King County House Prices

Linear Regression Models

● ● ●

Elizabeth Webster
March 2022

Introduction

This project uses linear regression analysis to understand how certain variables impact housing prices in King County.

We will be building different linear models in order to find the most accurately fitted line and it's coefficients..

Outline

- Stakeholder & Business Problem
- Data
- Methods
- Results
- Recommendations
- Next Steps

Stakeholder & Business Problem

- Real Estate Group in King County

- How zipcodes affect house prices

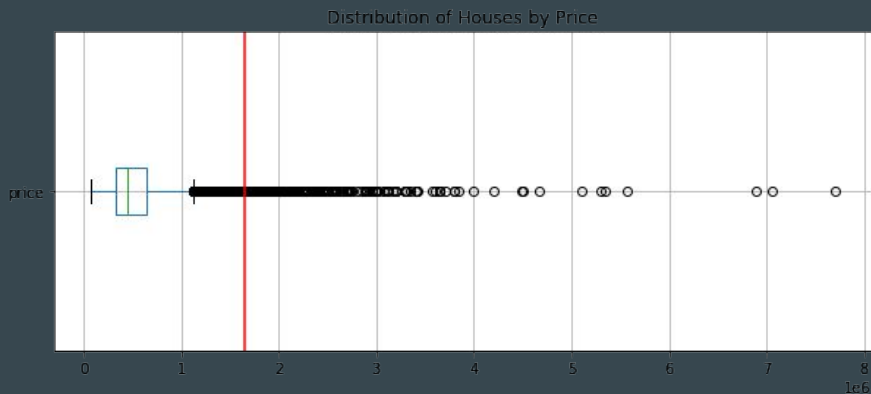- Narrow down house search to specific neighborhoods for potential buyers
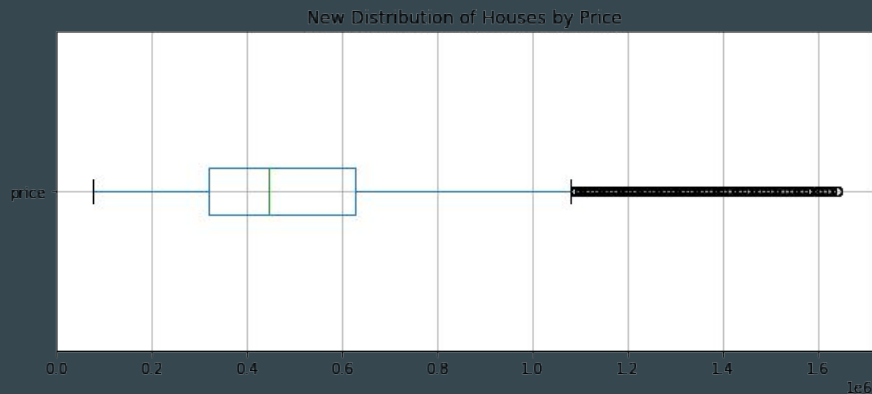
Data

King County House Sales Dataset

- Specifications from 21,597 house sales in King County

- We will use many variables to create our model

- Once we have our model, we will be looking into zipcodes

Methods - Removing Price Outliers

Price distributions with outliers

Price distribution after removing 406 outliers

Distribution of Houses by Price

New Distribution of Houses by Price

Methods - Data Cleaning

- Finding and Filling Missing Data
  - Waterfront and Year Renovated
- Categorical to Numeric Variables
  - Waterfront, Condition, Grade, Zipcode
- New Variable for House Age

Methods - Creating Models

- Baseline Model - Two Highest Correlated Variables
  - Accounts for 46.3% of price
- Model with All Variables
  - Removed variables that were not significant
  - Performed transformations on skewed data
  - Increased our model reliability to 82.4%

Results - Interpreting Coefficients

- Log Transformations
  - Coefficients in terms of percentage, not units
- Zipcodes:
  - High Price Impact - 98112, 98004, and 98039
  - Low Price Impact - 98168, 98030, and 98022
  - Negative Price Impact - 98032

Recommendations

- Client looking to live in a highly desirable neighborhood
  - 98112, 98004, and 98039
- Client more interested in saving money
  - 98168, 98030, and 98022
- Area to avoid
  - 98032

Next Steps

- Relationship between house condition, age of houses, and zipcodes
- Latitude and Longitude data for deeper insight on how location affects house prices

Thank You!

Email: eaw524@gmail.com

GitHub: https://github.com/elizabeth524