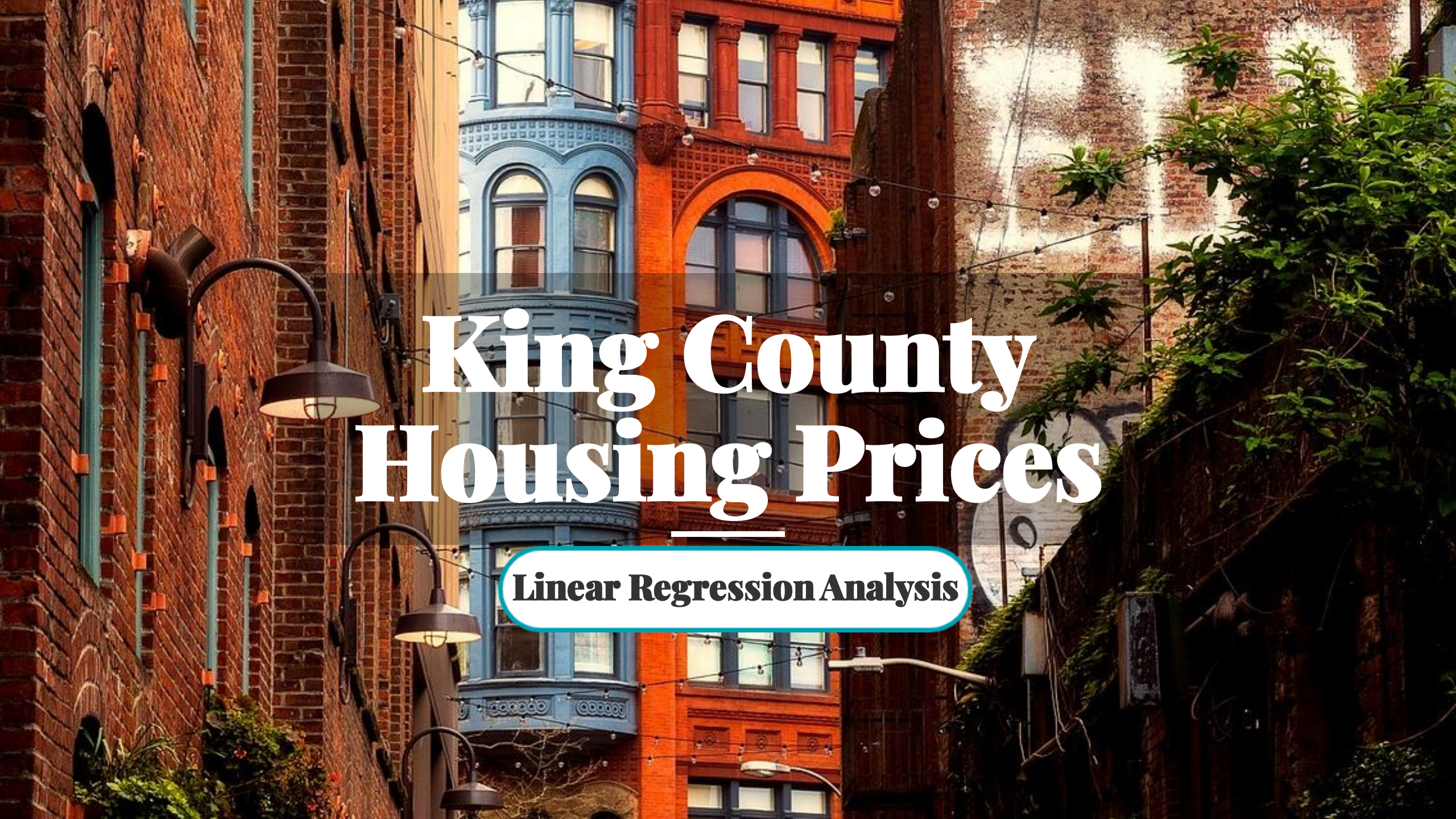
