King County Real Estate Model

•••

Chris Hollman
March, 2022

Overview

- A real estate group's marketing team has decided to invest in physical advertising (mailings, billboards, etc.) targeting homeowners who are looking to sell
- The marketing team is seeking insight into where to focus these ads.
- The emphasis is to be largely on location rather than individual attributes of a given home.

Project objective:

This project aims to model King County housing data, identifying areas and features that contribute to higher housing prices.

Advertisement Types

Mailings

Identify which areas in King County tend to be associated with higher house prices. As mailings are going to be a focus, there will initially be an emphasis on zipcodes.

Large Scale Ads

Within a given zipcode, determine where to place larger advertisement such as bus ads and billboards.

Examining Data

I looked into sales records for 21,500 houses in King County from 2014 to 2015.

Areas of focus:
- Zipcodes
- Neighborhood attributes

Initial Data Issues

Price Cutoff

- Significant variance in sale prices
- Settled on 1.1 million as limit.

Missing Data:

- Some missing values needed to be filled in or changed

Modeling The Data

Exclusions

- Information on neighboring properties

- Redundant data
- Latitude/longitude
- Data used for new features

Inclusions

- House basics
- Zipcode
- Condition
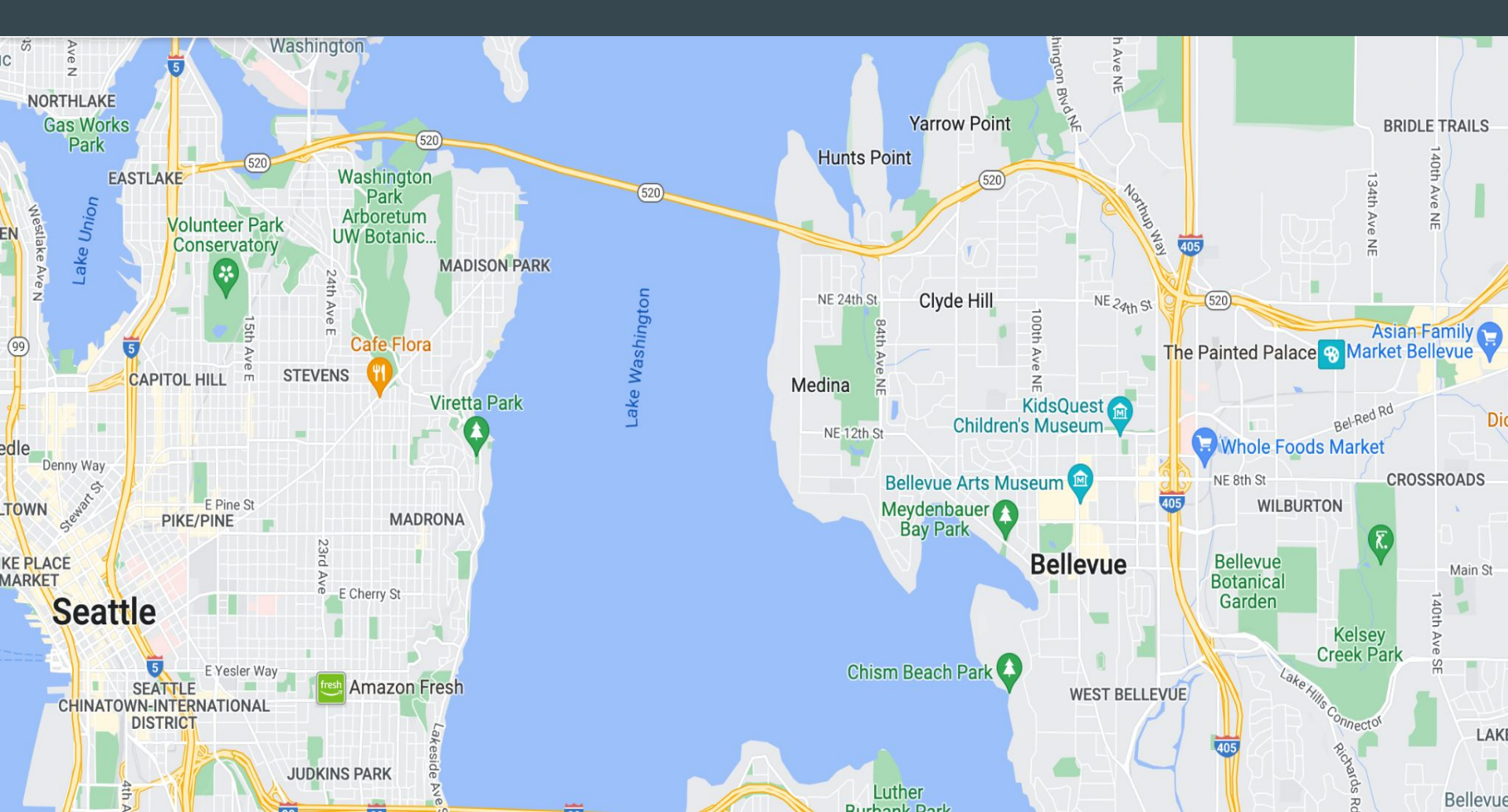- Renovations

