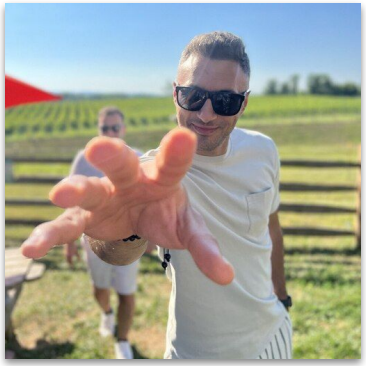
King County Housing Market Analysis

Predicting Sale Price using Regression Algorithms

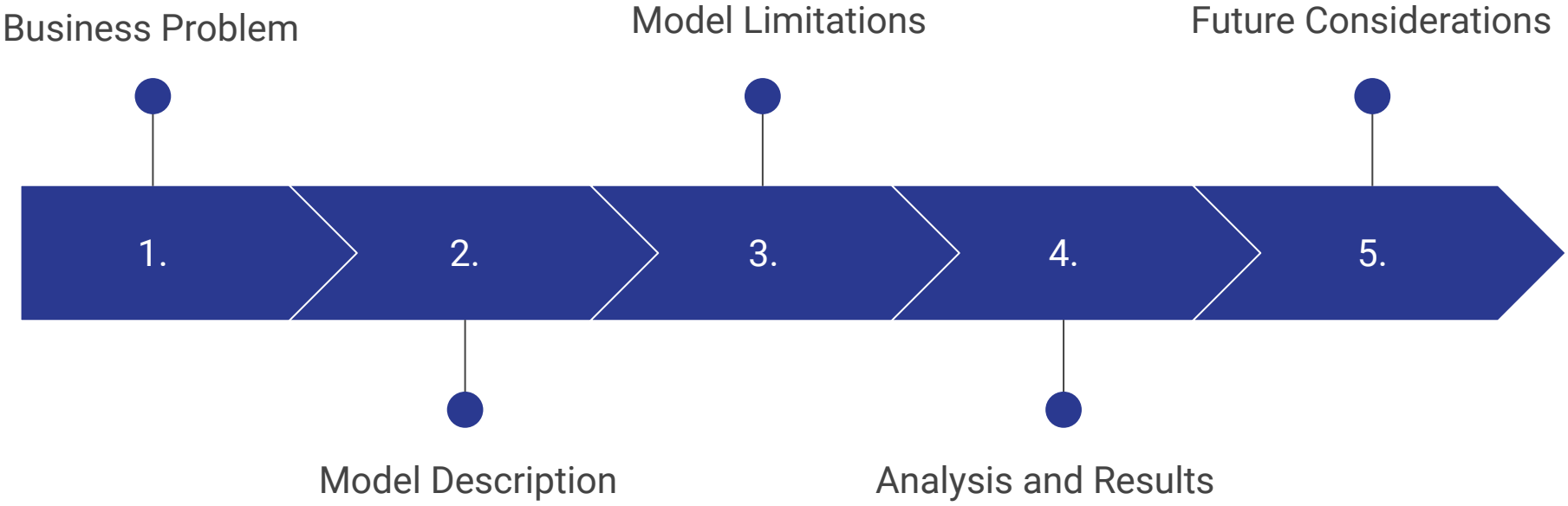
The Team



Milad Ghasemi

Ahmad Samiee

Kyle Lindstrom


Business Problem
Model Limitations
Future Considerations
1.
2.
3.
4.
5.
Model Description
Analysis and Results

