SEGMENTING HONG KONG MTR STATION AREAS

INTRODUCTION

- On average, more than 60 million tourists visit Hong Kong every year.

- MTR (Mass Transit Railway) is the most commonly used public transportation in Hong Kong.
  - Currently 95 heavy rail stations
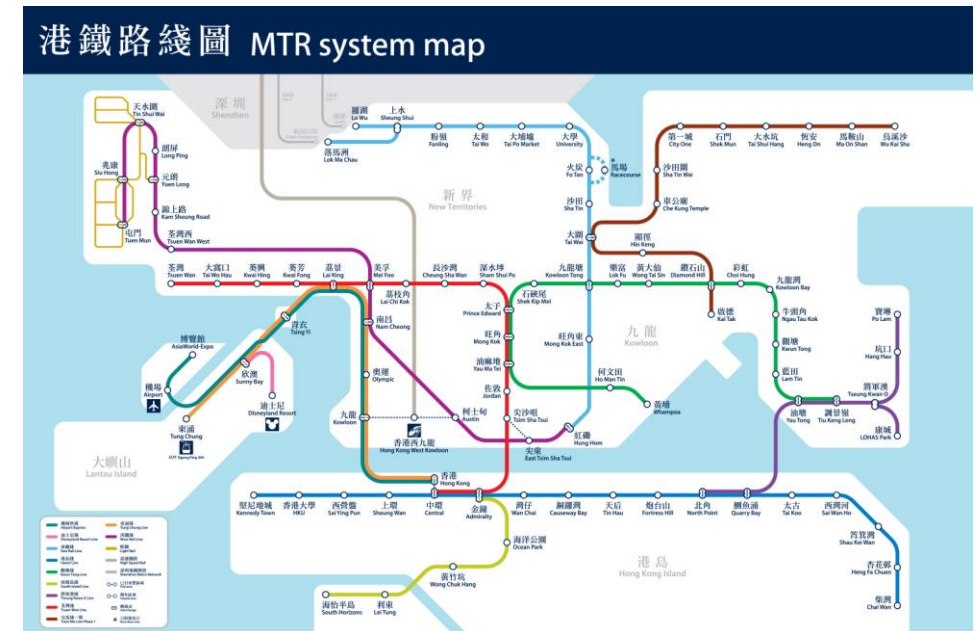 - Daily ridership is around five million

THE CURRENT PROJECT

- Segmenting MTR station areas and exploring the top venue categories in each area in Hong Kong.
  - Segmentation results were interpreted together with population density in different districts of Hong Kong

- Objective: To provide some guidance to tourists about using the MTR system to plan their visit in Hong Kong.

DATA

- A table of MTR heavy rail stations from Wikipedia (https://en.wikipedia.org/wiki/List_of_MTR_stations)

- Some stations are listed for more than once because they are in multiple lines; duplicate rows will be removed.

- Only the columns useful for the analysis (e.g., station name, district) will be retained

GEOLOCATIONS OF MTR STATIONS

- The geolocations (latitudes and longitudes) of the MTR stations were fetched with ArcGIS.

- These geolocation data were merged into the table of MTR stations.

	District	Station	Latitude	Longitude
0	Central and Western District	Admiralty Station	22.278980	114.164820
1	Central and Western District	Central Station	22.281690	114.158740

VENUES IN STATION AREAS

- The Foursquare API was used to explore the station areas.
  - At maximum 200 venues within the 500-meter radius of each station were fetched.

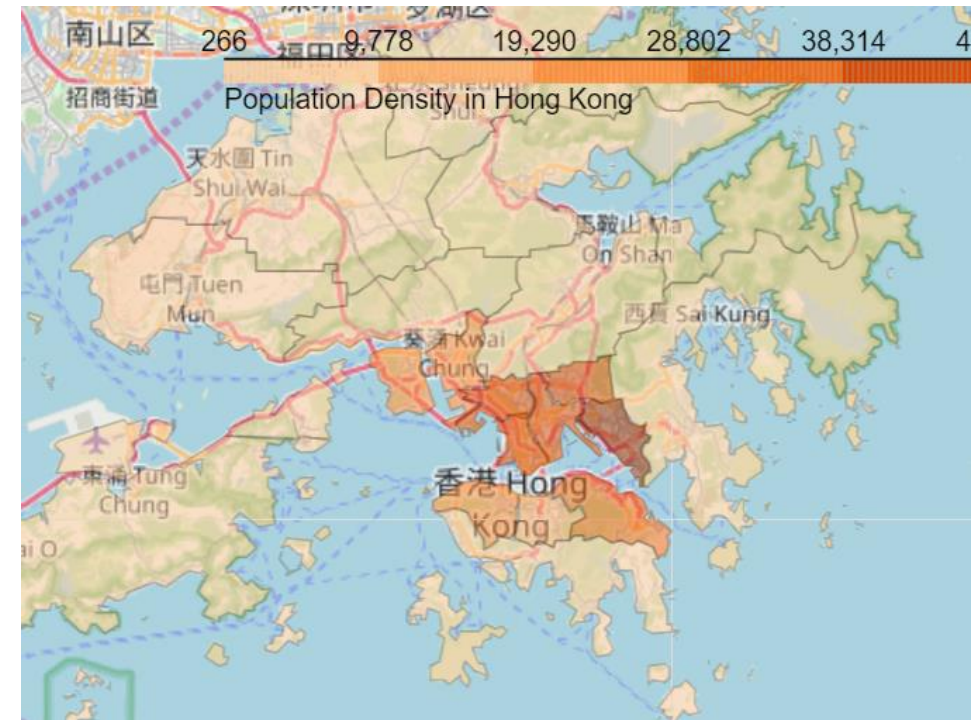|	Station	Station Latitude	Station Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Admiralty Station	22.27898	114.16482	The Upper House (奕居)	22.277499	114.166323	Hotel
1	Admiralty Station	22.27898	114.16482	Island Shangri-La (港島香格里拉大酒店)	22.277193	114.164310	Hotel

