

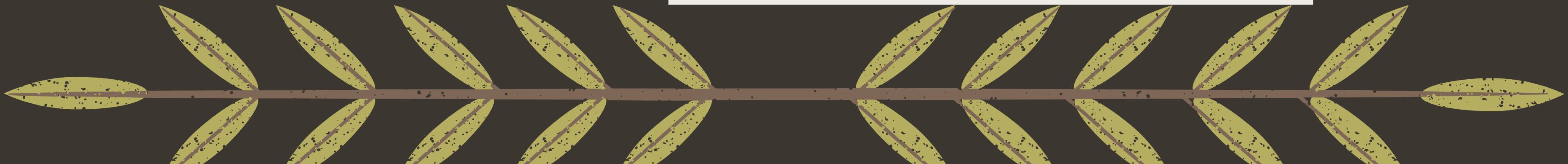


King County Housing Analysis

by Medrine Waeni

Overview

- Business Understanding
- Data Understanding
- Modeling
- Regression Results
- Recommendations
- Next Steps
- Thank You



Business Understanding

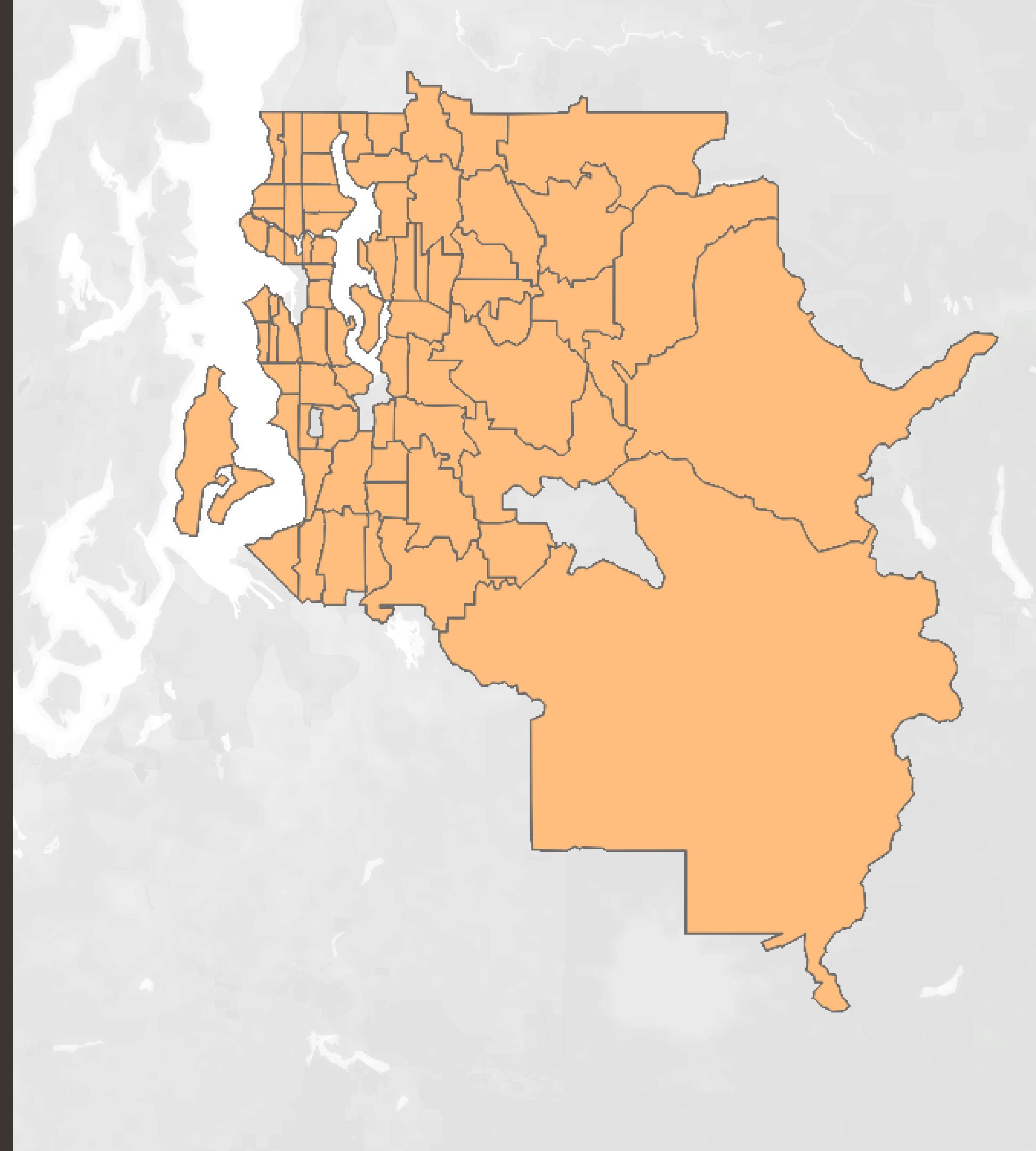
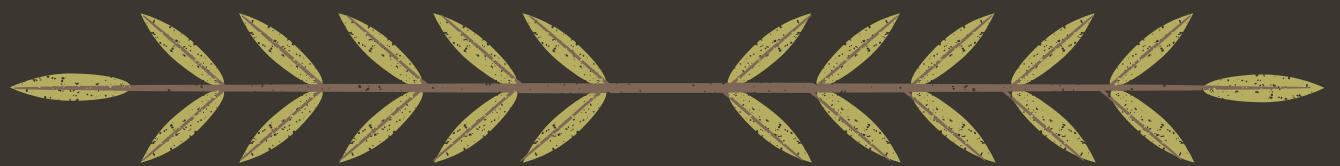
Business Problem

