King County Housing Data Analysis

Introduction

This project focuses on providing insight for a prospective homebuyer looking to purchase a home in King County, Washington.

The buyer intends to live in the home for 3-5 years and then sell, ideally for a profit.

The buyer also has a budget for renovation.

Business Problem

This analysis aims to answer the following questions:

- What is the impact of living in a particular section of King County versus other sections.
- What is the correlation between a house's square footage of living space and it's selling price?
- Where is

Libraries

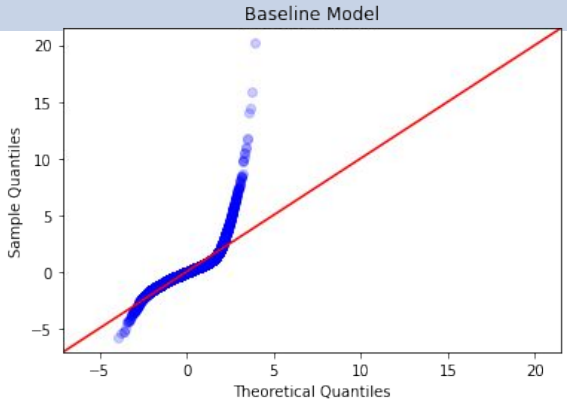
Data

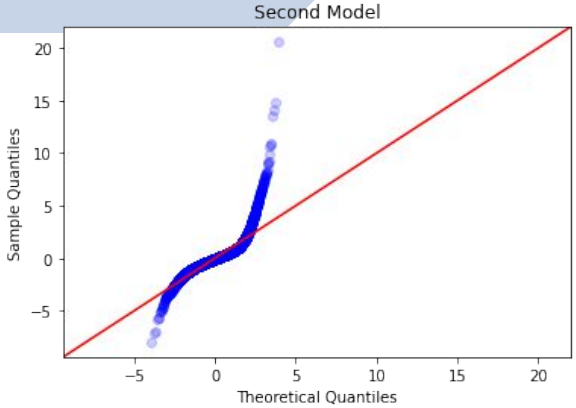
- Analysis focuses on the King County Housing dataset provide by Kaggle.com
- Data included 21,597 records of homes sold in 2014 and 2015
- Entries contained information such as selling price, square feet of living space, number of bedrooms, etc.

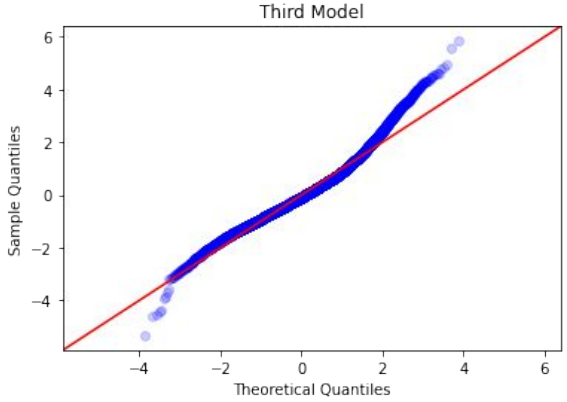
Methods

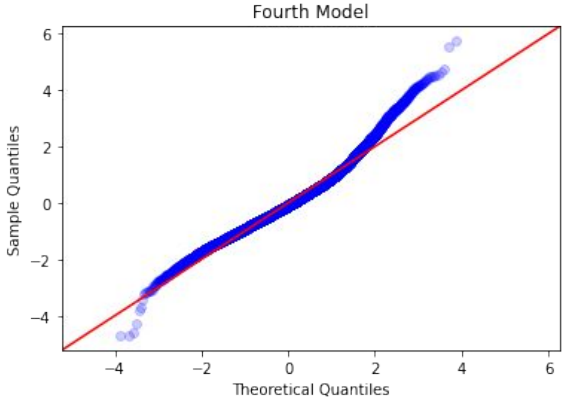
1. Cleaned dataset
2. Create dummies for all categorical variables
3. Remove continuous variables' outliers
4. Log-transform and scale continuous variables

