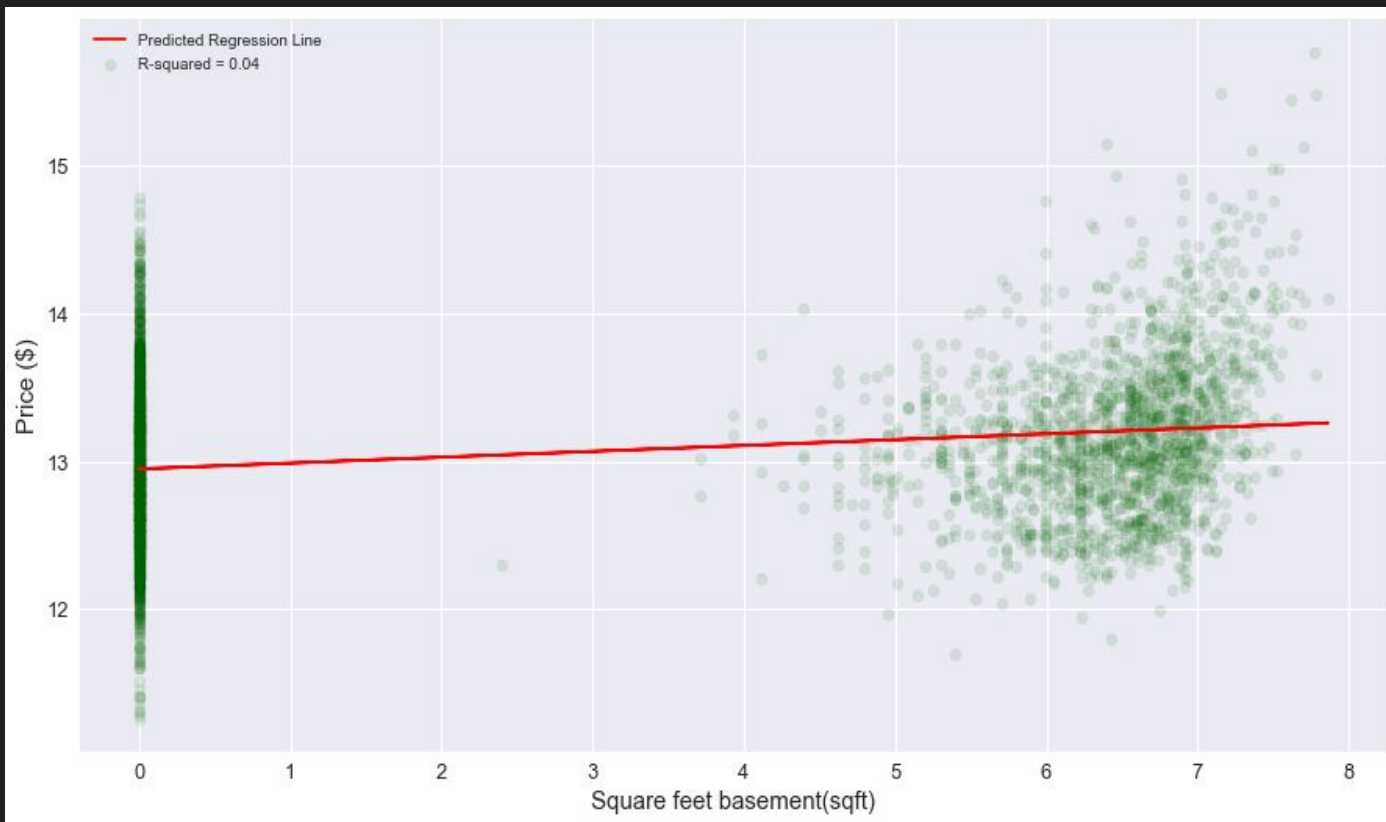Our model explains 85% of the variations in price of houses using the features in our data.

