

REAL ESTATE

CASE STUDY: REGRESSION
MAY 2023



IRON
HACK

Table Of Content

- **About Us**
- **Mission**
- **How we explored
the data**
- **Our properties**
- **Important
features**
- **Our model**
- **Conclusion**
- **Contact Us**

About Us

THREE DATA ANALYSTS WORKING FOR SOTHEBY, A LUXURY
REAL ESTATE COMPANY BASED IN SEATTLE.

FOR THOSE WHO SEEK AN EXCEPTIONAL HOME AND LIFE,
THERE IS ONLY SOTHEBY'S INTERNATIONAL REALTY.

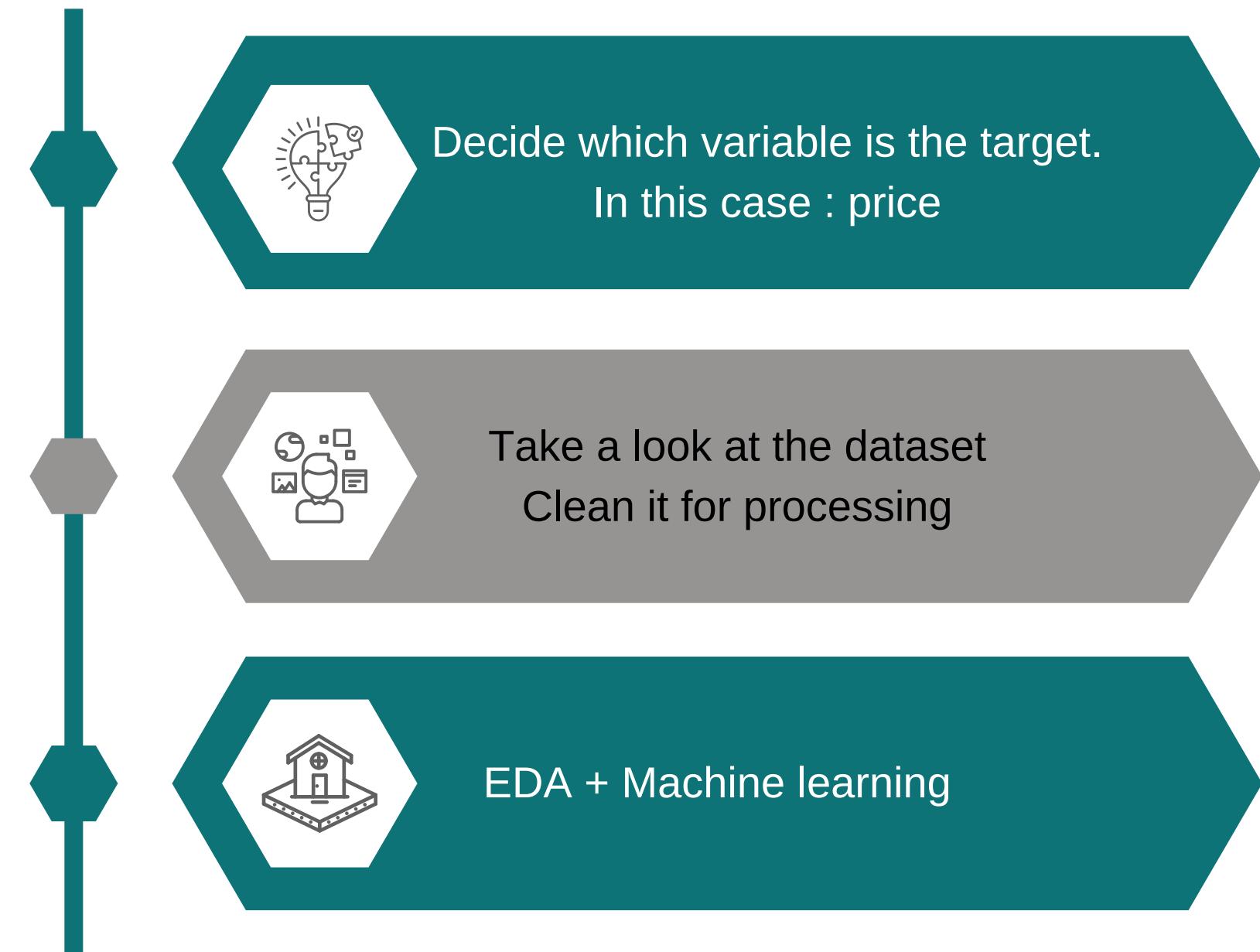


Mission

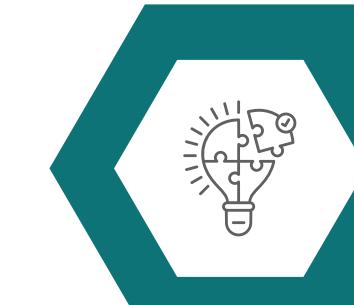
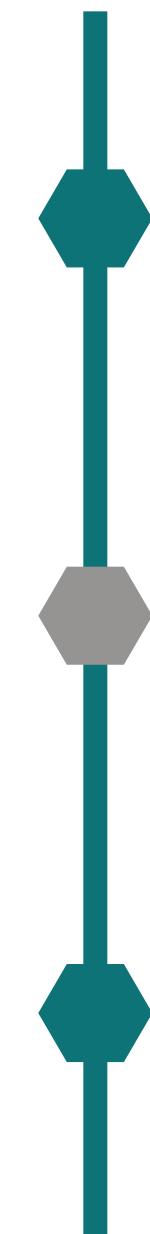
Our goal is to build a model that will predict the price of a house based on features provided in a dataset and to understand which factors are responsible for higher property value - \$650K and above



