

Predicting House Prices with the King County Housing Dataset

JULY 2023

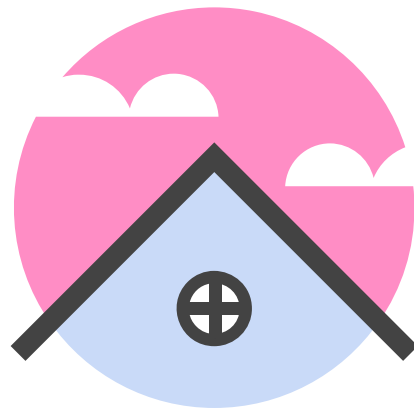


OUR DATA

The King County Housing Data Set contains information about the size, location, condition, and other features of houses in Washington's King County.

After data cleaning:

- 21,000 house sales
- 20 variables



MYHOME
KING COUNTY

BUSINESS UNDERSTANDING

How to fairly value homes in King County so that they may offer their clients fair recommendations when it comes to buying and selling homes.





PROCESS

01

**Data cleaning and
preprocessing**

02

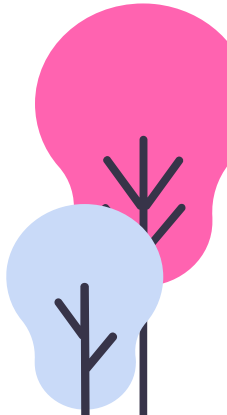
Exploring the data

03

Building linear models

06

Interpreting model results

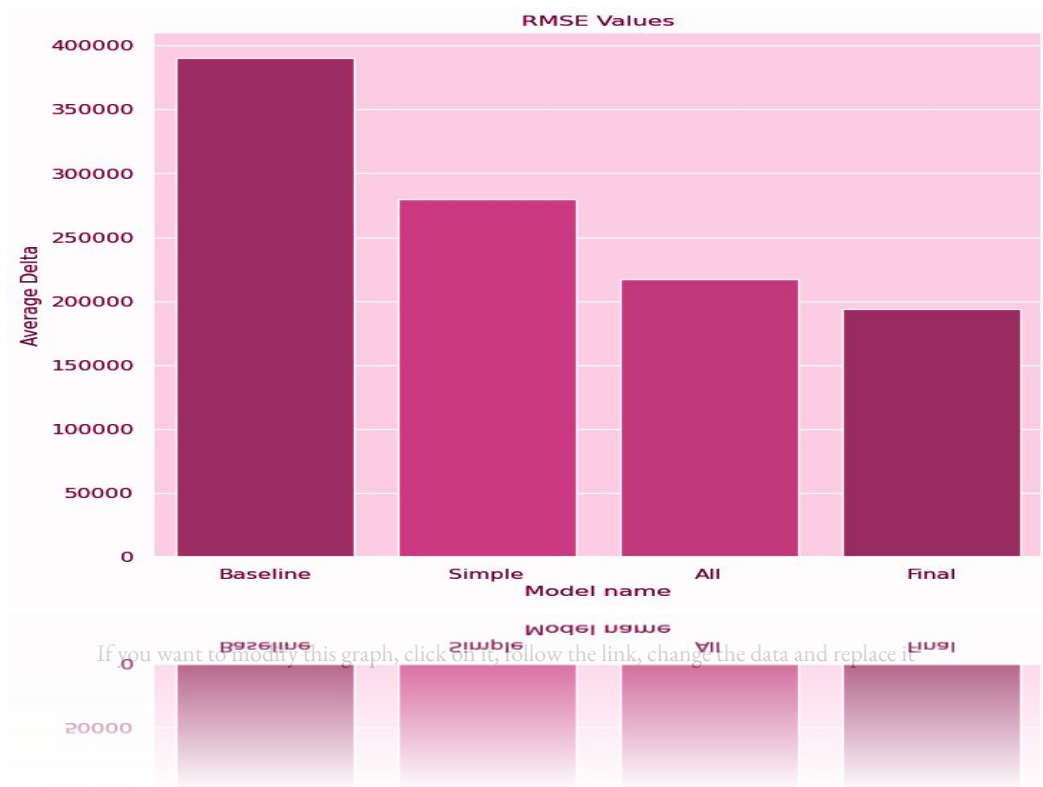


The difference between price and prediction

Consistent Improvement across models