LIST OF DISTRICTS IN HONG KONG

- A list of Hong Kong districts was also be scraped from Wikipedia (https://en.wikipedia.org/wiki/Districts_of_Hong_Kong)

- A dataframe was created to show the population densities of the 18 districts in Hong Kong.

	District	Density(/km²)
0	Central and Western	19983.92
1	Eastern	31217.67

GEOLOCATIONS OF HONG KONG DISTRICTS

- The GeoJSON file provided by Esri (http://opendata.esrichina.hk/datasets/eea8ff2f12b145f7b33c4eef4f045513_0) was used to create a choropleth map to visualize the population densities in different Hong Kong districts
  - The clustering results were shown on the choropleth map.
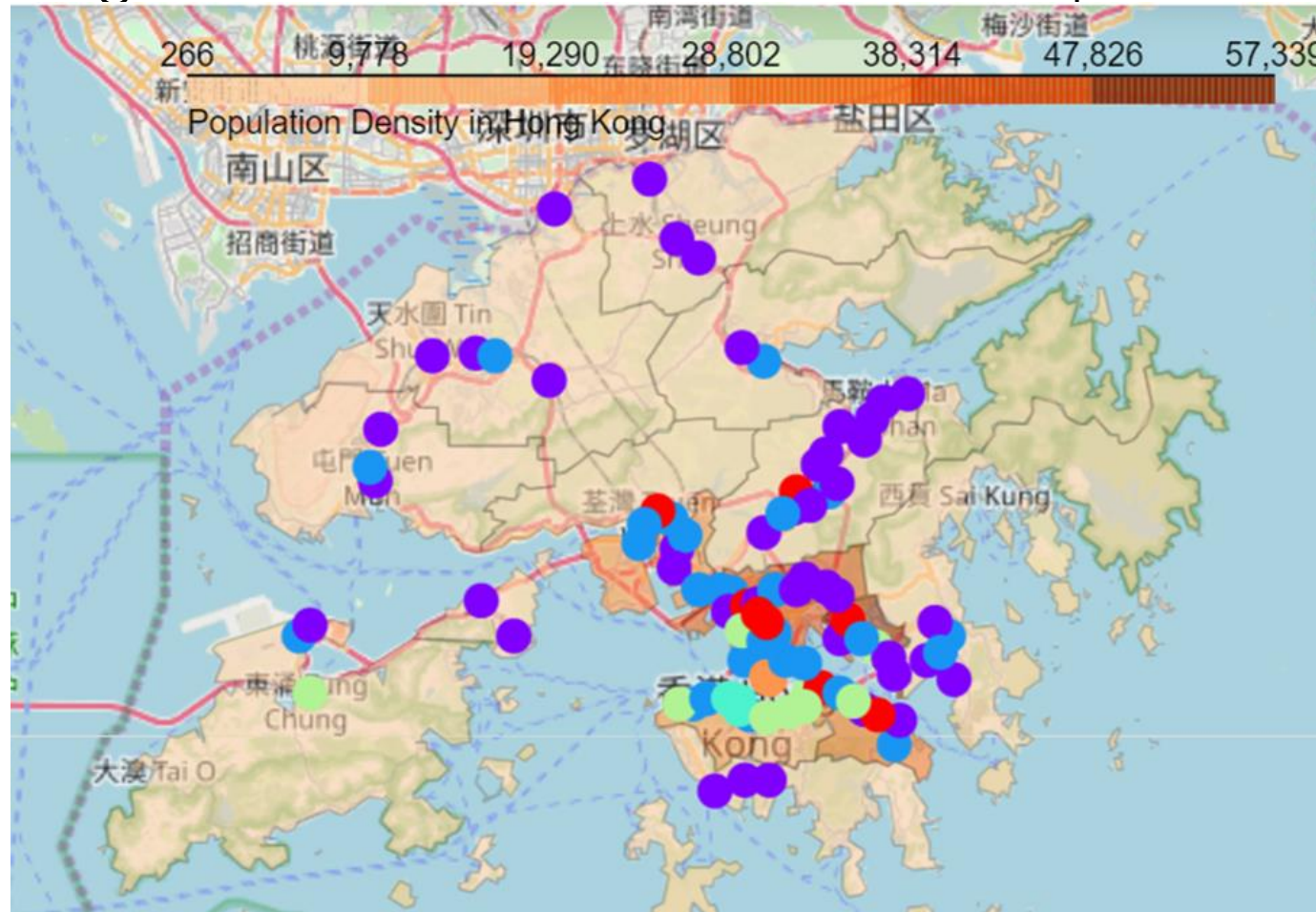
METHODOLOGY

