Title:
Reputation and Revenue in San Francisco Airbnb Superhosts.

Dataset:
I use the Inside Airbnb dataset for the San Francisco region (https://insideairbnb.com/get-the-data/), which provides detailed information about listings, reviews, and neighborhood boundaries.
Specifically, the project uses:

- **listings.csv.gz** – detailed listing attributes including revenue, occupancy, amenities, and host metadata
- **reviews.csv.gz** – review dates, reviewer information, and review comments
- **neighbourhoods.geojson** – polygon boundaries for San Francisco neighborhoods

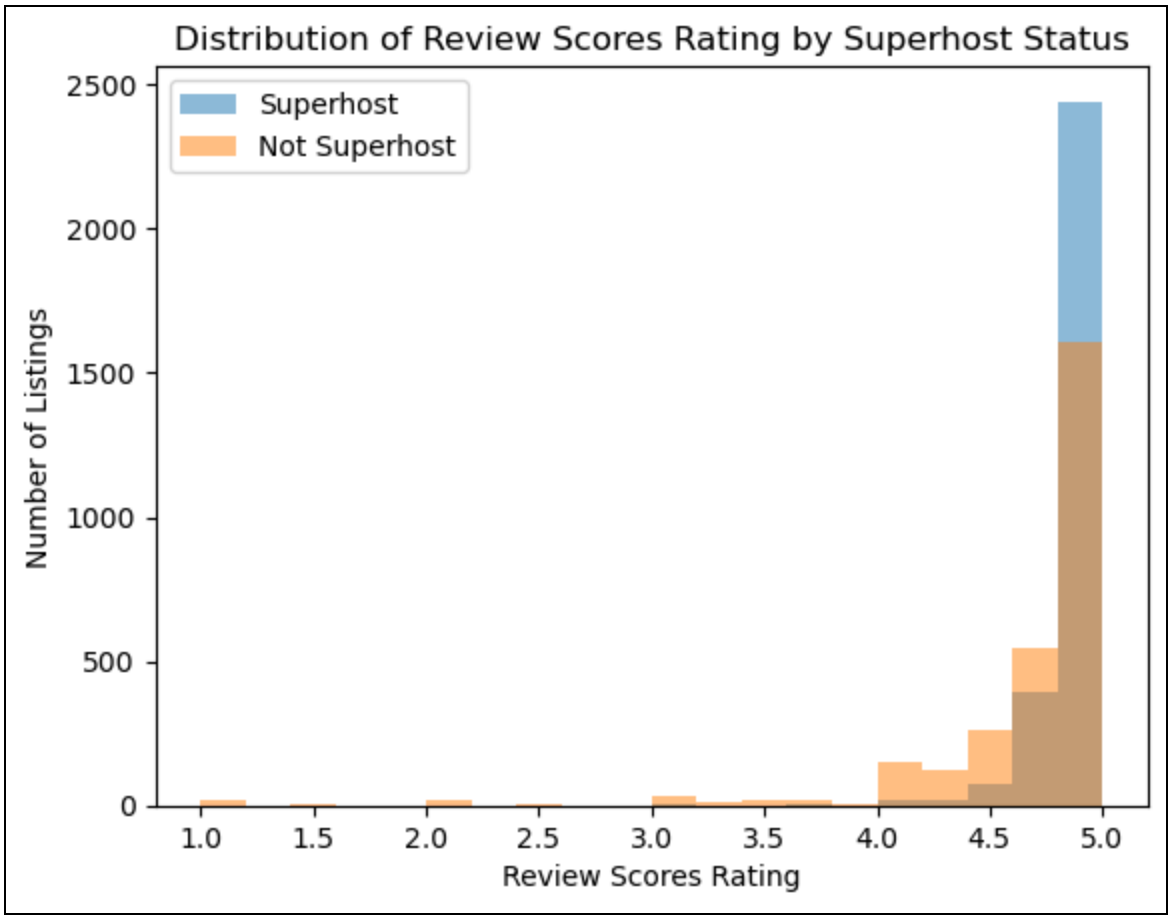
The dataset is publicly available to download.

