



Chapter 8 - ex 1

1. Cho district-boundary-hcm-city.geojson. Dùng geopandas để đọc và hiển thị dữ liệu dạng polygon
2. Đọc dữ liệu trong sheet 'location' của file HCMC_location. => tìm hiểu dữ liệu location
3. Vẽ scatterplot gồm lat.long của dữ liệu của câu 2
4. Kết hợp polygon ở câu 1 và scatterplot ở câu 3 để vẽ thành một biểu đồ chung

```
In [1]: import numpy as np # useful for many scientific computing in Python
import pandas as pd # primary data structure library
import matplotlib.pyplot as plt
import geopandas as gpd
```

```
In [2]: # Read in the services district shapefile and look at the first few rows.
service_district = gpd.read_file('data/district-boundary-hcm-city.geojson')
service_district.head()
```

Out[2]:

	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': ...}	(POLYGON ((106.6934435 10.7221295, 106.694747 ...
1	3797166	Binh Thanh District	Quận Bình Thạnh	2016-03-18T23:05:02	4326	6	{'name': 'Quận Bình Thạnh', 'name:en': 'Binh T...	(POLYGON ((106.6838586 10.8071124, 106.6838762...
2	2587287	District 1	Quận 1	2016-03-18T23:05:02	4326	6	{'name': 'Quận 1', 'name:en': 'District 1', 'n...	(POLYGON ((106.6816512 10.7654274, 106.6818712...
3	3819816	District 3	Quận 3	2016-03-18T23:05:02	4326	6	{'name': 'Quận 3', 'name:en': 'District 3', 'n...	(POLYGON ((106.6642244 10.7871435, 106.6645685...
4	2778323	District 4	Quận 4	2016-03-18T23:05:02	4326	6	{'name': 'Quận 4', 'name:en': 'District 4', 'b...	(POLYGON ((106.6863906 10.7518373, 106.6864121...

