

Exploring Rio de Janeiro Touristic Neighborhoods

IBM DataScience Capstone



Some information about Rio

- Is one of the most visited cities in the Southern Hemisphere
- Second largest municipal GDP in Brazil
- Is known for its natural settings, Carnival, samba, bossa nova, and beaches
- Population 6.7M

Here is the question

Where should I stay in Rio?



Data acquisition and sources

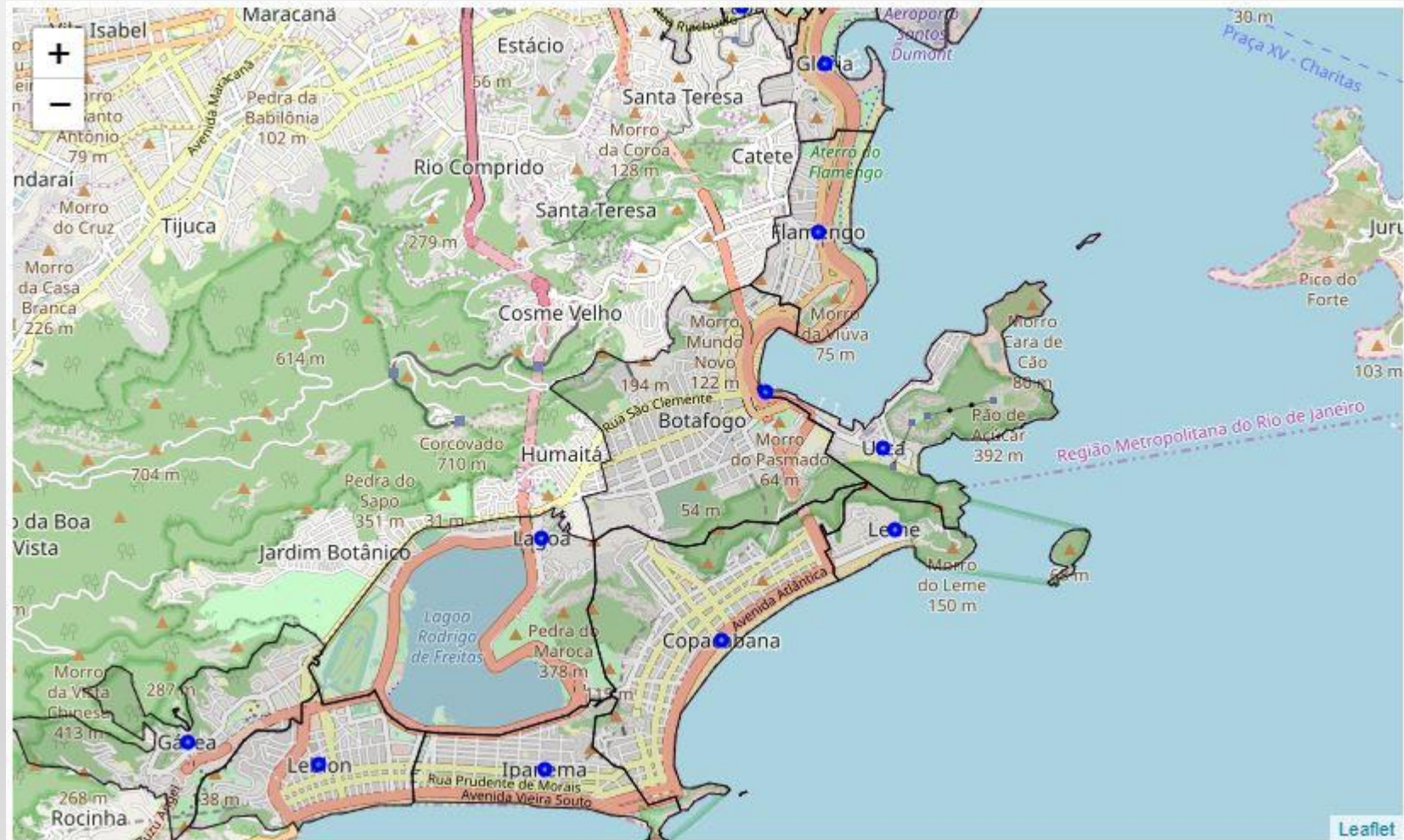
- Districts of Rio with their geo data
 - Source: <http://www.data.rio/>
- Venues data in each district. Manage more than 1000 venues in 30 categories
 - Source: Foursquare API

