

Joy Gao

#### Airbnb listings in New York

The dataset of New York airbnb listings was gathered through Inside Airbnb, a publicly available website that has been thoroughly analyzed to provide credible information. This is a secondary source, because the data is gathered from taking information from the city's airbnb website, which likely gathered the data from individual airbnbs. The format of the data is structured by name, year, neighborhood, property type, price, and reviews. The independent variables are neighborhood, year, or property type and the dependent variables are price and reviews. It is not up to date because the years range from 2008 - 2015. This makes the data out of date because prices could have changed in recent years and the number of reviews could have increased therefore lowering or raising the rating. The intended purpose of the information is to inform people of the different types of airbnbs in New York so they can make decisions on where to stay based on their preferences. As a result, the target audience is non-experts. There appears to be no ideological slants because there is only the presentation of facts. I can assume the data is credible, however, the usage of this data is not entirely relevant because prices have possibly gone up or reviews of the airbnb have changed.

The dataset of L.A. airbnb listings was gathered through a data scientist's project. This makes the data set a tertiary source because the scientist received the information from Inside Airbnb. The information was gathered in 2020, therefore, the data is more useful because it includes recent pricing and reviews. The author is also a student researcher so the evaluation of authority is fair. While the author may not be an expert in the field, he used information from Inside Airbnb. This is a trusted organization that is independent and not endorsed by Airbnbs. The dataset includes categorical variables including the name of listing, neighborhood, and room type. The dataset also includes quantitative data such as price, which is the dependent variable. The intended audience is people who wish to visit and stay at an Airbnb in Los Angeles. I can assume that the data will be helpful for determining the best Airbnb option for people visiting Los Angeles.

The dataset of hotel booking demand was gathered from the article Hotel Booking Demand. This is a secondary source because the data was collected for a research paper from an original source. The evaluation of authority is good because the paper is a peer reviewed journal. The information was gathered in 2019 which makes the data relevant. The data format is structured by hotel name, number of bookings cancelled, arrival year, month, week, day, number of nights, and number of adults. The intended audience is people who wish to know when the most and least popular times to visit hotels are. I can assume that the data is credible because of this evaluation.