

## DATA

For the present study, we will need the names of the boroughs of São Paulo and Rio de Janeiro, as well as the latitude and longitude of each one. For São Paulo, this information can be found on wikipedia ([https://pt.wikipedia.org/wiki/Lista\\_dos\\_distritos\\_de\\_S%C3%A3o\\_Paulo\\_por\\_popula%C3%A7%C3%A3o](https://pt.wikipedia.org/wiki/Lista_dos_distritos_de_S%C3%A3o_Paulo_por_popula%C3%A7%C3%A3o)) and for Rio, we can get it on the website “Data Rio” (<https://www.data.rio/datasets/limite-de-bairros/data?geometry=-44.752%2C-23.138%2C-42.139%2C-22.695&page=6>).

Geographic coordinates can be obtained using the Python Geocoder library or on the Batch Geo website (<https://batchgeo.com/>).

Based on this information, we will use the Foursquare API to discover the locations within a 500-meter radius and then organize the neighborhoods into clusters with the scikit-learn library.