- Cluster analysis was performed to segment the different MTR station areas in Hong Kong. The MTR station areas were clustered and segmented based on the similarities and differences in the venues around them.

- Procedure:
  1. First fetched the geolocations (i.e., latitudes and longitudes) of the MTR stations in Hong Kong;
  2. Used the Foursquare API to explore the venues around each MTR station;
  3. The frequencies of each venue categories around each station were calculated;
  4. **K-means clustering** was performed to segment the MTR station areas using the frequencies of different venue categories;
  5. The interpretations of the k-means clustering analysis were enriched by the additional consideration of population density.
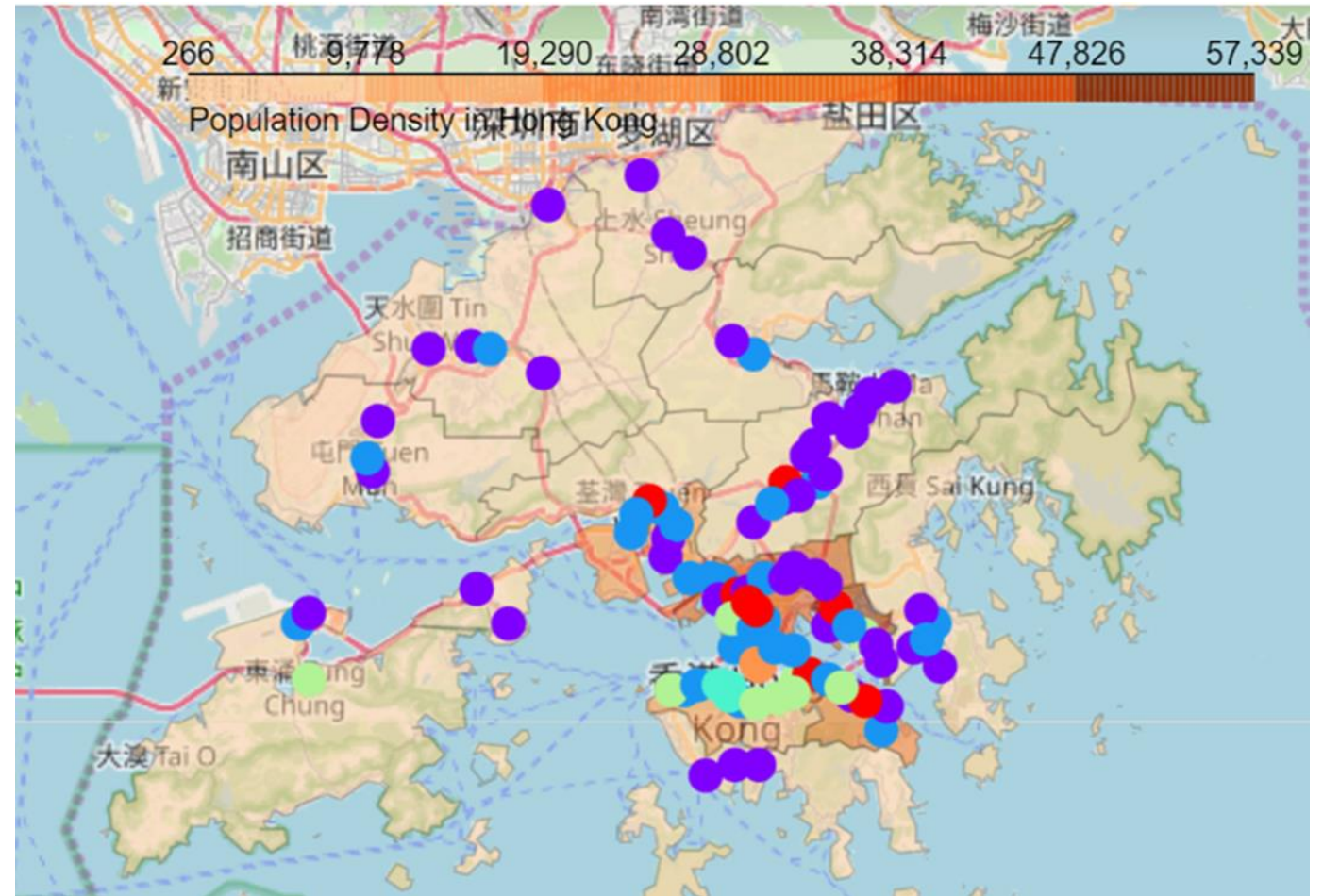
- Six clusters were generated and shown on the choropleth map.

