Phase 2 Project

Brad Blackwood, Ashley Eakland, Jalen Charles

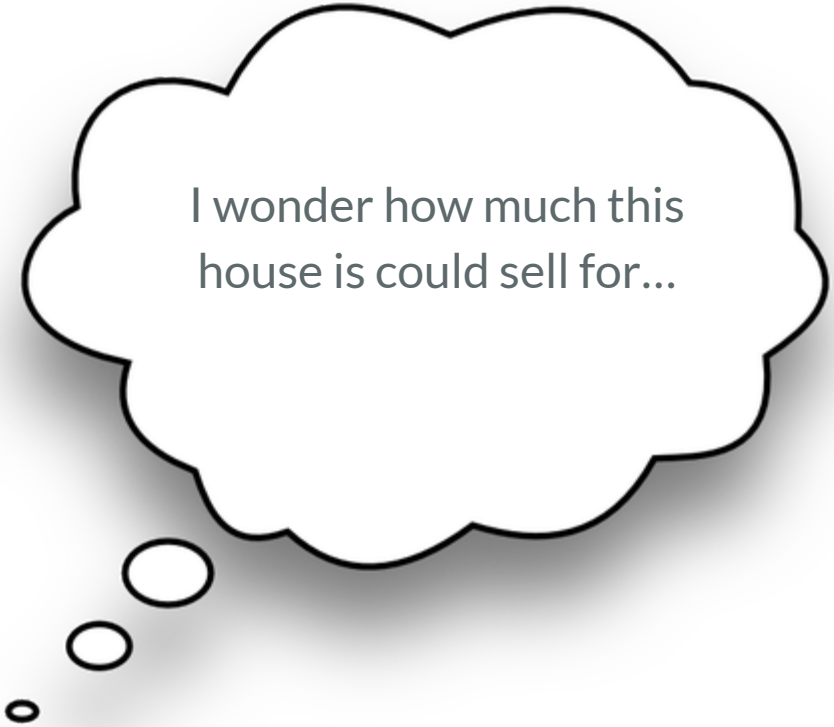
Welcome Home Buyers and Sellers!

Thanks for joining us to hear about how our user-friendly tool can help you predict the price of a home with just a few key metrics.

Agenda

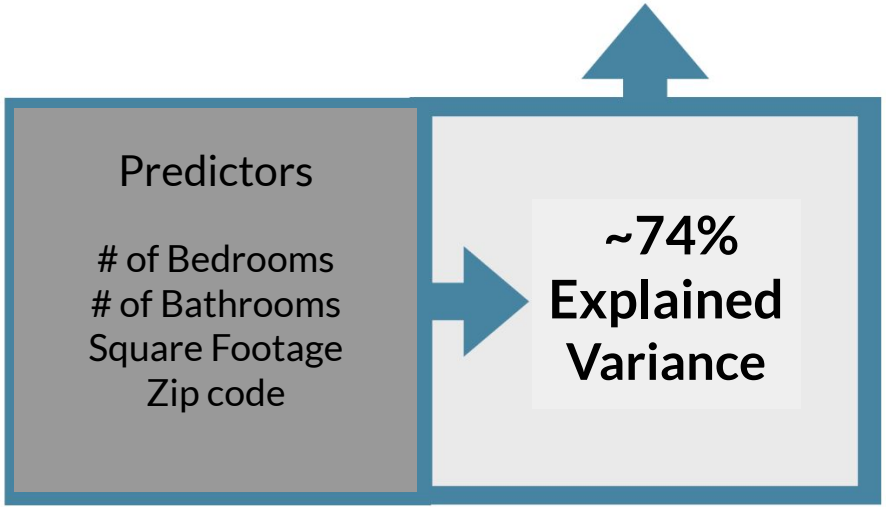
- ❏ Business Understanding
  - ❏ Data Overview
    - ❏ Modeling
      - ❏ Suggestions
        - ❏ Future Research

Have you ever thought to yourself:



Our user friendly tool can help!

Bottom Line


Predictors
# of Bedrooms
# of Bathrooms
Square Footage
Zip code
~74%
Explained
Variance

