House Sales in King County, USA

Fridah Kimathi

December, 2022

SUMMARY

❑ The project analyzes King County housing dataset.

❑ The aim is to aid a Real Estate Agency in King county, Washington-USA predict the house prices for single-family homes that their potential clients are looking buy and/or sell.

❑ House price prediction is done by analyzing the most influential features that affect house prices.

OUTLINE

- Business Problem
- Data
- Methods
- Results
- Conclusion
- Limitations
- Next Steps

BUSINESS PROBLEM

1 Predict house prices

2 Highlight influential factors affecting house prices

3 Reduce costs

DATA

The project analyzes about 21597 houses sold in King County, USA in the years 2014 and 2015.

The relevant features in the dataset used for analysis and linear regression in this project are;

- house id, date
- house price
- number of bedrooms and bathrooms
- size of living space and property, number of floors the house has
- the condition and grade of the house according to the King County System
- the year the house was built.

METHODS

Business Understanding
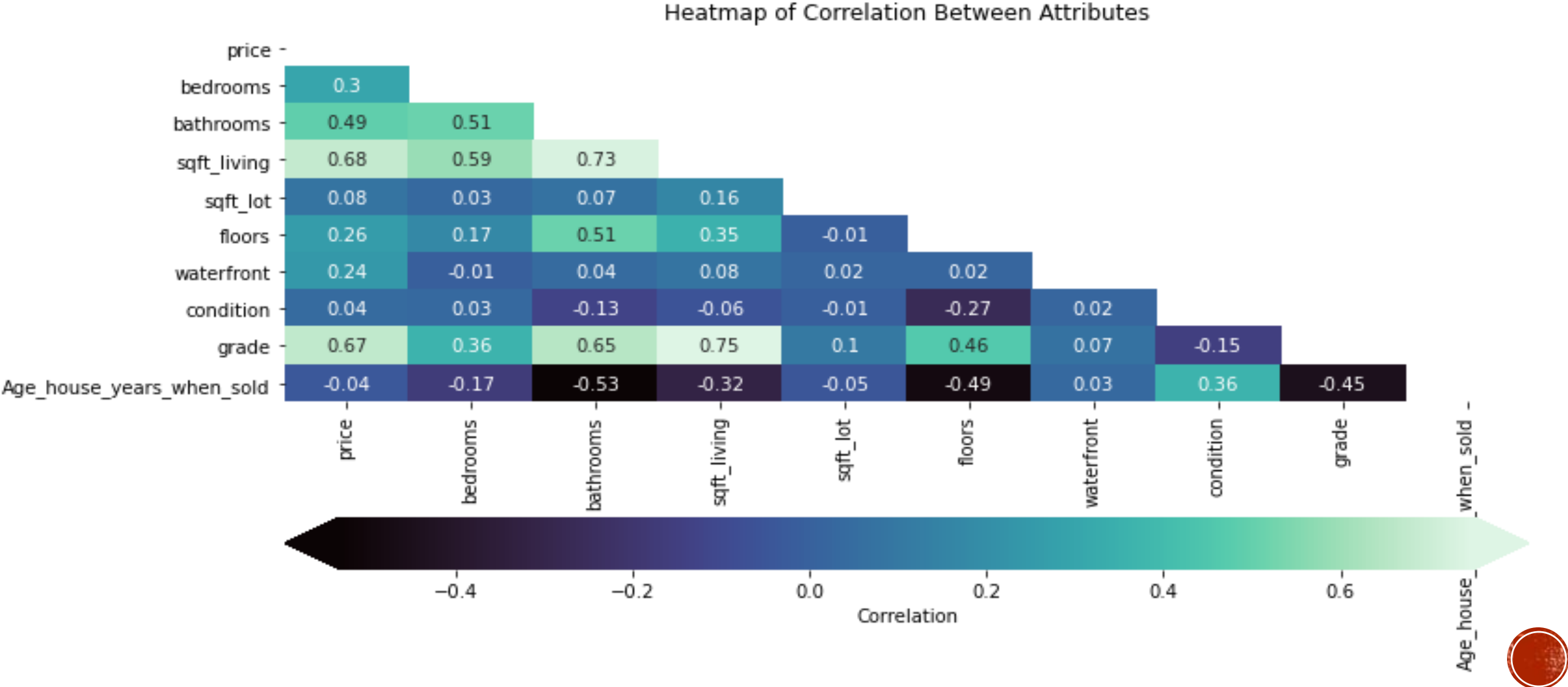
Data Understanding
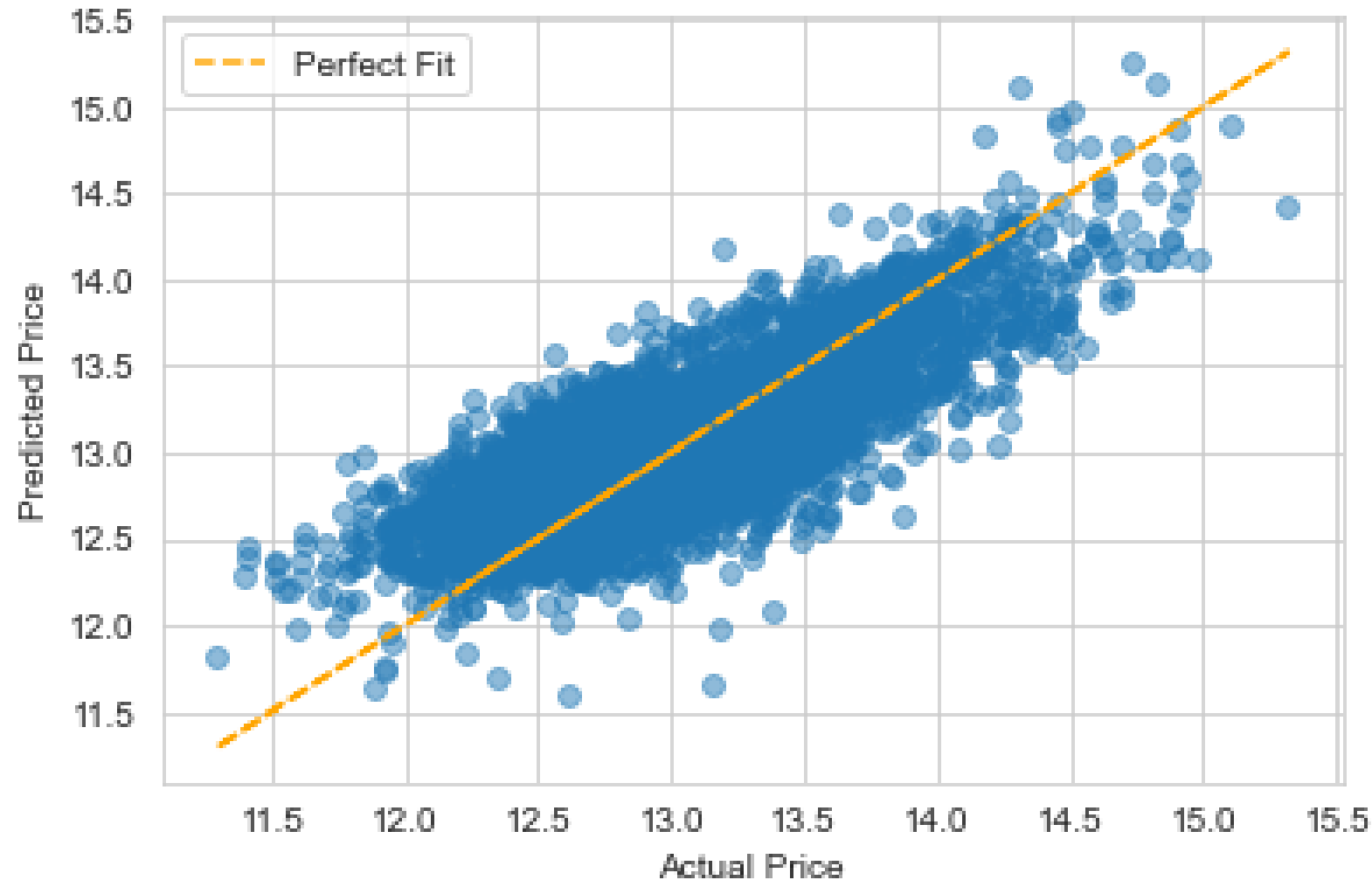
Data Preparation