CLUSTER 1 - LOCAL DINING SPOTS

- 10 station areas (red dots in the map)
  - Most in the more densely populated districts of Hong Kong
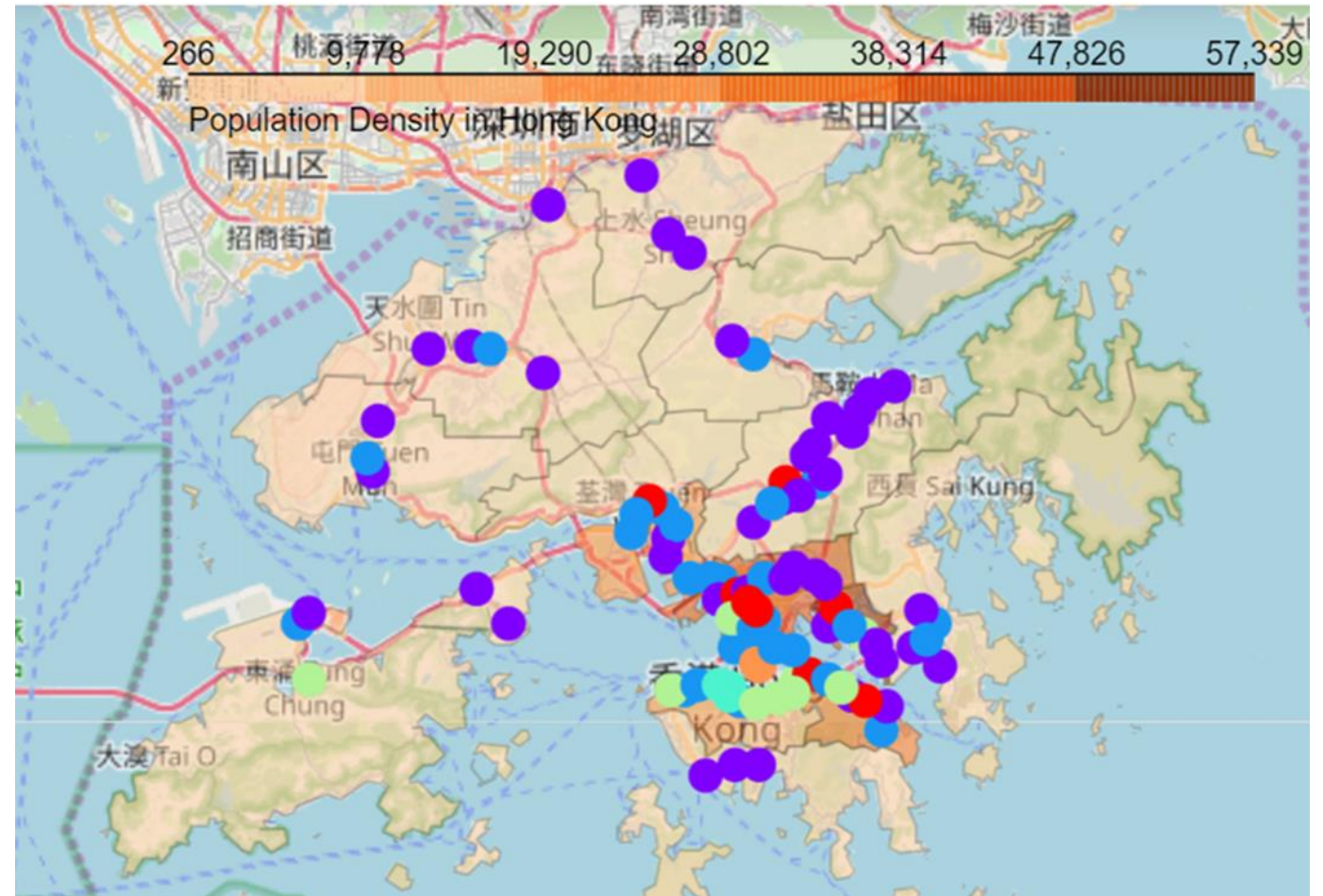
CLUSTER 1 - LOCAL DINING SPOTS

- Top 5 most frequent venue categories

Venue Category	Mean Frequency
Noodle house	3.70
Chinese restaurant	3.40
Dessert soup	3.10
Cha chaan teng (i.e., local diners)	2.40
Shopping mall	2.10

