Predicting King County Housing Prices

PREPARED BY;

Jeremy Nguyo

Overview

The King County Housing Data Set contains information about the size, location, condition, and other features of houses in Washington's King County.

The Data:
- 21,000 house sales
- 20 predictors

The goal is to explain some of the top features used to predict the sale price of homes and ultimately lead to higher sale price.

A Linear Regression Model was adopted for analysis.
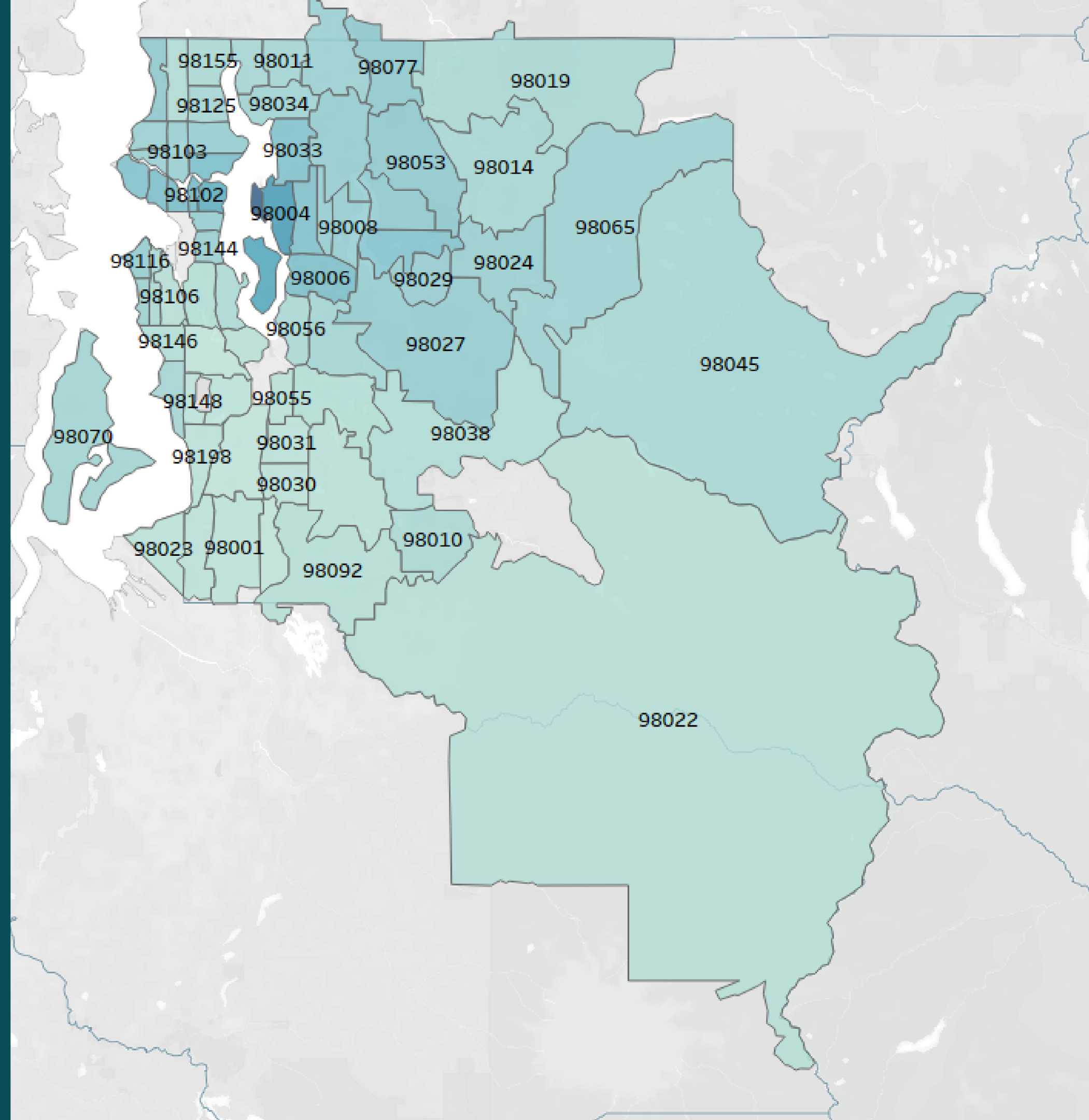
Linear Regression Analysis **Results**

P < 0.05

$R^2 = .547$

The model accounts for 54.7% of the variance in price. A p-value of less than 0.05 means we can reject the null hypothesis that there is no relationship between price and the predictor variables.

