Moringa School

Real Estate Analysis, King County

Presented by: MULEI MUTUKU

Overview

- *Introduction*
- *Problem Statement*
- *Objectives*
- *Data Understanding*
- *Analysis*
- *Conclusions*
- *Recommendations.*

Introduction

Real estate agencies are the bridge between buyers and sellers of real estate properties. They guide potential estate buyers and sellers by providing unbiased insights and advice. They also;

- Conduct a comprehensive market analysis to determine the property's market price
- Identify the best places to invest resources in improvements
- Identify the property's top selling points

Problem Statement

We will be conducting an analysis of king county property sales data so as to be able to guide potential customers/ property owners, on the best and most effective way to market their property.
In order for them to acquire favorable and high market prices for their properties.
This is aimed at improving the Agency's revenue collection.

Data Undestanding

The project used two datasets 'kc_house_data.csv' and 'City Names.csv' which were merged together and all un-relevant columns dropped to form one data frame.
The following variables were used to build the model;

- Numeric variables: bedrooms, bathrooms, sqft_living, sqft_basement, floors, and grade
- categorical variables: waterfront, view and condition

Main Objective

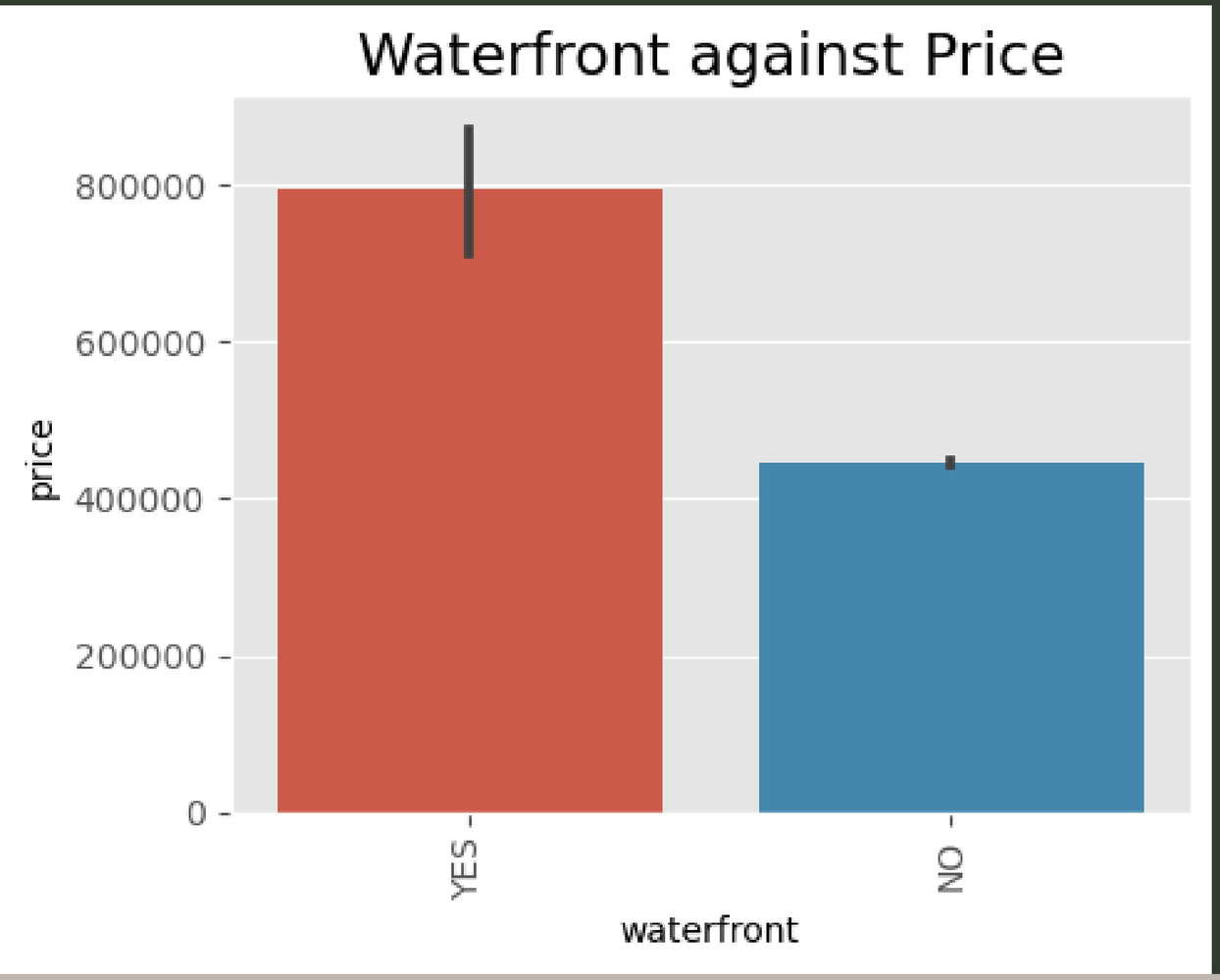
To develop a predictive regression model that accurately predicts the sale price of a property while identifying crucial areas for improvement in order to increase property prices.