- Mainly restaurants that primarily target local residents

CLUSTER 2 - TRANSIT SPOTS

- 43 station areas (the largest cluster; purple dots in the map)
  - Mostly in medium to low population density districts
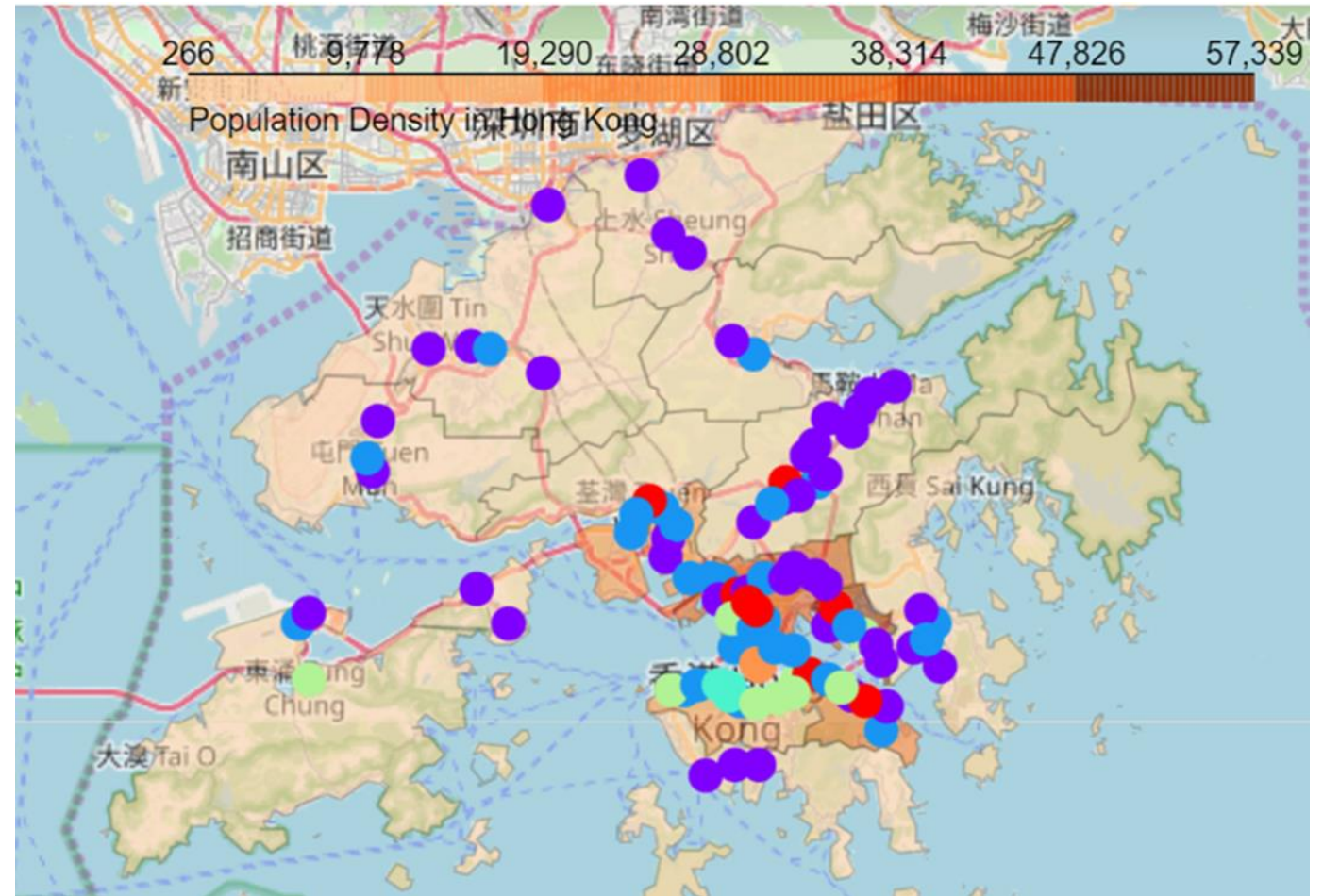
CLUSTER 2 – TRANSIT SPOTS

- Top 5 most frequent venue categories

Venue Category	Mean Frequency
Fast food restaurant	1.26
Chinese restaurant	0.88
Shopping mall	0.60
Coffee shop	0.56
Bus station	0.47