Data

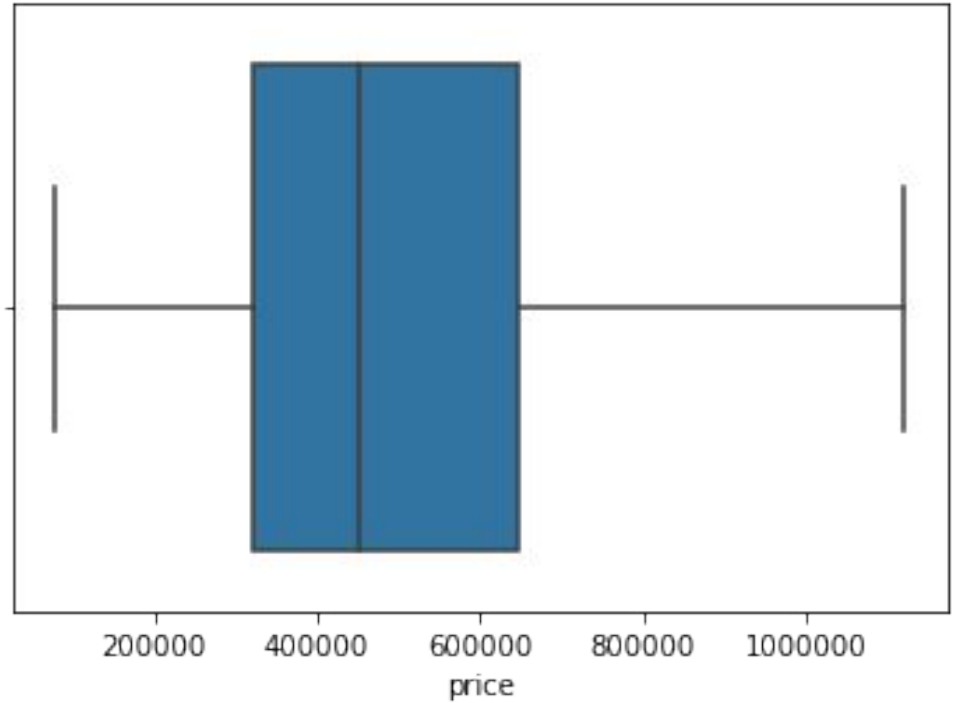
King County Home Sales Records

- May 2014 through May 2015
- 21,500+ home sales records
- 70 Unique Zip Codes

Exploring The Data

Target Metric: Price

- ❏ Mean Sale Price: $540,510
- ❏ Median Sale Price: $450,000

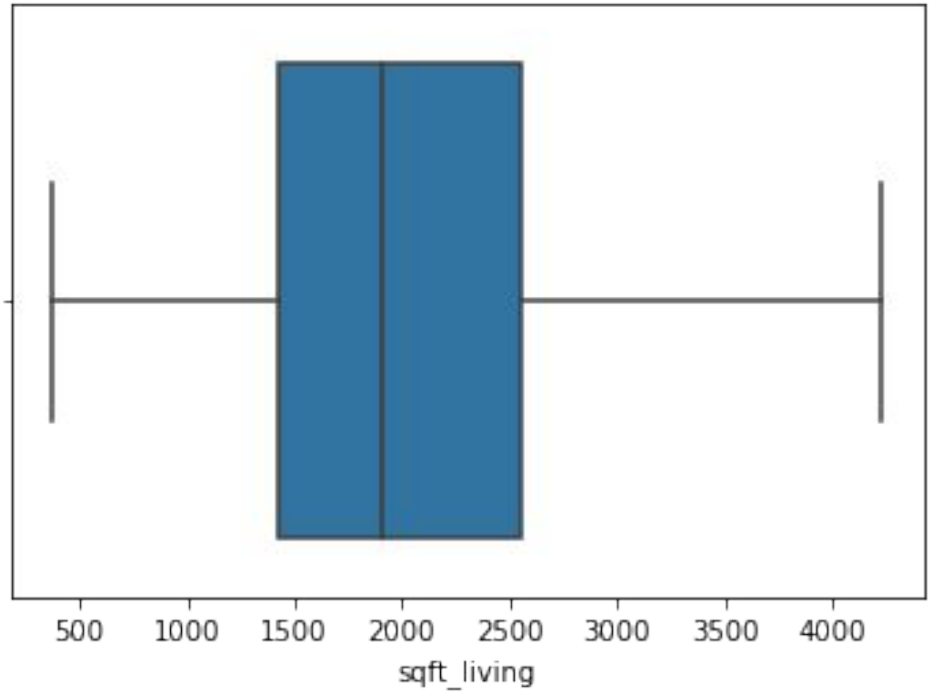


**Outliers are omitted from plots for clarity*

Exploring The Data Cont.

Square Footage

- ❏ Mean: 2080 ft²
- ❏ Median: 1910 ft²
- ❏ Mode: 1300 ft²

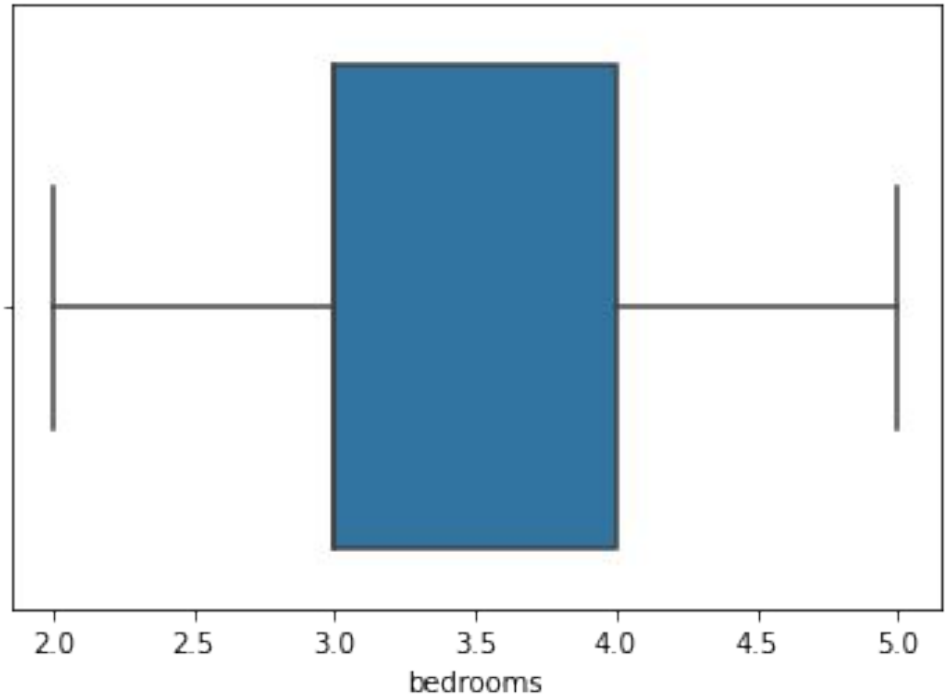


**Outliers are omitted from plots for clarity*

Exploring The Data Cont.