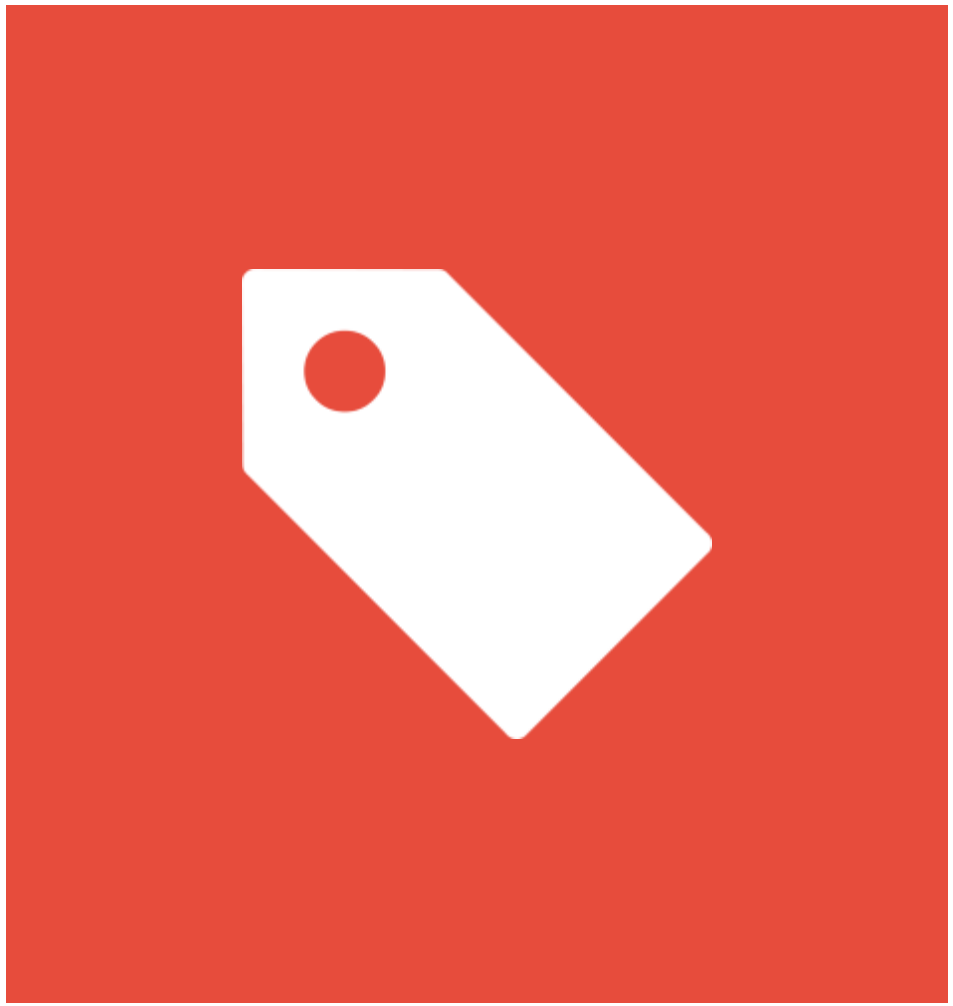