- Mainly targeting customers who are on the go

CLUSTER 3 - LOCAL CAFÉ SPOTS

- 28 station areas (the second largest cluster; blue dots in the map)
  - Mostly in medium to high population density districts
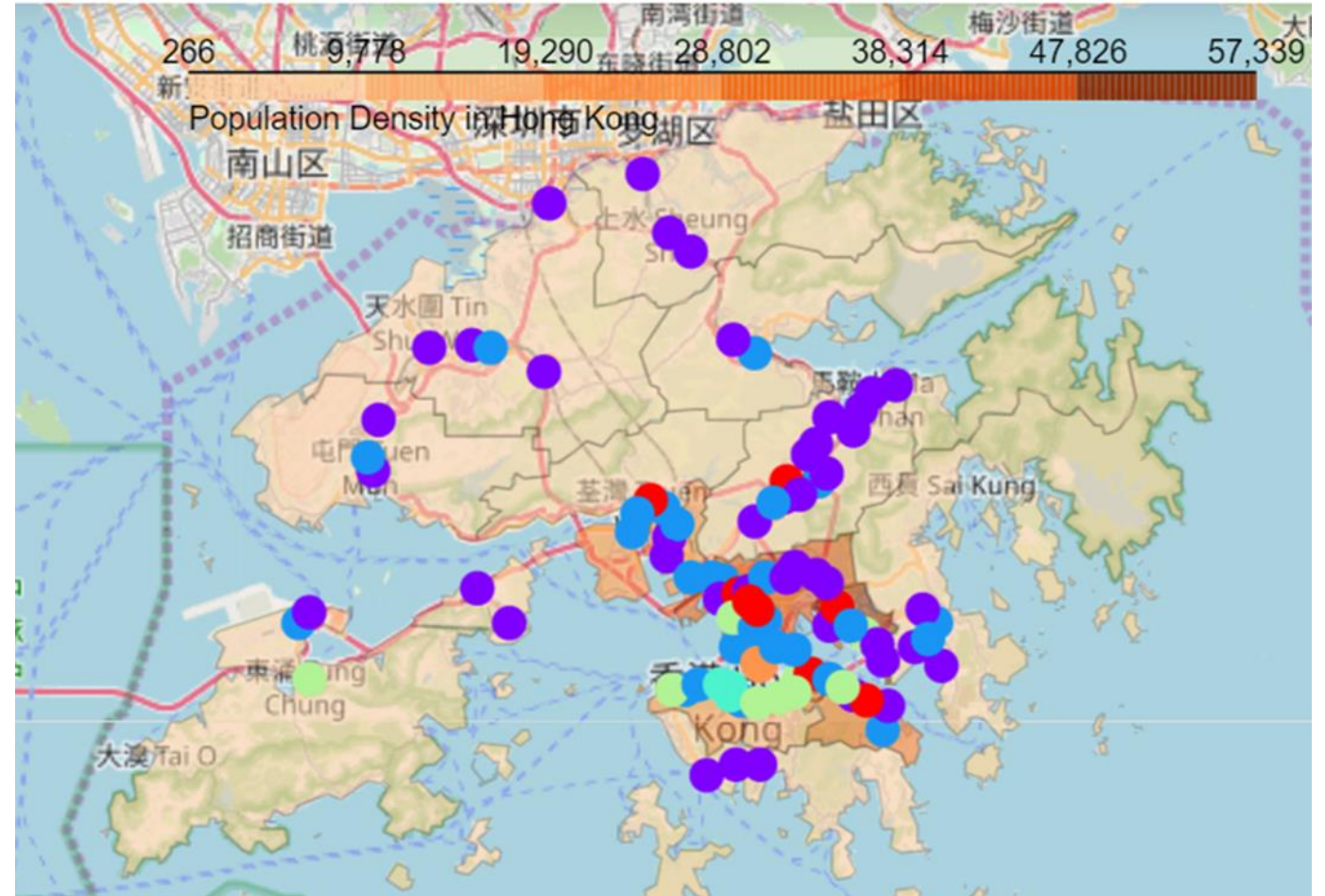
CLUSTER 3 - LOCAL CAFÉ SPOTS

- Top 5 most frequent venue categories

Venue Category	Mean Frequency
Chinese restaurant	2.18
Coffee shop	1.96
Café	1.54
Fast food restaurant	1.36
Cantonese restaurant	1.14

- Mainly targeting local customers
- Similar to Cluster 2 but have more coffee shops

