KING COUNTY HOUSING

MULTIPLE LINEAR REGRESSION ANALYSIS

NURGUL KURBANALI KYZY

May 02, 2022

CONTENTS

01

BUSINESS PROBLEM

Stakeholders: **Real Estate Agency**

The business problem is to provide advice to homeowners in King County on *which aspects of the house features they should invest in* to increase the estimated values of their homes , *and by what amount?*

02

DATA PREPROCESSING

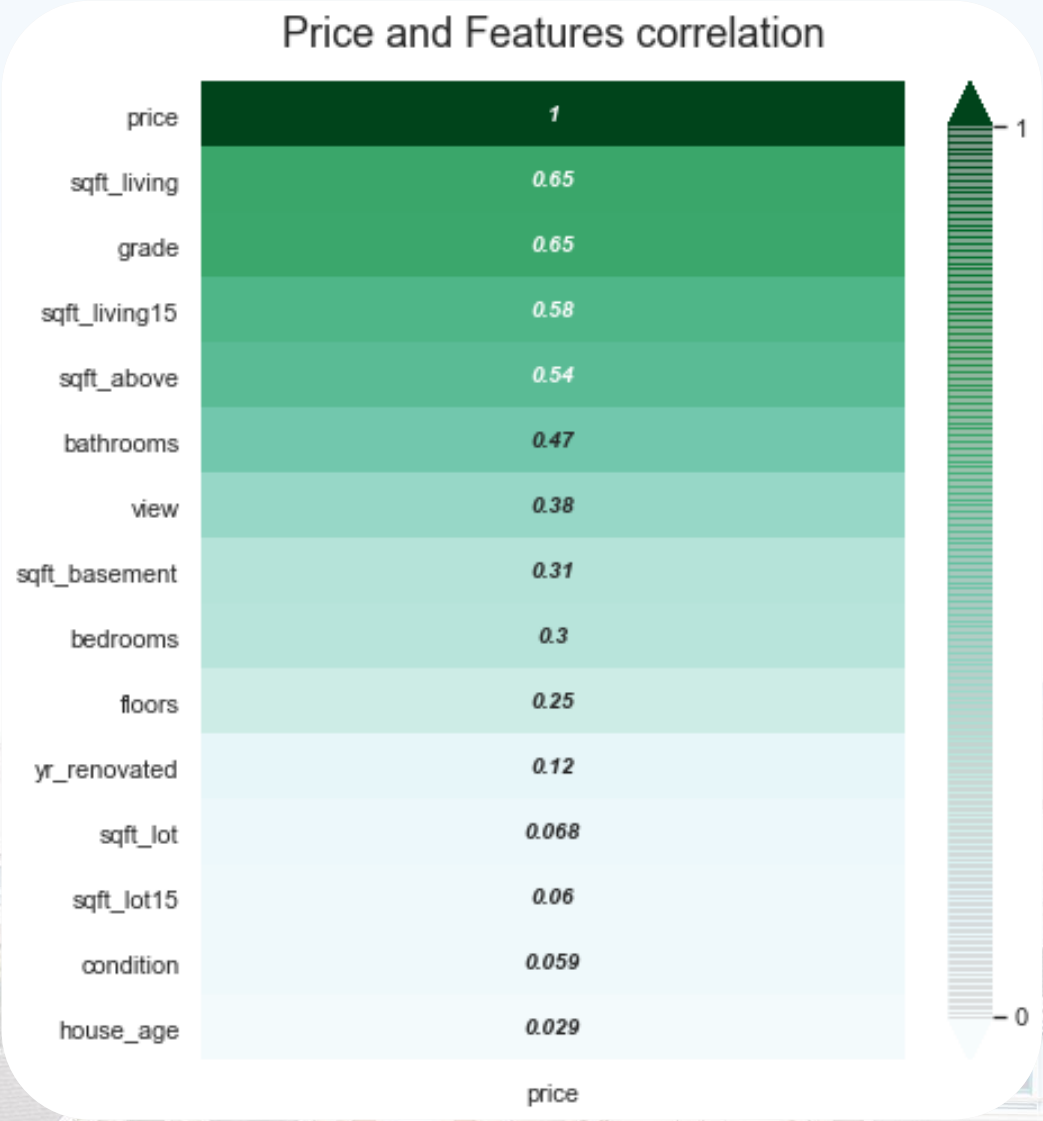
The data describes the sale of individual properties in King County from 2014 to 2015, comprises of 21 explanatory variables.

