King County Real Estate

Housing Analysis

Summary

Descriptive analysis and modeling reveal which factors contribute most to housing prices:

- Living Area(in square feet)
- Location (47.55°N)
- House Quality(grade)

Business Problem

Kings County Real Estate has hired us to investigate which features of a home have the greatest effect on price.

- They would like us to make a model to predict housing prices.
- From that model, they would like to know which factors have the largest effect on price.

Data and Methods

Data:

King County House Sales Data from 2014 to 2015.

The dataset "kc_house_data.csv" was obtained from the link below.
https://osf.io/twq9p/

Methods:

Linear regression models via predominantly ordinary least squares (OLS) regression method and descriptive analysis

Results

Prediction Accuracy

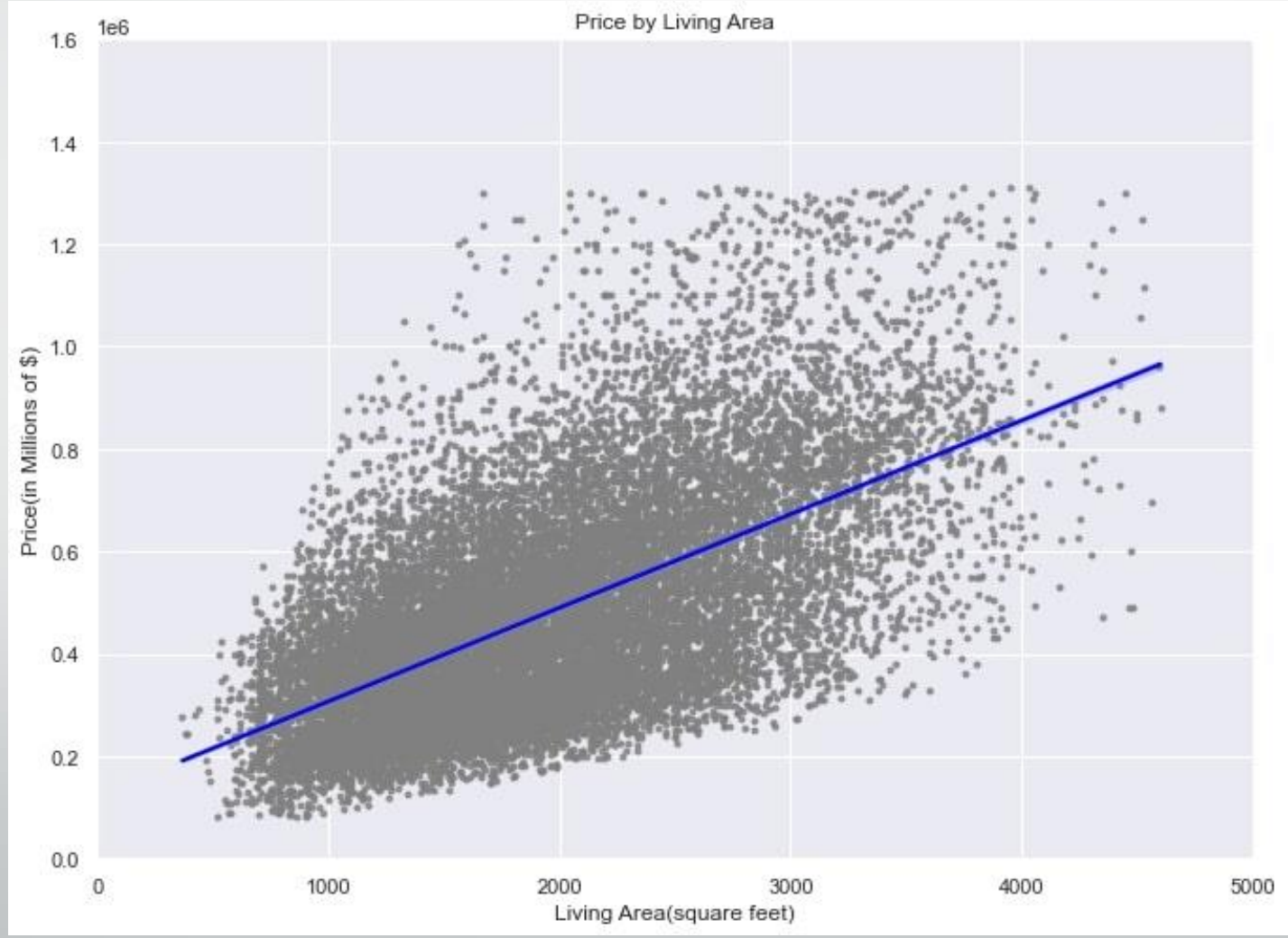
- Average home price: $476,027
- The price prediction was +/-$89,840 off the real price.

