

# INTROCUCTION

## Business problem

Rio de Janeiro is a Brazilian municipality, capital of the state of the same name, located in the Southeast of the country. It is also the largest international tourist destination in Brazil, Latin America and the entire Southern Hemisphere. In addition, it is the second largest metropolis in Brazil (after São Paulo), the sixth largest in America and the thirty-fifth in the world. Its population estimated by the Brazilian Institute of Geography and Statistics (IBGE) on July 1, 2020 was 6,747,815 inhabitants.

The city's territorial extension is 1,200,329 km<sup>2</sup>, generating a demographic density of 5,621 inhabitants per square kilometer. In relation to the city's wealth generation potential, in 2018 the GDP revolved around R \$ 364 million, which gives an average GDP per capita of R \$ 54,426.08.

According to the general information mentioned above, it appears that the city has great economic potential to be explored. However, for a neighborhood to be attractive for a company to set up, it is necessary to show the potential of that neighborhood, which are the most frequent locations, in addition to the average monthly income of resident residents.

Thus, this work aims to raise this information and create a visualization with the folium library, so that a company when analyzing the neighborhoods can easily identify which would be the best place to set up, based on its target audience and population income.

## Description of the data

The data used to solve the problem are listed below:

- I found the geojson file of the coordinates of the city's neighborhoods in a repository (<https://gist.github.com/esperanc/db213370dd176f8524ae6ba32433f90a>). As the data was in the correct format, I didn't need to do any cleaning on the base.
- Forsquare API will be used to find the most frequent locations in each neighborhood
- Socioeconomic data were obtained from the website of the Brazilian Institute of Geography and Statistics (IBGE)

## Methodology

The complete methodology is described in the notebook.

# Analysis of results