CLUSTER 4 - CBD SOCIAL SPOTS

- 3 station areas (turquoise dots in the map)
  - Close to each other in Hong Kong Island
  - At the heart of the CBD
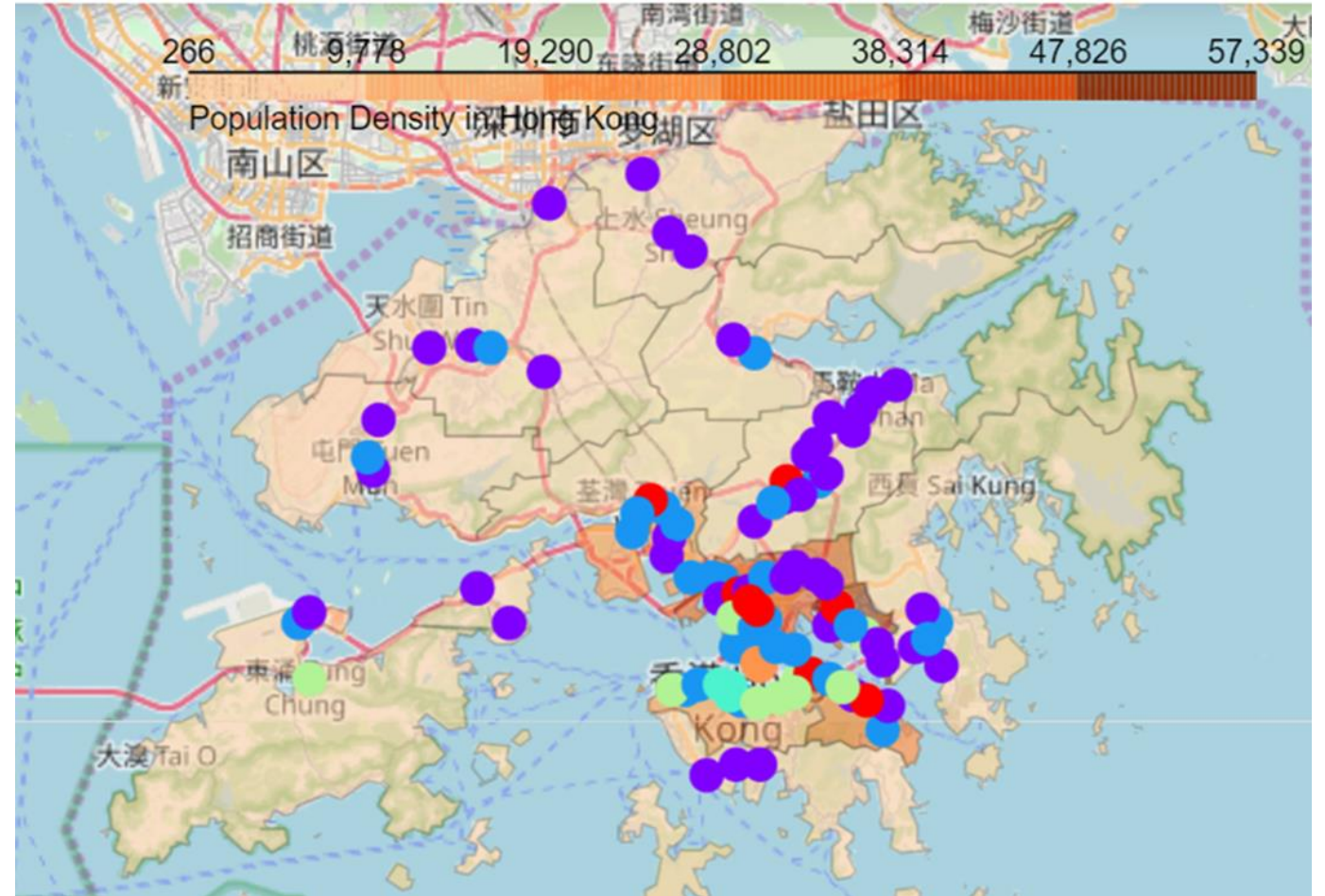
CLUSTER 4 - CBD SOCIAL SPOTS

- Top 5 most frequent venue categories

Venue Category	Mean Frequency
Japanese restaurant	4.67
Coffee shop	4.67
Cocktail bar	4.67
Café	4.00
Bar	4.00

