

KING COUNTY, WA HOUSING MARKET PRICE MODELING

January 2022

Project Team:

Syrvachev Sergey
Hannah Shurmann
Patrick Anastasio



"THE ACHE FOR HOME
LIVES IN ALL OF US, THE
SAFE PLACE WHERE WE
CAN GO AS WE ARE AND
NOT BE QUESTIONED."

MAYA ANGELOU



1

INTRODUCTION

2

RESEARCH

3

FINALIZE

- Business Problem
- Data Sources & Methods
- Market Research
- Models
- Regression Results
- Business recommendations
- Next Steps



INTRODUCTION



- Business problem
- Data sources & Methods



BUSINESS PROBLEM

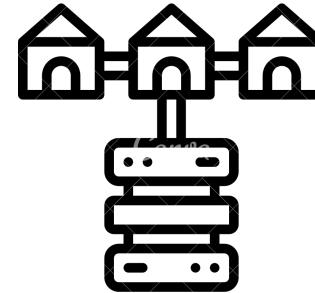
- Our team was hired by a major Seattle-based real estate agency **to create a model which predicts the prices of houses** in the King County, WA area based on certain property features.
- This model should provide customers with a predicted **house price using easy-fill questionnaires**



DATA UNDERSTANDING

Sources

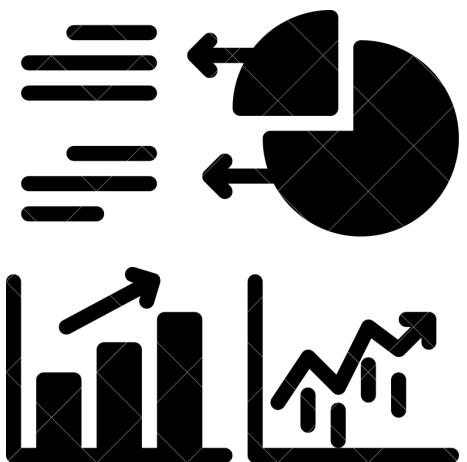
- King County, WA housing sales data
(2014-2015)
- Open Street Maps (www.openstreetmap.org)
Geographical maps and geolocation coding
- Washington Geospatial Open Data (<https://geo.wa.gov/>)
Washington State geospatial information



Methods

- Exploratory data analysis (EDA)
- Geoinformation scraping
- Determine statistically significant features
- Regression models - OLS, K-Fold, Train-Test Split

