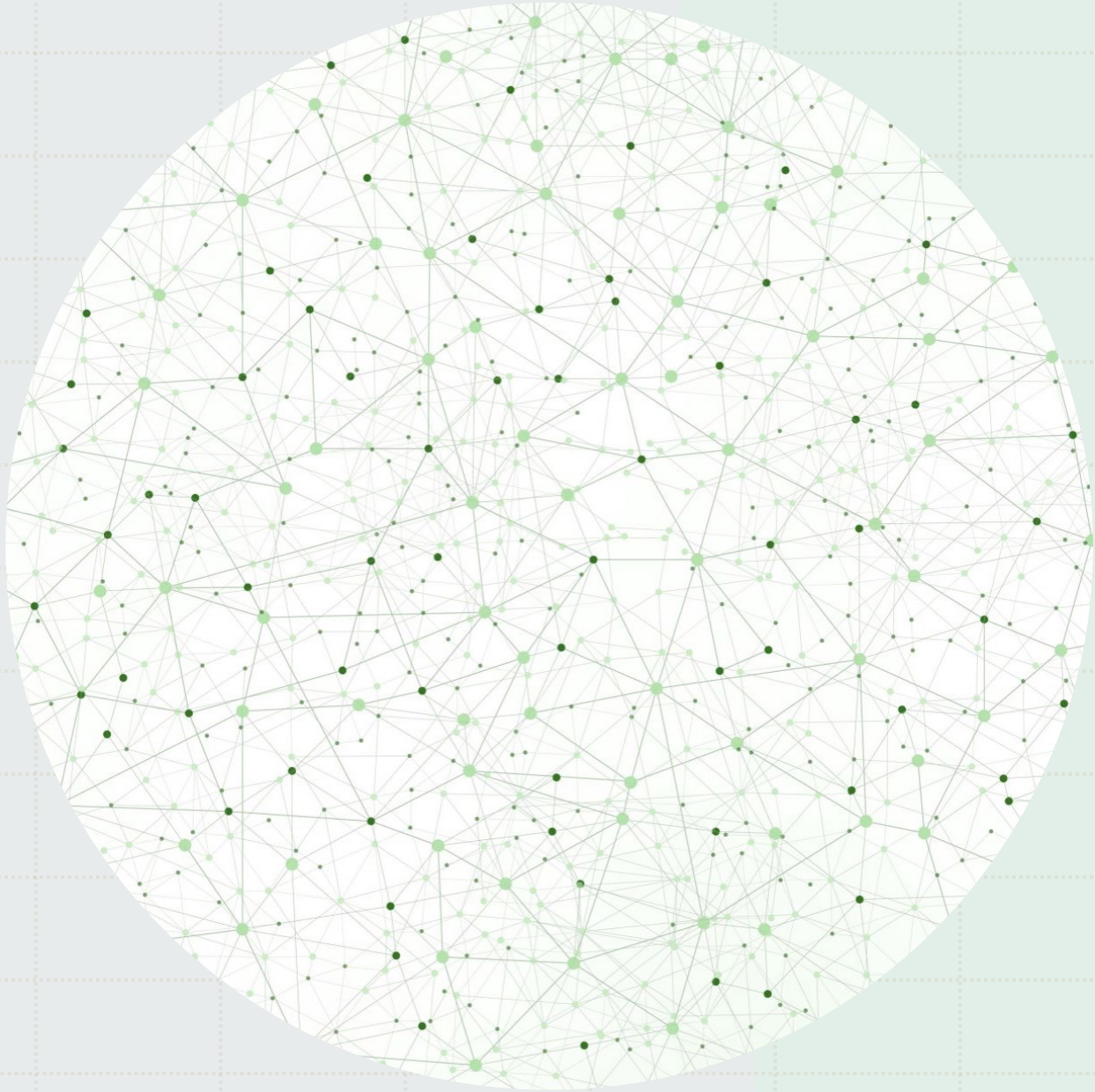
Home Sales Analysis

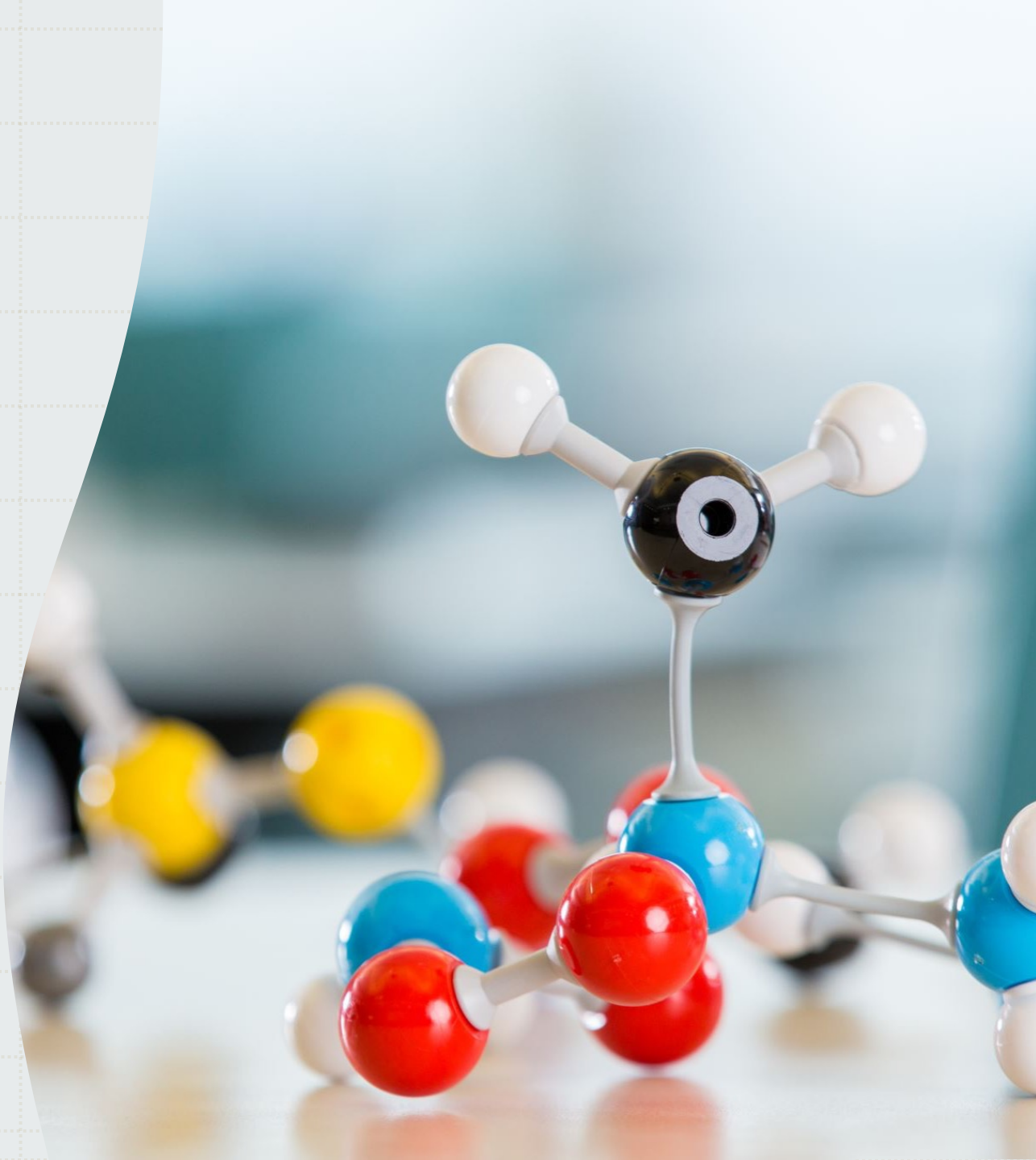
By Justin Giovatto



Business Problem

A King County Real Estate Agency is looking to create a linear regression model that will help provide potential clients who may be interested in selling their home determine the true value of their home.

