KING COUNTY, WA HOUSING MARKET PRICE MODELING

APRIL 20TH 2023

Project Team:

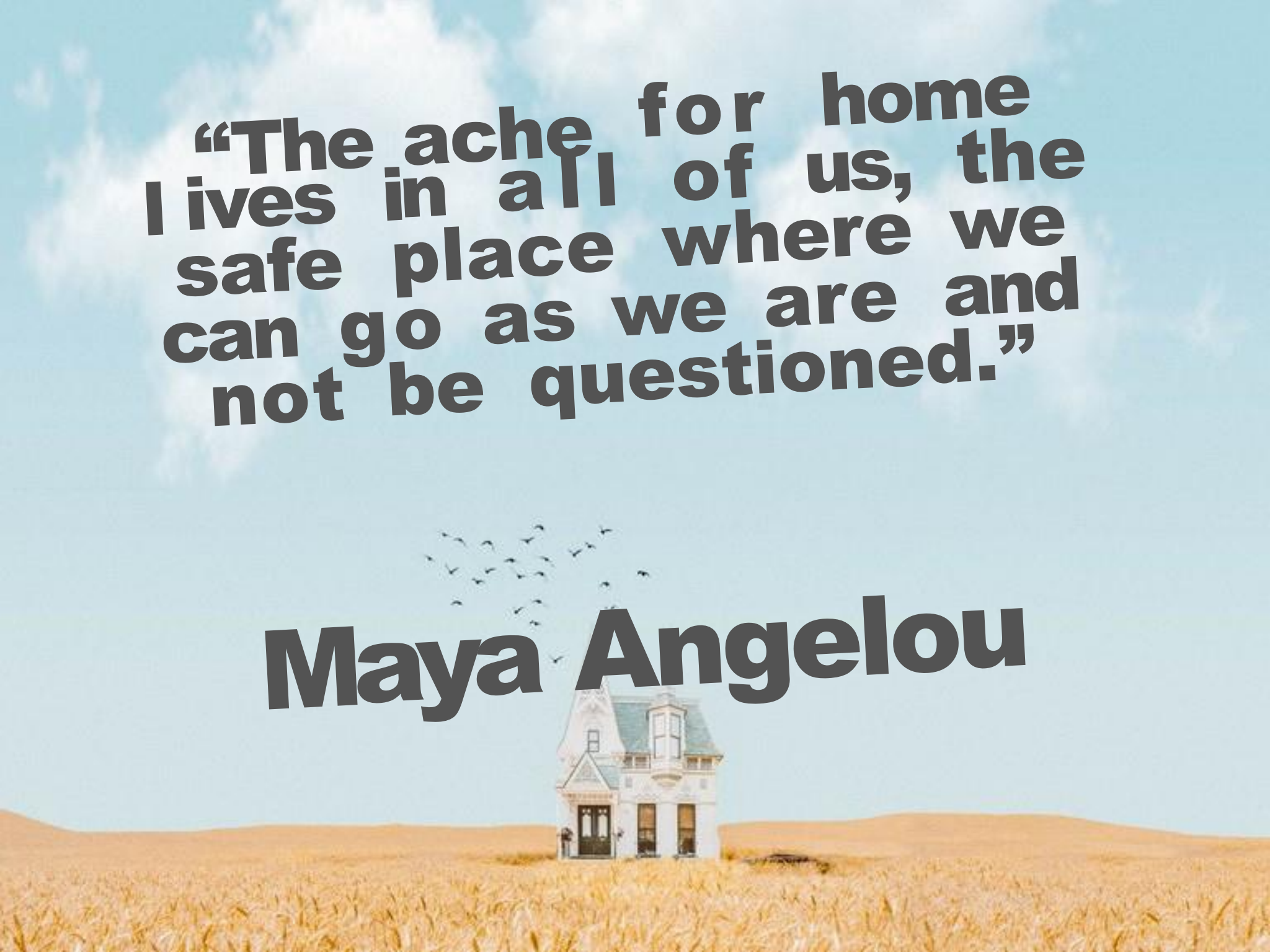
FLORENCE NGUUNI

JOSHUA RWANDA

- LEAH KALUMBA
- EDNA WANJIKU

STEPHANIE MBITHE

- KINOTI MWENDA


“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	2	3
INTRODUCTION	RESEARCH	FINALIZE

