

A scenic view of the Seattle skyline with snow-capped mountains in the background and a body of water in the foreground. The text is overlaid on the image.

DANIEL ROSS-LEUTWYLER, SEPT. 29, 2021

MODELING HOUSING PRICES IN KING COUNTY, WA.

Using Linear Regression to Describe House Prices

HOUSING SNAPSHOT

- Pop. 2.25M
- 16.6% Population Increase (2010-2019)
- \$779k median home price (2021)
- 43.9% Home Ownership Rate

King County Housing Market Trends

Median Sale Price

\$779,000

+13.5% year-over-year

of Homes Sold

3,850

+7.7% year-over-year

Median Days on Market

6

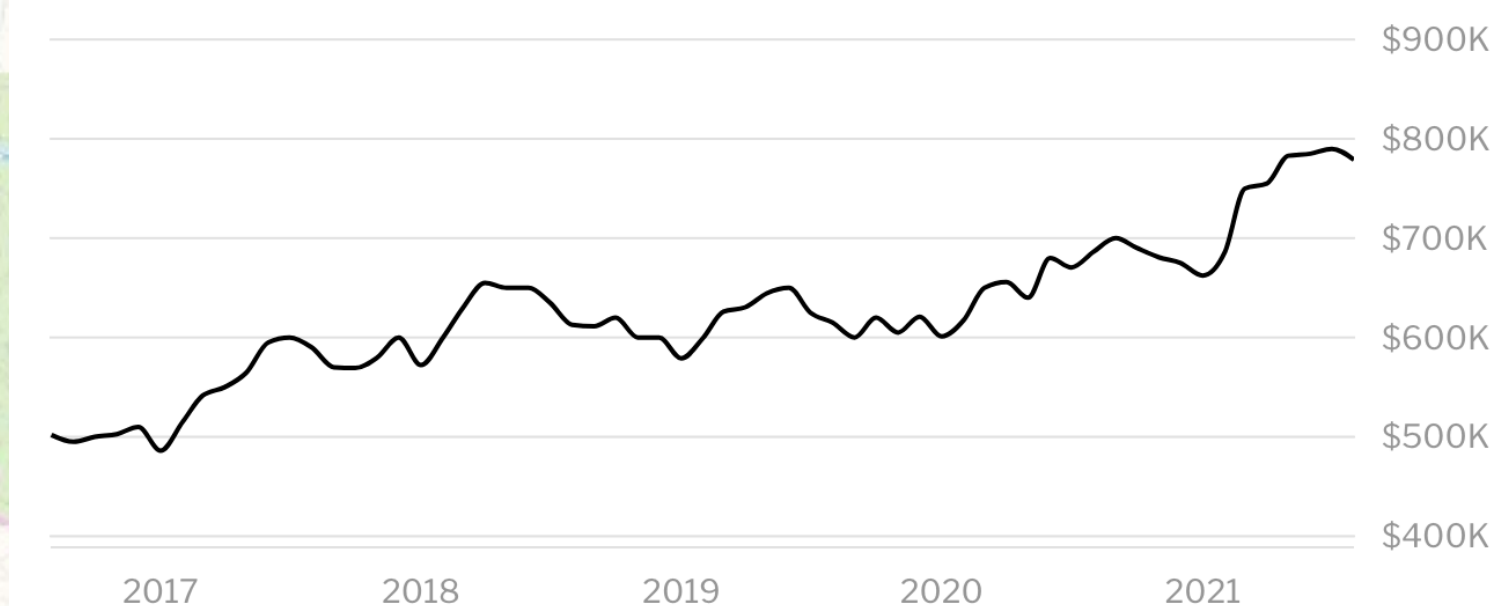
-14.3% year-over-year

All Home Types

1 year

3 years

5 years



Source :<https://www.census.gov/quickfacts/fact/table/kingcountywashington,WA/PST045219>
<https://www.redfin.com/county/118/WA/King-County/housing-market>

As a small real estate development company, we advise homeowners how to best make improvements to their property to increase home value,