Specific Objectives

- Develop a regression model to accurately predict the market price of a house based on its features/ properties.
- Evaluate the performance of the model in predicting market prices.
- From the model identify features that can be improved to increase the house price

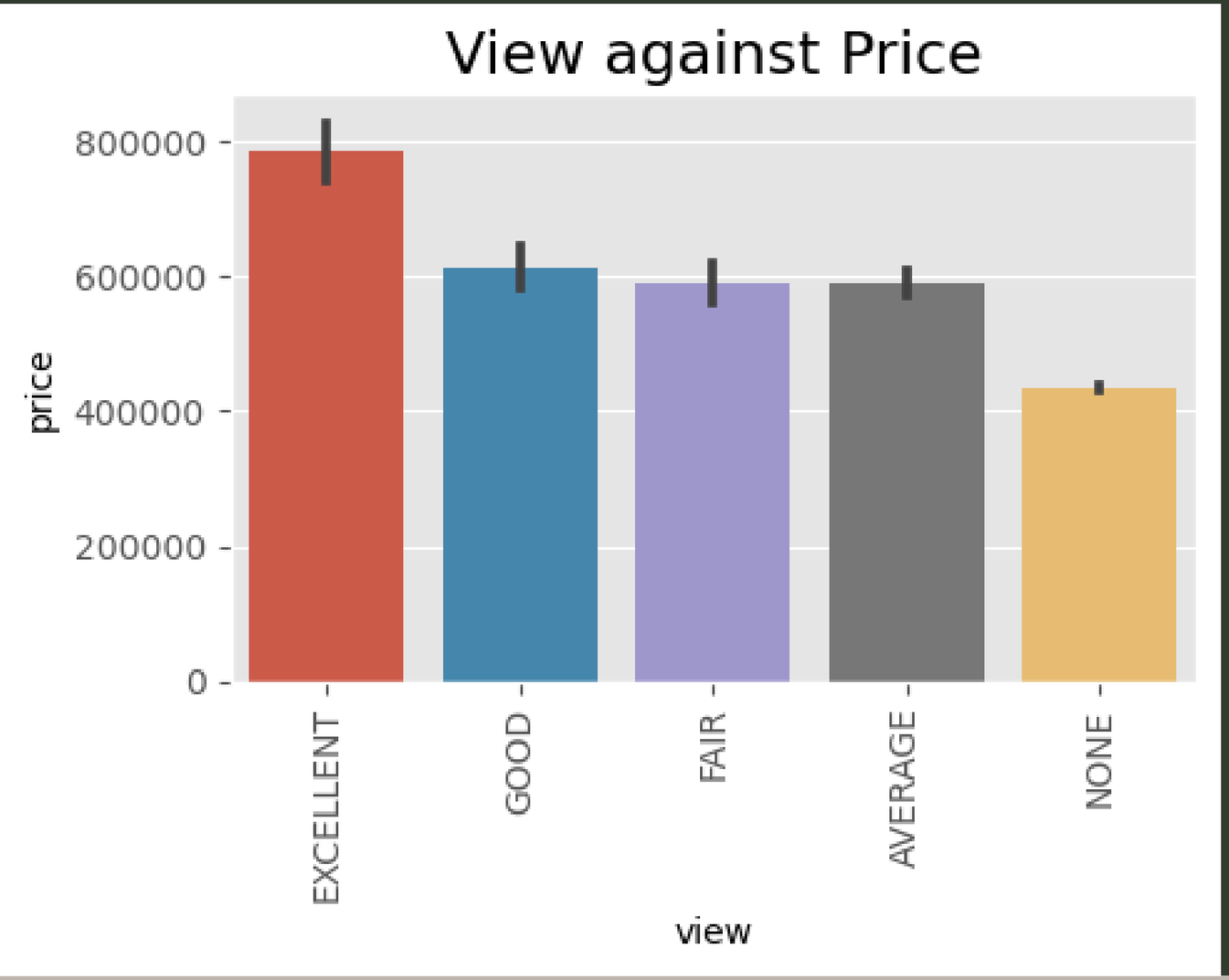
Exploratory Data Analysis

Analysis of Waterfront Against Price



Houses on a waterfront, that is houses around lakes, rivers and canals, fetch high house prices then those that don't.

