King County Housing Prices

A Multiple Linear Regression Model
**By: Jack Locke**

Objective:

To use the dataset to predict the price of a house.

Dataset:
- **Id** - unique identifier for a house
- **Date** - house was sold
- **Price** - is prediction target
- **Bedrooms** - number of Bedrooms/House
- **Bathrooms** - number of bathrooms/bedrooms
- **Sqft_living** - footage of the home
- **Sqft_lot** - footage of the lot
- **Floors** - floors (levels) in house
- **Waterfront** - House which has a view to a waterfront
- **View** - Has been viewed
- **Condition** - How good the condition is
- **Grade** - overall grade given to the housing unit
- **Sqft_above** - square footage of house apart from basement
- **Sqft_basement** - square footage of the basement
- **Yr_built** - Built Year
- **Yr_renovated** - Year when house was renovated
- **Zipcode** - zip
- **Lat** - Latitude coordinate
- **Long** - Longitude coordinate
- **Sqft_living15** - The square footage of interior housing living space for the nearest 15 neighbors
- **Sqft_lot15** - The square footage of the land lots of the nearest 15 neighbors

OSEMN - workflow

| OBTAIN | SCRUB | EXPLORE | MODEL | INTERPRET |

O **S** **E** **M** **N**
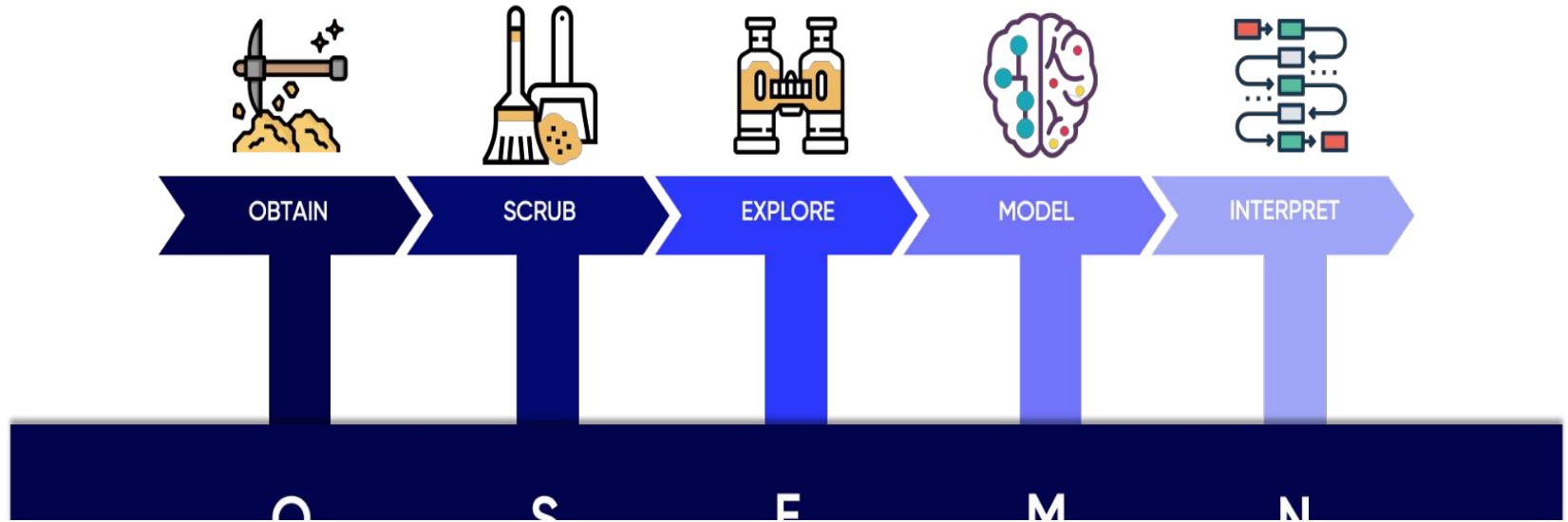
Overview

King County, Washington:

- Estimated population of 2,233,163 people
- 952,569 houses
- Median Value of $446,600
- https://www.census.gov/quickfacts/fact/table/kingcountywashington,US/PST045218
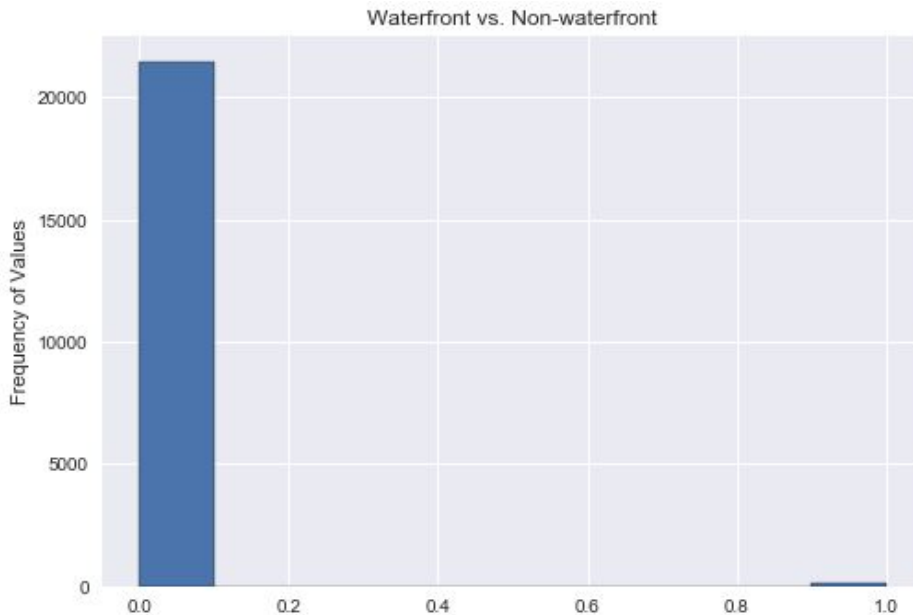
Our Dataset:

- 21,597 houses
- Median Value of $450,000

Question 1

How does the location affect the price of a house in King County?

- Waterfront (mean): $1,717,215
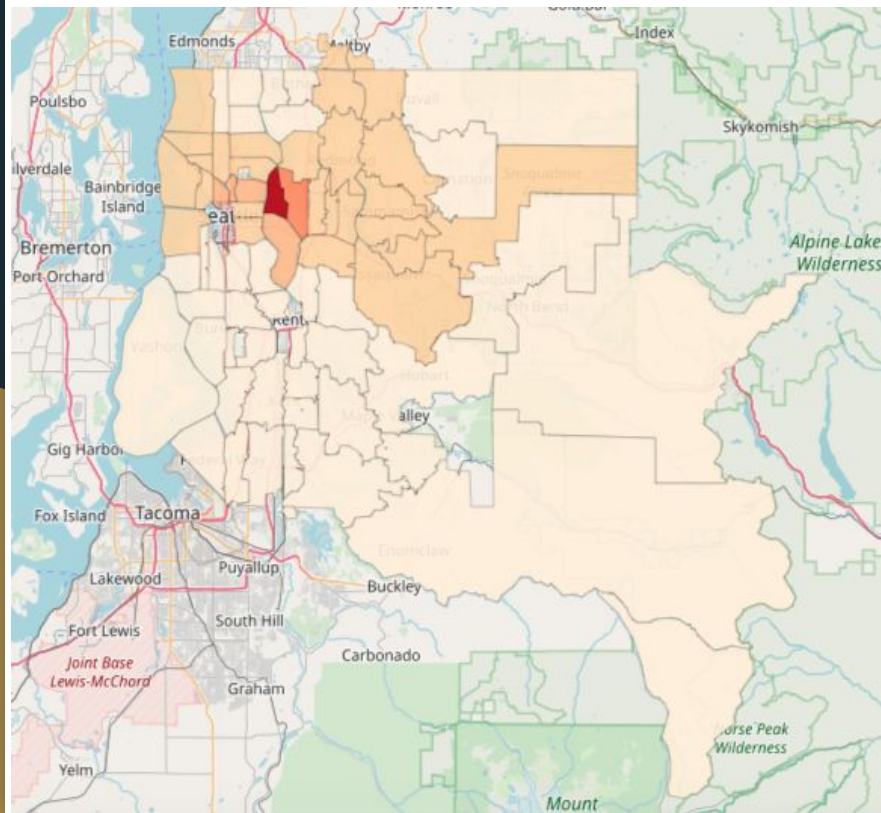- Non-waterfront (mean): $532,286
- $1,184,929 difference

Waterfront vs. Non-waterfront

Q 1 continued…

Zip Code:

- **98039**
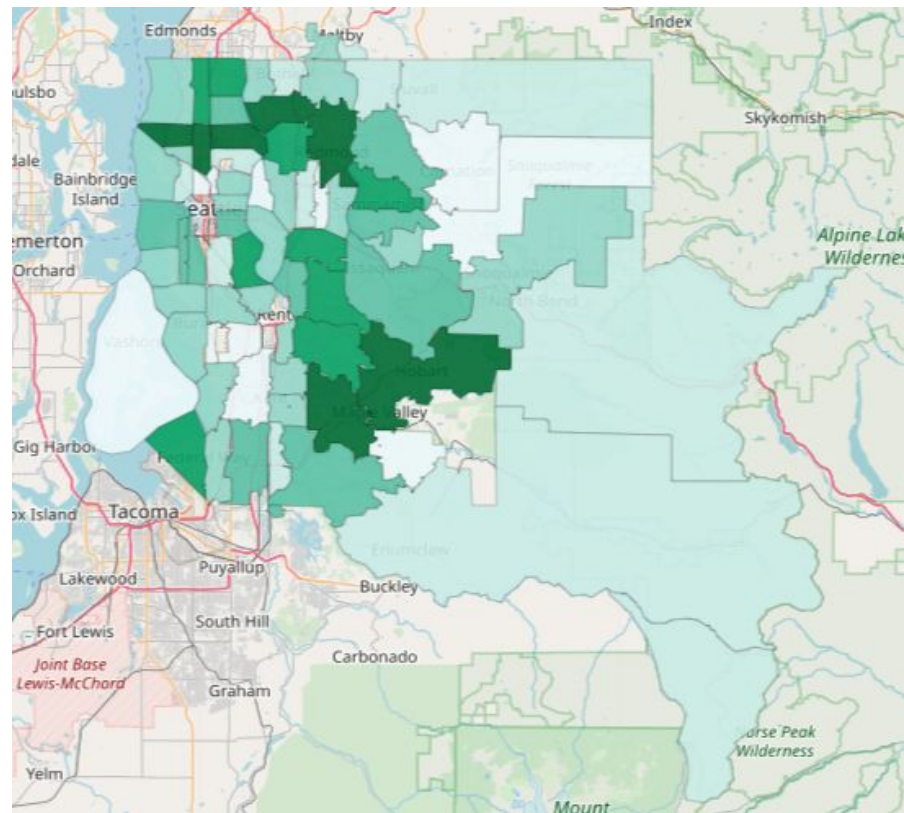- **Mean Price $2,161,300**

Rural/Urban:

- **Count of Houses**

Zip Code vs. Price

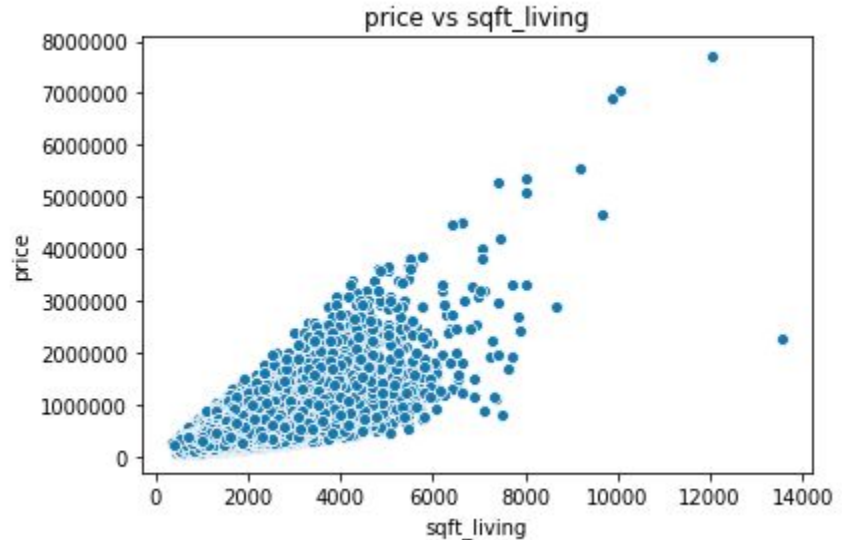## Zip Code vs. Count

Question 2

What is the relationship between the size of a home and the price?