Model A



Model A Evaluation

Ranked Features:

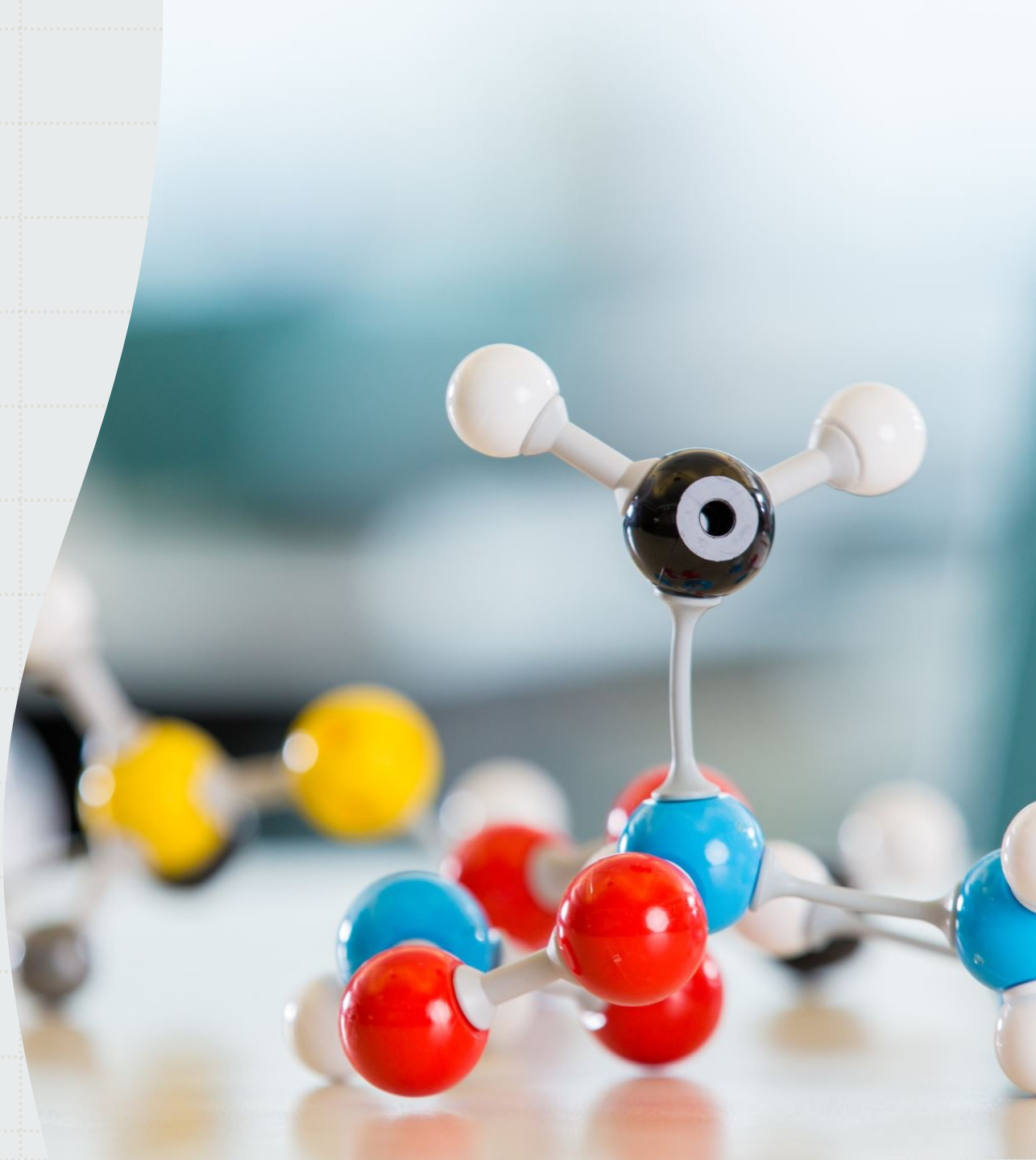
1. Grade
2. Bedrooms
3. Bathrooms
4. Total Square foot Living space

Model A Evaluation

Summary

- On average Model A predicts home prices within $160,590 of true home value.