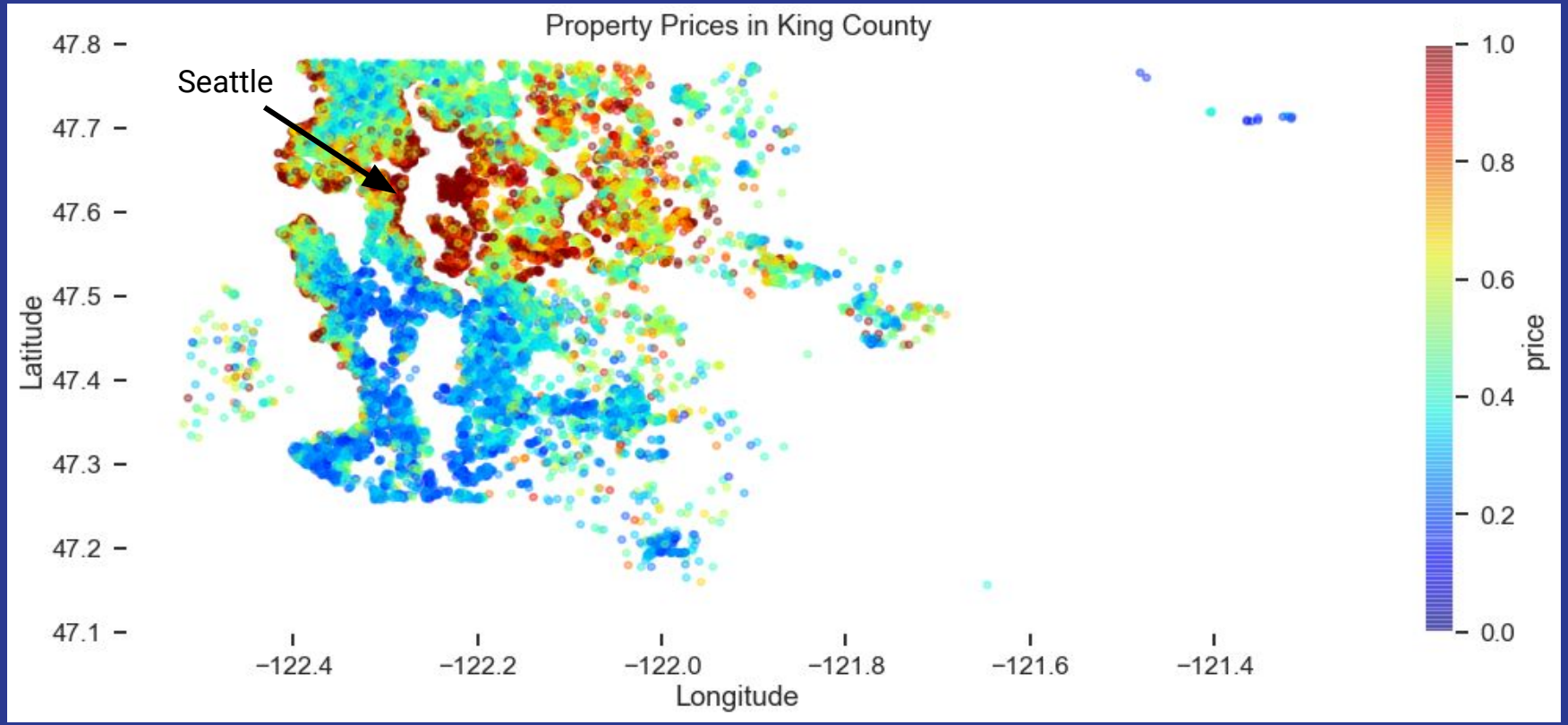
Dataset Overview

21,000 property sales in King County, Washington

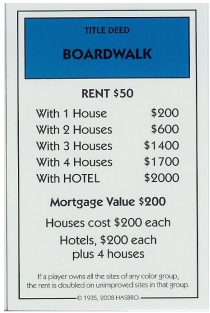
Sold between March 2014 and March 2015

Ranging in sale price between 78k — 7.7M

Overview of King County



The Business Problem



The Stakeholder	The Question	The Solution
Real Estate Tycoon in King County Washington	Which properties are undervalued?	Build a model that accurately predicts price

Bottom Line Up Front

Final Model

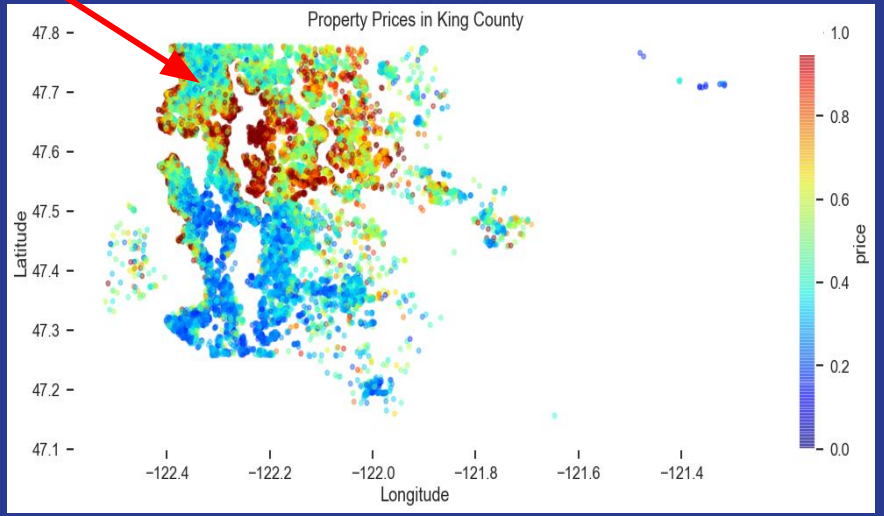
Median Absolute Percent Error of 11.67%

Use Case:

Property 10247 sold for $415k

Our model predicts a price of $449k

This property is likely undervalued!