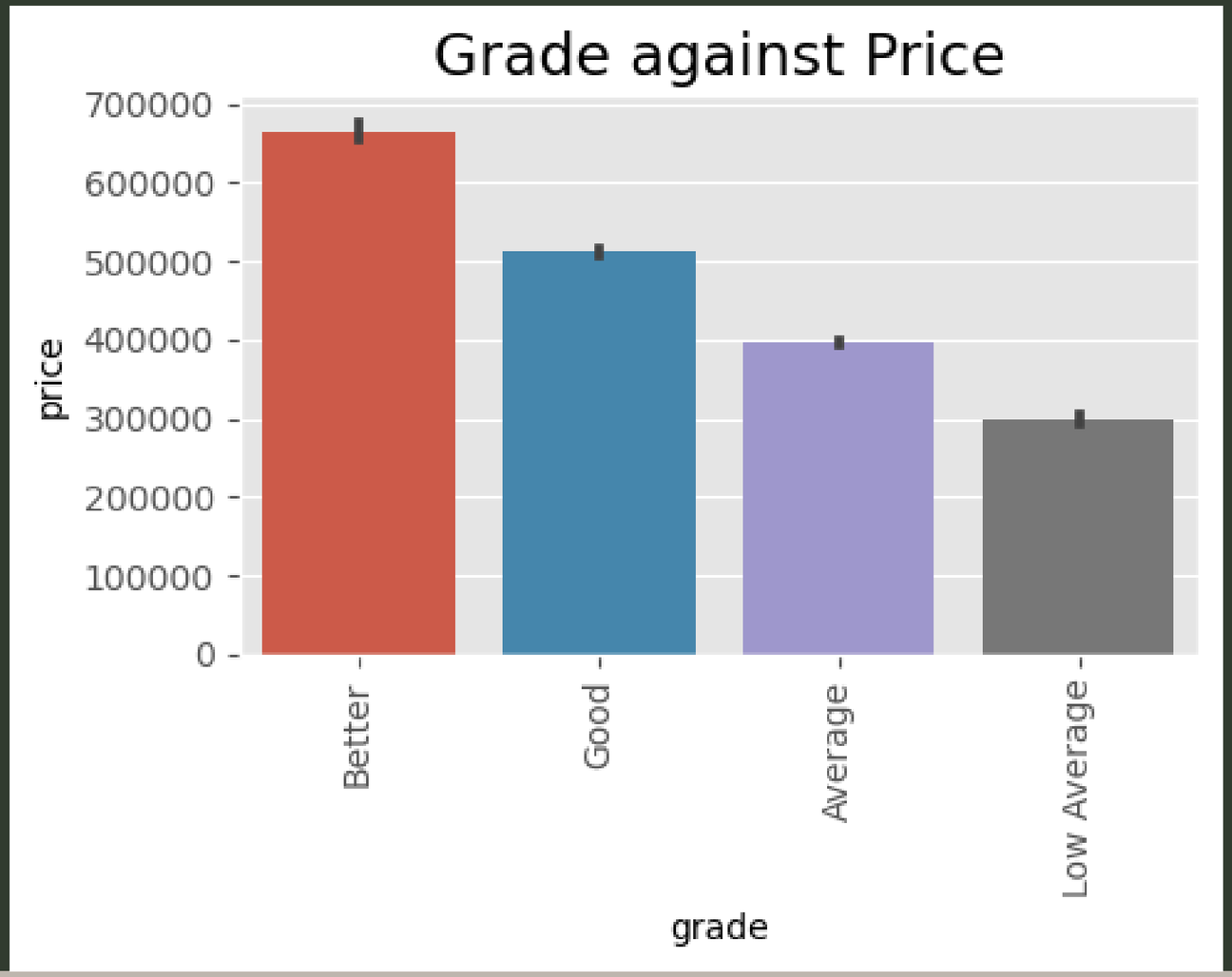
Analysis of View Against Price



View is the surrounding of the house, that is Mountains, Cascades, lake/river/creek, and others.

Houses with Excellent views from the house have higher prices.