Modelling

Model evaluation

Conclusion

RESULTS

Heatmap of Correlation Between Attributes

RESULTS

- All the linear model assumptions were met. That is;
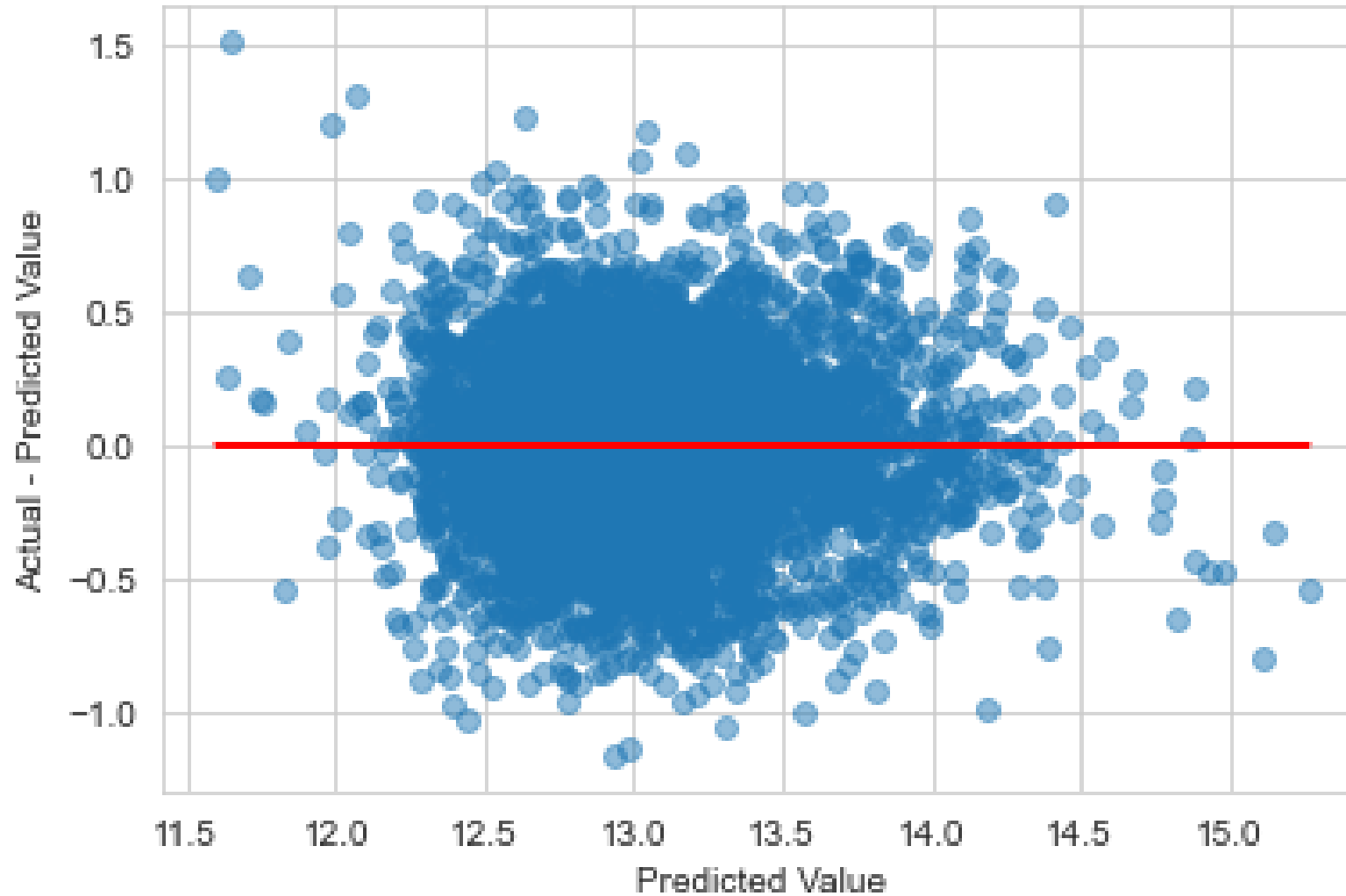
1. Linearity Assumption

RESULTS

2. Normality Assumption

Residuals QQ Plot

RESULTS

3. Homoscedasticity Assumption

4. Multicollinearity Assumption

CONCLUSION

Homeowners interested in selling their homes should focus on improving the design and quality of construction of their homes