Model Limitations

Focus on Minimizing Error

Individual predictor coefficients may be hard to interpret

Linear/Polynomial Regression

Not the most accurate prediction algorithm available

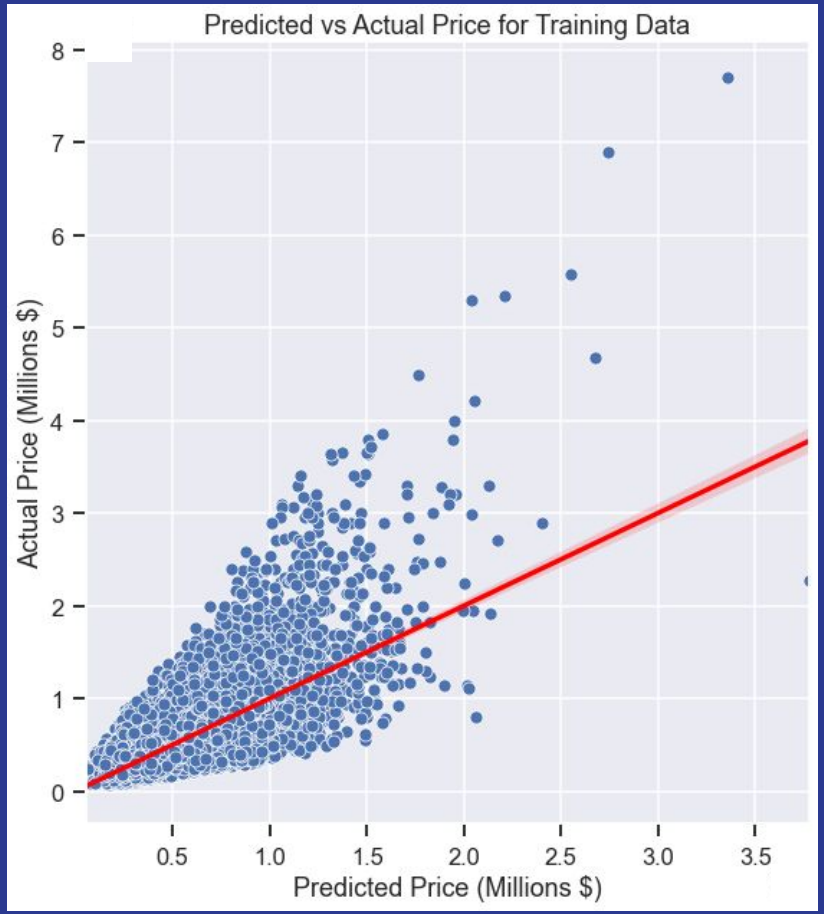
Baseline Model

Single Feature:

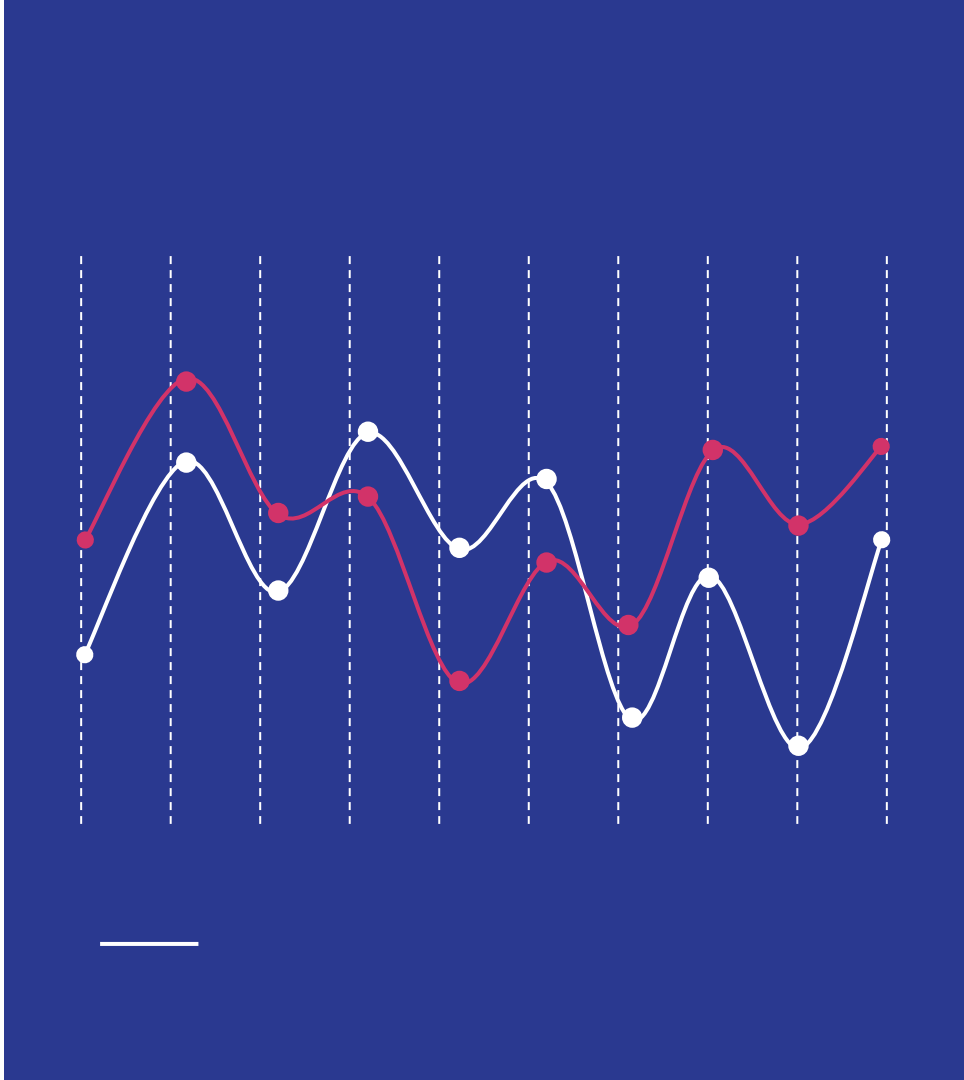
Square feet of living area

Percent Error:

28.51% — We can improve!



Feature Engineering



Features Added

Distance to Seattle

Grade

Median Income

HPI

Final Model

Additions:

Distance to Seattle
Overall Grade
Median Household Income
Housing Price Index
Polynomial Regression
Standard Scaling
Log Scaling

Percent Error:

11.67% — A major improvement!

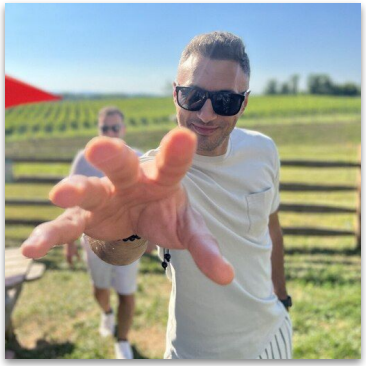
Future Considerations

Explore more accurate predictor algorithms

Investigate effect of demographics on sale price

Incorporate property data from more recent years

Thank you!