Methodology

- Loaded the districts geo data
- Using Foursquare API, got all the needed venues information
- Explore data analysis
- Create clusters of districts based on venues category
- Visualize cluster data pointing out their similarities and differences

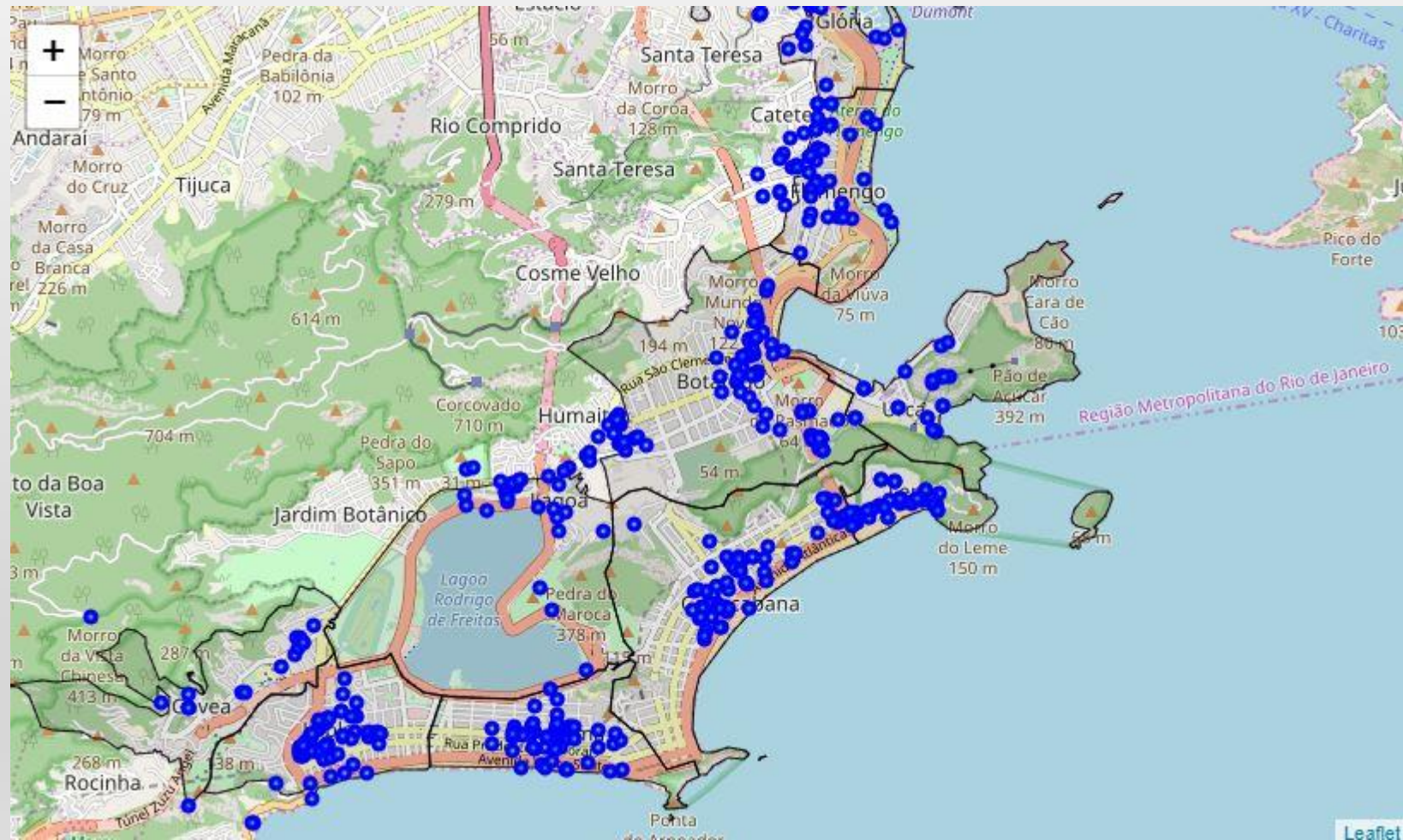
Exploratory Data Analysis

