



Module 1: Final Project

King County Housing Prices

Joe Wilkinson





OSEMN Methodology

Using the OSEMN methodology gives a strong framework in which we can insure accurate and actionable results.

- 01 | Obtain Data
- 02 | Scrub Data
- 03 | Exploratory Analysis
- 04 | Model Creation
- 05 | Interpret Results





Getting the Data to Work for Us

1

Our first step is to explore and understand the dataset

3

We make sure that all of the predictors we are using describe price and not each other

2

We need to deal with rows that are missing data

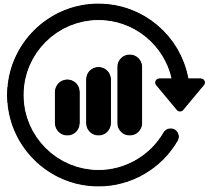
4

We want to be certain that our data fulfills all the assumptions that are necessary to create a model



Creating Our Model

0
1



Remove Inconsequential Predictors

We figure out which predictors don't actually influence our model and remove them from the equation to keep everything as simple as possible.

Test Our Model

We want to insure that our model is actually predicting results, so we run several tests to protect us from creating a model that works on the data we currently have.

0
2



0
3



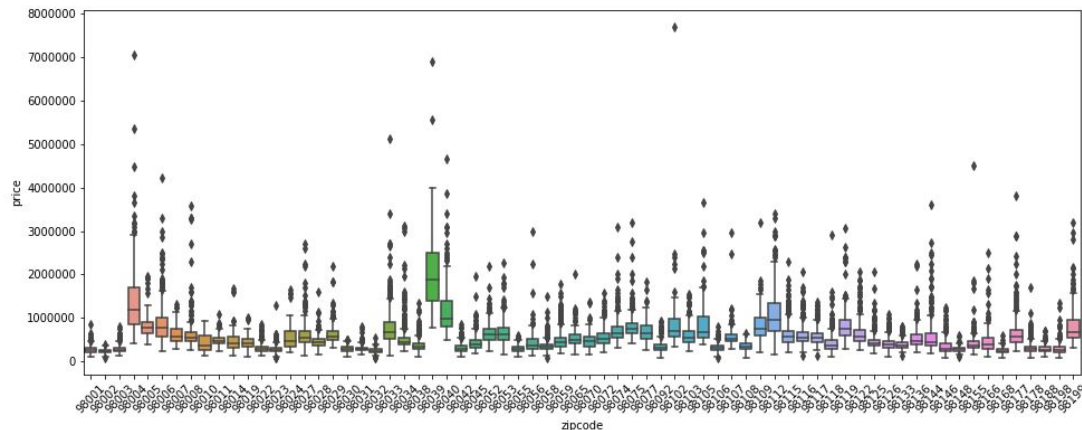
Final Results

98.9% of the variations in our dataset are explained by our final Model



Results

Our model clearly shows that location is king. Zip codes and Latitude were the biggest influencer of price. You will be able to increase profits by focusing on the zip codes that have the closest association with high prices. The image below shows the boxplots for each zip code and there are some that clearly outperform the others.





Predictors of Interest

Waterfront

2.65%

Increase in price if the property is waterfront..

Square Feet Above

2%

Increase in price for every 10% increase in square feet above.

Viewed

318%

Increase in price if the house has been viewed.