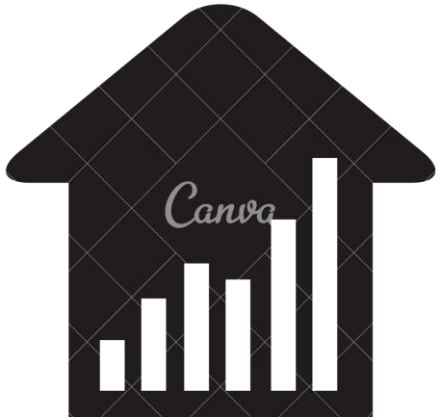
- Business Problem
- Data Sources & Methods

- Market Research
- Models
- Regression Results

- Business recommendations
- Next Steps

04

INTRODUCTION



- Business problem
- Data sources & Methods

BUSINESS PROBLEM

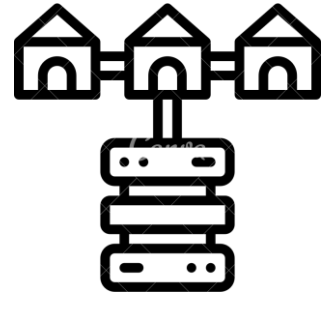
- Our team was hired by a major Seattle-based real estate agency **to create a model which predicts the prices of houses** in the King County, WA area based on certain property features.

- This model should provide customers with a predicted **house price using easy-fill questionnaires**

DATA UNDERSTANDING

Sources

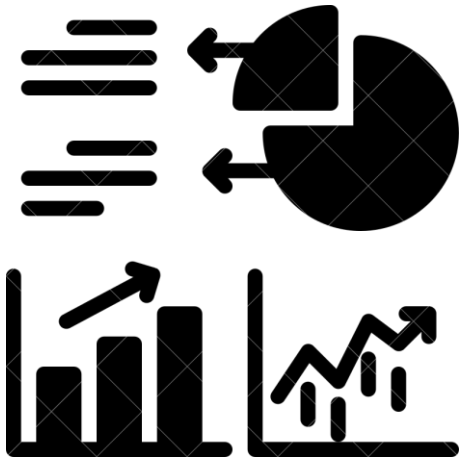
- **King County, WA housing sales data**
  (2014-2015)
- **Open Street Maps (www.openstreetmap.org)**
  Geographical maps and geolocation coding
- **Washington Geospatial Open Data (https://geo.wa.gov/)**
  Washington State geospatial information



