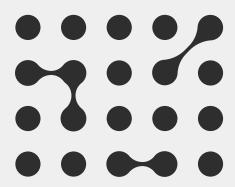
# Python Nordeste 2024

### **Tutorial**

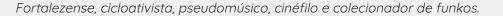
Visualização e Processamento de Dados Geoespaciais: Uma Introdução Prática

■ Pedro Florencio - Data Scientist @lpplan



# Quem sou eu?

- Graduando de Engenharia Elétrica na Universidade Federal do Ceará (UFC);
- Assessor Técnico na área de Ciência de Dados no Instituto de Pesquisa e Planejamento de Fortaleza (Ipplan Fortaleza);
- Pesquisador Voluntário de graduação no Biomedical Data Analytics Research Group (BioData UFC);





# Conteúdo Programático



Ao final deste tutorial, você será capaz de:

- 1. Compreender o modelo e estrutura de dados geoespaciais;
- 2. Realizar processamento de dados com Shapely, Pandas e GeoPandas;
- **3.** Produzir visualizações de mapas dos tipos:
  - a. Estática
    - i. **Matplotlib**
  - b. Interativa
    - i. Plotly
    - ii. **Folium** (Leaflet)
    - iii. **Datawrapper** (Plataforma *no-code*)
- **4.** Entender o processo de uma análise de dados direcionada ao suporte de decisões de políticas públicas do seu município.

# Resultados esperados



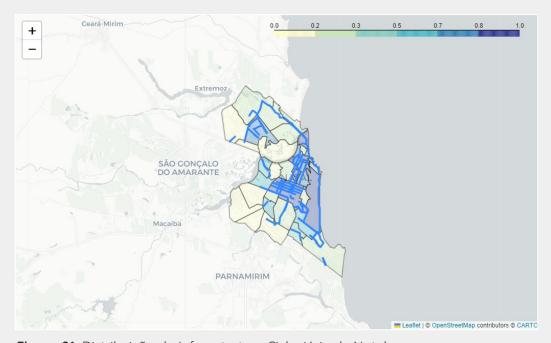


Figura 01: Distribuição da Infraestrutura Cicloviária de Natal

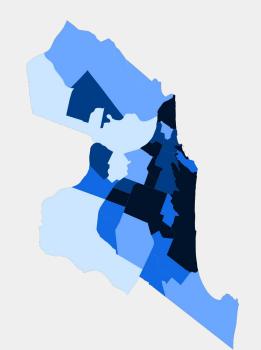


Figura 02: People Near Bike Lanes



Figura 03: Tabela (objeto DataFrame)

	id	bairro
0	9	Rocas
1	11	Praia do Meio
2	34	Neópolis
3	35	Capim Macio
4	26	Nova Descoberta

"Everything is related to everything else, but near things are more related than distant things, as a consequence of accessibility."

Waldo R. Tobler



Figura 04: Tabela Geográfica

	id	bairro	geometry
0	9	Rocas	POLYGON ((256848.065 9361403.003, 256791.200 9
1	11	Praia do Meio	POLYGON ((257073.088 9361374.537, 257074.693 9
2	34	Neópolis	POLYGON ((256774.934 9350774.711, 256732.016 9
3	35	Capim Macio	POLYGON ((256758.629 9353564.405, 256880.038 9
4	26	Nova Descoberta	POLYGON ((257075.549 9355457.924, 257105.533 9

"Everything is related to everything else, but near things are more related than distant things, as a consequence of accessibility."

Waldo R. Tobler

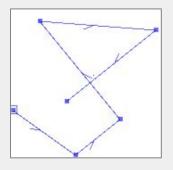


# **Geometry types** A geometry type that represents a single coordinate with x,y Point (\*args) and possibly z values.

**Point** 



Geometry type					
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.				
LineString ([coordinates])	A geometry type composed of one or more line segments.				



LineString



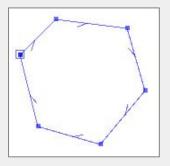
A geometry type that represents a single coordinate with x,y and possibly z values.	Geometry types	3
	Point (*args)	
nestring ([coordinates]) A geometry type composed of one or more line segments.	LineString ([coordinates])	A geometry type composed of one or more line segments.



