

Geospatial Analysis of Hong Kong for Optimal Sportswear Store Location

Author: Preston Chou

Date: February 26, 2021

1. Introduction

This project aims to assist in choosing the optimal location in Hong Kong for a sportswear store. The sportswear market has been growing in Hong Kong due to more active lifestyles, the proliferation of gyms and certain related fashion trends (e.g. athleisure, activewear). Starting a sportswear business seems promising, but assuming that one's business model involves brick-and-mortar stores, a sportswear store's location is a crucial success factor.

In this data analytics project, the primary metric used to assess an area's potential will be its number of fitness, clothing or shopping-related venues. Another metric will be each area's most common venue types. A greater number of venues in an area implies greater interest in fitness lifestyles and/or fashion, and the presence of a "fitness/fashion ecosystem."

The project intends to help existing sportswear companies looking to optimize their current location strategy, new start-ups looking to enter the market, and Hong Kong sportswear market analysts.

2. Data

2.1 Data Sources

This project will utilize the below data sources:

- Hong Kong Constituency Areas GeoJson: Border data for each of Hong Kong's Constituency Areas and corresponding Districts
 - o Sourced: Hong Kong Geodata Store <<https://geodata.gov.hk/gs/view-dataset?uuid=ddd39cbe-5f10-4dbd-96e7-7a4883e37388&l=en&sidx=0>>
- Nominatim Geocoder API: Coordinates (latitude & longitude) of Hong Kong areas
- Foursquare API: Venue data for each of Hong Kong's areas

2.2 Data cleaning & processing

2.2.1 Hong Kong Constituency Areas GeoJson

The Hong Kong Constituency Areas GeoJson provides a lot of useful data, including each of Hong Kong's Constituency Areas in 2019, the Districts they belong to, and the coordinates for their borders.

However, a disadvantage for using the Constituency Areas as the basis for analyzing Hong Kong is that some of the areas are used only in the context of being a constituency, rather than carrying significance in an economic or demographic sense. The names of some Constituency Areas may be obscure for our Nominatim Geocoder API, which is responsible for finding coordinates given the name of the area. We will find some ways to handle these issues.

We will extract the Constituency Areas and Districts from the file and put them into the columns of a new Pandas dataframe:

	Hong Kong Constituency Areas	Hong Kong Constituency District
0	Sai Wan Ho	Eastern District
1	Nam Fung	Eastern District
2	Kornhill Garden	Eastern District
3	Mount Parker	Eastern District
4	Fortress Hill	Eastern District
5	A Kung Ngam	Eastern District
6	Kam Ping	Eastern District
7	Fort Street	Eastern District
8	Tanner	Eastern District
9	Lei King Wan	Eastern District
10	Tai Koo Shing East	Eastern District

2.2.2 Nominatim Geocoder API Coordinate Data

The next step is to acquire the latitude and longitude coordinates for each Constitutional Areas. The Geocoder API returns the most accurate coordinates when the input address includes the constituency area, followed by the district and the city ('Hong Kong'). However, this input format yields more 'None' values, which we need to deal with later on. Alternatively, the input address can only include the constituency area and the city. This input format yields far less 'None' values, but the accuracy of the coordinates is poor, so let's go with the first format for now. We will then append the latitude and longitudes to the dataframe:

	Hong Kong Constituency Areas	Hong Kong Constituency District	Latitude	Longitude
0	Sai Wan Ho	Eastern District	22.2824	114.222
1	Nam Fung	Eastern District	22.4427	114.026
2	Kornhill Garden	Eastern District	22.2853	114.217
3	Mount Parker	Eastern District	None	None
4	Fortress Hill	Eastern District	22.2876	114.193
5	A Kung Ngam	Eastern District	22.2845	114.235
6	Kam Ping	Eastern District	None	None
7	Fort Street	Eastern District	22.2895	114.196
8	Tanner	Eastern District	22.2903	114.197
9	Lei King Wan	Eastern District	22.2837	114.223
10	Tai Koo Shing East	Eastern District	None	None

The locations for which the Geocoder API does not return coordinates can be split into three main categories.

The 1st category of these locations includes the names that contain very specific sub-area terms (e.g. North, South, East, West, Upper, Lower). For these instances, we can eliminate the sub-area terms and consolidate the names to one general name (if its corresponding sub-area location is located, then we simply delete the unidentifiable location).

The 2nd category conjoin two or more names (via commas or ampersands). We can form separate names for this category.

The 3rd (most overwhelming but most numerous) category includes names that are simply too obscure for the Geocoder API. Many of these names are only used in the context of being constituencies and lack geographically significance in business terms. We can try to input their address to the Geocoder API as comprising of the Constituency Area name followed by 'Hong Kong'.

But given the unreliability of the API with this input format, we will check to see if the coordinate is contained within the area boundaries as given by the GeoJson file. We can simply eliminate the rows that are outside the boundaries since most (if not all) of them are extremely small and closely neighbor other more significant and identifiable areas.

2.2.3 Foursquare API Venue Data

For each Constituency Area, we want to know which venues exist in the vicinity of the Area's coordinates. We will acquire this venue data through the FourSquare API. We will input the names and coordinates of each Area into the API, acquire up to 300 venues from a 500m radius. We create a new dataframe with a row for each venue.

