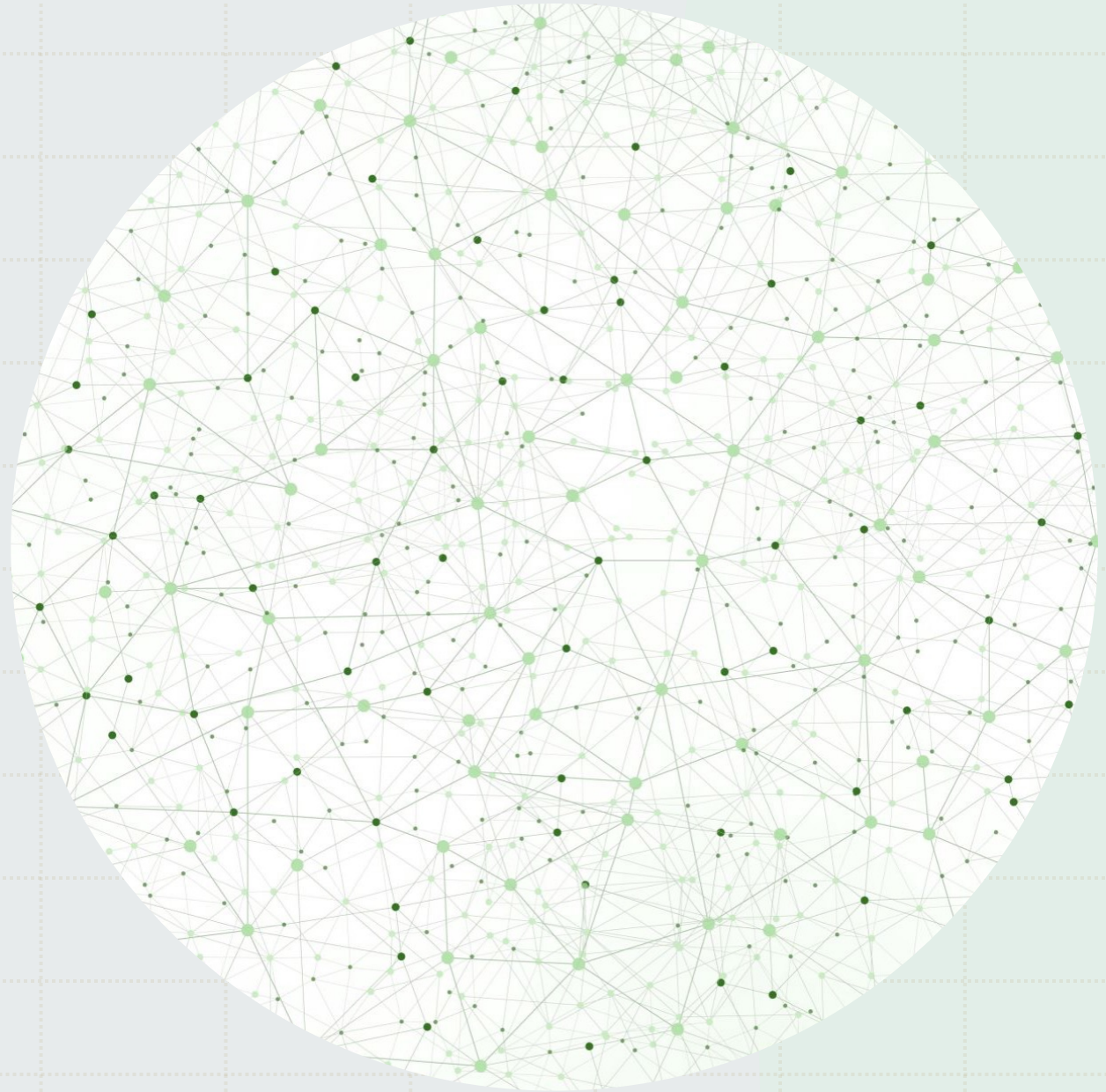


King County Home Sales Analysis

By Justin Giovatto



Introduction

- King County overview:
 - Estimated population of 2,252,782
 - Most populous county in WA
 - 12th most populous county in US
 - Two-thirds of population lives in Seattle suburbs
- Business Problem:
 - This project will create a series of linear regression models to help provide a Kings County real estate agency with information regarding the most important factors in determining home prices for potential clients interested in selling their homes.