Figura 03: Malha Cicloviária de Natal



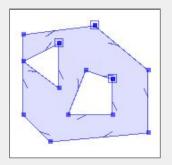
Geometry types			
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.		
LineString ([coordinates])	A geometry type composed of one or more line segments.		
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.		



LinearRing



Geometry types			
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.		
LineString ([coordinates])	A geometry type composed of one or more line segments.		
<u>LinearRing</u> ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.		
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.		



Polygon



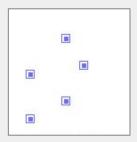
Geometry types			
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.		
LineString ([coordinates])	A geometry type composed of one or more line segments.		
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.		
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.		



Figura 04: Bairro Rocas



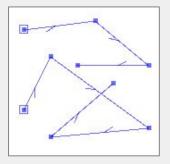
Geometry types			
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.		
LineString ([coordinates])	A geometry type composed of one or more line segments.		
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.		
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.		
MultiPoint ([points])	A collection of one or more Points.		



**MultiPoint** 



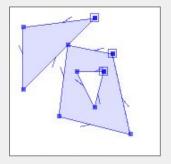
Geometry types		
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.	
LineString ([coordinates])	A geometry type composed of one or more line segments.	
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.	
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.	
MultiPoint ([points])	A collection of one or more Points.	
MultiLineString ([lines])	A collection of one or more LineStrings.	



MultiLineString



Geometry types			
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.		
LineString ([coordinates])	A geometry type composed of one or more line segments.		
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.		
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.		
MultiPoint ([points])	A collection of one or more Points.		
MultiLineString ([lines])	A collection of one or more LineStrings.		
MultiPolygon ([polygons])	A collection of one or more Polygons.		



MultiPolygon



Geometry types			
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.		
<u>LineString</u> ([coordinates])	A geometry type composed of one or more line segments.		
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.		
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.		
MultiPoint ([points])	A collection of one or more Points.		
<pre>MultiLineString ([lines])</pre>	A collection of one or more LineStrings.		
MultiPolygon ([polygons])	A collection of one or more Polygons.		

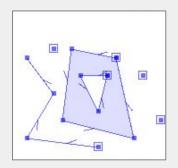




Figura 05: Bairros de Natal



Geometry types		
Point (*args)	A geometry type that represents a single coordinate with x,y and possibly z values.	
<u>LineString</u> ([coordinates])	A geometry type composed of one or more line segments.	
LinearRing ([coordinates])	A geometry type composed of one or more line segments that forms a closed loop.	
Polygon ([shell, holes])	A geometry type representing an area that is enclosed by a linear ring.	
MultiPoint ([points])	A collection of one or more Points.	
<pre>MultiLineString ([lines])</pre>	A collection of one or more LineStrings.	
<u>MultiPolygon</u> ([polygons])	A collection of one or more Polygons.	
GeometryCollection ([geoms])	A collection of one or more geometries that may contain more than one type of geometry.	



GeometryCollection



**Tabela 01:** Tipos de arquivos de dados geoespaciais

Formato	Descrição
Shapefile (.shp)	Conjunto de arquivos para armazenar dados geoespaciais, amplamente utilizado. Inclui .shp (geometria), .shx (índice) e .dbf (atributos).
GeoJSON	Formato baseado em JSON para representar dados vetoriais, como pontos, linhas e polígonos.
KML	Formato XML desenvolvido para o Google Earth, usado para representar dados como placemarks, polígonos e camadas de imagens.
GPKG (GeoPackage)	Formato baseado em SQLite para armazenamento de dados vetoriais e raster, alternativa ao Shapefile.



