Module 1

Final Project

- Objective -

The main objective of our research analysis was to produce a model that predicts the price of homes for King County.

APPROACH

In order to come up with an accurate model to predict housing prices, we looked at the King County Housing Data Set which had a sample of over 20,000 homes.
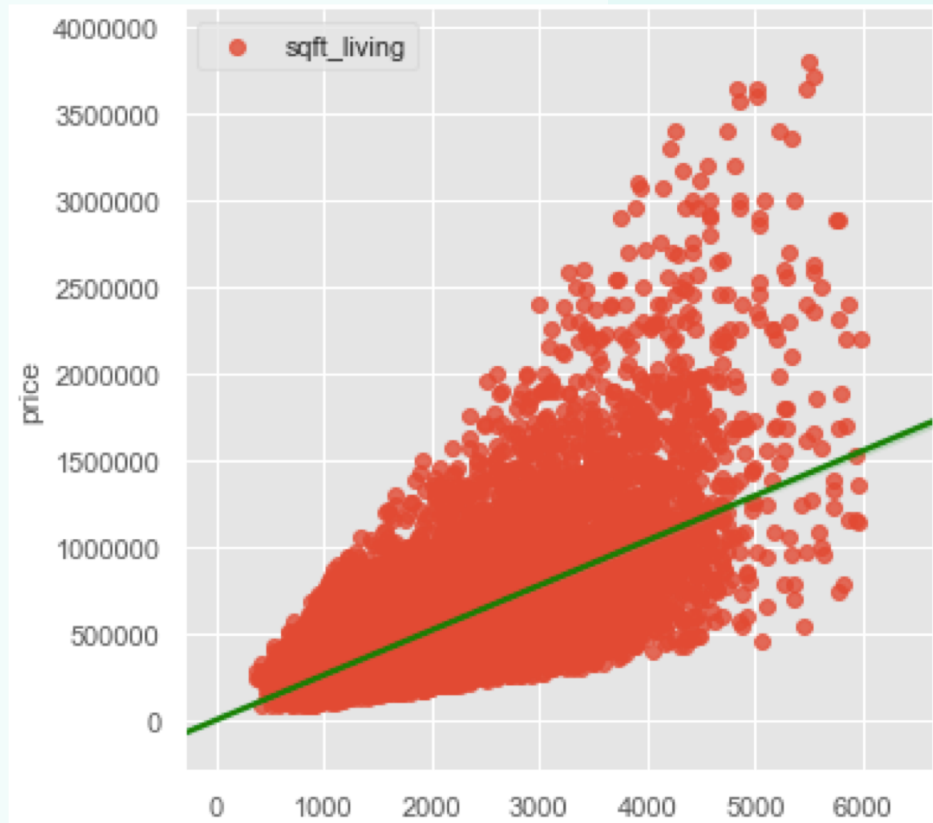
Our method was to use this data to fit a model with the most important features of a home in relation to its price.

The final model ended up with 3 main features.

Feature #1: Square Footage

The first feature is shown here demonstrating the strong relationship between square feet and price.

Our model predicts that for every **10%** increase in the square footage of your home results in about a **5%** increase in its price.

Feature #2: Building Grade

1-3 Falls short of minimum building standards. Normally cabin or inferior structure.

4 Generally older, low quality construction. Does not meet code.

5 Low construction costs and workmanship. Small, simple design.

6 Lowest grade currently meeting building code. Low quality materials and simple designs.

7 Average grade of construction and design. Commonly seen in plats and older sub-divisions.

8 Just above average in construction and design. Usually better materials in both the exterior and interior finish work.

9 Better architectural design with extra interior and exterior design and quality.

10 Homes of this quality generally have high quality features. Finish work is better and more design quality is seen in the floor plans. Generally have a larger square footage.

11 Custom design and higher quality finish work with added amenities of solid woods, bathroom fixtures and more luxurious options.

12 Custom design and excellent builders. All materials are of the highest quality and all conveniences are present.

13 Generally custom designed and built. Mansion level. Large amount of highest quality cabinet work, wood trim, marble, entry ways etc.