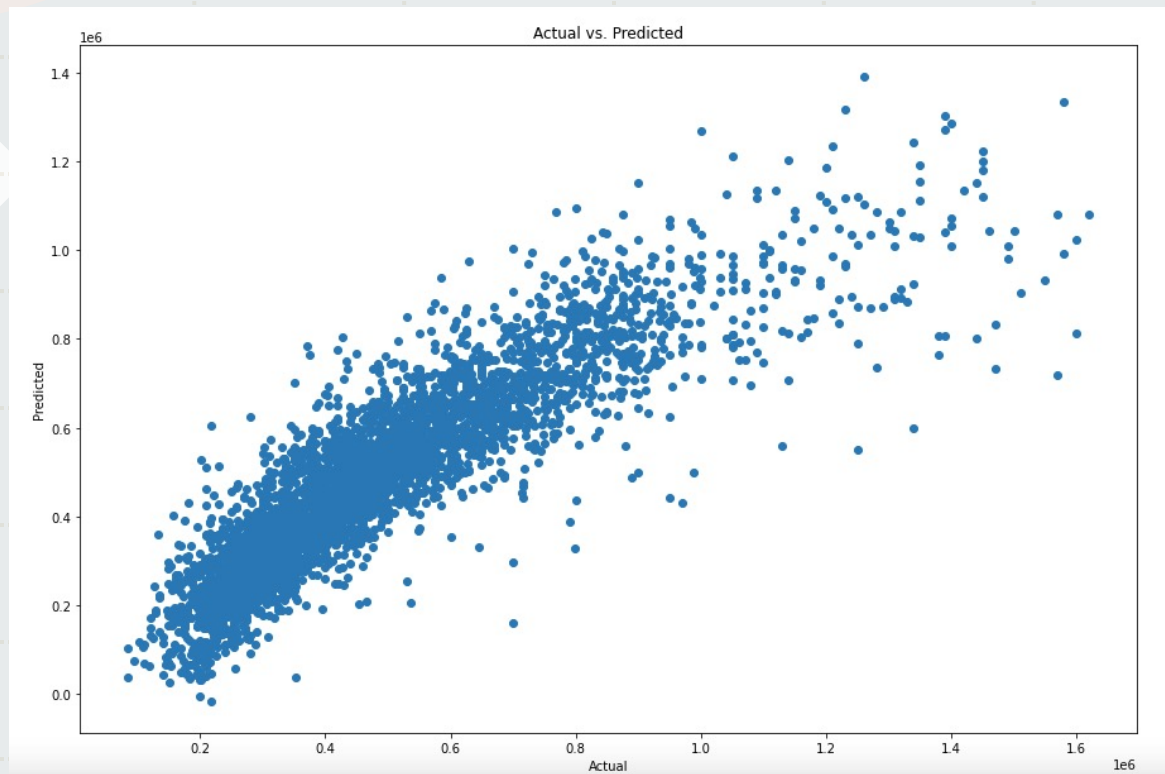


Dataset

- Contains information on 21,597 homes in Kings County, WA
- Models used will focus on the following features from the dataset:
 - Bathrooms
 - Square foot living
 - Square foot lot
 - Grade
 - Square foot above
 - Year built
 - Zip code
 - Average square foot living (15 nearest neighbors)
 - Latitude/Longitude



