

A Data Analysis of Houses in King County

for homeowners to maximize the price of their homes

by Steven Rosa

What was the data?

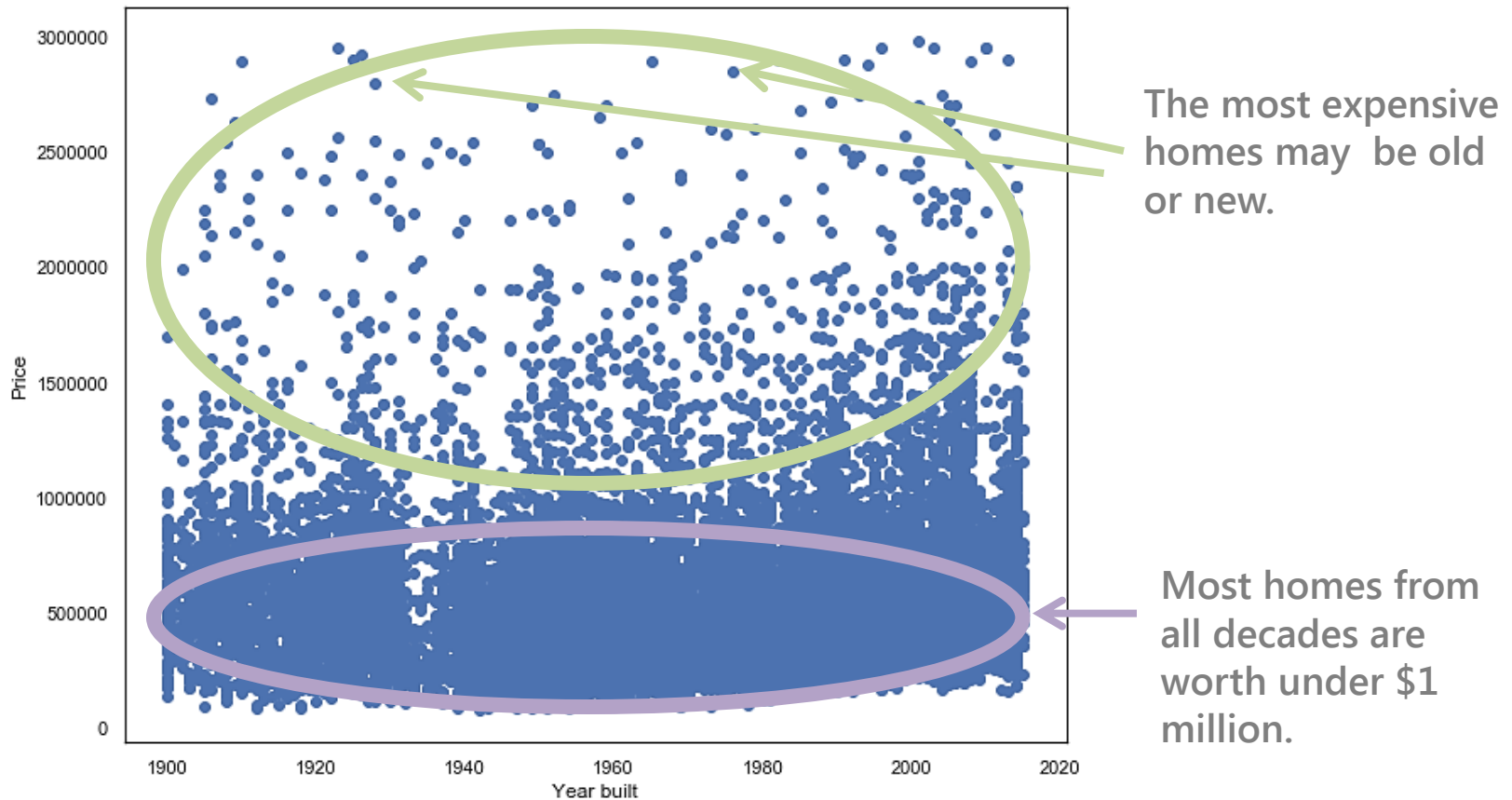
A set with information on 21,000 houses in King County.