How we explored the data



Machine Learning



First Model
R2: 0.63



Second Model
R2: 0.82



Third Model
R2: 0.94

Most Important Features



Waterfront



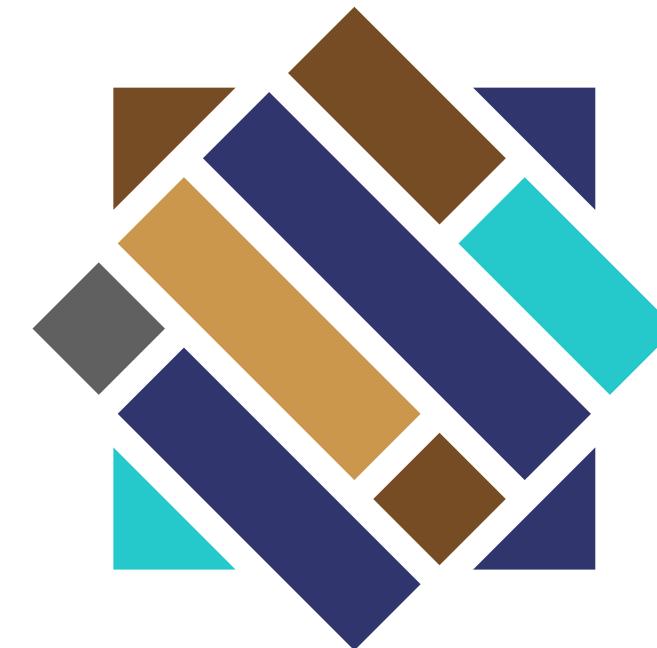
Grade



Bedrooms

