Real Estate Statistical Analysis

Lirui (Leroy) Guan

October 22, 2022

Overview

Statistical analysis of house data guides the evaluation the home sale price in King County, Washington by the most important physical features of the property:

- Areas that highly affect house prices.
- Overall house grade, on waterfront, excellent view that positively impact sale price
- Overall house grade and private sewer system that negatively impact sale price

Business Problem

- Fisher Real Estate Investments wants to invest properties in King County.

- King County is located in Washington State. The population was 2,269,675 in the 2020 census.

- Property evaluation is preliminary step for real estate investment. This involves evaluating the property location, physical features of property, the history of property.

King County Houses Data

- The dataset contains more than 30k houses data for 2021-2022 home sales in King County, Washington

- The data includes the house sale price, address, physical features, such as lot size, house size, bedrooms, bathrooms, grade, etc
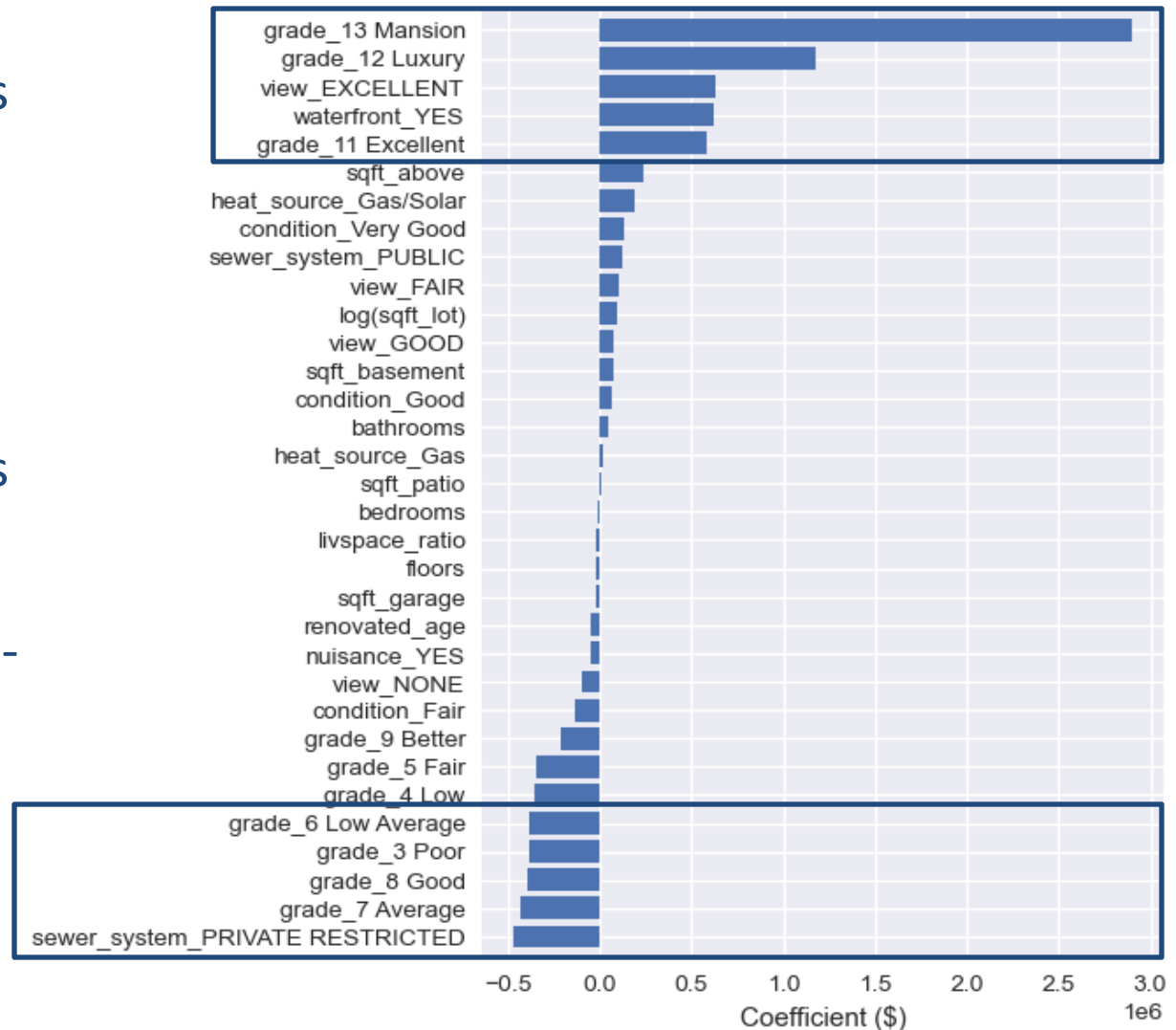
Multiple Linear Regression Model

