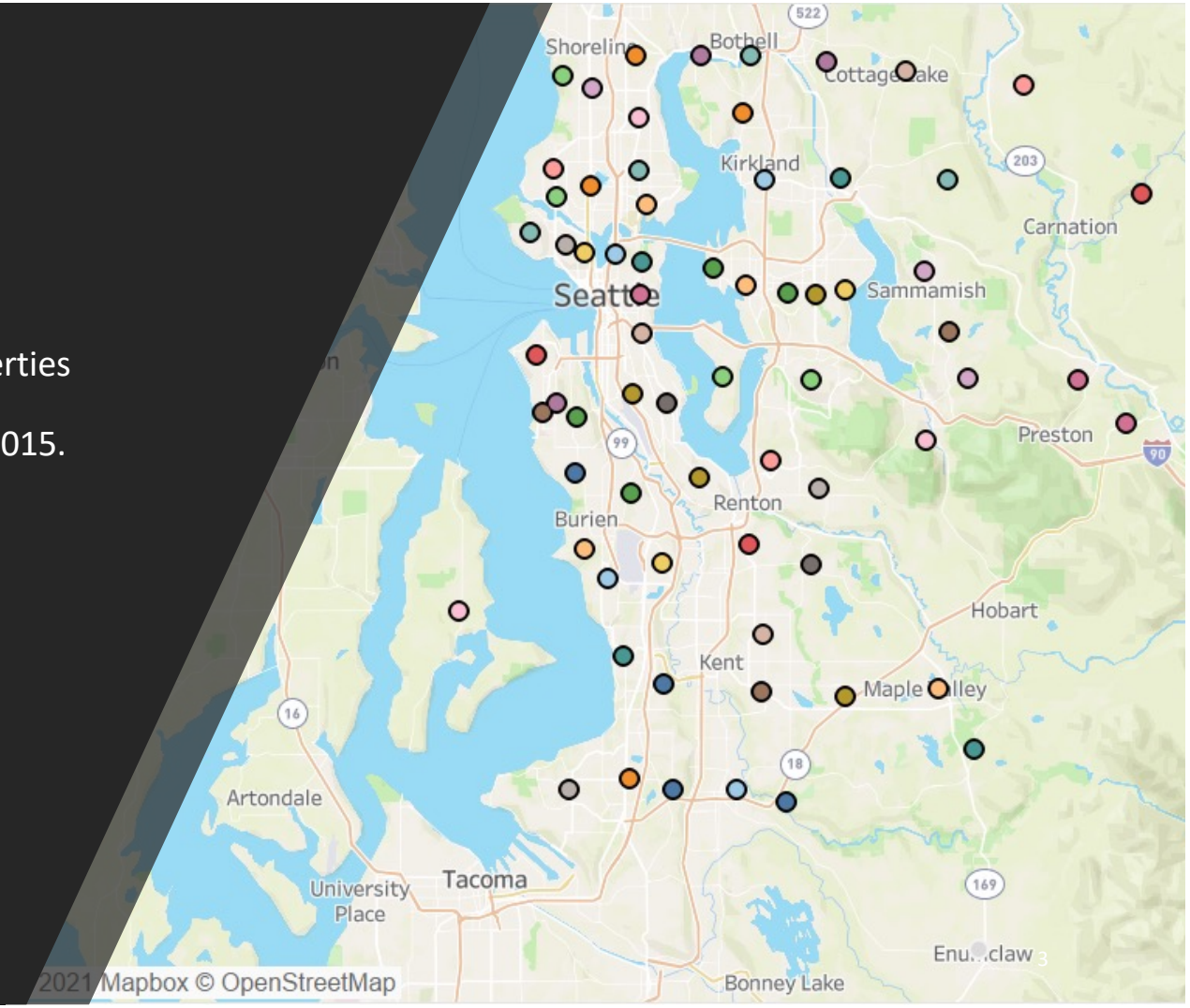

Seattle, Washington Real Estate
Presentation by Hermione Granger

Objectives

- Create a model to predict price based on features.
- Using BI tools, explore the property features in this dataset.
- Understanding the factors responsible for high property values.