and

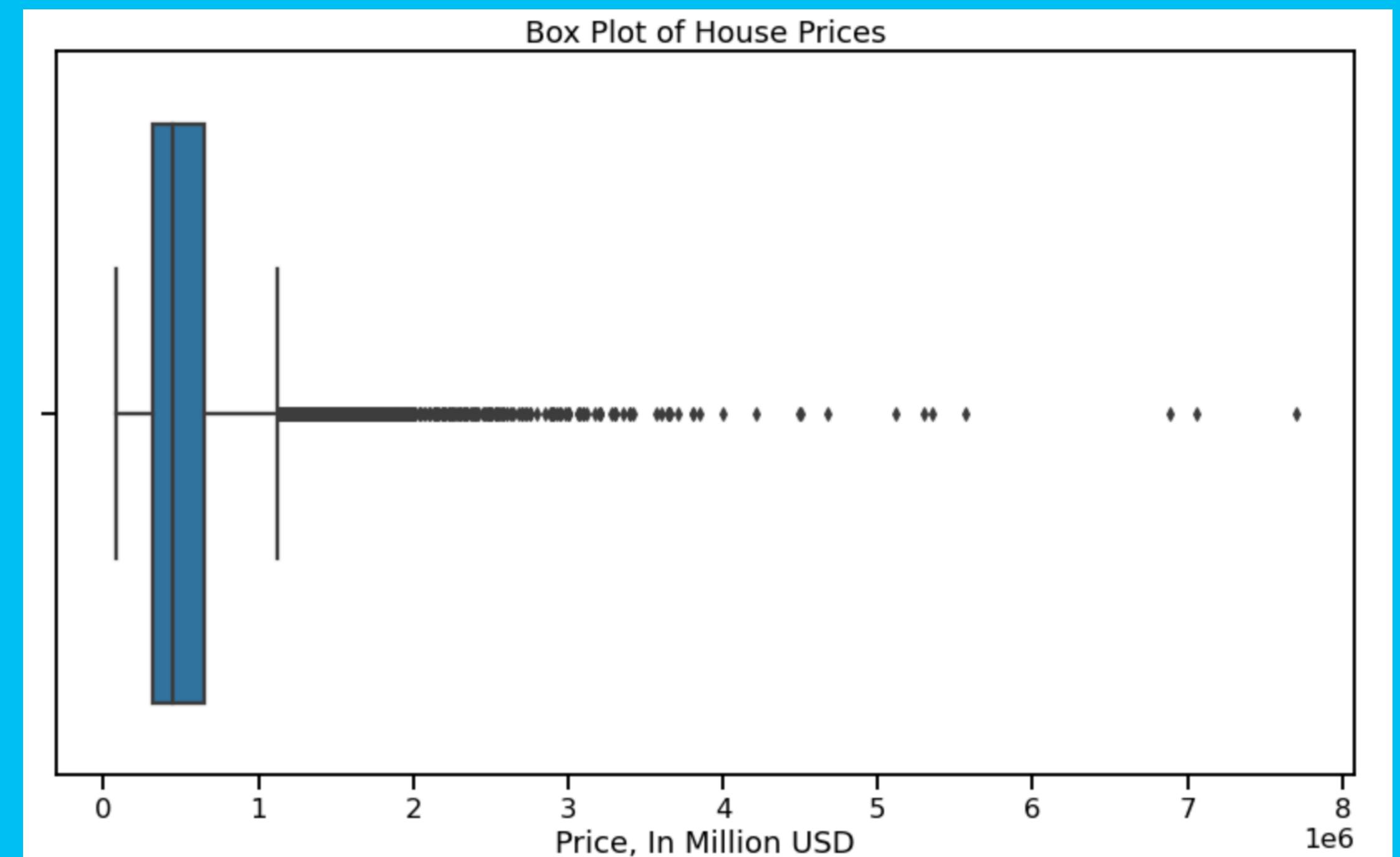
Work with small developers to identify investment opportunities in emerging housing markets.

SUMMARIZING THE KC HOUSING DATA SET

- Almost 21,600 entries.
- Contains all housing sales from May 2014 - May 2015.
- Includes data about house size, bedrooms, bathrooms, grade, condition, location, and lot size.
- Mean house price is \$540,296
- Median house price is \$450,000
- Most expensive house sold for \$7.7 million.
- Least expensive house sold for \$78,000.

SHORTCOMINGS

- Data is 5 years old.
- Some variables are missing a lot values.
- Price info contains a lot of outliers that skew data



INTRINSIC FEATURES

Immutable and unique to each house

- Waterfront view
- Location
- Age
- Size of lot

INCIDENTAL FEATURES

Able to be changed and improved

- Number of bedrooms and bathrooms
- Materials used for construction
- Size of house

MODEL PERFORMANCE:

The elimination of outliers reduced our data set to just over 11k entries. While working with fewer entries reduces our accuracy (as measured by R Squared) it helps satisfy the assumptions necessary to validate a linear regression model.