The predictor variables are:
   a. Sqft_Living
   b. View
   c. Sqft_Living15

PRICE ~ SQFT_LIVING + VIEW + SQFT_LIVING15

Question 1:

Which features are most highly correlated with price?

The features are listed below in descending order:
- Square Footage of the Living Space
- Square Footage of the House apart from the Basement
- Square footage of Interior Housing Living Space for the Nearest 15 Neighbors
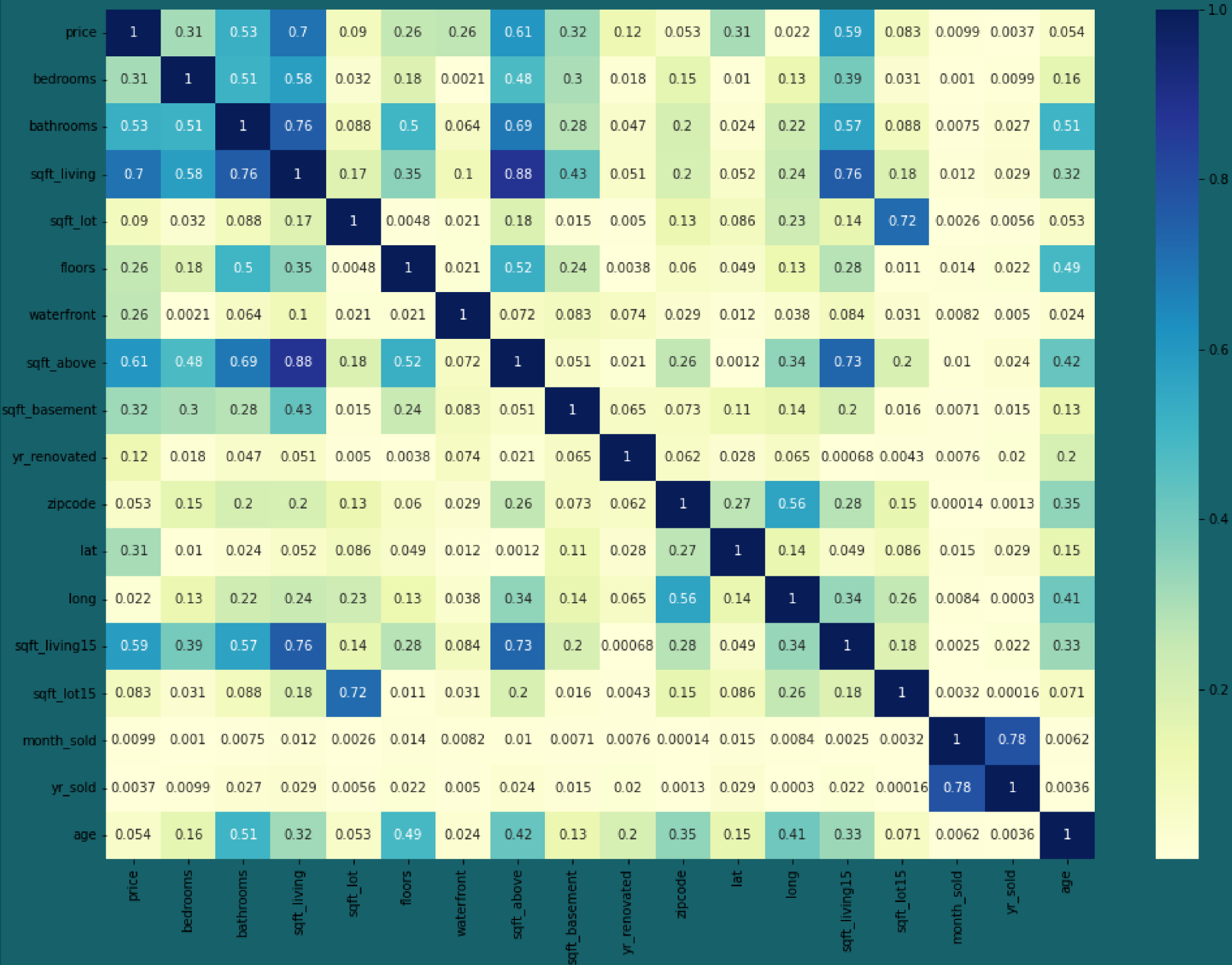- Bathrooms