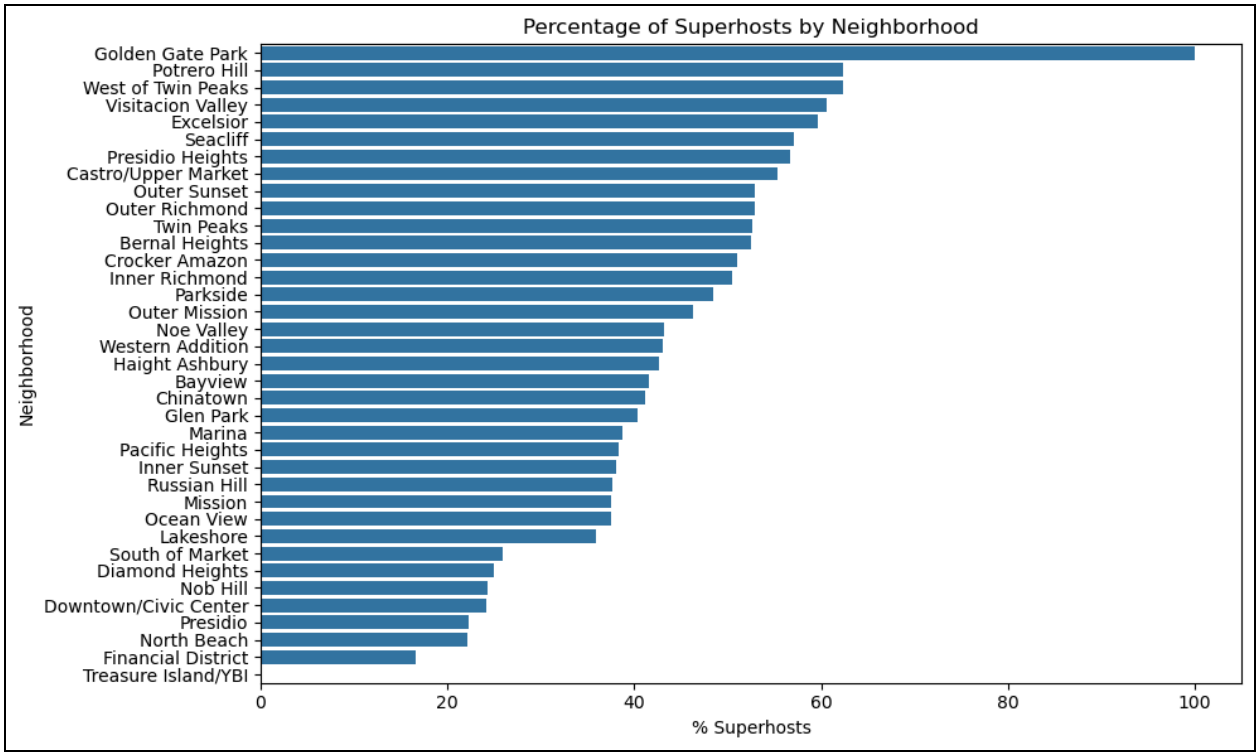
Brief writeup of what you intend to do (5-10 lines)
I would like to investigate the impact of host reputation (i.e., Superhost vs. Non-Superhost status) on the revenue generated by Airbnb listings in San Francisco. I will compare Superhosts and Non-Superhosts across multiple dimensions such as reviews, ratings, and revenue.

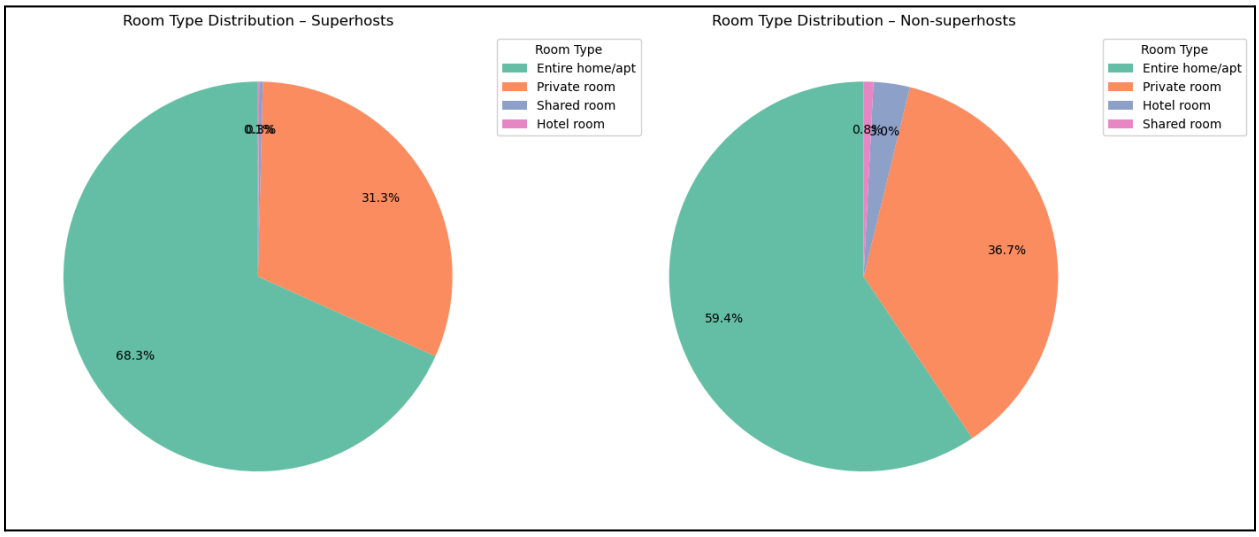
The main interactive component will be a San Francisco map where hovering over a neighborhood reveals a small dashboard (two pie charts for room type distributions, average review scores, average revenue, etc.). Clicking on each neighborhood will update two linked visualizations on the right side of the page: a histogram of review score distributions and a scatter plot showing the relationship between review scores and estimated revenue.

