

▼ Chapter 8 - Exercise 1: Geojson_hcmc

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import geopandas as gpd
```

```
# Câu 1: Đọc file dữ liệu ranh giới các quận - khu vực của TP.HCM (district-boundary-hcm-city)
district = gpd.read_file('data\district-boundary-hcm-city.geojson')
```

```
# Hiển thị 5 dòng đầu của dữ liệu
district.head()
```



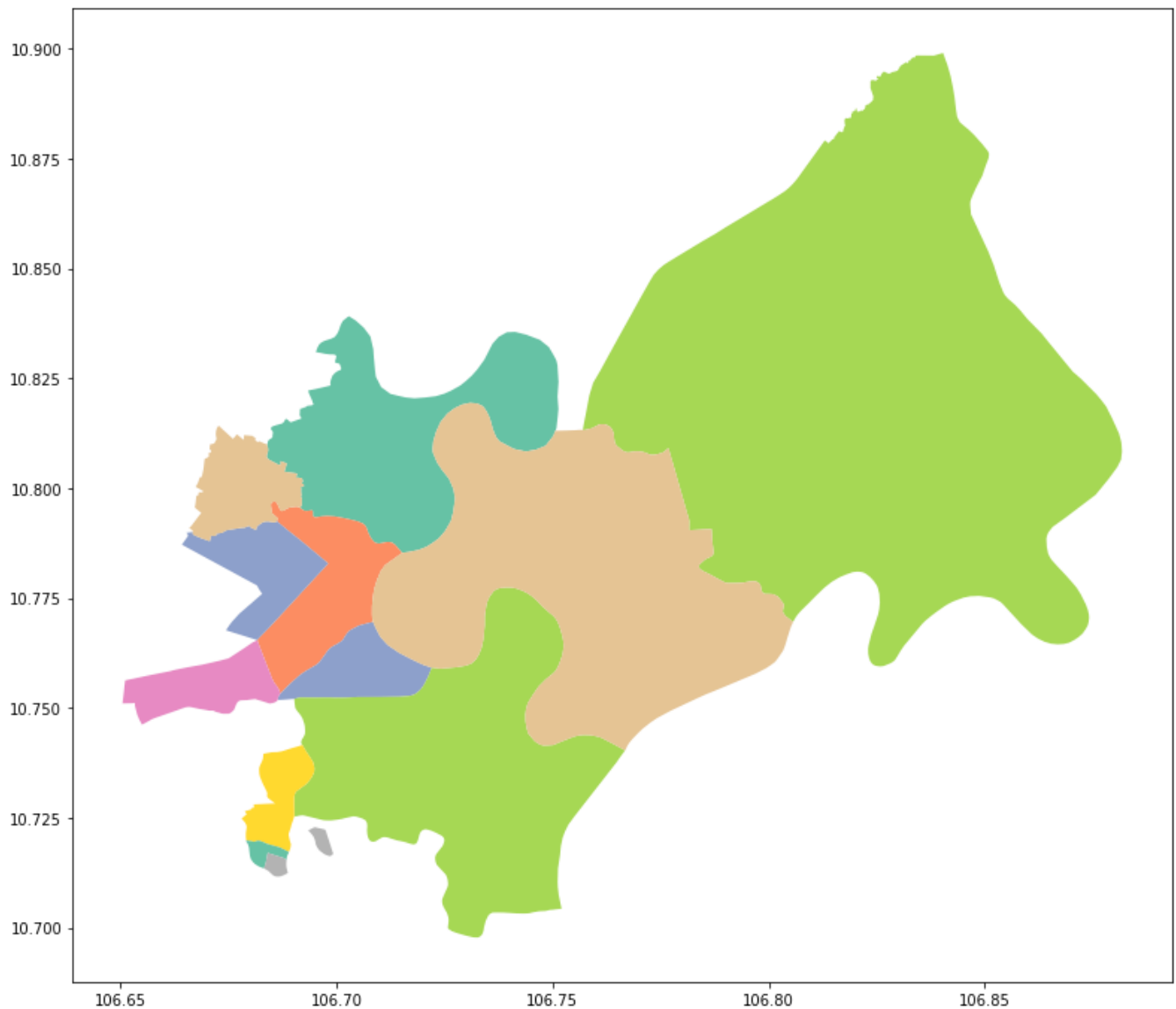
	id	name	localname	timestamp	SRID	admin_level	tags	geometry
0	3850184	Saigon South	Khu đô thị Nam Sài Gòn	2016-03-18T23:05:02	4326	5	{'name': 'Khu đô thị Nam Sài Gòn', 'name:en': ...}	MULTIPOLYGON (((106.69344 10.72213, 106.69475 ...
		Binh	Quận Bình	2016-03			{'name': 'Quận Bình	MULTIPOLYGON (((106.68286

```
# Hiển thị 5 dòng cuối của dữ liệu
district.tail()
```

	id	name	localname	timestamp	SRID	admin_level	tags	
8	3799817	Quận 2	Quận 2	2016-03-18T23:05:02	4326	6	{'name': 'Quận 2', 'boundary': 'administrative...	MULT ((1(
9	3851694	Quận Phú Nhuận	Quận Phú Nhuận	2016-03-18T23:05:02	4326	6	{'name': 'Quận Phú Nhuận', 'boundary': 'admini...	MULT ((1(