The methods included:

- Cleaning the dataset
- Removing nonimportant features
- Checking for the completeness of data (missing values)
- Removing Outliers

03

Multiple Linear Regression
MODEL BUILDING



03

RESULTS

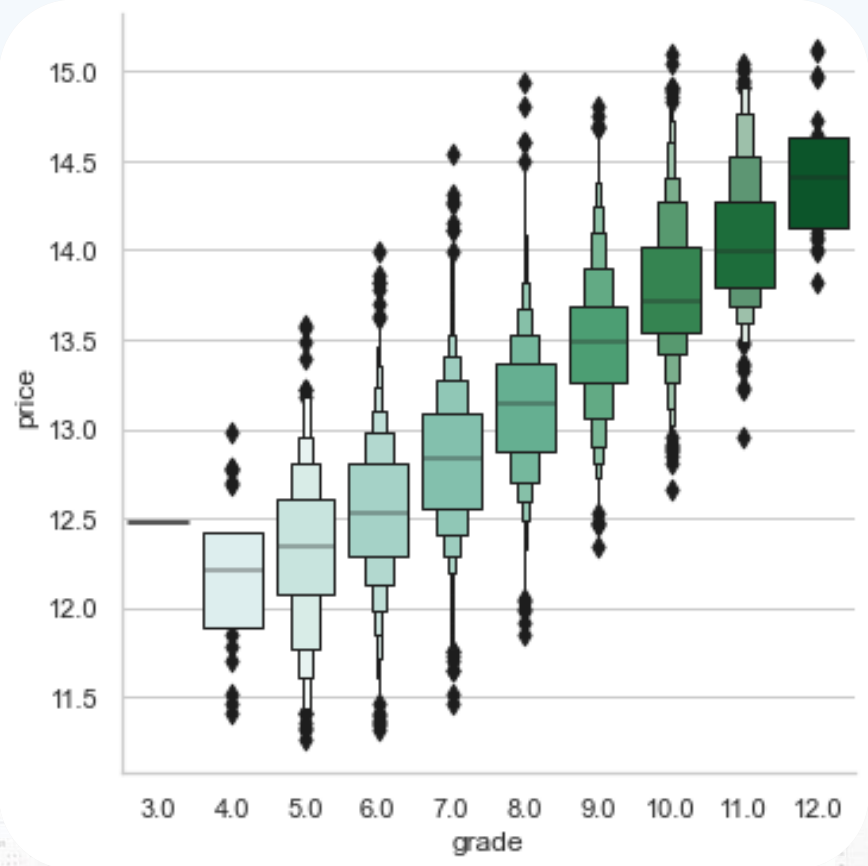
All assumptions of the multiple linear regression were checked, and an optimal final model was achieved by keeping the most influential features only. Base on the final model 54.6% of the variation in the price can be explained from the independent variables :

- *Grade*
- *View*
- *Square footage of the living area*
- *Square footage of the basement*

All feature variables have p-values less than 0.05, meaning all these features are statistically significant and there is no correlation between the target and the feature variables.

