

Data Section

In order to find the best neighbourhoods in RIO for Pharmacy Retail business our data Science team will take an in-depth look in the local market on each neighbourhood. That means finding statistical data about the neighbourhoods, close hospitals/clinics , population, life expectancy, average income and the amount of pharmacies 'The Competition'. In our process, there will be 4 datasets that will be merged into one at the end of extraction:

- The first dataset consists of data about all official neighbourhoods in Rio de Janeiro. This dataset is composed of the names of each neighbourhood, region, subprefecture, latitude and longitude. This data will be gathered through web scraping of this link on Wikipedia:

[“https://pt.wikipedia.org/wiki/Lista_de_bairros_da_cidade_do_Rio_de_Janeiro”](https://pt.wikipedia.org/wiki/Lista_de_bairros_da_cidade_do_Rio_de_Janeiro)

[illegible]

- The second dataset consist of data about the demographics of each neighbourhood. The dataset is composed of Total Population, Total Area and Population Density. This data will be gathered through web scraping of all the **LINKs** of each neighbourhood of the previous wikipedia link. One example of scraped page:

["https://pt.wikipedia.org/wiki/G%C3%A1vea_\(bairro_do_Rio_de_Janeiro\)"](https://pt.wikipedia.org/wiki/G%C3%A1vea_(bairro_do_Rio_de_Janeiro))

[Página principal](#)

[Portal comunitário](#)

[Mudanças recentes](#)

[Manutenção](#)

[Criar página](#)

[Páginas novas](#)

[Contato](#)

[Doações](#)

[Imprimir esta página](#)

[Criar um livro](#)

[Descargar como PDF](#)

[Versão para impressão](#)

Ferramentas

[Páginas afiluentes](#)

[Alterações relacionadas](#)

[Cancelar histórico](#)

[Páginas especiais](#)

[Hiperligação permanente](#)

[Informações da página](#)

[Elemento vinculado](#)

[Outra esta página](#)

Nóveis línguas

Esportif

Norsk bokmål

Svenska




[⚙ Editar hiperligações](#)

2 Zonas não-oficiais

- 2.1 Zona Central
- 2.2 Zona Norte
- 2.3 Zona Sul
- 2.4 Zona Oeste

3 Referências

Bairros oficiais [[editar](#) | [editar código-fonte](#)]

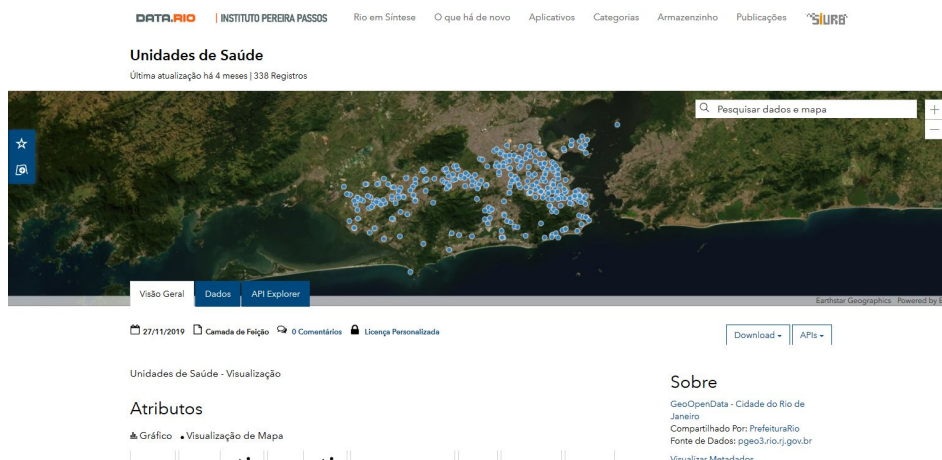
Zona	Subprefeitura(nº de bairros)	Bairros ^[1]	Imagens
Central	Centro Histórico e Zona Portuária (17) ^[1]	Bairro Imperial de São Cristóvão • Benfica • Caju • Camburi • Centro • Cidade Nova • Estácio • Gamboa • Lapa • Mangueira • Paqueta • Rio Comprido • Santa Teresa • Santo Cristo • Saúde • Vasco da Gama	
Sul	Zona Sul (17) ^[1]	Botafogo • Catete • Copacabana • Cosme Velho • Flamengo • Glória • Humaitá • Ipanema • Jardim Botânico • Lagoa • Laranjeiras • Leblon • Leme • São Conrado • Urca • Vidigal	
	Barra da Tijuca e Baixada de Jacarepaguá (20)	Anil • Barra da Tijuca • Camomim • Cidade de Deus • Curitica • Freguesia de Jacarepaguá • Gardênia Azul • Grumari • Itanhangá • Jacarepaguá • Joá • Praça Seca • Pechincha • Recreio dos Bandeirantes • Tanque • Taquara • Vargem Grande • Vargem Pequena • Vila Valqueire • Jardim Sulacap	

- The third dataset consists of data about Health, Education and Income about each neighbourhood in Rio. The data is composed of life expectancy, income, education index and human development index of each neighbourhood. This data will be gathered through web scraping of this link on Wikipedia:

[‘https://pt.wikipedia.org/wiki/Lista_de_bairros_do_Rio_de_Janeiro_por_IDH’](https://pt.wikipedia.org/wiki/Lista_de_bairros_do_Rio_de_Janeiro_por_IDH)

- The Fourth Dataset consists of data about the location of all Clinics and Hospitals of SUS “Health Unic System” . The data is composed by types of clinics, latitude, longitude, neighborhood. This dataset is downloaded at the website data.rio that is a website with open data about the city of Rio de janeiro:

[“http://www.data.rio/datasets/unidades-de-sa%C3%BAd”](http://www.data.rio/datasets/unidades-de-sa%C3%BAd)



After extracting all datasets and cleaning them, our team is going to merge them into one to analyse and select a group of neighbourhoods to further analysis. In the next step our team is going to use the foursquare API to find all pharmacies on each neighbour finding how competitive each one is. In the end our team is plans to suggest few strategies and locations for Drogasil S.A.