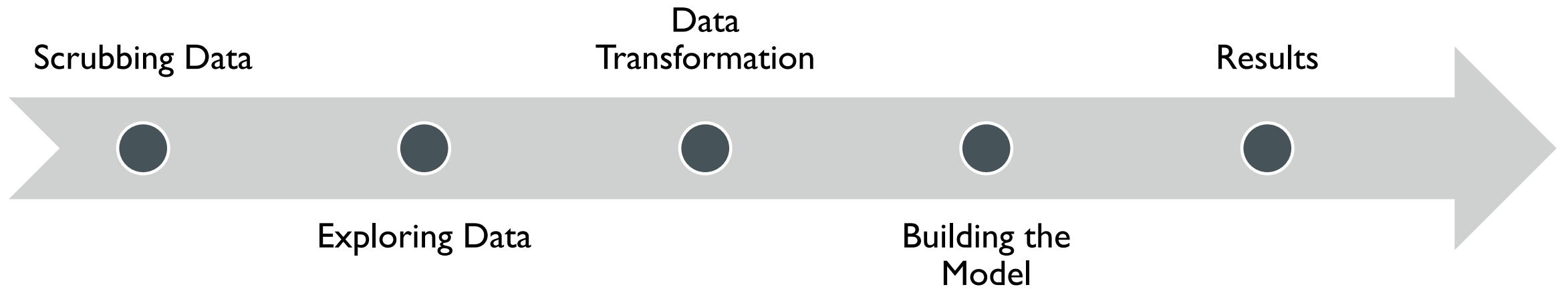

PROJECT I – PRESENTATION KC HOUSE PRICES

RAYMOND WILLEY



METHODOLOGY



SCRUBBING DATA

Missing Data



Waterfront View

- Missing values replaced with most common result: No waterfront view



Times Viewed

- Very few items (less than 1%): omitted instances with missing data



Year Renovated

- Used change in square footage and year renovated to determine if ever renovated

