CPSC 447 Milestone # 1

Basic Info

Project Title

- When In Rome

Team Members and Student Numbers

- Kristine Gonzales 40270654
- Michael Woolsey 87234621
- Ruchir Malik 19053065

Overview

With over 24,000 Airbnbs in Rome, it can be overwhelming to find an Airbnb to stay at when visiting the city. There are many factors to consider when selecting an Airbnb, such as the neighborhood, room type and price depending on your stay length. To help with this decision, we propose a visualization that will allow travelers to make smart accommodation decisions, both financially and logistically. Our users will be able to filter by stay-length, look up the number of different types of airbnbs in every neighborhood and pick the most suitable accommodation for them based on price and rating.

Data and Preprocessing

Dataset:

- Description: The original dataset (accessible by the link below) is a collection of all the airbnbs available online in Italy till September 2022. **However, we will be working with the Rome dataset specifically, which is a subset of the original dataset**. The Roma dataset has 24,783 airbnbs listings, where each listing has 75 attributes. For this project, we will be using 8 existing attributes and 1 derived attribute for our visualization. They have been summarized in the table below.

- The derived attribute (number_of_airbnbs) is the sum of the number of airbnbs of each type in a given neighborhood. For instance, if we are trying to calculate the number of airbnbs in 'XIII-Aurelia', it would be as follows:

   number_of_airbnbs = private room + shared room + hotel room + entire apt.

- Dataset link:
   https://www.kaggle.com/datasets/alessiocrisafulli/airbnb-italy

Attributes to visualize:

Attribute	Type	Cardinality	Range	Derived?
neighbourhood_cleansed	categorical	15		No
room_type	categorical	4		No
price	quantitative		0-91,000	No
minimum_nights	ordinal	19		No
maximum_nights	ordinal	166		No
review_scores_rating	quantitative		0-5	No
accommodates	ordinal	0-16		No
number_of_airbnbs	quantitative		thousands	Yes
listing_url	categorical	thousands		No

Tasks

1. Compare the number of airbnbs between neighborhoods
2. Summarize the types of airbnbs available in Rome (and within its neighborhoods)
3. Discover airbnbs within a certain stay length
4. Summarize the airbnbs by their rating and price

Visualizations

The must-have components are the scatterplot, bar chart and choropleth views, as well as, the stay-length filter and the tooltip that shows the number of airbnbs in a neighborhood. Furthermore, when a neighborhood is selected in the map, it should turn a different color.

We chose the scatterplot with rating on the x-axis and price on the y-axis because we want to be able to summarize the airbnbs by their rating and price.
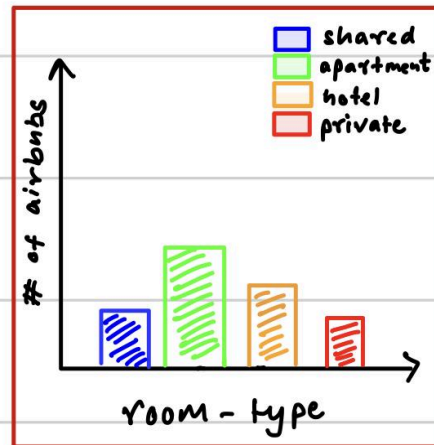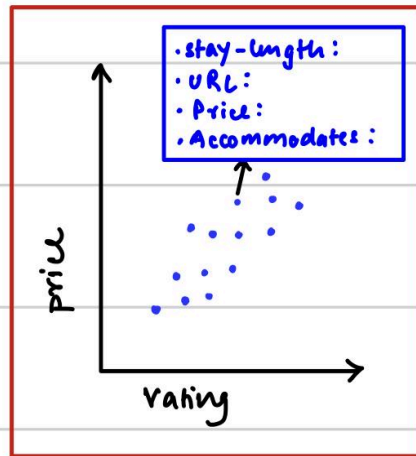
Furthermore, we have a bar chart that depicts the number of airbnbs per room type because we are interested in summarizing the types of airbnbs available in Rome and its neighborhoods. A categorical color palette will be used for each room type: Entire home/apt, private room, hotel room, and shared room. As for the choropleth map, each neighborhood is coloured depending on the number of airbnbs there and a sequential color channel will be used. This will also help with comparing the number of airbnbs between neighborhoods.

When a room type bar(s) is selected, the scatterplot will update to display only the airbnbs for that particular type. Multiple room type bars can be selected. If a bar is not selected, its color fades. When a point in the scatterplot is selected, a border will appear around the room type of that airbnb. Only one point can be selected.