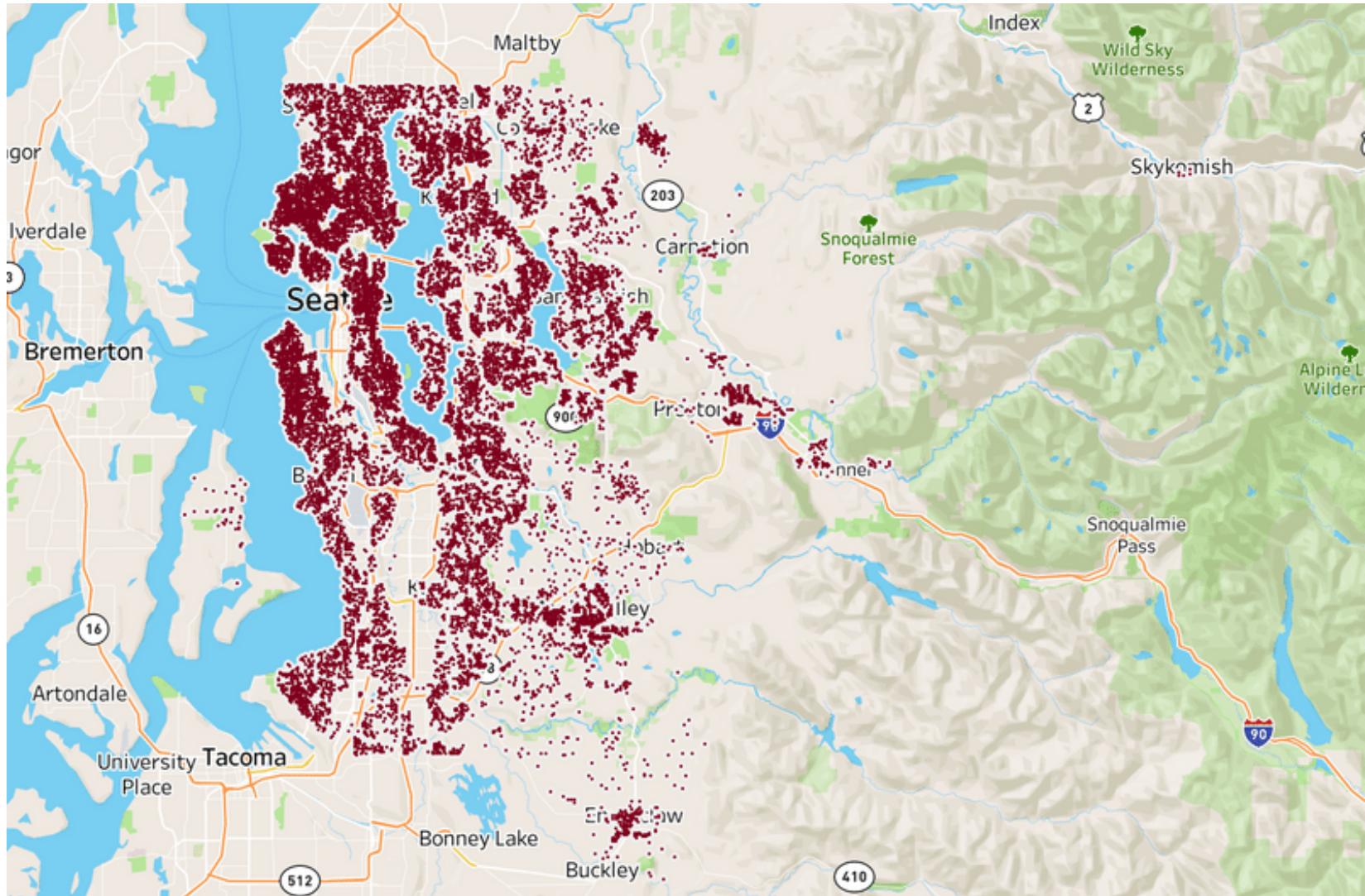
RESEARCH



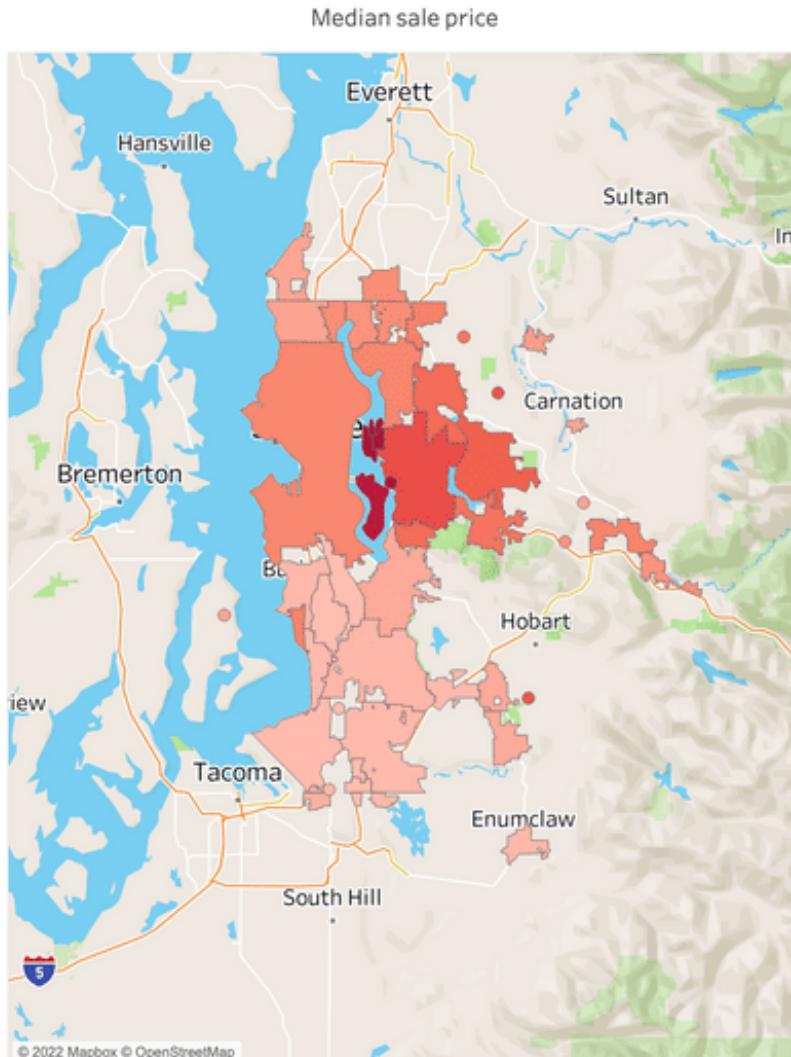
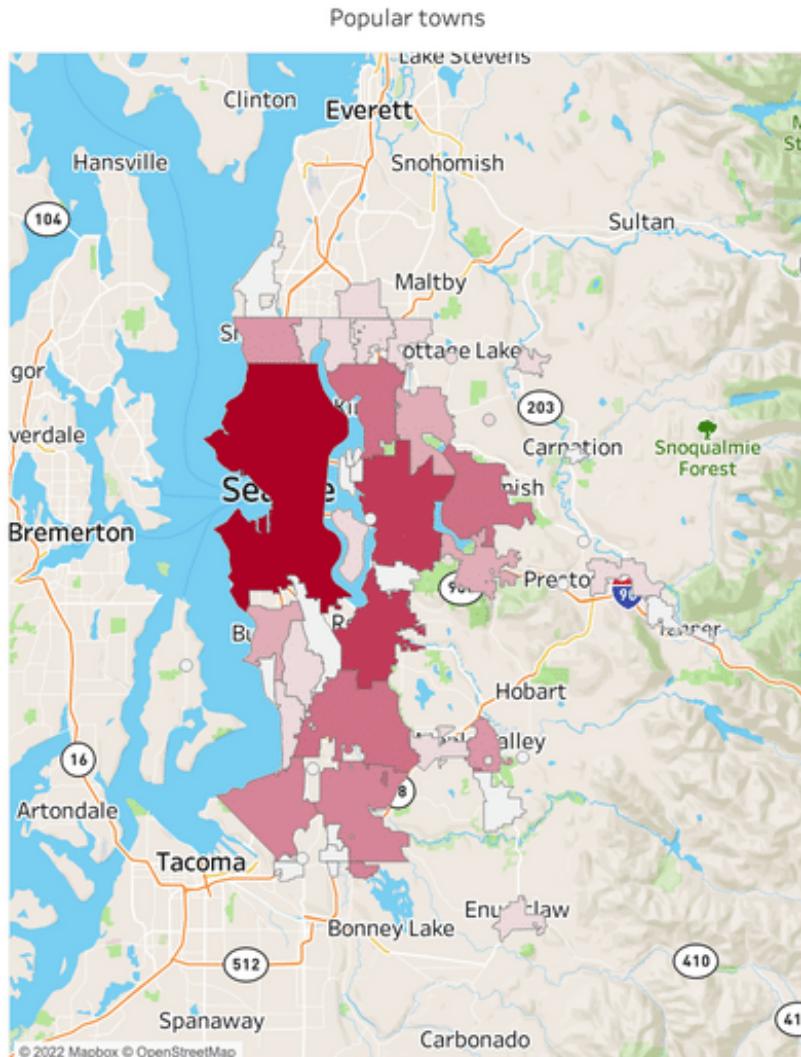
- Market Research
- Preliminary Models
- Final Model



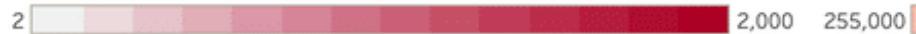
SALES TRANSACTIONS THROUGHOUT KING COUNTY. (2014 / 2015)



LOCATION HAS SIGNIFICANT IMPACT ON PRICE.