Model B



Model B Evaluation

Ranked Features:

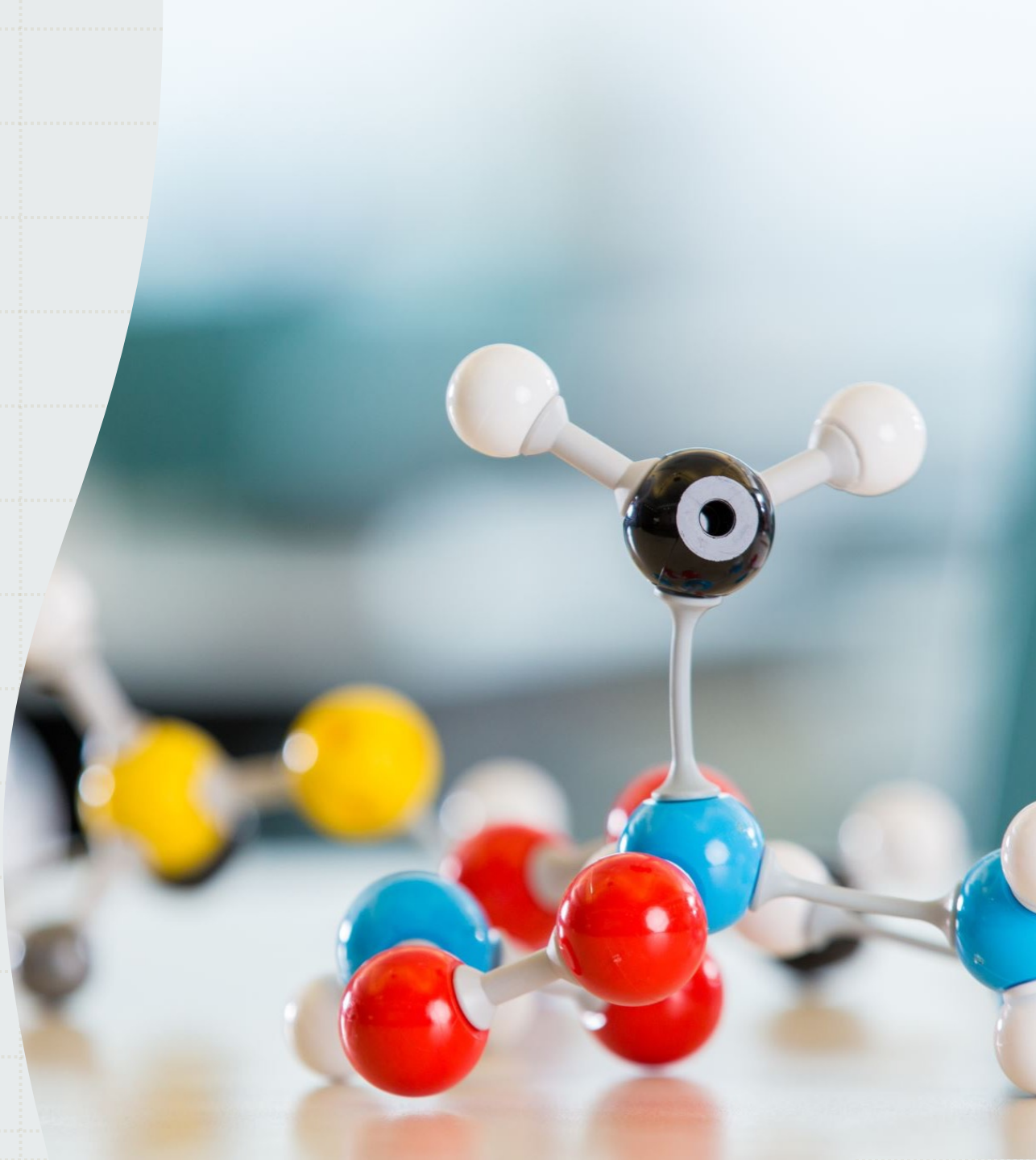
1. Zip codes
2. Square foot living
3. Bathrooms
4. Bedrooms
5. Year built

Model B Evaluation

Summary

- On average Model C predicts home prices within $122,502 of true home value.

Model C



Model C Evaluation

Ranked Features:

1. Zip codes
2. Square foot living
3. Bathrooms
4. Bedrooms
5. Year built

Model C Evaluation

Summary

- On average Model C predicts home prices within $91,518 of true home value.