What were the results?

We built a model that can predict the price of a house in King County up to \$3 million. It uses variables like square footage, zip code, the condition of the house, &c.

We can use the model to answer questions about what's important and what's not important in predicting the price of a house.

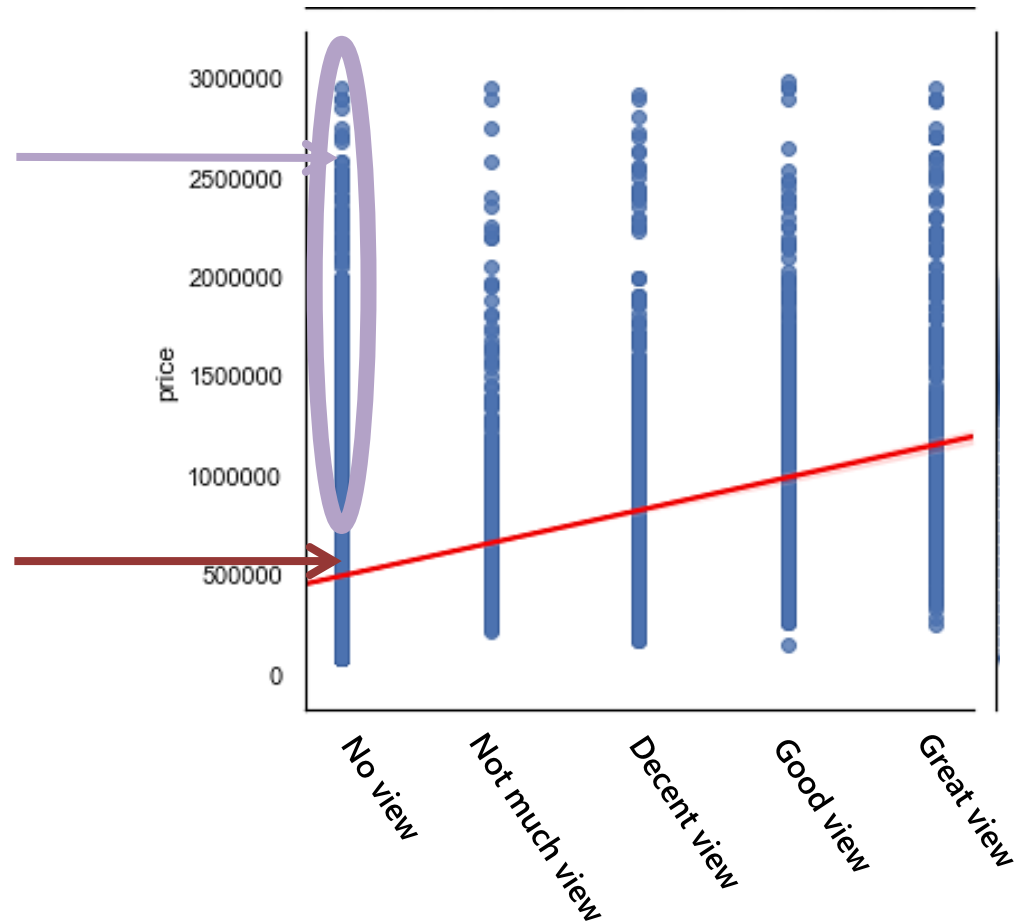
For example, the year your house was built does not predict its value.



Or, to take another example, notice that house price does rise a little as the view from your house improves. But you don't need to worry about it.

Most houses have no view, including plenty of the highest-value houses.

