

Presentation for: *We Invest in KC Houses, LLC*

# TAKE IT EASY AFTER CHRISTMAS . . .

---

FACTS YOU SHOULD KNOW WHEN SELLING A PROPERTY



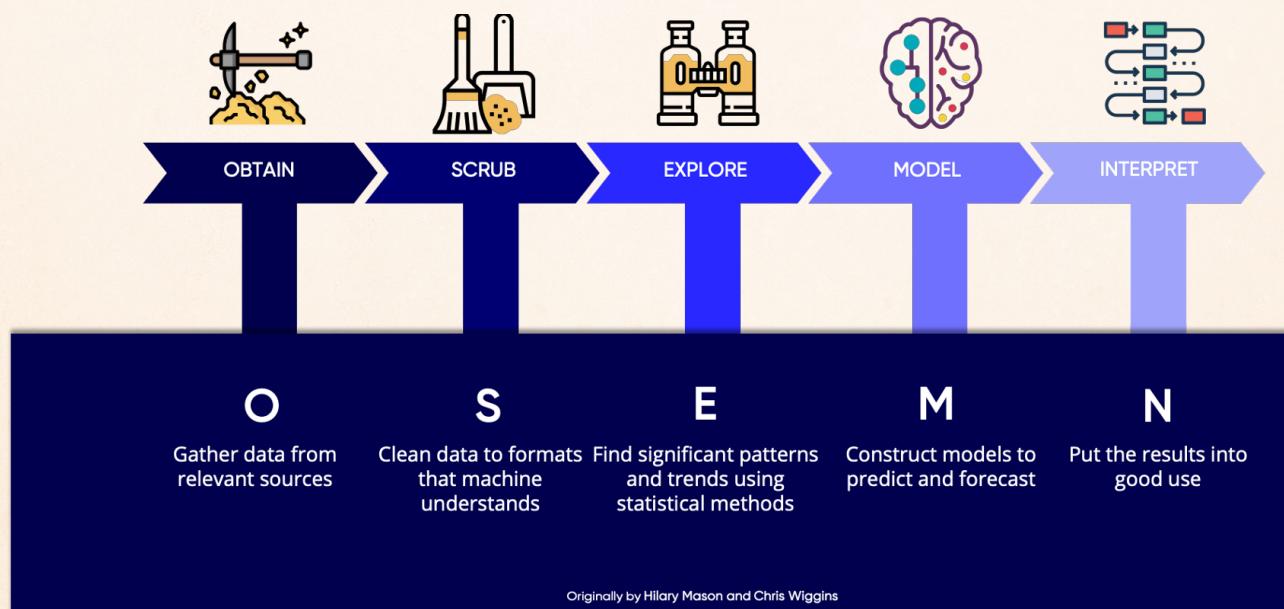
# PLANNING TO MAKE A PROFITABLE REAL ESTATE INVESTMENT?

- ❖ We obtained comprehensive housing data of King County and carefully examined the different elements that should be considered when selling a property in order to maximize your profit.



# OSEMIN

- We aim to predict house prices in the King County area. With machine learning, we utilize the OSEMiN approach to achieve our goals. We selected the most influential elements of housing to successfully sell at a higher price. Here are the steps we take with our data interpretation:



Source *Towards Data Science*

# REQUIRED HOUSEKEEPING

- ❖ We obtain the official, detailed data from King County housing records.
- ❖ We clean up the data which allows us explore relationships among the variables.

Scrub:

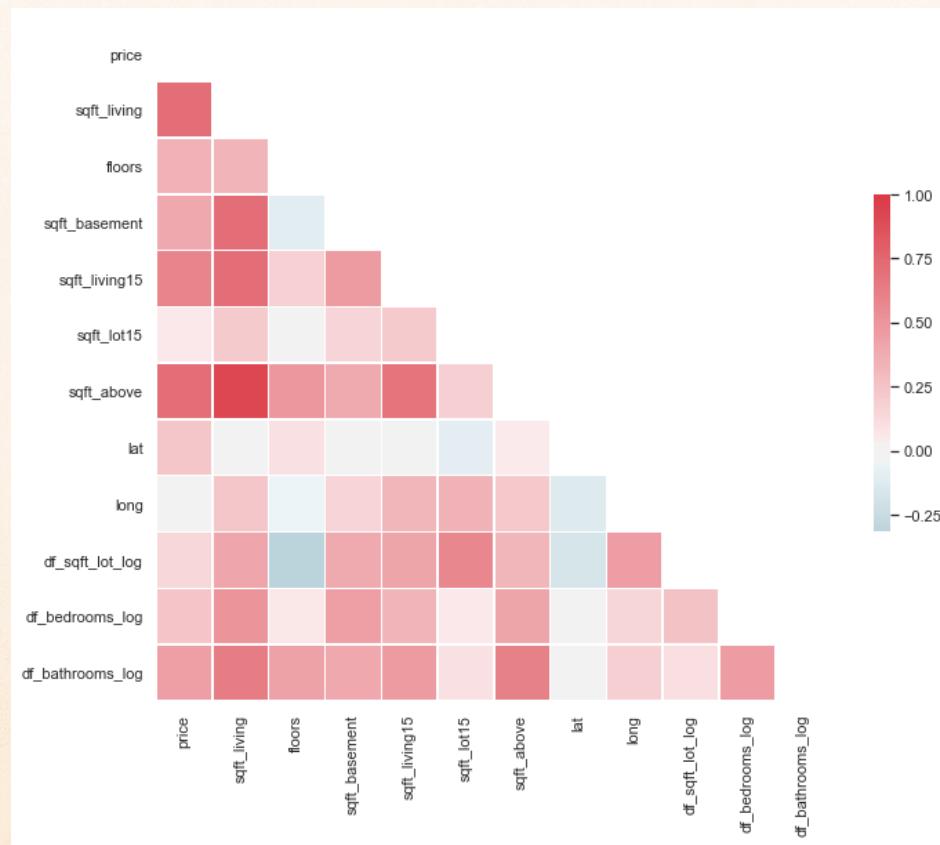


Scrubbing

Figure 1

# EXPLORING RELATIONSHIPS

We examine the correlation between house pricing and other important variables. We noted a few strong correlations (see figure 2 below)



- Sqft.  
Living &  
Price
  - Sqft.  
Basement  
& Price
  - Sqft.  
Above &  
Price

# GETTING READY TO BUILD OUR MODEL

- ❖ Certain variables were not normally distributed and therefore we normalized them using log transformation. (*See example of the square feet variable in figure 4a and 4b*).
- ❖ We also ensure that variables are on the same scale before we start building our model as it helps us interpret the data better.