Bedrooms

- ❏ Mean: 3.37
- ❏ Median: 3
- ❏ Mode : 3

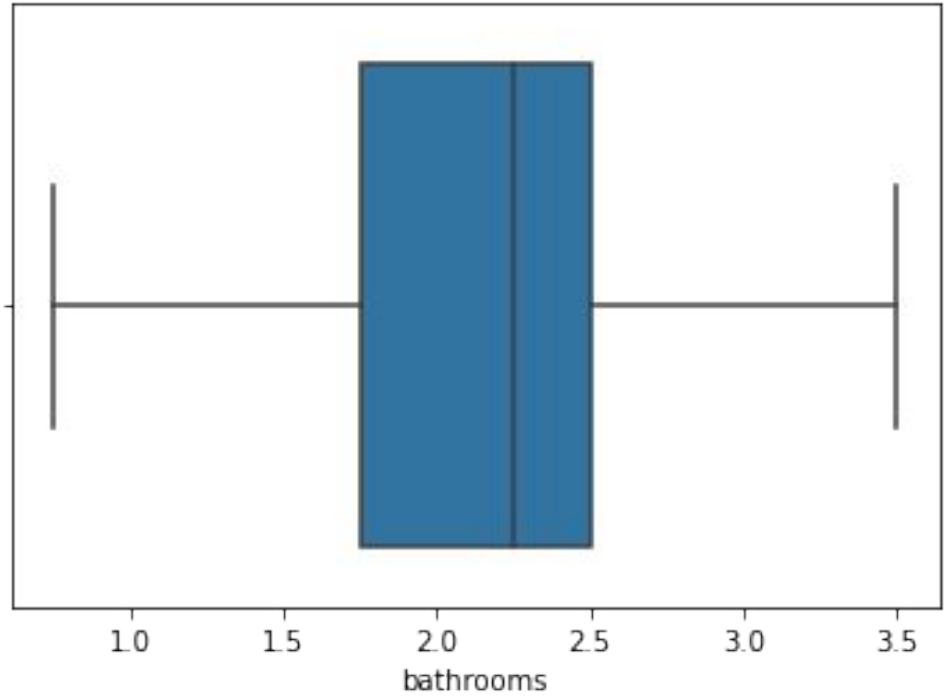


*Outliers are omitted from plots for clarity

Exploring The Data Cont.

