

An aerial photograph of Seattle, Washington, showing the city skyline and waterfront. The image is used as a background for the title slide. The text is overlaid on the upper left portion of the image.

# *Module 2 Project*

## *Predicting King County House Prices*

Joe Buzzelli

March 6, 2020



## Project Scenario

This presentation is a mock project in support of **Redfin Now**

- Redfin began its direct buying program in **Q1 of 2017**
- This presentation presents the mock results of a pricing effort in support of Redfin Now's effort to enter the **King County, WA** market

# Predicting housing prices in King County Washington

**Statement of Work:** Empower Redfin Now to predict housing sales prices in a pilot focusing on King County, WA. Key questions include:

- What housing features drive selling prices?
- What level of accuracy can we anticipate?
- How can Redfin Now reliably forecast profitable housing purchases with as few data elements as possible?

# Assumptions and data sources

## Assumptions

- The analysis in this deliverable is scoped to **King County, WA** specifically and the resulting model is not intended for usage outside of this geography
- Only houses are included in this analysis, apartments and condos are not in scope of this effort

## Data sources

- Housing data King County, WA from **9/9/2014** to **1/10/2015**

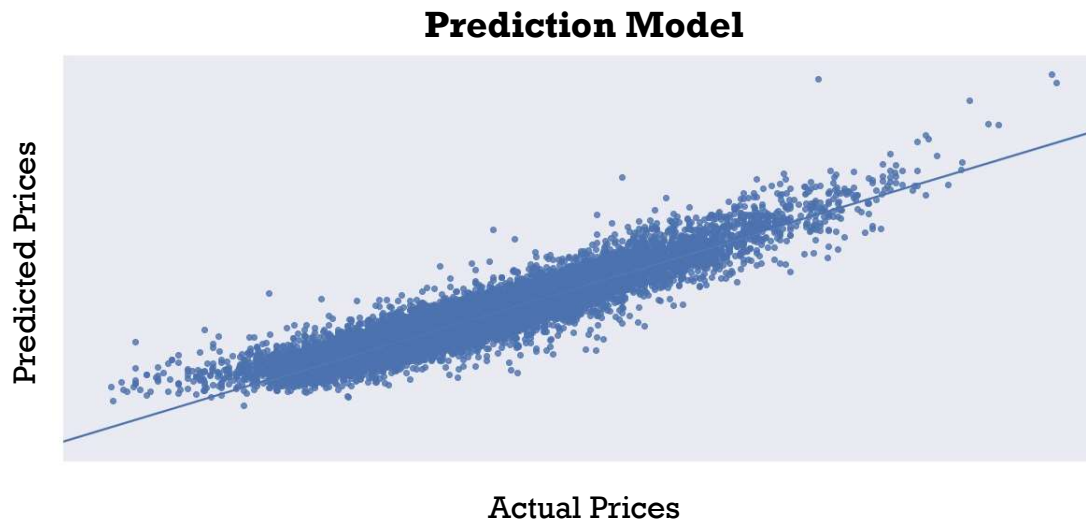
# 16 data elements provide the initial data for price predictions

The initial data set included **21** data elements, but only **16** are relevant to this analysis to include:

Number of Bedrooms	Number of Bathrooms	Living Space	Lot Size
Floors	Waterfront Location	Times Viewed	Condition
Grade	Living Space Less Basement	Basement Size	Year Built
Year Renovated	Zipcode	Living Space for 15 Closest Neighbors	Lot Size for 15 Closest Neighbors

## Price prediction model describes **87%** of price variance

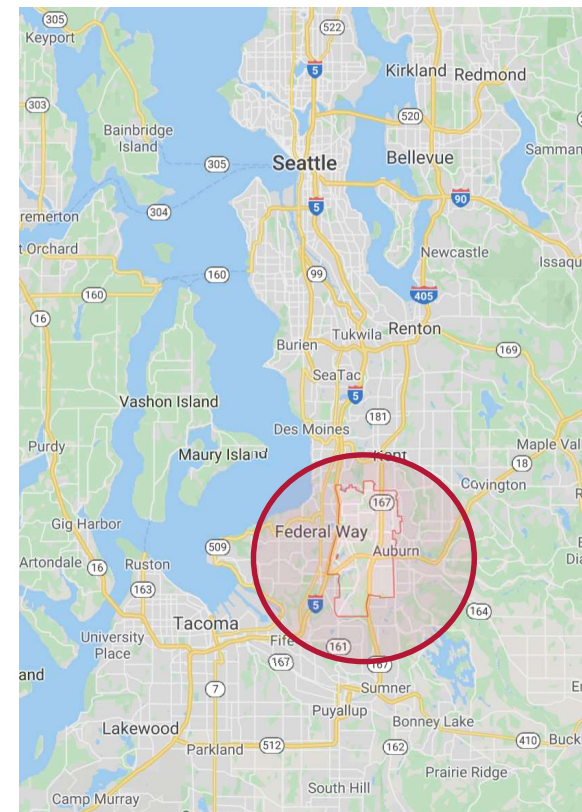
- The model predicts what a house price ***should be*** based on pricing data but there is variance with real prices