Model D



Model D Evaluation

Summary

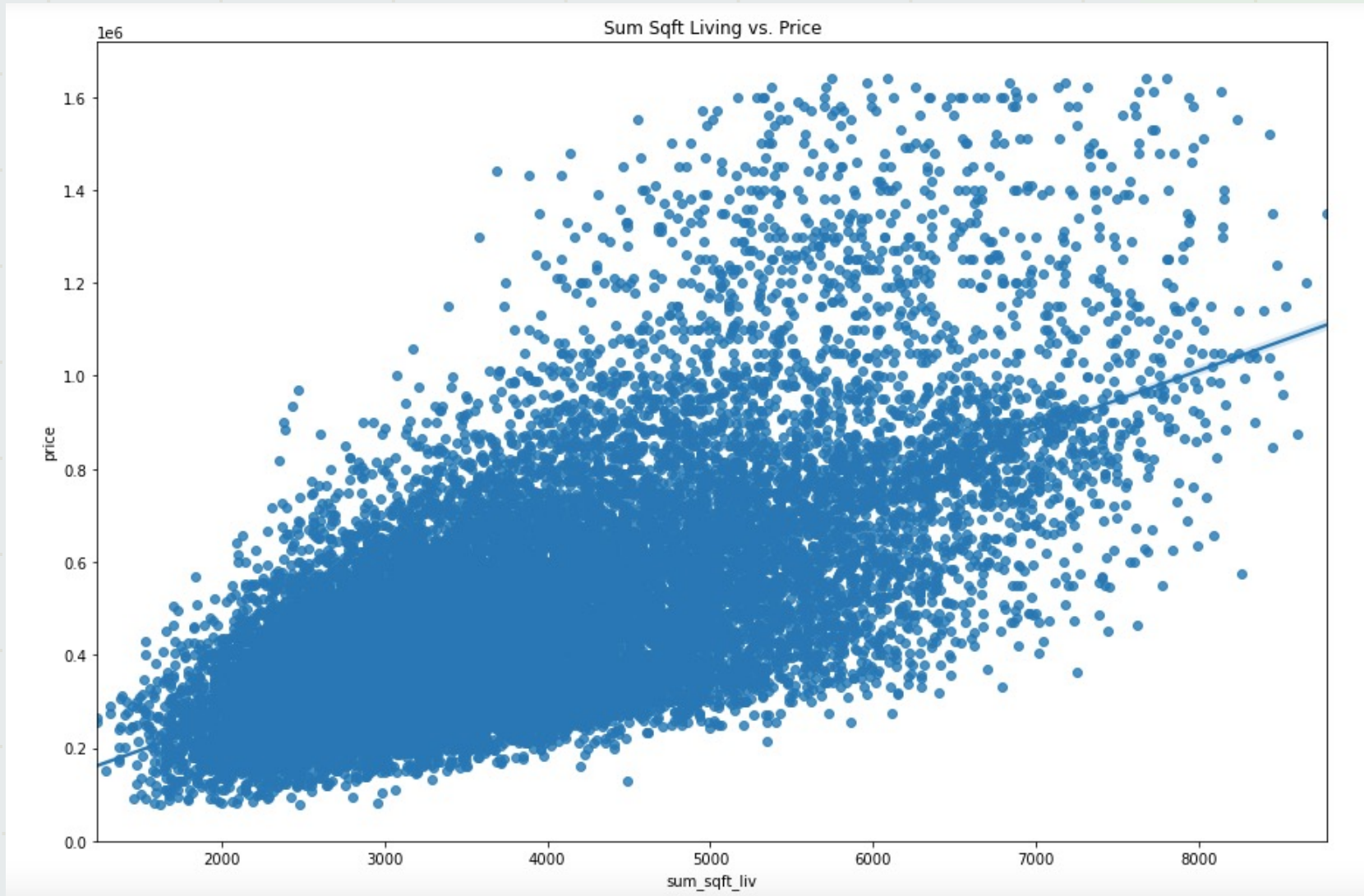
- On average Model D predicts home prices within \$77,473 of true home value.