King County Housing Prices

Linear Regression Analysis

Our Goal



1. Use data analysis to gain insights about the housing market
2. Determine the features of a house that best predict the price
3. Build a model to predict the price using the features

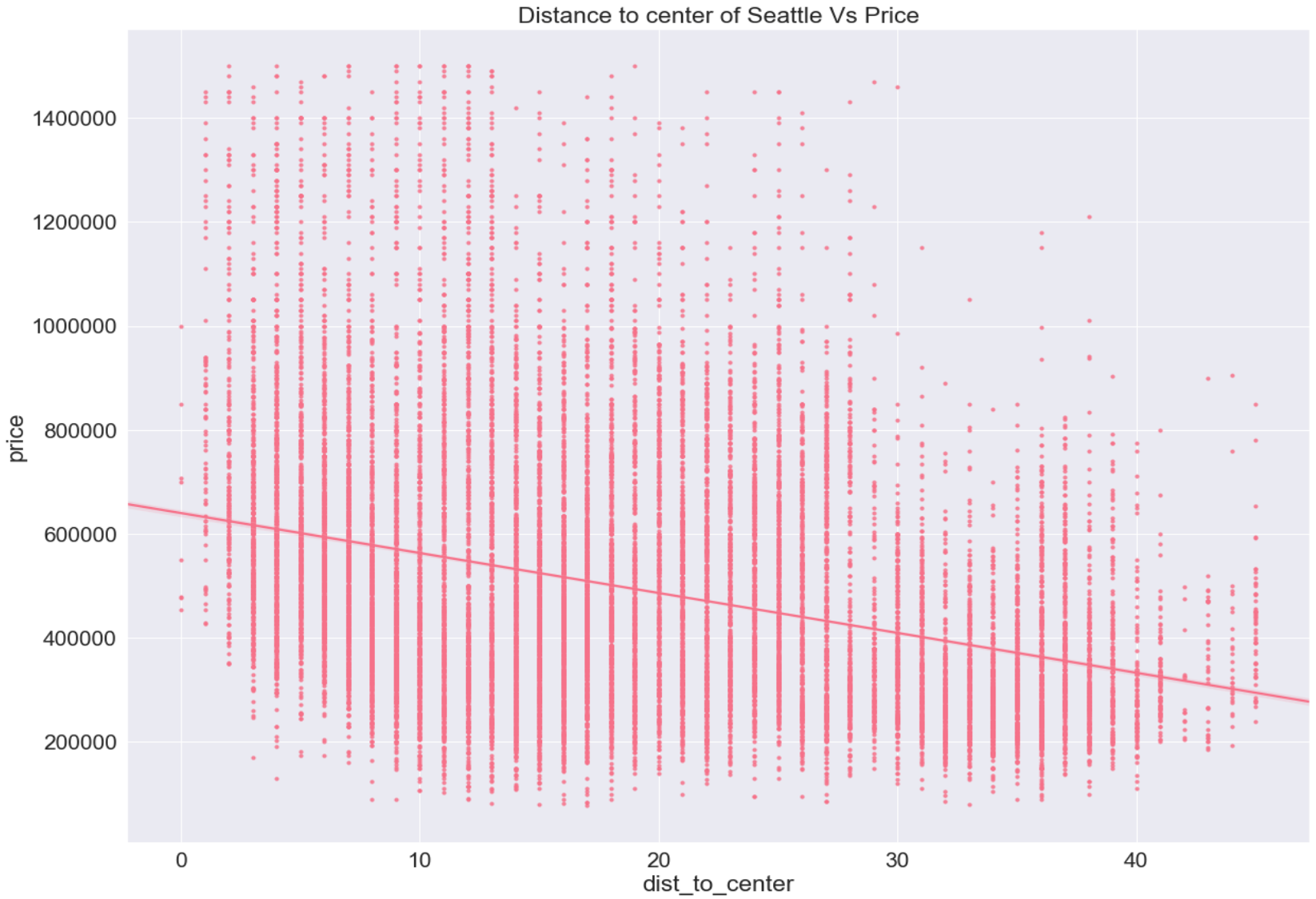
The Data

- 21597 houses in the original dataset
- 20084 after cleaning data and removing outliers
- Distance to Center added
- Categories created