Hong Kong Constituency Areas	Constituency Area Latitude	Constituency Area Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	
0	Sai Wan Ho	22.282446	114.221506	Little Kitchen	22.281781	114.221516	French Restaurant
1	Sai Wan Ho	22.282446	114.221506	海南雞專門店	22.283032	114.221937	Hainan Restaurant
2	Sai Wan Ho	22.282446	114.221506	Aldrich Bay Park (愛秩序灣公園)	22.282663	114.224877	Park
3	Sai Wan Ho	22.282446	114.221506	華記士多	22.283214	114.220707	Cantonese Restaurant
4	Sai Wan Ho	22.282446	114.221506	Ashoka Indian Restaurant (皇子印度餐廳)	22.280257	114.224753	Indian Restaurant
5	Sai Wan Ho	22.282446	114.221506	Tai Hing (太興)	22.283081	114.222627	Hong Kong Restaurant

Based on this dataframe, we will produce a new dataframe with dummy values for the venue categories. Scanning through the venue categories, we will filter by only including categories with fitness, clothing or shopping-related keywords that indicate relevancy for our objective – opening a sportswear store. The chosen keywords are:

- Sports
 - Sport
 - Sporting
 - Fitness
 - Gym
 - Studio
 - Field
 - Pool
 - Martial art
 - Court
 - Clothing
 - Fashion
 - Outdoor
 - Shoes
 - Shopping

We then add a column listing the name of the Constituency Area, group the dummy values by sum according to the Area names. Additionally, we add a column listing the Total number of venues for each Constituency Area – this total value will be useful for our choropleth map.

3. Data Analysis & Results

Our data analysis phase has 2 objectives. The 1st objective is determining which areas in Hong Kong have greater numbers of fitness, clothing or shopping-related venues. Areas with higher numbers of venues are more likely to be promising sportswear store locations, since they probably have relevant economic-activity and higher consumer demand.

The 2nd objective is understanding the areas' venue-makeup. For example, some areas may have more gyms, others may have more clothing stores, while others may have many shopping malls. This provides another dimension to assess the suitability of an area.

The 1st objective is more straightforward, since we already have 'Total' venue numbers for each area. For the 2nd objective, we will employ K-means clustering to categorize the areas by their venue makeup. For this purpose, we will identify the top 10 most common venue categories for each area.

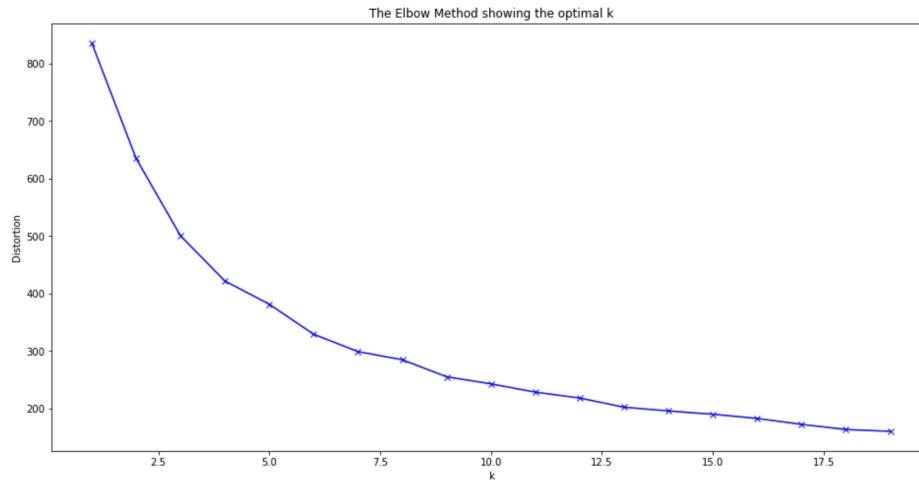
Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0 A Kung Ngam	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center
1 Aberdeen	Athletics & Sports	Shopping Mall	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym
2 Aldrich Bay	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center
3 Allway	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center
4 Ap Lei Chau Estate	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym

3.1 K-Means Clustering

To conduct K-Means Clustering analysis (through the Scikit-learn library), we need to decide on an optimal number of clusters – the K-value.

3.1.1 Elbow method

We can determine the optimal number of clusters through a couple methods, the first being the Elbow Method. In the Elbow Method, we run K-Means clustering for a range of K-values. For each K-value, we find the average of the squared distances for each point from the center of its selected cluster. In the plot, we choose the K-value after which the line appears to descend in a more linear pattern.



Unfortunately, the curve does not have a very sharp elbow (an 'ideal' elbow) at which we can easily choose a K-value. However, the line begins to resemble in a straight line in the 10-15 range. We should use another method to better decide on the precise K-value to use.

3.1.2 Silhouette Analysis

The second method is the Silhouette Analysis, also acquired through the Scikit-learn library. This method measures the degree of separation of the clusters from each other. We calculate a score for each sample point, reflecting its distance from other clusters.