- Quite a few places for people to mingle and socialize

CLUSTER 5 - CONVENIENT ACCOMMODATION SPOTS

- 9 station areas (green dots in the map)
  - Most of which are located on the Hong Kong Island
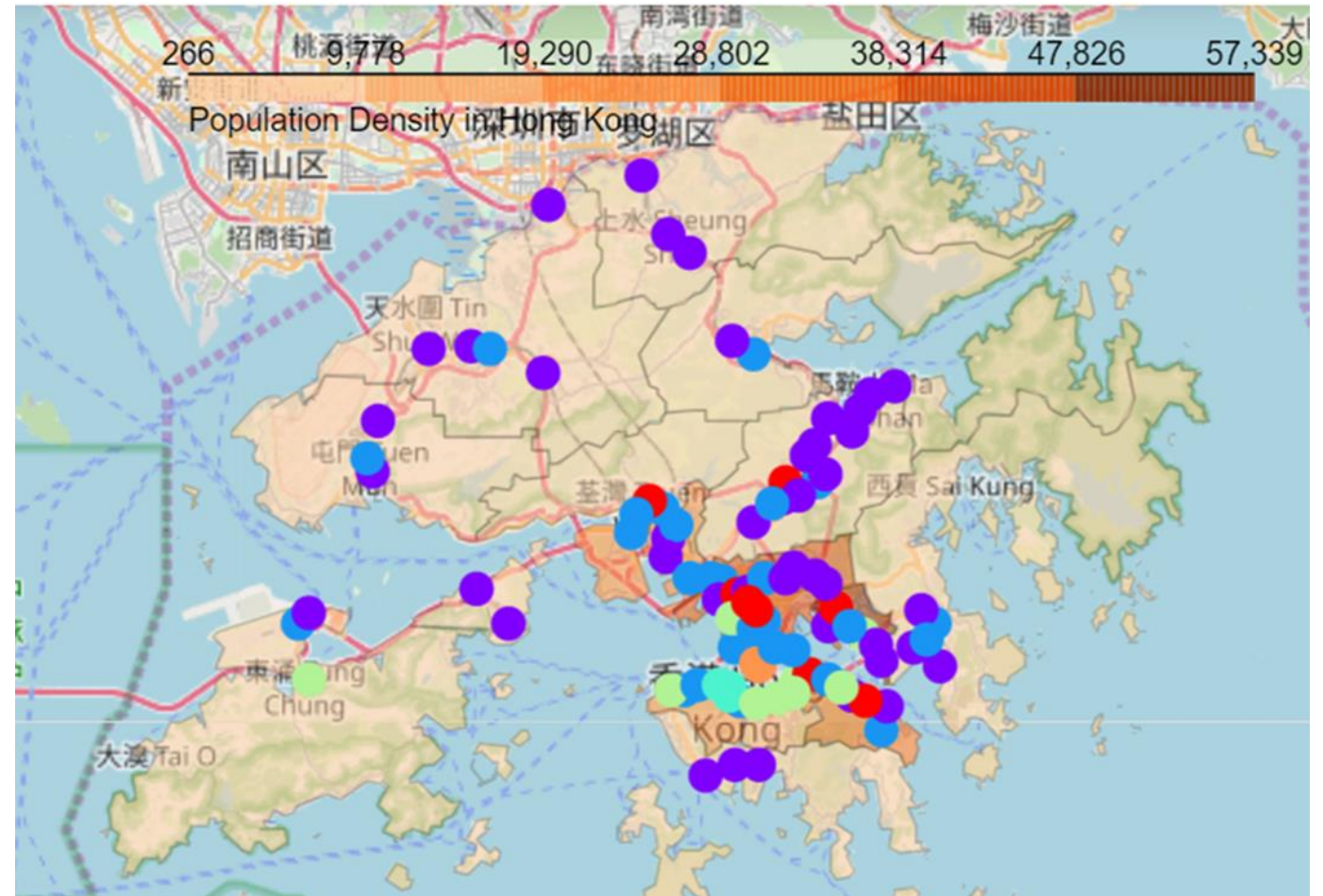
CLUSTER 5 - CONVENIENT ACCOMMODATION SPOTS

- Top 5 most frequent venue categories

Venue Category	Mean Frequency
Coffee shop	4.78
Japanese restaurant	3.78
Café	3.44
Chinese restaurant	3.44
Hotel	2.56

- Similar to Cluster 4 in that both contain quite a few Japanese restaurants and coffee shops
- But Cluster 5 have more hotels

CLUSTER 6 - CENTRAL TOURISTY SPOTS

- 2 station areas (orange dots in the map)
  - Very close to each other in Tsim Sha Tsui in Kowloon on the Victoria Harbor

CLUSTER 6 - CENTRAL TOURISTY SPOTS

- Top 5 most frequent venue categories

Venue Category	Mean Frequency
Hotel	9.00
Japanese restaurant	8.00
Hotel bar	5.00
Steakhouse	4.00
Shopping mall	3.50

- Many hotels and restaurants that target foreign travelers

DISCUSSION

- Six clusters of MTR station areas in Hong Kong, that is, *Local Dining Spots, Transit Spots, Local Café Spots, CBD Social Spots, Convenient Accommodation Spots,* and *Central Touristy Spots*

- **Local Dining Spots:** Many restaurants or eateries that serve local residents
  - Visitors who are interested in authentic local food at affordable prices should visit these station areas.

- **Transit Spots:** Restaurants and cafés that target customers on the go
  - Visitors will find it easy to grab a bite quickly if they happen to visit these areas on the way to other places.

DISCUSSION

- *Local Café Spots:* Many cafés that mainly serve local customers
  - Visitors to these places can enjoy coffee at reasonable prices

- *CBD Social Spots:* More high-end restaurants and cafés
  - Visitors can visit these places if they prefer something fancier

- *Convenient Accommodation Spots*: Many hotels on Hong Kong Island; close to CBD

- *Central Touristy Spots:* Many hotels in Tsim Sha Tsui, a major tourist attraction on the Victoria Harbor (the Kowloon side)

CONCLUSION

- The clustering results will help visitors to Hong Kong to better plan their stay in Hong Kong using the MTR system.

- Future directions: More focused clustering analysis
  - For example, the different MTR station areas can be clustered based on the density of sports facilities or the density of parks in order to provide more tailored recommendations for travelers who are interested in visiting areas with particular characteristics.

Thank you!