



# NEIGHBORHOODS IN DISTRICT 1 AND DISTRICT 3 IN HCMC VIETNAM

VU HOANG HA LINH



# CONTENTS

- Introduction
- Objectives
- Data
- Methodology
  - Analyze District 1
  - K-mean Cluster District 1
  - Analyze District 3
  - K-mean Cluster District 3
- Results
- Discussion
- Conclusion

---

# 1. INTRODUCTION

Ho Chi Minh City, formerly known as Saigon, once “The Pearl of the Far East” is the second largest city in Vietnam (the largest being Hanoi). It is the most crowded city in the country with the official population of over 8 million people on the total area of over 2,095 km<sup>2</sup>. The city now comprises 19 districts and 5 suburban districts. District 1, along the Saigon River, where downtown Saigon is located, is the commercial center and contains most of the city’s monuments and landmarks. 8 km away in district 5 resides a big market or Cho Lon where Chinese community lives. With favorable weather the whole year round and the convenient access from other countries by air, by road and by sea, Ho Chi Minh city is a busy and dynamic metropolitan.

# 1.1 DISTRICT 1

District 1 (Vietnamese: Quận 1) is the central urban district of Ho Chi Minh City, the largest city in Vietnam. With a total area of 7.7211 km<sup>2</sup> (2.9811 sq mi) the district has a population of 204,899 people as of 2010. The district is divided into 10 small subsets which are called wards (phường). District 1 contains most of the city's administrative offices, consulates, and large buildings. District 1 is the busiest district in the city with the highest living standards. Đồng Khởi street and Nguyễn Huệ boulevard in District 1 are the city's two main commercial centers. Đồng Khởi street is an area in high demand for real estate, hitting a record price of USD50,000 per square meter in 2007.

(source: [https://en.wikipedia.org/wiki/District\\_1,\\_Ho\\_Chi\\_Minh\\_City](https://en.wikipedia.org/wiki/District_1,_Ho_Chi_Minh_City))



---

## 1.2 DISTRICT 3

- District 3 (Vietnamese: Quận 3) is an urban district of Ho Chi Minh City, the largest city in Vietnam. Together with District 1, District 3 is considered the bustling heart of the city, with several businesses, religious sites, historical buildings and tourist attractions.
- Source: [https://en.wikipedia.org/wiki/District\\_3,\\_Ho\\_Chi\\_Minh\\_City](https://en.wikipedia.org/wiki/District_3,_Ho_Chi_Minh_City)



# 3. OBJECTIVE

- In this project, I will study in details the area classification using Foursquare data and machine learning segmentation and clustering.
- The aim of this project is to segment areas of District 1 and District 3 based on the most common places captured from Foursquare.
- I will determine the similarity or dissimilarirty of both district and classification residential/tourism places/others of area located inside District 1 and District 3 by using segmentation and clustering

## 4. DATA

- Link of data: [https://github.com/vhhlinh/Study\\_data/blob/master/HCMC\\_District.csv](https://github.com/vhhlinh/Study_data/blob/master/HCMC_District.csv)

|   | No. | District code | District name | Ward_name             | Ward code |
|---|-----|---------------|---------------|-----------------------|-----------|
| 0 | 1   | 760           | District 1    | Tan Dinh Ward         | 26734     |
| 1 | 2   | 760           | District 1    | Da Kao Ward           | 26737     |
| 2 | 3   | 760           | District 1    | Ben Nghe Ward         | 26740     |
| 3 | 4   | 760           | District 1    | Ben Thanh Ward        | 26743     |
| 4 | 5   | 760           | District 1    | Nguyen Thai Binh Ward | 26746     |