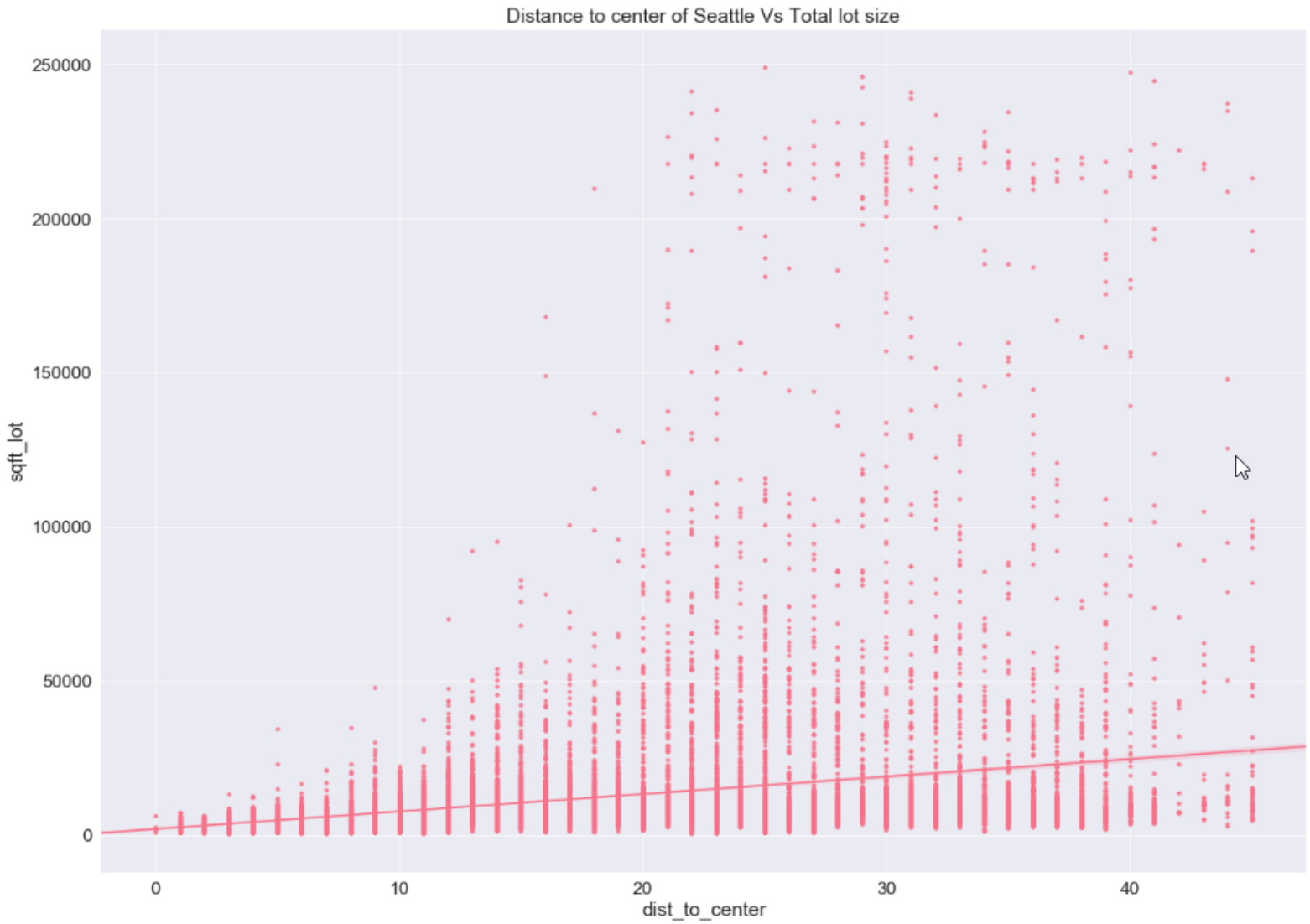
Explorative Data Analysis

What stands out ?

How does distance affect price?