An association of home owners are seeking to know what features in a house could be remodelled to increase the house prices and what features they should add as they list their houses in the market.



Data Understanding

- King County is located in Washington, USA
- The dataset contains 21 features
- From the data, the houses are values at:
 - mean = \$ 540,000
 - maximum = \$7,000,000
 - minimum = \$78,000



Method

Data Exploration



Data Preparation and cleaning

- Null values were replaced with the modes
- View, condition and grade columns were changed to numbers
- Extreme outliers in grade were removed i.e grade 3 and 13

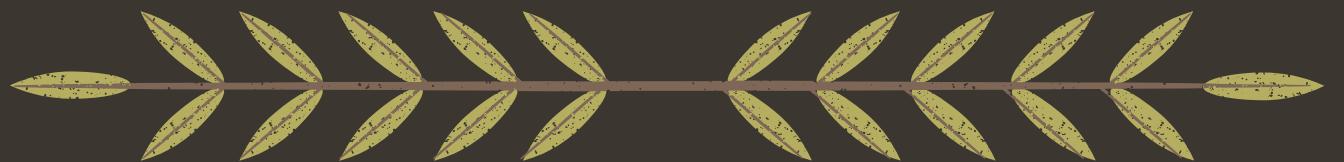
Modeling and analysis

- An iterative process to select the model was undertaken
- The best model was evaluated