```
# Hiển thị bản đồ ranh giới từ dữ liệu district, cmap = 'Set2'
plt.figure(figsize=(8,8))
district.plot(column = 'name', cmap = 'Set2', figsize=(16,12))
plt.show()
```

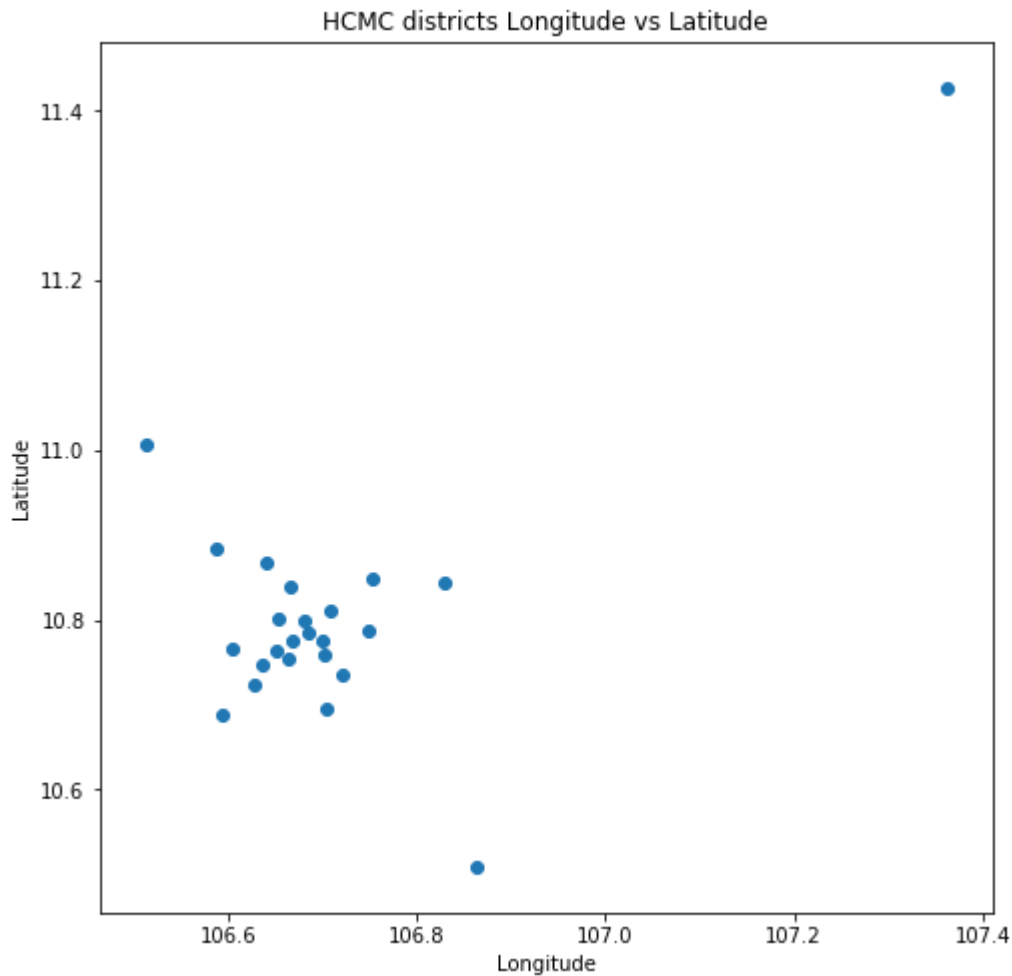
<Figure size 576x576 with 0 Axes>



```
# Câu 2: Đọc dữ liệu trong sheet 'location' của file HCMC_location.xlsx
df_hcm = pd.read_excel('data\HCMC_location.xlsx', sheet_name='Location')
# Hiển thị 5 dòng đầu của dữ liệu
df_hcm.head()
```

	STT	ID	Name	Bourough	Postal cost	Latitude	Longitude	Population	Population
0	1	760	Quận 1	Vietnam, Quan 1	NaN	10.775659	106.700424	193.632	193.632
1	2	761	Quận 12	Vietnam, Quan 12	NaN	10.867153	106.641332	510.326	510.326
2	3	762	Quận Thủ Đức	Vietnam, Thu Duc	NaN	10.849409	106.753705	528.413	528.413

```
# Câu 3: Vẽ scatterplot với dữ liệu Longitude, Latitude của dữ liệu của câu 2
plt.figure(figsize=(8,8))
plt.scatter(df_hcm.Longitude, df_hcm.Latitude)
plt.xlabel("Longitude")
plt.ylabel("Latitude")
plt.title('HCMC districts Longitude vs Latitude')
plt.show()
```



```
# Câu 4: Vẽ bản đồ thể hiện vị trí các quận-khu vực tại TP.HCM
plt.figure(figsize=(16,12))

district.plot(column = 'name', cmap = 'Set2', legend=True, figsize=(16,12))
plt.scatter(df_hcm.Longitude, df_hcm.Latitude)
plt.xlabel("Longitude")
plt.ylabel("Latitude")
plt.title('Districts in HoChiMinh in City')

plt.grid()
plt.show()
```

<Figure size 1152x864 with 0 Axes>