The scoring range runs from -1 to +1. -1 implies that the sample point is much closer to another cluster and is therefore incorrectly assigned. 0 implies that the sample point is right on the border between 2 clusters. +1 implies that the sample point is very close to its own assigned cluster. So the higher the score, the more accurate the clustering!

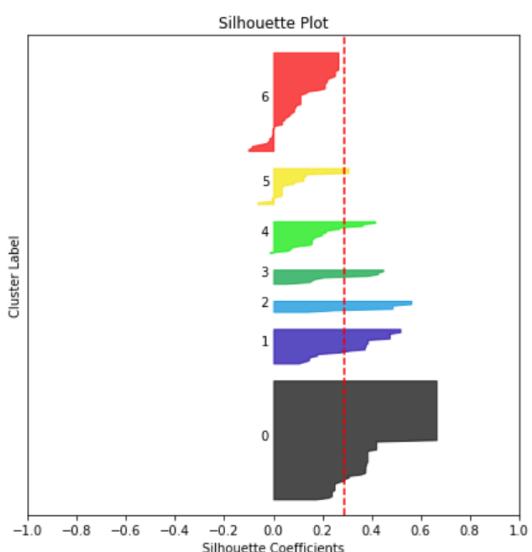
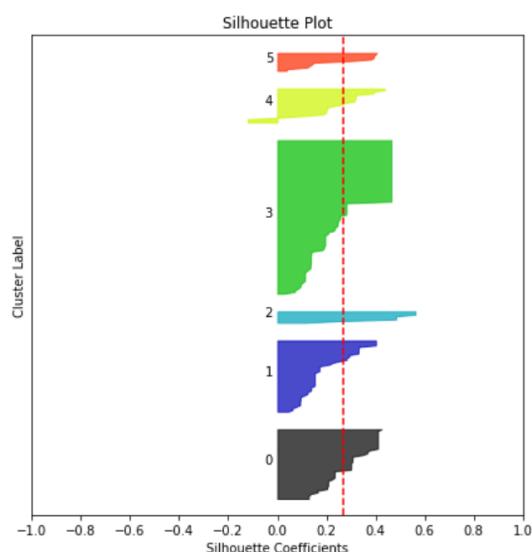
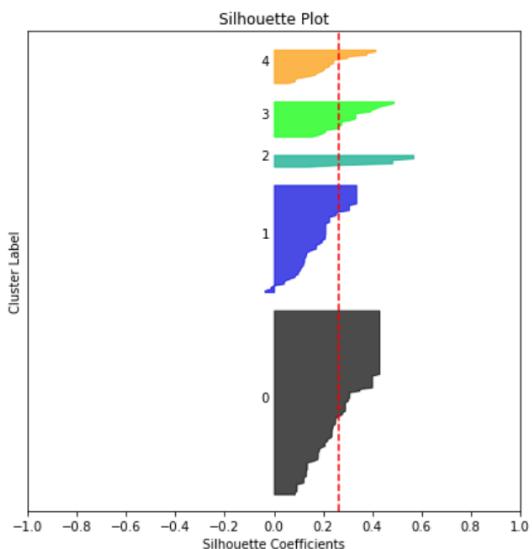
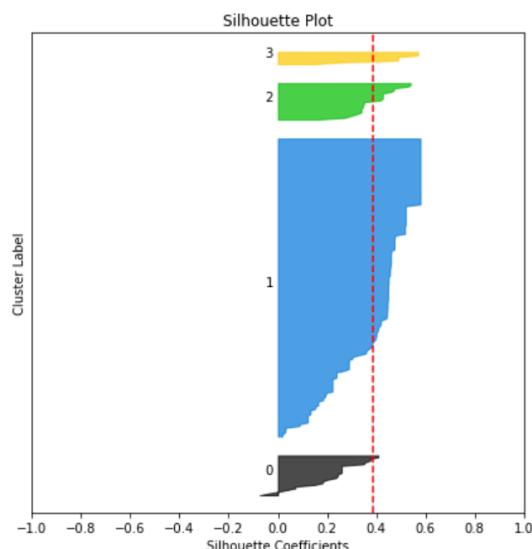
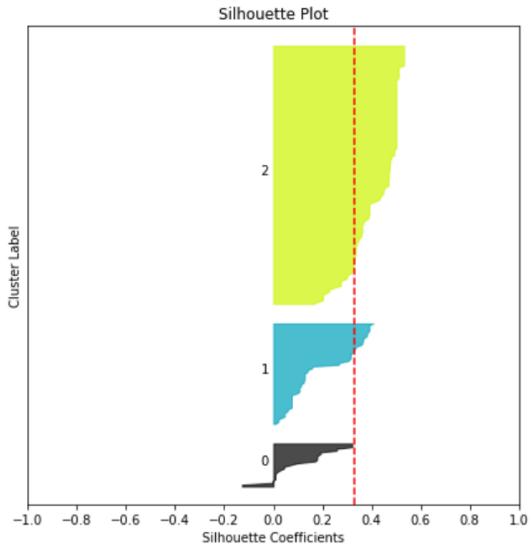
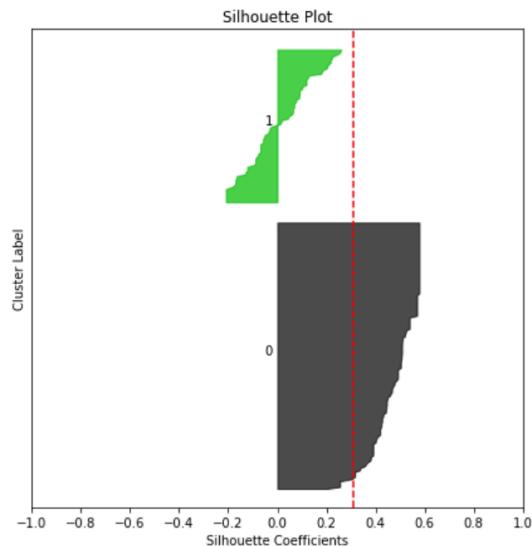
We will use a range of K-values from 2 to 15 (since the elbow curve shows the curve resembling a straight line in the 10-15 range). Each cluster has its own silhouette coefficient, and an average silhouette score is calculated for a given K-value.

