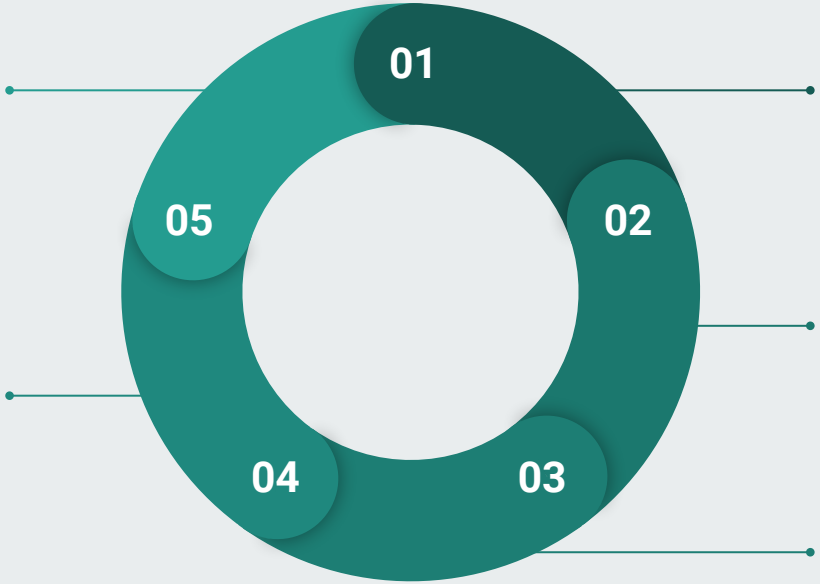
INLS 641 NYC Airbnb Visualization Project Presentation

Suchi Parekh, Yuan Wen, Cody(Yuxuan) Xu, Rui Hou

Presentation Outline

Conclusion

Live Demo



Introduction to the data and research questions

Our process to develop solution design

Our final design

Research Questions

- Which neighborhoods in NYC are popular for Airbnb service?
  Which neighborhood has more available rooms for listing?
- What are the average prices of these neighborhoods?
  Which areas are pricier compared to others?
- Which room types are booked more frequently?
  Do people have preference for short term or long term stay in NYC?

Tasks

- Price Trends
- Popularity Trends
- Listings Availability

Dataset from Kaggle

https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data

Process to Develop Solution Design

- Raw dataset eg. listing ID, host_name, neighbourhood_group, latitude…
- Pick up some of the fields, based on our objectives.

Process to Develop Solution Design

Objective 1: Price in all locations

- Fields: “*latitude*”, “*longitude*”, “*price*”, “*room_type*”, “*neighbourhood_group*”
- Visualization: **Scatter Plot**, color_coded range

Process to Develop Solution Design

Objective 2: Price in five neighborhood areas- Brooklyn, Manhattan, Queens, Staten Island, Bronx

- Fields: "*price*", "*neighbourhood_group*"
- Visualization: **Violin Plot**: show density and distribution of prices in the five neighborhood groups

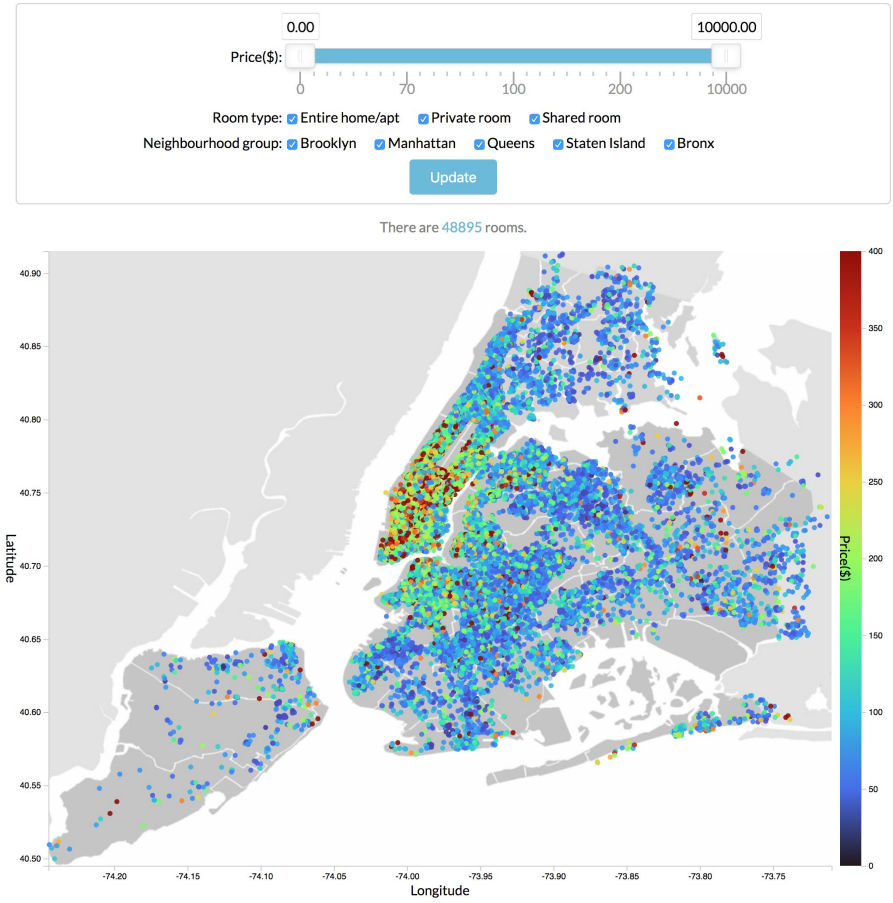
- Fields: "*minimum_nights*", "*id*"
- Visualization: **Histogram**: show number of short term/long term rooms