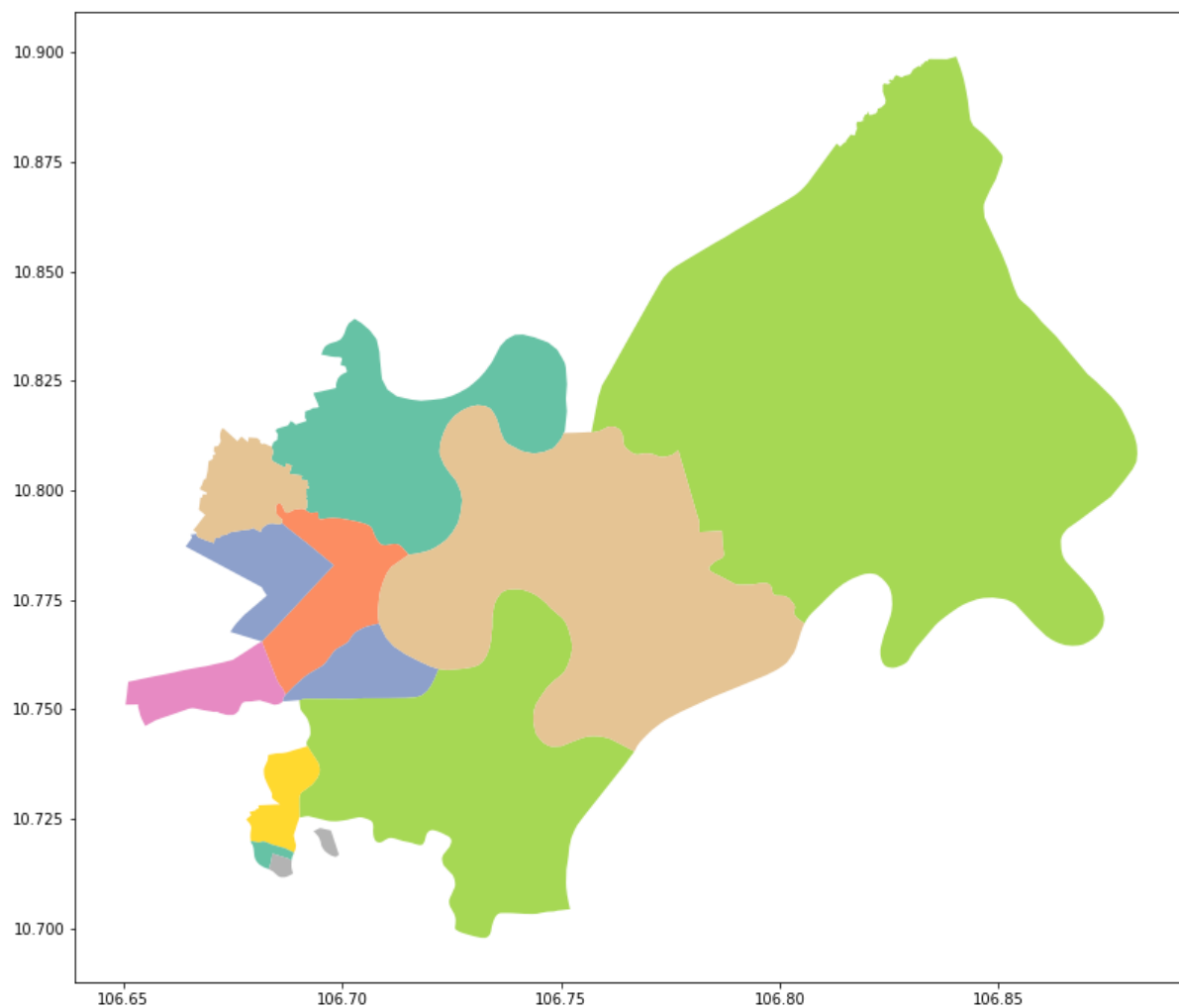
In [3]: `service_district.tail()`

Out[3]:

	id	name	localname	timestamp	SRID	admin_level	tags	geometry
8	3799817	Quận 2	Quận 2	2016-03-18T23:05:02	4326	6	{'name': 'Quận 2', 'boundary': 'administrative'...	(POLYGON ((106.7081991 10.7714064, 106.7082205...
9	3851694	Quận Phú Nhuận	Quận Phú Nhuận	2016-03-18T23:05:02	4326	6	{'name': 'Quận Phú Nhuận', 'boundary': 'admini'...	(POLYGON ((106.6659201 10.7908067, 106.6660183...
10	3853748	Khu 6	Khu 6	2016-03-18T23:05:02	4326	9	{'name': 'Khu 6', 'boundary': 'administrative'...	(POLYGON ((106.6780478 10.7248501, 106.6792251...
11	3854476	Dai Phuc Residences	Khu Nhà ở Rạch Bà Tánh	2016-03-18T23:05:02	4326	10	{'name': 'Khu Nhà ở Rạch Bà Tánh', 'name:en': ...	(POLYGON ((106.6791201 10.719912, 106.6809446 ...
12	3854477	T30 Residential Area	Khu Dân cư T30	2016-03-18T23:05:02	4326	10	{'name': 'Khu Dân cư T30', 'name:en': 'T30 Res...	(POLYGON ((106.6832112 10.7134866, 106.6836886...

```
In [4]: plt.figure(figsize=(8,8))  
service_district.plot(column = 'name', cmap = 'Set2', figsize=(16,12))  
plt.show()
```

<Figure size 576x576 with 0 Axes>



```
In [6]: df_hcm = pd.read_excel('data/HCMC_location.xlsx',  
                                sheet_name='Location')
```