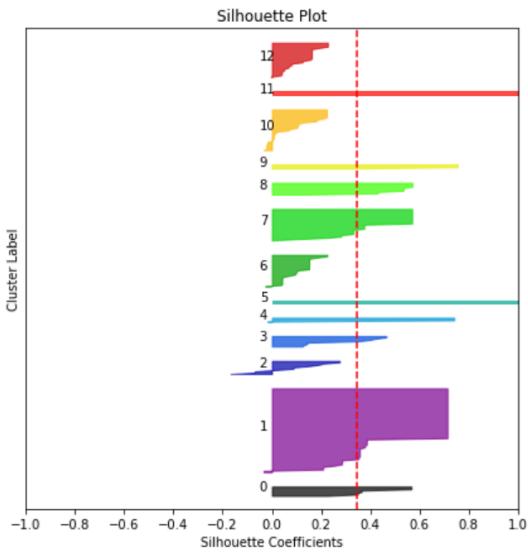
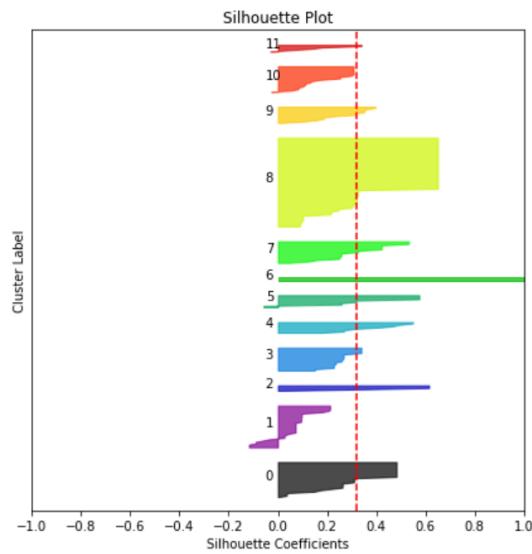
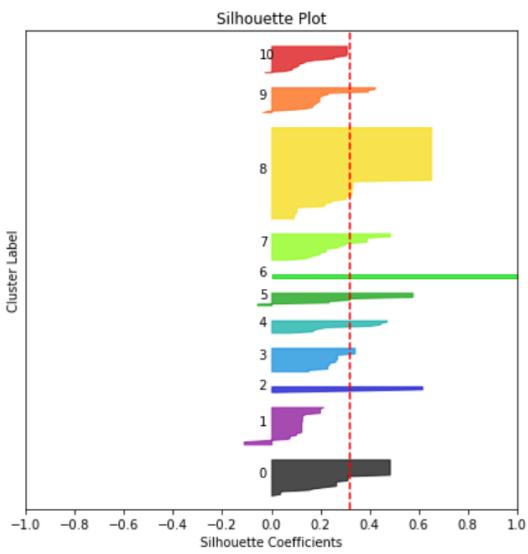
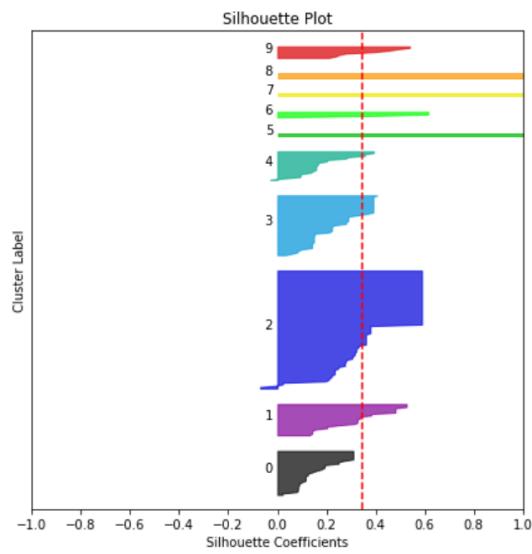
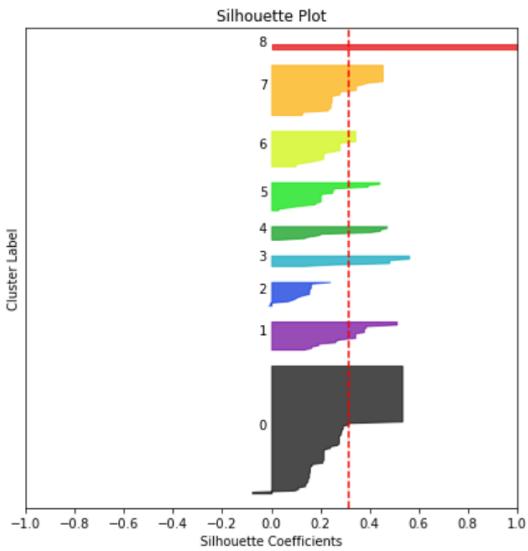
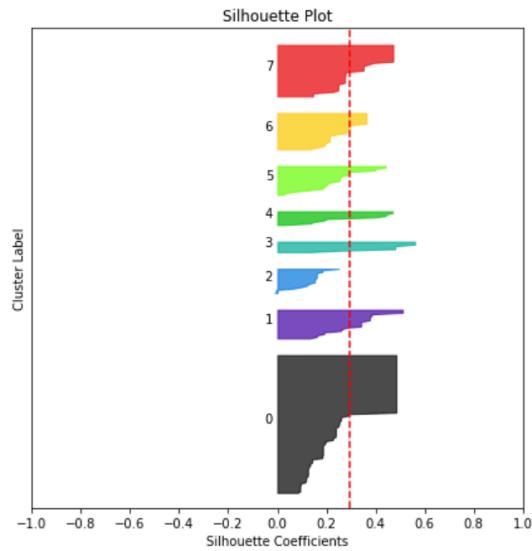
```

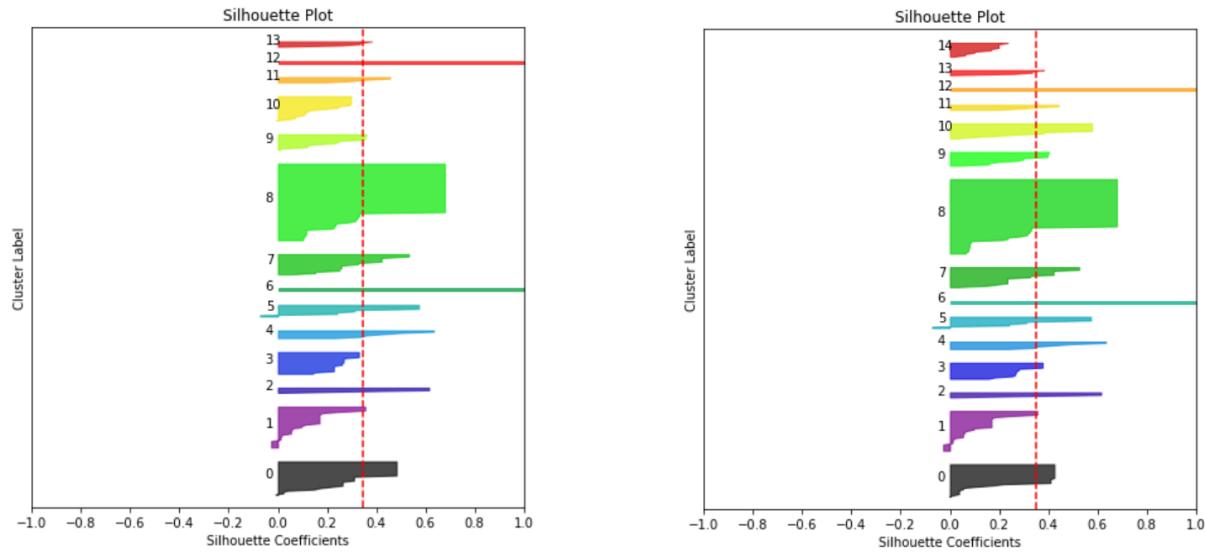
For n_clusters = 2 Average silhouette_score = 0.30989660701186794
For n_clusters = 3 Average silhouette_score = 0.3324970097497605
For n_clusters = 4 Average silhouette_score = 0.3856071570985365
For n_clusters = 5 Average silhouette_score = 0.26582083042547267
For n_clusters = 6 Average silhouette_score = 0.2684904915157421
For n_clusters = 7 Average silhouette_score = 0.2917072742360841
For n_clusters = 8 Average silhouette_score = 0.2941787921071116
For n_clusters = 9 Average silhouette_score = 0.31486471614870215
For n_clusters = 10 Average silhouette_score = 0.34565090362621814
For n_clusters = 11 Average silhouette_score = 0.32243710050245344
For n_clusters = 12 Average silhouette_score = 0.3218181470577988
For n_clusters = 13 Average silhouette_score = 0.3463772608405005
For n_clusters = 14 Average silhouette_score = 0.34345664554056665
For n_clusters = 15 Average silhouette_score = 0.3491730880071366