Home buyers looking to buy a house should consider older homes with a high grade

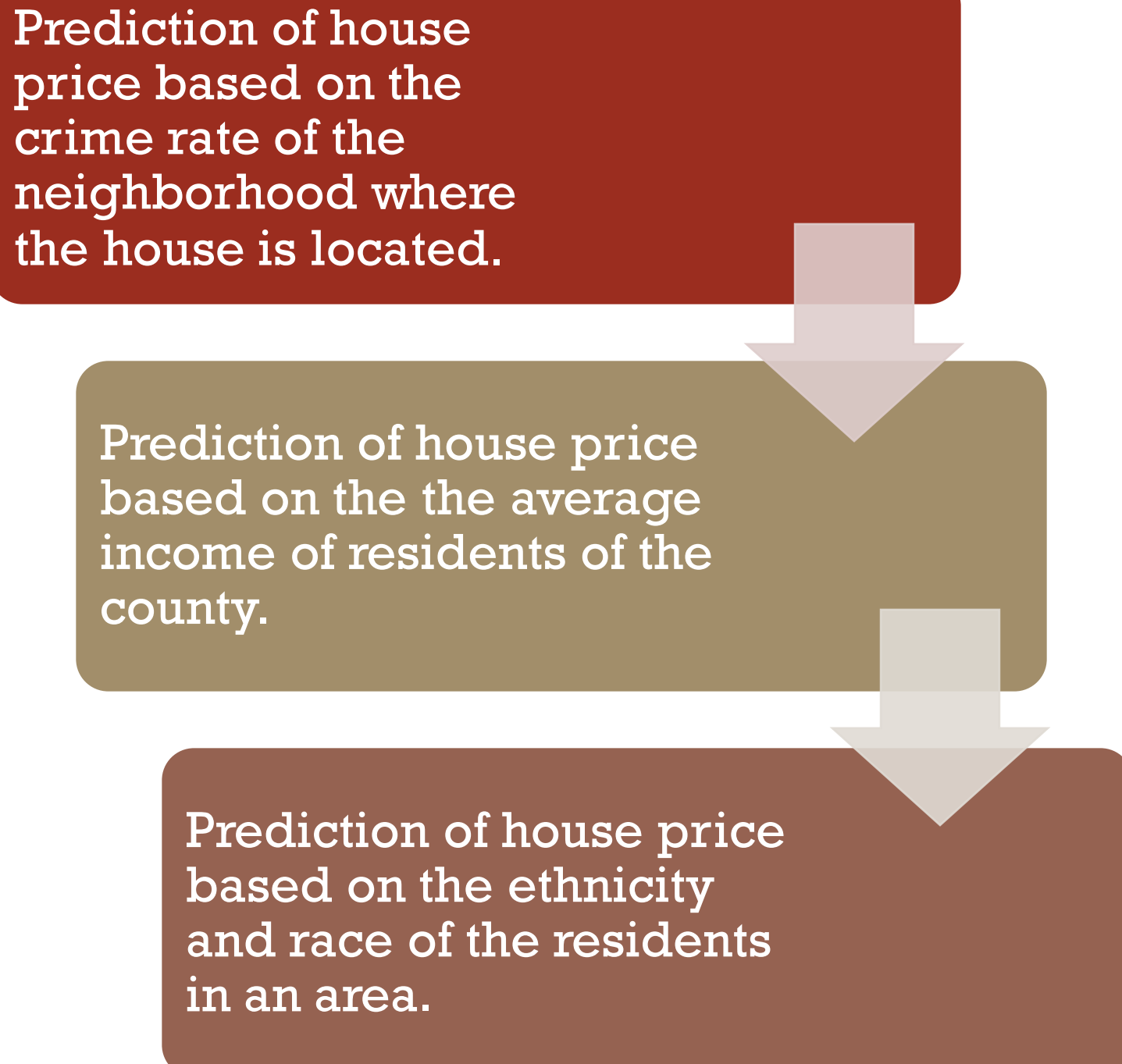If possible, homeowners should also expand the square footage of living space on the lot

LIMITATIONS

SOME OF THE VARIABLES NEEDED TO BE LOG-TRANSFORMED TO SATISFY REGRESSION ASSUMPTIONS HENCE ANY NEW DATA USED WITH THE MODEL WOULD HAVE TO UNDERGO SIMILAR PREPROCESSING.

GIVEN THAT OUTLIERS WERE REMOVED, THE MODEL MAY ALSO NOT ACCURATELY PREDICT EXTREME VALUES.

Prediction of house price based on the crime rate of the neighborhood where the house is located.

Prediction of house price based on the the average income of residents of the county.

Prediction of house price based on the ethnicity and race of the residents in an area.

NEXT STEPS

THANKYOU

- Email : fridahnkirotekimathi@gmail.com
- GitHub : Fridah Kimathi
- LinkedIn : Fridah Kimathi