Neighborhoods in Rio



Exploratory Data Analysis

Venues in Rio



Exploratory Data Analysis

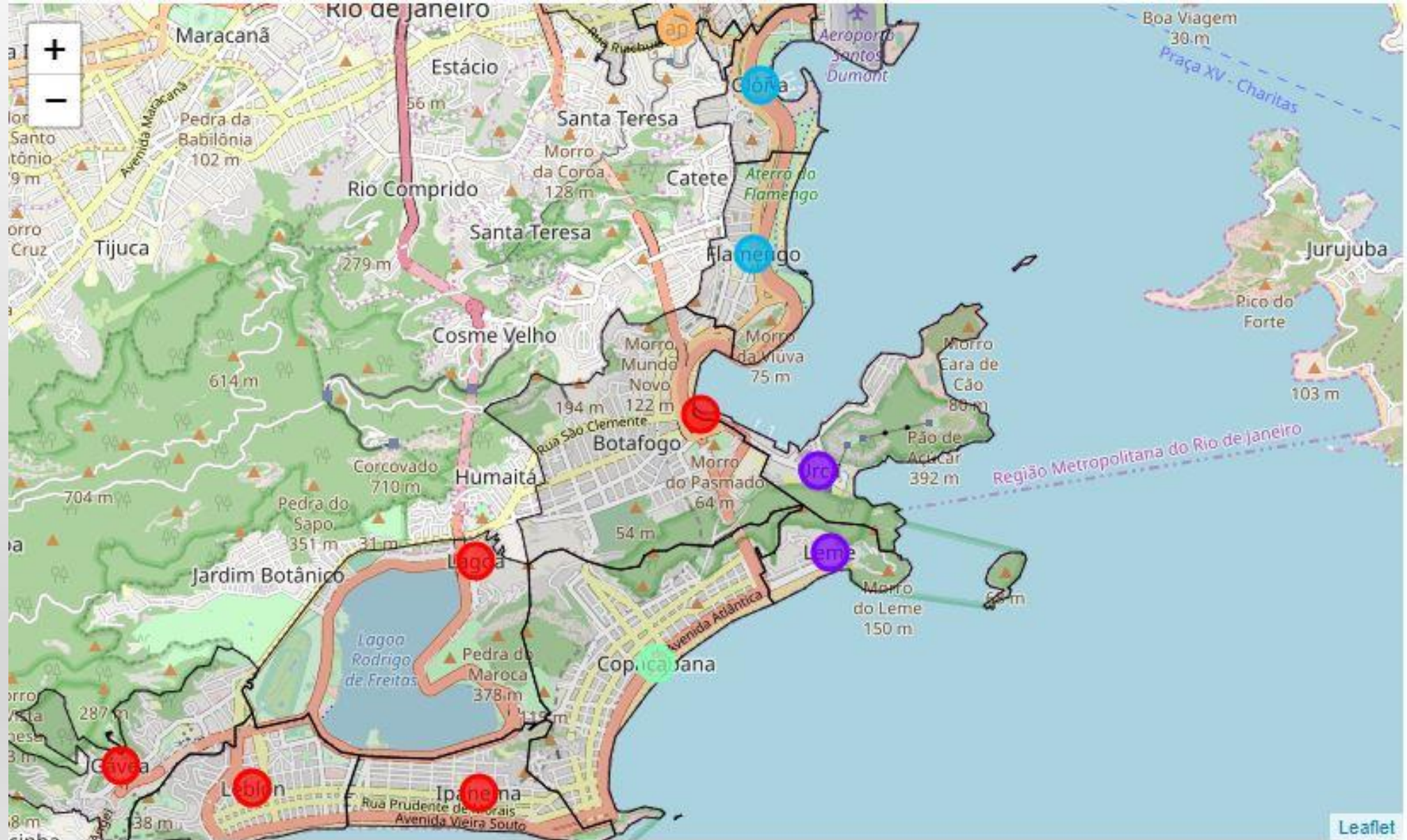
Most popular venue categories

Venue Category	
Brazilian Restaurant	41
Bar	36
Hotel	25
Coffee Shop	17
Beach	17
Pizza Place	16
Gym / Fitness Center	14
Japanese Restaurant	14
Bookstore	13
Scenic Lookout	12