Model D

Model D Evaluation

Ranked Features:

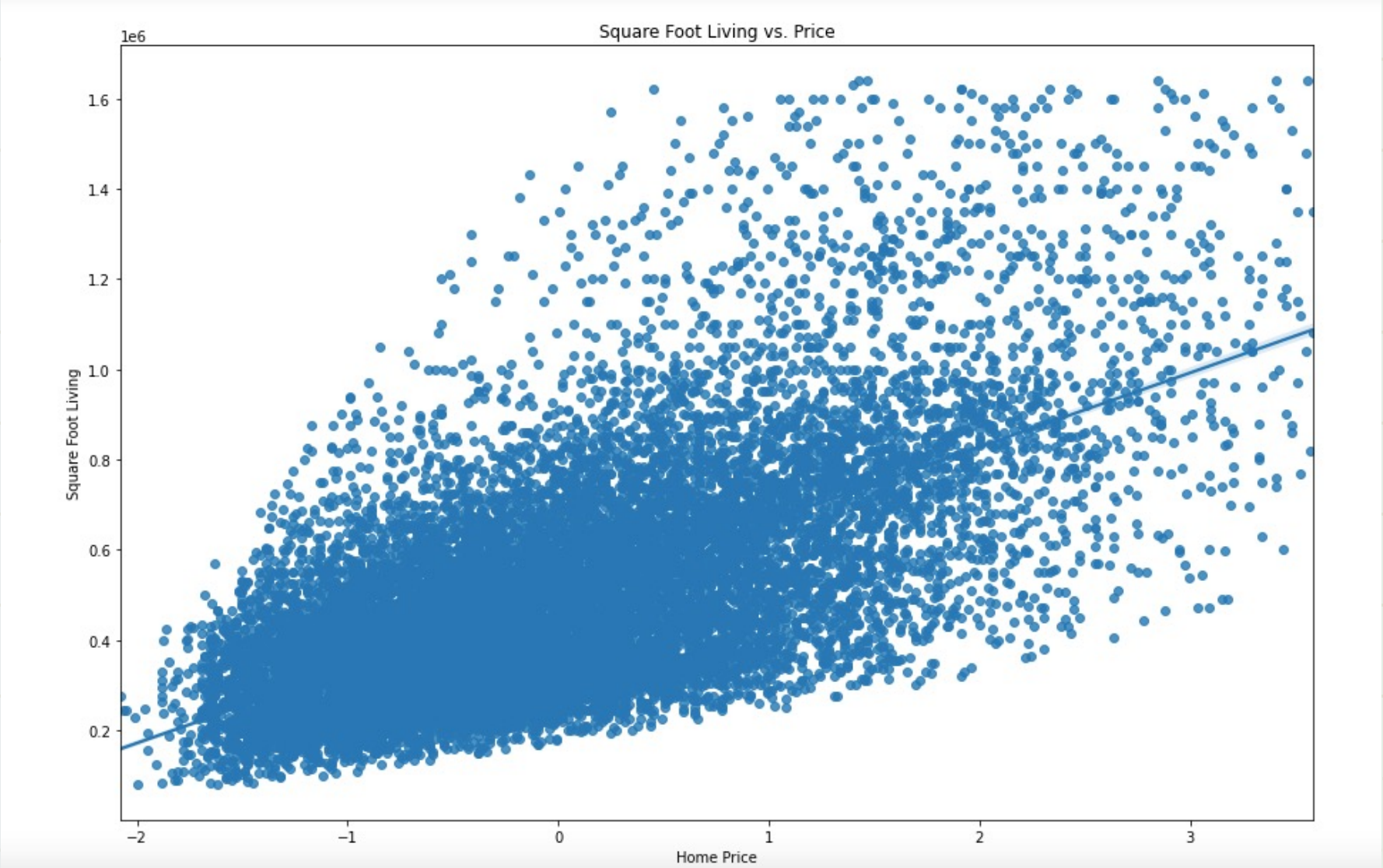
1. Zip codes
2. Square foot living
3. Grade
4. Year Built
5. Bathrooms
6. Bedrooms

