Analysis of the King County Housing Market

BY KYLE JOHNSON

Summary

The model created can explain 86% of a home's value based on its measurable characteristics
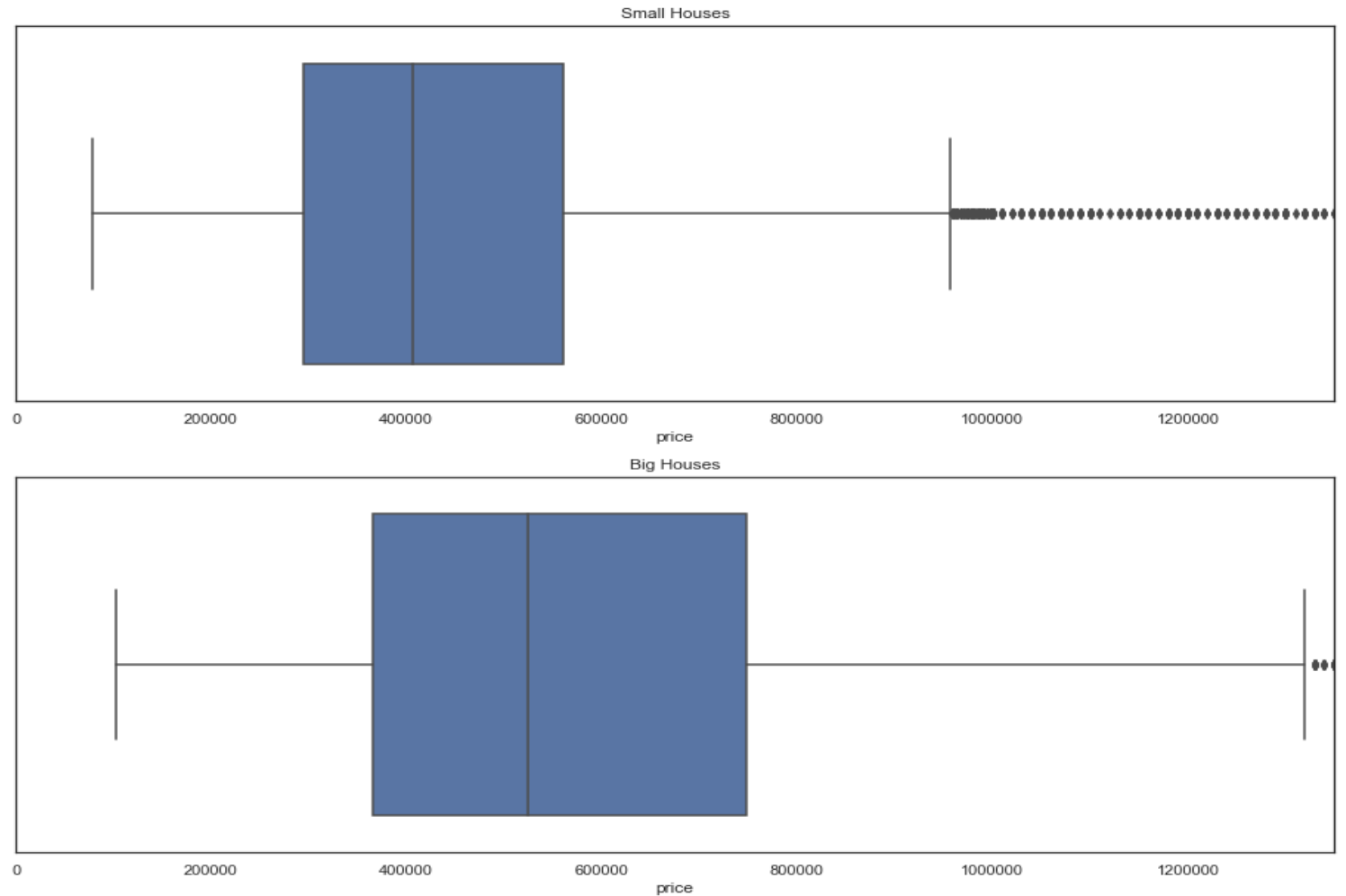
Methodology