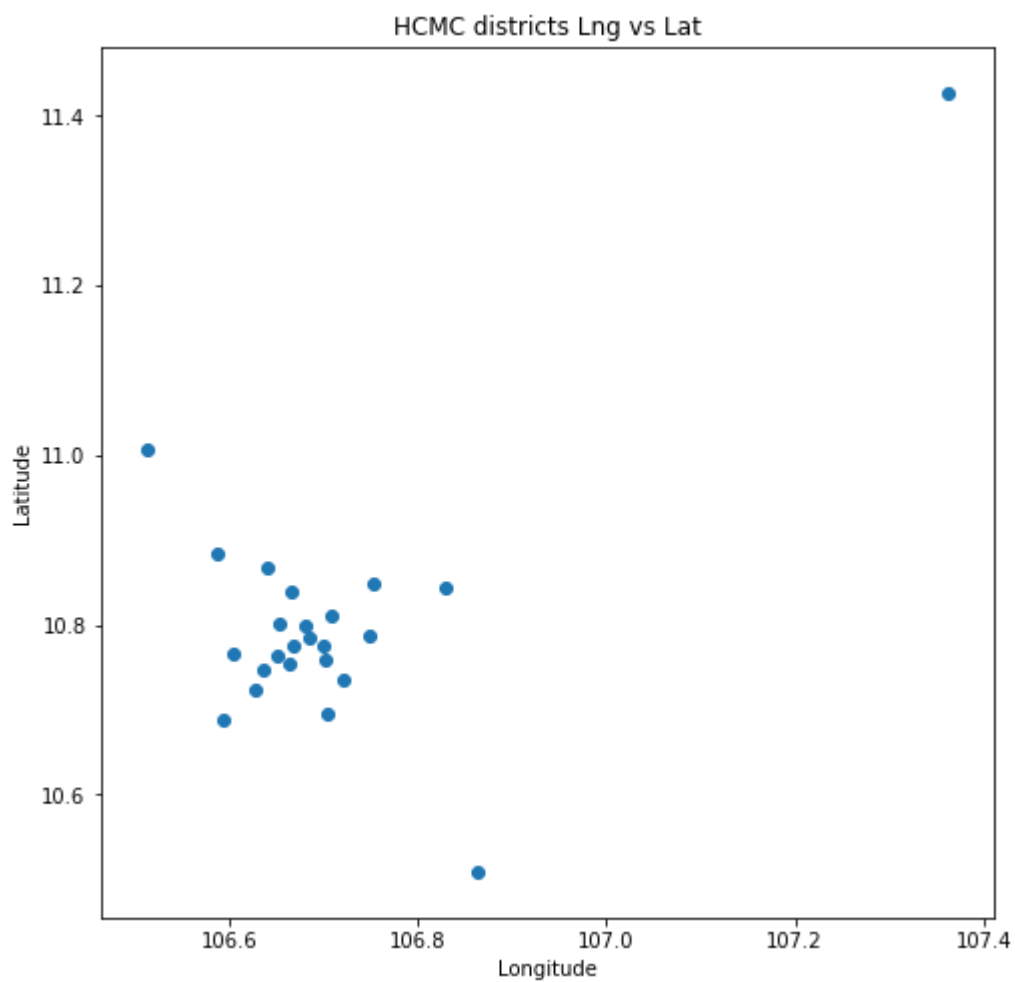
In [7]: `df_hcm.head()`

Out[7]:

	STT	ID	Name	Bourough	Postal cost	Latitude	Longitude	Population	Population_all	Avg_lanc
0	1	760	Quận 1	Vietnam, Quan 1	NaN	10.775659	106.700424	193.632	193632	59
1	2	761	Quận 12	Vietnam, Quan 12	NaN	10.867153	106.641332	510.326	510326	3
2	3	762	Quận Thủ Đức	Vietnam, Thu Duc	NaN	10.849409	106.753705	528.413	528413	5
3	4	763	Quận 9	Vietnam, Quan 9	NaN	10.842840	106.828685	290.620	290620	3
4	5	764	Quận Gò Vấp	Vietnam, Go Vap	NaN	10.838678	106.665290	634.146	634146	9