The Data

- The data set consist of information on 22,000 properties sold in Seattle Washington, between May 2014 to May 2015.

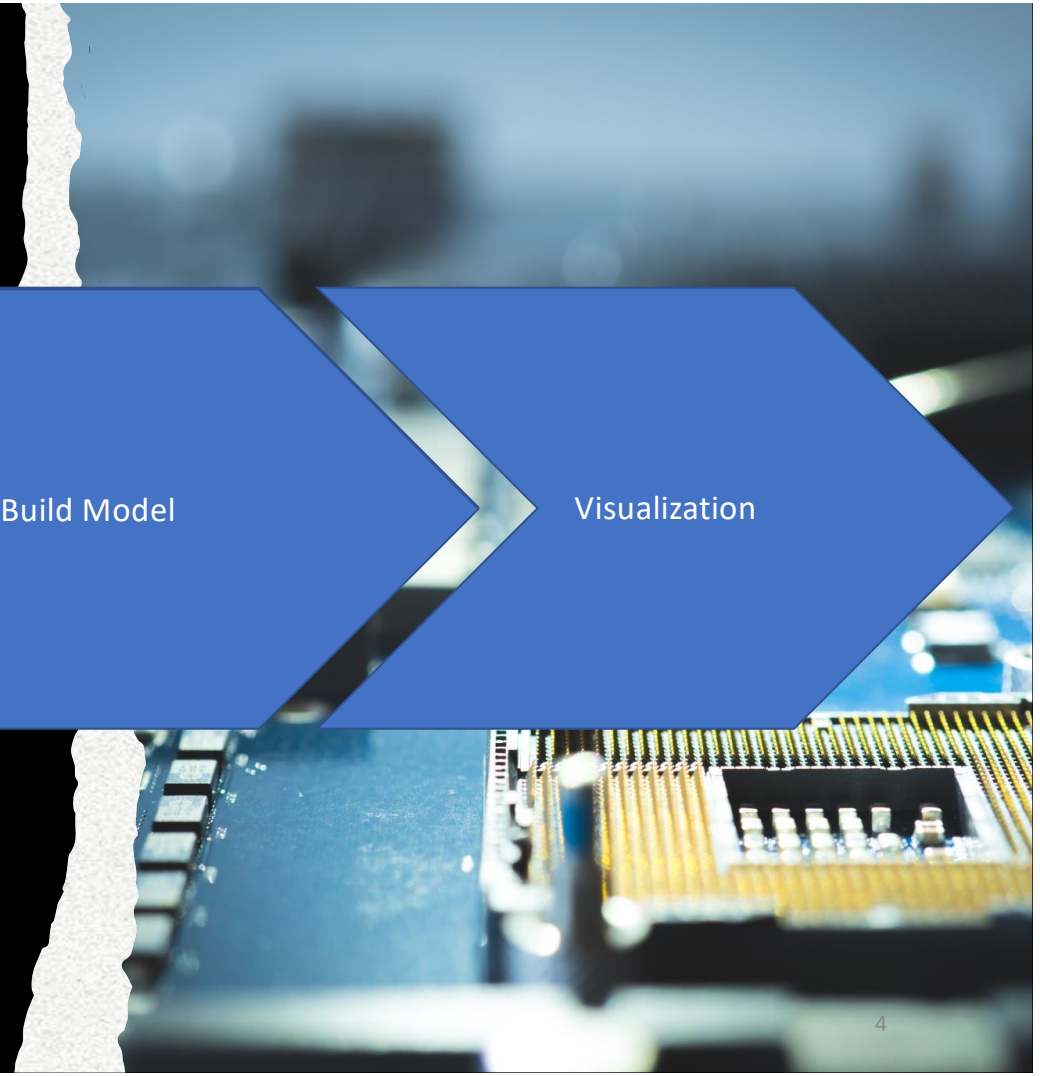


Workflow

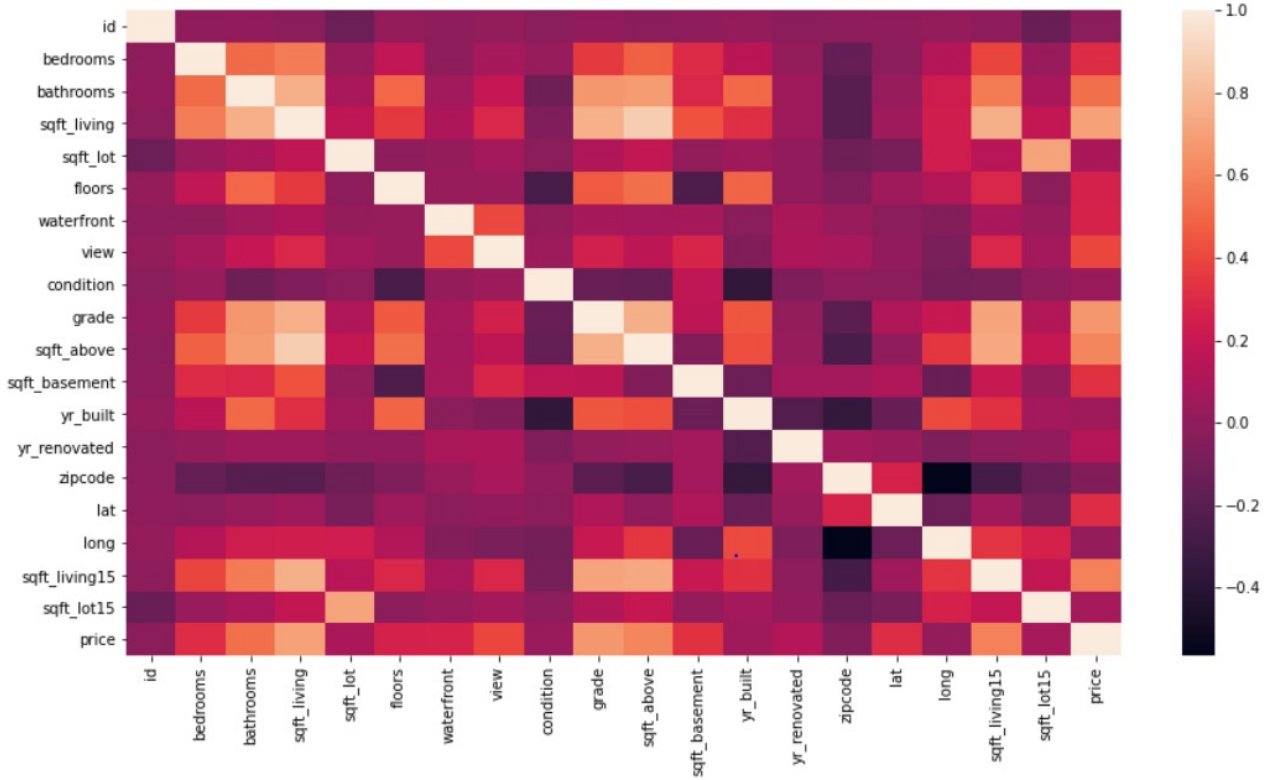
- Explore data

Build Model