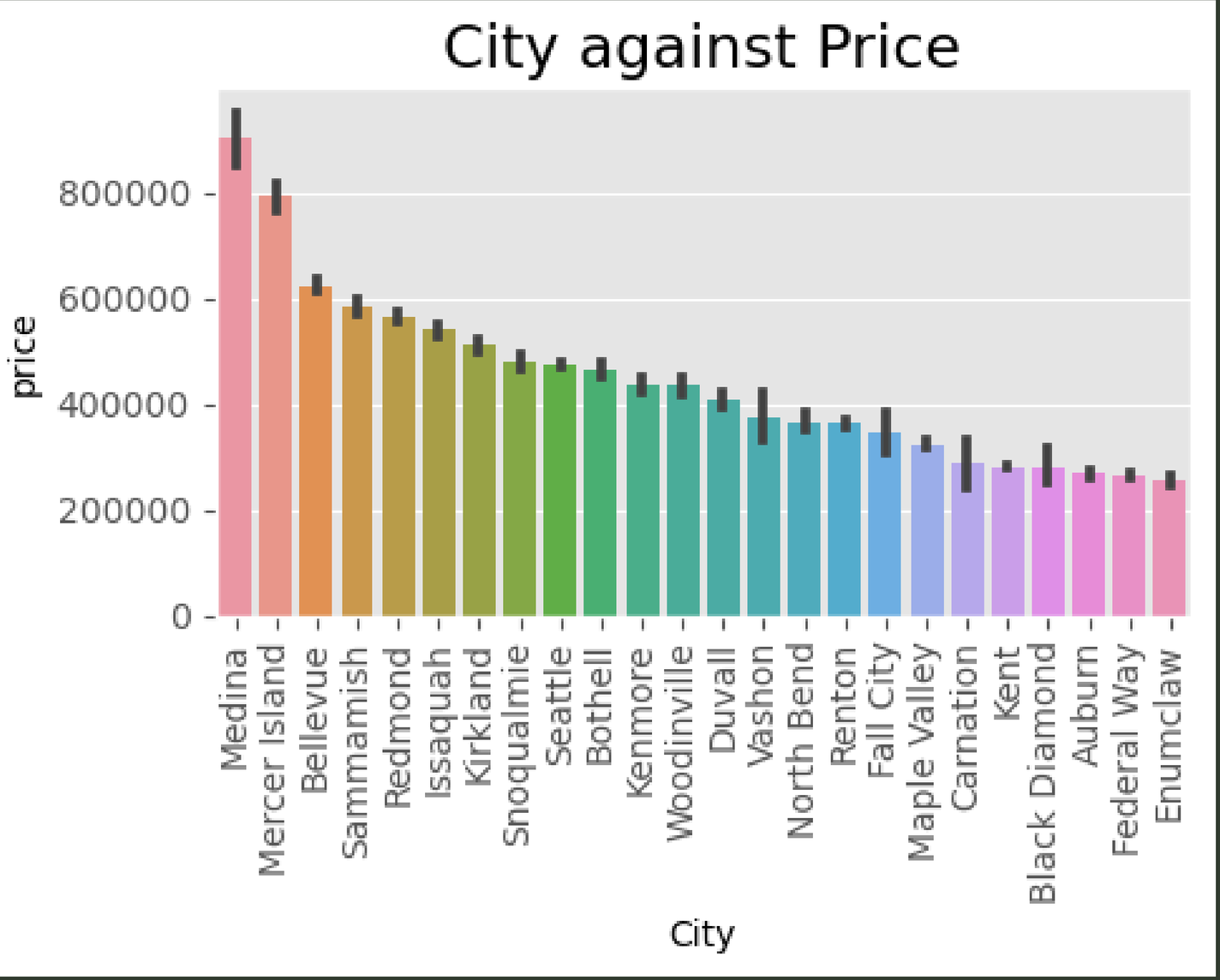
Analysis of Grade Against Price



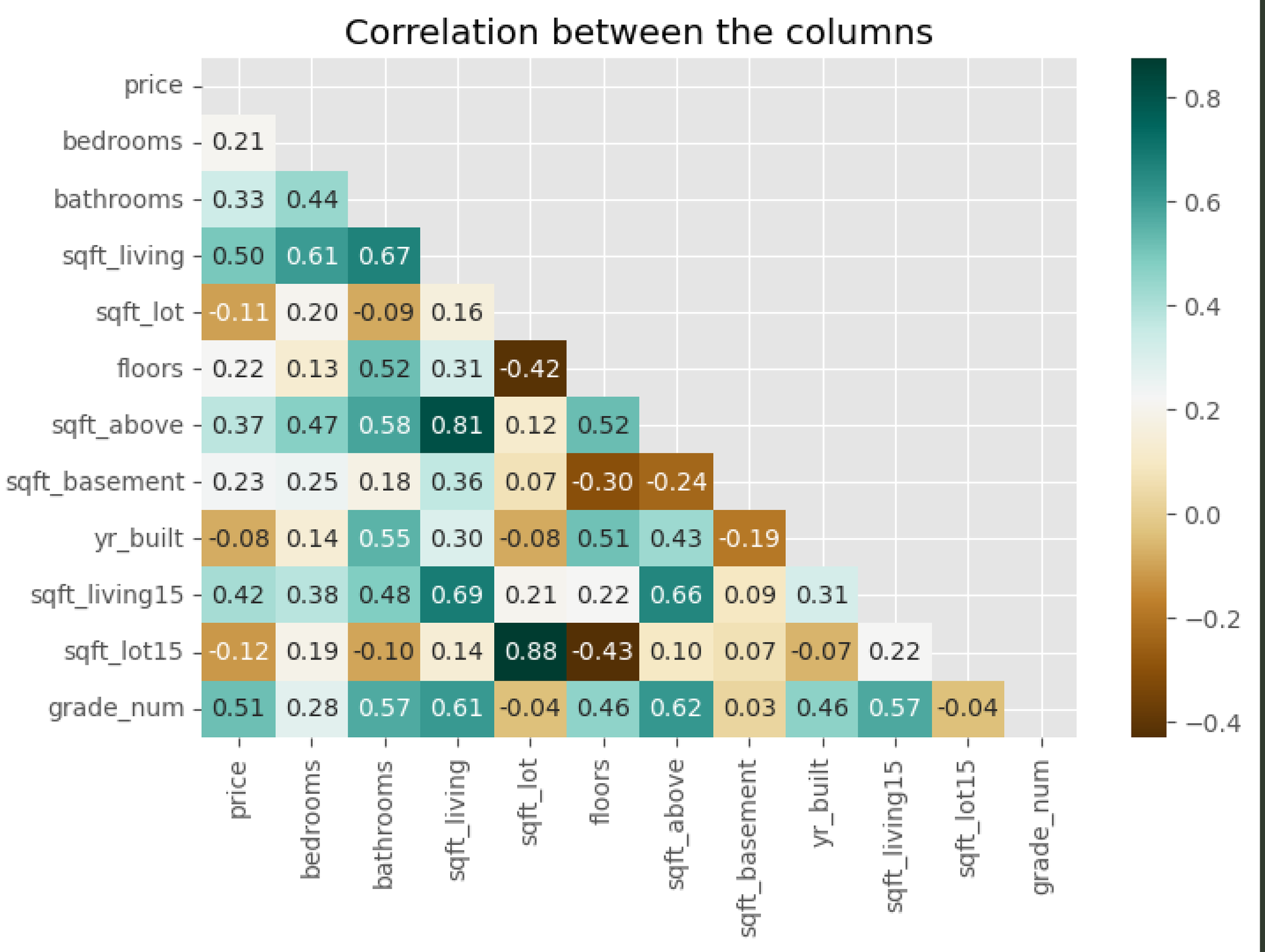
Grade is the house construction in terms of interior and exterior design quality.