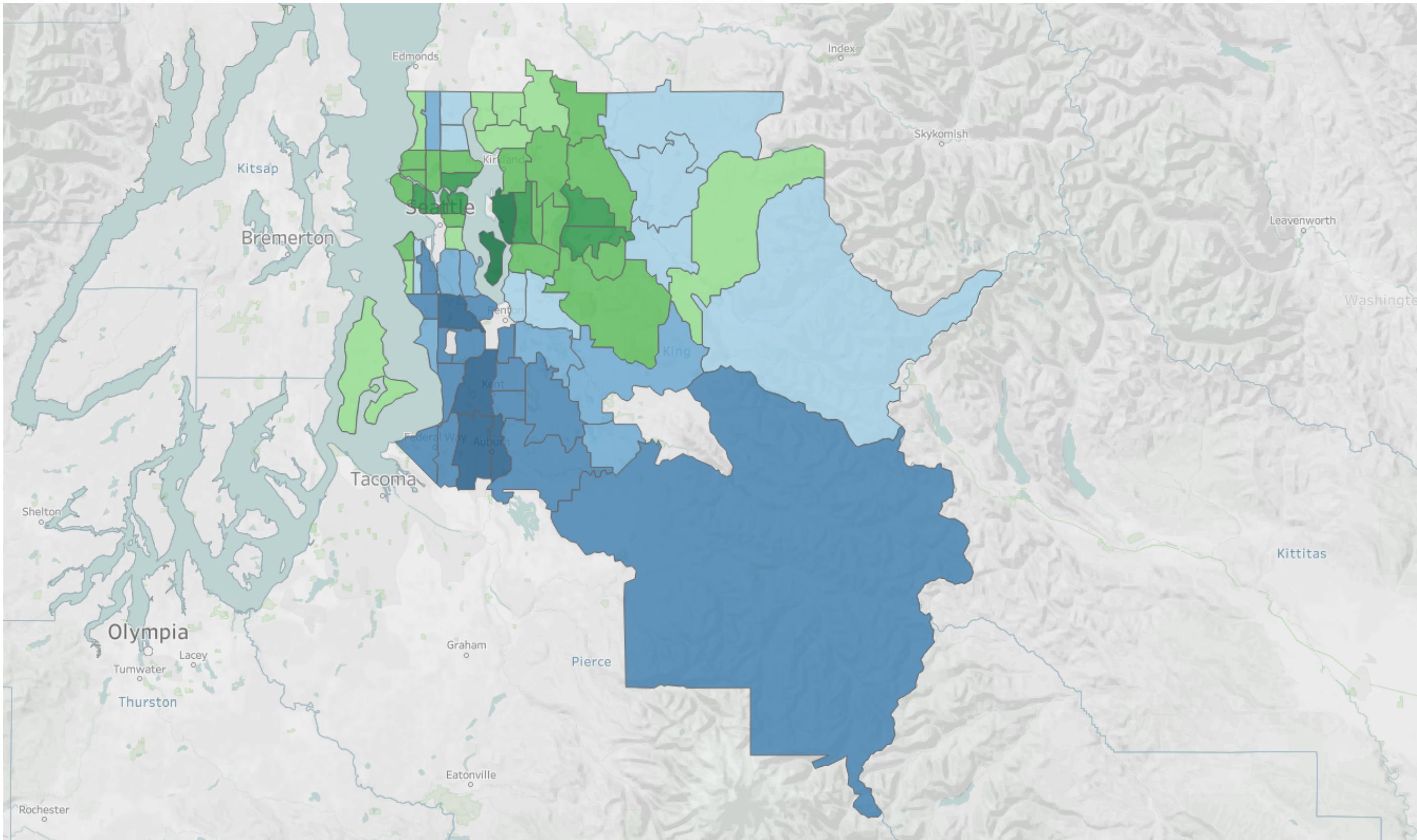


Bathrooms

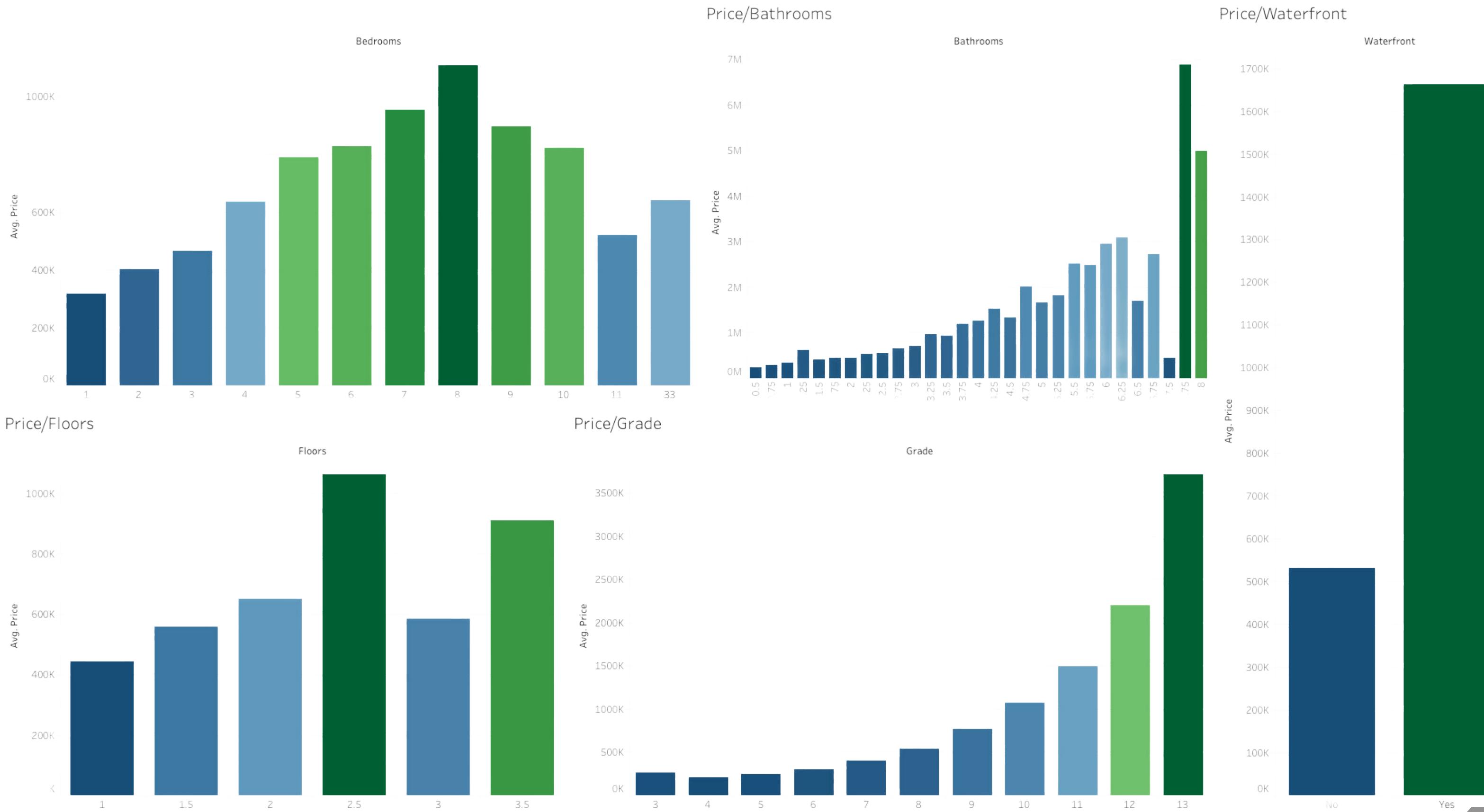


Floors

Average prices

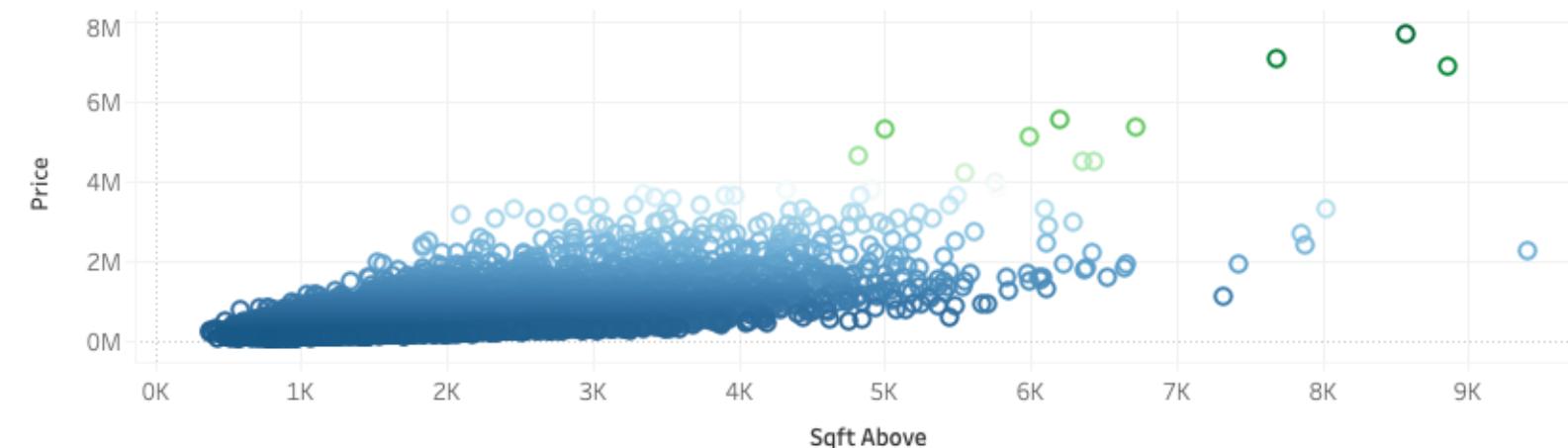


Average price by features

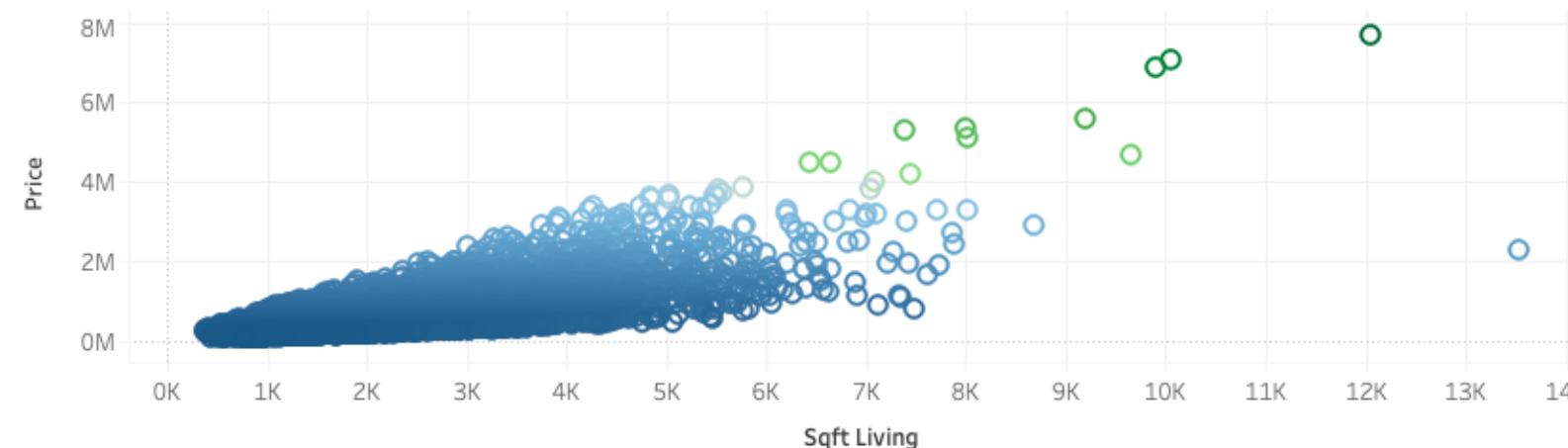


Average price by features

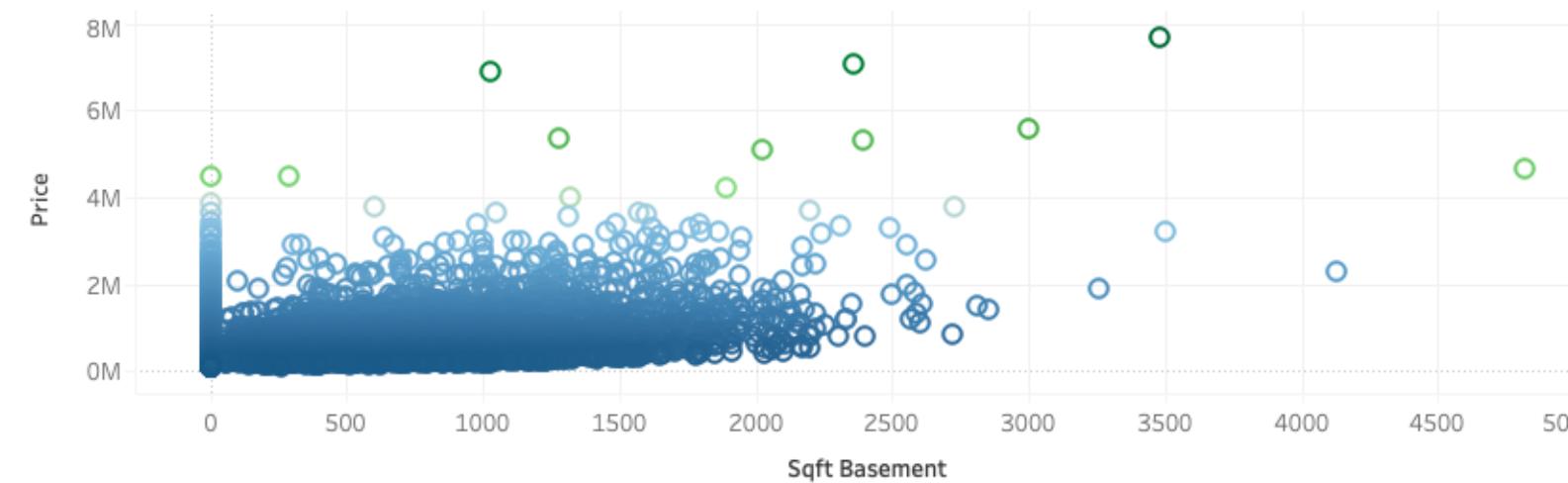
SQFT Above



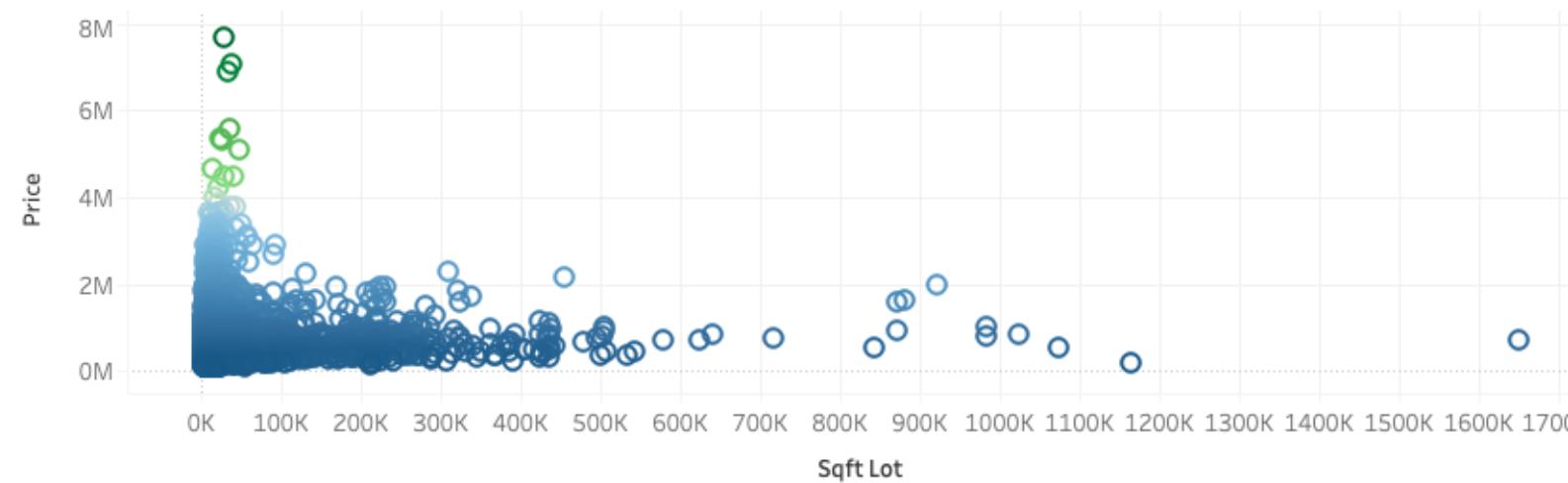
SQFT Living



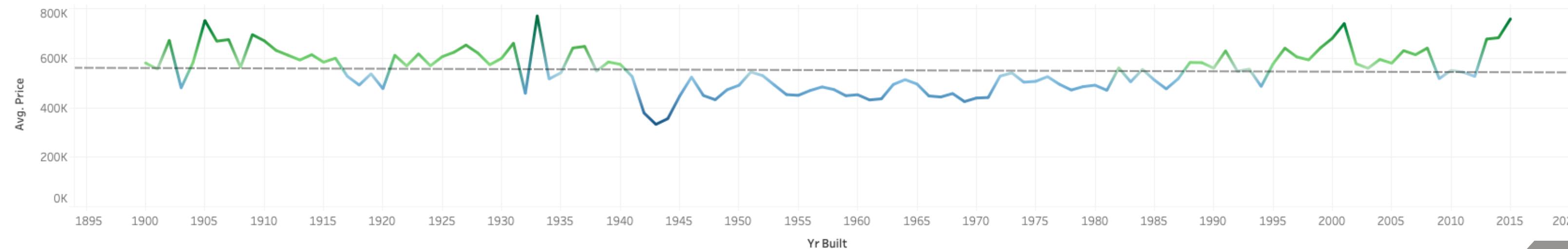
SQFT Basement



SQFT Lot



YrBuilt Tendency



Most Important Features

Higher Property Value



Waterfront



Grade



Bathrooms



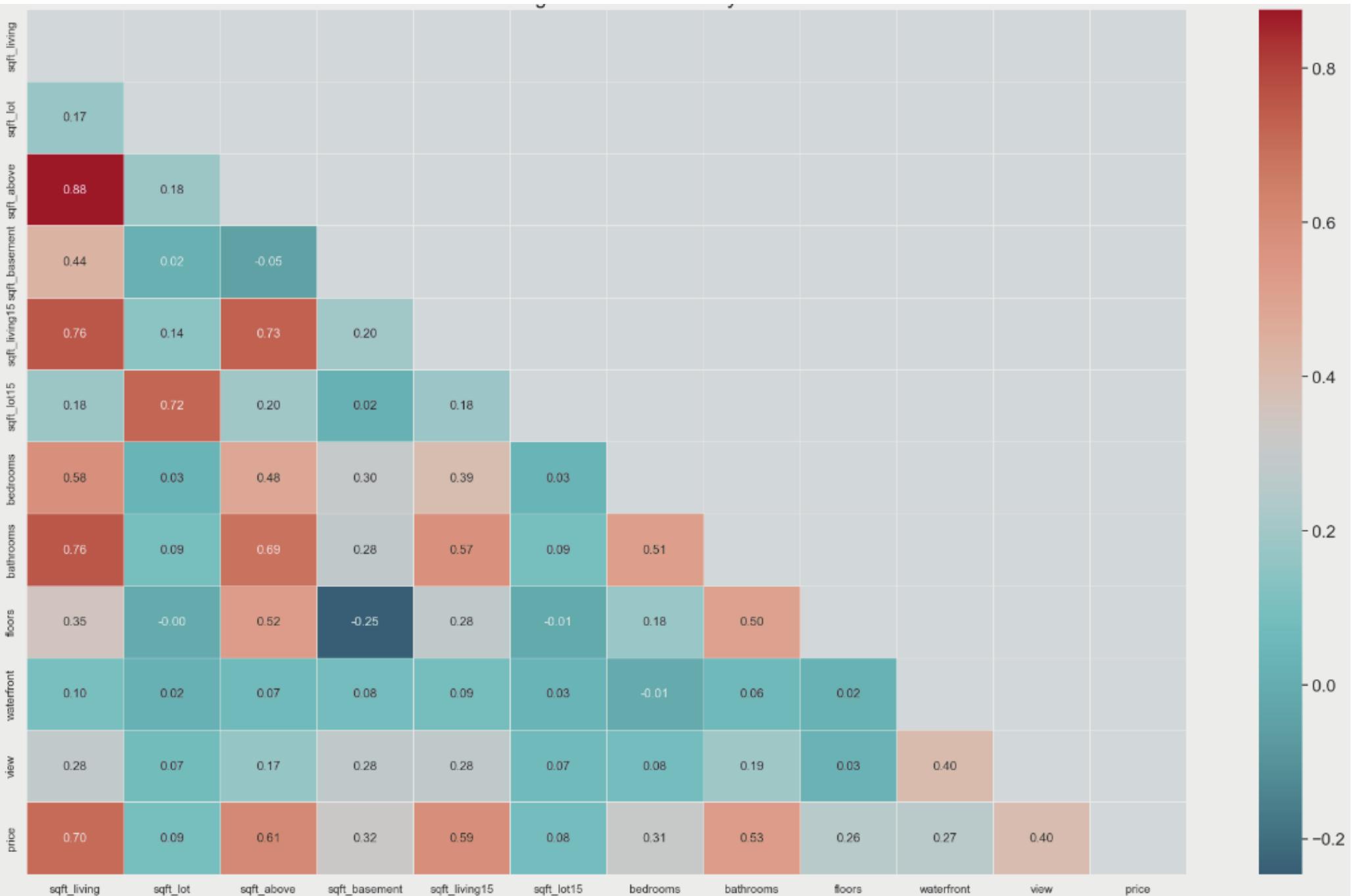
Floors



Bedrooms

Improving the Model

Pearson

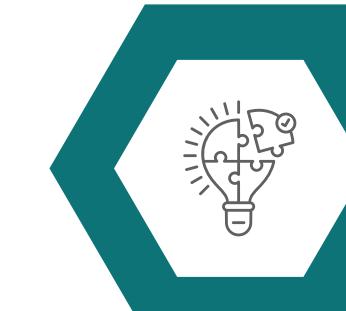
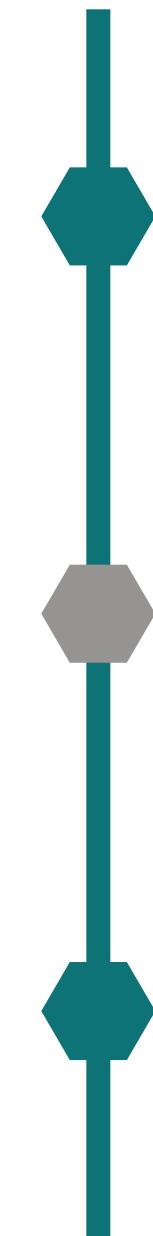


Spearman



Conclusions

We worked on improving the model to get a perfect measure of how well our regression line approximates the actual data