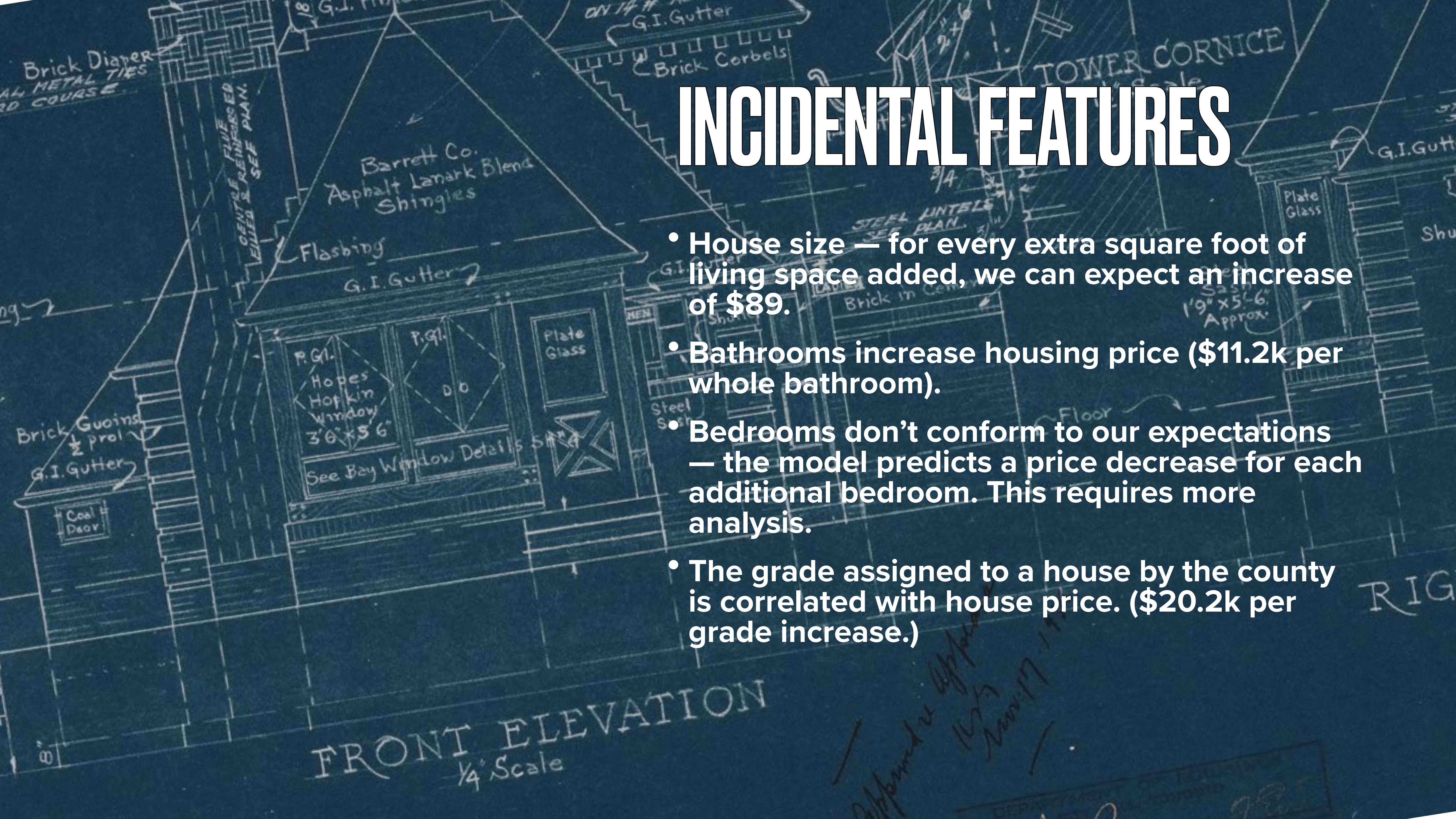
The Mean Absolute Error (MAE) between our training and testing data was very small — with less than a 1% difference.

Our R Squared value, indicating how much of the change in price is accounted for by the change in our independent variables is a respectable .67.

We have elected to keep our results as interpretable as possible. The only transformation done on the data has been to scale it to eliminate errors of magnitude between variables.

INCIDENTAL FEATURES

- House size — for every extra square foot of living space added, we can expect an increase of \$89.
- Bathrooms increase housing price (\$11.2k per whole bathroom).
- Bedrooms don't conform to our expectations — the model predicts a price decrease for each additional bedroom. This requires more analysis.
- The grade assigned to a house by the county is correlated with house price. (\$20.2k per grade increase.)



