Predicting Home sale Prices in King County, Washington

Identifying the most important features that can help increase the sale price of your home

Business Goals

➢ Develop A model to estimate Real Estate Prices in King County, Washington from housing features

➢ Create a Tool for sellers and agents to maximize the sale price of a home based on housing features

    ➢ Is location important to my sale price?

    ➢ Will the size of my home help determine its price?

    ➢ Should I consider renovating before I sell my house?

    ➢ Should I aggressively show my house during the selling process?

Features of the Kings County Data set

- **price** - is prediction target
- **bedrooms** - of Bedrooms/House
- **bathrooms** - 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 (Overall)
- **grade** - overall grade given to the housing unit, based on King County grading system
- **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

Multiple Linear Regression Model

➤ Our model depends on identifying housing features that have increasing values that correspond with increasing house price

➤ For example it makes sense that larger houses cost more. This should be a good predictor of price. Indeed this is the case. Square footage is a good predictor of price and was included in our model.

➤ Location is also extremely important factor in real estate pricing.

The two key elements for size and location used in our model

Square footage and Latitude plotted vs Price(log)
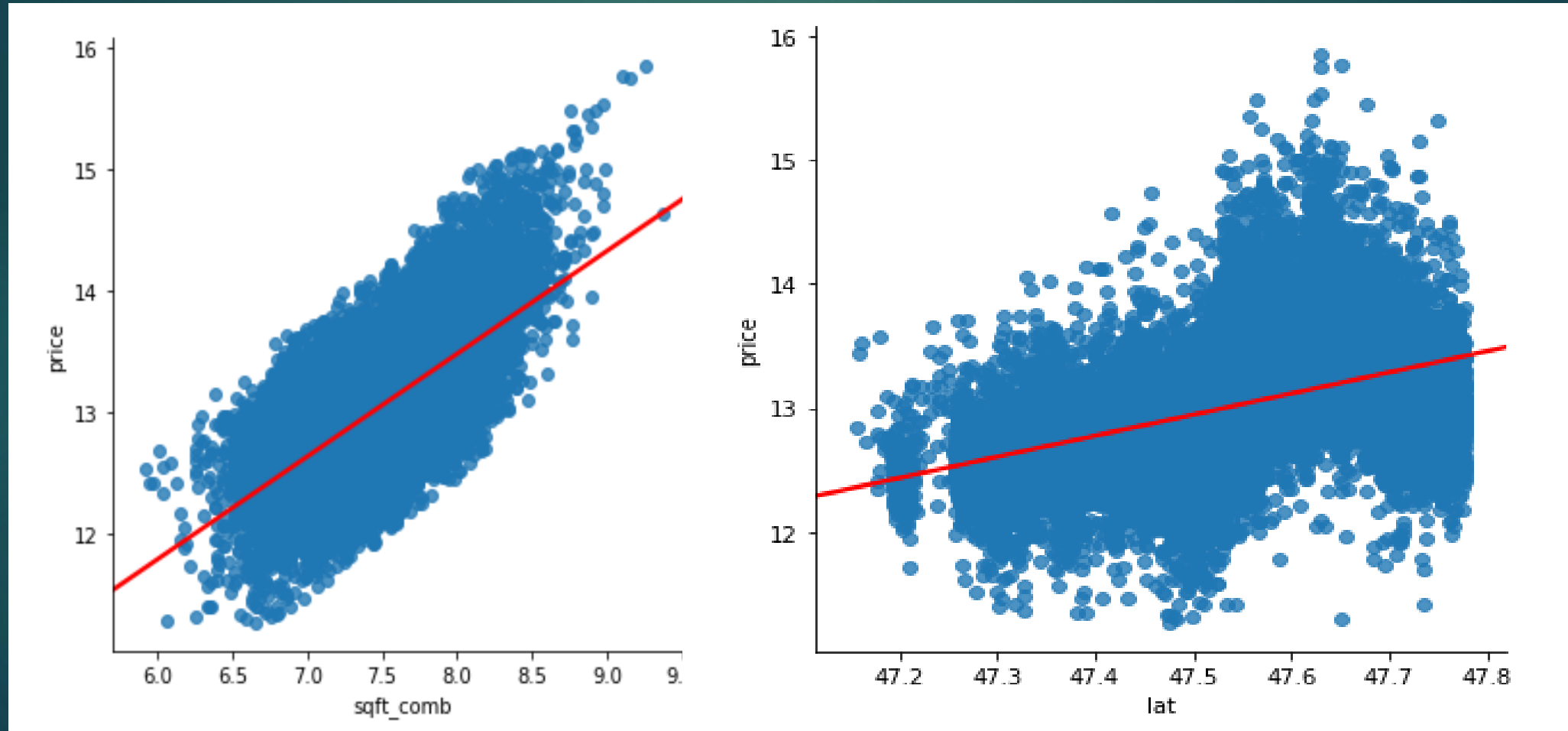
Features of the Model

➤ After analysis it turns that the most important features for predicting price were, in order of importance:

  ➤ Latitude- our feature representing location

  ➤ Sqft Combined- our feature representing size

  ➤ Waterfront

  ➤ Grade

  ➤ View

|	price	Predicted_Sales
0	221900.0	313846.0
1	538000.0	596932.0
2	180000.0	314335.0
3	604000.0	357803.0
4	510000.0	504882.0
5	1230000.0	1359621.0
6	257500.0	272676.0
7	291850.0	255788.0
8	229500.0	344070.0
9	323000.0	311770.0
10	662500.0	614833.0

As can be seen our model has predictive power

However there are also clear limitations

Business Recommendations

- Know the market for your Location. Maximize your price based on your location

- If there is a renovation budget consider additions. Size is an important driver of price.