Bathrooms

- ❏ Mean: 2.12
- ❏ Median: 2.25
- ❏ Mode: 2.5

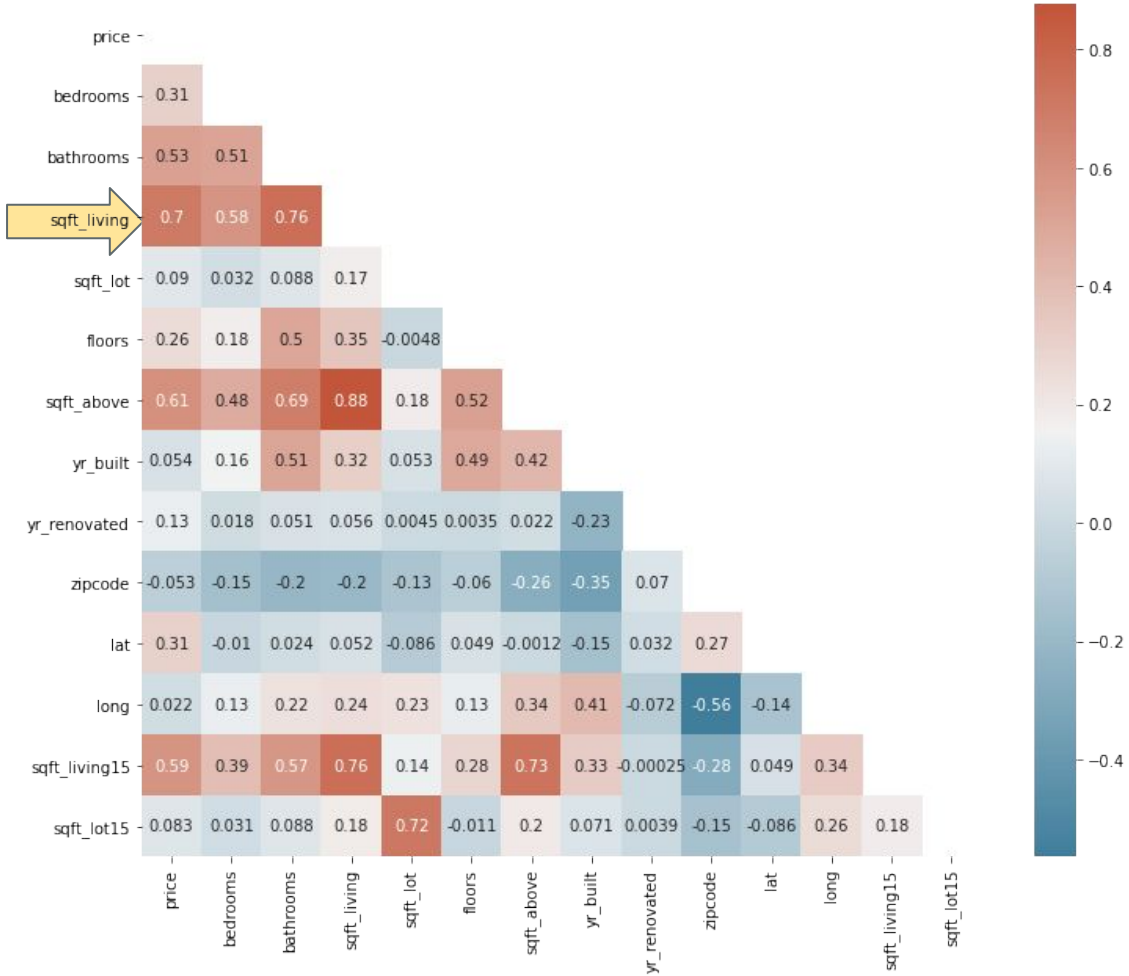


*Outliers are omitted from plots for clarity

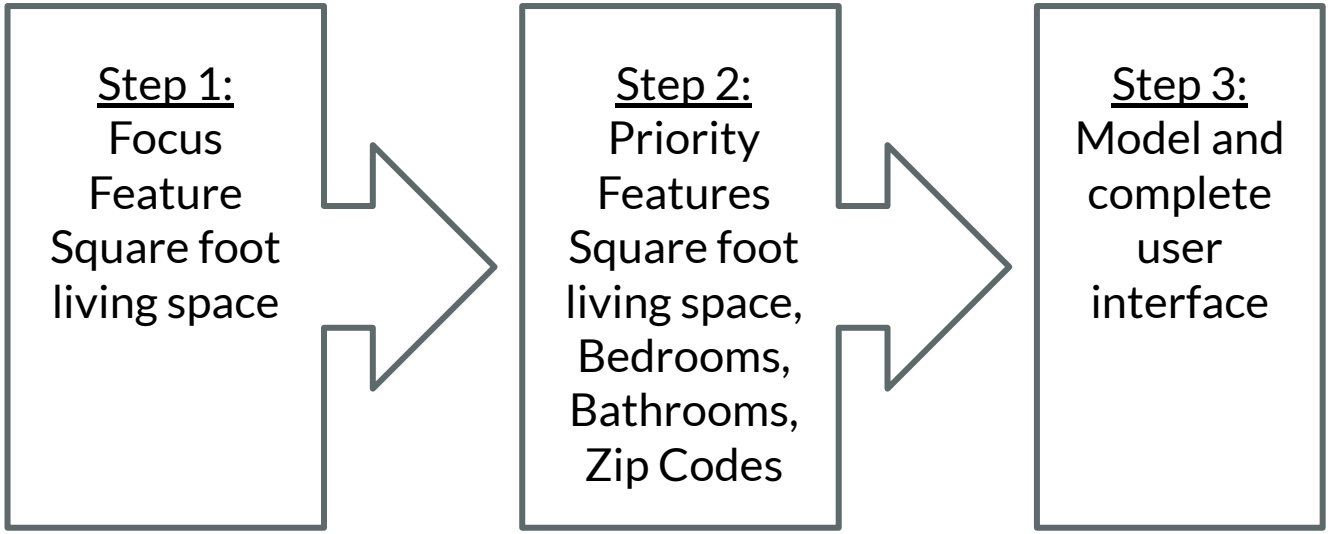
Correlations

Our Objective is to predict home ‘price’.

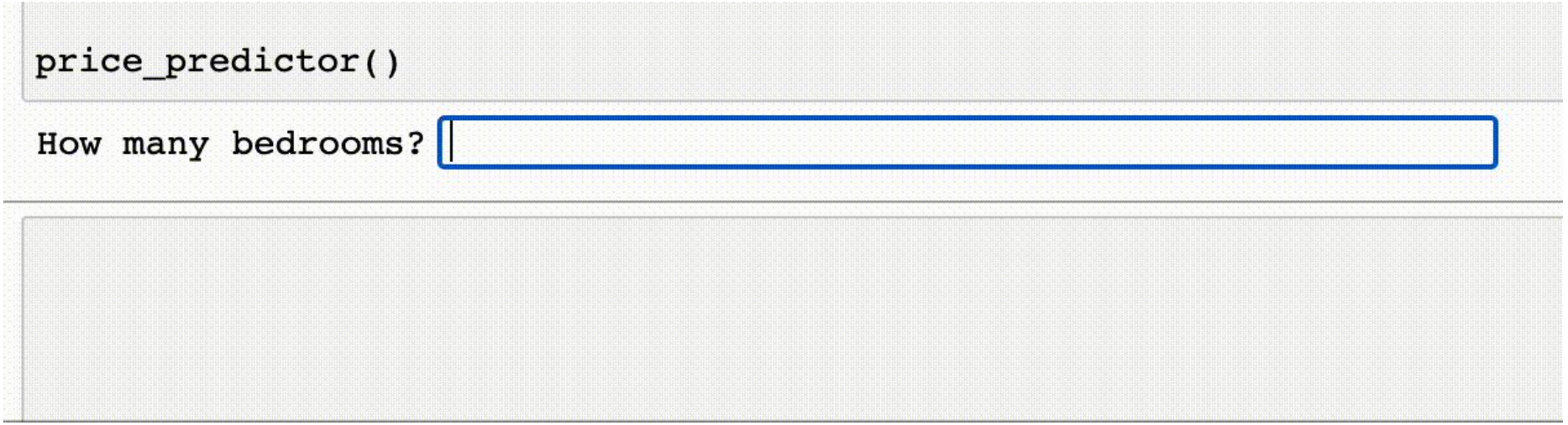
Square footage is the strongest predictor of price.