Houses with a Better grade fetch high house prices.

Analysis of City Against Price

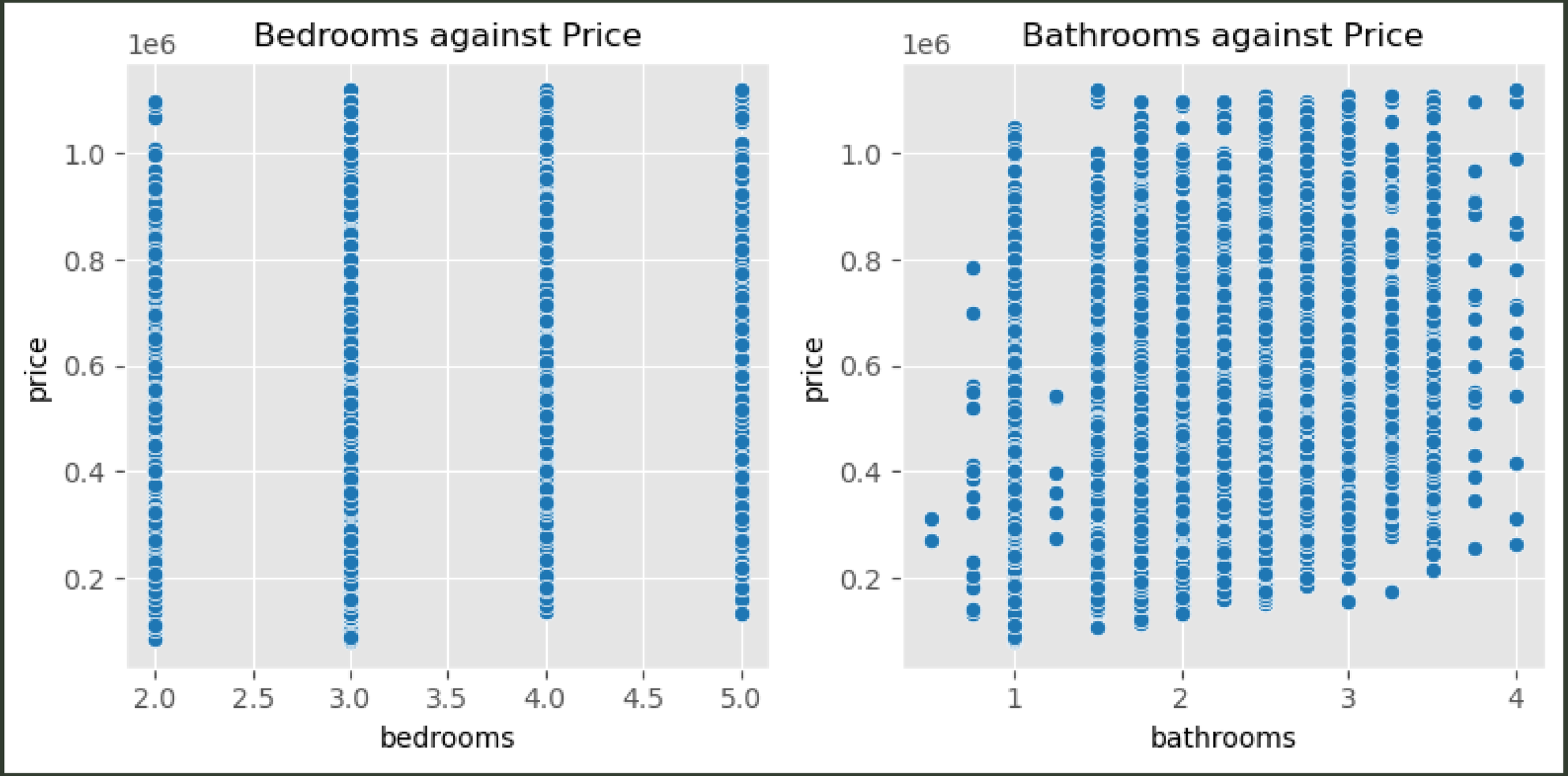


House locations such as houses from Medina, Merce Island, and Bellevue, are more expensive than those from the other locations/cities.
This is because they possess waterfront properties and scenic views that as discussed above fetch high market prices for houses.