- The multiple linear regression model uses sale price as target and various house features as predictors
- The model overall is statistically significant and can explain about 72.5% variance in sale price
- The intercept and most predictors are statistically significant

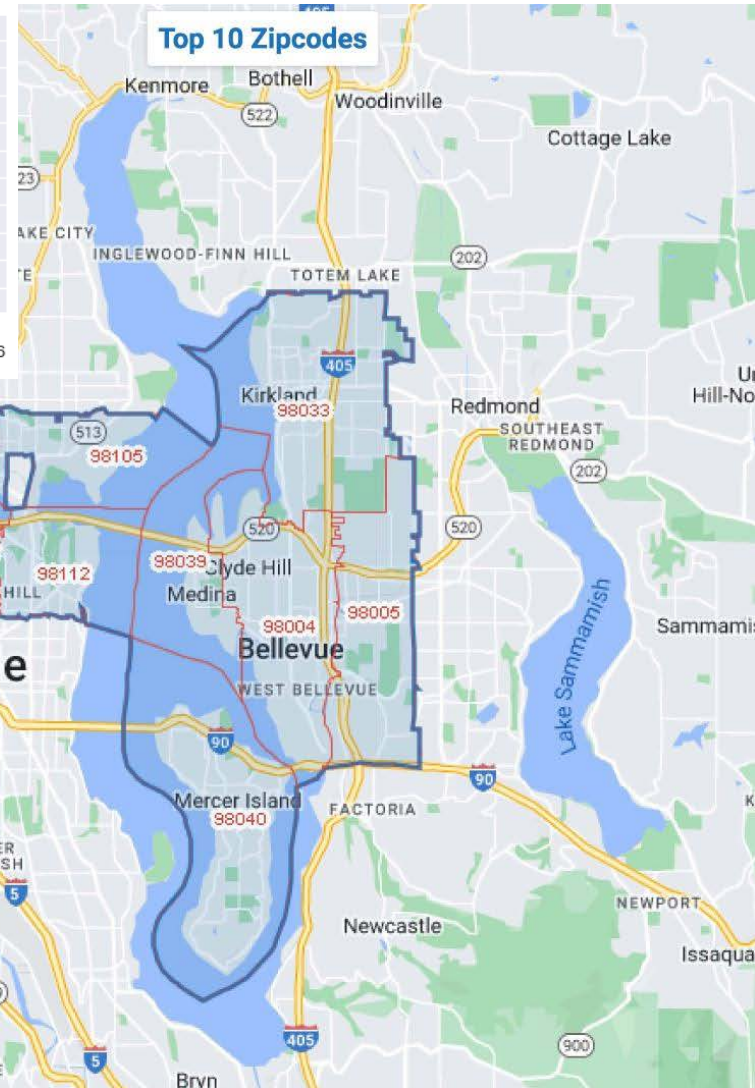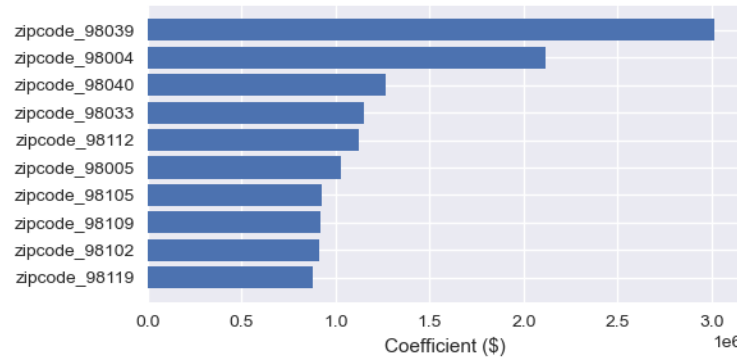| Labels	Adjusted R^2	Regression coefficients	Standardised regression coefficients	p-values
Overall	72.50%			0
Intercept		-382,400	937,200	0

Physical features

- Top 5 most important features that positively impact sale price (Coefficients from $0.6M to $2.9M)

- Top 5 most important features that negatively impact sale price (Coefficients from -$0.4M to -$0.5M)

Top 10 zip codes

Conclusion

- Houses located in the top 10 zip codes

- Excellent or higher grade houses with excellent view and on waterfront positively impact price

- Lower grade houses and those with private restricted sewer system negatively impact price

Next steps:

- Source addition data such as perceived safety of a neighborhood and the number of services or amenities nearby.

Thank you!

Email: liguan698@gmail.com

GitHub: @lerguan

LinkedIn: linkedin.com/in/leroyguan/