



DERIVING HOUSE SALES PRICES IN KING COUNTY

JAN ABDULLAH

Introduction

- ▶ We take a look at the King County Housing sales prices dataset and create an inference model where we can deduce some features that drive higher sales prices.
- ▶ We'll focus on a few categorical and continuous attributes that drive higher sales prices.

Additional questions to consider:

1. Taking a zipcode-centric approach, which zipcode generates the highest sales prices?
2. Does having a waterfront property drive up the price?
3. What are other general features that affect the price?

Which zipcode generates the highest prices?

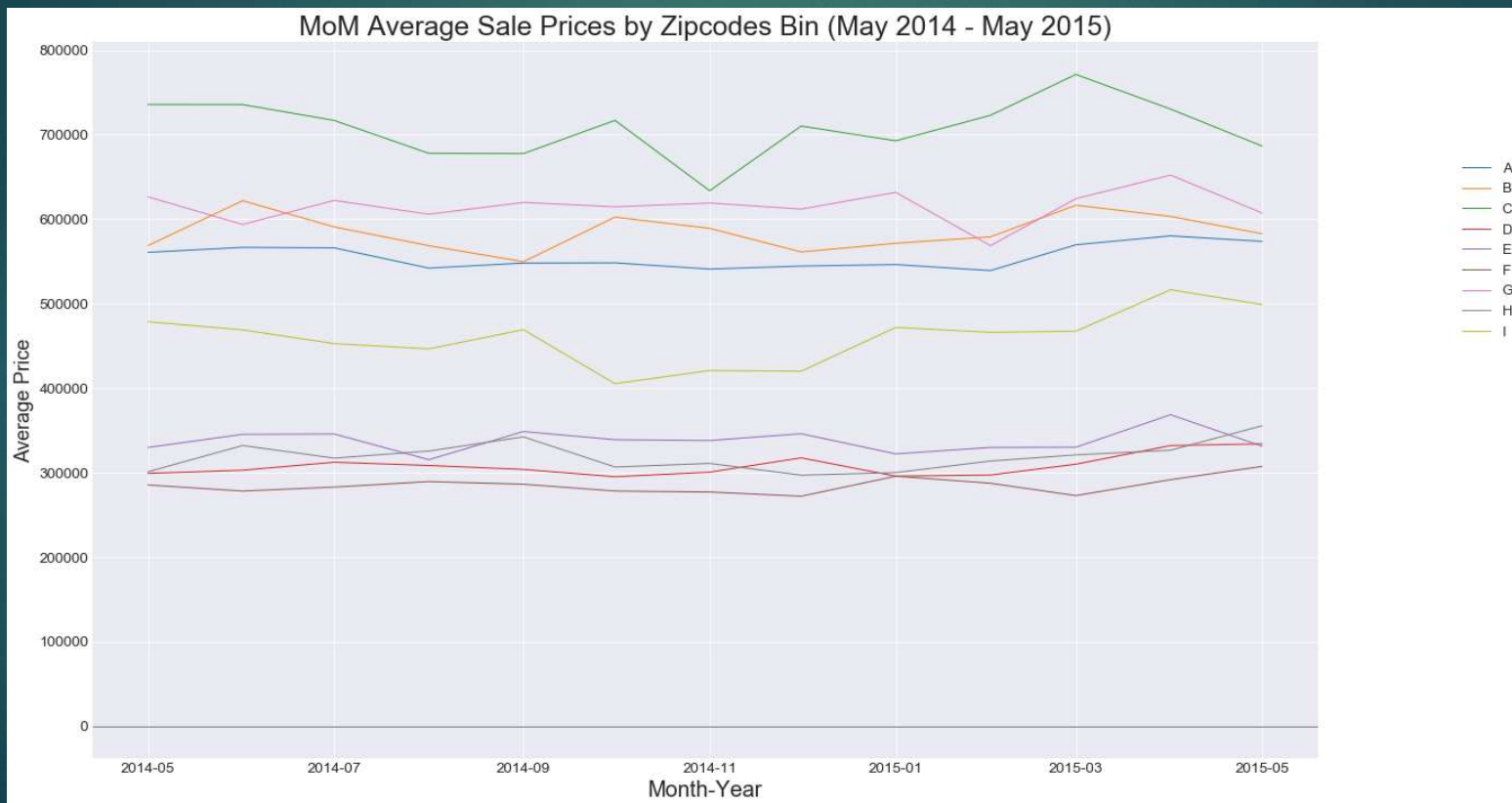
There are a lot of different zipcodes in our dataset. Thus, for simplicity, let's bin zipcodes of similar neighborhood attributes into values "A" through "I".