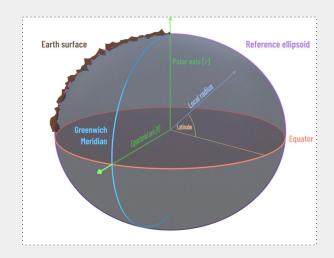
Sistemas de Referência de Coordenadas (SRC)

Um sistema de georreferenciamento de coordenadas, também conhecido como Sistema de Referência de Coordenadas (CRS - Coordinate Reference System), é um sistema que utiliza **coordenadas** (geralmente latitude e longitude) para **identificar** e **localizar** pontos na superfície da Terra. Esses sistemas podem ser baseados em **modelos diferentes da Terra** (esferoides ou elipsoides) e podem ser projetados para diversas finalidades, como mapeamento global, regional ou local.



Tipos de Sistemas de Coordenadas

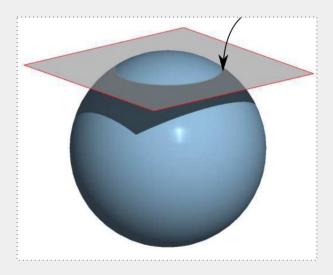
**Geodésico:** Utiliza a superfície do **elipsoide** para representar a Terra. As coordenadas são expressas em latitude, longitude e, às vezes, altitude. Um exemplo comum é o WGS 84, que é utilizado pelo GPS.

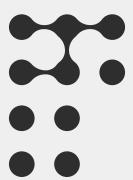




#### Tipos de Sistemas de Coordenadas

Projetado: Transforma coordenadas esféricas (latitude e longitude) em um sistema cartesiano (x, y), permitindo que a superfície curva da Terra seja representada em uma superfície plana, como em mapas. Exemplos incluem UTM (Universal Transverse Mercator) e projeções como a de Mercator, Lambert, entre outras.







**Hands-On:** Onde investir em infraestrutura cicloviária na minha cidade?

Estudo de Caso

### People Near Bike Lanes (PNB)

#### Acesso à Infraestrutura Cicloviária

- > Desenvolvido pelo Institute for Transportation and Development Policy (ITDP), foi criado para medir a acessibilidade das infraestruturas de ciclovias e a proximidade da população a essas vias.
- > O **PNB** é definido matematicamente como a proporção da população que vive a pelo menos **300 metros de uma infraestrutura cicloviária**, sendo um valor de varia de 0 a 1, isto é, de 0 a 100% da população.

