

# Predicting House Prices

## In King County, Washington



Kevin Velasco - Data Scientist  
Alex Shropshire - Data Scientist

*Overview > Exploration > Methodology > Insights*

# Overview



## Goals

To accurately predict the market price of homes given historical sales data

To understand the effect of unique property attributes on price

## Dataset

2014 -2015 King County, WA House Sales data

Over 20,000 home sales

19 unique variables describing location, size, condition, time, & special features

## Tools

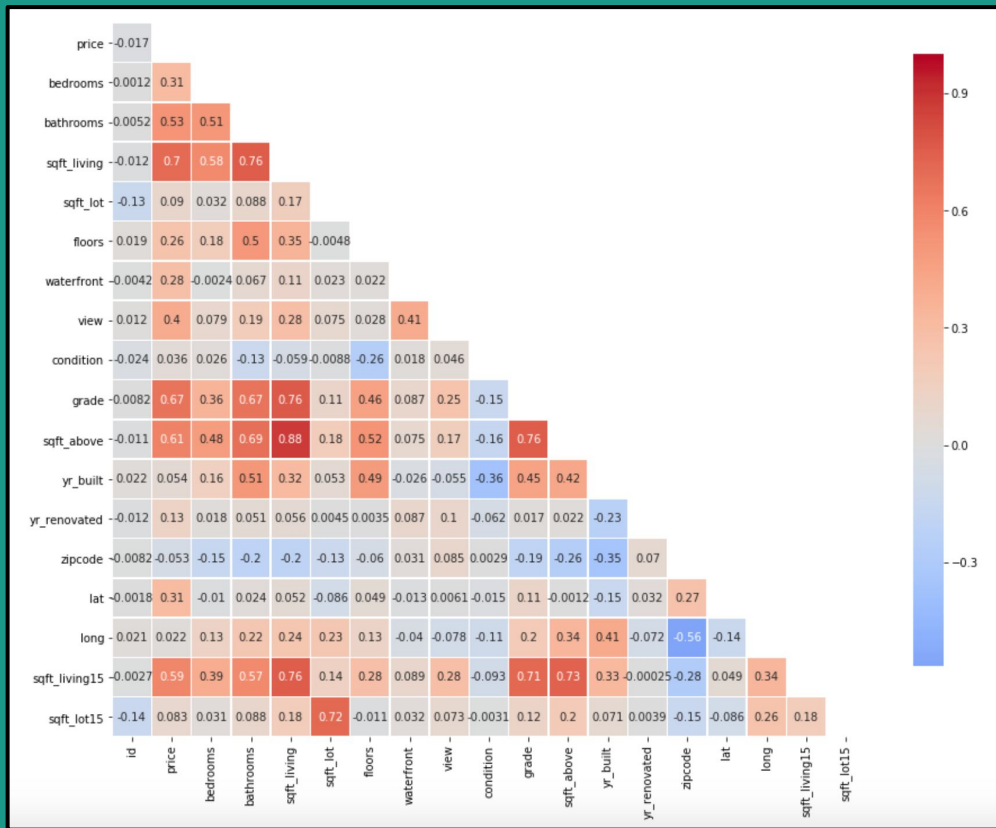
Python: Numpy, Pandas, Statsmodels, Sci-kit learn, Matplotlib, Seaborn