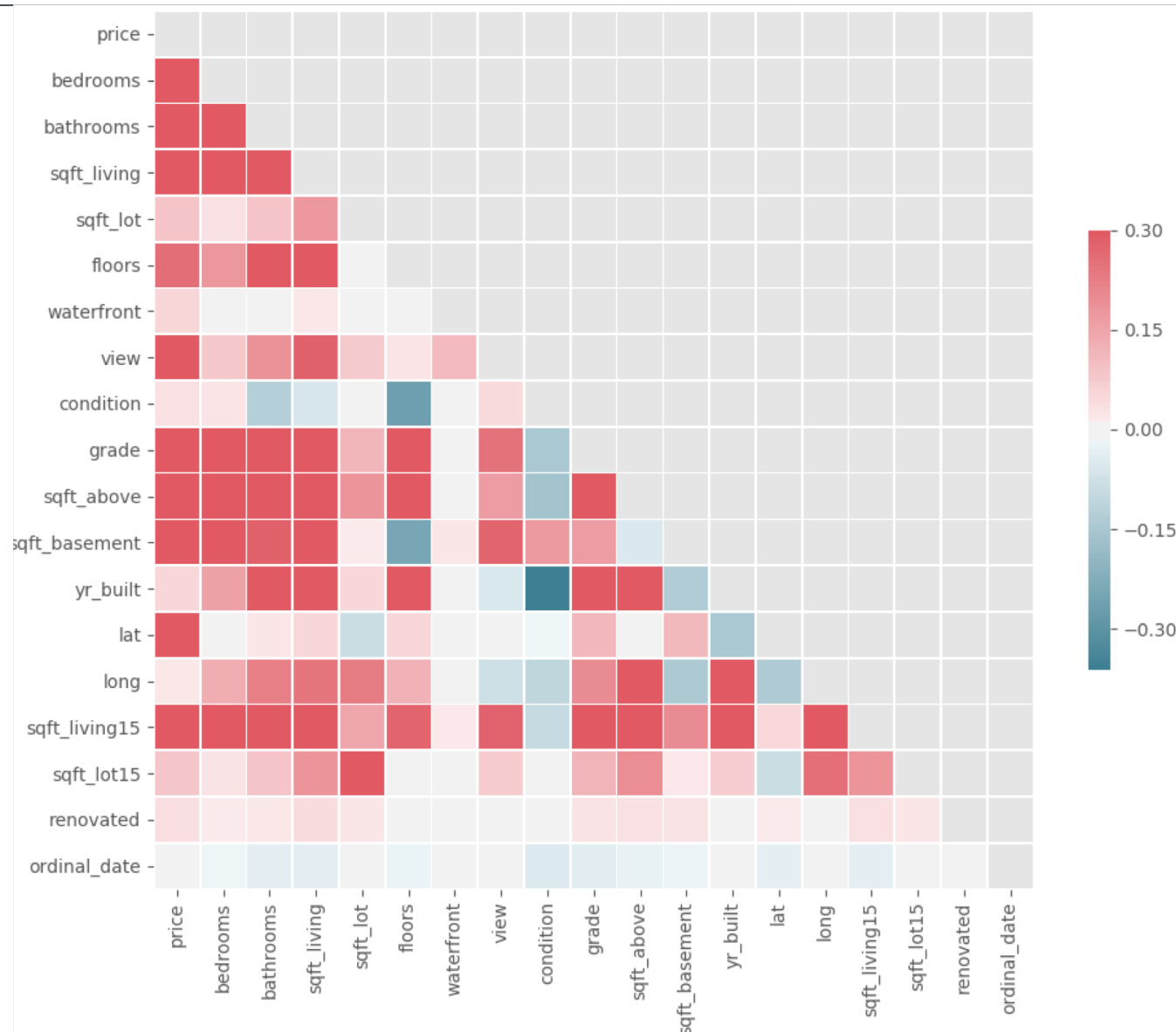


Basement Square Footage

- Derived from difference between total square footage of living area and living area above basement

- Items in bright red or bright blue indicate strong correlation

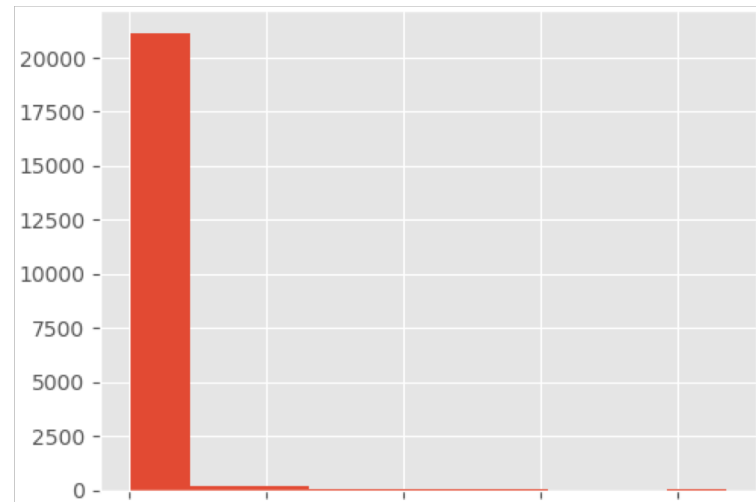
- Items dropped:
 - Number of bathrooms, bedrooms, specific square footage calculations, etc. strongly correlated with total, current square footage
 - Year built, longitude, and grade not sufficient predictors (grade too ambiguous)