Clustering

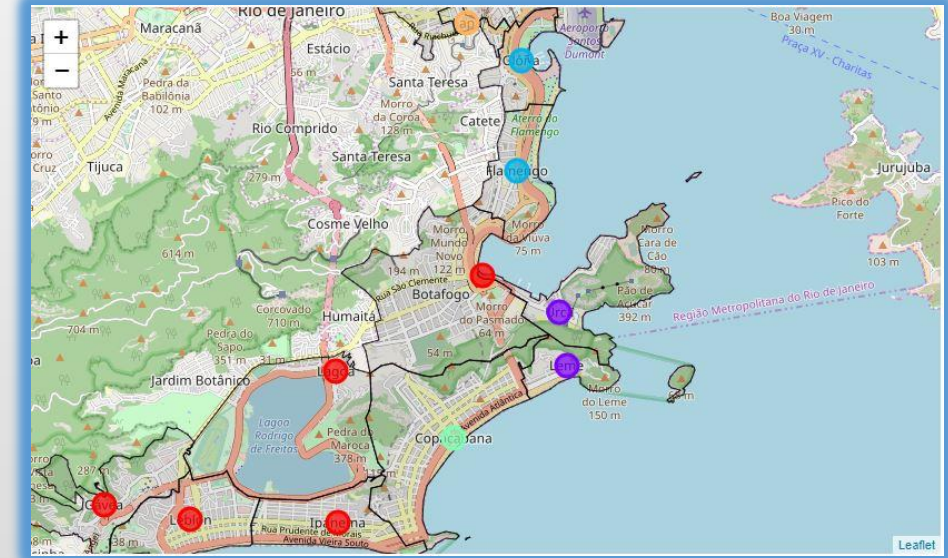
- Based on neighborhood geo data and Foursquare venue information, I created top 10 venues types for K-Means input.
- K-Means clustering with 5 clusters to understand what each neighborhood has to offer.

Clustering on map



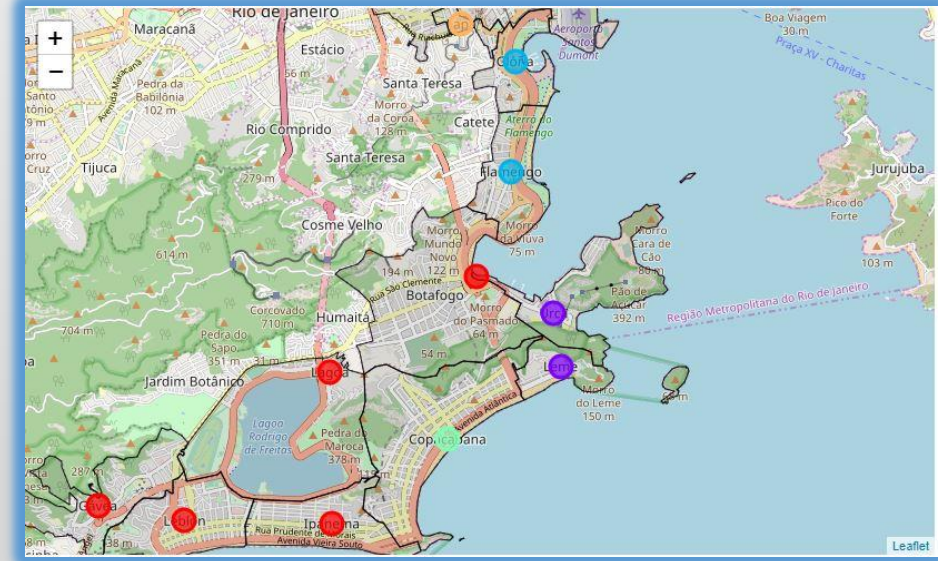
Understand each cluster

- **Red:** Park, gym, fitness center
- **Cyan:** Beach bar, Gym, Resorts
- **Purple:** Scenic Lookout, Mountain, Trail
- **Blue:** Historic Site, Theater, History Museum
- **Brown:** Music Venue, Nightclubs



