

The Prediction of Airbnb Price in NYC: the effect of COVID-19 in the Airbnb market

Nicolas Oriol Guerra (no227) Linna Yue (ly389) Sophie Zhao (sbz24)

1. Motivation and introduction

The housing market has been a rising problem due to the steady increase of housing costs[1] during the past decade. The problem of affordable housing is present in many cities in both America and Europe[2] and it has also made its way to the front seat of political debate in recent years. However, after the Covid-19 pandemic, reports[3] indicate that housing prices are decreasing and that the trend may be reversed in the coming months.

The aim of this project is to understand what factors may affect housing prices and how those factors have been altered by the pandemic. In order to do so, a database of Airbnb listings will be used comparing the same city at different points in time (pre and post pandemic).

Some questions that could help to investigate the problem are: Can we predict Airbnb prices before and after Covid-19? Is the model the same or have the factors changed importance? Have popular housing locations changed in that time?

Answering this problem could be important to identify underlying trends in renting which could be applied as a benchmark to the more general housing market. It could also point out important relationships between the different variables and how they affect price and provide a narrow overview of the effect that Covid-19 has had on the economy.

2. Presenting the data set

The Data source used is from Inside Airbnb, <http://insideairbnb.com/get-the-data.html>.

The Airbnb listing data of New York before and after Covid-19 lockdown has 16 columns and 46512 observations. It has information about the price, location, room type, number of reviews, availability_365, minimum_nights, the number of host listings, etc. We can use machine learning methods to predict the price with room types, reviews, location, etc. Airbnb has data for different cities and dates. Thus, we can compare the models before and after pandemic. We plan to use folium to map the listings in NYC colored by the level of price and popularity, and figure out if the popular housing locations have changed or not due to the pandemic.. Furthermore, NLP will be used to conduct sentiment and subjectivity analysis on the reviews. Then visualize the high frequent words in two word clouds to compare whether there is some difference between the reviews before and after Covid-19.

3. Hypothesis of results

Before testing our model, we predict that due to the state of the economy after the Covid-19 pandemic, the Airbnb prices compared to the prices before Covid-19. Since tourism will have declined after the pandemic, we also hypothesize that the average number of Airbnb rentals will have declined since before the pandemic. And because more people are practicing social distancing, it is also likely that there will be more rentals for private Airbnbs and less rentals shared Airbnbs.

[1] Median Sales Price of Houses Sold for the United States. Fred Economic Data – Federal Reserve Bank of St. Louis

[2] TSENKOVA, S., FRENCH, M. *Affordable Land and Housing in both Europe and North America*. UN Habitat

[3] WILLIAMS, D. *Where Does The U.S. Housing Market Go From Here? Five Experts Share Predictions For The Rest Of 2020*. Forbes