Model: Ordinary Least Squares

OSEMN model

Obtain	Scrub	Explore	Model	iNterpret
• From other location • Query from database or API • Extract from another file • Generate data (e.g. Sensors)	• Filtering lines • Extracting columns or words • Rreplacing values • Handling missing values • Converting formats	• Understanding data • Deriving statistics • Creating visualizations	• Clustering • Classification • Regression • Dimensionality reduction	• Drawing conclusion from data • Evaluating meaning of results • Communicating result

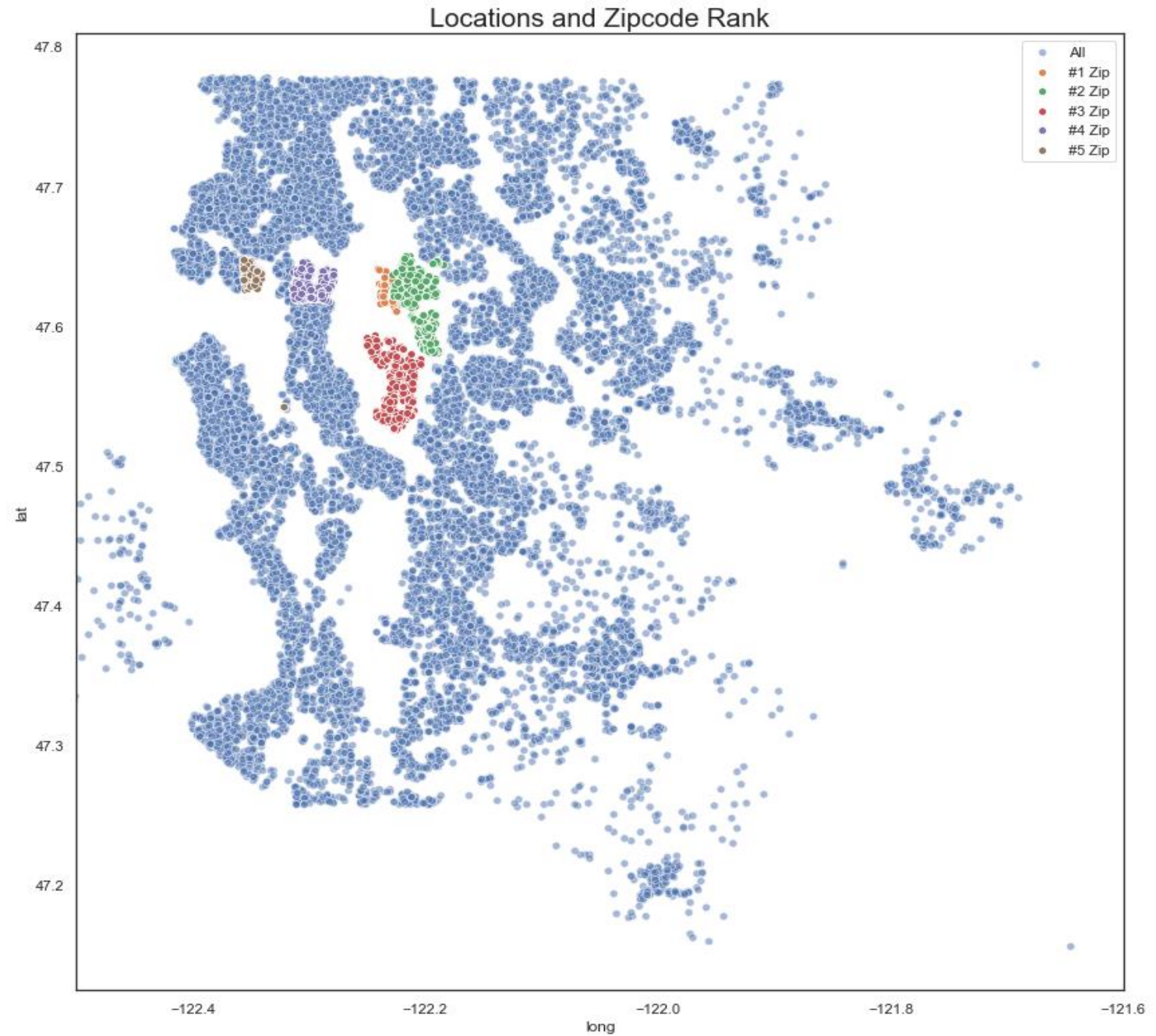Source: A Taxonomy of Data Science

