



# Predicting King County House Prices Using Linear Regression

---

DOUG STEEN

10/30/2019

# Project Approach

---

King County housing dataset includes home sale prices from May 2014 – May 2015

Includes other features:

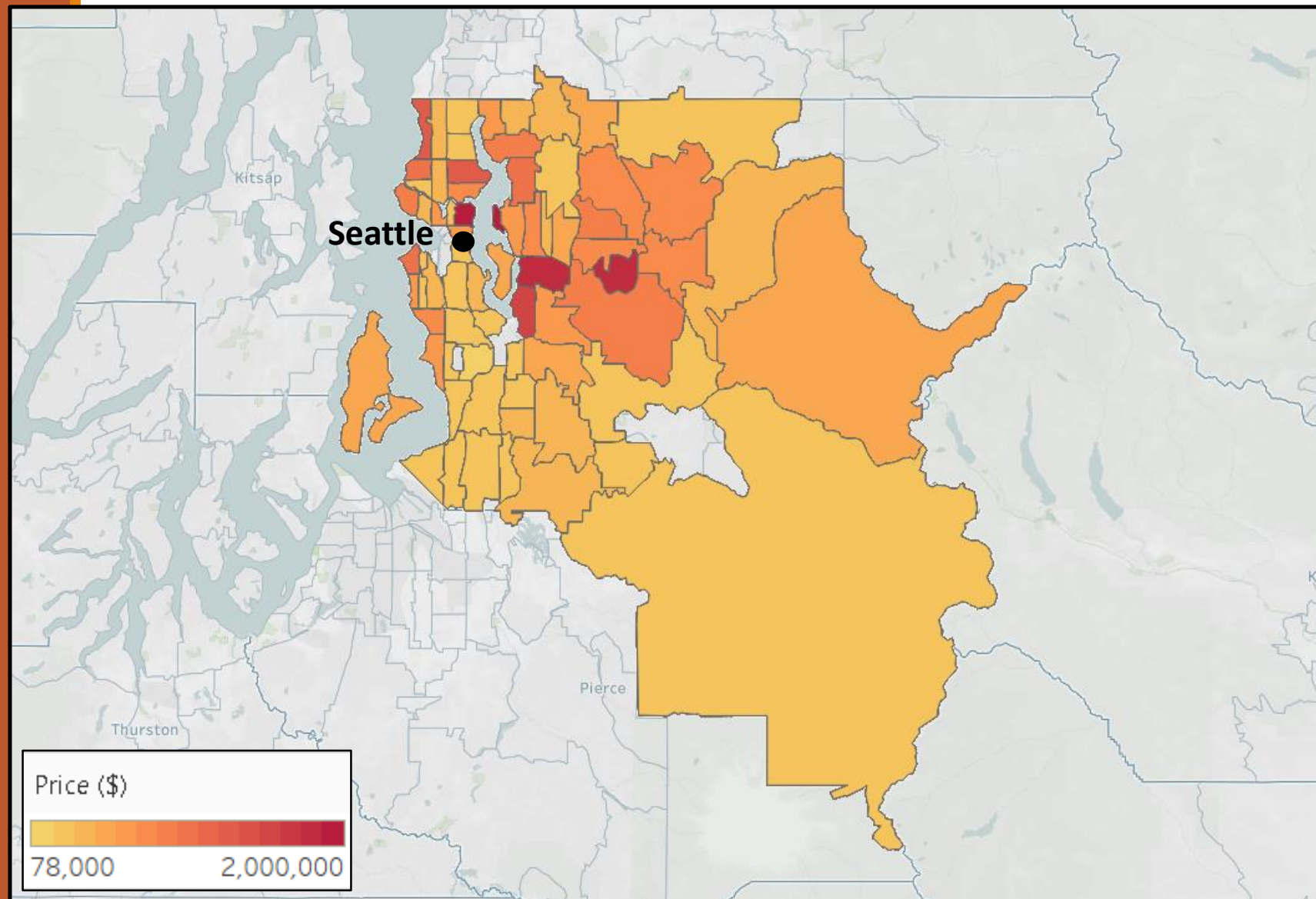
- Location - zip code, latitude and longitude
- Size - ft<sup>2</sup> ... above ground, basement, living area
- Size of neighborhood homes
- Number of floors, bedrooms, bathrooms
- Condition, grade
- Year built, year renovated
- Setting - waterfront, view

Objective: Produce linear regression model to predict sale price using most critical features