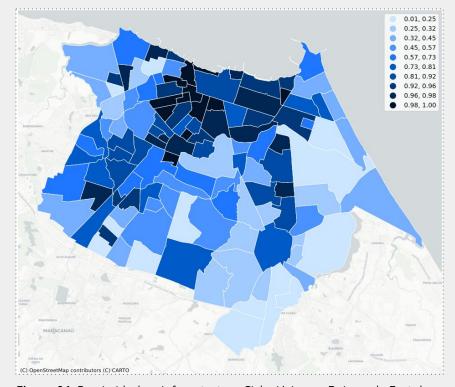


Figura 06: Proximidade a Infraestrutura Cicloviária nos Bairros de Fortaleza





### Passo-a-Passo



### Aquisição de Dados

Leitura dos dados e formação do objeto *GeoDataFrame* 





#### E.D.A.

Análise Exploratória dos Dados com *Pandas*, *Matplotlib* e *Seaborn* 

#### Pré-Processamento

Limpeza e preparação dos dados usando Shapely e GeoPandas



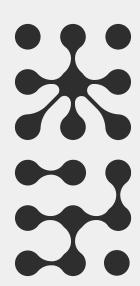


### Visualização

Composição dos mapas como produto



# Aquisição de Dados



Leitura dos dados e formação dos objetos *GeoDataFrame* 



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
   import contextily as cx
   import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca que realiza o mapeamento do pacote *Leaflet.JS* para o Python.



```
import folium
   import numpy as np
 3 import pandas as pd
4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
 7 import contextily as cx
 8 import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca responsável
por funções
matemáticas, geração de
números aleatórios,
álgebra linear,
Transformada de Fourier,
entre outros.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
   import contextily as cx
8 import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca que realiza a manipulação e permite a análise de dados estruturados.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
6 import geopandas as gpd
   import contextily as cx
   import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca de visualização de dados baseada na Matplotlib.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
   import contextily as cx
 8 import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca de manipulação de dados geoespaciais, que se baseia nos métodos do **Pandas** e faz uso de métodos do **Shapely** e **Matplotlib**.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
 7 import contextily as cx
8 import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```

#### Contextily

Pacote Python que permite a inserção de *tilemaps* (mapas de textura) como uma camada base nas visualizações.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
   import contextily as cx
 8 import plotly.express as px
 9 from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca de plotagem interativa que permite a visualização de gráficos e mapas. O módulo express permite criar figuras inteiras de maneira prática.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
   import contextily as cx
   import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```



Biblioteca de plotagem estática, interativa e visualizações animadas no Python. Além disso ela permite a visualização de **mapas**.



```
import folium
   import numpy as np
 3 import pandas as pd
 4 import seaborn as sns
 5 import folium.plugins
 6 import geopandas as gpd
 7 import contextily as cx
 8 import plotly.express as px
   from google.colab import drive
   import matplotlib.pyplot as plt
   from shapely.geometry import Point
   import matplotlib.colors as mcolors
13
   drive.mount('/gdrive', force_remount=True)
15
   %matplotlib inline
```

#### Shapely

Pacote para análise teórica e manipulação de modelos geoespaciais.

### Acesso aos Datasets



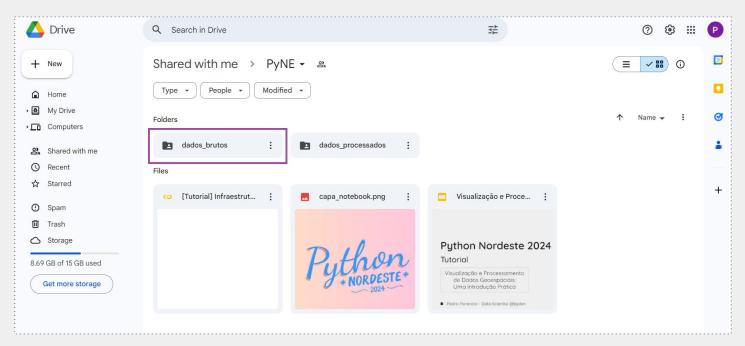


Figura 07: Local de armazenamento dos datasets no Google Drive

### Acesso aos Datasets



Name	<b>↑</b>
	bairros_natal

#### bairros\_natal

Shapefile (.shp) dos polígonos dos bairros do município.

Figura 08: Conjuntos de dados necessários para o cálculo.



Name	<b>↑</b>
	bairros_natal
	limites_municipio

### limites\_municipio

Limites municipais de Natal no formato .shp

Figura 08: Conjuntos de dados necessários para o cálculo.





### agregados\_preliminares

Dados sobre os setores censitários do Rio Grande do Norte

Figura 08: Conjuntos de dados necessários para o cálculo.



<b>↑</b>
bairros_natal
limites_municipio
agregados_preliminares_por_setores_censitarios_RN.csv
Dicionario_CNEFE_Censo_2022.xls
infraestrutura_cicloviaria_natal.geojson 😃

infraestrutura\_cicloviaria

Dados sobre a infraestrutura cicloviária de Natal

Figura 08: Conjuntos de dados necessários para o cálculo.



Name	<b>↑</b>
	bairros_natal
	limites_municipio
	agregados_preliminares_por_setores_censitarios_RN.csv 😃
X	Dicionario_CNEFE_Censo_2022.xls
	infraestrutura_cicloviaria_natal.geojson 😃
	RN_malha_preliminar_2022.gpkg 🚢

RN\_malha\_preliminar

Geolocalização dos setores censitários do Rio Grande do Norte

Figura 08: Conjuntos de dados necessários para o cálculo.



<b>↑</b>
bairros_natal
limites_municipio
agregados_preliminares_por_setores_censitarios_RN.csv 🚢
Dicionario_CNEFE_Censo_2022.xls 🕰
infraestrutura_cicloviaria_natal.geojson 😃
RN_malha_preliminar_2022.gpkg 🐣

Métodos