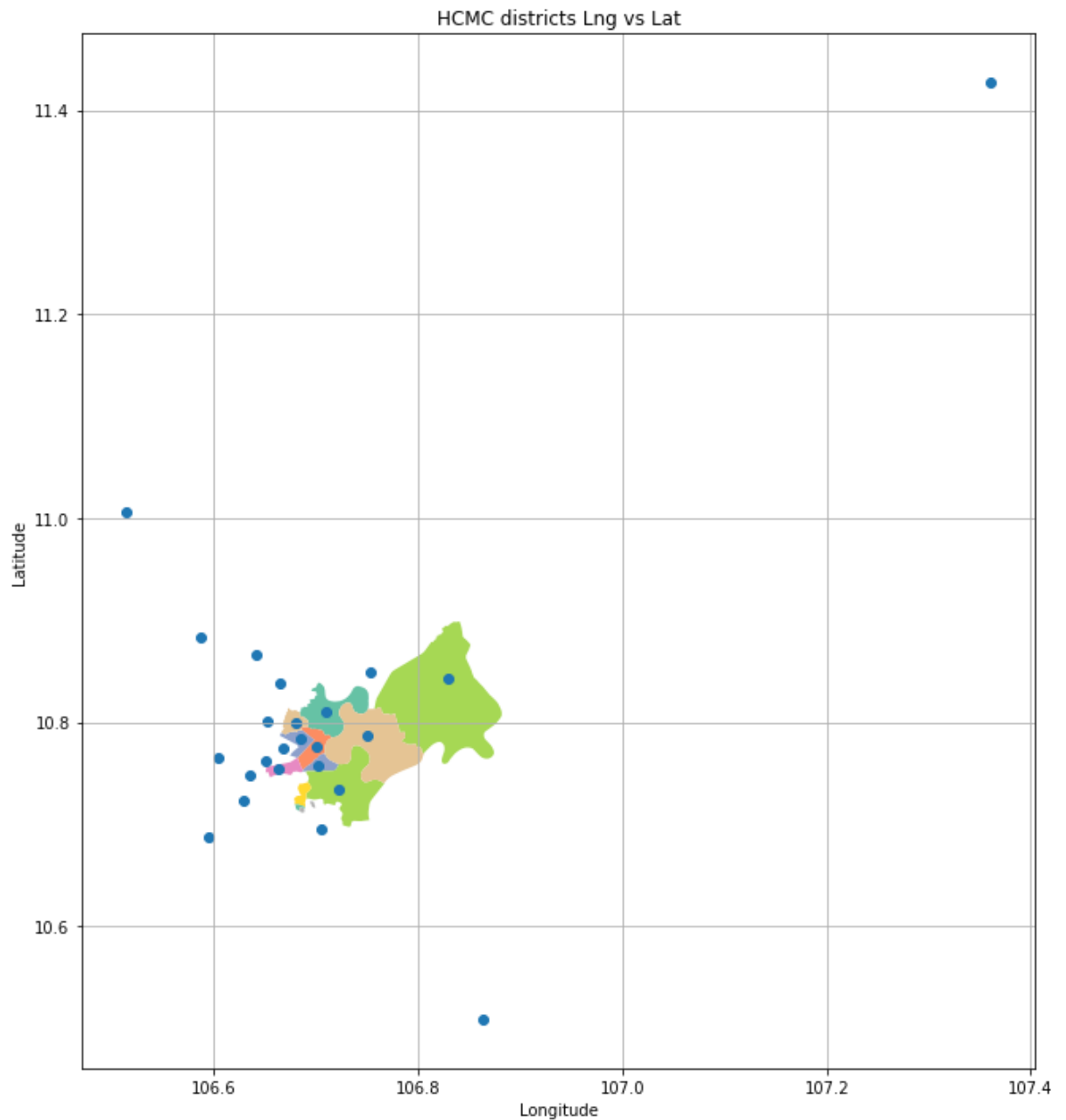


```
In [9]: plt.figure(figsize=(8,8))
plt.scatter(df_hcm.Longitude, df_hcm.Latitude)
plt.xlabel("Longitude")
plt.ylabel("Latitude")
plt.title('HCMC districts Lng vs Lat')
plt.show()
```



```
In [12]: plt.figure(figsize=(16,12))
# Plot the service district shapefile
service_district.plot(column = 'name', cmap = 'Set2', figsize=(16,12))
# Add the school locations
plt.scatter(df_hcm.Longitude, df_hcm.Latitude)
plt.xlabel("Longitude")
plt.ylabel("Latitude")
plt.title('HCMC districts Lng vs Lat')
# Add grid lines and show the plot
plt.grid()
plt.show()
```

<Figure size 1152x864 with 0 Axes>



In []:



