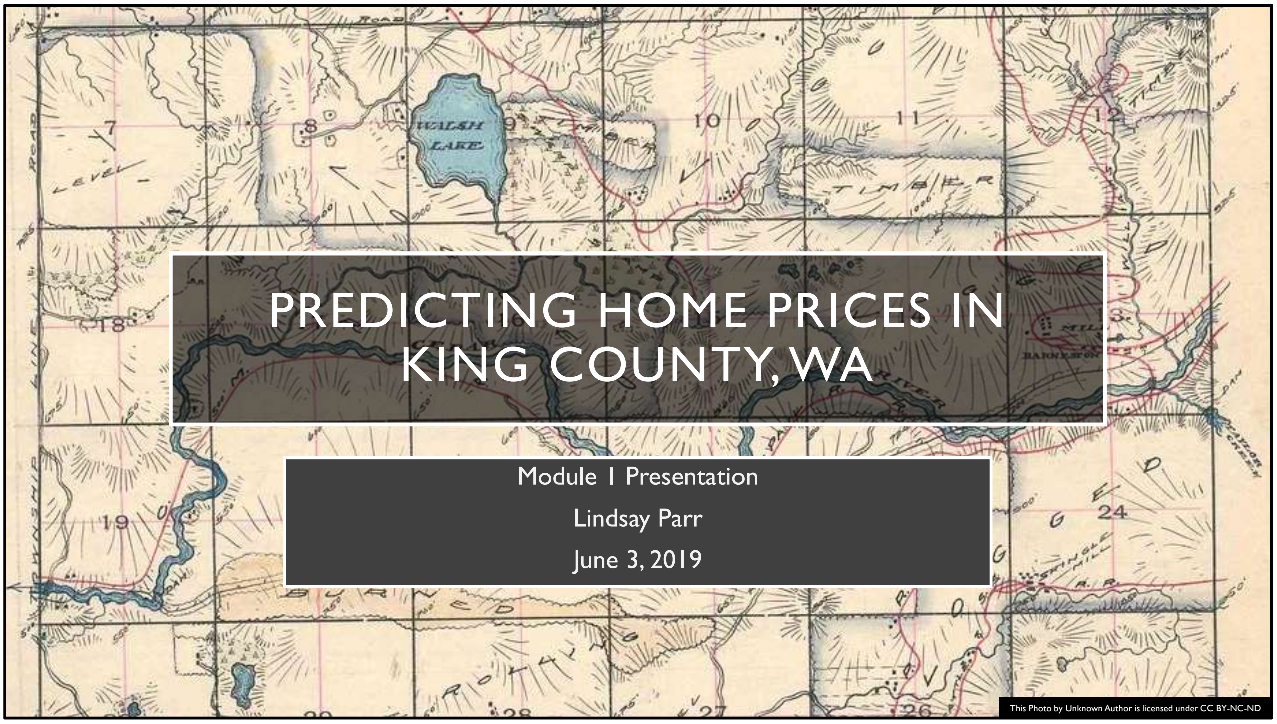
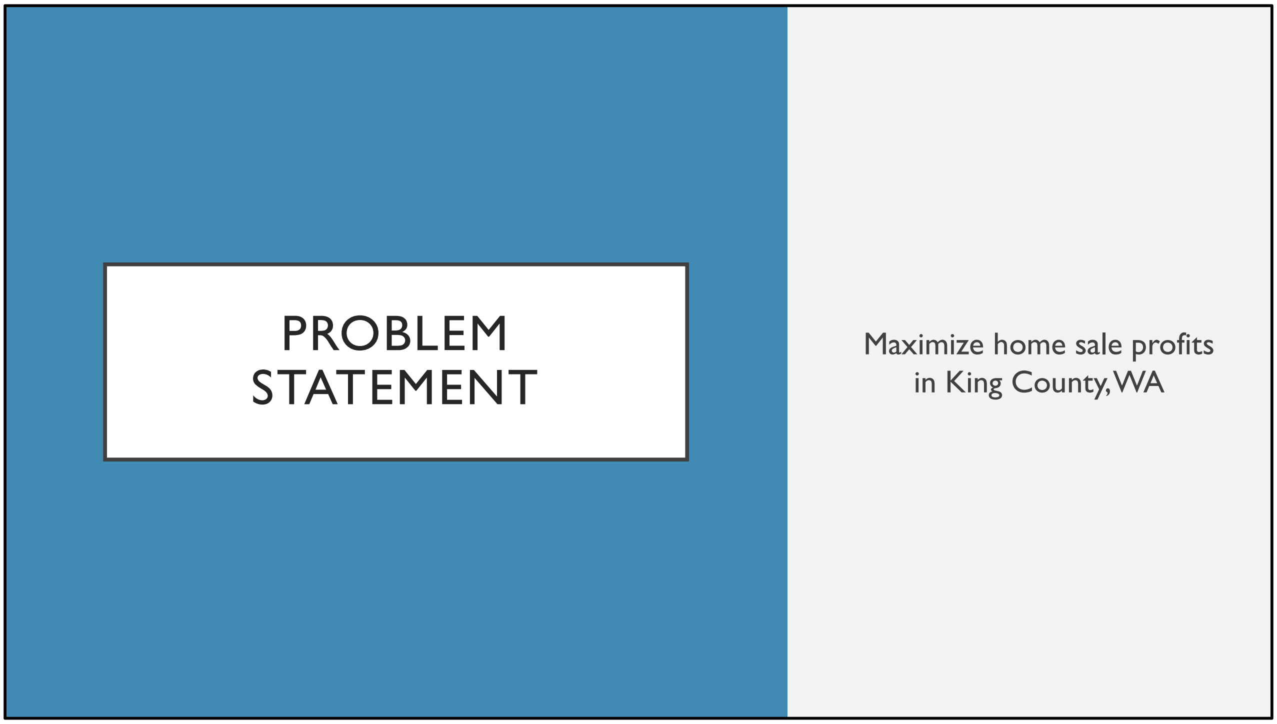