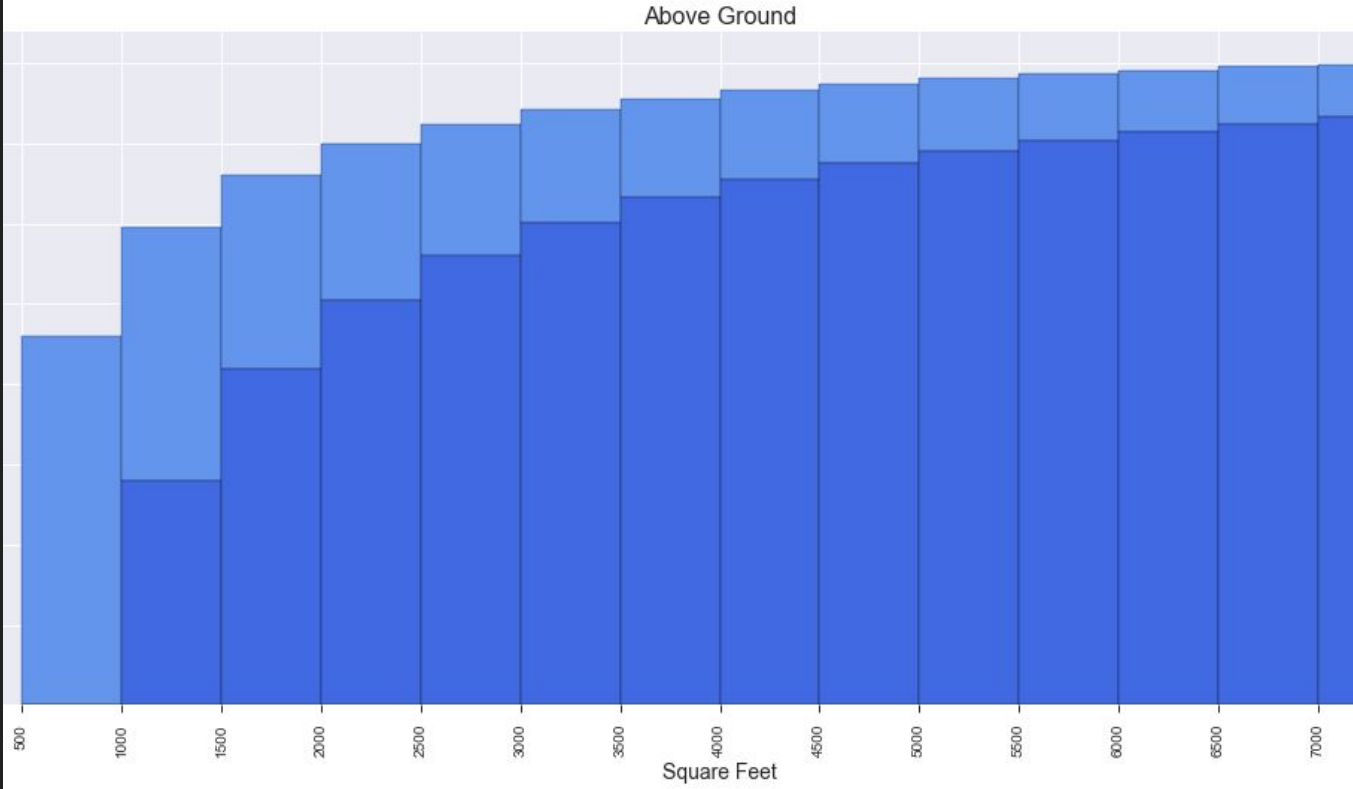Square footage is very important when determining the price of a house.