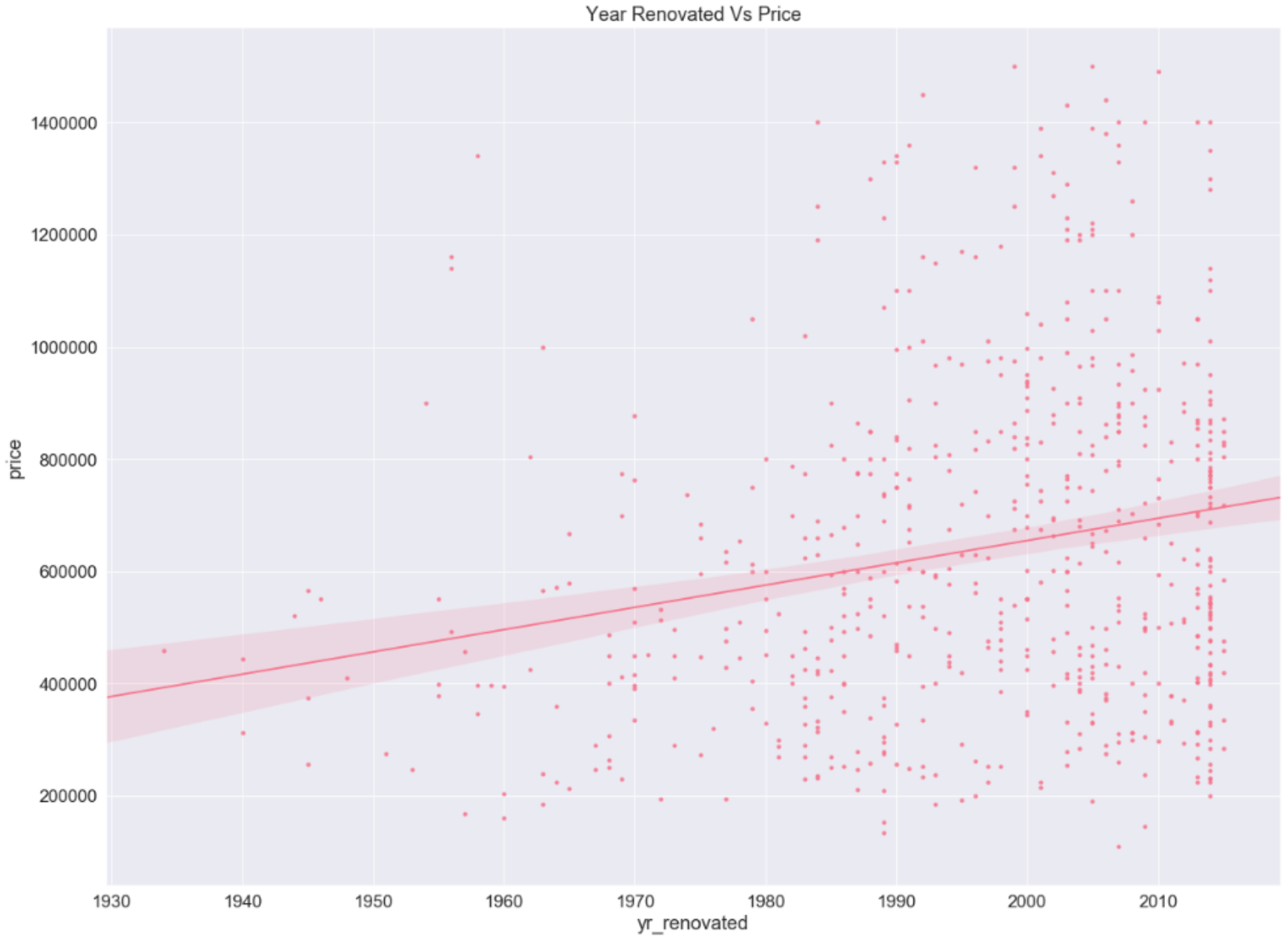


Distance to Center vs Lot size



Year renovated vs Price

What's the relationship between renovated date and price?



Best predictors for the Model

Error rate of 6.6% on properties up to 1.5M - $98k variance.

- Year renovated
- Distance to center of Seattle
- Lot size
- Living Space
- Zipcode

Recommendations

1. Renovate your home, the more recent the higher the price of the property.

2. For more outside space move further from the center of Seattle

Future Work

1. Get more data relating to places of interest, schools, crime rates, etc and improve model predictions

2. Further analysis on outliers, the homes with more than 6 rooms, price higher than 1.5m and are unusual


THANK
YOU