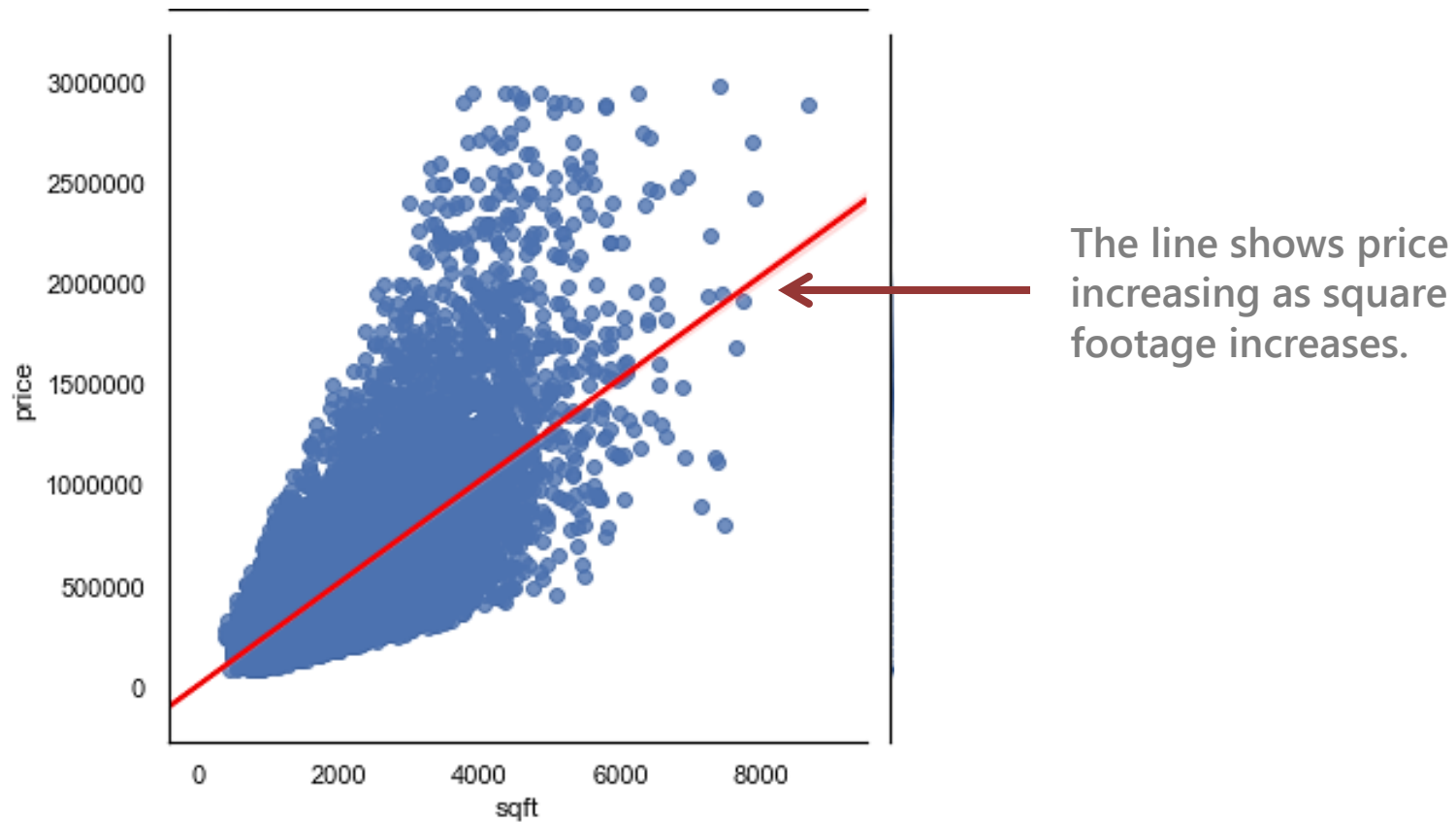
Price increases very gradually as view improves.



So which variables *are* important?