Model D Evaluation

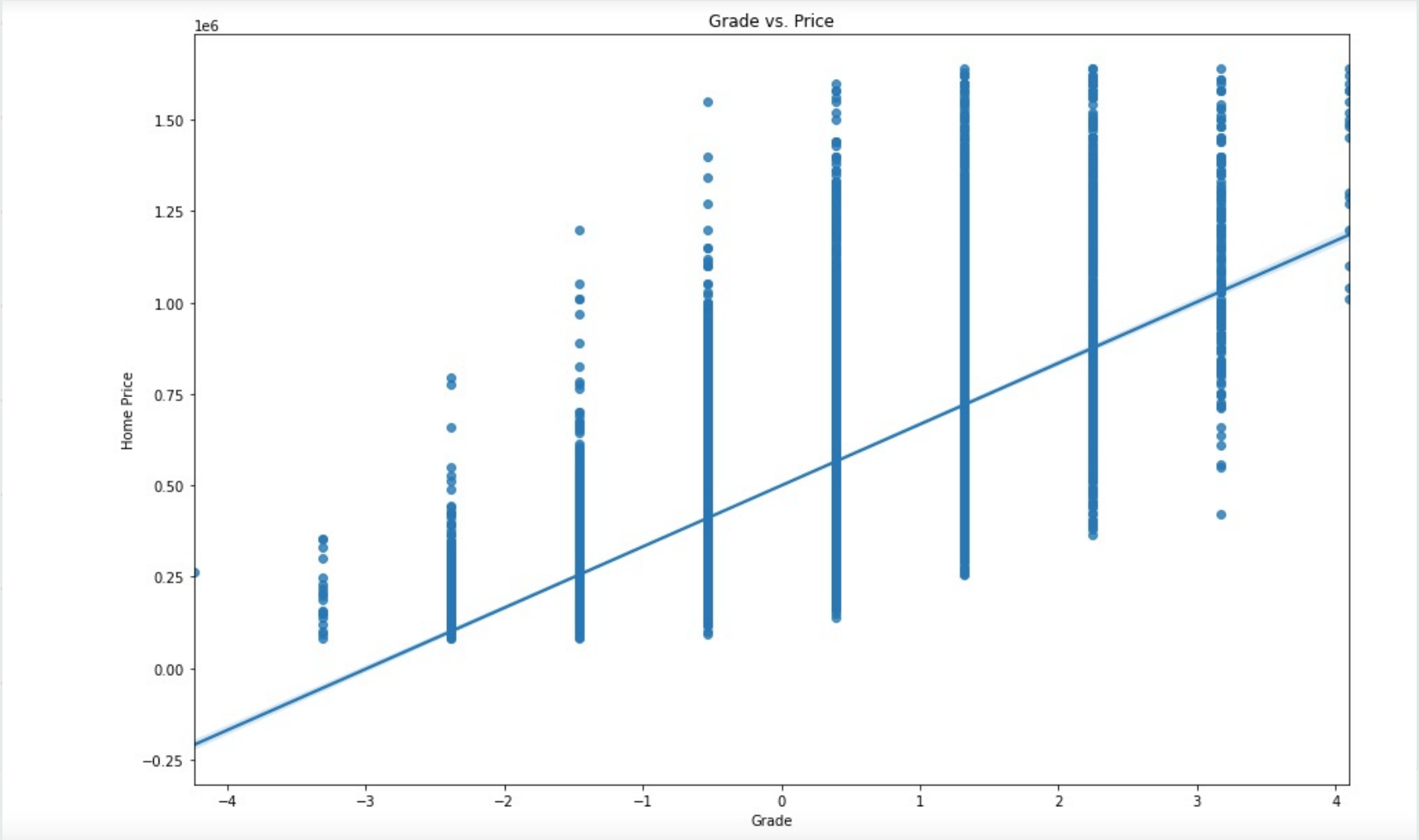
Summary

- On average Model C predicts home prices within $78,281 of true home value.