Margin of Error

18.9%

Results

Price increases with increased living area (sqft).



Results

Location: Most expensive homes are here.

Results

House Quality(grade):

Price in $Thousands USD

$800,000

$700,000

$600,000

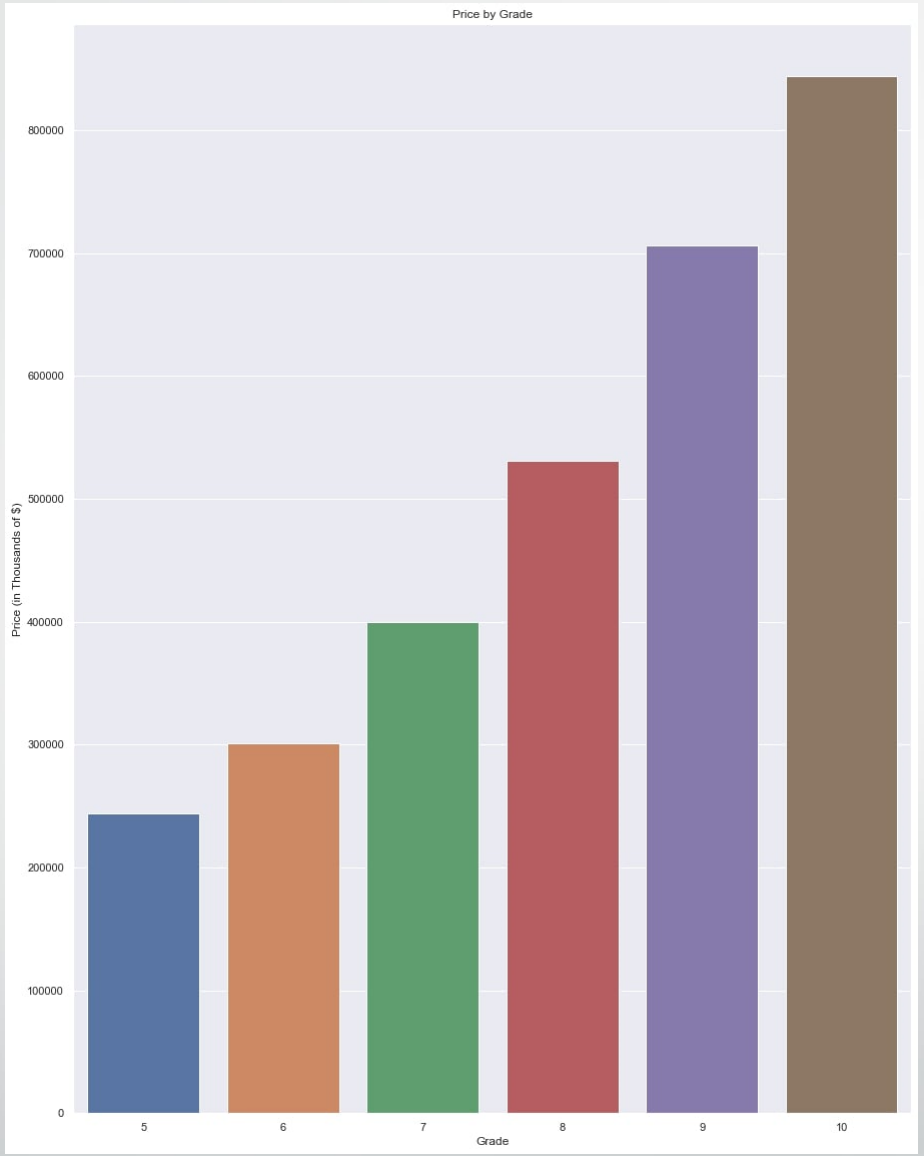
$500,000

$400,000

$300,000

$200,000

$100,000



5 6 7 8 9 10

Grade

Conclusion

Descriptive analysis and modeling reveal which factors contribute most to housing prices:

- Increase Living Area(in square feet)
- Flexible price point buy homes
- Upgrade the quality of your home:
  (curb appeal, efficiency, interior renovation, e.g. kitchen)

https://www.bankrate.com/loans/home-improvement/cheap-fixes-to-boost-the-value-of-your-home/

Future Research

- Outdated Data
- More from location data, incorporating zipcode
- Streamlining, better model
- New modeling techniques e.g. Polynomial Regression and Weighted Least Squares

Thank you!

Email: bennyzhu@buffalo.edu
Github: bigbenx3