Modeling Process

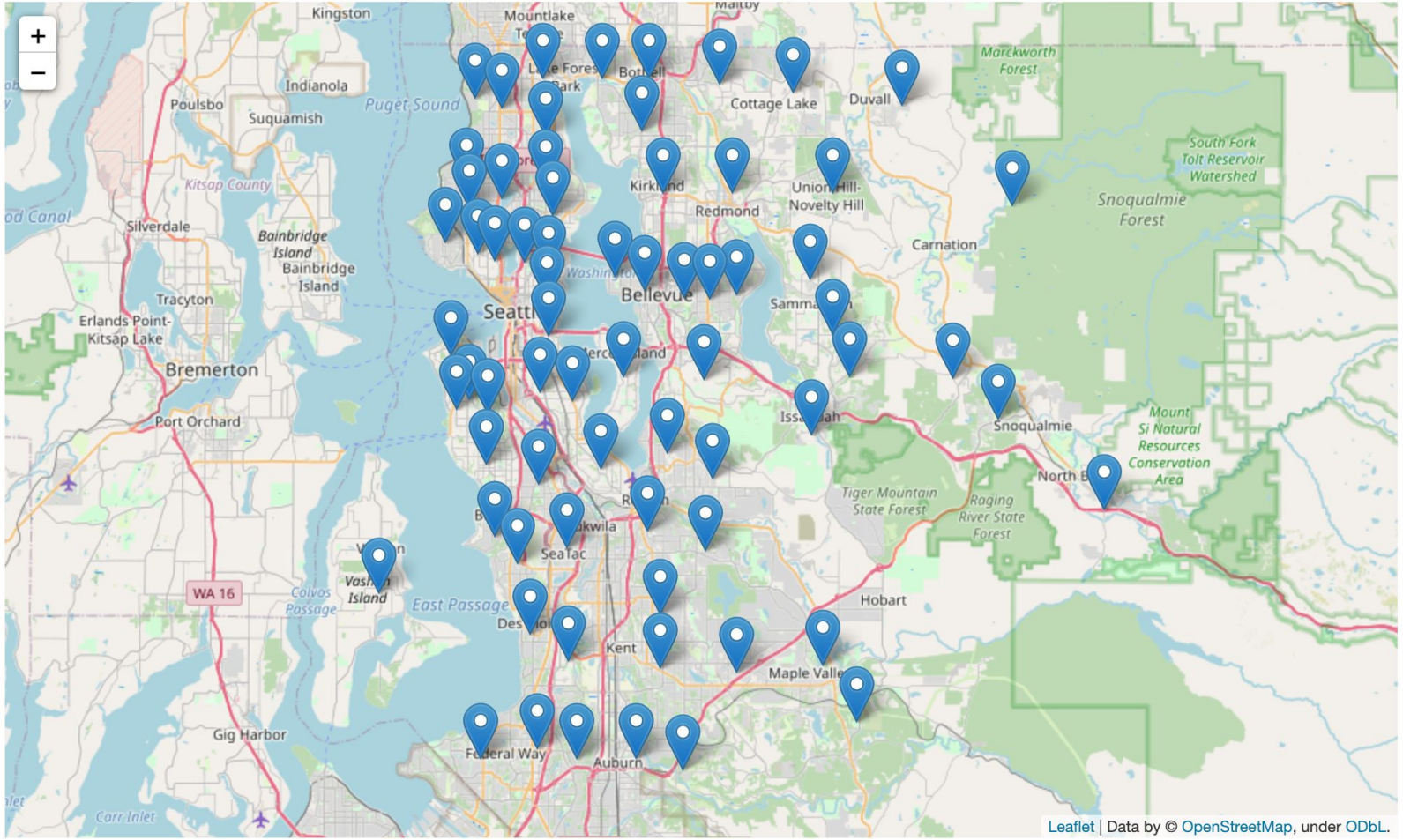


Visuals

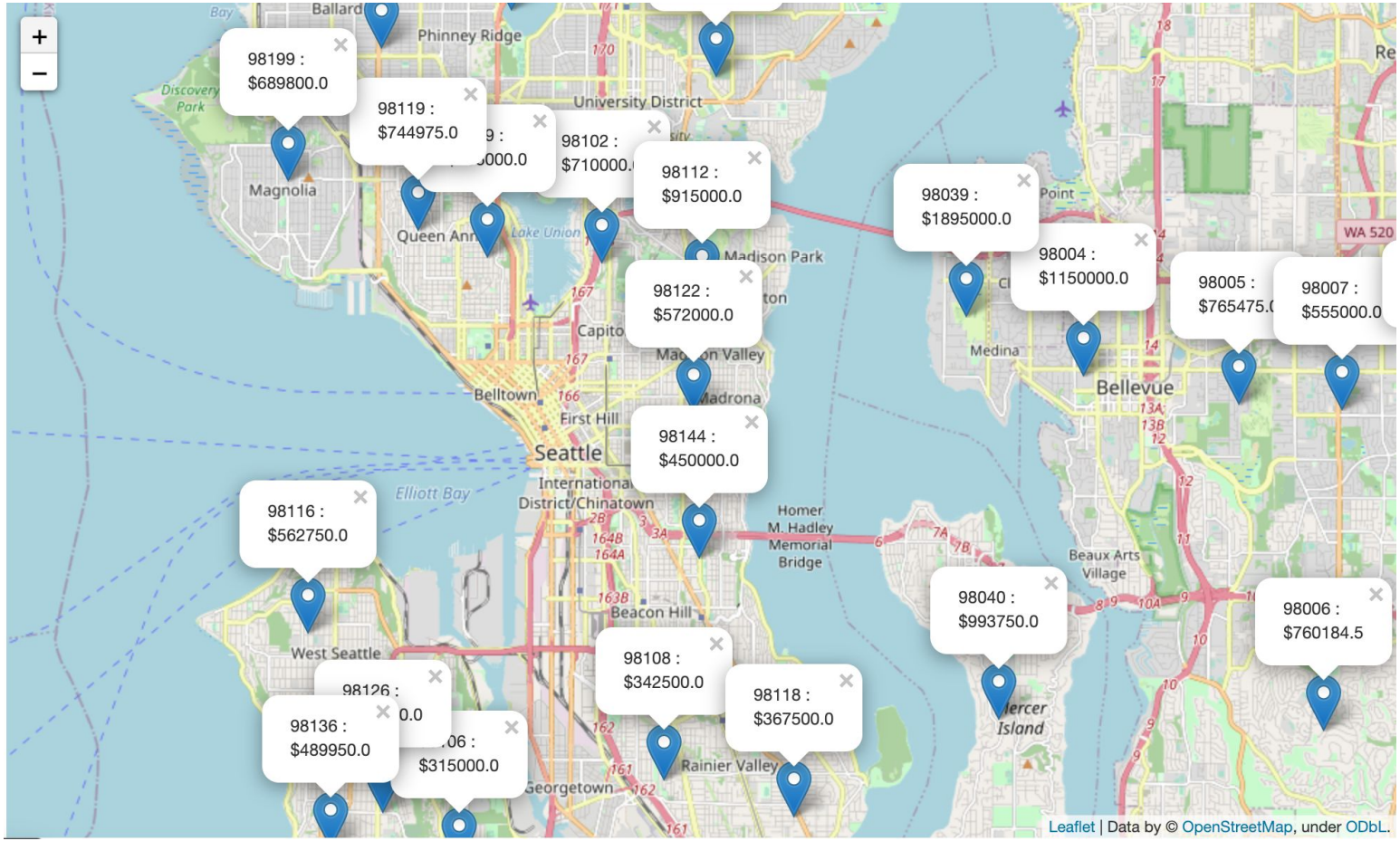


Simple user input fields allow quick predictions of home value to be produced based on key metrics.

Interactive Map of Median Home Prices in King County, WA by Zip Code



Interactive Map of Median Home Prices in King County, WA by Zip Code (zoomed in with flags showing)