See the bin layout below:

```
"A": [98177, 98133, 98125, 98155, 98107, 98115, 98105, 98109,  
      98112, 98117, 98199, 98122, 98144, 98118, 98108, 98106,  
      98136, 98126, 98116, 98102, 98103, 98119, 98134  
    ],  
"B": [98053, 98052, 98077, 98072, 98033, 98034, 98011, 98028],  
"C": [98059, 98056, 98006, 98040, 98008, 98007, 98005, 98004, 98039],  
"D": [98057, 98055, 98058, 98032, 98031, 98030, 98178],  
"E": [98070, 98146, 98168, 98166, 98198, 98148, 98158, 98188],  
"F": [98023, 98003, 98063, 98047, 98001, 98002, 98093, 98071, 98002, 98092],  
"G": [98074, 98075, 98029, 98027, 98019, 98014],  
"H": [98022, 98042, 98010, 98038, 98051],  
"I": [98065, 98045, 98024]
```

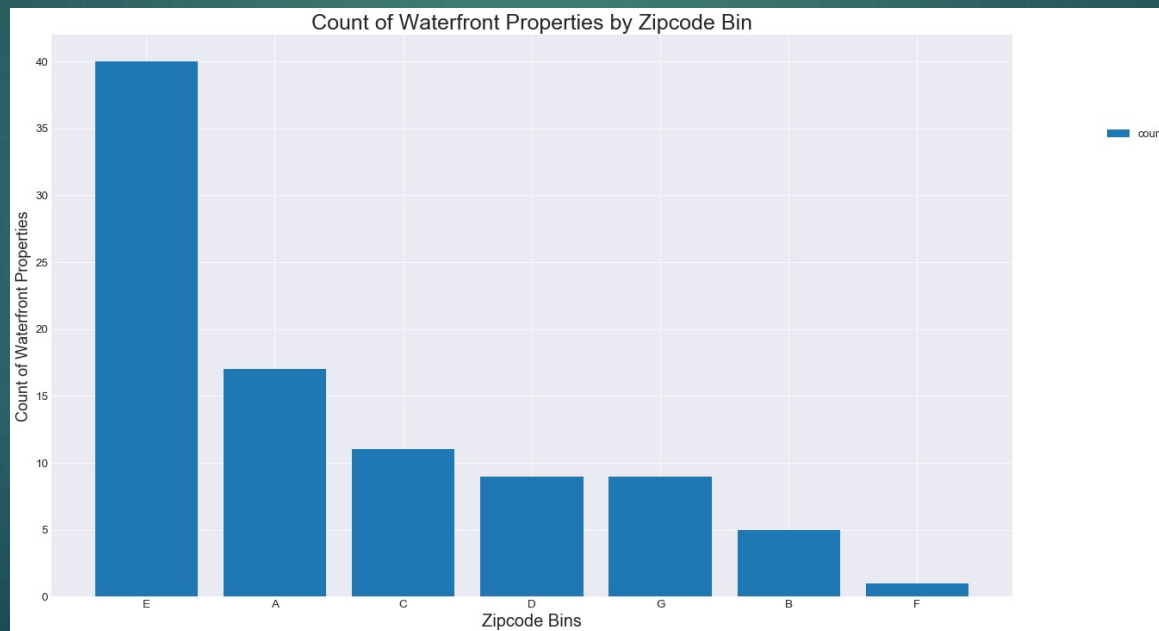
Which zipcode generates the highest prices? (cont.)