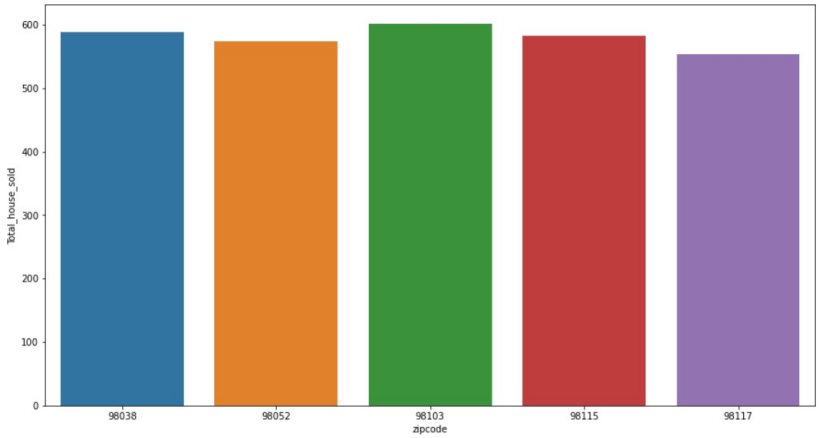
Visualization



Correlation between dependent and independent variables

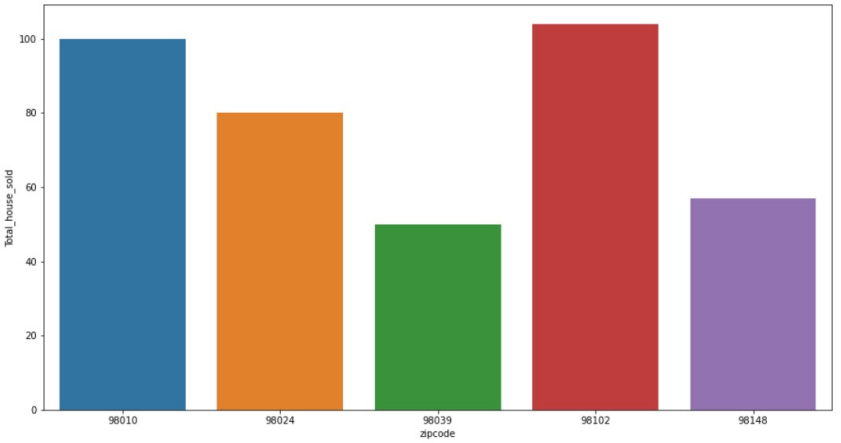


Top 5 highest selling zipcode by median price



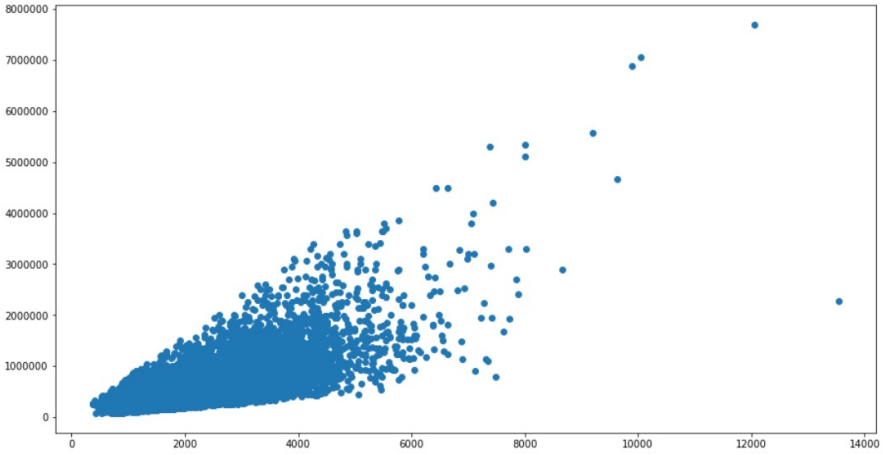
You will notice that it is an even distribution. After doing some research on line, it is in conclusion that this is due to low morgage rates and favourable remote work by employers in this area and also that it is a home-owner middle class area.