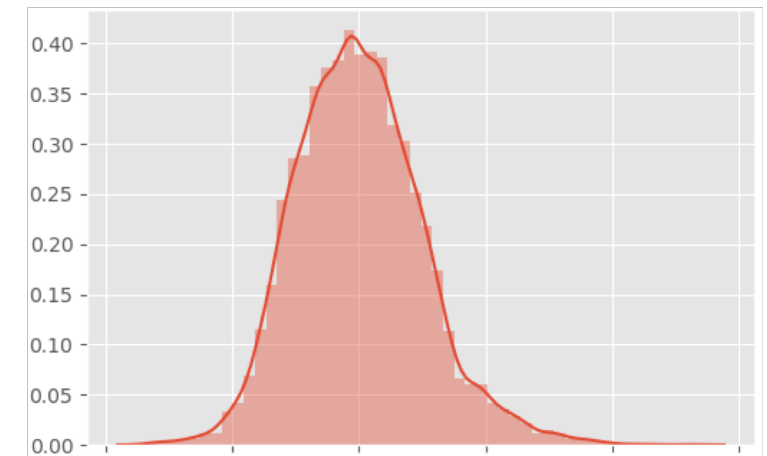
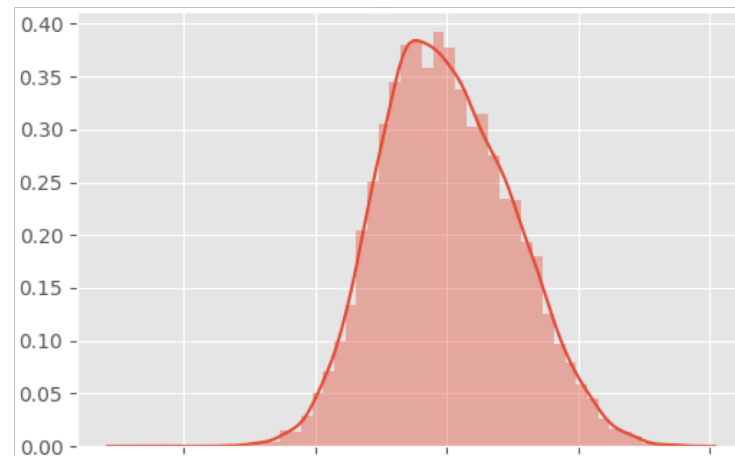
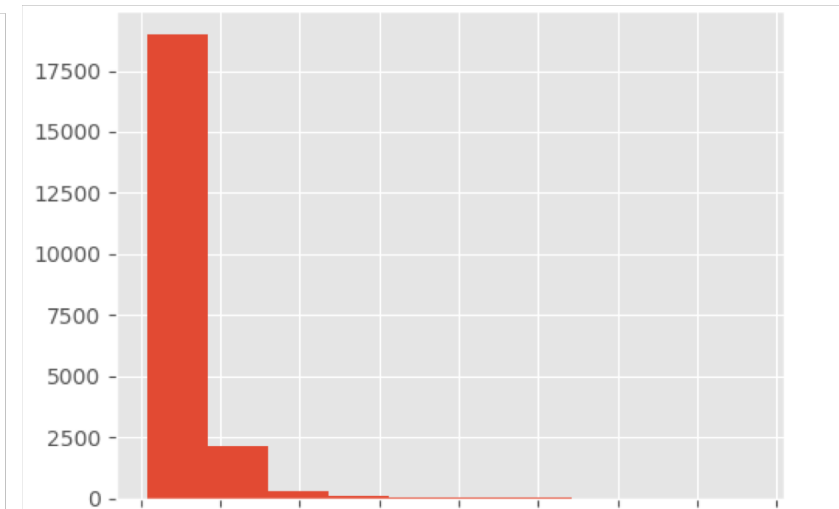
TRANSFORMING THE DATA

- Distribution of data suggests changes increase on a percentage basis rather than absolute basis
- Example: percentage increase in square footage results in a percentage increase in price

Square Footage (2015)