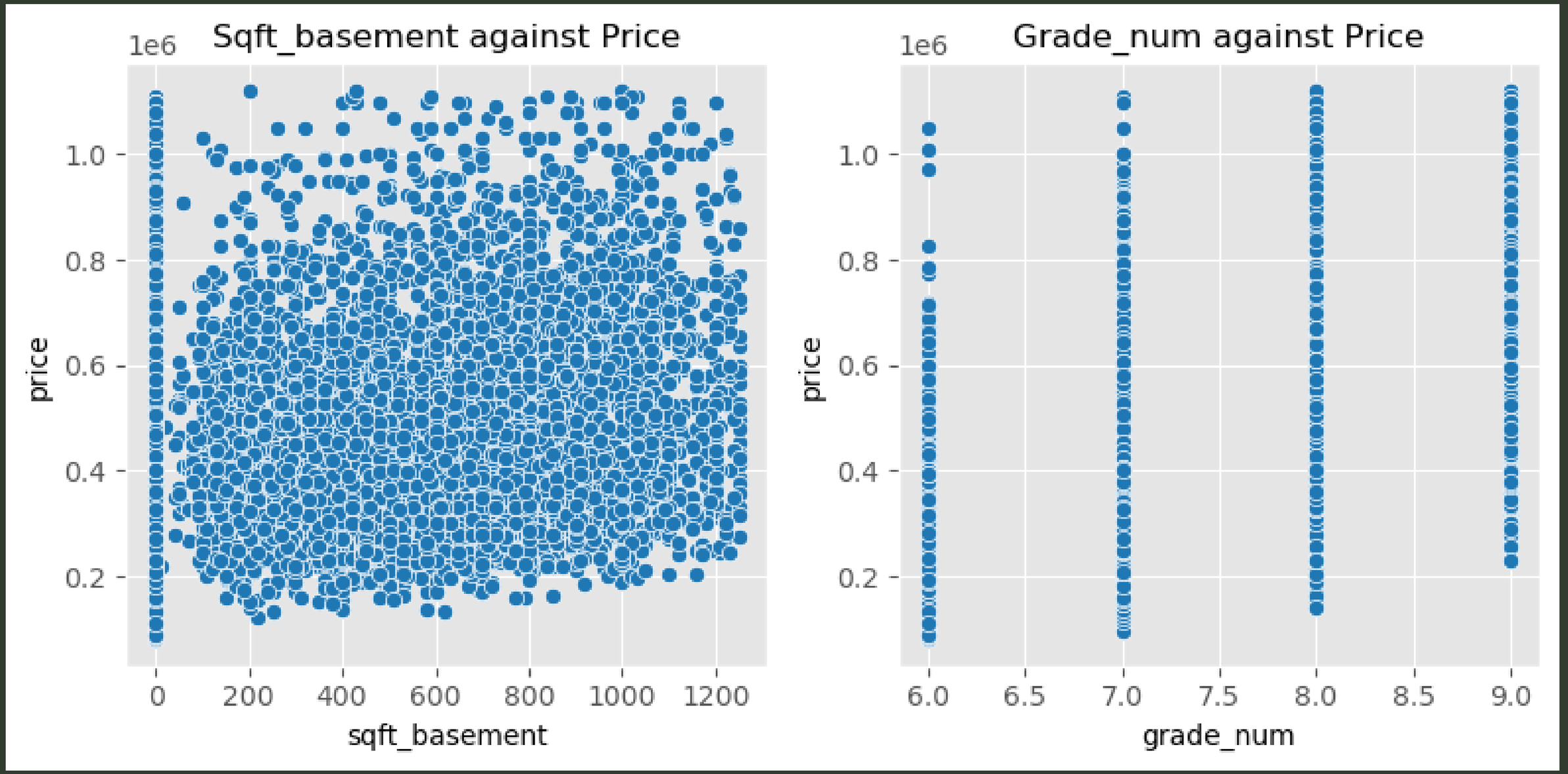


Heatmap of correlation between the variables