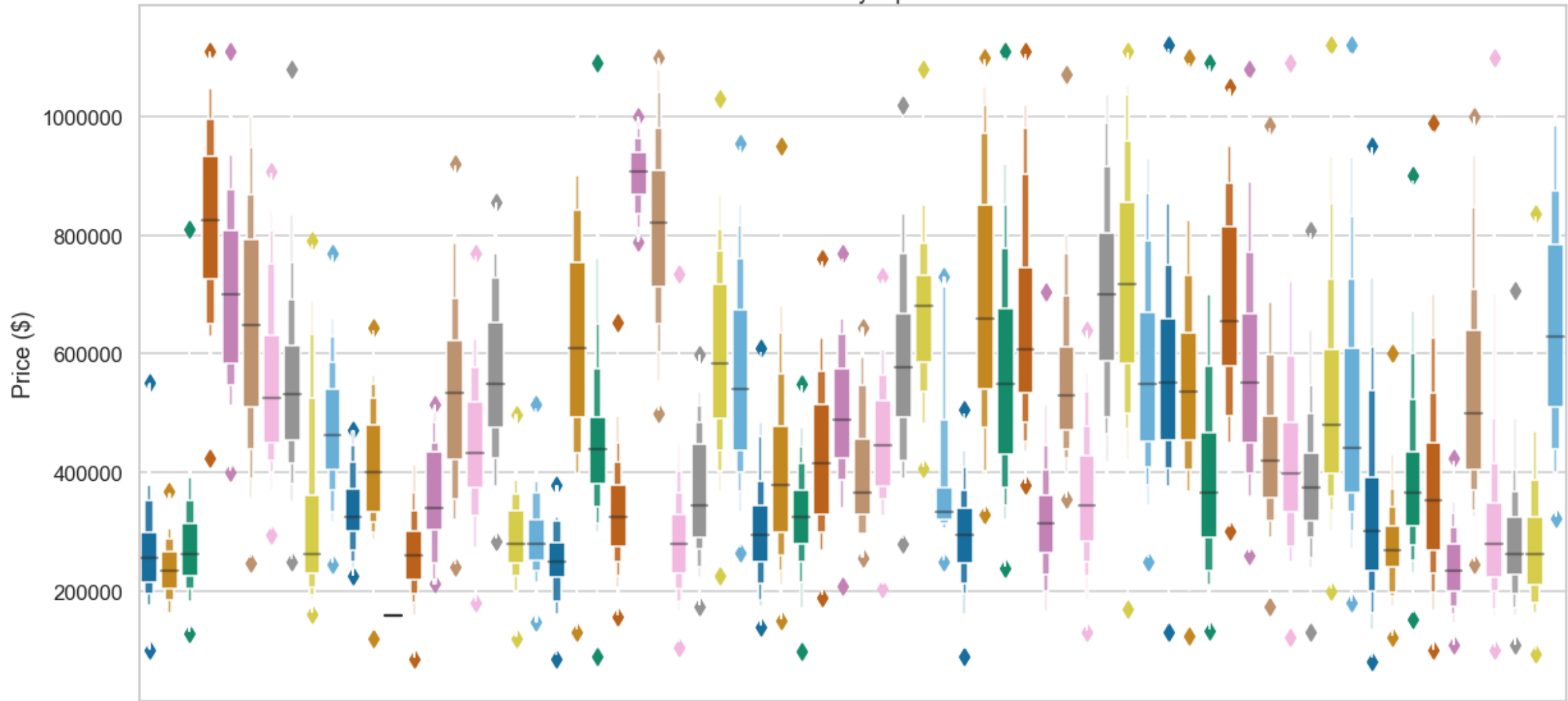
# Results

- Five most important features for predicting sale price:
  - Zip code
  - Construction/Design Grade
  - Neighborhood Home Living Area
  - Distance from Seattle
  - Home Size (excluding basement)
- Model explains 77% of the overall variability in King County house prices