WALSH LAKE
7
8
9
10
11
12
18
19
24
TIMBER
RIVER
BURNED
ROLLING

# PREDICTING HOME PRICES IN KING COUNTY, WA

Module 1 Presentation

Lindsay Parr

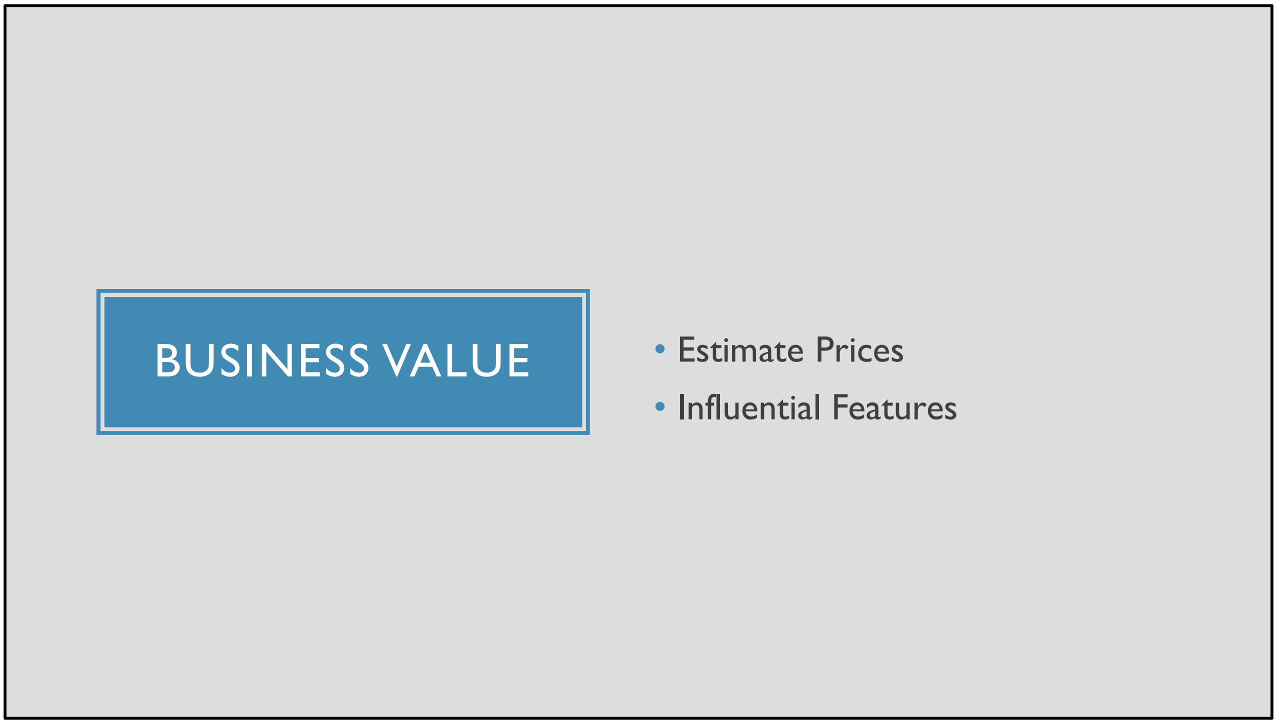
June 3, 2019



## PROBLEM STATEMENT

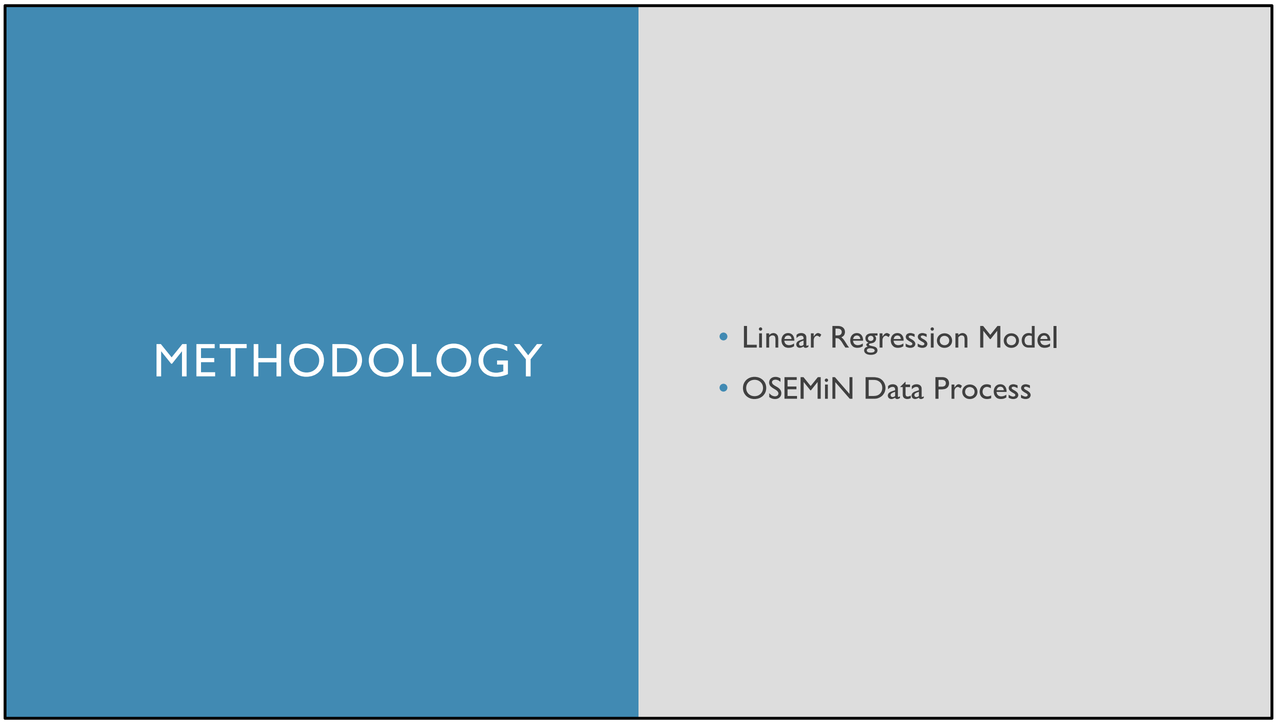
Maximize home sale profits
in King County, WA