However, Square feet above is more important.
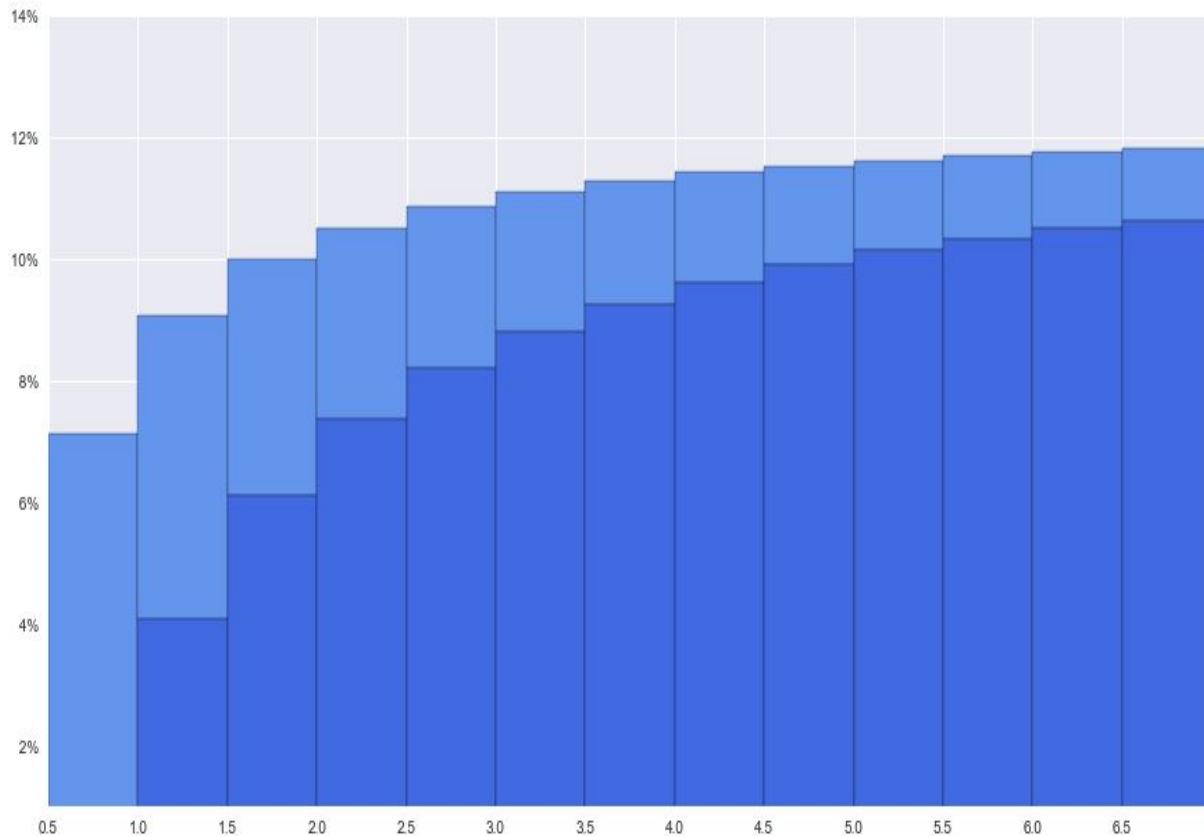
Compare with basement square feet.

As we can see, it's not a good predictor of price at all.

Percent Increase in House Price with Increasing Square Footage

To maximise profit, houses between 500 and 4000 square feet above ground should be prioritised.

Percent Increase in House Price by Increase in Number of Bathrooms

Additionally, building between 1-3 bathrooms increases has the most value for money based on money spent and increase in price.

Limitations

The project was limited by inability to change features in the model while keeping others constant in order to determine the effect on price of the houses. This can be done in future as in this scenario, it would take too many iterations of the same model to achieve.

Conclusions

In summary, King County Real Estate Services should focus on:

1. Building houses with more square footage above ground and more square footage in general.
2. Building houses with 1 - 3 bathrooms.

Thank You!

Email: moses.kuria@student.moringaschool.com
GitHub: @moseskuriia
LinkedIn: linkedin.com/in/moseskuria/