Methods

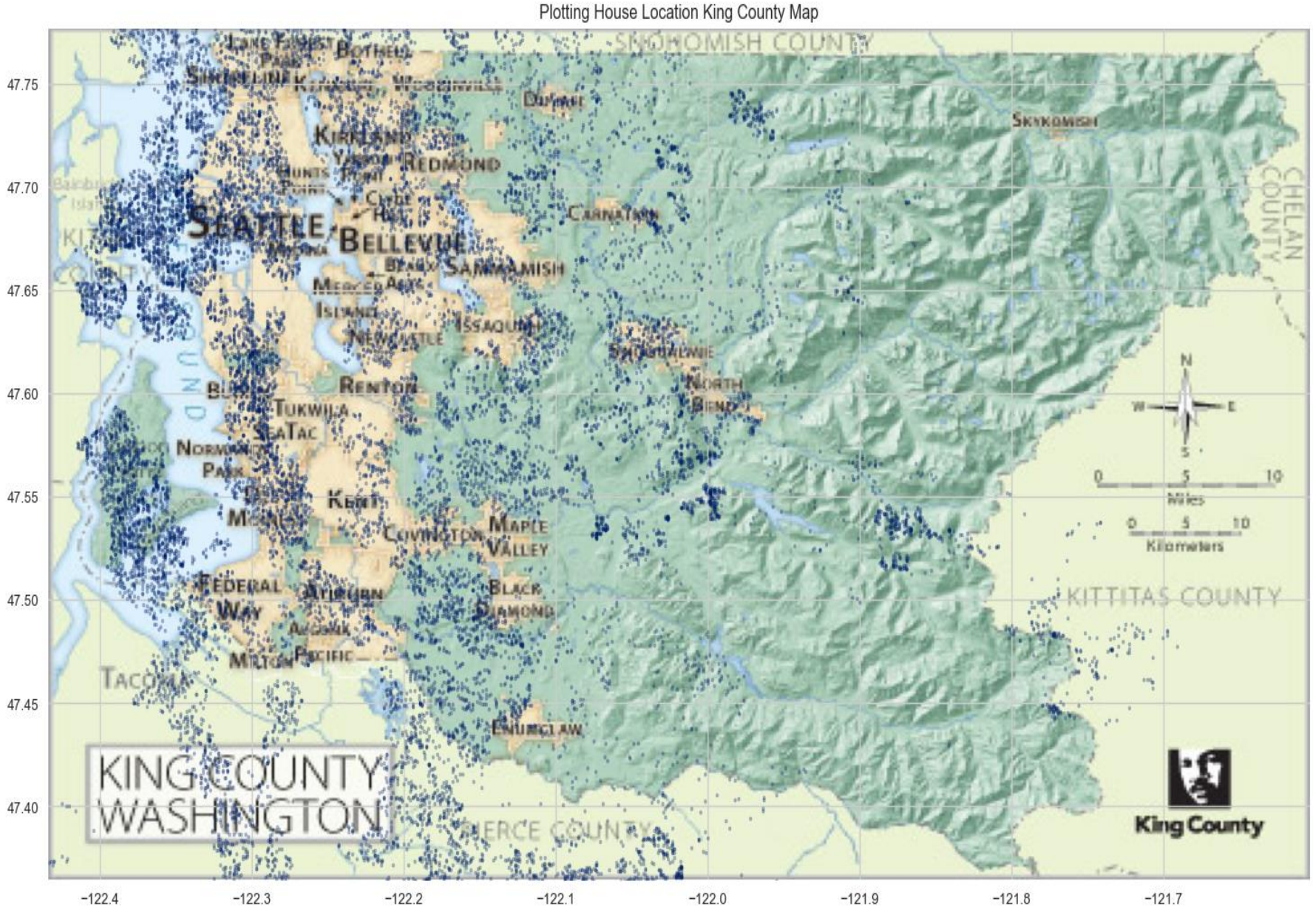
- Exploratory data analysis (EDA)
- Geoinformation scraping
- Determine statistically significant features
- Regression models - OLS, K-Fold, Train-Test Split

RESEARCH



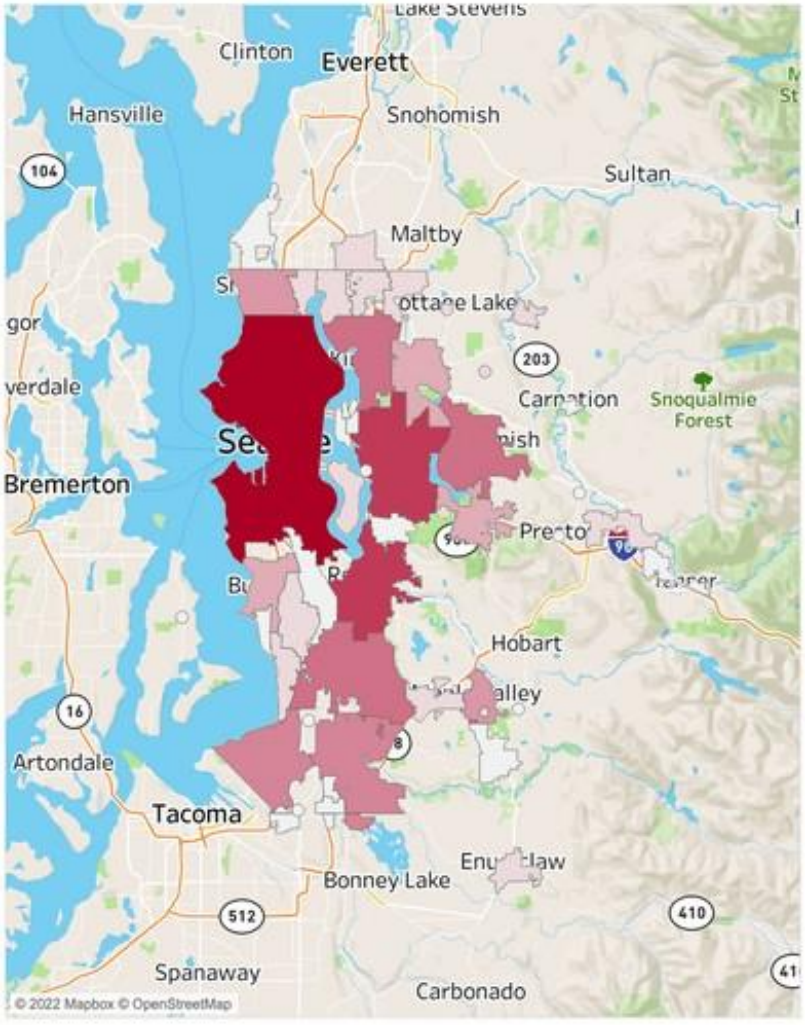
- Market Research
- Preliminary Models
- Final Model

SALES TRANSACTIONS THROUGHOUT KING COUNTY.



LOCATION HAS SIGNIFICANT IMPACT ON PRICE.

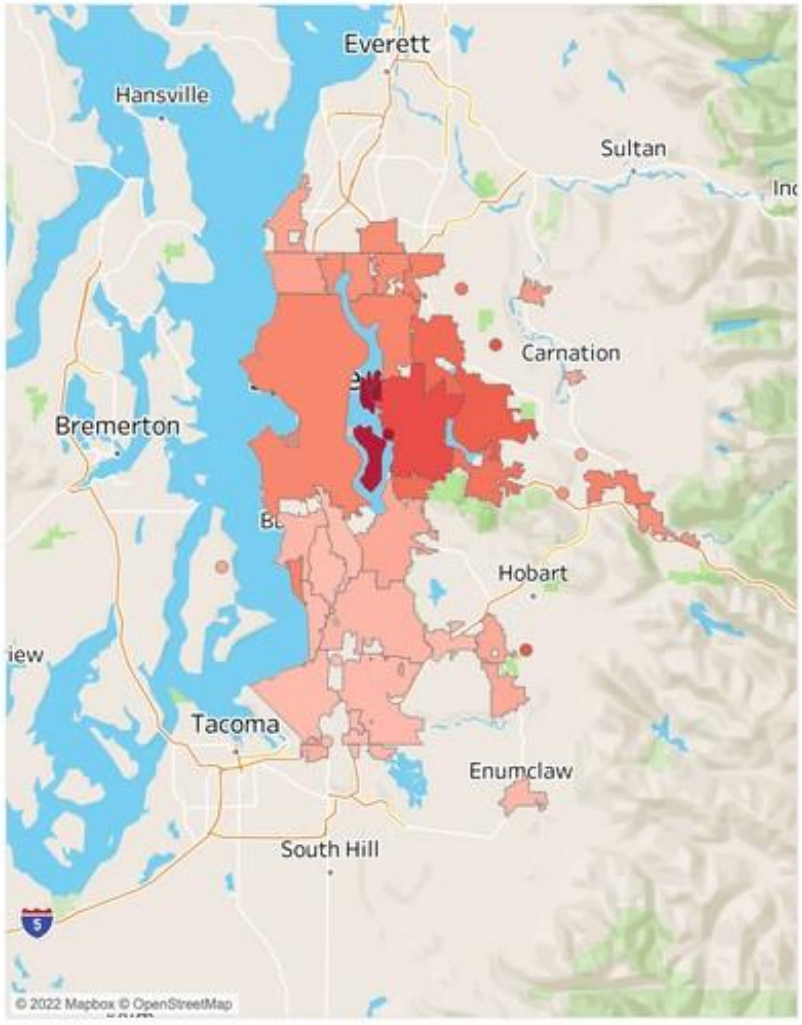
Popular towns



The number of transaction

2 — 2,000

Median sale price