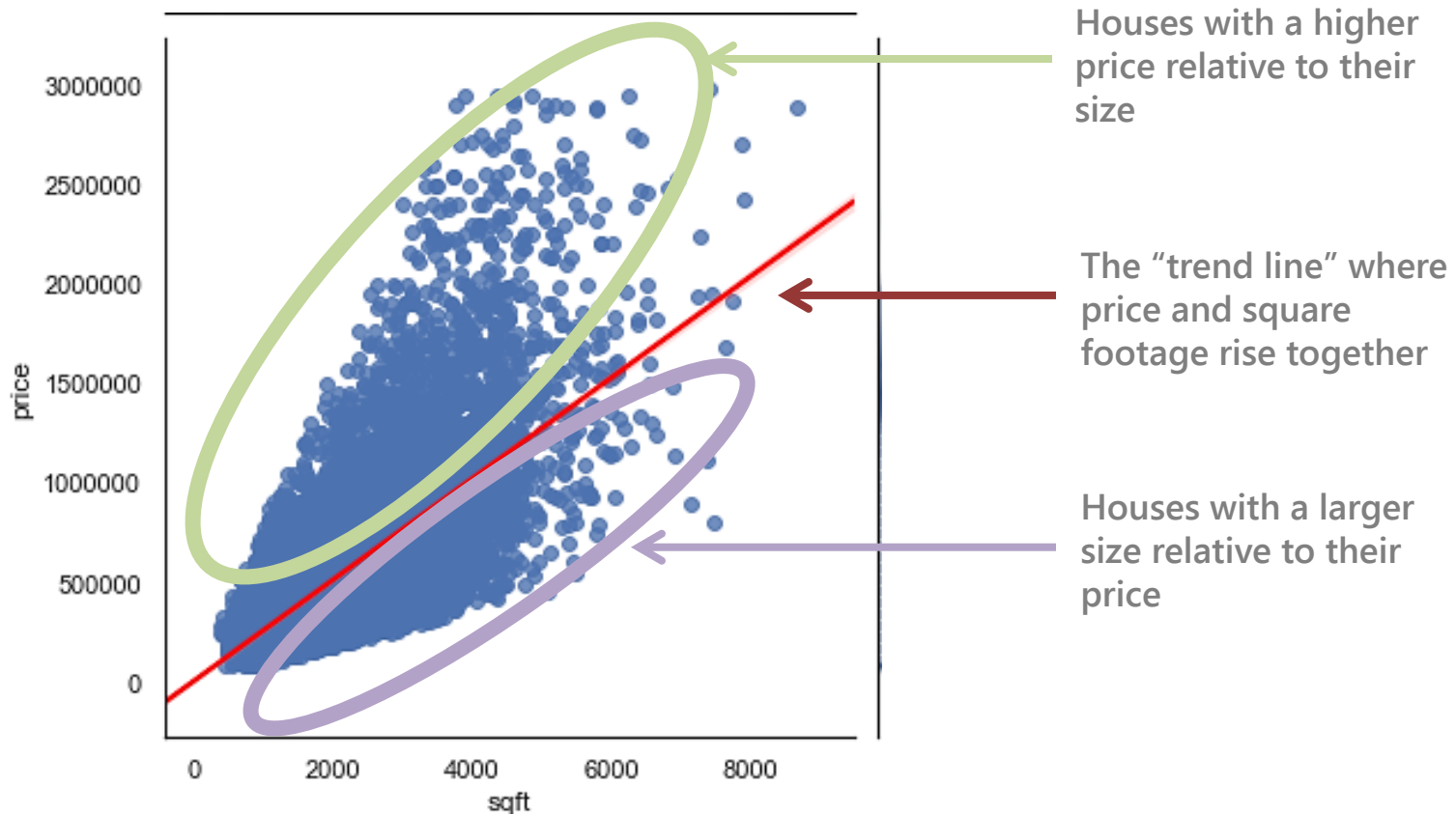
Size matters.

In general, price increases as square footage increases.



But it's not a perfect relationship.