Process to Develop Solution Design

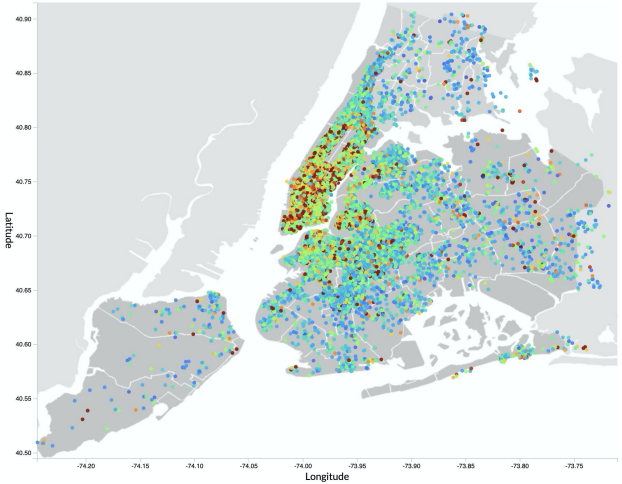
Objective 3: details for availability and popularity of the 200 neighborhoods

- Fields: “*availability_365*”, “*neighbourhood*”
- Visualization: **Bar Chart**: calculate average available days in a year for every neighborhood; show the number of rooms for every neighborhood.