Sqft_living

- 25%   1,430 $322,000
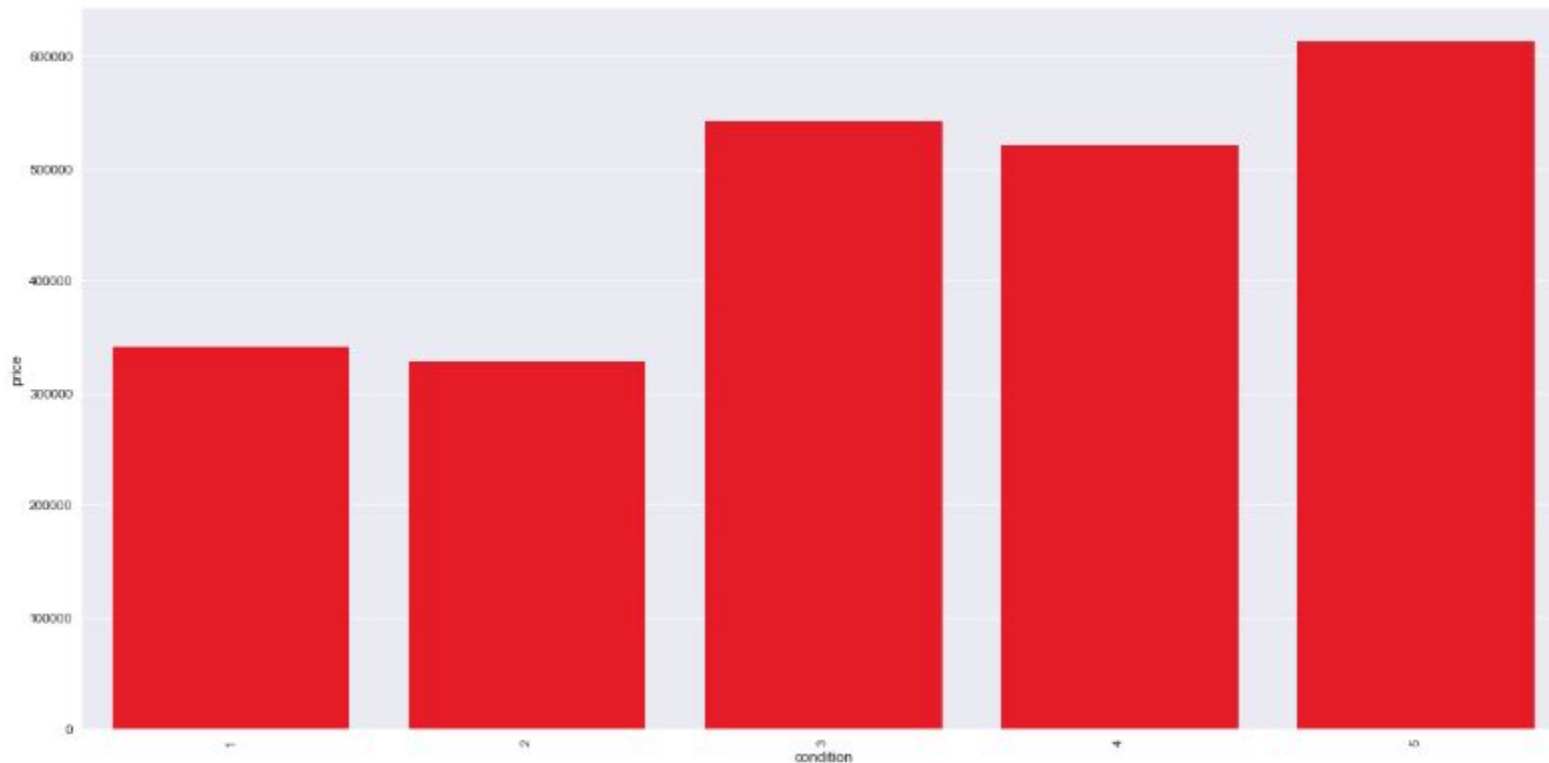- 50%   1,910 $450,000
- 75%   2,550 $645,000

- $450,000 / 1,910 =

    $235.60 per square foot

Question 3

How does the condition of the house affect price?

Model

- **Best Features:**
  - Sqft_living - 0.4919
  - Zipcode - 0.3499
  - Sqft_lot - 0.181
- **Adjusted R-Squared**: 0.862
- **86.2% predictive power**

- **Train Mean Squared Error**: 0.0022
- **Test Mean Squared Error**: 0.0023
- **Train RMSE**: 0.047
- **Test RMSE**: 0.048

Questions and Answers