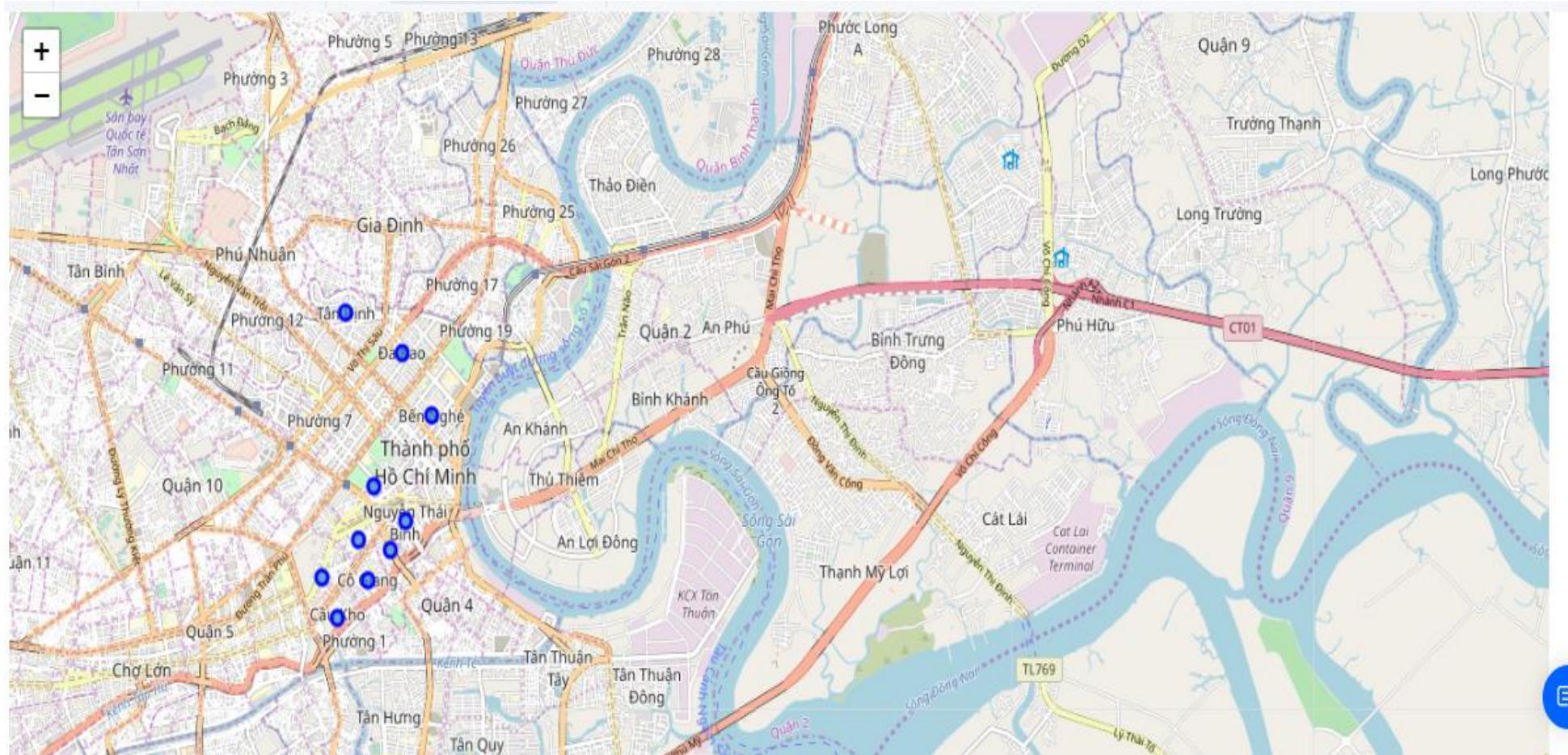
# DATA OF DISTRICT 1

- District 1 has 10 Wards include in table as below:

|   | No. | District code | District name | Ward_name             | Ward code | Latitude  | Longitude  |
|---|-----|---------------|---------------|-----------------------|-----------|-----------|------------|
| 0 | 1   | 760           | District 1    | Tan Dinh Ward         | 26734     | 10.793203 | 106.690203 |
| 1 | 2   | 760           | District 1    | Da Kao Ward           | 26737     | 10.788476 | 106.698318 |
| 2 | 3   | 760           | District 1    | Ben Nghe Ward         | 26740     | 10.781234 | 106.702650 |
| 3 | 4   | 760           | District 1    | Ben Thanh Ward        | 26743     | 10.772866 | 106.694300 |
| 4 | 5   | 760           | District 1    | Nguyen Thai Binh Ward | 26746     | 10.768828 | 106.698797 |
| 5 | 6   | 760           | District 1    | Pham Ngu Lao Ward     | 26749     | 10.766707 | 106.692001 |
| 6 | 7   | 760           | District 1    | Cau Ong Lanh Ward     | 26752     | 10.765459 | 106.696588 |
| 7 | 8   | 760           | District 1    | Co Giang Ward         | 26755     | 10.762010 | 106.693365 |
| 8 | 9   | 760           | District 1    | Nguyen Cu Trinh Ward  | 26758     | 10.762397 | 106.686651 |
| 9 | 10  | 760           | District 1    | Cau Kho Ward          | 26761     | 10.757525 | 106.688900 |



# MAP OF DISTRICT 1



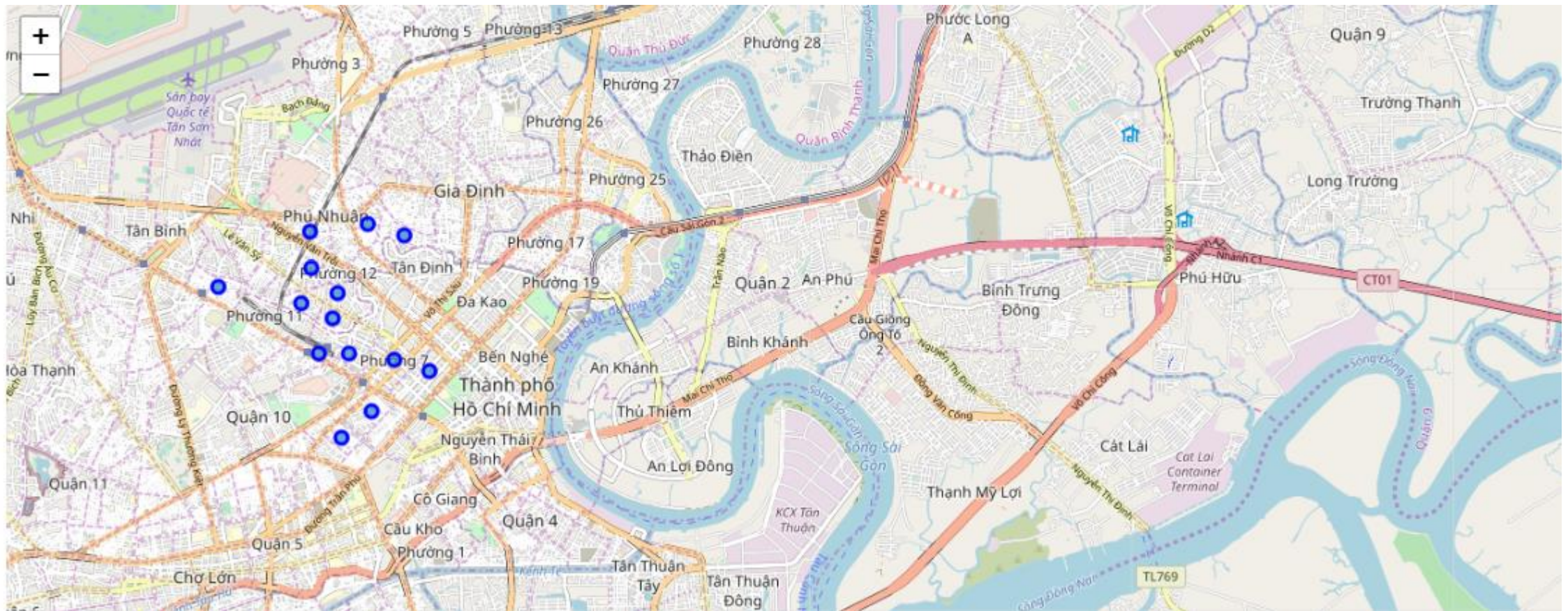
# DATA OF DISTRICT 3

- District 3 has 14 Wards include in table as below:

| No. | District code | District name | Ward_name  | Ward code | Latitude | Longitude |            |
|-----|---------------|---------------|------------|-----------|----------|-----------|------------|
| 0   | 135           | 770           | District 3 | Ward 8    | 27121    | 10.797914 | 106.674852 |
| 1   | 136           | 770           | District 3 | Ward 7    | 27124    | 10.780620 | 106.686400 |
| 2   | 137           | 770           | District 3 | Ward 14   | 27127    | 10.789566 | 106.678634 |
| 3   | 138           | 770           | District 3 | Ward 12   | 27130    | 10.788233 | 106.673674 |
| 4   | 139           | 770           | District 3 | Ward 11   | 27133    | 10.792952 | 106.674998 |
| 5   | 140           | 770           | District 3 | Ward 13   | 27136    | 10.786314 | 106.677848 |
| 6   | 141           | 770           | District 3 | Ward 6    | 27139    | 10.779185 | 106.691148 |
| 7   | 142           | 770           | District 3 | Ward 9    | 27142    | 10.781610 | 106.680207 |
| 8   | 143           | 770           | District 3 | Ward 10   | 27145    | 10.781472 | 106.676094 |
| 9   | 144           | 770           | District 3 | Ward 4    | 27148    | 10.773762 | 106.683225 |
| 10  | 145           | 770           | District 3 | Ward 5    | 27151    | 10.790434 | 106.662371 |
| 11  | 146           | 770           | District 3 | Ward 3    | 27154    | 10.770193 | 106.679057 |
| 12  | 147           | 770           | District 3 | Ward 2    | 27157    | 10.797398 | 106.687658 |
| 13  | 148           | 770           | District 3 | Ward 1    | 27160    | 10.798837 | 106.682654 |



# MAP OF DISTRICT 3



## 4. METHODOLOGY

- After converting addresses into their equivalent latitude and longitude values. I will use the Foursquare API to explore neighborhoods in both District 1 and District 3 with steps as below:
  - Explore function to get the most common venue categories in each neighborhood
  - Use the Folium library to visualize the neighborhoods in District 1 and District 3 and their emerging clusters
  - Use feature to group the neighborhoods into clusters
  - Use K-means clustering algorithm to analyst. And also, the Folium library to visualize the neighborhoods in Kuala Lumpur and Johor Bahru and their emerging clusters.
- Based on dataframe analysis above, I can display the highest number of area within both district.

