Which factors influence a home's sales price?

Author: Samuel Rahwa
May 5, 2021

Summary

- I have been tasked to work for RE/MAX, a real estate agency that helps homeowners buy and/or sell homes.
- I need to provide advice to homeowners about which home renovations might increase the estimated value of their homes, and by what amount.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- Do the various predicting factors which have been chosen initially really affect a home's price?

- Which top factors might increase the estimated value of a home, for the buyer or the seller, and by what amount?

- Which top factors might decrease the estimated value of a home, for the buyer or the seller, and by what amount?

Data

- This project uses the King County House Sales

- From the King County House Sales dataset, I used the following columns:
  - sqft_living
  - sqft_lot15
  - sqft_living15
  - bedrooms
  - bathrooms
  - floors
  - waterfront
  - view
  - condition
  - grade
  - age
  - year_sold
  - renovated
  - month_sold

Methods

- Data Exploration
- Data Cleaning
- Creating a subset dataframe to answer business questions
- Using statsmodels to run our Multiple Linear Regression
  - Logging continuous variables
  - Scaling non-continuous variables
- Checking the assumptions of our linear regression
- Creating Visualizations to confirm if the mathematically tests are in line with the "eye test"

Results

- We are able to account for about 60% of the discrepancy within the King County House Sales dataset by using the 14 columns mentioned previously
- We able to prove a relationship existed between Home Prices and the 14 columns
- We can visually provide strong evidence that these numbers truly did come from a normal distribution.
- We are able to visually see that the home price column had the same random disturbance amongst the other 14 columns
- We were able to see the observations were independent of each other

Conclusion

- The three factors that would increase price of a home:
  - 1.) The footage of the home
  - 2.) The square footage of interior housing living space for the nearest 15 neighbors
  - 3) The overall grade given to the housing unit, based on the King County grading system

- The top three factors that would decrease price of a home:
  - 1.) Not having a Renovation within the last 10 years, based off of the year 2021
  - 2.) Not having a Waterfront View
  - 3.) The square footage of interior housing living space for the nearest 15 neighbors

Questions to Consider After the Results

- What are some reasons why your analysis might not fully solve the business problem?
  - Multicollinearity affect the coefficients and p-values, but it does not influence the predictions, precision of the predictions, and the goodness-of-fit statistics. If your primary goal is to make predictions, and you don't need to understand the role of each independent variable, you don't need to reduce severe multicollinearity.

- What else could you do in the future to improve this project?
  - The use of an expanded dataset with more home features/engineered or our various datasets included from the surrounding counties.

Thank You!

Email: samuelaaronrahwa@email.com
GitHub: @SamuelRahwa
LinkedIn: www.linkedin.com/in/samuelrahwa/