```

We will also produce a silhouette plot that shows the silhouette coefficient, and the thickness of the plot reflects the size of the cluster size.







A K-value of 15 offers a high silhouette score, but we are also concerned with the risk of overfitting. 13 clusters seems to also yield a good silhouette score, so we will use K-Means Clustering to segment the areas into 13 different clusters.

3.1.2 K-Means Clustering Analysis

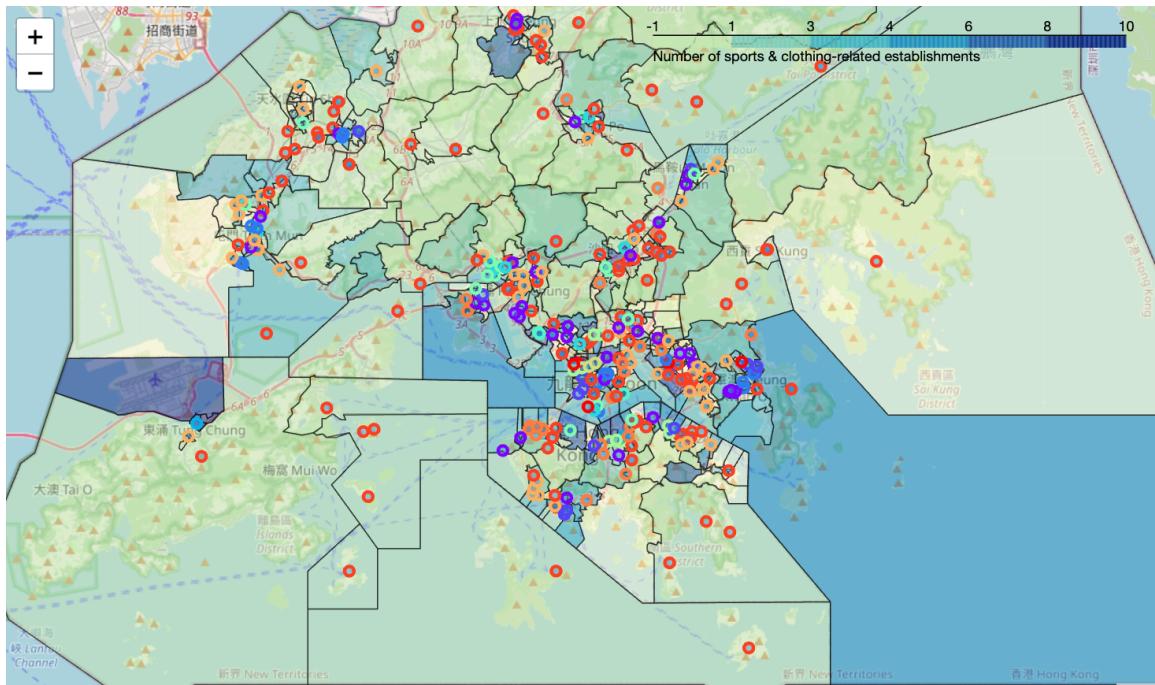
Using the K-Means Clustering algorithm, we assign a cluster label (from 1-13) to each Constituency Area that has a non-0 value for its total number of venues. For Constituency Areas that have 0 relevant venues, we will assign a cluster label of -1.

	Hong Kong Constituency Areas	Hong Kong Constituency Areas for Geocoder	Hong Kong Constituency District	Latitude	Longitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total	Cluster
0	Sai Wan Ho	Sai Wan Ho	Eastern District	22.282446	114.221506	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center	0.0	-1
1	Nam Fung	Nam Fung	Eastern District	22.442702	114.026020	Shopping Mall	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	5.0	3
2	Kornhill Garden	Kornhill Garden	Eastern District	22.285267	114.216780	Clothing Store	Shopping Mall	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	3.0	2
3	Fortress Hill	Fortress Hill	Eastern District	22.287563	114.193315	Gym	Dance Studio	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Gym / Fitness Center	3.0	11

3.2 Geospatial Analysis

3.2.1 Map generation

We present the results of our analysis through maps via Folium. With choropleth maps, we will use shadings of blue to indicate different ranges of venue numbers in each area (darker meaning more numerous). We will then use circle markers that shows each area's name, district and cluster label. The circle marker's color corresponds to its cluster.



3.2.2 Cluster observation

We will filter the dataframe to only include areas belonging to each separate cluster. This way, we can observe the most important and discriminating features of each cluster.

Cluster 1

	Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
23	Healthy Village	Athletics & Sports	Climbing Gym	Shopping Mall	Gym	Gym Pool	Boxing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	4.0
27	Hiu Lai	Athletics & Sports	Shopping Mall	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	2.0
34	Jordan Valley	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center	1.0
35	Shun Tin	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center	1.0
68	San Hui	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	Gym / Fitness Center	1.0

Cluster 2

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
2 Kornhill Garden	Clothing Store	Shopping Mall	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	3.0
26 Upper Ngau Tau Kok Estate	Shopping Mall	Athletics & Sports	Clothing Store	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	Gym	4.0
84 Charming	Clothing Store	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	Gym	3.0
96 Tai Nan	Clothing Store	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	Gym	2.0
99 Kowloon Station	Clothing Store	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	Gym	2.0
108 Tsing Yi South	Shopping Mall	Athletics & Sports	Sporting Goods Shop	Clothing Store	Gym / Fitness Center	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	5.0