Milad Ghasemi

Ahmad Samiee

Kyle Lindstrom

Appendix

Model building

Teach our model how to predict prices

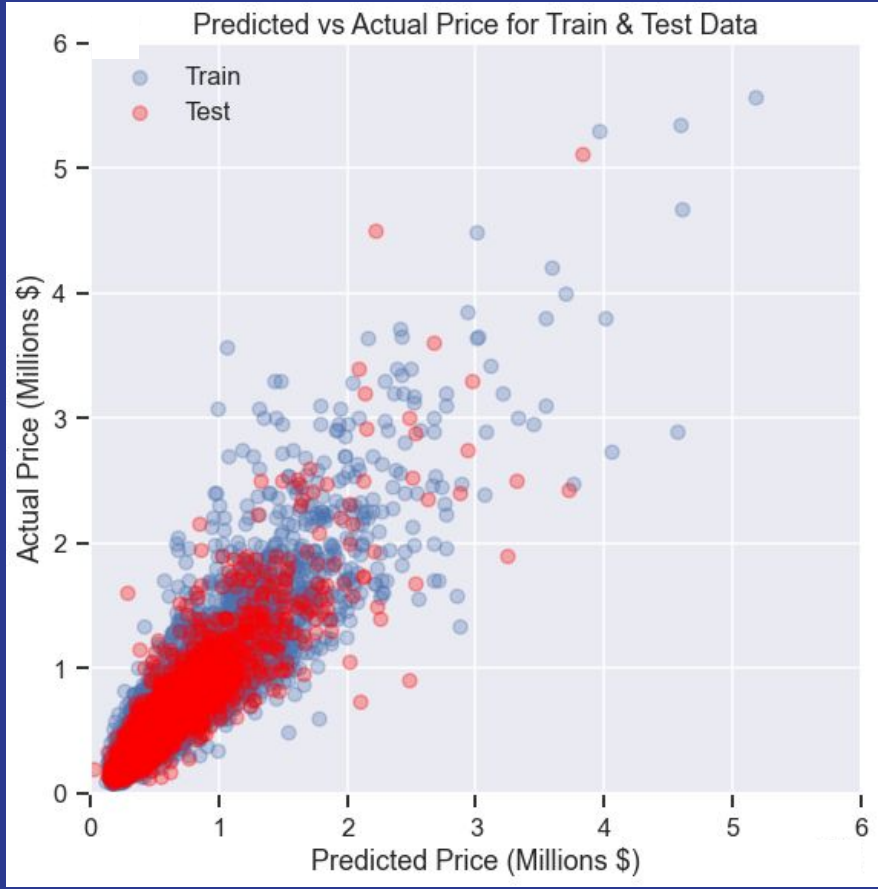
Ensure our model will generalize well

Predict new home prices, calculate Median Absolute Percent Error

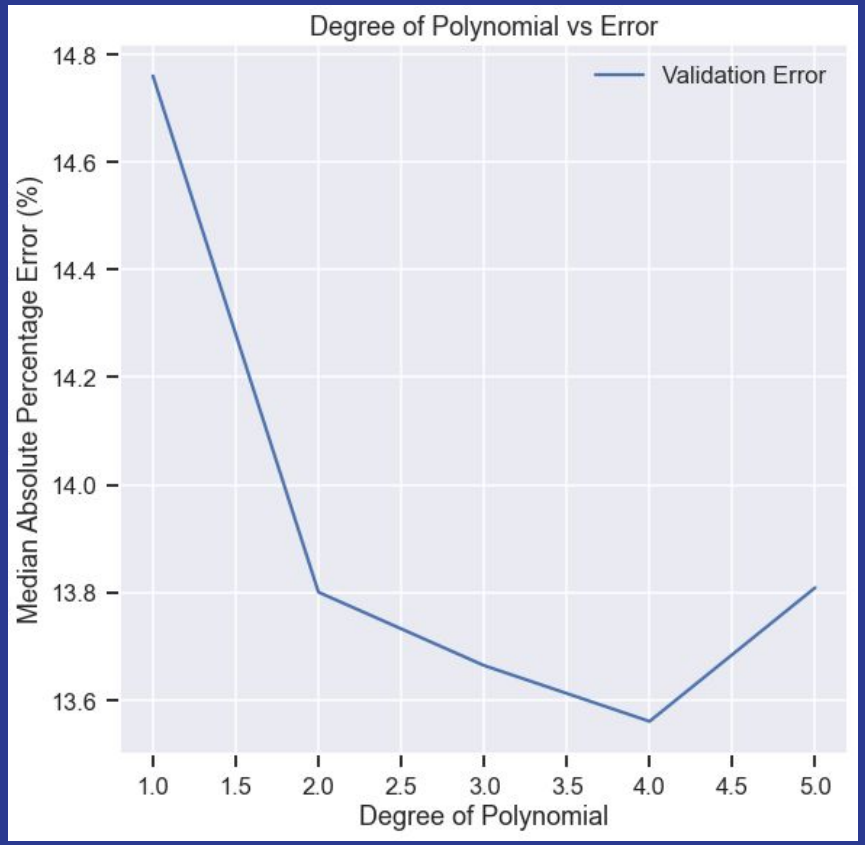
Testing on Unseen Data

Model Performance:

Median absolute percent error of 12.63%



Linearity vs Nonlinearity

What is the best fit for our data?

The relationship between price and our variables is nonlinear!