Model Findings/Limitations

Limitations

- 84% "accurate"
- 18% average error
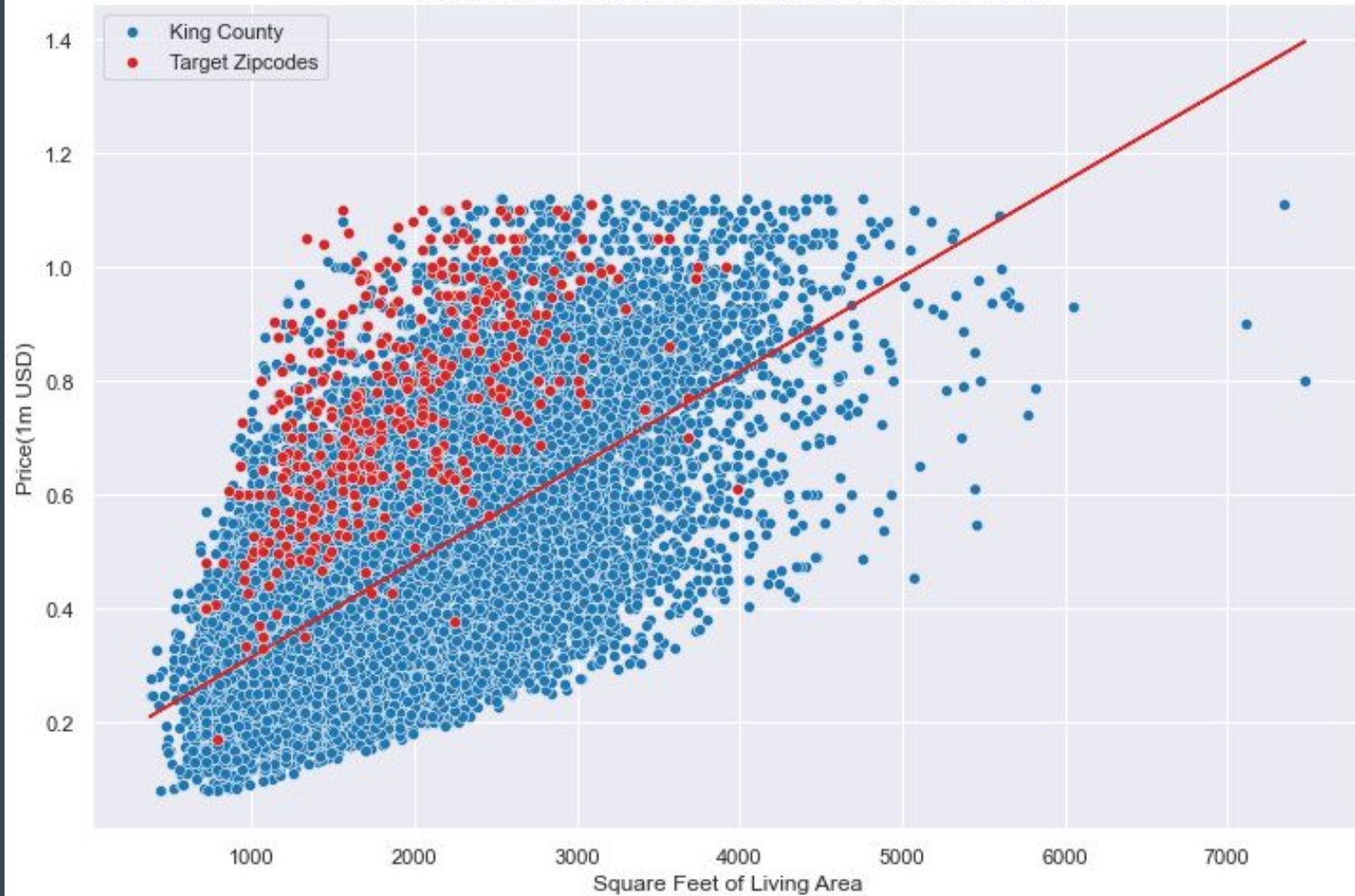- Still influenced by outliers

Findings

- 98004 (Bellevue) +72%
- 98112 (Madison Park/Arboretum) +57%
- 98102 (Capitol Hill) +55%
- Waterfront +36%

Comparison of Price/Square Square Foot By Zipcode

Recommendations

Mailings

Focus mailed advertisements into the top areas from the study.

Larger Ads:

Focus larger advertisements around residential areas with waterfront property.

Next Steps

Latitude/Longitude

Map sales data using latitude/longitude points to gain insight into specific areas that are associated with higher price. Possible feature engineering tied to this concept.

Turnover Rates

Assess which zipcodes are associated with more sales per year. Areas with high turnover rates provide more opportunities for agents to earn commissions.

Improving Model

Model accuracy and insights could be made more reliable by looking deeper into outliers

Thank You!

Please feel free to ask any questions now.

You may also reach me via email:

Chris Hollman
chollman91@gmail.com