Cluster 3

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
1 Nam Fung	Shopping Mall	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	5.0
15 Lok Hong	Shopping Mall	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	5.0
33 Kowloon Bay	Athletics & Sports	Shopping Mall	Clothing Store	Gym / Fitness Center	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
70 Tuen Mun Rural	Shopping Mall	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	4.0
74 Yan Tin	Shopping Mall	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	4.0
85 East Tsim Sha Tsui & King's Park	Shopping Mall	Athletics & Sports	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	3.0

Cluster 4

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
45 Tung Chung Central	Sporting Goods Shop	Clothing Store	Yoga Studio	Gym / Fitness Center	Shopping Mall	Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	10.0
48 Tung Chung North	Sporting Goods Shop	Clothing Store	Yoga Studio	Gym / Fitness Center	Shopping Mall	Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	10.0
79 Tung Chung South	Sporting Goods Shop	Clothing Store	Yoga Studio	Gym / Fitness Center	Shopping Mall	Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	10.0
327 Victoria Park	Sporting Goods Shop	Clothing Store	Martial Arts School	Pool	Gym / Fitness Center	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	7.0

Cluster 5

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
60 Tuen Mun Town Centre	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	4.0
148 Sha Tin Town Centre	Shopping Mall	Clothing Store	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	Gym	7.0
207 Un Chau	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	5.0
229 Prince	Shopping Mall	Sporting Goods Shop	Clothing Store	Yoga Studio	Gym / Fitness Center	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	7.0
244 Tai Po Central	Shopping Mall	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	5.0
345 Tsuen Wan South	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	5.0
346 Fuk Loi	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	6.0
358 Luk Yeung	Shopping Mall	Pool	Yoga Studio	Gym / Fitness Center	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	5.0

Cluster 6

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
170 Whampoa West	Gym / Fitness Center	Shopping Mall	Pool	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	3.0
195 Lai Kok	Gym / Fitness Center	Shopping Mall	Pool	Tennis Court	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	5.0
199 Mei Foo South	Gym / Fitness Center	Shopping Mall	Pool	Tennis Court	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	5.0
202 Mei Foo Central	Gym / Fitness Center	Shopping Mall	Pool	Tennis Court	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	5.0
216 Whampoa East	Gym / Fitness Center	Shopping Mall	Pool	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	3.0
219 Mei Foo North	Gym / Fitness Center	Shopping Mall	Pool	Tennis Court	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	5.0

Cluster 7

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
10 City Garden	Shopping Mall	Gym / Fitness Center	Sporting Goods Shop	Gym	Yoga Studio	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	5.0
65 Kin Sang	Shopping Mall	Sports Club	Shopping Plaza	Yoga Studio	Gym / Fitness Center	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	4.0
80 Tsim Sha Tsui Central	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	2.0
87 Tsim Sha Tsui West	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	2.0
110 Cheung On	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	2.0
140 Chui Tin	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	2.0
147 Tai Wai	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	2.0
169 Ma On Shan Town Centre	Shopping Mall	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	3.0

Cluster 8

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
8 Tai Koo Shing East	Clothing Store	Gym / Fitness Center	Yoga Studio	Shopping Mall	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
9 Tai Koo Shing West	Clothing Store	Gym / Fitness Center	Yoga Studio	Shopping Mall	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
20 Kornhill	Clothing Store	Shopping Mall	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	4.0
24 Quarry Bay	Clothing Store	Gym / Fitness Center	Yoga Studio	Shopping Mall	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
90 Cherry	Clothing Store	Shopping Mall	Gym / Fitness Center	Sporting Goods Shop	Yoga Studio	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0

Cluster 9

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
89 Mong Kok West	Sporting Goods Shop	Shopping Mall	Athletics & Sports	Gym / Fitness Center	Outdoor Supply Store	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	10.0
91 Mong Kok South	Sporting Goods Shop	Shopping Mall	Athletics & Sports	Gym / Fitness Center	Outdoor Supply Store	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	10.0
95 Mong Kok North	Sporting Goods Shop	Shopping Mall	Athletics & Sports	Gym / Fitness Center	Outdoor Supply Store	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	10.0
97 Mong Kok East	Sporting Goods Shop	Athletics & Sports	Clothing Store	Shopping Mall	Outdoor Supply Store	Gym / Fitness Center	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	7.0

Cluster 10

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
12 Yue Wan	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
16 Upper Yiu Tung	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
19 Heng Fa Chuen	Sporting Goods Shop	Shopping Mall	Yoga Studio	Gym / Fitness Center	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	2.0
21 Sau Mau Ping Central	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
22 Lam Tin	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
30 Sau Mau Ping North	Shopping Mall	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0

Cluster 11

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
3 Fortress Hill	Gym	Dance Studio	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Gym / Fitness Center	3.0
36 Kai Yip	Paintball Field	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
58 Yau Oi North	Soccer Field	Pool	Yoga Studio	Gym / Fitness Center	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	2.0
101 Cheung Hong	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
112 Cheung Hang	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	Gym	1.0
139 Keng Hau	Pool	Gym	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	Clothing Store	College Gym	Cycle Studio	Dance Studio	2.0