Variable Correlations

Question 2:

Which features have the strongest correlations with other predictor variables?

The features are listed below in descending order:
- Square Footage of the Living Space
- Square Footage of the House apart from the Basement
- Square footage of the basement


Variable Correlations

Question 3:

What combination of features is the best fit, in terms of predictive power, for a multiple regression model to predict house prices?

The features are listed below in descending order:

- **Square Footage of the Living Space.** For each additional square footage of the house, the price is projected to increase by USD 200

- **Square footage of Interior Housing Living Space for the Nearest 15 Neighbors.** The price of a house is projected to increase by USD 52 for every increment of square footage of the nearest 15 neighbors.

- **The View**. House with the best views will always go for a very high price. Below is a cost comparison to houses with no view:
    a. One with an excellent view will cost USD600k more
    b. One with a good view will cost USD 200k more
    c. One with fair view will cost USD 160k more
    d. One with an average view will cost USD 119k more

Scatter Plot of Price vs Sqft Living
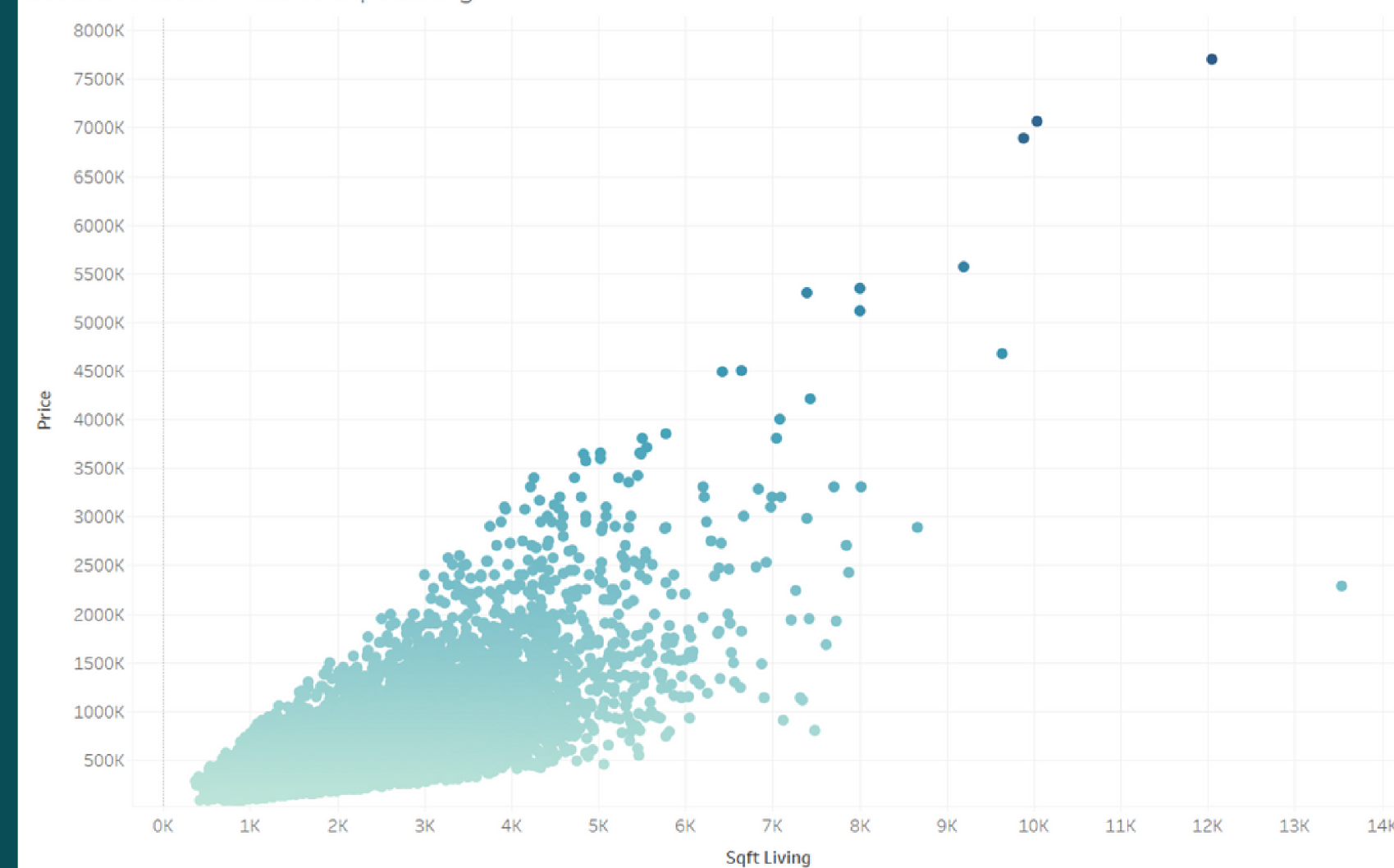
Price vs View Bar Graph
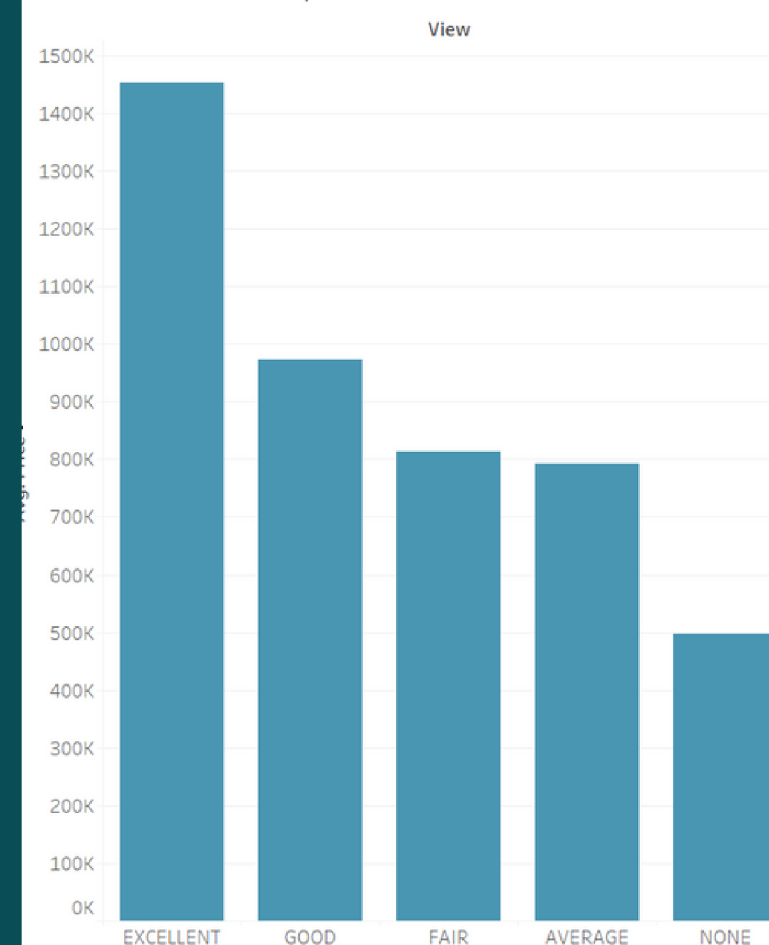
Conclusion

1. The houses with the best views always fetch a high price
2. A larger living space will increase the price of a house
3. The size of the neighboring houses affects the cost of a house. The larger the houses, the higher the price.