

King County Housing

Data Analysis and Model

By: Mohammed Siddiqui
March 25, 2021



SUMMARY



Microsoft

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

- **Living Area(in square feet)**
- **Location**
- **Waterfront Property**
- **House Quality(grade)**

OUTLINE

- Business Problem
- Data and Methods
- Results
- Conclusions



BUSINESS PROBLEM

- Microsoft would like to to help new hires find housing.
- Create model to predict house prices
- Identify factors which most affect housing price

DATA & METHODS

Data

- King County housing sales data in 2014 and 2015
- Median household incomes by Zipcode

Methods

- Linear Regression

RESULTS

Prediction Accuracy

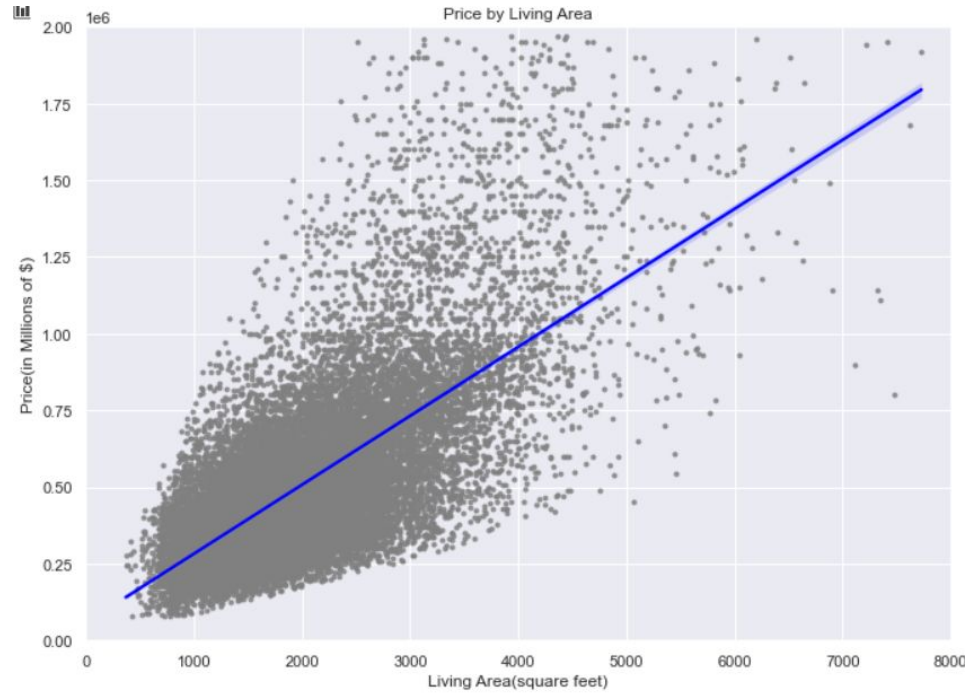
- On average, our predictions were off from the real price by +/- **\$81,500**
- The average home price in the area is **\$518,500**

Margin of Error

15.7%

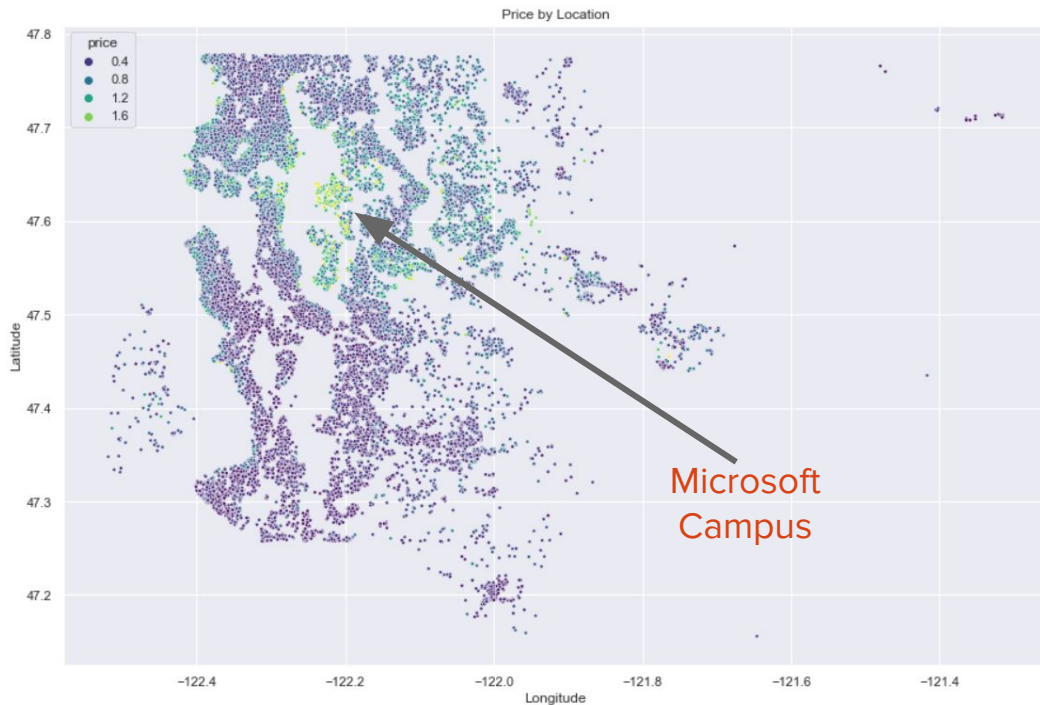
RESULTS

- Price increases proportionally with living area.