```
pandas.read_csv(...)
geopandas.read_file(...)
```

Figura 08: Conjuntos de dados necessários para o cálculo.



### 000

```
# caminho para os dados brutos
caminho_dados = '/gdrive/MyDrive/PyNE/dados_brutos/'

# setores censitarios do Rio Grande do Norte
df_setores = pd.read_csv(caminho_dados+'agregados_preliminares_por_setores_censitarios_RN.csv', sep=';')

# malha georreferenciada de setores censitarios
gdf_setores = gpd.read_file(caminho_dados+'RN_malha_preliminar_2022.gpkg', crs='EPSG:4326', driver='GPKG')

# limites dos bairros de Natal
gdf_bairros = gpd.read_file(caminho_dados+'bairros_natal/bairros.shp', crs='EPSG:4326')

# infraestrutura cicloviaria
gdf_infraestrutura = gpd.read_file(caminho_dados+'infraestrutura_cicloviaria_natal.geojson', crs='EPSG:4326')
```



Figura 08: Pandas DataFrame

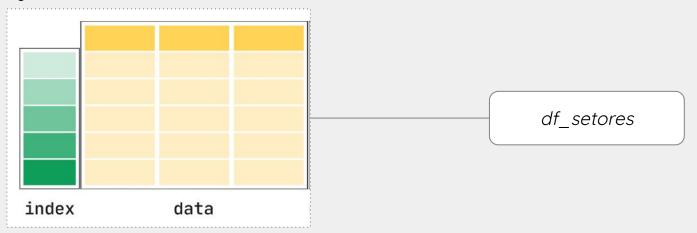
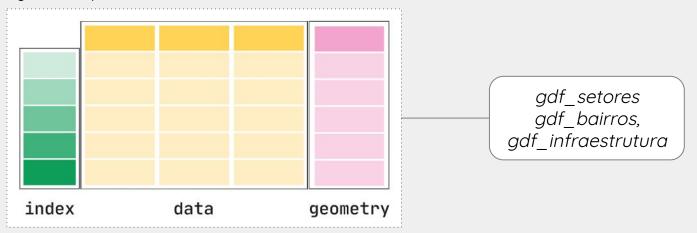




Figura 09: Geopandas GeoDataFrame





### Entendimento dos Dados

	Name	description	timestamp	begin	end	altitudeMode	tessellate	extrude	visibility	draw0rder
0	Ciclofaixa da Av. Pref. Omar O'Grady	Administrada pelo Governo do Estado;   Mã	None	None	None	None	1	0	-1	None
1	Ciclofaixa da Via Costeira	Administrada pelo Governo do Estado;   hã	None	None	None	None	1	0	-1	None
2	Ciclovia Anel Viário da UFRN	> Administrada pela UFRN;  br>> Mão-única.	None	None	None	None	1	0	-1	None
3	Ciclovia do Anel Viário da UFRN	> Administrada pela UFRN; 	None	None	None	None	1	0	-1	None
4	Ciclovia Interna da UFRN I	Administrada pela UFRN;  br>> Mão dupla.	None	None	None	None	1	0	-1	None

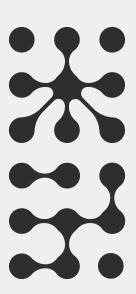
### Métodos





# Pré-Processamento

Limpeza e preparação dos dados usando *Shapely* e *GeoPandas* 



# Pré-Processamento de Dados



### Seleção de colunas

```
O O O

1 gdf = gdf[['coluna_1','coluna_2',...,'coluna_n']]
```

### Conversão de tipos

```
1 gdf.nome_coluna = gdf.nome_coluna.astype('nome_tipo').copy()
```

### Substituição de nomes

### Filtro de linhas