- If you are waterfront take advantage. Your property is worth more.

- If you do renovate. Have your house assessed after renovations. Renovations and a higher grade will both add value

- Aggressively market and show your house
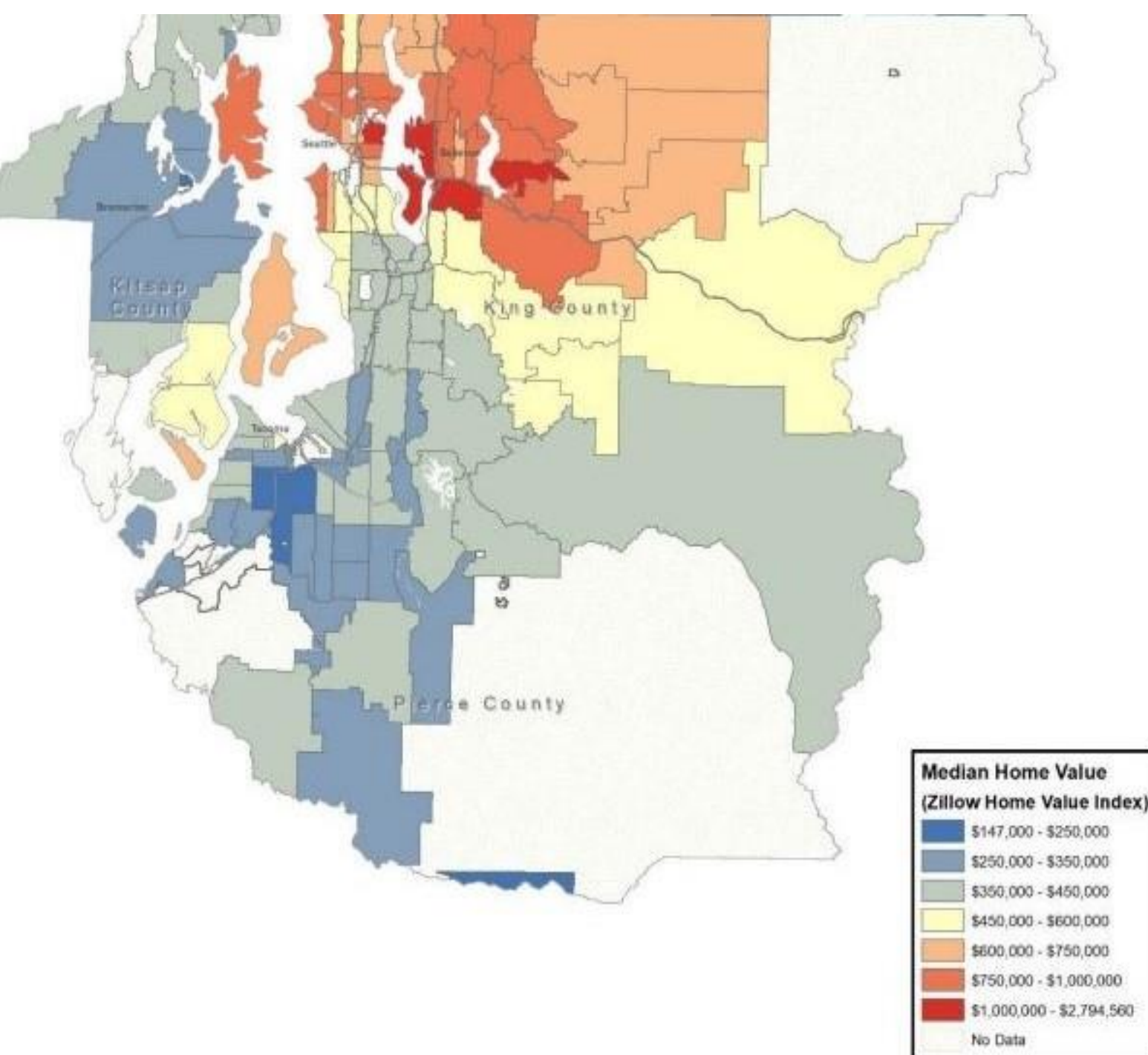
Median Home Values Around King County (from Zillow)

IT IS CLEAR FUTURE MODELING WILL BE ENHANCED WITH MORE PRECISE LOCATION DATA

Future Direction

- Digging deeper into location

  - Obtain better geographical and location data
  - Use more detailed geographical price data for clustering and more powerful Machine learning tools

- Add multi-year data

Thank You!