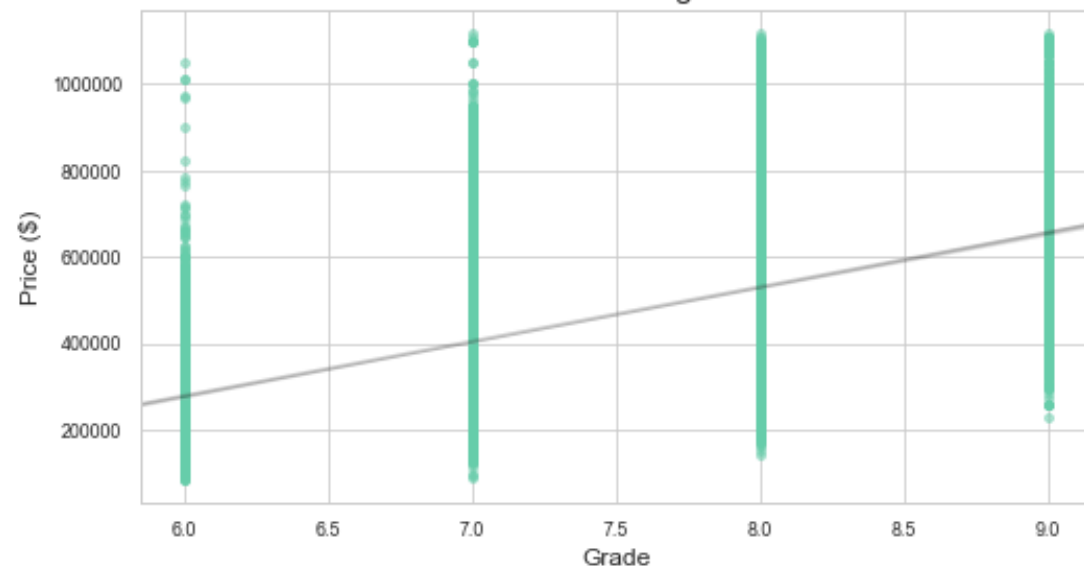
## Median Sale Price by Zip Code in King County



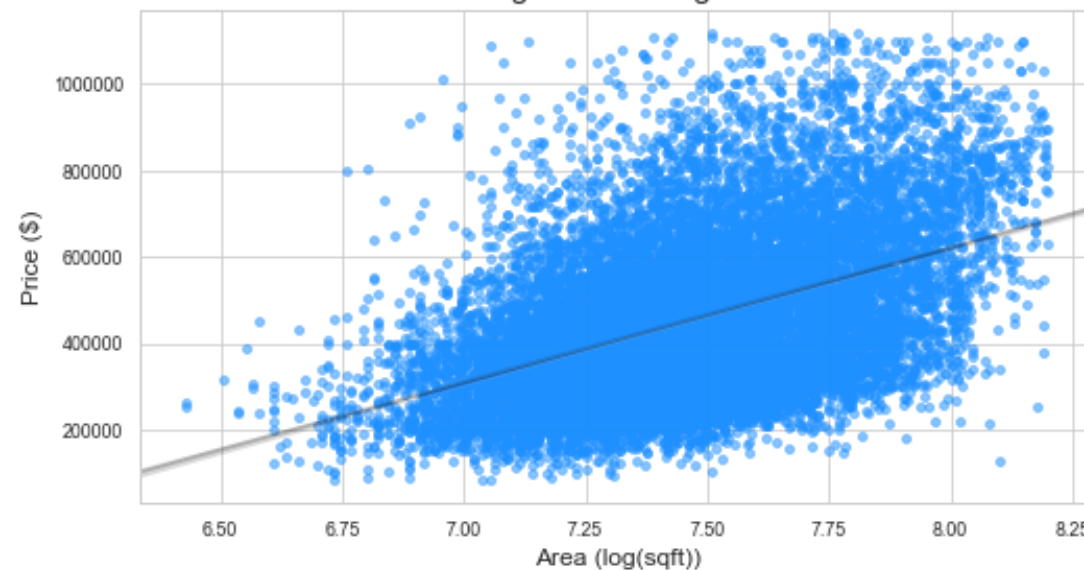
A box plot showing the distribution of the number of nodes in the largest component for different values of  $\alpha$ . The x-axis represents  $\alpha$  with values 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. The y-axis represents the number of nodes, ranging from 0 to 100. The plot shows that as  $\alpha$  increases, the number of nodes in the largest component generally increases, with a significant jump around  $\alpha = 0.5$ . The distributions are highly variable, with many outliers shown as individual points.