03

RESULTS /RECOMMENDATION

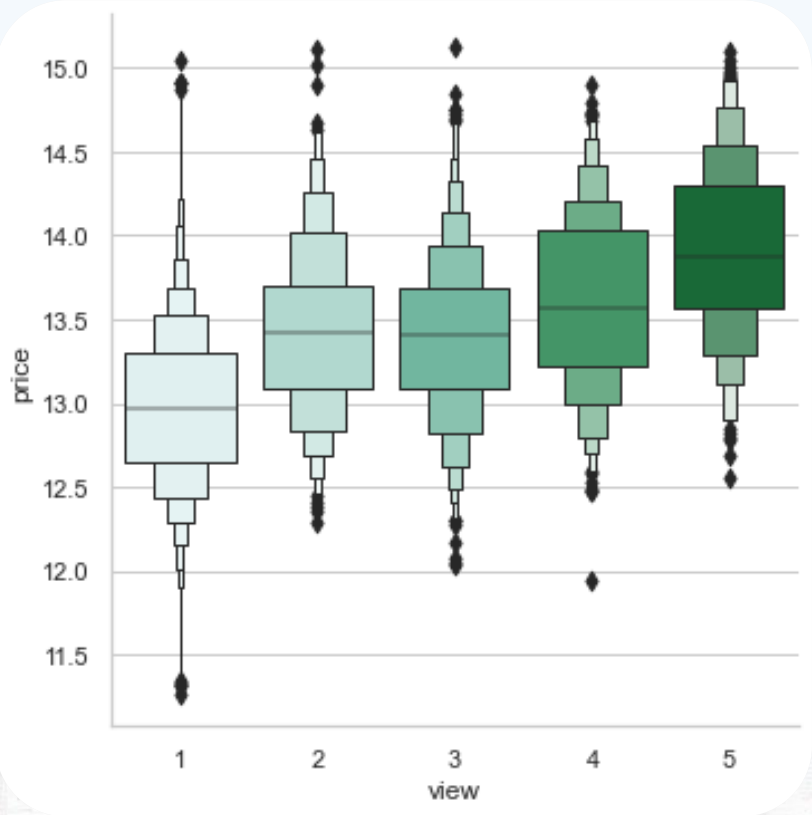


One unit change in <u>Grade</u> can increase house price at <u>18.77%</u>

Aim to get at least grade #8 by using better materials in both the exterior and interior finishes.

03

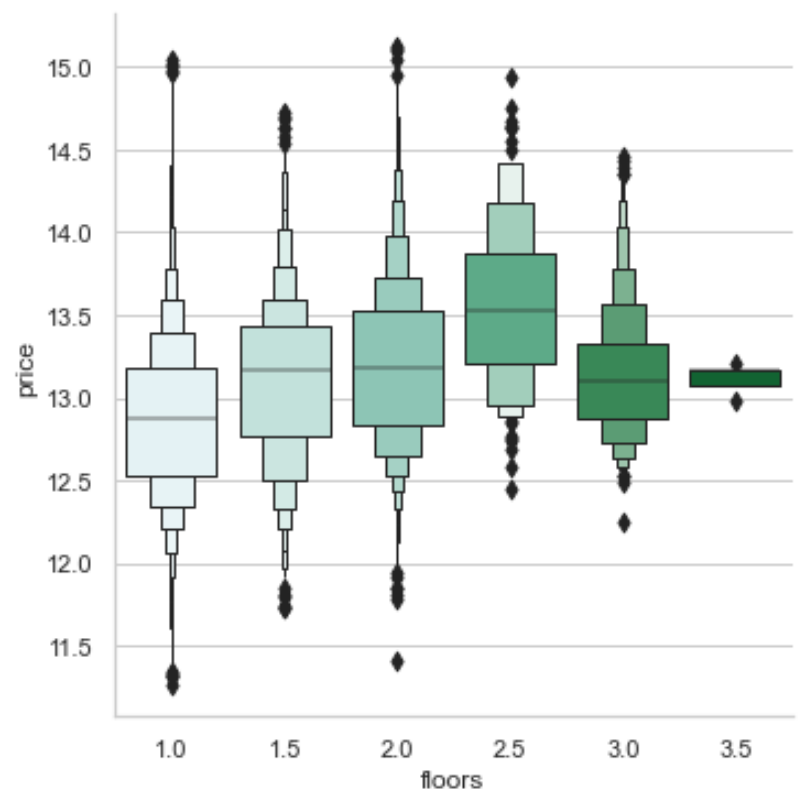
RESULTS /RECOMMENDATION



One unit change in <u>View</u> can increase house price at <u>9.20%</u>

Homes with a partial view may be a good investment because often trimming a few trees or building addition will turn the view from partial to full. The best vie w you can ask for is panoramic