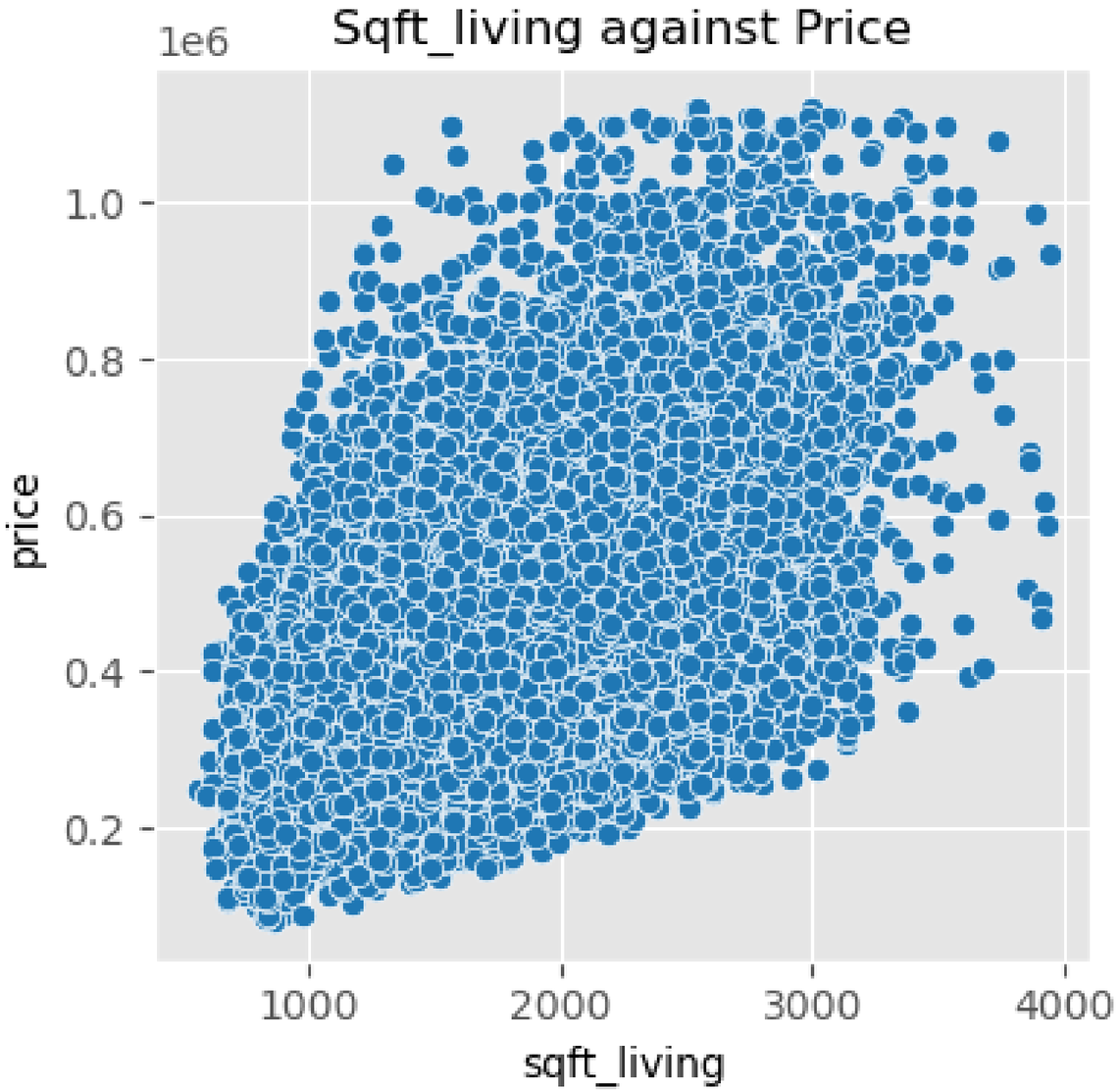
Scatter Plot of Bedrooms and Bathrooms against Price