Model D Evaluation

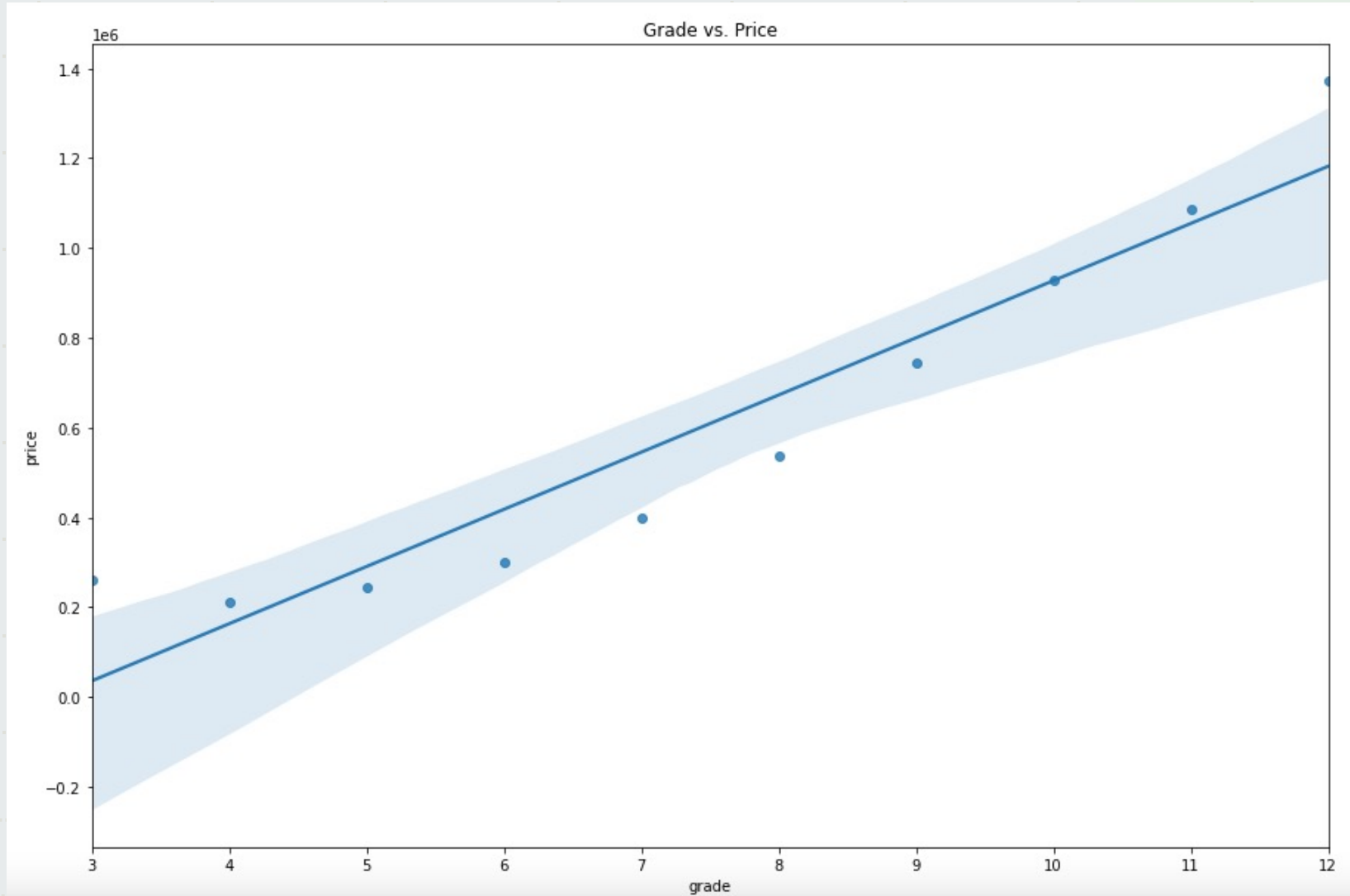
Ranked Features:

- 1.Zip codes
- 2.Sum square foot living
- 3.Grade
- 4.Year Built
- 5.Bathrooms
- 6.Distance from Seattle

