DSC520 Final Project Step3

Anjani Bonda

March 4th 2022

Project: Impact of AirBnB on rental home prices in Columbus, OH

Introduction

Airbnb is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. Airbnb does not own any of the listed properties; instead, it profits by receiving commission from each booking. The company has been criticized for possibly driving up home rents and a direct correlation between listings and rental prices and creating nuisances for those living near leased properties.

I would like to address the impact of Columbus(OH) Airbnb rentals on the nearby rental prices using Data Science algorithm in predicting the rental prices of Columbus and understand the correlation between them.

Problem Statement

The company has been scrutinized for a direct correlation between increase in the number of its listings and increase in nearby rental-prices and creating nuisance for those living near leased properties.

Research Questions

1. How are the Airbnb rental prices for different neighborhoods in Columbus ?
2. What's the correlation between Airbnb prices and Columbus neighborhood rentals ?
3. What kind of homes are most rented on Airbnb ?
4. What are the average rental prices on Airbnb by neighborhood ?
5. What are the average rental prices by Columbus neighborhood ?
6. How much profit does Airbnb possibly make ?

How you addressess the problem statement

Data Research and collection:

I have used inside airbnb website to gather the data for AirBnb listings, affordable rental housing data and average rent Columbus neighborhooddata. Each row of the data represents property listed on Airbnb, their prices, neighborhood, zip codes, average rents in that neighborhood. The data is focused on Columbus city.

Data Preparation and Cleansing:

Identified there were some missing data in the datasets. I have removed those records which have NA. Also dropped some of the fields that are not used for the analysis. Merged the datasets for the analysis.

Analysis

EDA (Exploratory Data Analysis):

I did the correlation analysis between the variables and noted the strength and weakness of relationships. 1. I found that zip codes, neighborhood, price, average rent, property type are good predictors for the analysis. Once these predictors are finalized, I looked into R2, Adjusted R2 stats and p-value. 2. Visualized different aspects of data to gain knowledge. 3. Calculated betas for the predictors in the regression model. This showed me how the 1 standard deviation change in predictor will impact dependent (response) variable. 4. Calculated confidence intervals which estimates how the model are likely to be representative of the true population values. 5. Performed an analysis of variance on all models to compare their performance. 6. Calculated standardized residuals, the leverage, cooks distance and covariance ratios. 7. Check if the regression model is unbiased.

Implications to consumer

The implication of the research is that the prices of the housing rental have direct impact on Airbnb listing in that same neighborhood. The recommendation from research is that there should be federal rule on how much the housing prices should increase year by year. Airbnb should consider the neighborhood prices when deciding prices for their listings.

Limitations of the analysis

1. The research is limited to Columbus.
2. The research needs more sample size for accurate analysis.
3. The research datasets are gathered from only one source.
4. The research is limited based on neighborhood. There are other factors to be considered for much detailed analysis.
5. The research is done using linear regression. There is room for improvement by fitting other ML algorithms.

Concluding Remarks

The research helped in applying the concepts and knowledge gained from this course. This project provided hands-on experience on a real life case study. Methods defined in the course were very helpful in applying to this case study. Visualization techniques learned thru this course can be used for other research areas.