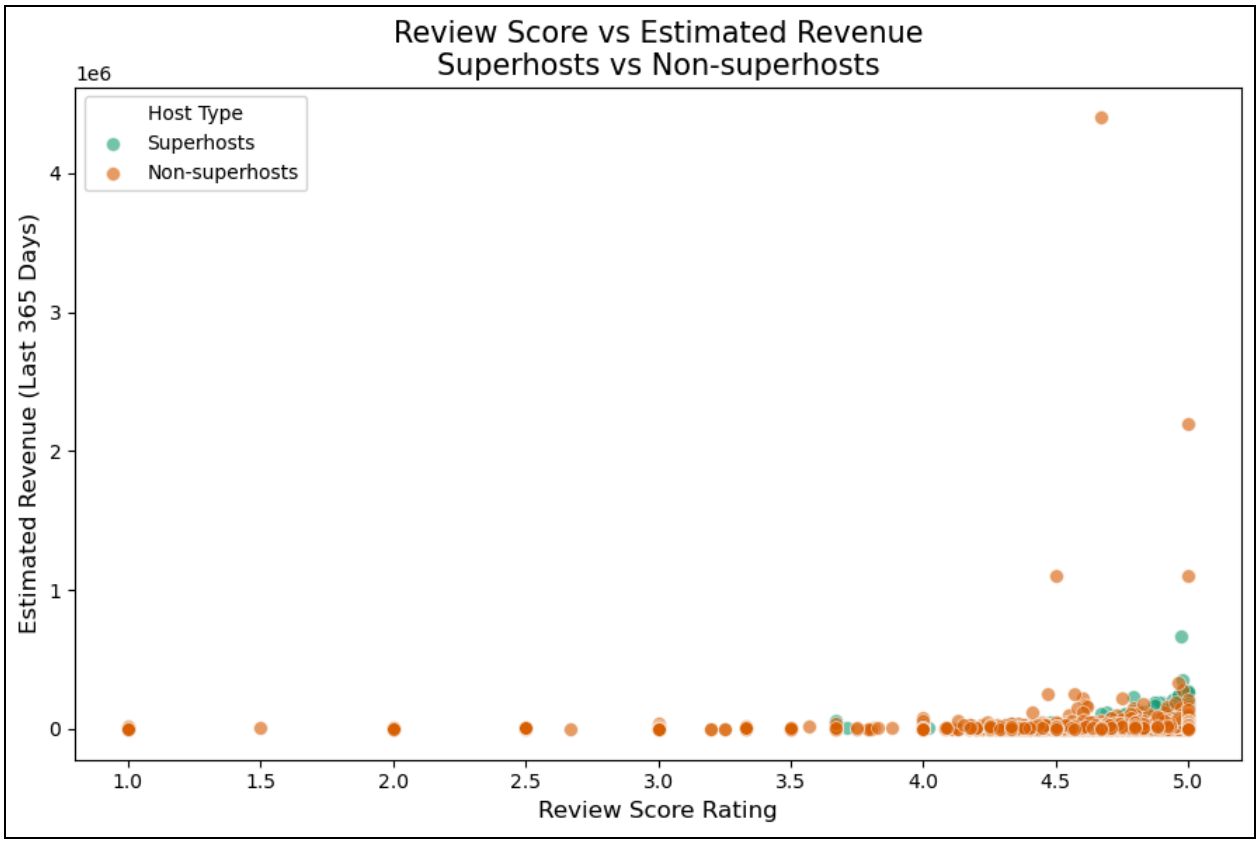
Users can also explore the graphs by selecting different sliders/options to see how room type, reputation, and price affect the reviews.