Normality

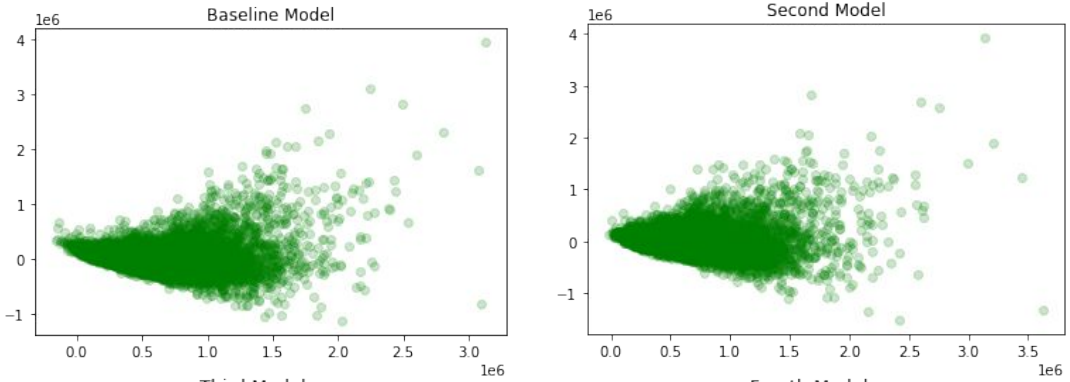

Baseline Model
Sample Quantiles
Theoretical Quantiles

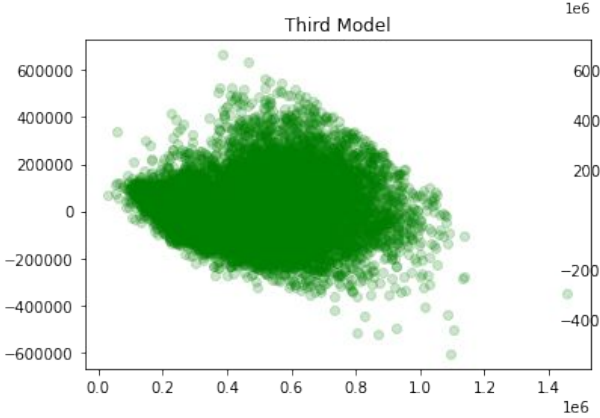

Second Model
Sample Quantiles
Theoretical Quantiles

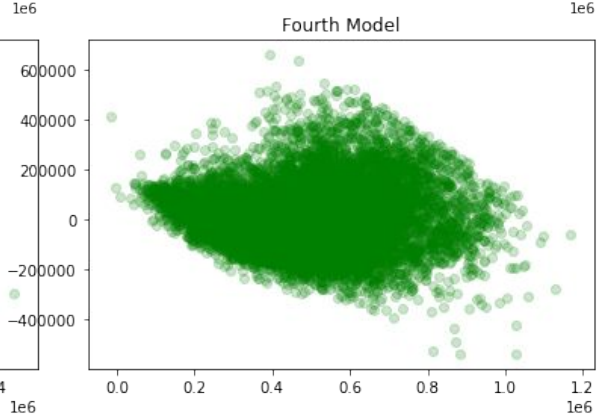

Third Model
Sample Quantiles
Theoretical Quantiles


Fourth Model
Sample Quantiles
Theoretical Quantiles

Homoscedasticity



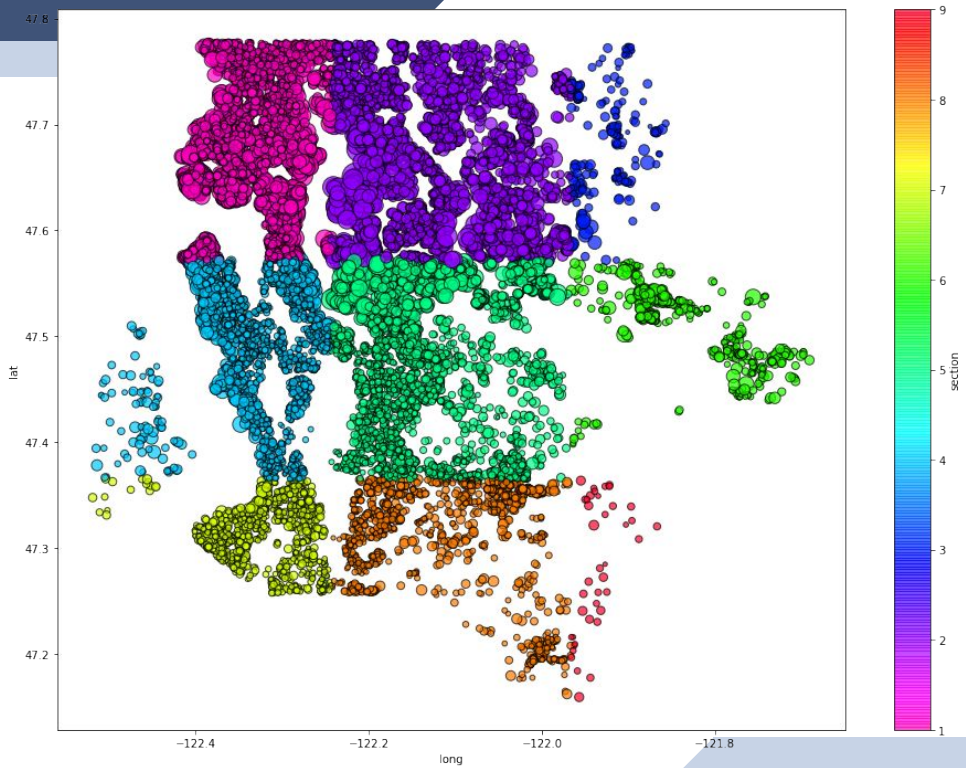




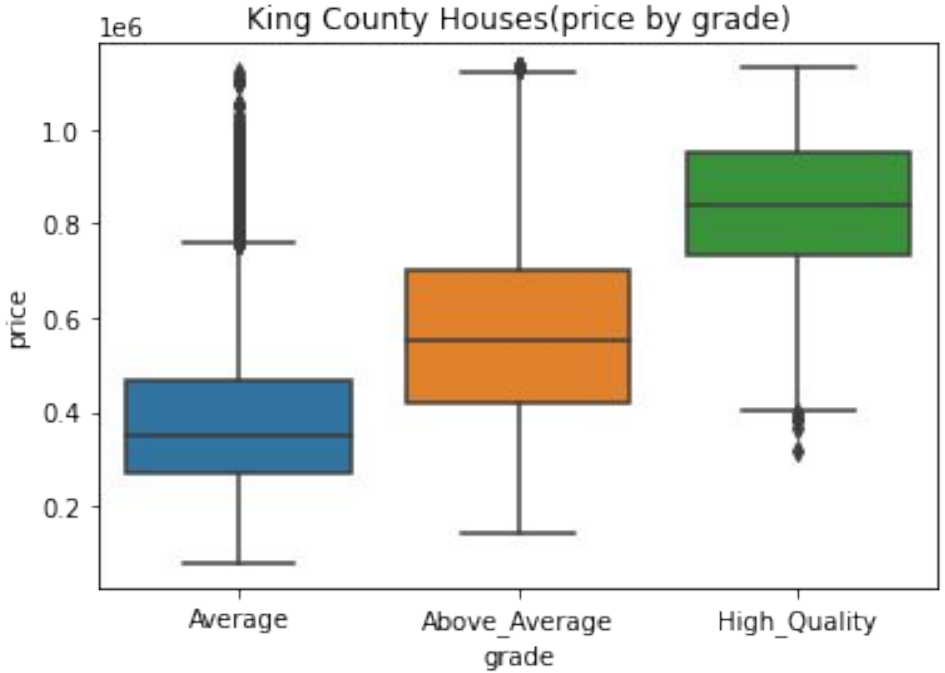
Conclusions

Model 4 advises purchasing homes in sections 1(pink) and 2(purple)

Homes North of 47.57 Lat. and West of -121.97 Long.



Conclusions



Renovations should be made using highest quality contractors and materials

Conclusions

	coefficient
sqft_living	676146.942065
Grade_High_Quality	253667.108839
View_4	218207.762842
Intercept	166583.865718
Condition_5	161292.797051
Condition_4	133805.608255
View_3	108943.847766
Condition_3	105616.408877
Grade_Above_Average	96526.407549
View_1	93599.373751

Future Analysis

- Incorporate other datasets that contain information schools, services, etc.
- Learn and employ geospatial libraries

Thank you
for reading

GitHub: Benjamin2817
Email: benjamin.a.dean.17@gmail.com