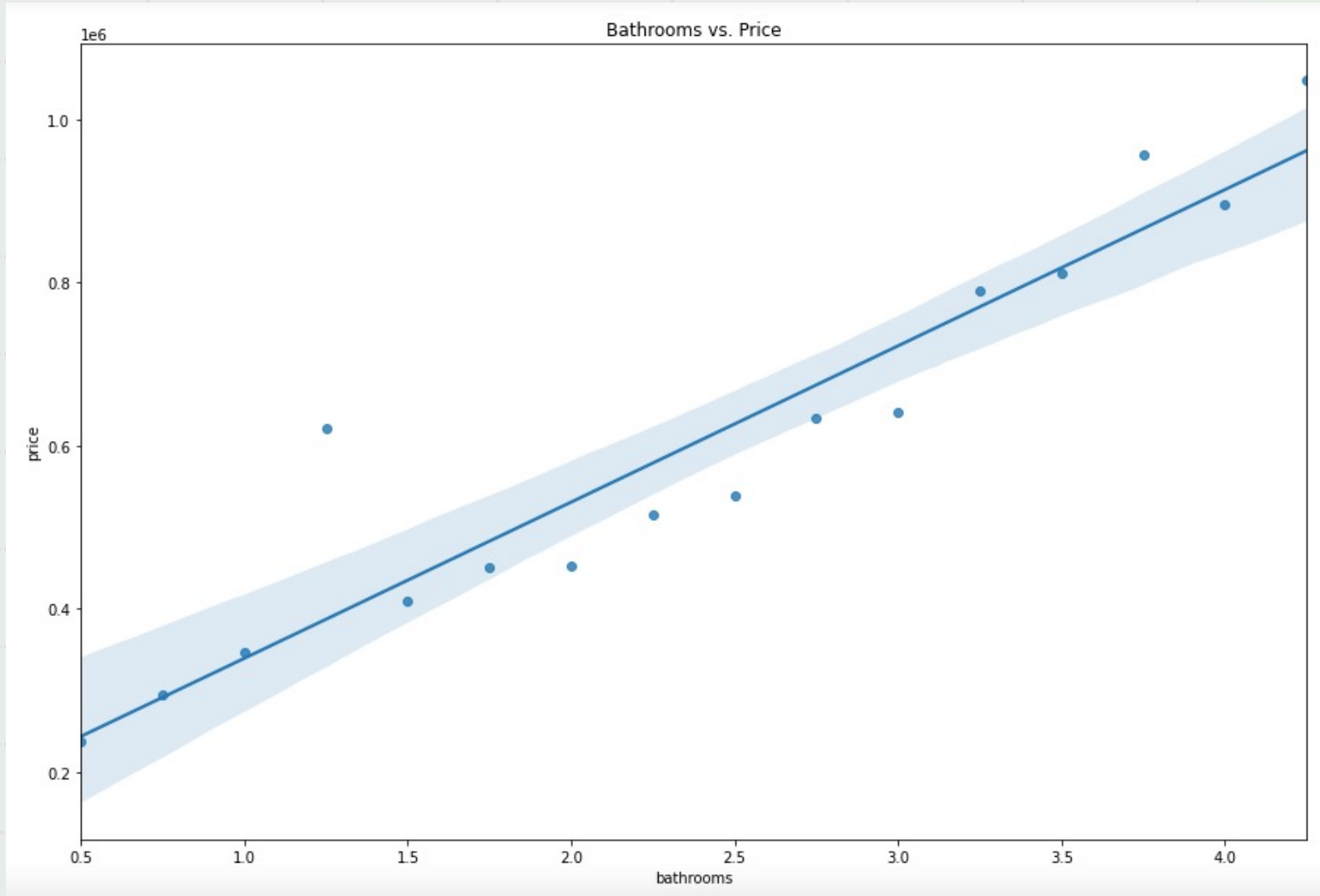
Model D Features (Sum Square Foot Living vs. Price)



Model D Features (Grade vs. Price)