We need to create a predictive model that allows us to accurately estimate home prices in King County, WA based on certain variables.  Our initial data included 21,597 home sale records with 20 predictor variables and target variable – price.  Predictor variables included square footage of the home and lot, number of bedrooms, bathrooms, and floors, condition of the home, year built and renovated, number of views, if the property was on the waterfront, and an average square footage of the 15 nearest neighbors.



## BUSINESS VALUE

- Estimate Prices
- Influential Features



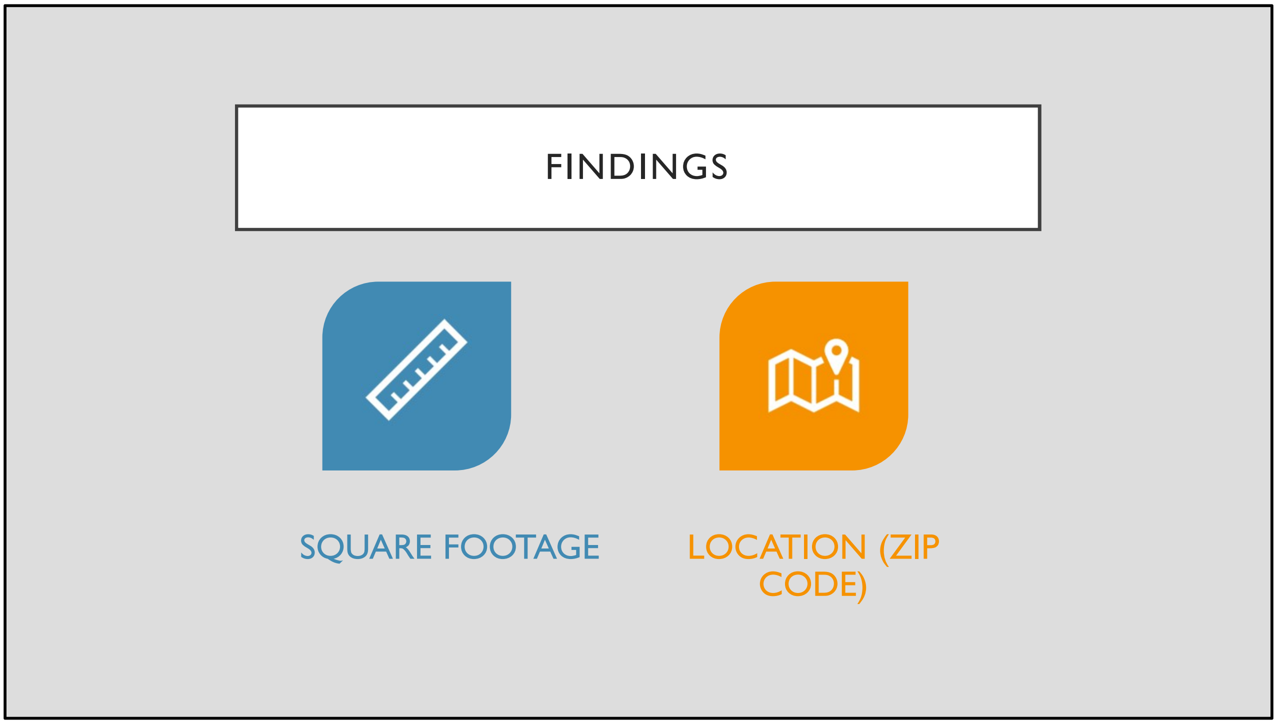
## METHODOLOGY

- Linear Regression Model
- OSEMiN Data Process

- Removed placeholder values (?) and null values
- Removed outlier values
- Binned categorical (non-numeric) data into groups to identify potential relationships
- Scaled data to capture full magnitude of changes

The revised, clean data set has 19,468 home records and 15 predictor variables to estimate the target (price).