lowest selling zipcode by median price

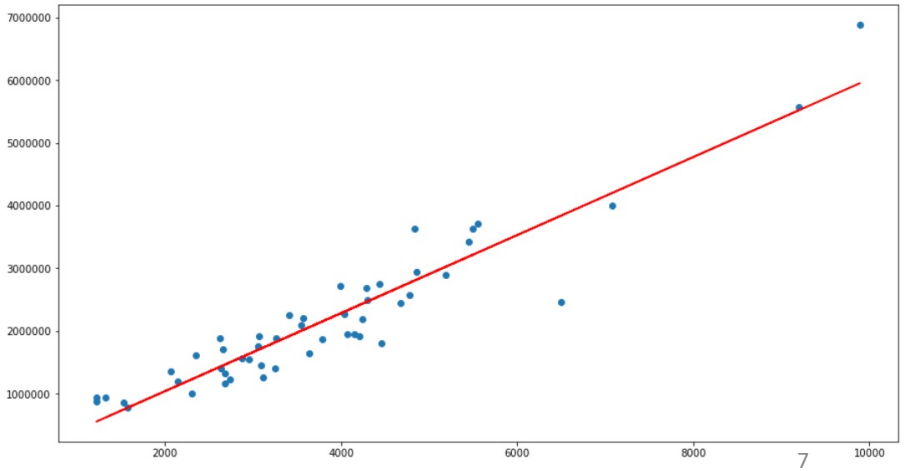


The least lowest number of properties sold are in Zipcode 98039. If we check back on our most expensive neighhoods, the same zipcode records highest. Further research revealved it called Medina and is home to some of the worlds richest billionaires; Jeff Bezos and Bill Gates.

According to the data represented, one zipcode is an outlier.



The most expensive neighbourhood

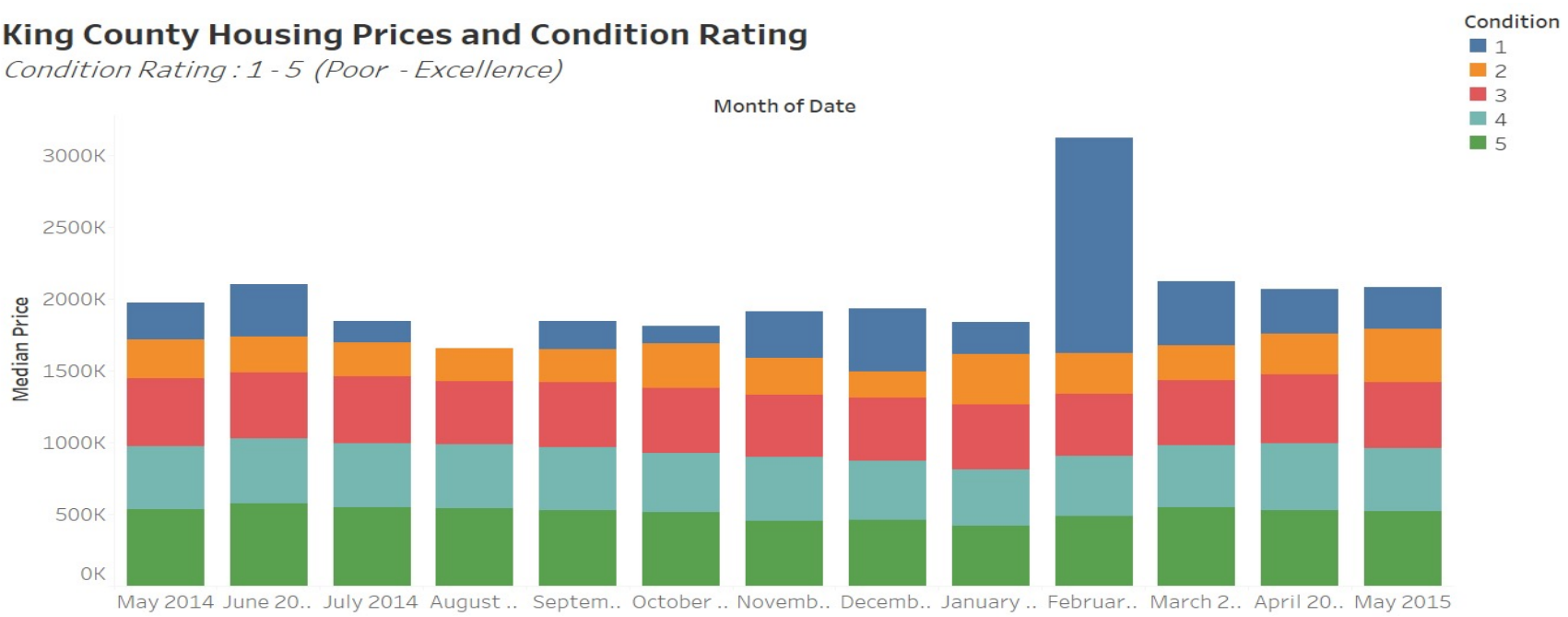


The Algorithm

- **Simple Linear Regression was used to analyze the relationship between dependent and independent variable with a prediction score of 0.49%.**
- **Multiple Linear Regression was used to predict the outcome using more independent variables with 0.88% and improved to 0.92% accuracy score.**