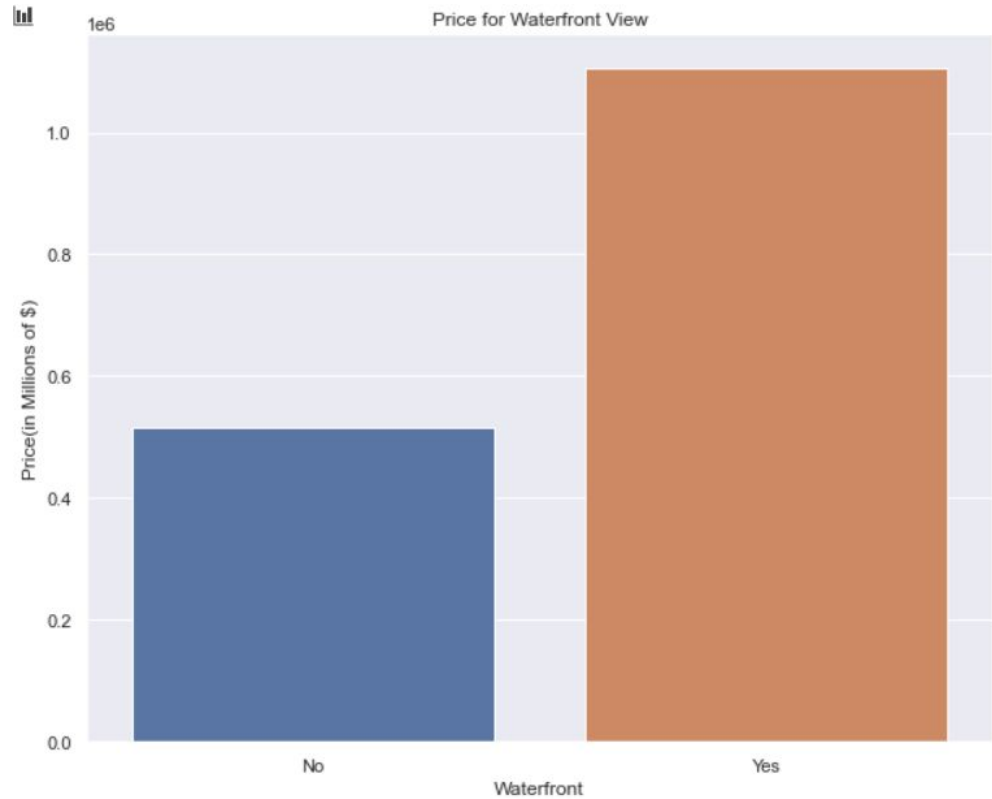
RESULTS

- Most expensive houses are in the same general area.
- Proximity to Microsoft campus plays a big factor



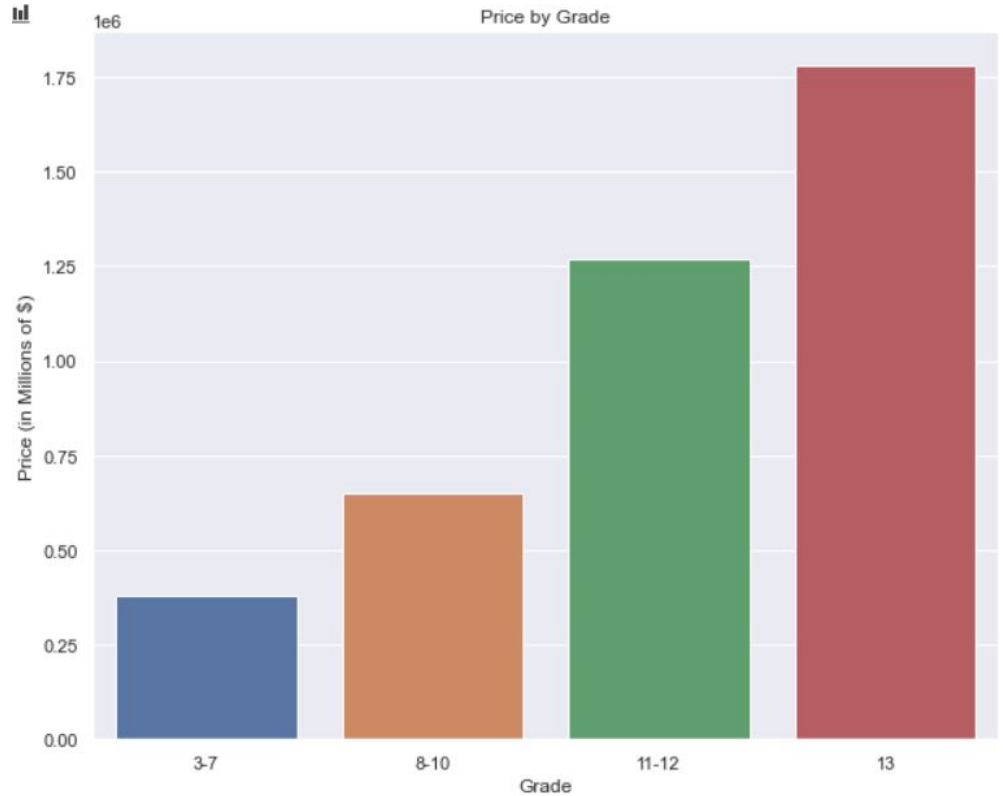
RESULTS

- Having a waterfront view has a substantial increase in price.



RESULTS

- Quality of the house(grade) has a significant impact on price.
- This is most apparent in the grades 11 and higher



CONCLUSION

FACTORS AFFECTING HOUSE PRICE

- Living Area(in square feet)
- Location(proximity to Microsoft Campus)
- Waterfront View
- House Quality(grade)

NEXT STEPS

- Find more recent data
- Better system for quantifying effect of location
- Use models more robust than Linear Regression

Thank You

Email: siddiqui@mohammed@gmail.com

GitHub: [@saifword](#)