

NYC Airbnb Data Walkthrough

Anthony Chau

UCI Center for Statistical Consulting

2021/01/12 (updated: 2021-07-23)

Background

- `airbnb_nyc.csv` contains Airbnb listing data for New York City
- The original dataset can be found on [Tableau Public](#)
- The data has 30,478 rows and 13 columns (variables)
- The variables are a mix of qualitative and quantitative data
- We don't know if each listing is for the same length of stay.

Variable Descriptions

- `Host.Id`: ID number of the host
- `Host.Since`: Date since the host
- `Name`: Name of the listing
- `Neighborhood`: Borough name
- `Property Type`: Property type
- `Review.Scores.Rating..bin.`: Binned review score rating (0 - 100, in increments of 5)
- `Room.Type`: Room type
- `Zipcode`: Zipcode of listing
- `Beds`: Number of beds

Variable Descriptions

- `Number.of.Records`: Number of records per listing*
- `Number.Of.Reviews`: Number of reviews
- `Price`: Price for listing
- `Review.Scores.Rating`: Review score (0 - 100)

Goals

1. Explore data - show general process in R
2. Produce summary statistics for some of the variables
3. Visualize the data with common plots: barplots, histogram, boxplots