Findings

There is a premium of $91,000 for being double the size of the average neighbor.
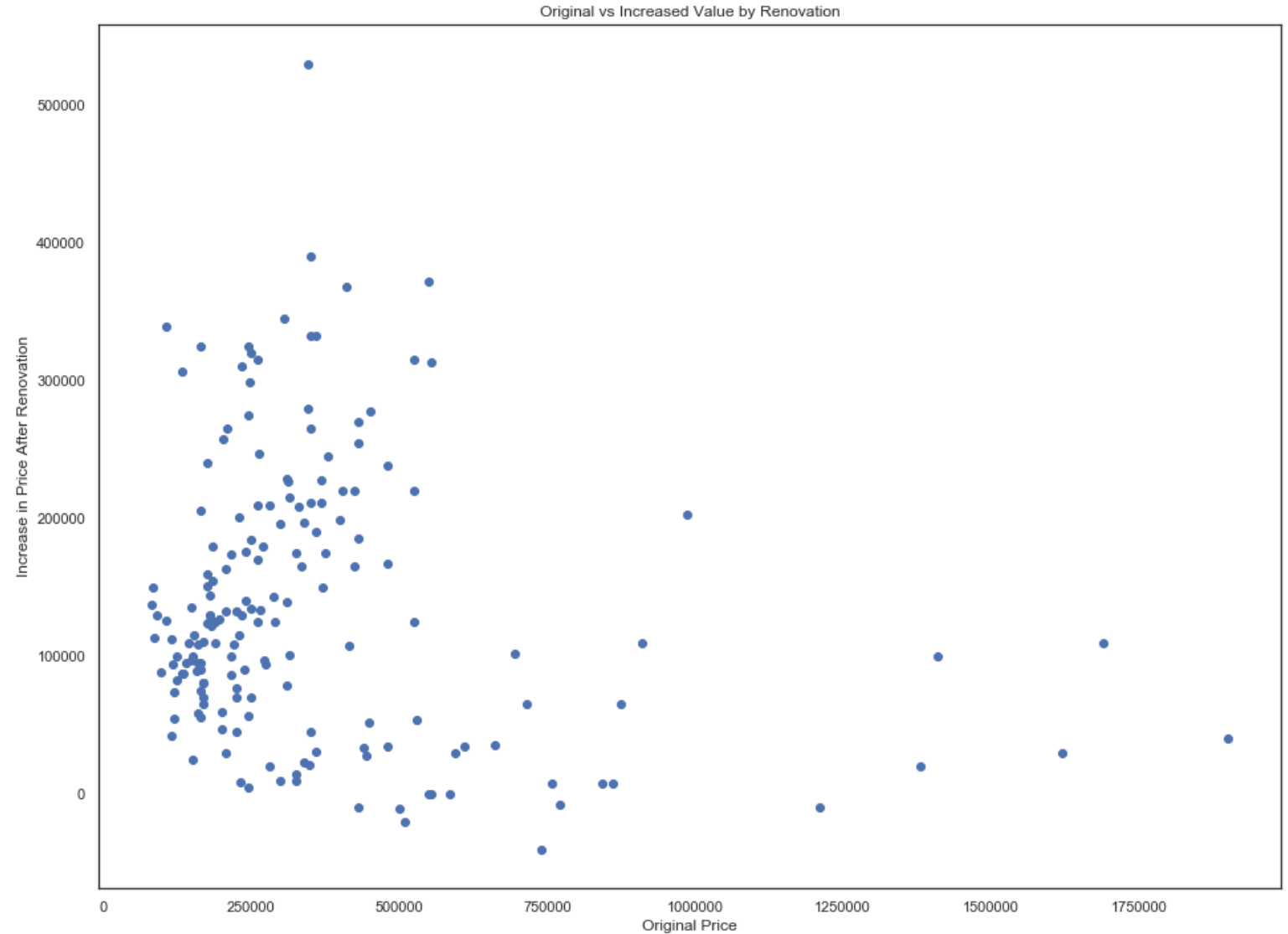
Findings

The premium zip codes are all close to downtown Seattle and near the water

Findings

Renovation increases a home's value by an average of 28%

Original vs Increased Value by Renovation

Findings

Being on the water accounts for a $116,000 increase in home value

Next Steps

Further Exploration should involve obtaining new variables such as school quality and crime rates or deploy more sophisticated modeling techniques.