Example:Figure 4a

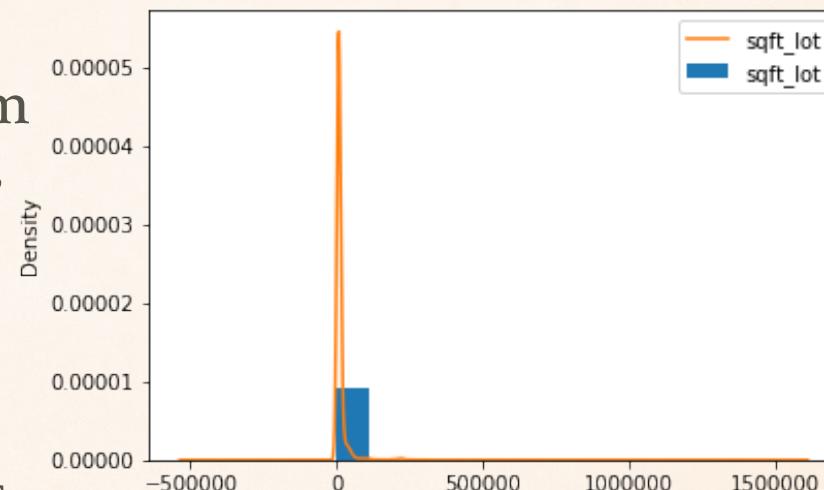


Figure 4a

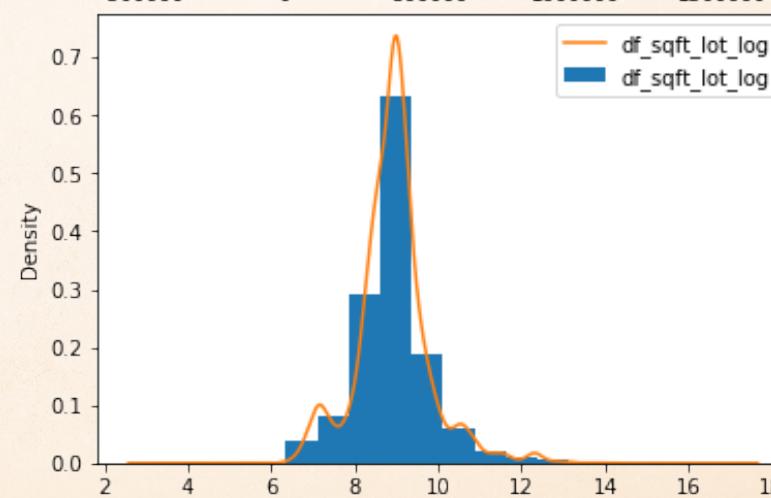


Figure 4b

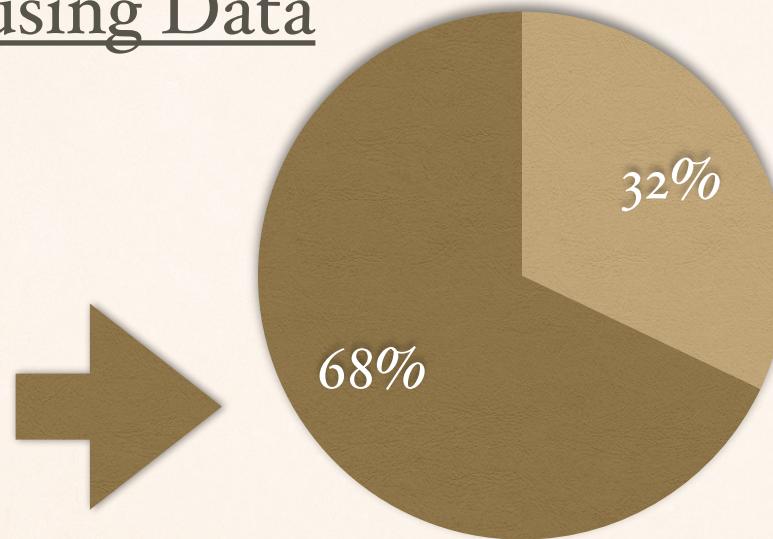
# HOW DO WE BUILD OUR MODEL?

Building a model:

- ❖ We add the variables that were suggested to us using a specific programming code.
- ❖ Our goal is to see how well the different variables (square foot of living, grade of the house, view, or floors for example) explain the house pricing.

# HOW WELL DID WE PREDICT HOUSE PRICING ?

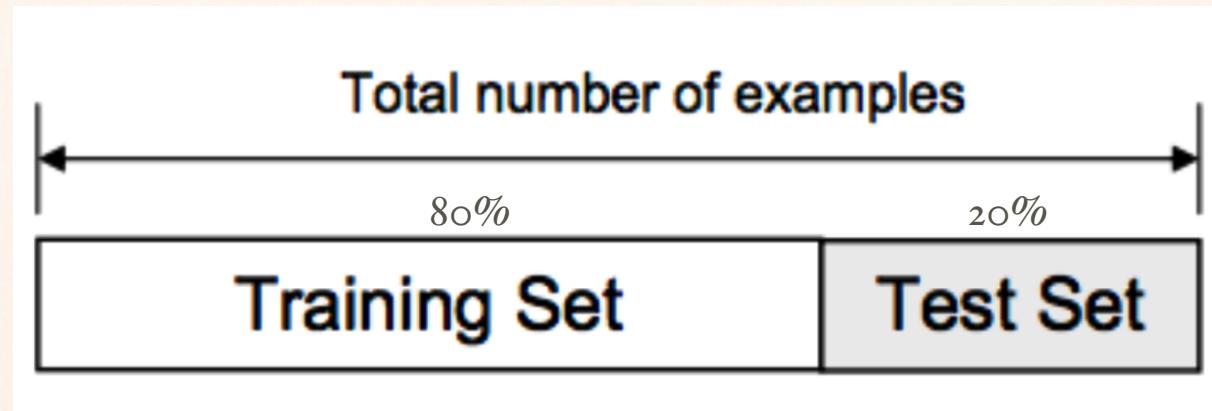
## KC Housing Data



- 68% of the variability of housing price can be explained by the variation in the independent variables (e.g. Sqft. living/lot/basement, grade 9-13, floors) \* pls note that we only included the main variables. Further details are available in the notebook.