## Zip code location significantly affect house prices

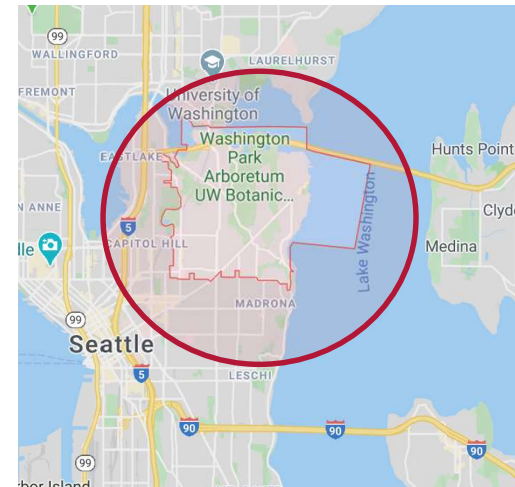
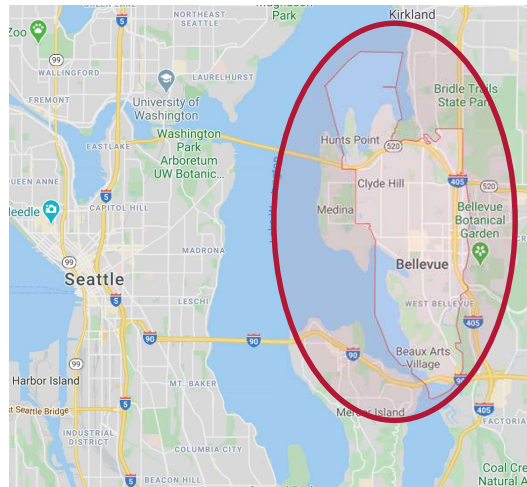
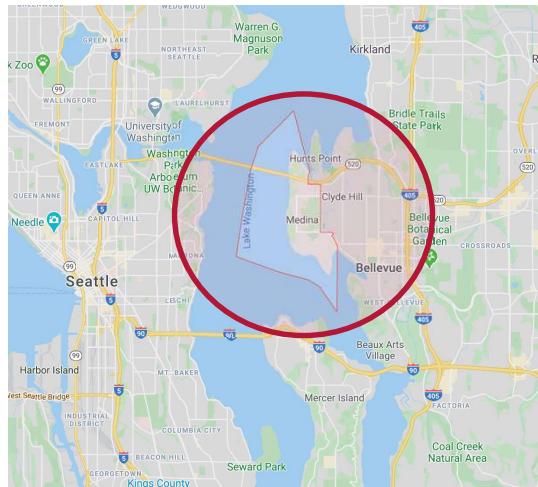
Zip codes represent the top **31** most impactful features with the top **16** exhibiting at least a **2x** from the baseline zip code

- Due to the backend statistical processes, the zip code predictions derive from Federal Way, WA 98001
- The average price of a house in Federal Way equals approximately **\$289,000**
- For reference, the average house price across the entire data base equals **\$561,000**



# Location, location, location

Top 3 Zip Codes	Price Factor (vs Federal Way)	Average House Price in Federal Way	Forecasted Price in Other Location
Medina, WA 98039	3.3x	\$289,000	\$953,700
Bellevue, WA 98004	2.9x	\$289,000	\$838,100
Seattle, WA 98112	2.7x	\$289,000	\$780,300





# Features affect house prices drastically less than zip codes

The top three housing features affect house prices far less than zip codes

Feature	Price factor
Waterfront Location	1.59x
Grade	1.10x
Times Viewed	1.06

## Redfin Now can use this model to automate house purchases

With the pricing model integrated into its technology stack, Redfin Now can assess the profitability on a house by house basis and offer beneficial offers to perspective sellers



# Next Steps



**REDFIN** | NOW

- Continue to reduce the total number of features required for the model
- Change baseline zip code to one of Redfin's choosing
- Increase the explained variance higher than the current **87%**
- Integrate machine learning to enable the model to learn as Redfin Now purchases and sells more homes

Thank you  
for your time



Are there any  
questions?