Location, Location, Location!

Zip code value is extremely variable:

- For example:
  - Zip Code 98039 = $1,335,000
  - Zip Code 98059 = - $1962

~ +$275 per square foot living space

~ +$10,275 for each additional bathroom

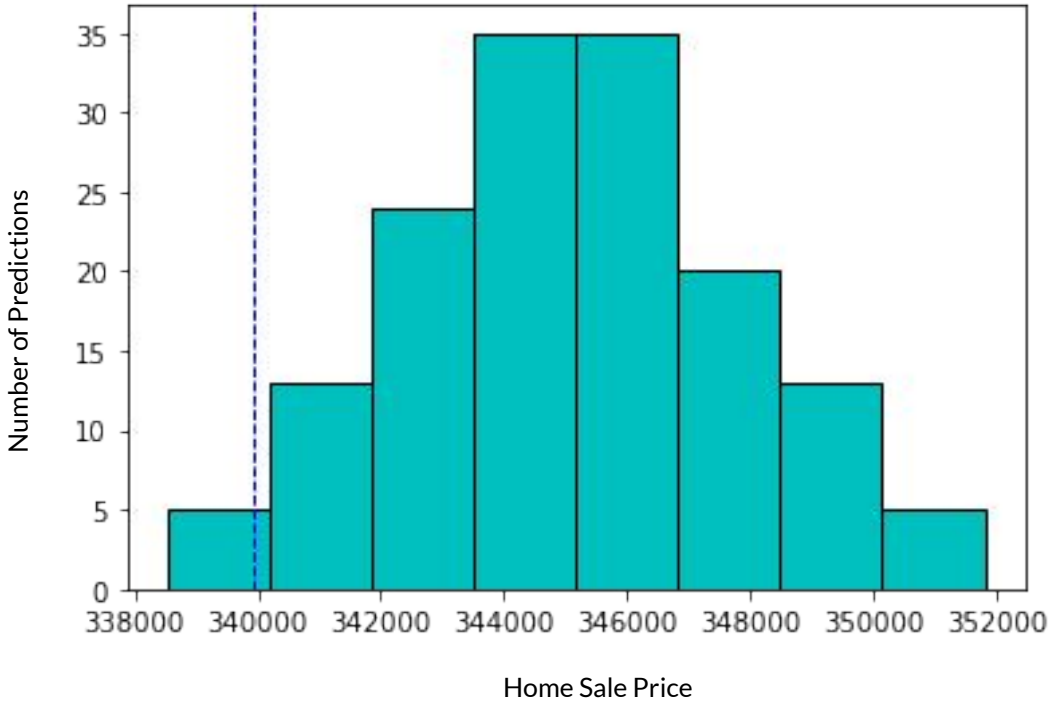
~ -$47,700 for bedrooms

Model Performance

Model Training is precise.