Jupyter Lab/Notebook

Git & GitHub

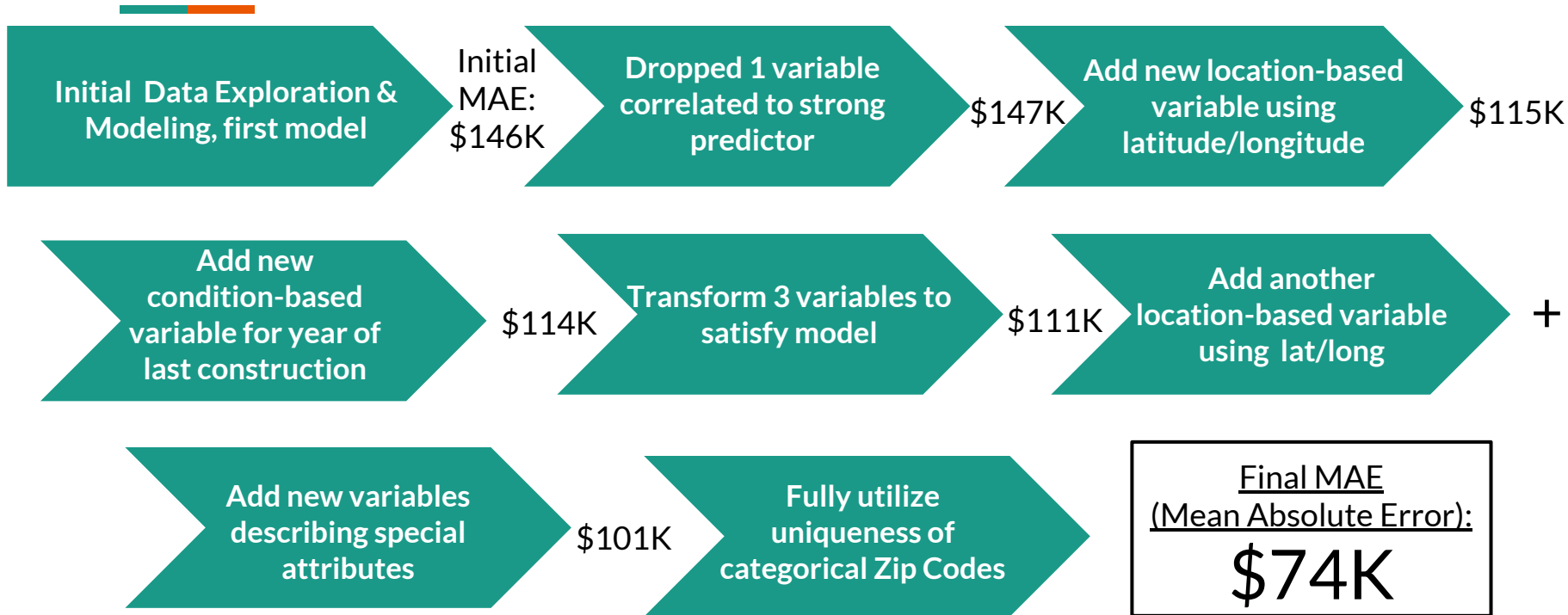
# Data Exploration: making sense of it all

- Understanding context
- Cleaning and filling in the dirty or messy data
- Engineering new features & incorporating new data
- Transforming data to meet model assumptions



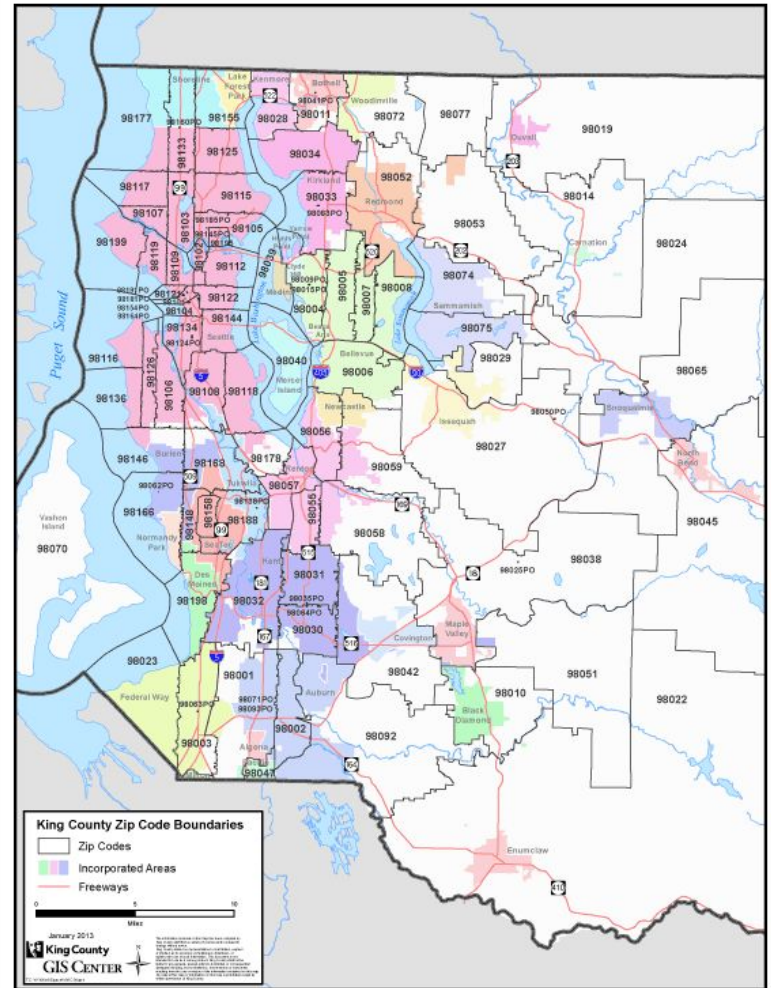
*Understanding initial data relationships via correlation.  
Bolder red & blue indicate stronger relationships*

# Methodology & Model Improvement Effect on MAE



# Key Predictors

- Zip Codes
- Distance from economic hubs





**Thank You**