Modeling

Feature Exploration

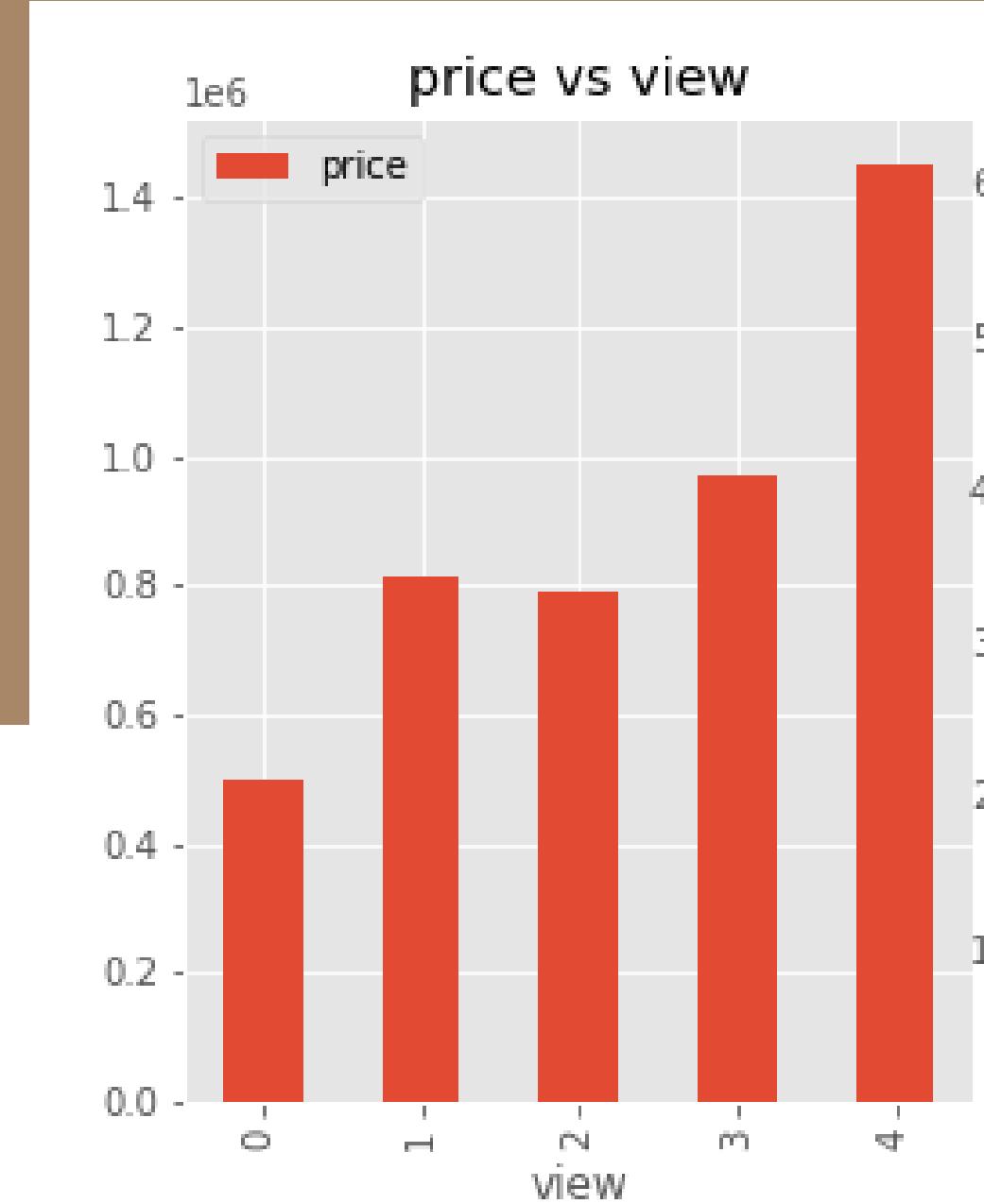
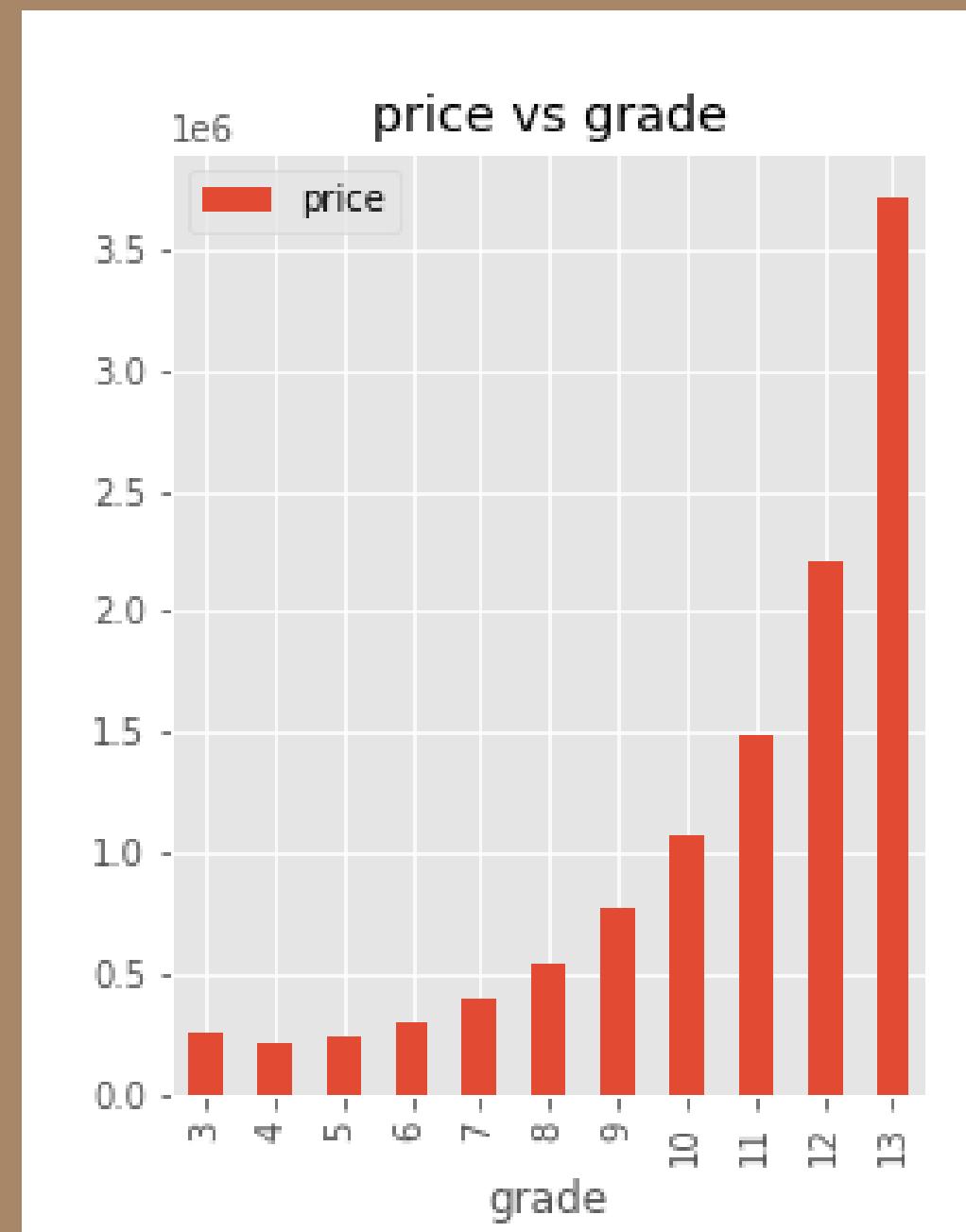
- The feature with the highest correlation with house price was the sqft_living.
- This means that the square footage of the living room in the house had a big impact on the house prices



