Hello

Hello, my name is Mike Erb

TheKeys

Today, I will be giving a short presentation about the status of the the housing price prediction and housing recommendation portions of our startup, TheKeys.

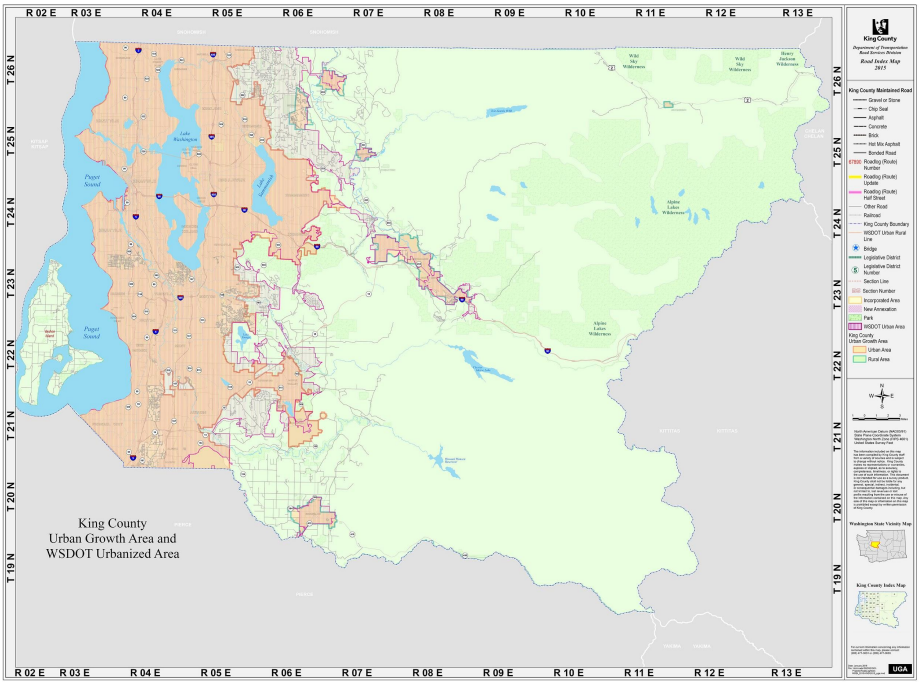
I completed an exploratory data analysis of our initial housing dataset and created a linear regression model to predict a house's price depending on its characteristics and location.

In this presentation I will speak about how location and two example house characteristics influence price, then give my recommendations for how to proceed.

King County, Washington

As you may know, we chose to launch our beta service in King County Washington.

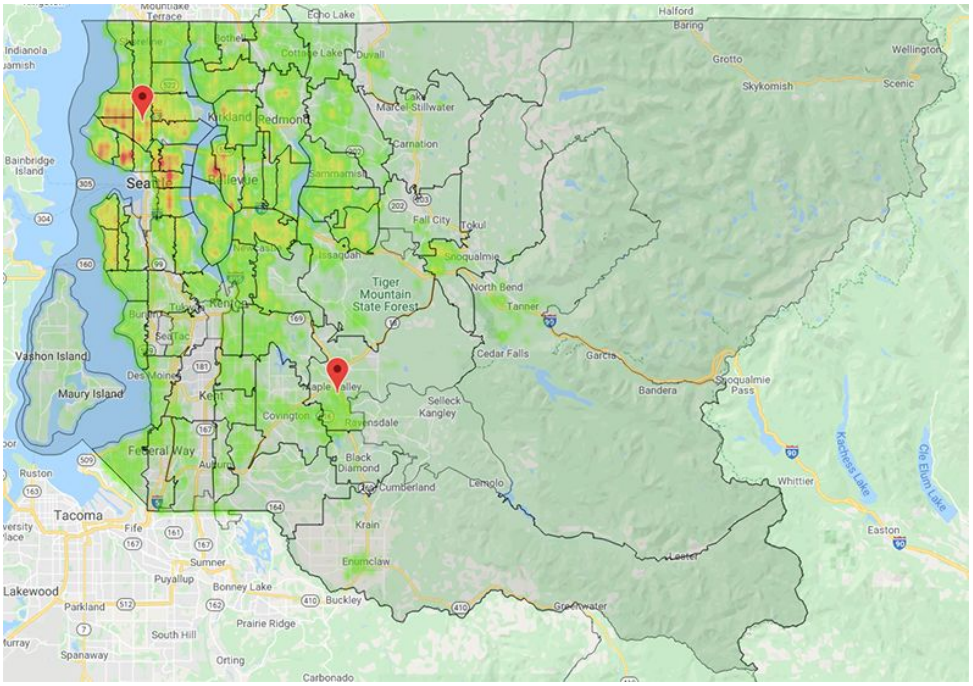
The first thing we need to understand is the area itself.



As you can see in this map provided by the King County DOT...

King County is predominantly rural...Urban area (approximately ⅓ of the total) is designated in orange and rural area is designated in green.

We wouldn't expect rural areas to have active housing markets...



and they don't...