The number of transaction



Median Price





USED FEATURES:



FOOTAGE



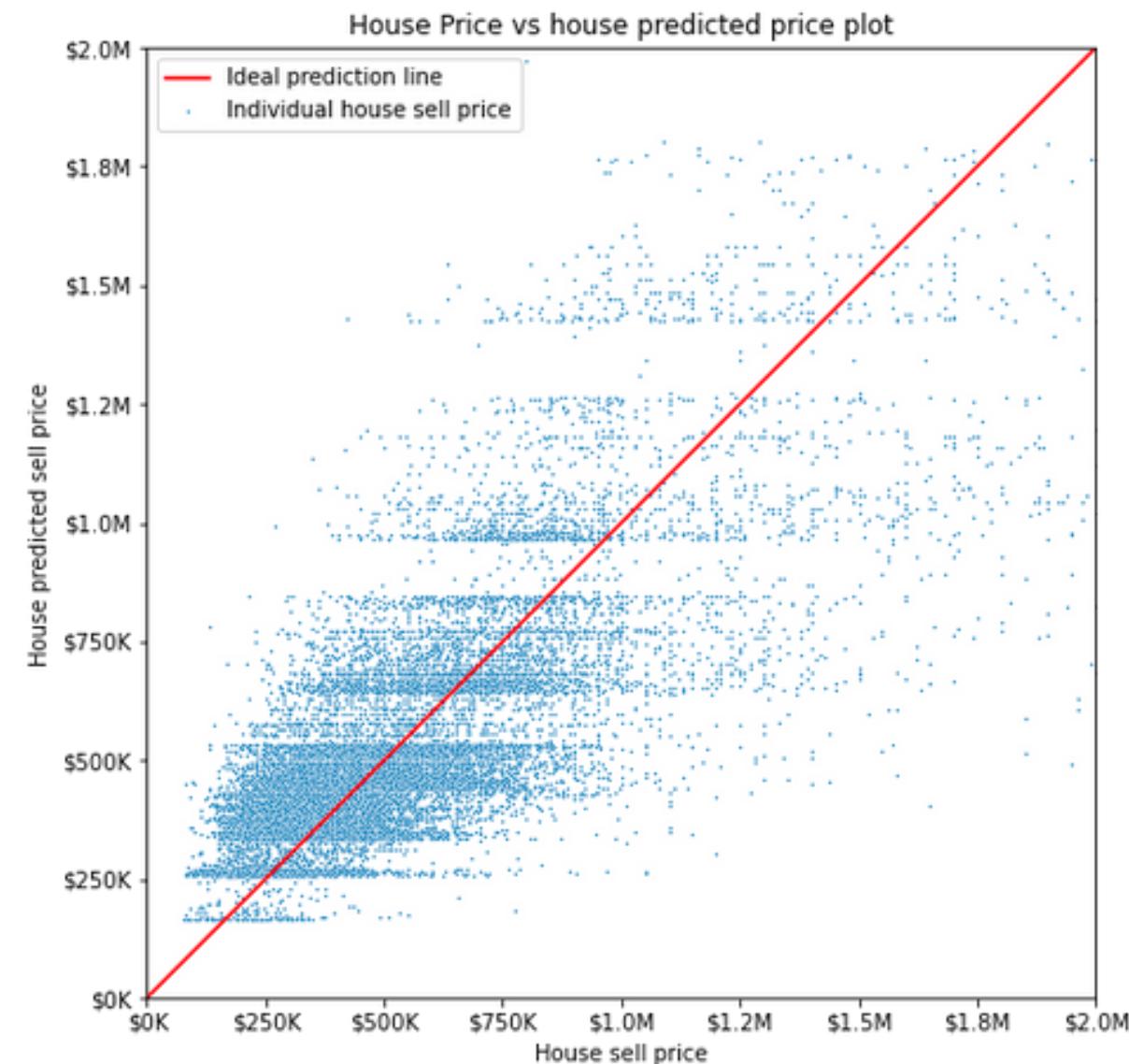
BEDROOMS



WATERFRONT



VIEW



USED FEATURES:



AGE



GRADE



INCLUDED FEATURES:



SQUARE FOOTAGE



BEDROOMS



WATERFRONT



VIEW



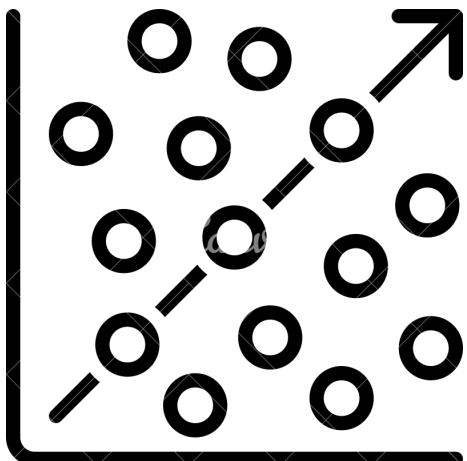
AGE



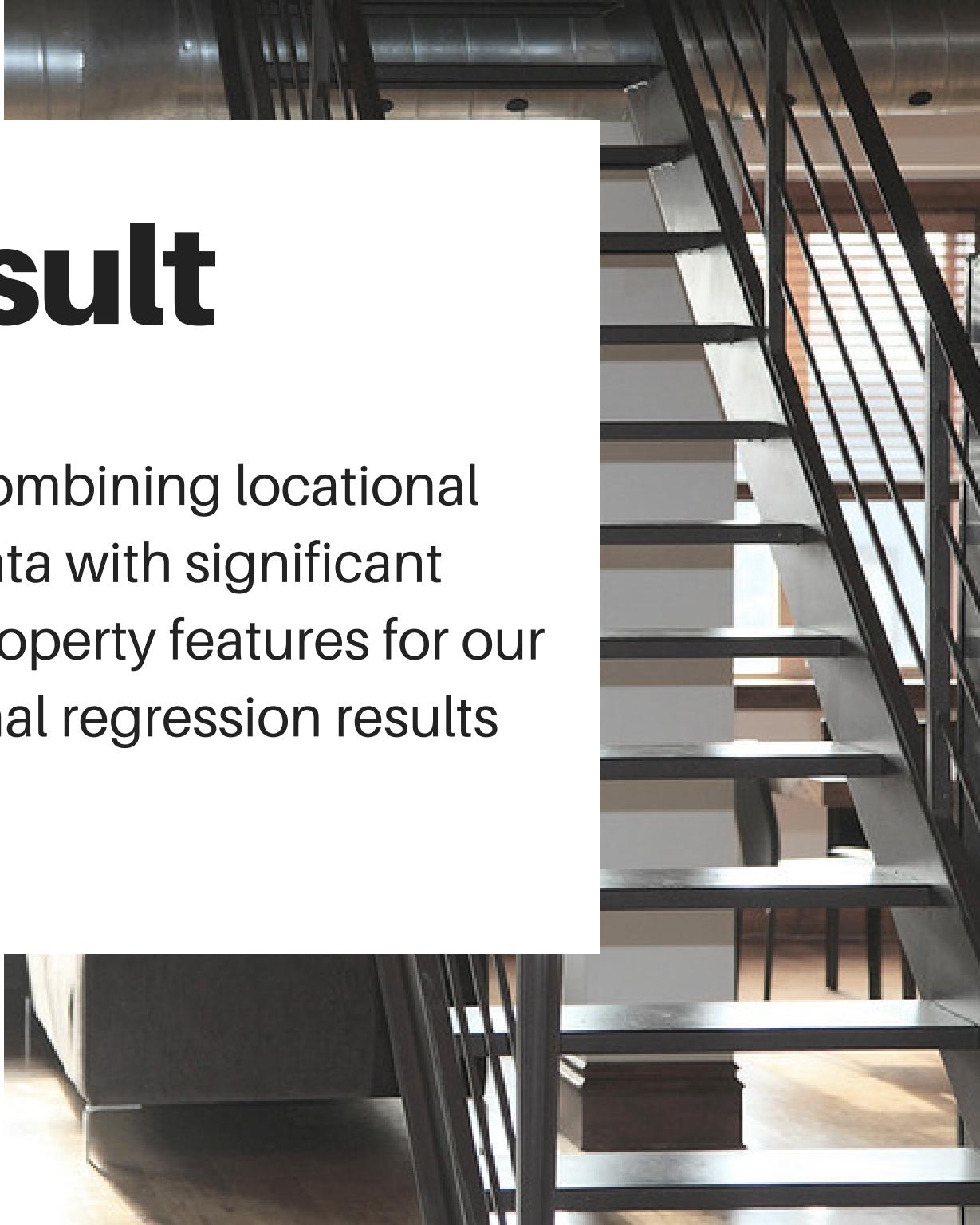
GRADE

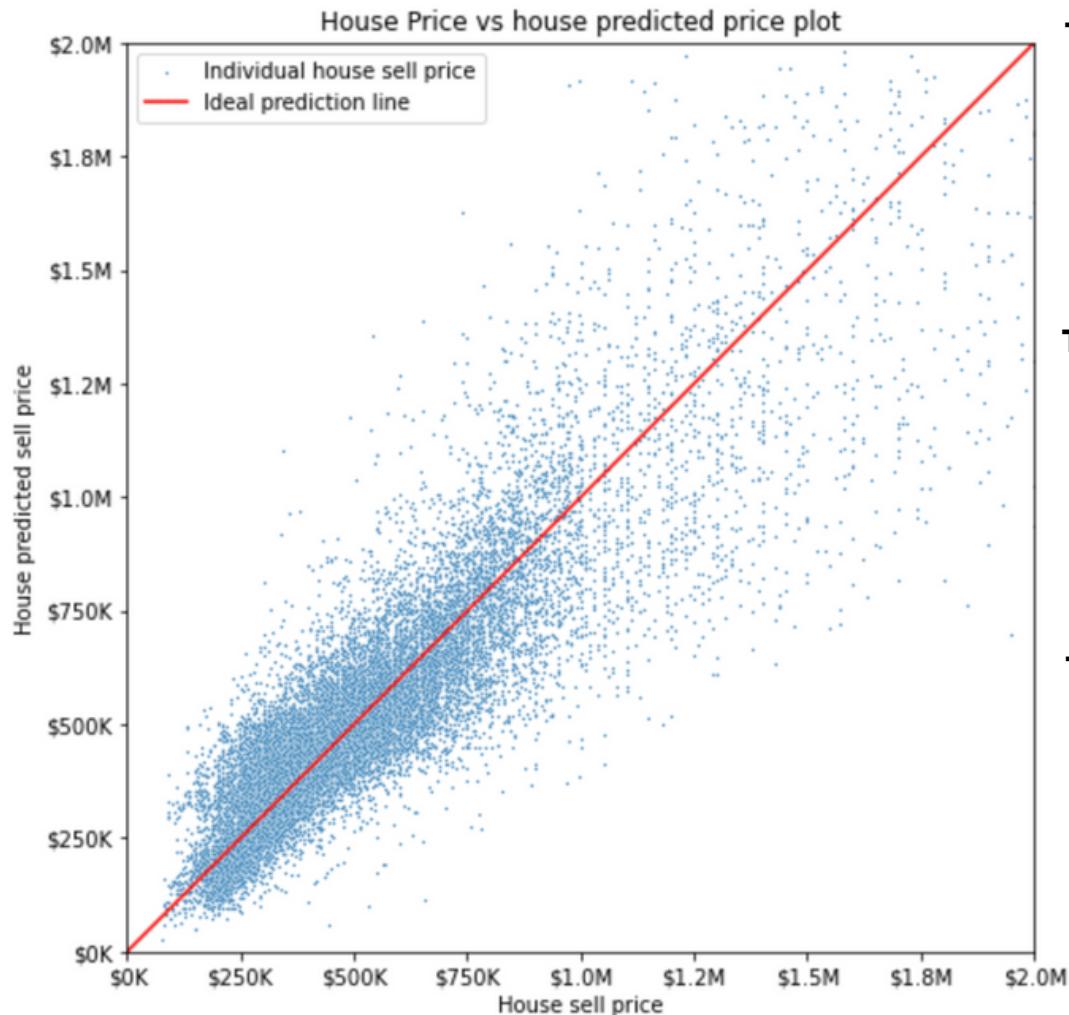
EXPLAINED VARIANCE = **69.2%**
(R² SCORE)