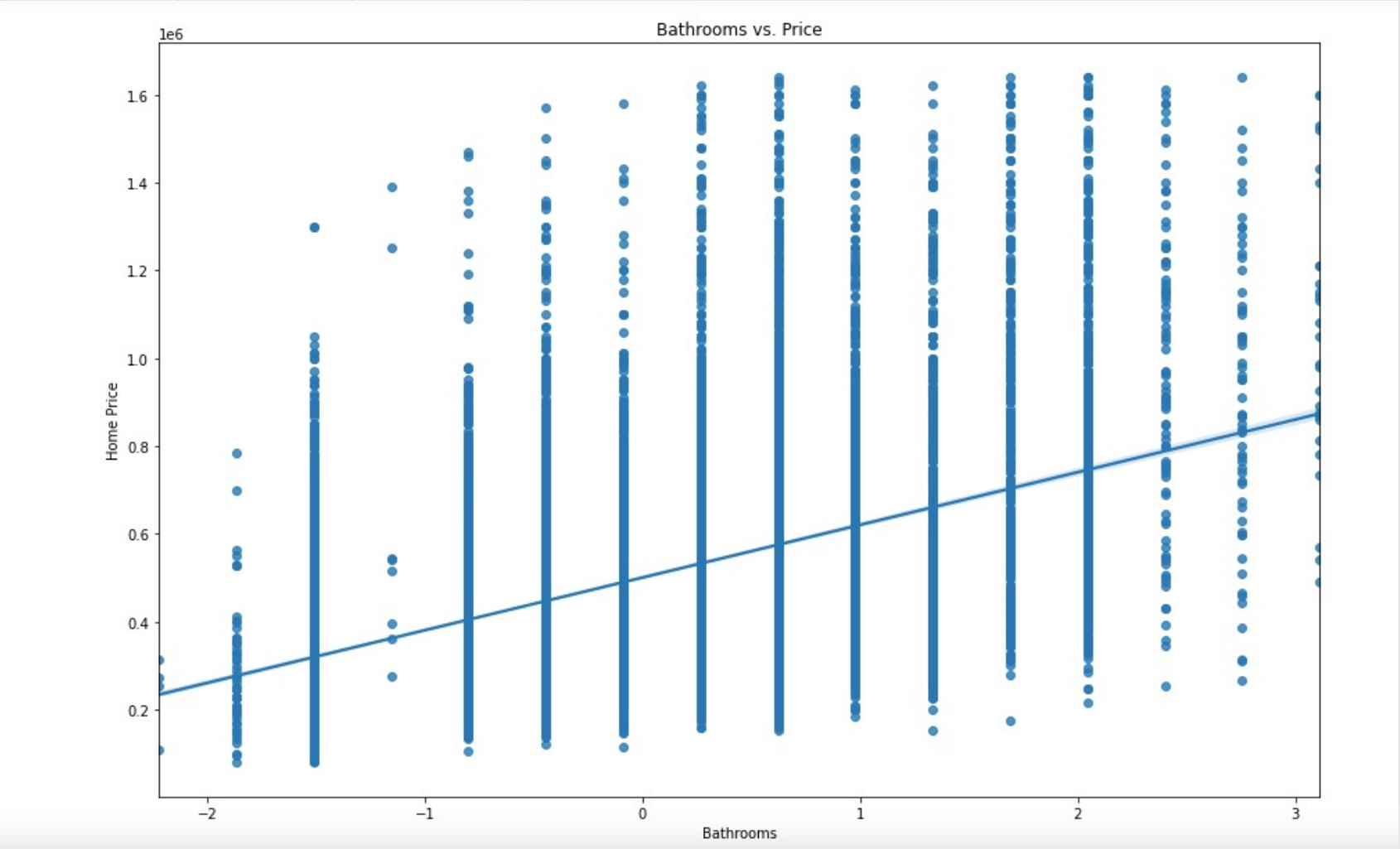
Model D Features (Square Foot Living vs. Price)



Model D Features (Grade vs. Price)



Model D Features Bathrooms vs. Price)



Conclusion

1. Two most important features according to model analysis:

1. Zip Codes (Most Important)

2. Square foot living/Square foot above (2nd Most Important)

3. 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.

Thank You!

Questions?