# TESTING OUR MODEL

We conduct train-test-split in order to confirm that our data has learned the relationship between the variables. We tested 20% of our data



Our goal in the test is to achieve very close Mean Square Error (“MSE”) among the training set and the test set

**Test MSE:** 0.30

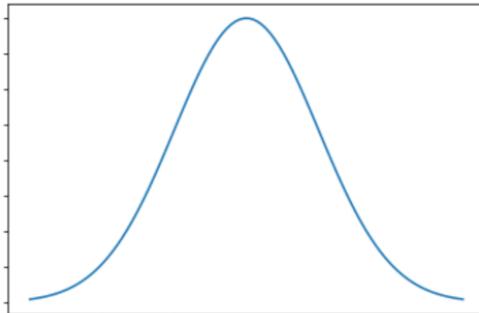
**Train MSE:** 0.33



# CAN WE TRUST OUR RESULTS?

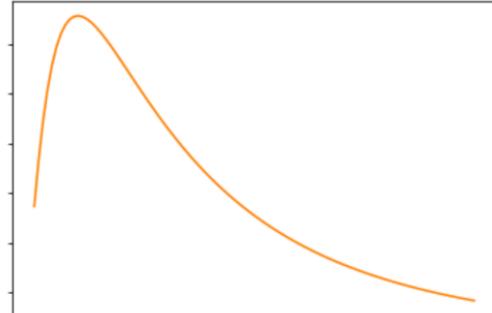
Example: normal distribution  
vs. not-normal distribution

Yes



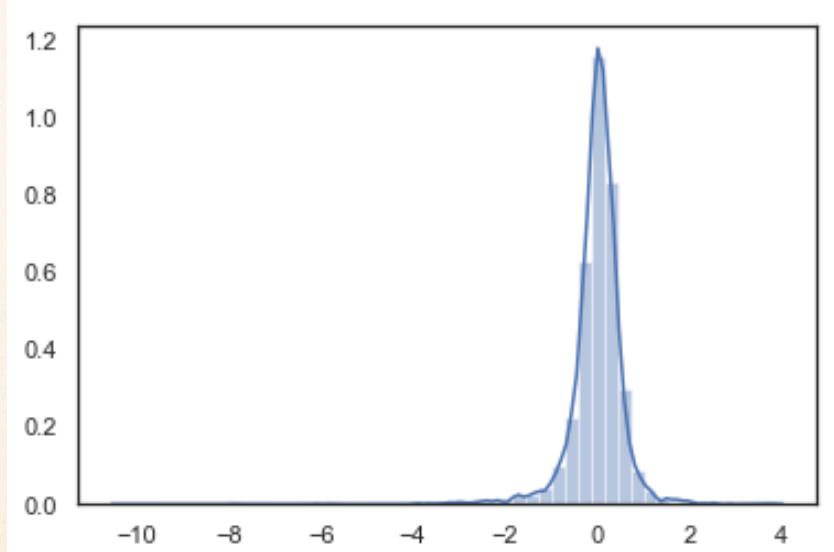
Normal  
distribution

No



Non - normal  
distribution

We confirm that the  
\*residuals of our a data  
normally distributed  
- we can trust our results.



\*(RESIDUAL=Price - PREDICTED Price)

# CONCLUSION

In order to sell the property at a higher price, you should consider the following three main elements

## i. House with **Grade**

**10-13:** High-end finishes.



Renovate before selling if the property needs to be updated.



# CONCLUSION (CONTINUED)

2. **Waterfront View:** the price tends to increase with view of water.



Invest in houses by  
the water  
- you won't regret it



# CONCLUSION (CONTINUED)

3. **Wait:** Both January and February have a negative effect on prices. Therefore, take it easy after christmas.



Wait until March before selling the house. You'll be much more likely to get your money's worth for your investment.



# FUTURE WORK

---

WHEN IS IT BEST TO BUY A HOUSE  
TO MAXIMIZE YOUR INVESTMENT?

We will continue to explore the data on the best timing to purchase a house(the lowest price) so when we sell it we make the most profitable investment.



“Thank You“

– *Natalia Edelson*