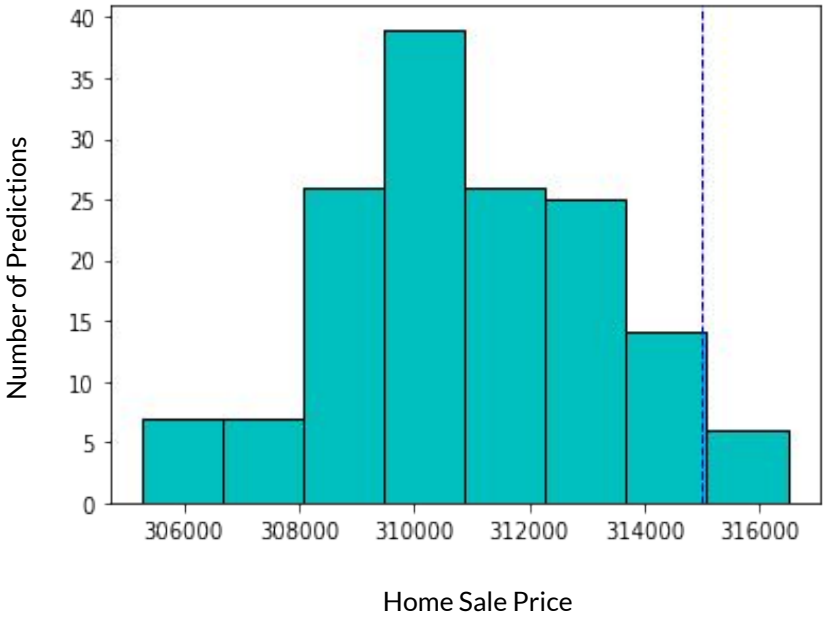
- ❏ It performs consistently
- ❏ Normal distribution
- ❏ Predictions generally are in range of observed values