This is a heat map of King County showing the gross value of housing sales and the zip code boundaries.

As the color changes from green to yellow to red, the value of the total housing sales for an area increases.

From this map we can easily see that the most valuable and active housing market is concentrated up here close to Seattle.

What is not readily apparent from the heat map is that down in Maple Valley is the 2nd most active zip code for housing sales.

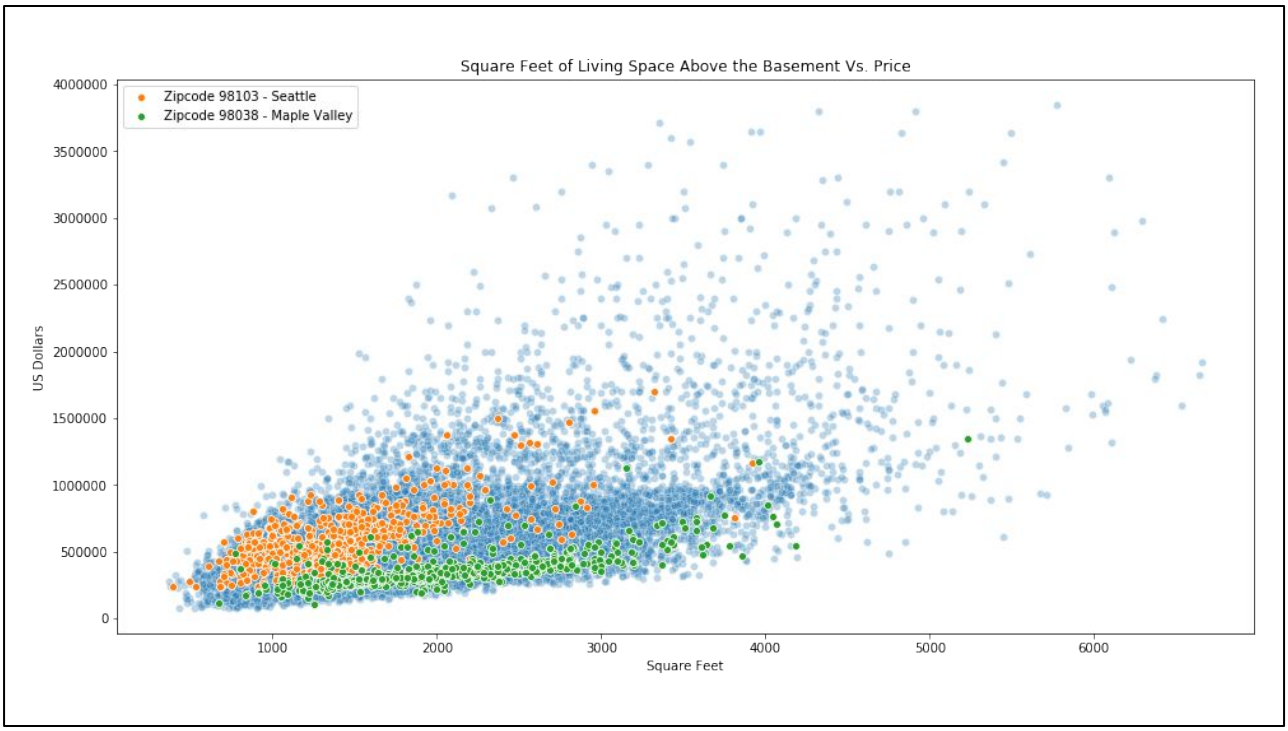
Despite it being the 2nd most active, its housing prices are below average, and consequently not as hot as Seattle which has above average prices.

Location Matters

Which brings us to the truism confirmed by our data and model -- “Location Matters”

Given our currently available data, the smallest level of analysis we have is by zip code, but as the previous map showed, it is informative.

Depending on the zip code alone, the predicted price of a particular house can vary from a base amount of $390K by a decrease of $47K all the way up to an increase of $620K



Let's return to our two example zip codes, one part of Seattle, shown in orange, and the other Maple Valley, shown in green, with all the houses in King County shown in blue.

This scatterplot shows how as the square footage of living space above the basement of a house increases so does the price of the house. Our model predicts that for every additional sqft foot of living space, the mean house value would increase by about $144.

However, as can been seen in the two example zip codes, *location matters* not just for prices, but also
when considering the size of the houses available.

For Example, we can observe that:

In this Seattle zip code the houses are generally under 2000 sqft and have a wider range in prices at a given house size, while in Maple Valley there are plenty of houses up to 3000 sqft but the range of prices for a given house size is relatively narrow.