Modeling

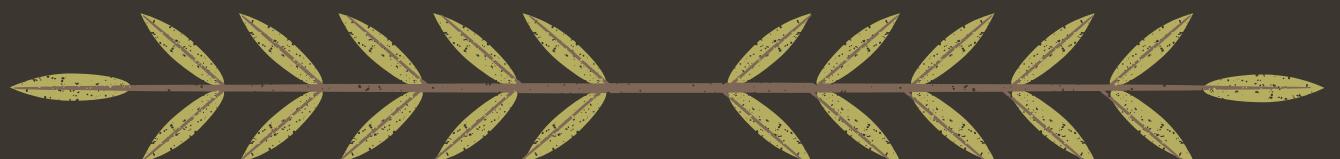
Feature Exploration

- The features highlighted in the model are views and grade
- Their distribution with price is as shown
- Insert percentages for the price increase



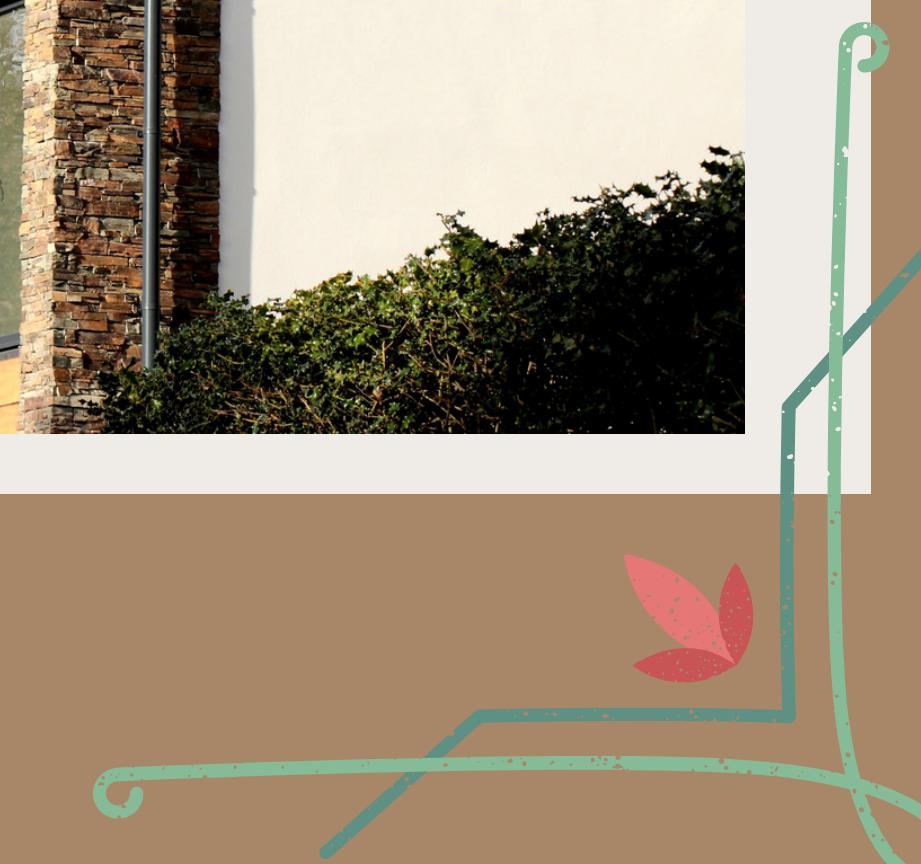
Regression Results

- The model selected explained about 65.4% of the variance in price
- The model was off by \$141,119 in price
- Views that have been rated as 4 or 1 i.e Excellent or Fair respectively increased house prices by more than \$122,500
- Houses that built with excellent quality material were found to have price increase of \$1,923,000



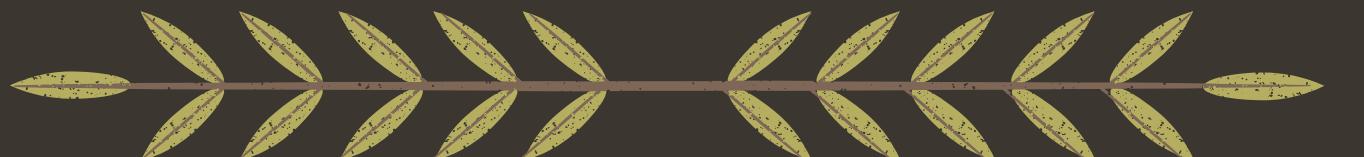
Recommendation

- The home owners, should they intend to renovate or list their houses for sale, should focus on the following features:
 - Ensure the house has a waterfront
 - Ensure the living space square footage is large
 - The grade of the building materials should be excellent
 - The house should have a fair or excellent view



Next Steps

- The data needs to be analysed while zoning in on specific zipcodes so as to get more accurate and fine tuned results
- Using other types of models aside from linear regression to analyse the data



Thank You

Any questions?