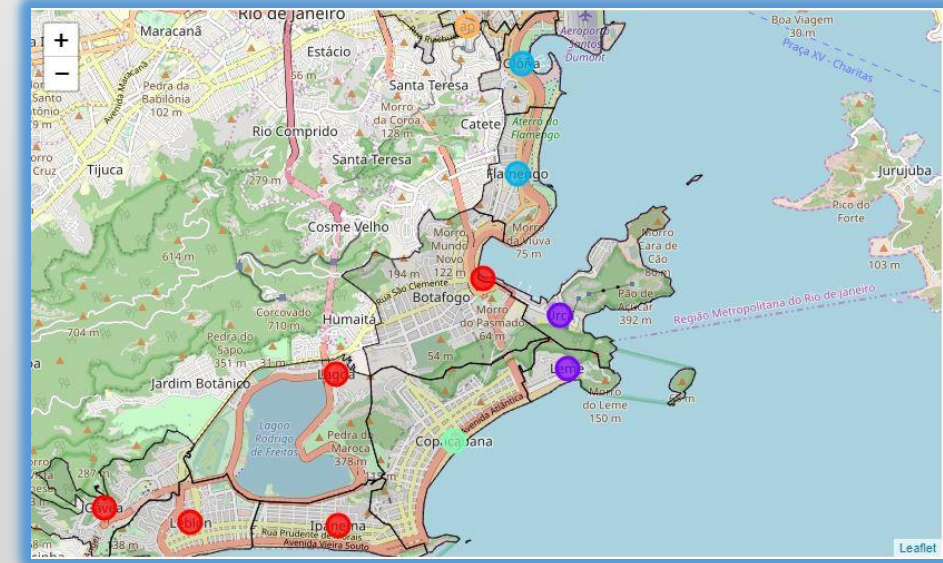
Understand each cluster

Red: Park, gym, fitness center



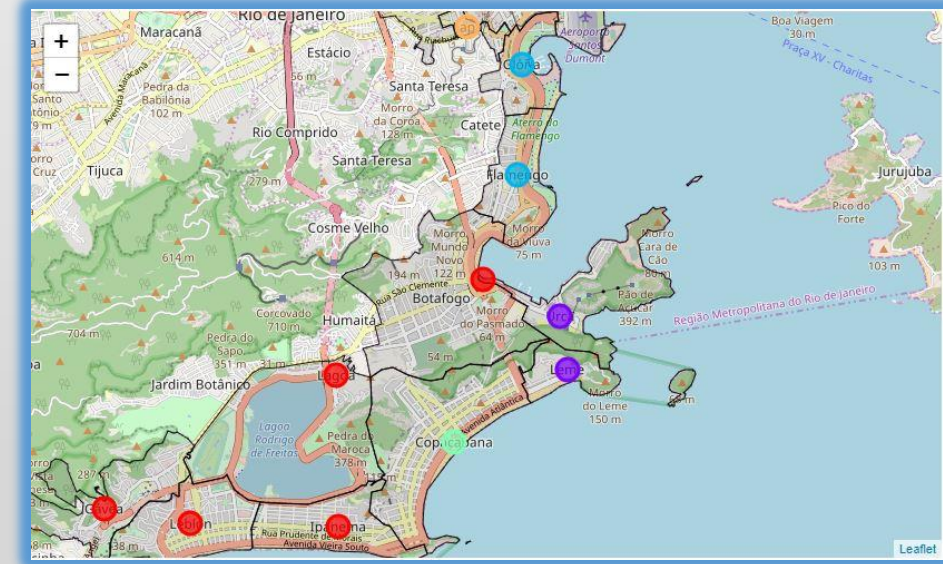
Understand each cluster

Cyan: Beach bar, Gym, Resorts



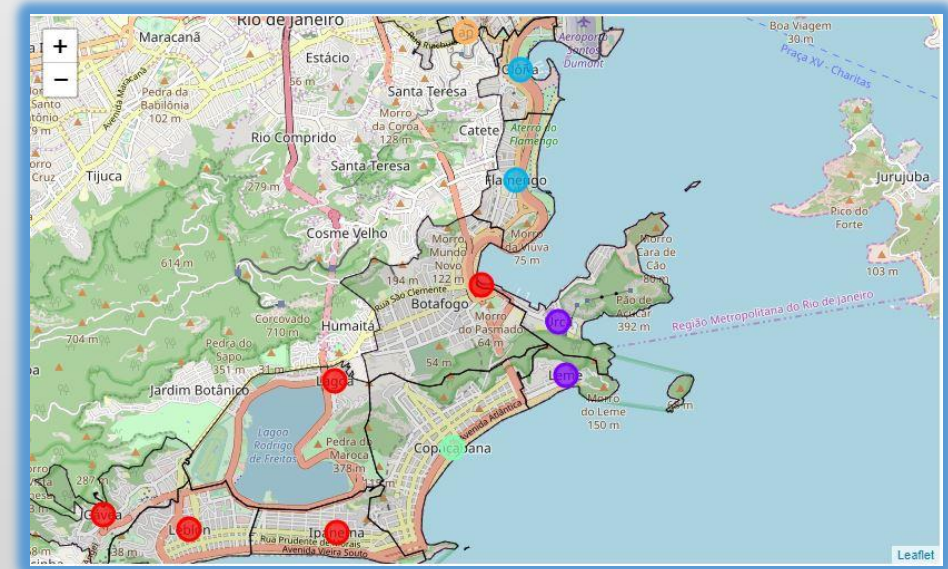
Understand each cluster

Purple: Scenic Lookout, Mountain, Trail



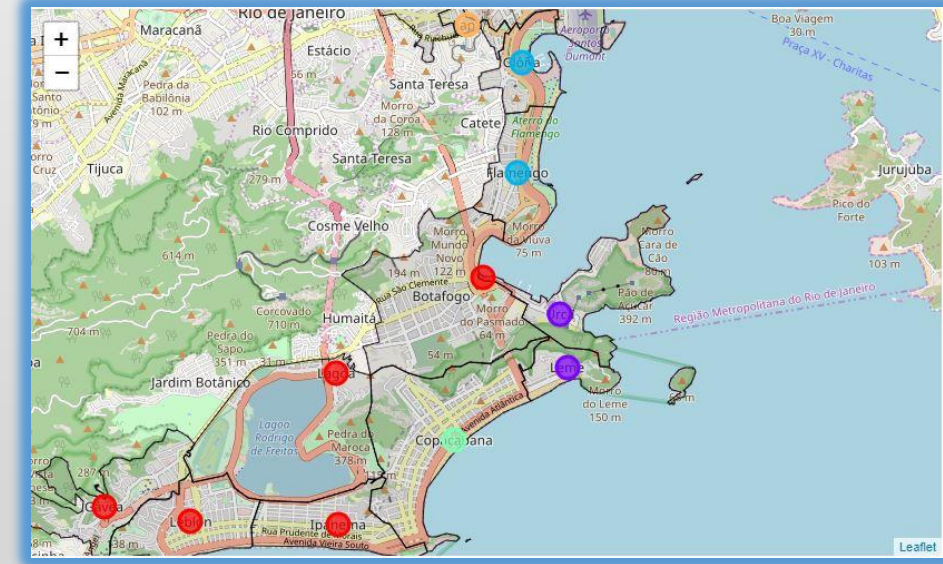
Understand each cluster

Blue: Historic Site, Theater, History Museum



Understand each cluster

Brown: Music Venue, Nightclubs



Conclusion and future directions

- Rio proves to be a city for all tastes, offering various attractions for different audiences
- There is a potential for this kind of analysis in real life business. One way to improve this report would be to include ratings of the venues and collect more data with a paid Foursquare account.

End