The second feature is the building grade that a housing unit receives. Here you can see the different factors that go into the grading system.

In regards to grade, our model predicts that each increase in grade a housing unit receives results in a 11% increase in price.

The final feature that our model uses is the home's zip code. Although this feature cannot be used to help increase price, it is very useful in determining the price of your home in order to maximize profit.

The chart below shows the King County zip codes along with their associated average price. After accounting for square footage and grade of the house, the model then uses the home's zip code to further predict the price.

For example, the model shows an increase in price of about 111% for homes in Medina, WA (98039). It's easy to see that Medina is the most expensive part of King County with an average home price of over $2 million.

Feature #3: Zip Code

Price of Homes by Zipcode
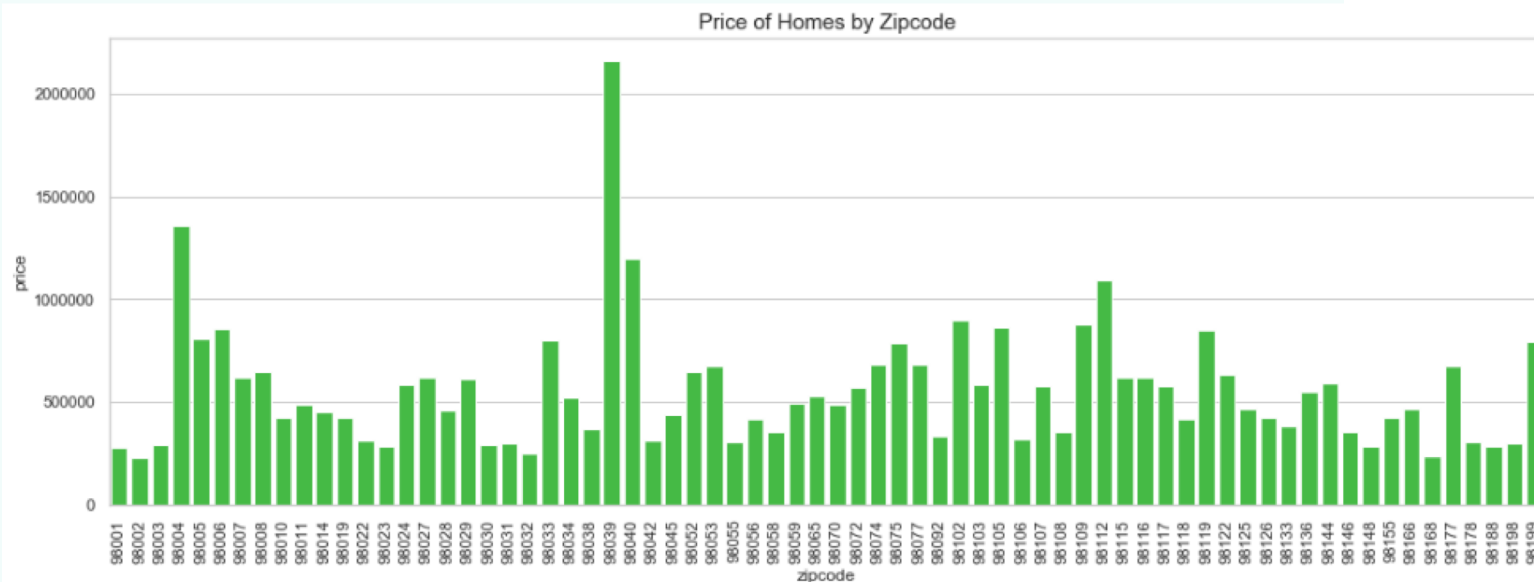
Final Thoughts

Overall our model can be used in many valuable ways in order to maximize the profit of your home.

In the end the model finished with an R2 value of about 0.838 meaning that about 84% of the spread of the data around its average can be explained by the model.

Again using the three main features of the home in order to predict the price:

- Square Footage
- Building Grade
- Zip Code

- T H A N K   Y O U -

Sabatino M. Chen