Cluster 12

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
32 To Tai	Shopping Mall	Clothing Store	Sporting Goods Shop	Soccer Field	Climbing Gym	Pool	Yoga Studio	Gym	Boxing Gym	College Gym	8.0
50 Yuet Wah	Clothing Store	Gym	Sporting Goods Shop	Soccer Field	Shopping Mall	Pool	Yoga Studio	Boxing Gym	Climbing Gym	College Gym	7.0
53 Po Lok	Shopping Mall	Clothing Store	Sporting Goods Shop	Soccer Field	Climbing Gym	Pool	Yoga Studio	Gym	Boxing Gym	College Gym	8.0
57 Kwun Tong Central	Clothing Store	Sporting Goods Shop	Soccer Field	Climbing Gym	Shopping Mall	Gym	Yoga Studio	Gym / Fitness Center	Boxing Gym	College Gym	7.0
262 Wah Ming	Clothing Store	Sporting Goods Shop	Climbing Gym	Shopping Mall	Paintball Field	Gym	Yoga Studio	Gym / Fitness Center	Boxing Gym	College Gym	7.0

Cluster 13

Hong Kong Constituency Areas	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Total
13 Fei Tsui	Shopping Mall	Clothing Store	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	7.0
31 Lower Ngau Tau Kok Estate	Shopping Mall	Athletics & Sports	Clothing Store	Gym / Fitness Center	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
92 Tai Kok Tsui South	Shopping Mall	Clothing Store	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
94 Tai Kok Tsui North	Shopping Mall	Clothing Store	Gym / Fitness Center	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	6.0
127 Po Lam	Shopping Mall	Clothing Store	Yoga Studio	Gym Pool	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	Gym	4.0
360 Kwun Lung	Shopping Mall	Clothing Store	Sporting Goods Shop	Yoga Studio	Gym / Fitness Center	Boxing Gym	Climbing Gym	College Gym	Cycle Studio	Dance Studio	9.0

4. Discussion

The results of the geospatial analysis yield many useful insights for understanding the business potential of the various areas in Hong Kong and the optimal sportswear business strategy for these areas.

Through the choropleth map, we can tell that areas along the northern edge of Hong Kong island have a darker shading. This is expected, as the areas are widely known to be extremely urban, affluent, commercial and densely populated. The high occurrence of cluster 7 and 8 labels (in which shopping-related outlets are most common) shows that the nature of these areas translates into the high occurrence of shopping malls and gyms to meet the needs of higher-income, white-collar people. Sportswear stores in these areas may ideally be smaller, higher-end, and focused on “athleisure” and general-purpose products. As these downtown areas of Hong Kong tend to be more expensive, costs and prices are a particularly important consideration for a successful store location.

In addition, cluster 11 labels on the western end of the northern edge (e.g. Sai Ying Pun, Hong Kong University, Kennedy Town) reflect a demand for fitness venues amongst people who can afford to live in the more residential areas on Hong Kong island. Stores in these areas should still cater for the demands of consumers with more spending power, but their product offerings can be more fitness-focused compared to their downtown counterparts.

On Kowloon side, we also see darker shading in areas that are densely populated and popular amongst younger crowds (residents and consumers alike). Cluster 2 and 9 labels imply that younger demographics have demand for fashion and sports clothing outlets. By extension, we can hypothesize that areas with younger audiences (often due to lower rent and vibrant "hip" environments) are promising sportswear store locations. These stores should reflect the tastes of their residents, such as by being more trendy, affordable and fitness-focused.

Another area with darker shading is Kwun Tong, with many cluster 12 labels indicating shopping centers and clothing stores. This may be attributed to its dense and residential population. The relatively affordable property prices make Kwun Tong a less expensive store location, but the lower average income of its residents will reduce their purchasing power. A unique characteristic of cluster 12 areas is the higher number of soccer fields. We may consider offering more soccer-related sportswear products to cater to local demands.

Tsing Yi is an interesting area with different characteristics. The cluster 1 and 11 labels reveal a high number of fitness-related venues (e.g. Pools, Boxing Gyms, Fitness Centers, Climbing Gyms). This is likely because of the area's more spacious and residential nature. Stores in this area can be larger, fitness-focused and family-oriented. Partnerships with local fitness-outlets can also be promising.

5. Conclusion

This project's geospatial analysis is a helpful first step to understanding what areas to open store locations. By analyzing the number of relevant venues in each area, we now have a solid understanding of which areas have fitness and fashion-ecosystems. Areas with more venues imply greater interest in fitness and fashion, as well as more economic potential. In addition, we also understand the specific types of venues in each area. Some areas emphasize on fashion and shopping, others on sports and fitness.

Going forward, we have many ways of furthering our analysis. We have not yet considered factors such as population density, average income for residents, foot traffic, economic trends over time, and more granular features of the venues (e.g. high or low-end, specific types of sports practiced). All these factors would strengthen our understanding of which areas are promising locations, and how to optimize our stores for each location.