When a neighborhood is selected, both the scatter plot and bar chart are filtered to show only airbnbs within that neighborhood. Only one neighborhood can be selected at a time. This will help summarize the types of airbnbs available in Rome (and within its neighborhoods).
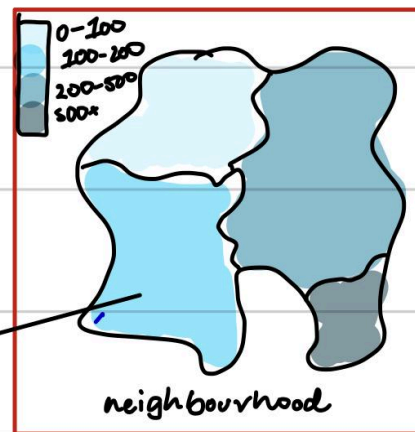
When users filter by stay length, all three views will update. This will help users discover airbnbs within a certain stay length.

The scatter plot tooltip will show information about the airbnb that the mouse hovers on. The information it will display is the stay length range, URL, price and number of people that it accommodates. The choropleth map tooltip will show the number of airbnbs in that neighborhood.

- stay-length:
- URL:
- Price:
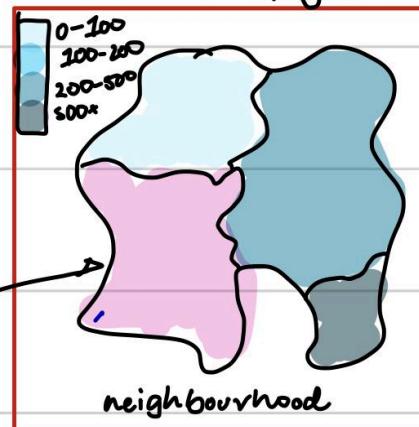- Accommodates:

price

rating

# of airbnbs

shared
apartment
hotel
private

room - type

0-100
100-200
200-500
500+

stay - length ▽

#airbnb:

neighbourhood

↓ click neighbourhood

0-100
100-200
200-500
500+

turns pink when selected

neighbourhood

Usage Scenario

George is a current CPSC 447 student who's in his last term at UBC. He wants to travel to Europe after graduation because he's jobless and wants to explore the world before submitting the rest of his life to corporate slavery. But George is smart and likes to plan in advance. He is looking for an airbnb for his stay in Rome. He opens our application and first clicks on one of the 15 neighborhoods that he would prefer staying in. As soon as he does that, our bar chart displays the number of airbnbs of each type in that neighborhood and our scatter plot shows all airbnbs plotted based on their price and rating. He can also choose to hover over a point (i.e. an airbnb) to find out more information (for instance, URL of the airbnb, the number of people it can accommodate, about that accommodation. And if he chooses to click on one of the points on the scatterplot, a border would be added to the bar on the barchart that that airbnb belongs to. Moreover, based on the type(s) of airbnb George prefers, he can click on the bar(s) and that will filter airbnbs on our scatterplot.

Team Communication Plan

Include a description of your plan for how you will communicate during the duration of the project. Describe your team meeting strategy in terms of when, how often, and how/where you will meet (whether online or in person). Also describe how you will asynchronously communicate outside of meetings (which platform you will use for messaging) and what the expected response time is for those messages. Finally, explain the steps you will take should any of your team members stop complying with the agreed-on communication plan

When: Thursdays or Fridays

How Often: 1x/week

How we meet: in person

Where you will meet: at UBC

Asynchronous communication
- Platform: Instagram
- Expected response time: 24 hours

Stop complying with agreed-on communication plan
- Remind team member about the plan
    - E.g., "Hey, it's been 24hours, could you please reply to the message in the group chat"
- Talk to the team member in person about it
- After the first two steps don't work, ask a TA for their opinion

Work breakdown and schedule

ToDo	Assigned to	# number of hours it will take	Target completion date
create initial static version of scatter plot	Kristine	4	March 22nd
create initial static version of bar chart	Ruchir	4	March 22nd
create initial static version of choropleth map	Michael	5	March 22nd
Link scatter plot to bar chart	Kristine	10	March 29th
Link bar chart to scatter plot	Ruchir	10	March 29th
Link choropleth map to bar chart	Michael	10	March 29th
Link choropleth map to scatter plot	Kristine	10	April 5th
Create stay length filter	Ruchir	3	March 29th
Tooltip for scatter	Kristine	1	April 5th

plot			
Tooltip for choropleth map	Michael	1	April 5th
Optional: if scatter plot point is selected, show its position in choropleth map (will need lat & long. attributes)		5	
Total hours:	58 + (5) = 63		