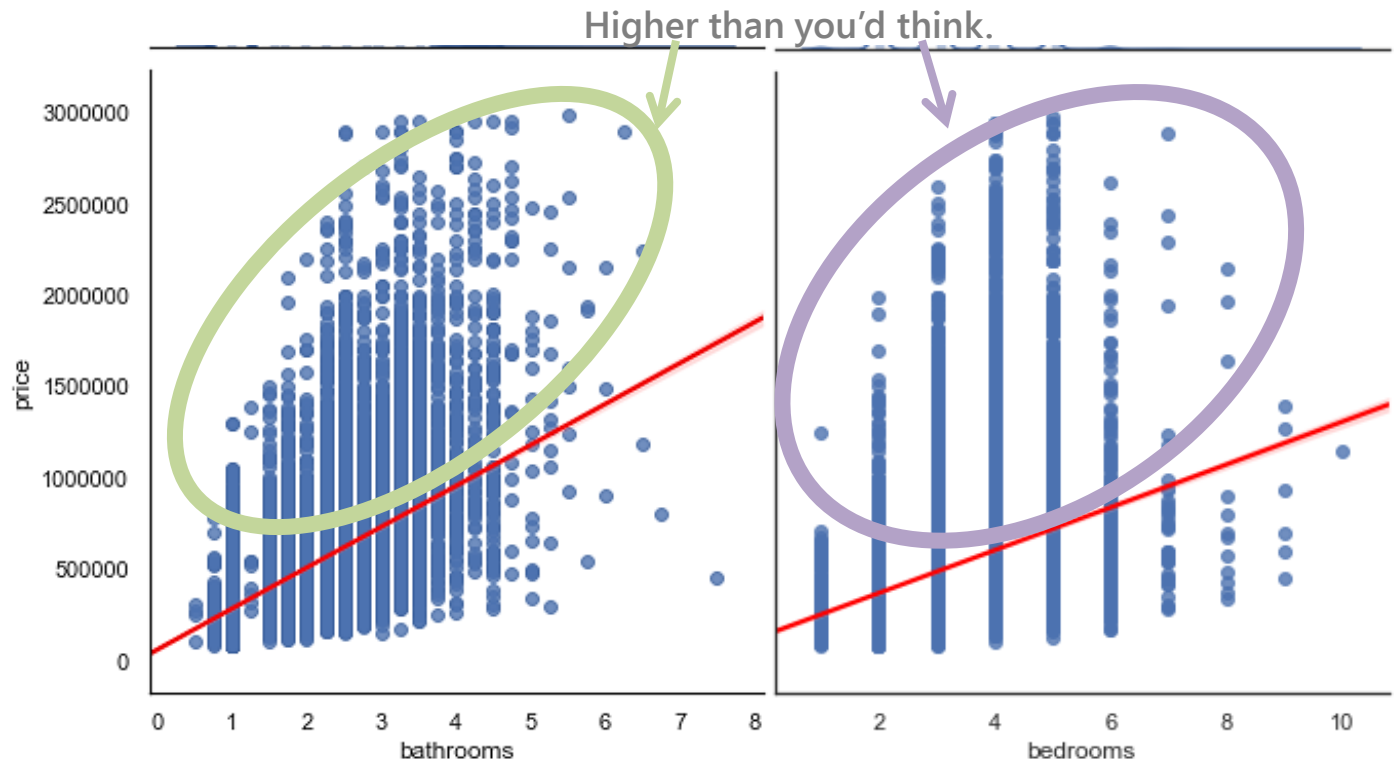
Square footage on its own cannot accurately predict house price. Other factors must be taken into account.



Rooms in the house

The number of bathrooms and bedrooms also contributes to your house price. But, like square footage, these factors can't predict the price on their own.

A lot of houses have higher prices than you'd think, based just on their number of bathrooms or bedrooms.



Location.