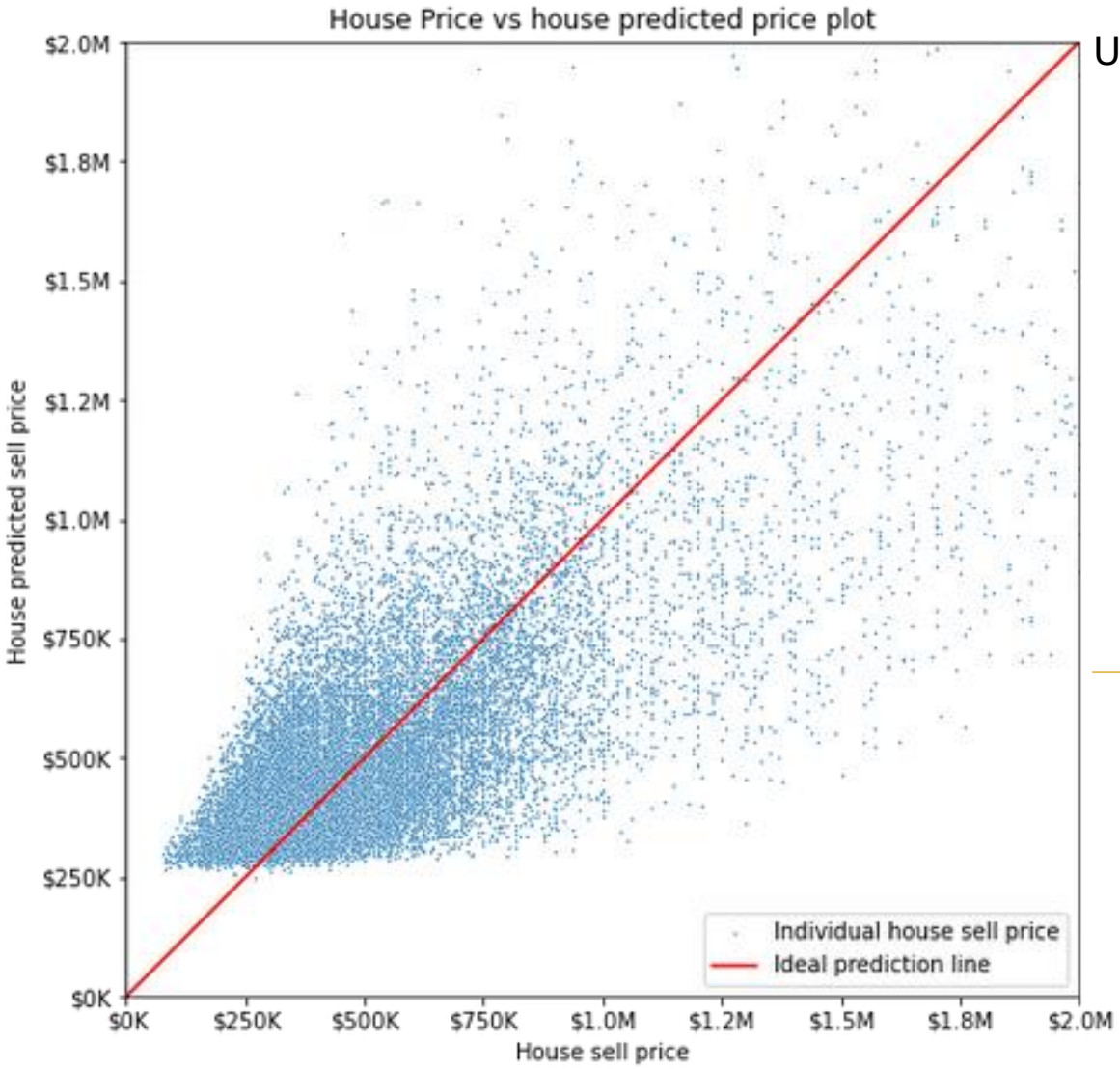


Median Price

255,000 — 1,050,000

MODEL 1

STRUCTURE FEATURES



USED FEATURES:

FOOTAGE

BEDROOMS

WATERFRONT

VIEW

EXPLAINED VARIANCE =**47.7%**

(R2 SCORE)

TRAIN MODEL

USED FEATURES:

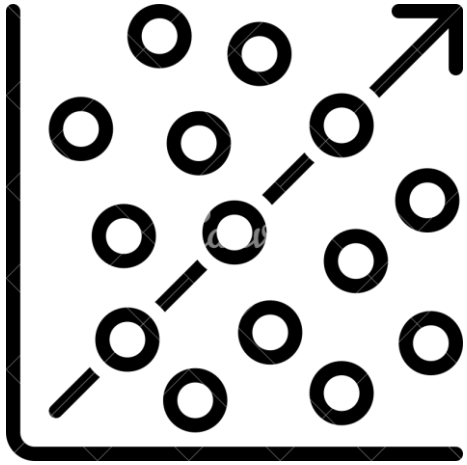
ACTUAL PRICE

PREDICTED

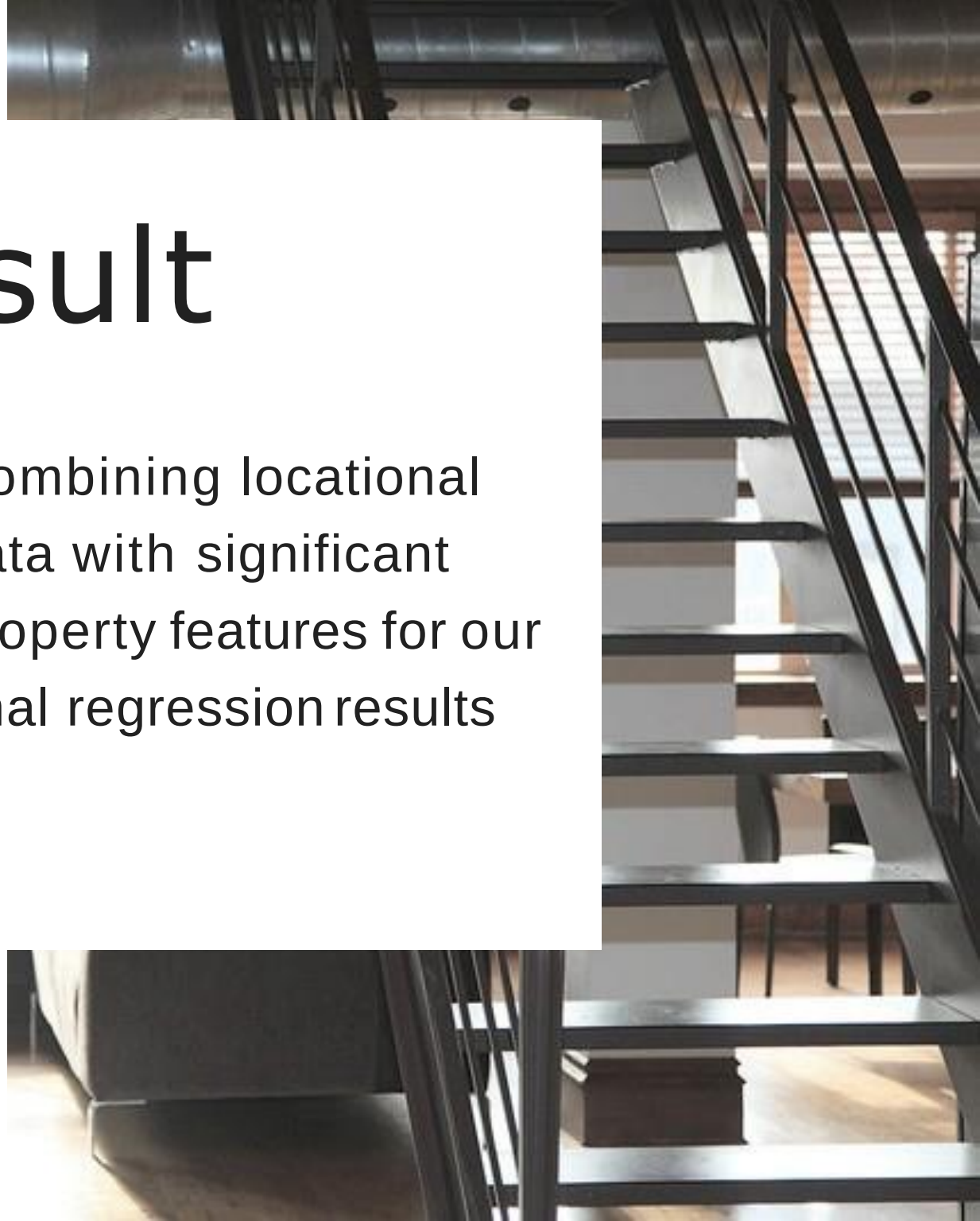
EXPLAINED VARIANCE =**51.1%**

(R 2 SCORE)