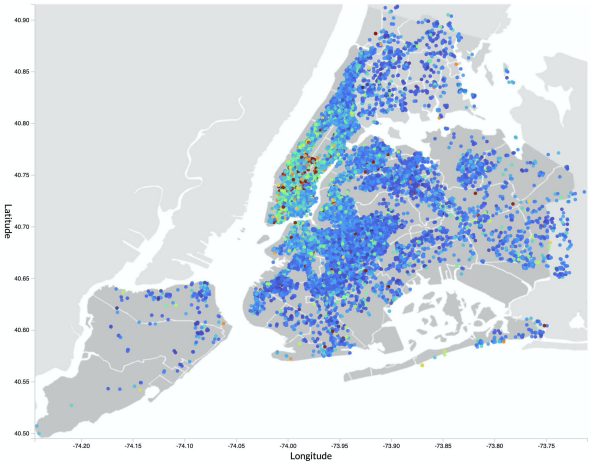
The visualization of NYC Airbnb



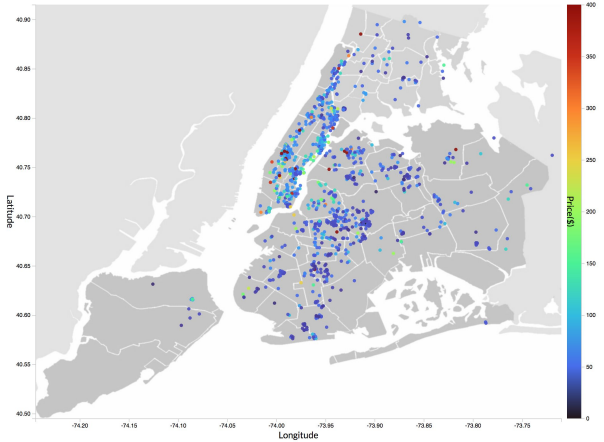
Entire Apartment vs. Private Room



Entire Apt: 25409 listings

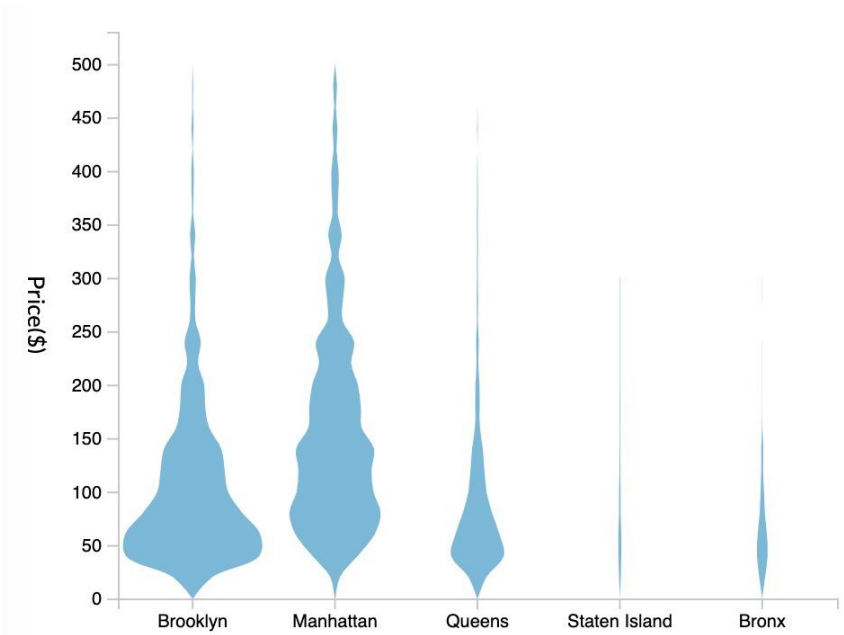


Private Room: 22326 listings

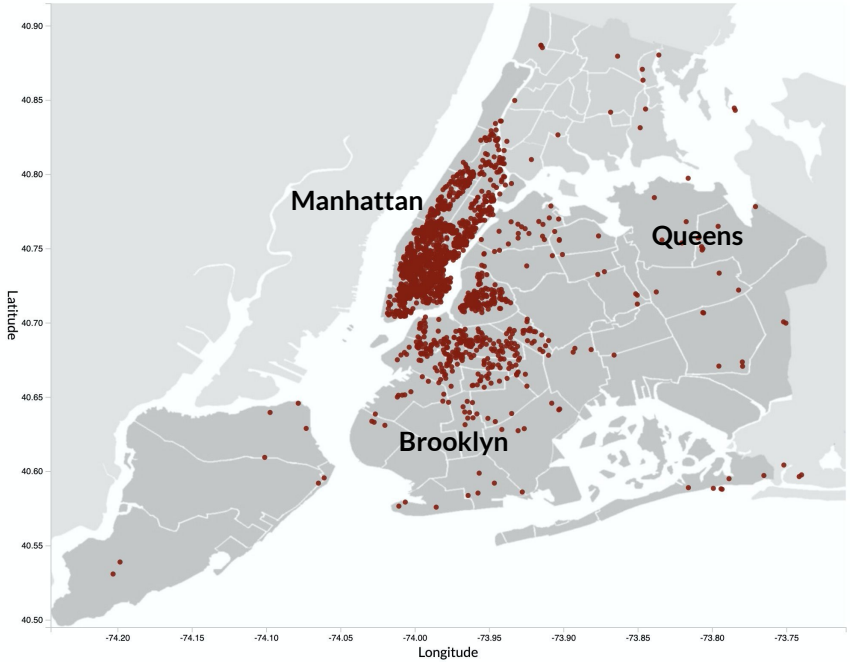


Shared Room: 1160 listings

Price Distribution



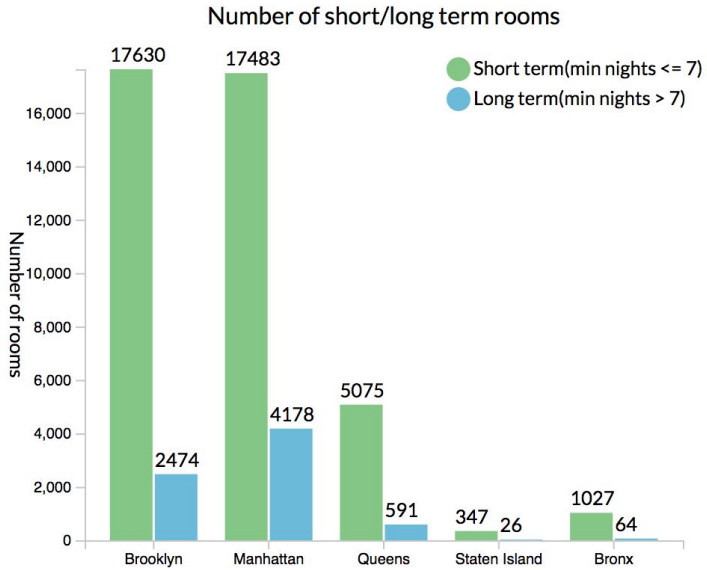
Violin plot of price distribution



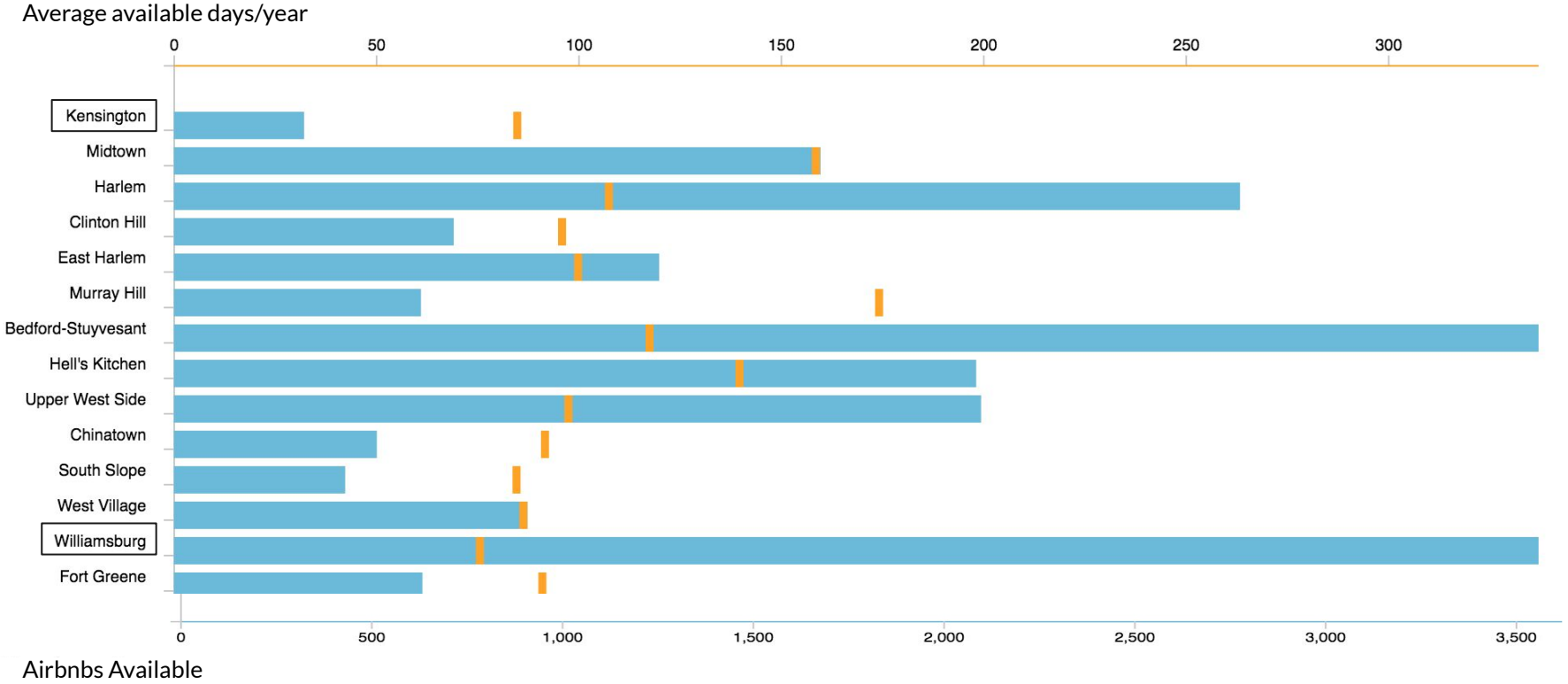
Listings of more than $400/night

Manhattan is pricey!!

Short and long term rooms by Neighbourhood



Neighbourhoods and their availability



Demo

https://github.com/imyuanwen/nyc-airbnb

https://imyuanwen.github.io/nyc-airbnb/

- We put our code, data and assets on a Github repository.
- Use Github pages to present our prototype.

Summary of Contribution

- Suchi
- Rui
- Eddie
- Cody

Conclusion & Lesson Learned

- D3 & SimpleStats
- Different visualizations for different data
  - (scatter plot, violin plot, bar chart)
- Usability

Thank You