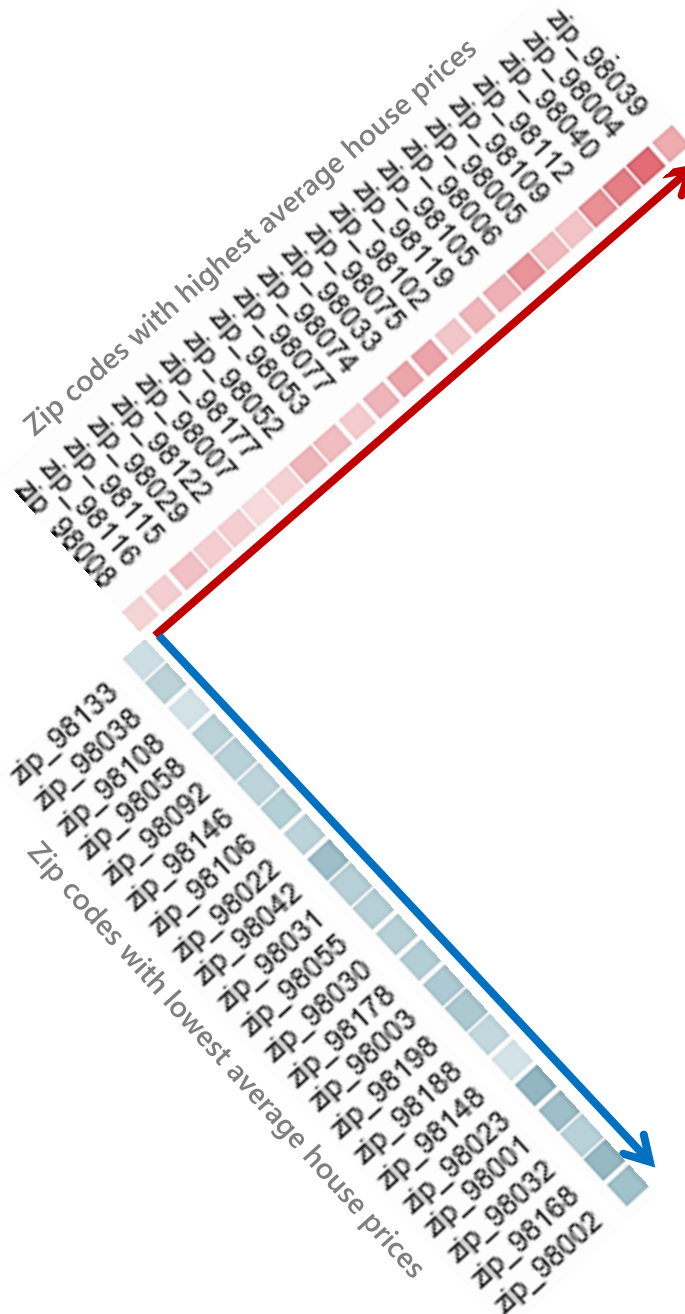
Apart from those other variables, it turns out that location matters a lot.

High-value houses tend to be grouped together in certain zip codes; likewise for low-value houses. In fact, all other things being equal, your zip code will predict your house's price quite well.

A list of the top 10 zip codes by average house price will look familiar to those who've explored the fancy areas of King County:

1. 98039: Medina
2. 98004: Bellevue
3. 98040: Mercer Island
4. 98112: Madison Park
5. 98109: Queene Anne
6. 98005: Bellevue
7. 98006: South Bellevue
8. 98105: U District
9. 98119: Queene Anne
10. 98102: East Lake

Location,
Location,
Location.



As average house price increases above the middle of the full range of prices, the predictability of house price in that zip code increases.

As average house price decreases below the middle of the full range of prices, the predictability of house price in that zip code increases.

What to do?

If you're a homeowner, the price of your house is likely to rise or fall in tandem with the rest of your neighborhood. By working together with your neighbors to make improvements to the buildings, conditions, and services in your area, you may raise the average house price, and thereby increase the likelihood of raising the price of your house.

And note that undertaking renovations to your own home may not increase its selling price if the average house price in your neighborhood remains low.

If you're looking to buy, consider that a home with your desired size, number of rooms, view, &c. will likely cost more in a higher-valued zip code than it will in a lower-valued zip code, even though all its other features may be the same.

If you want the best possible house and don't care about the surroundings, look in a zip code with low average house prices.