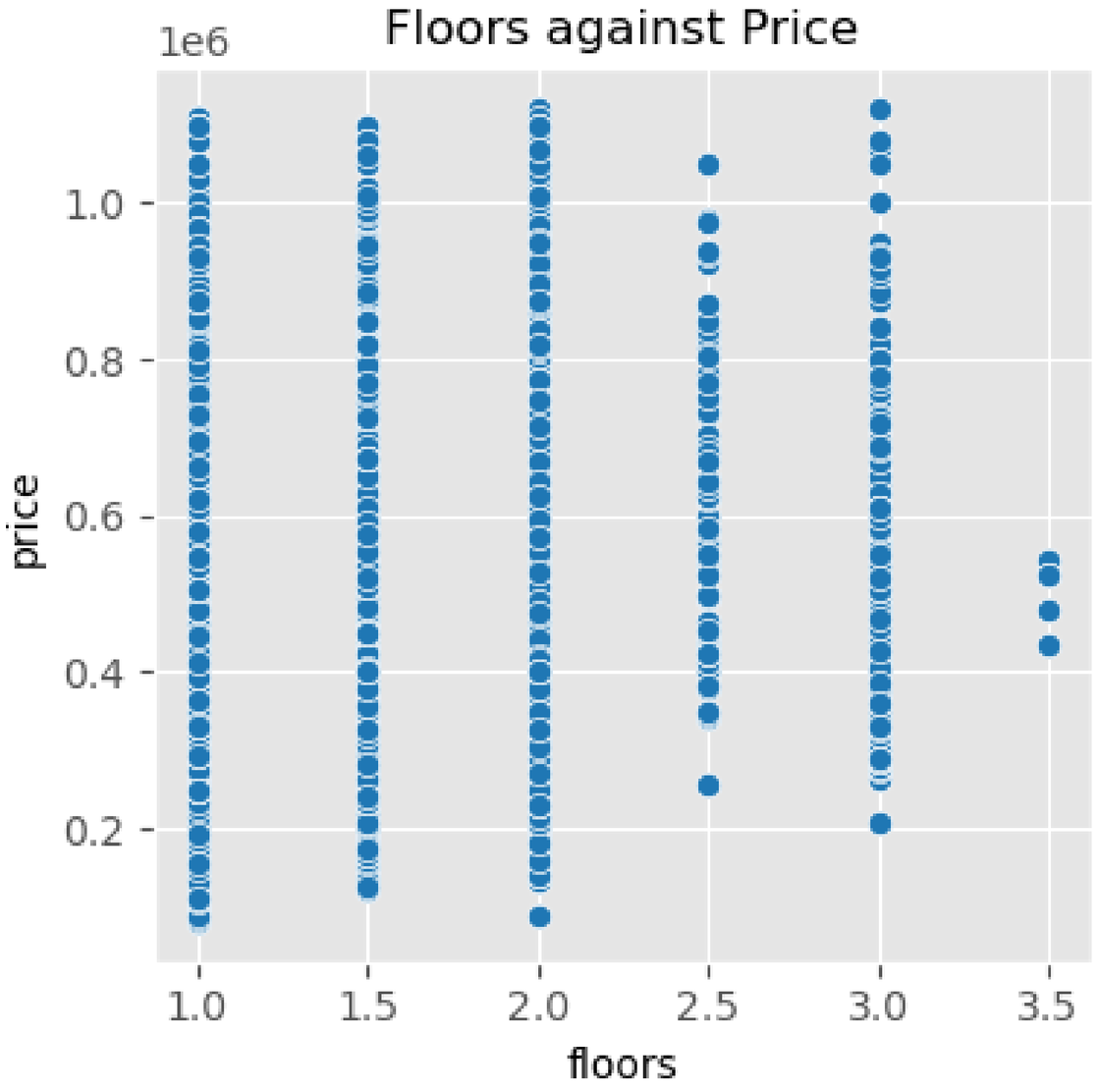


Scatter Plot of Sqft_basment and Grade against Price



Scatter Plot of Sqft_living and Floors against Price





Modeling

The final model was much better than the the baseline as it explains 37% of variations in price as compared to 35% in the baseline model. Also using the root mean square error , our model is off by $143,789 USD in a given prediction which is better than the baselines root mean square error of $148,607.

Model Parameters

The model is statistically significant at an alpha of 0.05.

- **As we increase the number of bedrooms by 1, the house price drops by $18,564.**
- **As we increase the number of bathrooms by 1, the house price drops by $37,731.**
- **As we increase the living space in the home by 1 square foot the price will rise by $101.**
- **As the house floors increase by one, the house price rises by $25,548.**
- **As we increase the basement space by 1 square foot the price rises by $90.**
- **As the grade increases by 1 level, the price rises by $87,890.**

Model Parameters

- The price of a house on a `waterfront` is `$104,200` higher as compared with one that is not.
- The price of a house with an `Average` view is `$88,000` higher as compared to without a view.
- The price of a house with an `Excellent` view is `$213,700` higher as compared to without a view.
- The price of a house with a `Fair` view is `$89,050` higher as compared to without a view.
- The price of a house with a `Good` view is `$85,080` higher as compared to without a view.
- The price of a house with a `Very Good condition is `$98,780` higher as compared to a house with poor condition

Model Assumptions

Normality: The assumptions of normality of model residuals were met, histogram and a qqplot were used to test this assumption.

Linear relationship: Linear Rainbow test was used to test for this assumption and returned a large p-value hence the assumption was met.

Multicollinearity: VIF was used to test for this assumption and the variables had small VIF's hence the assumption was met

Homoskedasticity: Breusch-Pagan test was used to test for homoscedasticity and returned a small p-value hence the assumption was violated. The estimates of the coefficients are still unbiased but the coefficients are not precise.

CONCLUSIONS

- The price model explains 39% of the variations in price.
- The model is off by $143,789 in a given prediction
- Some of the house's top selling points are; waterfront, view, grade, and condition.
- An increase in living spaces, basement space, and the number of floors leads to an increase in the market prices of a house.
- An increase in the number of bathrooms and bedrooms leads to a drop in the price of a house.
- Grade has the most impact on the house price.

Recommendations

- Improving the grade of a house through maintenance and renovation increases the price of a house immensely.
- Houses with large living spaces, basement space, and many floors fetch high market prices.
- The number of bathrooms and bedrooms should be kept on average, that is 3 bedrooms and 2 bathrooms, since as they increase by one the price of a house drops drastically.

Moringa School

Thank You

Presented by: MULEI MUTUKU