We take a look at the Month-on-Month average sales prices in each zipcode bin between May 2014 and May 2015.



Does having a waterfront property drive up the price?

- ▶ The average sales price of a waterfront property in our dataset is \$1,144,775.54, whereas the average for a non-waterfront property is \$509,600.52.
- ▶ Thus, we do expect, given all else is constant, that waterfront properties will drive up sales prices.
- ▶ Lets take a look at the concentration of waterfront properties by zipcode.

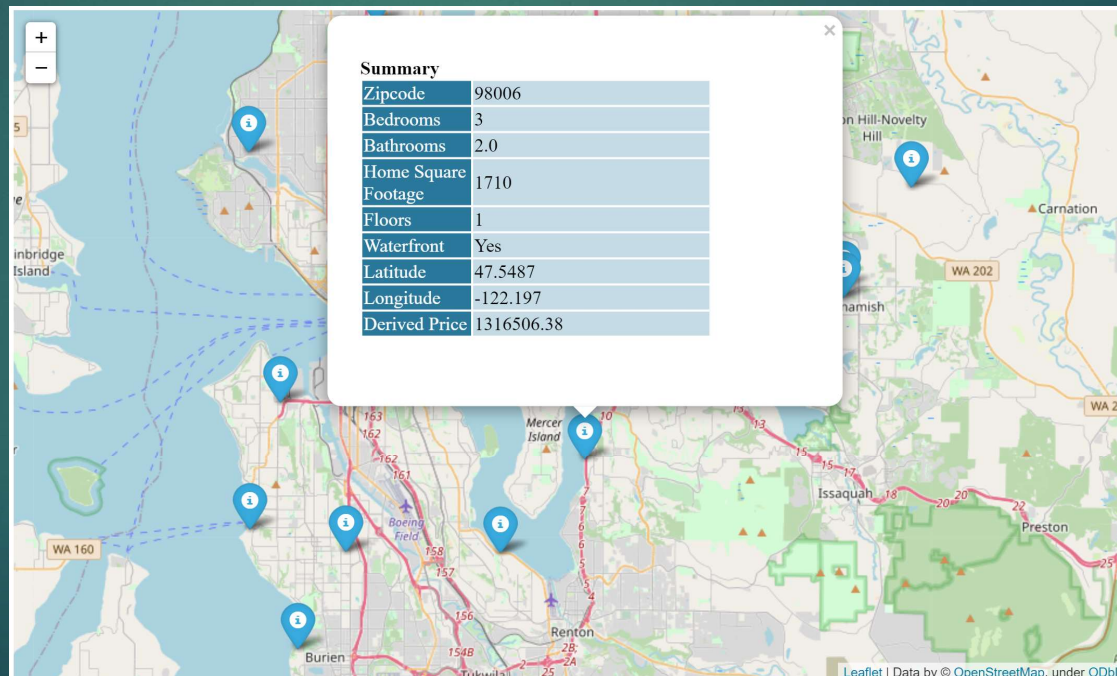


What are other general features that affect the price?

- ▶ So far we know that a waterfront property and a property in zipcode bin 'C' will generate higher sales prices.
- ▶ To quickly explain the other features and coefficients, here are some other variables that increase the sales prices, according to the model:
 1. More bathrooms.
 2. Higher grade.
 3. Higher house square footage.
 4. Higher condition.
 5. More floors.

Conclusion

- ▶ In summary, you're likely to sell a house at a much higher price if it is a waterfront property in a zipcode in zipcode bin 'C'.
- ▶ Let's look at an example of an input with these features, derive the price, and plot it on a map.



Thank you!