```
0 0 0

1 df = df[df['coluna'] == numero]
```

Com estes métodos você será capaz de realizar até o item 6.6

# Pré-Processamento de Dados



### Junção de tabelas

	lef	ft		right				Result					
	key	А	В	. [	key	С	D	- [	key	A	В	С	D
0	KD	AD	BO	0	KD	0	DO	0	KD	AD	BO	co	DO
1	кі	AI	B1	1	кі	а	D1	1	кі	A1	B1	а	DI
2	K2	A2	B2	2	K2	Q	D2	2	K2	A2	B2	Œ	DZ
3	КЗ	АЗ	B3	3	КЗ	в	D3	3	кз	АЗ	В3	з	D3

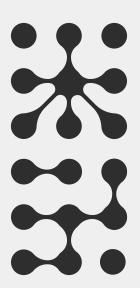
```
OOO

1 gdf = pd.merge(df_1, df_2[['key', 'geometry']], on='key', how='left')
```



E.D.A

Análise Exploratória dos Dados com Pandas, Matplotlib e Seaborn



### Limites dos Bairros de Natal

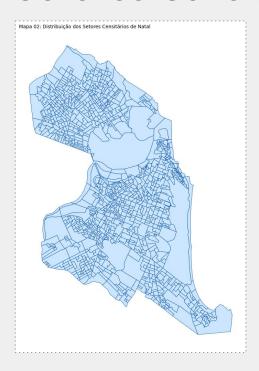




```
000
 1 # parametros do mapa
 2 titulo_do_mapa = 'Mapa 01: Limites dos Bairros de Natal'
 3 resolucao_dpi = 'figure'
 4 tamanho_do_mapa = (10,10)
 6 # criando figura base
 7 fig, ax = plt.subplots(figsize=tamanho_do_mapa)
 9 # criando camada de limites dos bairros
10 gdf_bairros.boundary.plot(ax=ax, color='black', linewidth=.7, alpha=.5)
12 # configurando nome
13 plt.title(titulo_do_mapa, loc='left', fontsize=10)
14
15 # removendo eixos da visualização
16 plt.axis('off')
18 # adicionando cartografia base
19 cx.add_basemap(ax=ax, crs=gdf_bairros.crs, zoom=13, source='CartoDB positron')
20
21 plt.show()
```

# Setores Censitários de Natal





### 000

```
titulo_do_mapa = 'Mapa 02: Distribuição dos Setores Censitários de Natal'
resolucao_dpi = 'figure'
tamanho_do_mapa = (12,12)

# criando figura base
fig, ax = plt.subplots(figsize=tamanho_do_mapa)

# malha de setores censitarios
gdf_setores_natal.plot(ax=ax,lw=.6, edgecolor='#004BA6', color='#CCE6FF')

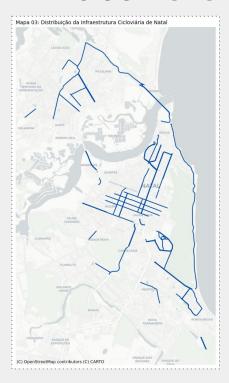
# configurando nome
plt.title(titulo_do_mapa, loc='left', fontsize=10)

# removendo eixos da visualizacao
plt.axis('off')

plt.show()
```

### Infraestrutura Cicloviária de Natal





```
000
 1 titulo_do_mapa = 'Mapa 03: Distribuição da Infraestrutura Cicloviária de Natal'
 2 resolucao_dpi = 'figure'
   tamanho_do_mapa = (12, 12)
 5 # criando figura base
 6 fig, ax = plt.subplots(figsize=tamanho_do_mapa)
 8 # adicionando camada de infraestrutura cicloviaria
   gdf_infraestrutura.plot(ax=ax, color='#004BA6')
11 # configurando nome
   plt.title(titulo_do_mapa, loc='left', fontsize=10)
13
14 # removendo eixos da visualização
15 plt.axis('off')
   cx.add_basemap(ax=ax, crs=gdf_infraestrutura.crs, zoom=13, source='CartoDB positron')
18
19 plt.show()
```



# Cálculo do PNB



Vamos para o tópico 7.4 do notebook =)

# Obrigado!

Ficou com dúvidas?

### pedroflorencio@alu.ufc.br



 CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