RESULTS /RECOMMENDATION



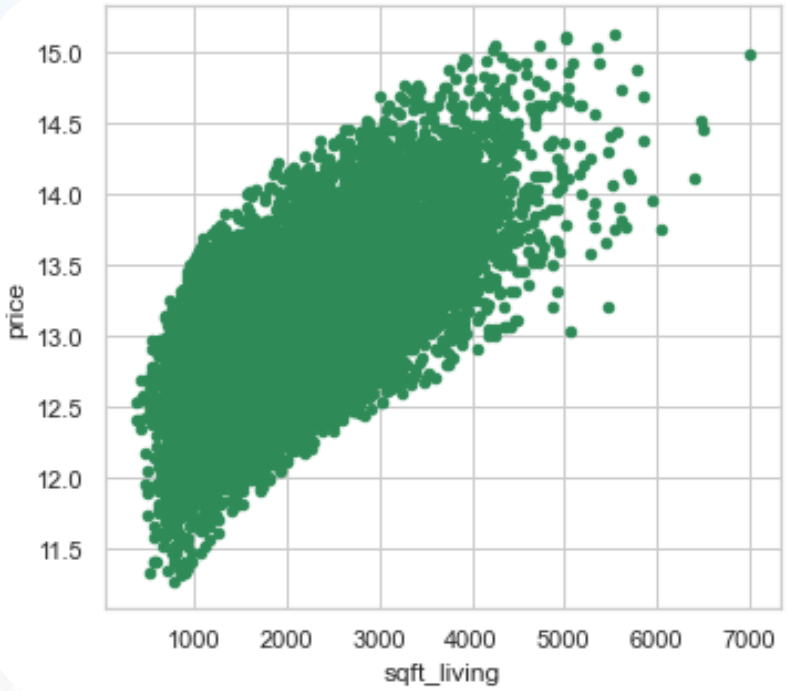
One unit change in floor(level) up to 2.5 increase house price at 4.64%

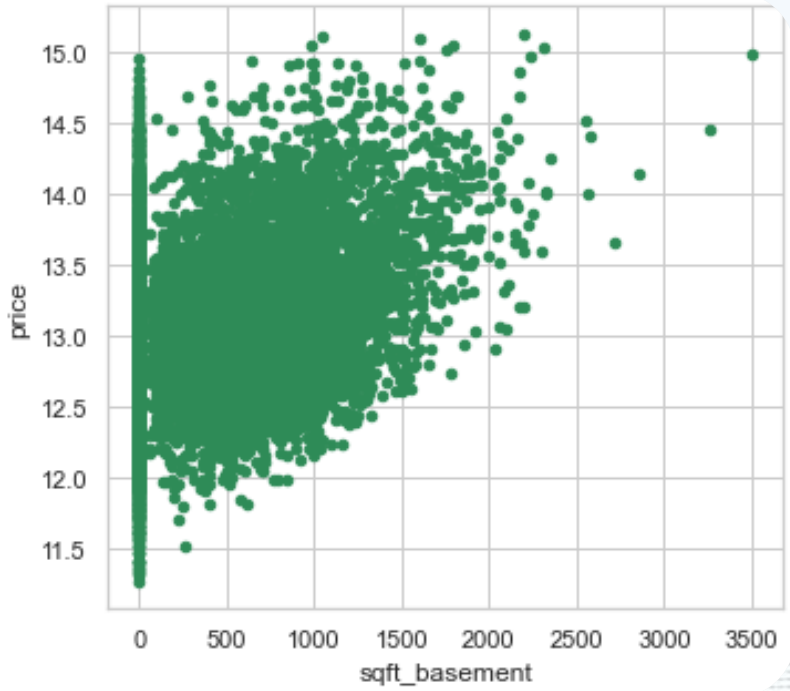
Most ideal and preferable house floor(level) is around 2.5

Houses with floors(levels) between 3-3.5 are not desirable since prices getting sharply decreasing.

03

RESULTS /RECOMMENDATION





One unit change in sqft_living can increase house price at 0.02 %

One unit change in sqft_basement can increase house price at 0.01%

Increasing the square footage of the living area along with the square footage of the basement will also tend positively effects the price increase

FUTURE CONSIDERATION

Based on the adjusted R-squared we got more than 45% of the variance in housing prices cannot be explained by the selected principal components. Given that some of the variables needed to be log-transformed and scaled to satisfy re gression assumptions. Other models like logistic regression, decision trees can be constructed and compared by results. Future analysis should include other predictors such house locations, demographics , security of a neighborhood etc.

THANK YOU