---

## 4.1 VENUES AT SURROUNDING AREA OF BOTH DISTRICT 1

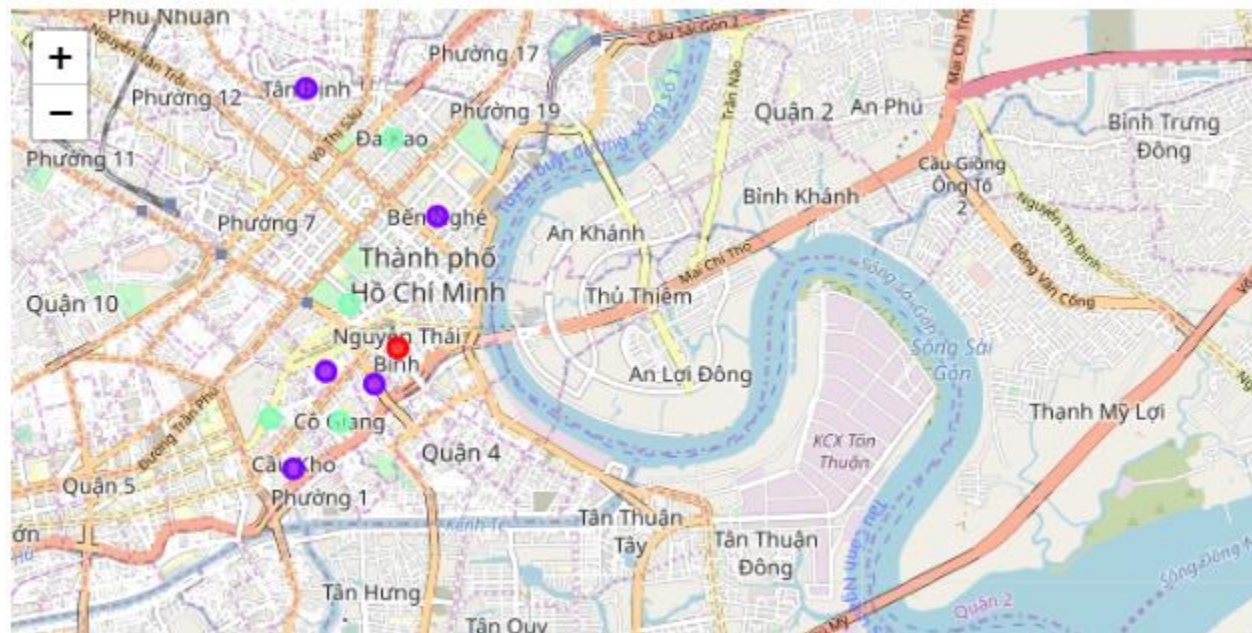
## 4.1 VENUES AT SUROUNDING AREA OF DIST 1

- 34 venues were returned by Foursquare for District 1, sample as below:

|   | name                     | categories                    | lat       | lng        |
|---|--------------------------|-------------------------------|-----------|------------|
| 0 | Cuc Gach                 | Vietnamese Restaurant         | 10.792957 | 106.689020 |
| 1 | Cuc Gach Quan            | Vietnamese Restaurant         | 10.790773 | 106.691795 |
| 2 | Buddha Chay              | Vegetarian / Vegan Restaurant | 10.792762 | 106.688252 |
| 3 | Bánh canh cua 87         | Vietnamese Restaurant         | 10.794697 | 106.690917 |
| 4 | Cơm Tấm Nguyễn Phi Khanh | Breakfast Spot                | 10.791676 | 106.692159 |