First Model
R2: 0.63



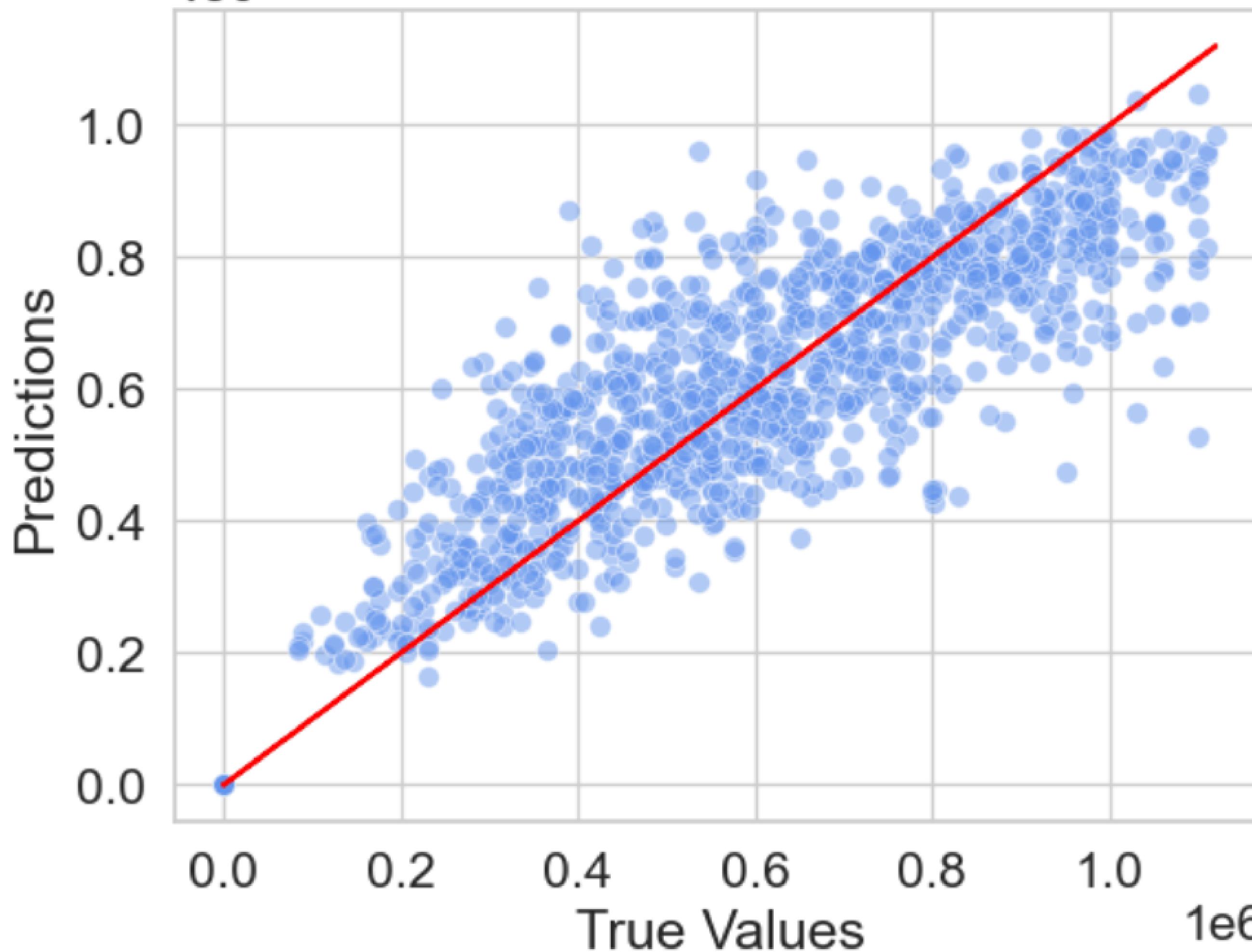
Second Model
R2: 0.82



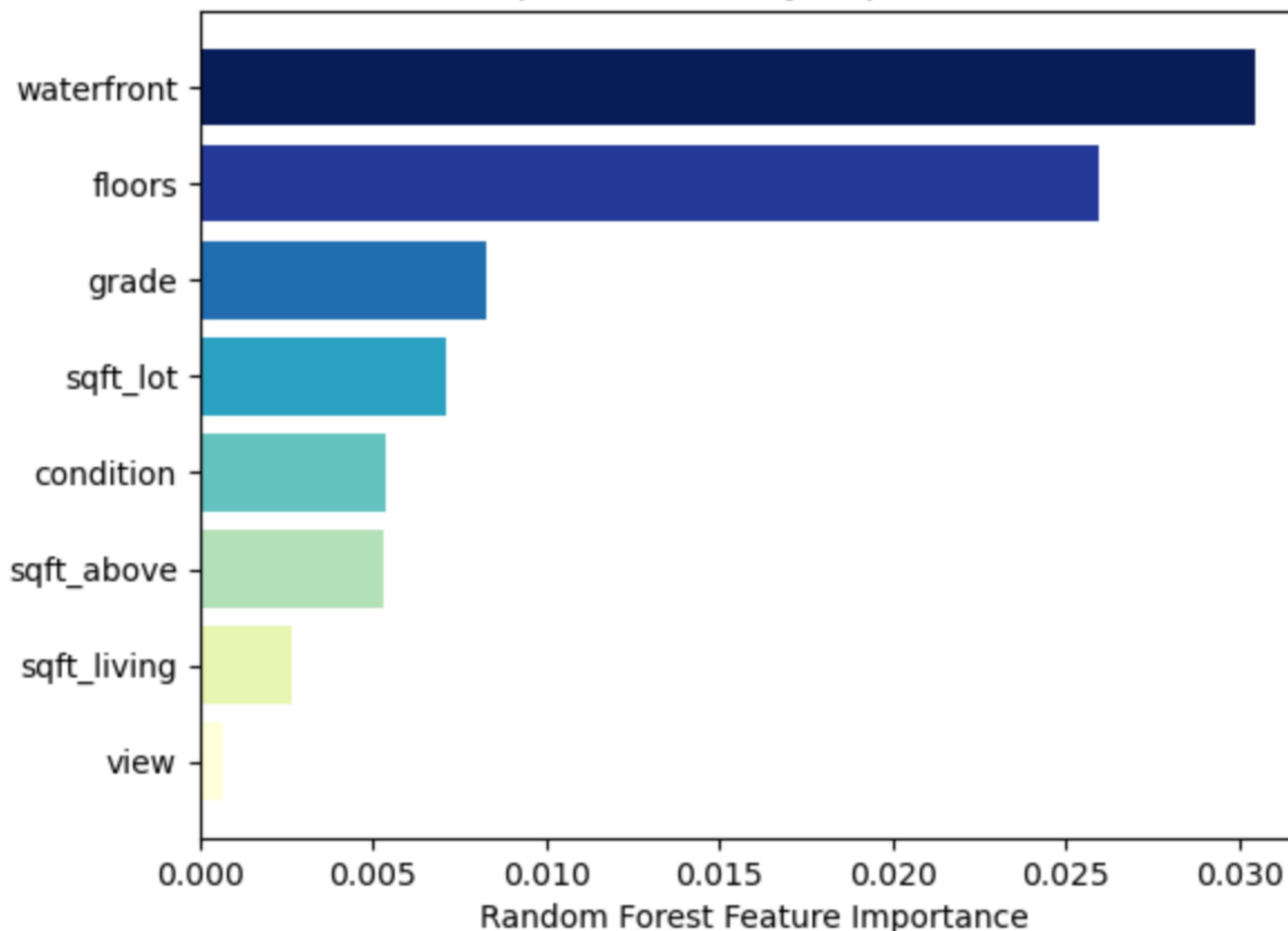
Third Model
R2: 0.94

*This models are too expensive for the company as they take a lot of features to work

Final Model Predictions



Top 8 Features by Importance



Performance

After some improvements we noticed that the model could perform as good as before with less features



MLP Model
R2: 0.86



KNN Model
R2: 0.82

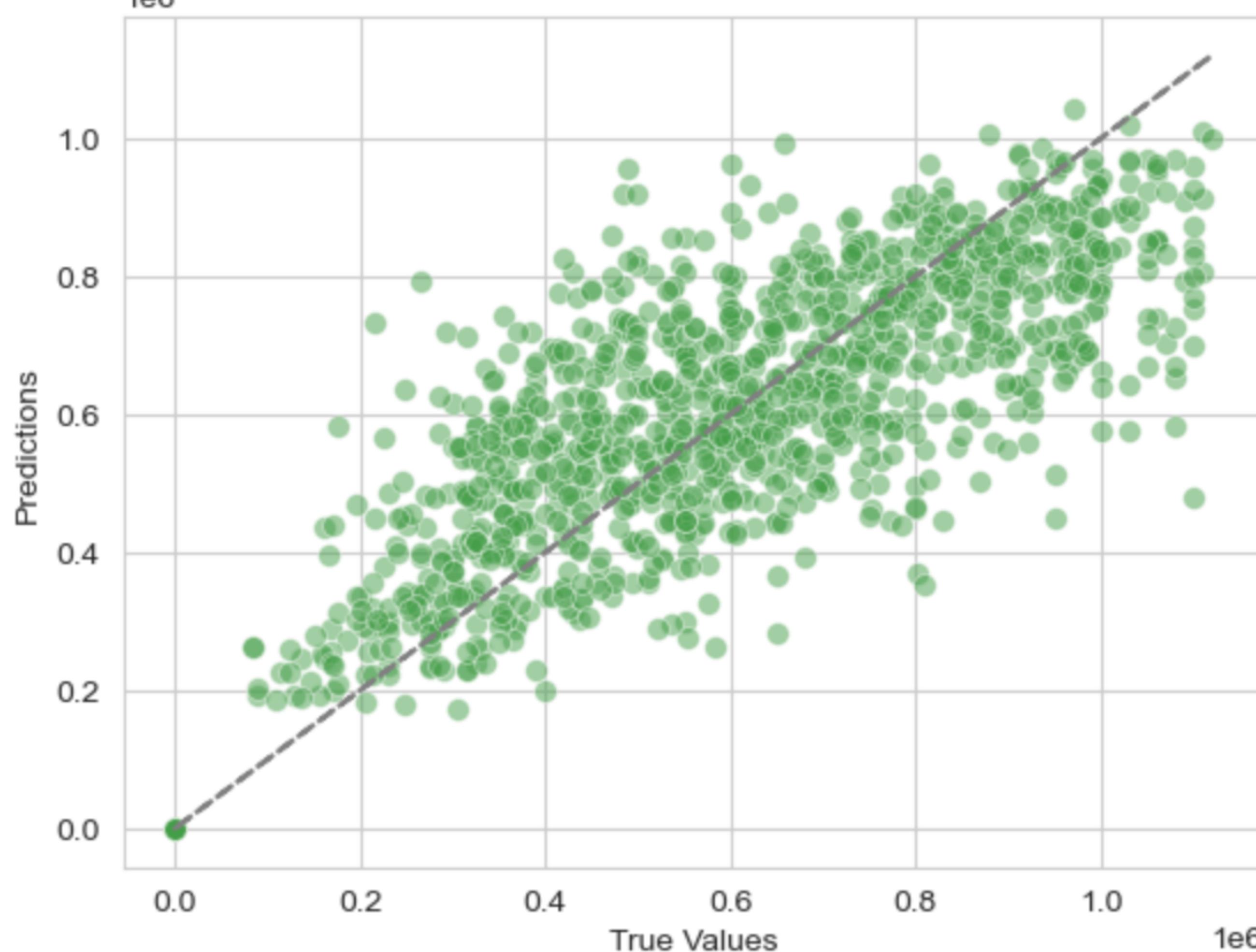


Improved RF
R2: 0.94



RF Reduced
R2: 0.92

8 Features Model Predictions



Takeaways / Conclusions

- The final model just takes **8 features** into the account, making it **less expensive**.
- The most important **high-value properties features** are; waterfront, grade, bathrooms, floors and bedrooms.
- Using **Machine Learning techniques** can decently improve **property investment** and informed and data-driven decisions.

Contact Us !