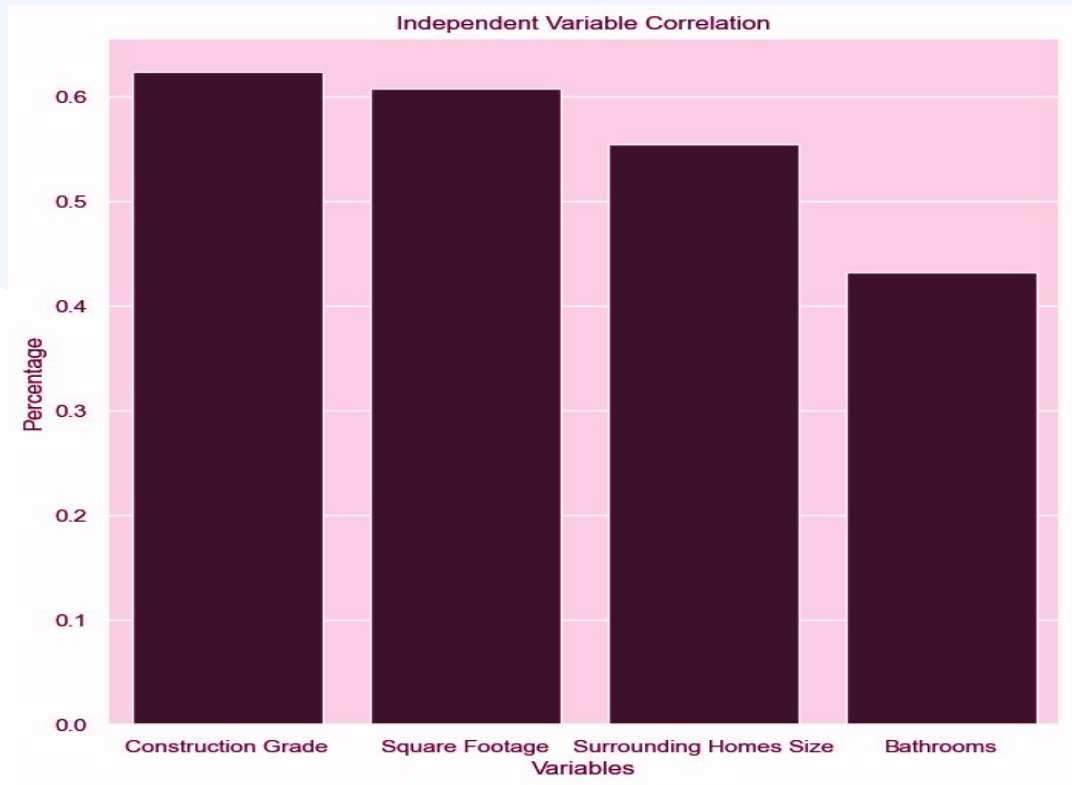
Driven by:

- Statistically significant values
- Following linear assumptions



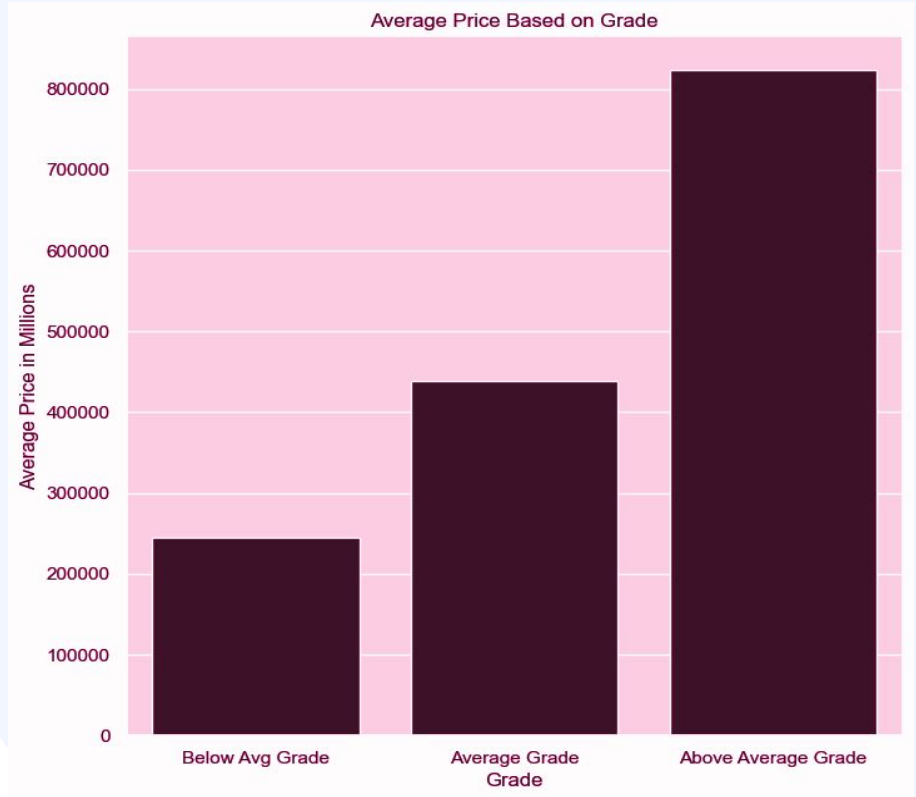
Biggest Factors for Sale Price

- Square footage
Square footage of
surrounding homes
- Construction grade



Construction Grade

Above average graded homes are much more expensive



In Conclusion

Square Footage and Grade matter

- **Uneffective model**
- **RMSE: 194K**
- **R-Squared: 48%**



Next step

- It is evident that linear regression may not have been the most suitable tool for extracting the maximum accuracy from this dataset. In future iterations, we intend to explore more powerful machine learning techniques to improve the predictive capabilities of our model and achieve a higher level of accuracy



THANKS

Does anyone have any questions?