# MAP OF DISTRICT 1 SURROUNDING AREA



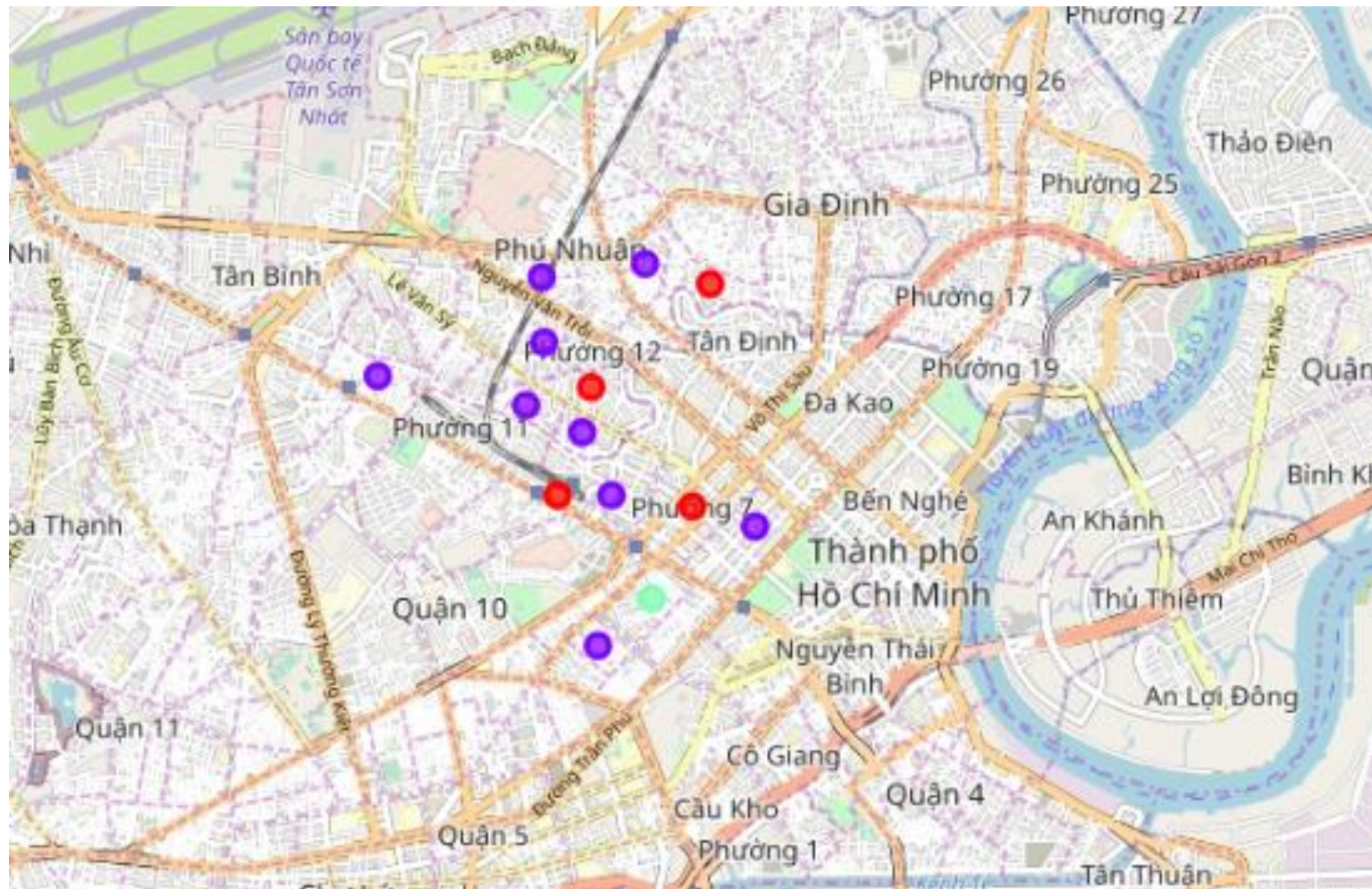


## 4.1 VENUES AT SURROUNDING AREA OF DIST 3

- 32 venues were returned by Foursquare for District 3, sample as below:

|   | name                      | categories           | lat       | lng        |
|---|---------------------------|----------------------|-----------|------------|
| 0 | EASTIN Grand Hotel Saigon | Hotel                | 10.796807 | 106.673363 |
| 1 | Tung Garden               | Cantonese Restaurant | 10.796832 | 106.673386 |
| 2 | The Open Space            | Coffee Shop          | 10.796959 | 106.674827 |
| 3 | The Fig                   | Café                 | 10.797200 | 106.676869 |
| 4 | Yeebo                     | Chinese Restaurant   | 10.795833 | 106.675274 |

# MAP OF DISTRICT 3 SURROUNDING AREA



## 6. DISCUSSION

- Based on cluster for each district above, I believe that classification for each cluster can be done better with calculation of venues categories (most common) in each district in Hochiminh city. Referring to each cluster, I can't determine clearly what represent in each cluster by using Foursquare - Most Common Venue data.
  - Assumed each cluster as follow:
  - Cluster 1: District 1: Tourism/Business place
  - Cluster 2: District 1: Residential
  - Cluster 3: District 1: Mix
  - Cluster 1: District 3: Tourism/Business placeResidential
  - Cluster 2: District 3: Residential
  - Cluster 3: District 3: Mix
- What is lacking at this point is a systematic, quantitative way to identify and distinguish different district and to describe the correlation most common venues as recorded in Foursquare. The reality is however more complex: similar cities might have or might not have similar common venues. A further step in this classification would be to find a method to extract these common venues and integrate the spatial correlations between different of areas or district.



THANK YOU