After cleaning ran regression model with 0.88 r-squared, but was not able to cross-validate the results. Our features were paired down using the stepwise selection and a p-value threshold of 0.05. There were 107 features selected out of the initial 145. I ran the model again with the significant features and got a 0.88 r-squared again and was able to cross-validate it with 87% accuracy.

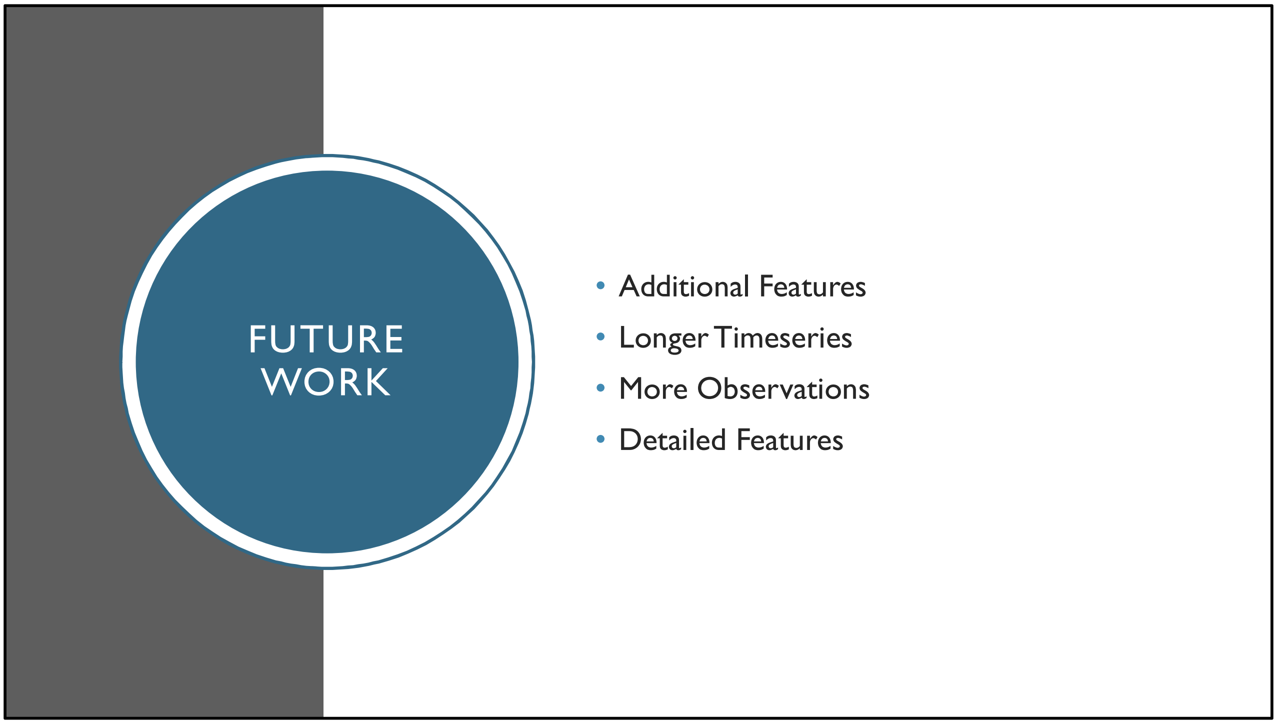


# FINDINGS

SQUARE FOOTAGE

LOCATION (ZIP CODE)

Square footage of the home is the most influential factor on price. Various zipcodes were the next most influential factor. Size of the home and a prime location are the biggest drivers in housing prices.



FUTURE WORK

- Additional Features
- Longer Timeseries
- More Observations
- Detailed Features

**Recommendations for Future Work**

More records to observe

Longer timeframe

Additional Features

- Green Areas
- School Districts
- Tax Brackets
- Crime Rates
- Distance to Highways
- Distance from Major City
- Roof/HVAC/Windows Age
- Types of Finishes

Renovation Descriptors

- Aesthetic vs. Structural