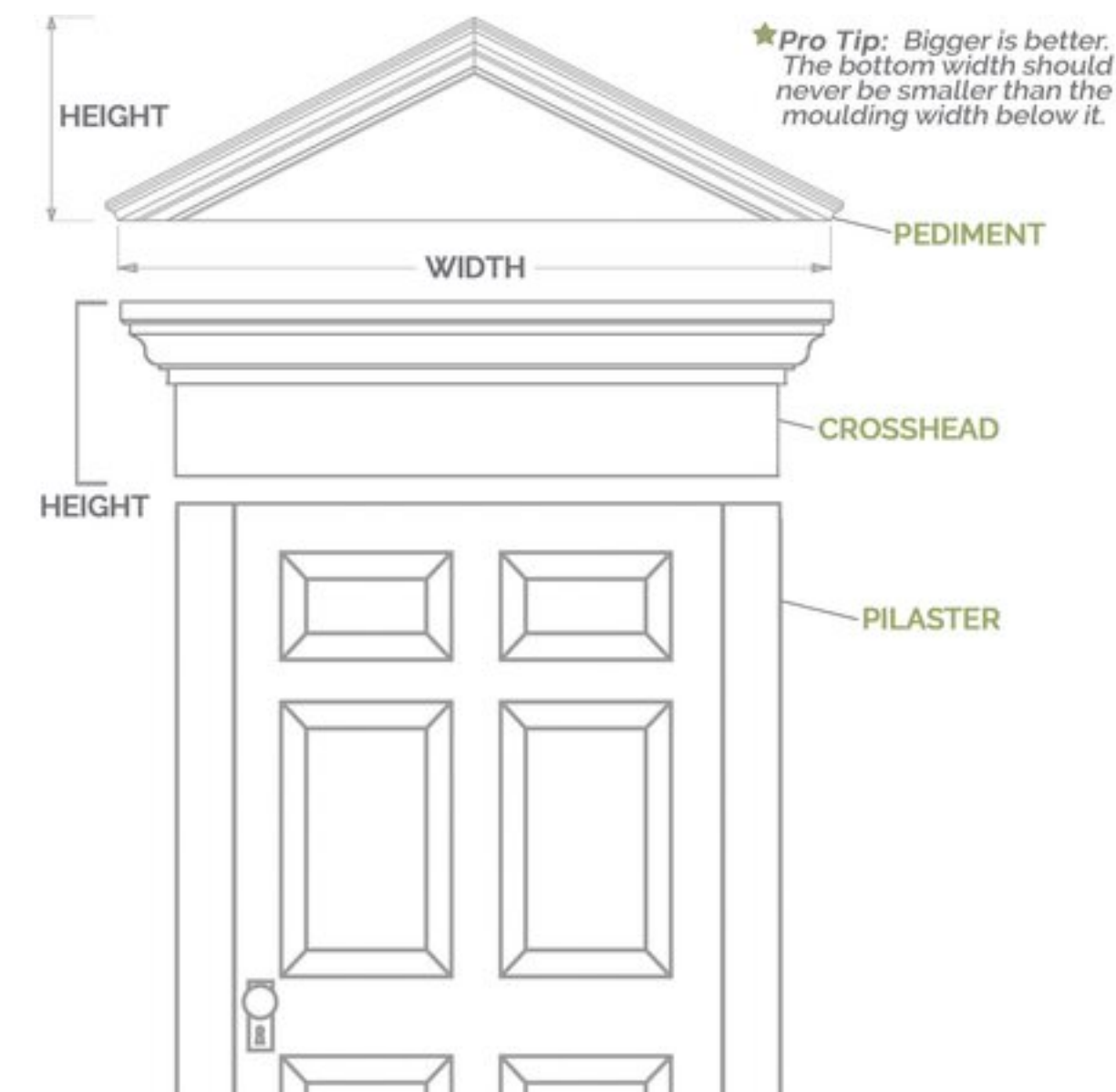
INTRINSIC FEATURES AKA, LOCATION! LOCATION!, LOCATION!

- Houses with waterfront views sold for an average of \$209k more than houses without a waterfront view.
- Houses in the 98039, 98004, 98112, and 98040 zip codes have a very high house price compared to other zip codes.
- Further investigation in adjacent zip codes could yield potential deals as demand for housing in exclusive zip codes grows.



RECOMMENDATIONS:

- Increase the square footage of your house by building a whole or half bathroom.
- Make improvements in the quality of building materials and craftsmanship whenever possible. You will have to get a reassessment of the grade of your house, but the upside is tremendous.
- If looking for investment properties, look in the aforementioned zip codes, and adjacent zip codes, for houses in need of renovations where improving the grade is relatively painless.





THANK YOU!

Full Jupiter notebook and dataset available at:
https://github.com/dan-rl1976/kc_housing_regression