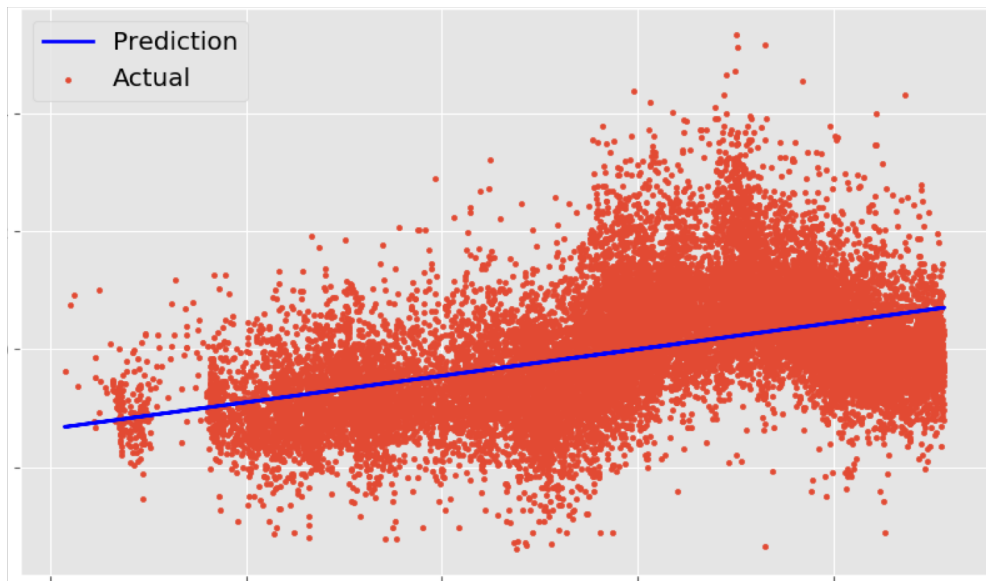
Price



Upper figures show absolute distribution of Price & Square Footage; lower figures show percentage scaling

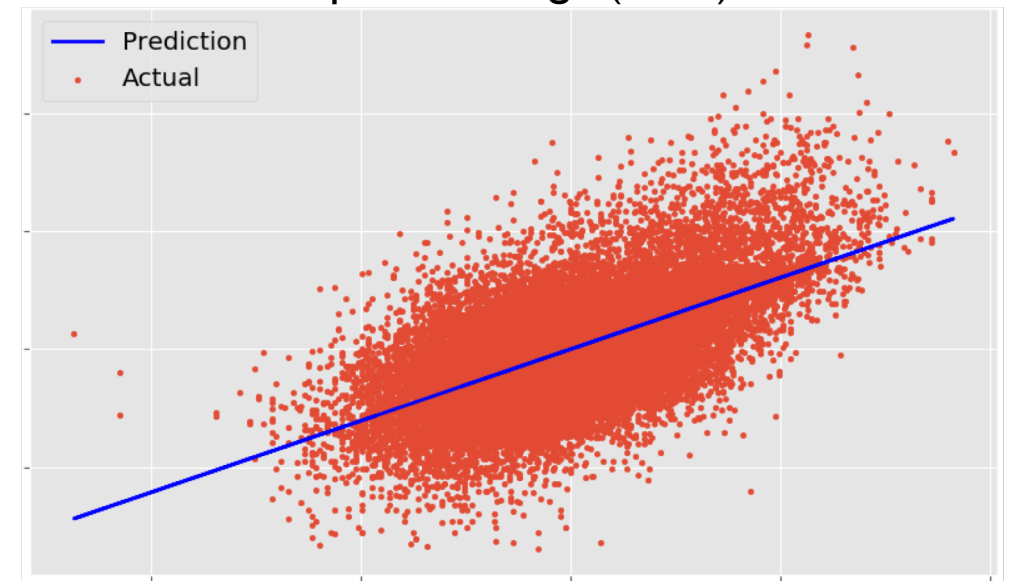
STRONGEST PREDICTORS OF PRICE

Latitude



Prices higher in the north

Square Footage (2015)



Larger size leads to higher price

RESULTS

PRIMARY DRIVERS OF PRICE

- For every percentage point increase in the following, price increases by the given percentage:
 - Square Footage (Living Area): 0.52%
 - Latitude: 0.09%
 - Square Footage (Lot): 0.11%
 - Zip Code
 - Can lower price by as much as 1.71%
 - Can raise price by as much as 0.98%
- Approximately 74% (+/- 3%) of the variance in price can be explained by these 4 variables.

RESULTS

RECOMMENDATIONS

- Maximize square footage
- Build on large plots (or combine lots where possible)
- Build in the northern part of the county



QUESTIONS?