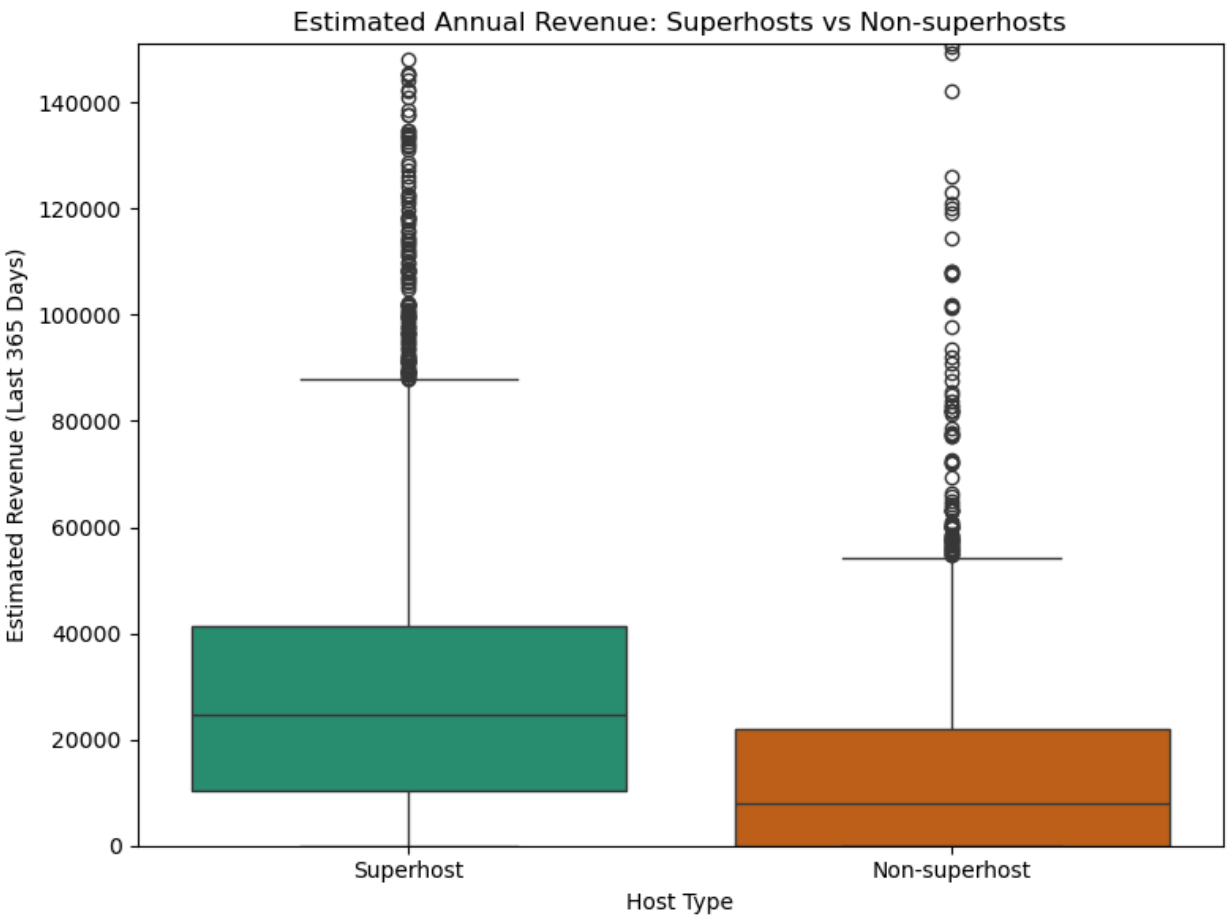
5-6 static visualizations on your dataset


Distribution of Review Scores Rating by Superhost Status
Superhost
Not Superhost
Number of Listings
2500
2000
1500
1000
500
0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Review Scores Rating


Percentage of Superhosts by Neighborhood
Golden Gate Park
Potrero Hill
West of Twin Peaks
Visitacion Valley
Excelsior
Seacliff
Presidio Heights
Castro/Upper Market
Outer Sunset
Outer Richmond
Twin Peaks
Bernal Heights
Crocker Amazon
Inner Richmond
Parkside
Outer Mission
Noe Valley
Western Addition
Haight Ashbury
Bayview
Chinatown
Glen Park
Marina
Pacific Heights
Inner Sunset
Russian Hill
Mission
Ocean View
Lakeshore
South of Market
Diamond Heights
Nob Hill
Downtown/Civic Center
Presidio
North Beach
Financial District
Treasure Island/YBI
Neighborhood
0
20
40
60
80
100
% Superhosts


Room Type Distribution – Superhosts
Room Type
Entire home/apt
Private room
Shared room
Hotel room
31.3%
68.3%
Room Type Distribution – Non-superhosts
Room Type
Entire home/apt
Private room
Hotel room
Shared room
36.7%
59.4%


Review Score vs Estimated Revenue
Superhosts vs Non-superhosts
1e6
Host Type
Superhosts
Non-superhosts
Estimated Revenue (Last 365 Days)
0
1
2
3
4
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Review Score Rating


Estimated Annual Revenue: Superhosts vs Non-superhosts
Estimated Revenue (Last 365 Days)
0
20000
40000
60000
80000
100000
120000
140000
Superhost
Non-superhost
Host Type