And

In the high value location of Seattle, even a small 1000 sqft house may be more expensive than a 3000 sqft house in Maple Valley.

bathrooms **>** bedrooms

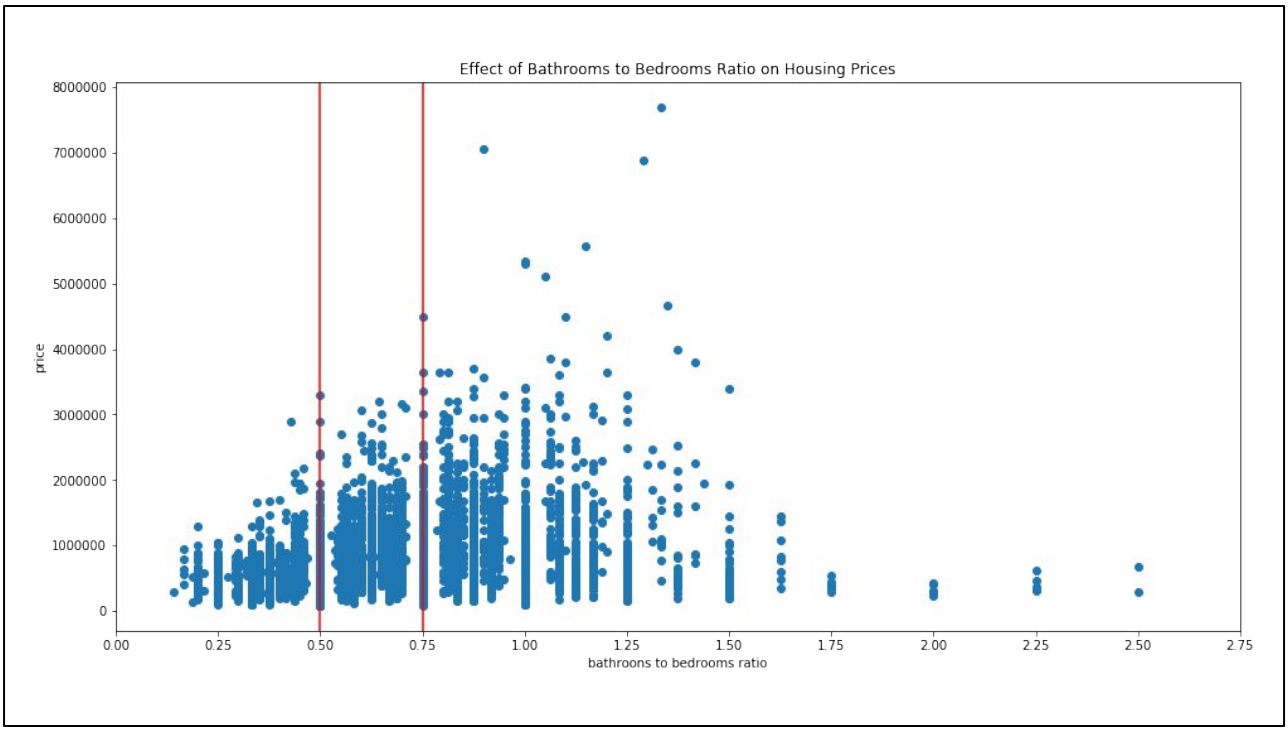
Moving on to some more specific housing features...bathrooms and bedrooms.

When considering a house, a buyer will want to consider how many people are expected to live in that house. This often means a bedroom for the parents, plus one each for the children, and maybe a guest room.

But then how many bathrooms? Is there a bathroom for each bedroom, maybe more, maybe less?

In our dataset, the number of bathrooms ranged from a partial bathroom up to 8 full bathrooms, while the number of bedrooms ranged from 1 up to 11. We have to consider this the set of realistic houses. We don't have houses without bathrooms and we don't have houses with a lot more bathrooms than bedrooms.

So why are bathrooms greater than bedrooms? Because for realistic houses adding a bathroom will always increase the price, while depending on the number of bathrooms, increasing the number of bedrooms may lower or may raise the price of the house.



Let's look at how you can lower the price of a house by adding a bedroom (assuming everything else stays constant)

This visualization compares the ratio of bathrooms to bedrooms vs. housing price.

You can see that as the ratio of bathrooms to bedrooms approaches 1, the housing prices are increasing, but after one the price generally starts to decrease.

Now for an example of how to lower the price of your house. 50% of the houses are in the 0.5 to 0.75, or 1 bathroom per 2 bedrooms to 1-½ bathrooms to 2 bedrooms section bounded by the two vertical red lines , so I will take house from that set.

Say I have a house with 2 bedrooms and 1-½ bathrooms with a ratio of 0.75. A family is living there and has a another child and they want to give that child their own room, so they turn some living space into a bedroom. Now the house has 3 bedrooms and 1-½ bathrooms, with a ratio of 0.5. Instead of increasing the mean sale price of the house, our model predicts it will go down by over $11,500.

Conclusions

In conclusion:

Housing Recommendations

Ready

We should move forward with the housing recommendation portion of our service. We can offer suggestions on a general price range for a location or housing characteristic that would help a homebuyer or employer make salary or employment decisions.

Housing Price Prediction

Not Ready

Our housing price prediction model needs improvement. While the predictions we can make for price changes on a particular house characteristic or location are significant, the overall house prices predicted by our model are not accurate enough yet to launch a public beta of our service.

MORE DATA

To make improvements to our model, we need MORE DATA.

To start, better, neighborhood level, location information, historical sale and assessment information for each house, and real time localized information about the housing market.

Thank You