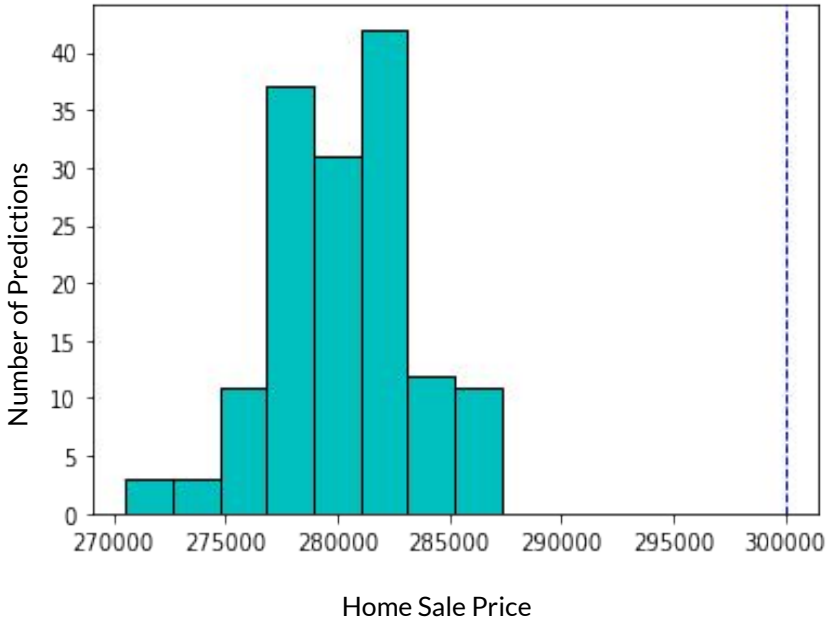
*Blue Dotted Line = Actual Value



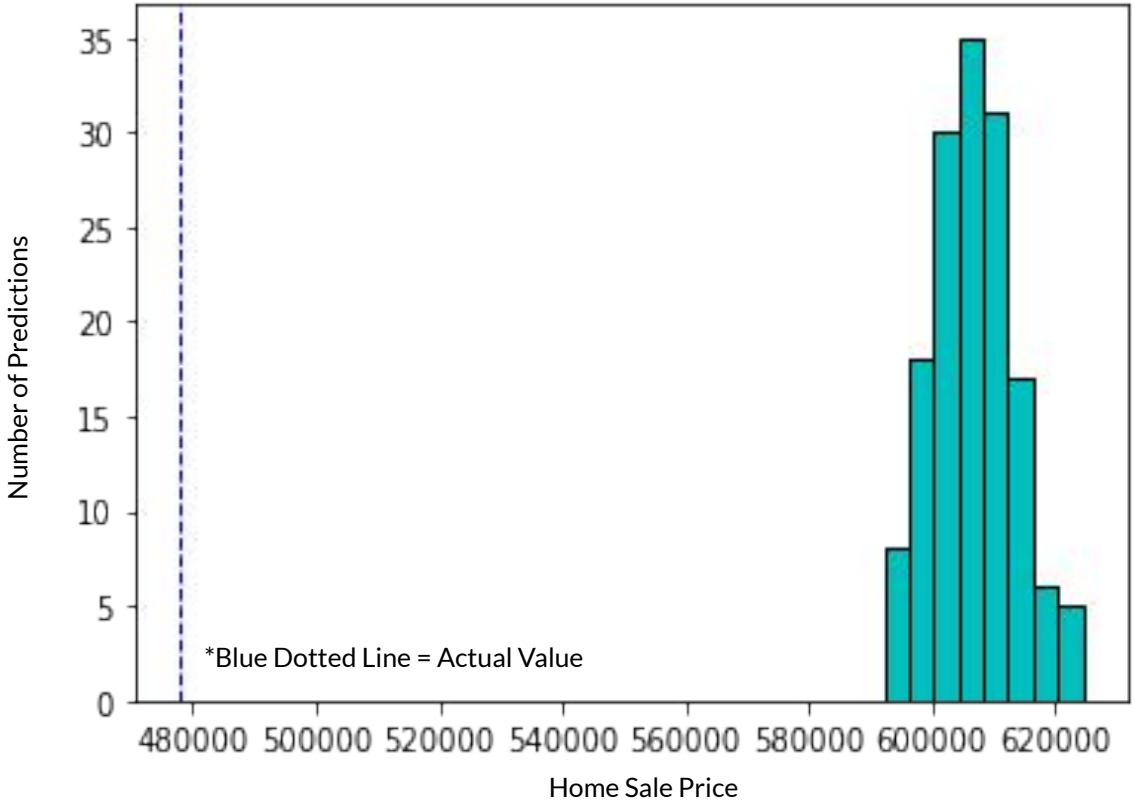
Model Performance Cont.

*Blue Dotted Line = Actual Value





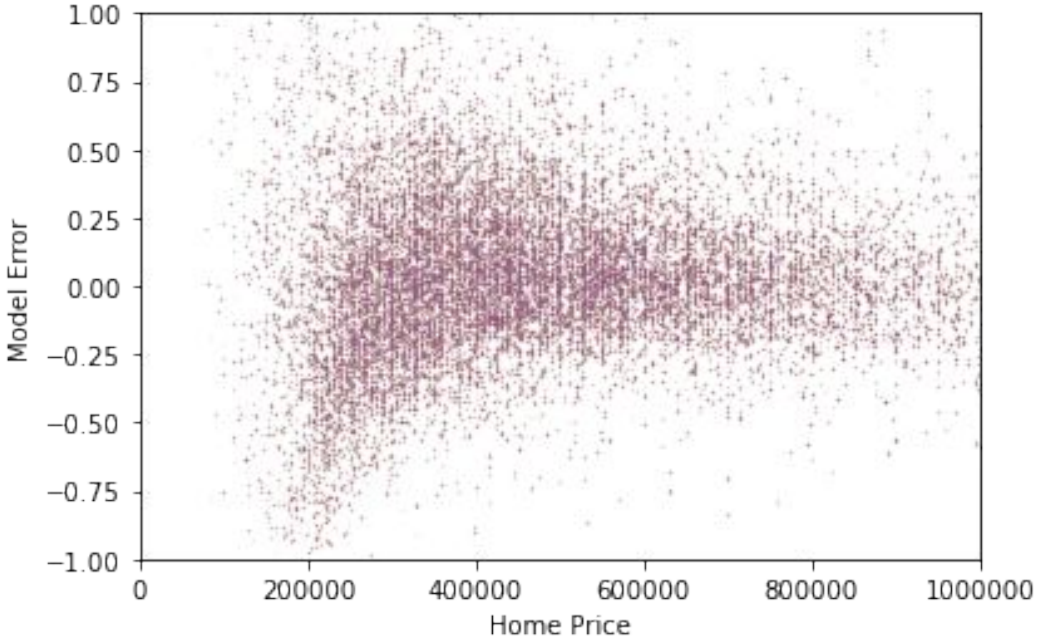
Model Performance Cont.



Future Analysis

While the model performs well at and above the median of overall home sales, we see a decline in accuracy on lower priced homes.

To be used with discretion.



Future Analysis - Cont.

Future Analysis would be centered around refining this model to:

- ❏ Reduce error
- ❏ Incorporate additional zip codes in King County
- ❏ Improve impact of bedrooms

Appendix

What if I do not know the zip code. Will the predictor still work?

-*The predictor tool does require a zip code to predict a price.*

Would this work in my town of Puyallup?

-*This tool works only for towns falling in King County zip codes present in our data set. Expansion to other counties and zip codes would be open to exploration in future analysis.*

Please reach out with questions to the collaborators of this project.

Ashley Eakland: ashley@eakland.net // ***LinkedIn:*** *https://www.linkedin.com/in/ashleyeakland/*

Brad Blackwood: brad.blackwood.03@gmail.com // ***LinkedIn:*** *https://www.linkedin.com/in/blackwoodbs/*

Jalen Charles: jalenacharles@gmail.com // ***LinkedIn:*** *https://www.linkedin.com/in/jalen-charles*

For technical information and to see the Jupyter Notebook:

GitHub Repository: https://github.com/smashley-eakland/phase-2-project

Thank You!