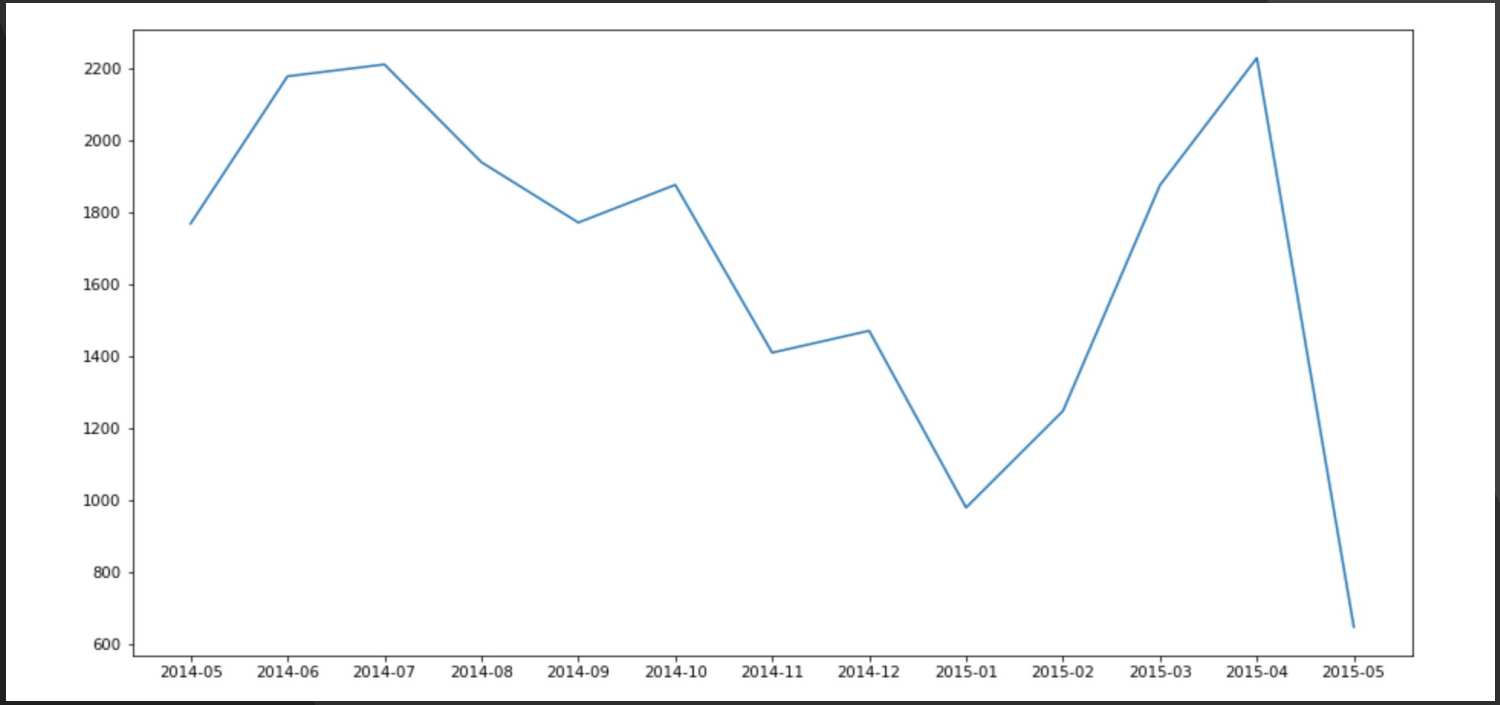
King County Housing Prices and Condition Rating

Condition Rating : 1 - 5 (Poor - Excellence)