Final Result



- Combining locational data with significant property features for our final regression results





TEST 1

R² Score: 80.1% +-3.9%
Root Mean Squared Error: 164282.0
Random house price: 858000.0
Predicted price: 922740.0
Difference: 7.55 %

TEST 2

R² Score: 80.1% +-3.9%
Root Mean Squared Error: 164282.0
Random house price: 225000.0
Predicted price: 214094.0
Difference: -4.85 %

TEST 3

R² Score: 80.1% +-3.9%
Root Mean Squared Error: 164282.0
Random house price: 420000.0
Predicted price: 429121.0
Difference: 2.17 %

EXPLAINED VARIANCE = 80.1%
(R2 SCORE)



Business Recommendations



01

PROPERTY GRADE & SQUARE FOOTAGE



Property Grade and Square Footage have the strongest correlation with price

02

RENOVATIONS



Renovations did not significantly add to sales price. Consider minimally when pricing a property.

03

NEIGHBORHOOD



Prices vary widely based on neighborhood. Location is a major predictor of price.



Next steps

- Pinpoint Seattle district information
- Explore other statistical model types
- Expand to other regions and cities

"HOME ISN'T WHERE
YOU'RE FROM, IT'S
WHERE YOU FIND LIGHT
WHEN ALL GROWS
DARK"

PIERCE BROWN





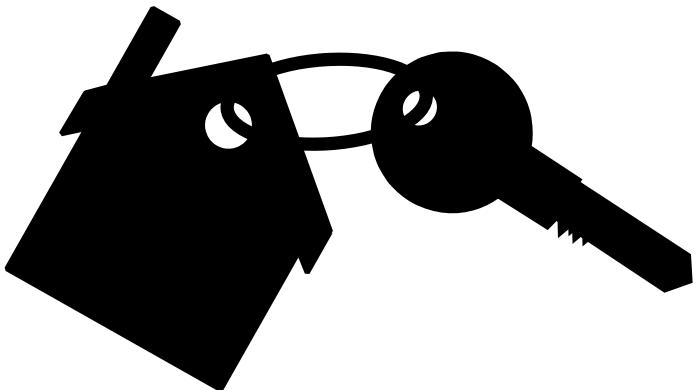
Q & A:



**Thank you for joining
today's presentation.**

SYRVACHEV SERGEY

- DEST.STUDIO@GMAIL.COM
- LINKEDIN: /SSYRVACHEV
- GITHUB: 314KA4Y
- MEDIUM: @SERGEYSYRVACHEV



HANNAH SHURMANN

- HANNAH.SCHURMAN1@GMAIL.COM
- LINKEDIN: /HANNAH-SCHURMAN-706391188
- GITHUB: HANNAH-SCHURMAN
- MEDIUM: @HANNAH.SCHURMAN1

PATRICK ANASTASIO

- SUDOMAKECOFFEE1@GMAIL.COM
- LINKEDIN: /PATRICKANASTASIO
- GITHUB: PATRICK-ANASTASIO
- MEDIUM: @PATRICKANASTASIO