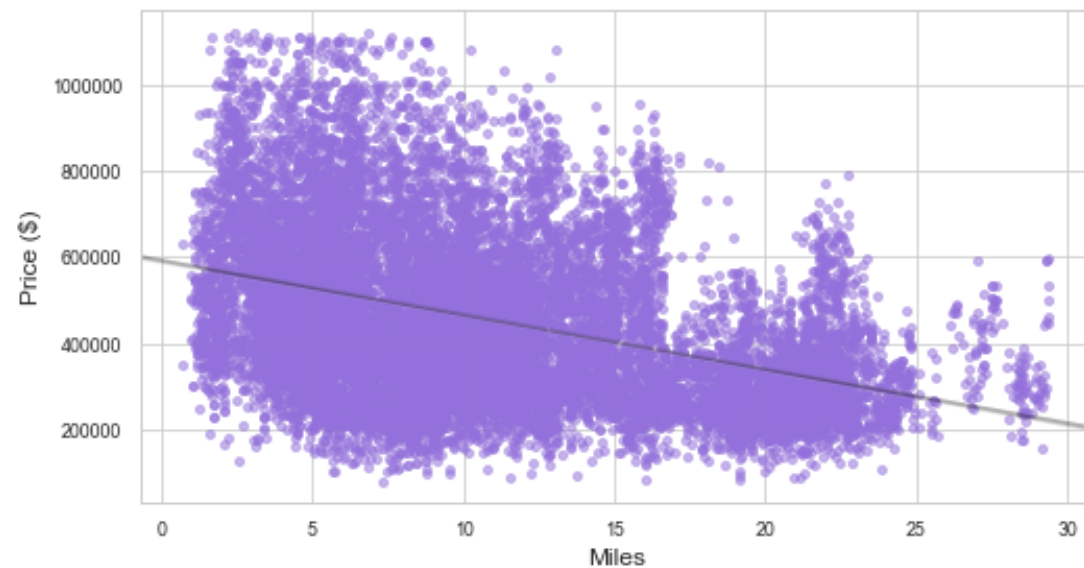
Construction/Design Grade



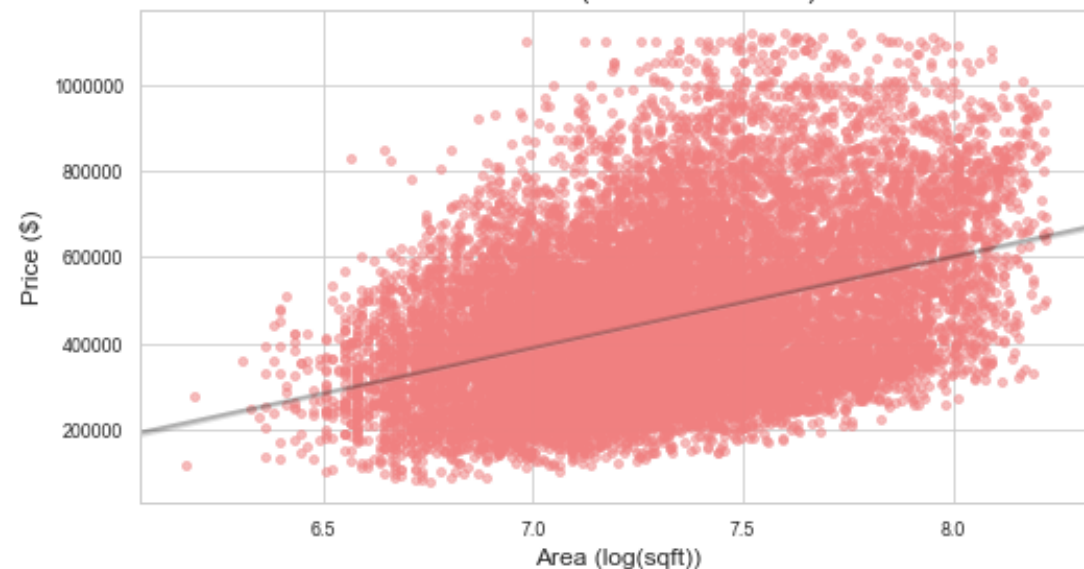
Living Area of Neighbors



Distance from Downtown Seattle



Home Size (Excl. Basement)



# How to maximize profits when selling your home?

---

**Consider an addition to increase the square footage of your home**

- Encourage your neighbors to do the same!



**Stick to quality construction and design for home building and renovations**



**For current homeowners, location factors are largely out of your control**

- If you live in a desirable zip code or close to the city, know your worth!



# Future Work

---

- Explore more regional features driving house prices:
  - Are there unique temporal trends in certain areas?
- Explore different methods for addressing outliers:
  - Predictors may change with different treatment of very cheap & very expensive properties

Thank you!