Final Result

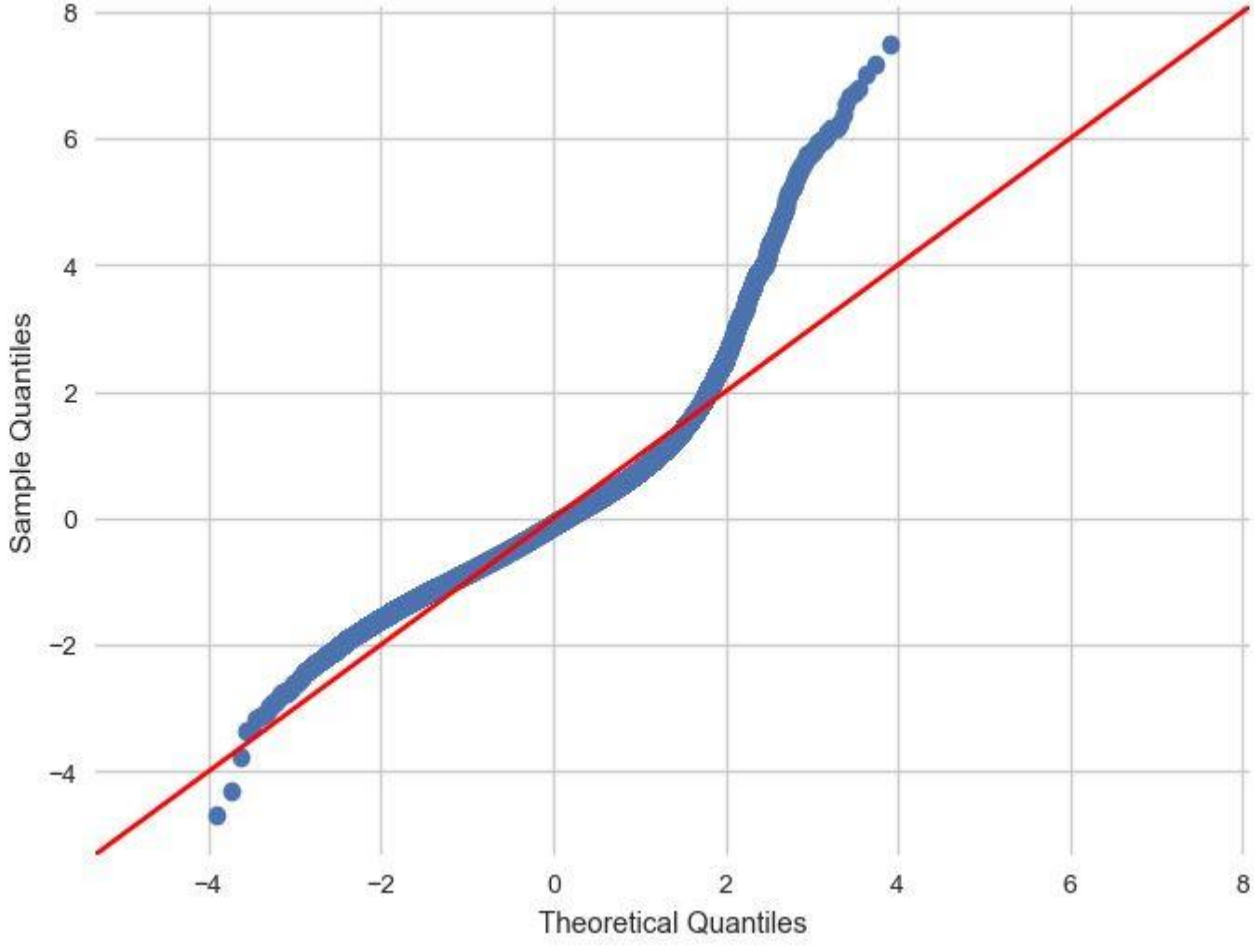


- Combining locational data with significant property features for our final regression results



POLYNOMIAL REGRESSION

INCLUDED FEATURES:



- SQUARE LIVING
- RENOVATED GRADE(LOG TRANSFORMED GRADE)

EXPLAINED VARIANCE = **50.5%**

(R 2 SCORE)

Business Recommendations

01

PROPERTY GRADE &SQUARE F

Property Grade and Square Footage have the strongest correlation with price

02

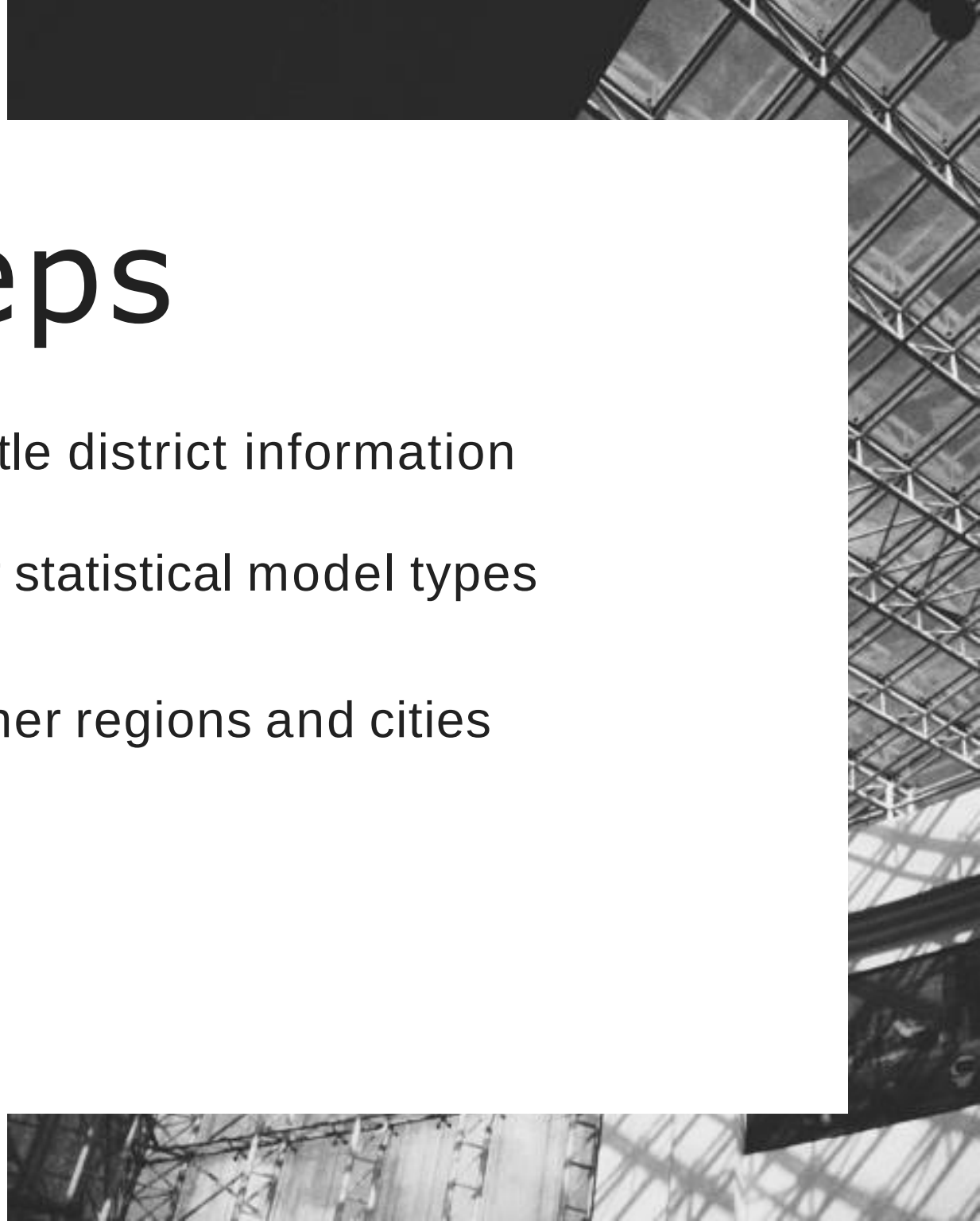
RENOVATIONS

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.