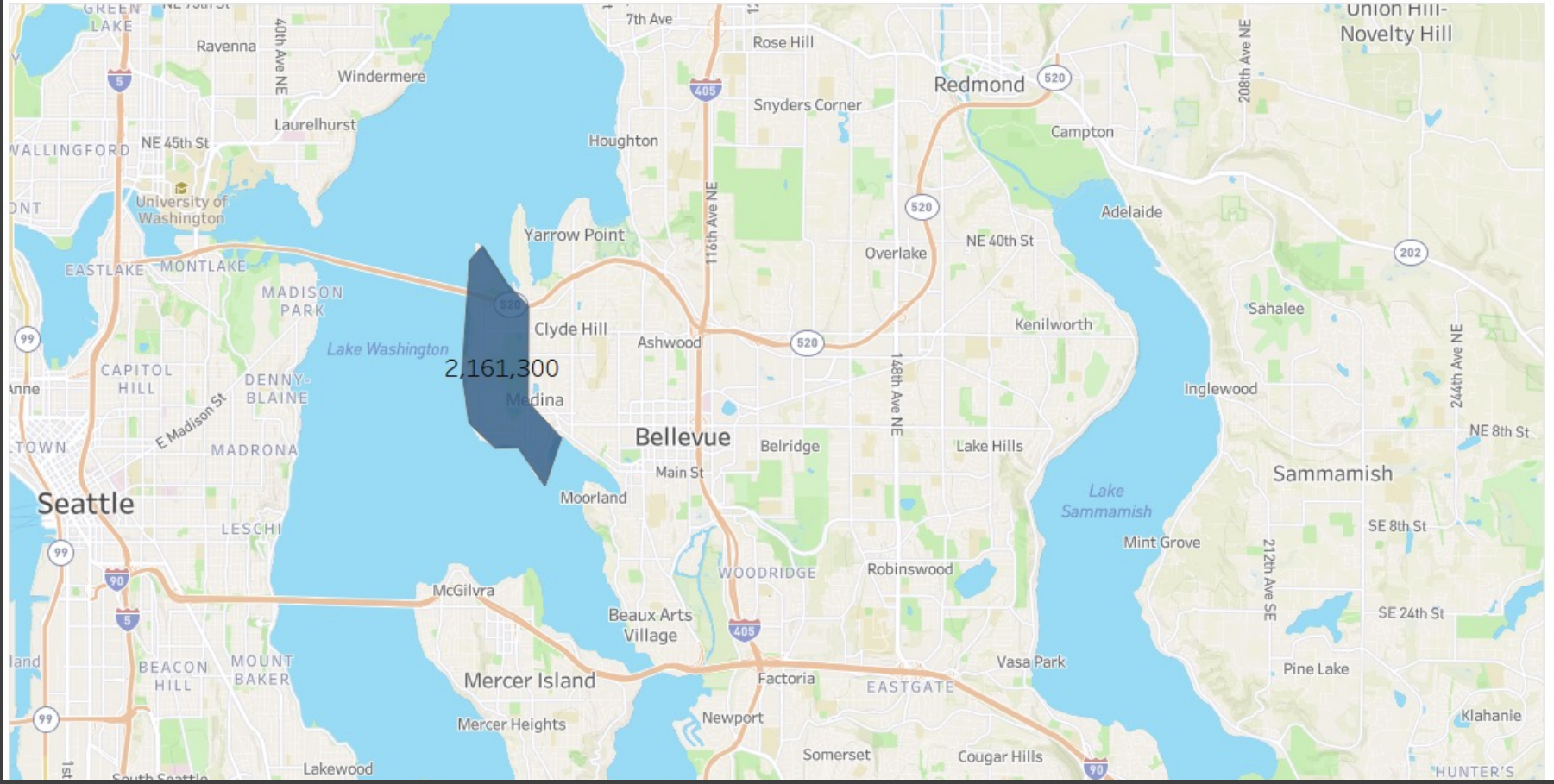
The plot shows that there is correlation between price and condition even though there is bit of inconsistency. The green represent excellence and has more consistency, which signifies a parameter for the purchase decision while the blue represent poor and has less significance in the purchase decision.

Visualizing the trend of number of properties sold for each month




Insights

Most expensive houses based on zipcode



Challenges

Converting CSV

Fixing Tableau Dashboards

Accuracy Score

Thank you