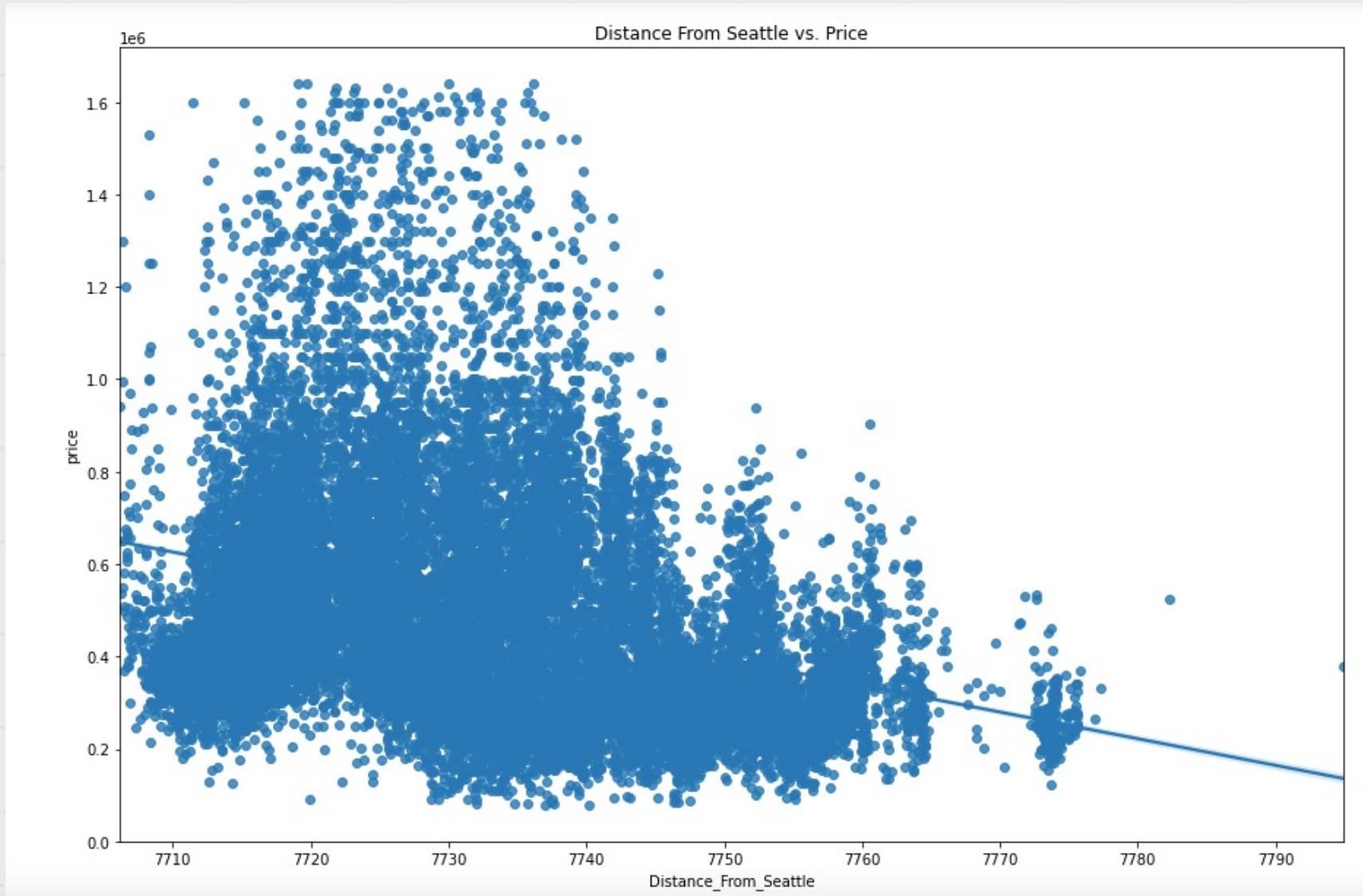


Model D Features (Bathrooms vs. Price)



Model D Features

(Distance from Seattle vs. Price)





Conclusion

- Two most important features according to model analysis:
 - Zip Codes (Most Important)
 - Square foot living/Square foot above/Sum square foot living (2nd Most Important)
- Bathrooms and Grade also appear to be fairly significant features in determining price.



Future Work

In order to better improve the models, will likely need to try more features for comparison.

Will also need to engineer more new features in order to improve model accuracy.

Would also like to test model on datasets from other Washington counties.



Thank You!

Questions?