Recommending Location for new Bookstore

The Battle of Neighborhoods : IBM Coursera Applied Data Science Capstone Project

Student: Sridharan Sadagopan

Recommending Location for new Bookstore

- Opening a new bookstore in a new city requires
  - Selection of the right business location
    - Near Schools., Colleges, Universities
    - In Shopping mall so as easy for reachability for the right target customers
  - Competition landscape understanding
- Objective is to recommend top location in the City of Kuala Lumpur in Malaysia

Data Collection and Data Understanding

- We got the list of suburbs from Wikipedia for Kuala Lumpur (the city of interest) from
  - https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur
- And GeoJson Data from Federal Territory Map Kuala Lumpur
  - https://raw.githubusercontent.com/TindakMalaysia/Federal-Territories-Maps/master/KL/2016/MAP/MIGRATED/result/09-WPKL-New-DM-4326.geojson
- With Geo Coordinates for the suburbs, we proceed to use FourSquare Service to explore and search venues of interest (e.g school)

Analytical Approach

- Selection of right neighborhood for a new bookstore could potentially be done by clustering the neighborhoods, understanding the clusters, picking the cluster that has right characteristics that match our business understanding and requirements.
  - We could use KMeans clustering algorithm to cluster the neighborhoods
  - Further analysis of clusters to understand the pattern
  - Select the cluster that is close to our business understanding and requirement
  - Pick the top 5 neighborhoods and choose the location for bookstore near them.

Data Understanding

- FourSquare API returns lot of different venues, so filter results using the category ids as defined by FourSquare API [https://developer.foursquare.com/docs/build-with-foursquare/categories]

  - College_iniversity='4d4b7105d754a06372d81259'
  - Library='4bf58dd8d48988d12f941735'
  - School='4bf58dd8d48988d13b941735'
  - Shopping_mall='4bf58dd8d48988d1fd941735'
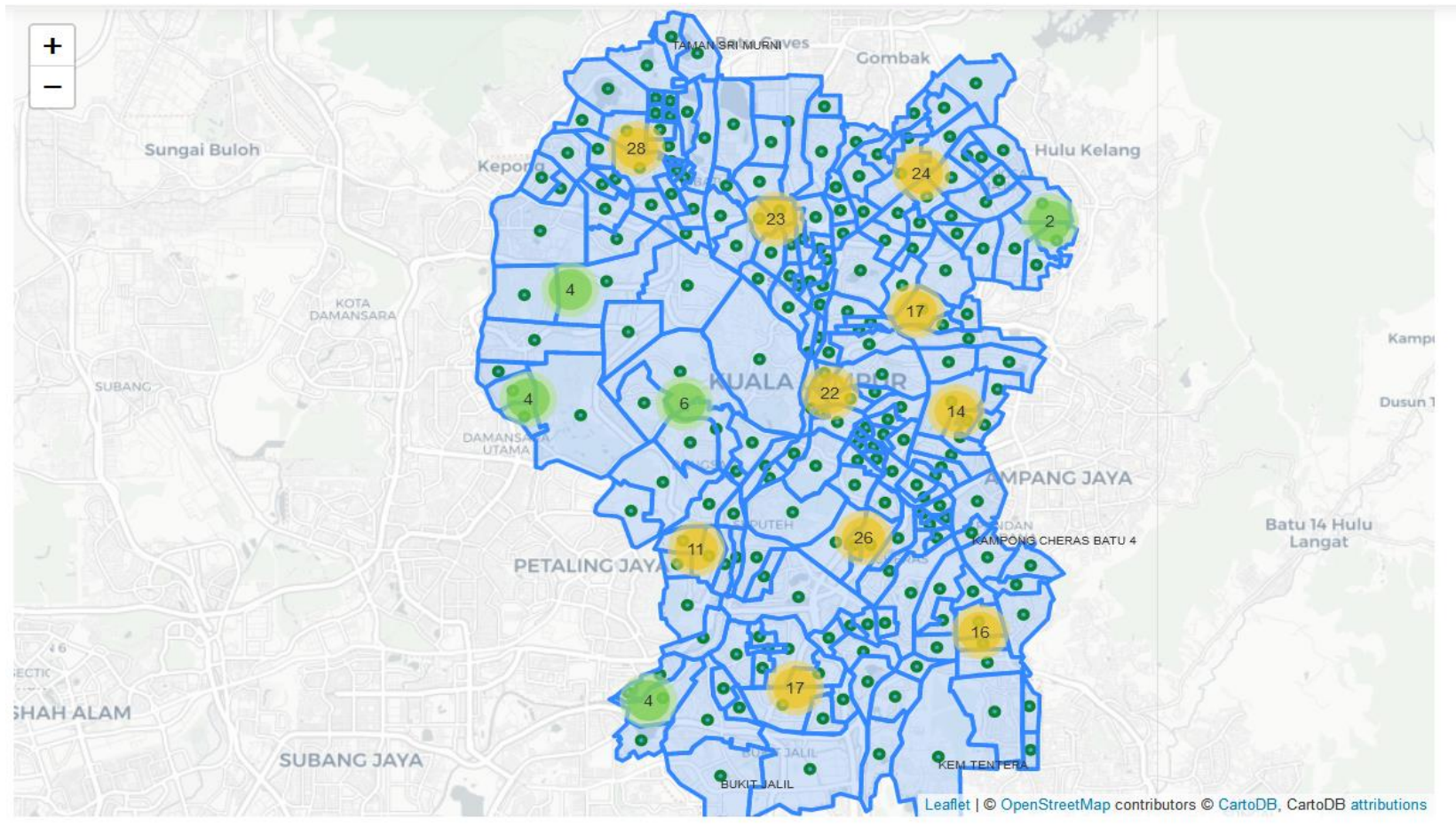 - Shopping_plaza='5744ccdfe4b0c0459246b4dc'

Data Understanding

- FourSquare API returns lot of different venues, so filter results using the category ids as defined by FourSquare API [https://developer.foursquare.com/docs/build-with-foursquare/categories]

  - College_iniversity='4d4b7105d754a06372d81259'
  - School='4bf58dd8d48988d13b941735'
  - Shopping_mall='4bf58dd8d48988d1fd941735'

	Unnamed: 0	Suburb	Suburb Latitude	Suburb Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
21052	21052	KEM TENTERA	3.057482	101.725049	Royal Military College, Malaysia	3.044258	101.723056	High School
20730	20730	KAMPONG SUNGAI BESI	3.054618	101.694689	Asia Pacific University of Technology & Innova...	3.048224	101.692856	University
20704	20704	KAMPONG SUNGAI BESI	3.054618	101.694689	Bukit Jalil Sports School	3.050039	101.694600	School
20752	20752	KAMPONG SUNGAI BESI	3.054618	101.694689	SK Bukit Jalil	3.050485	101.686807	Elementary School
20760	20760	KAMPONG SUNGAI BESI	3.054618	101.694689	Sekolah Sukan Bukit Jalil	3.050677	101.690249	Middle School

Data Understanding

- Visualize the neighbourhoods data in Kuala Lumpur

Data Exploration

- Group Venue Category by Neighbourhood to understand the venue count

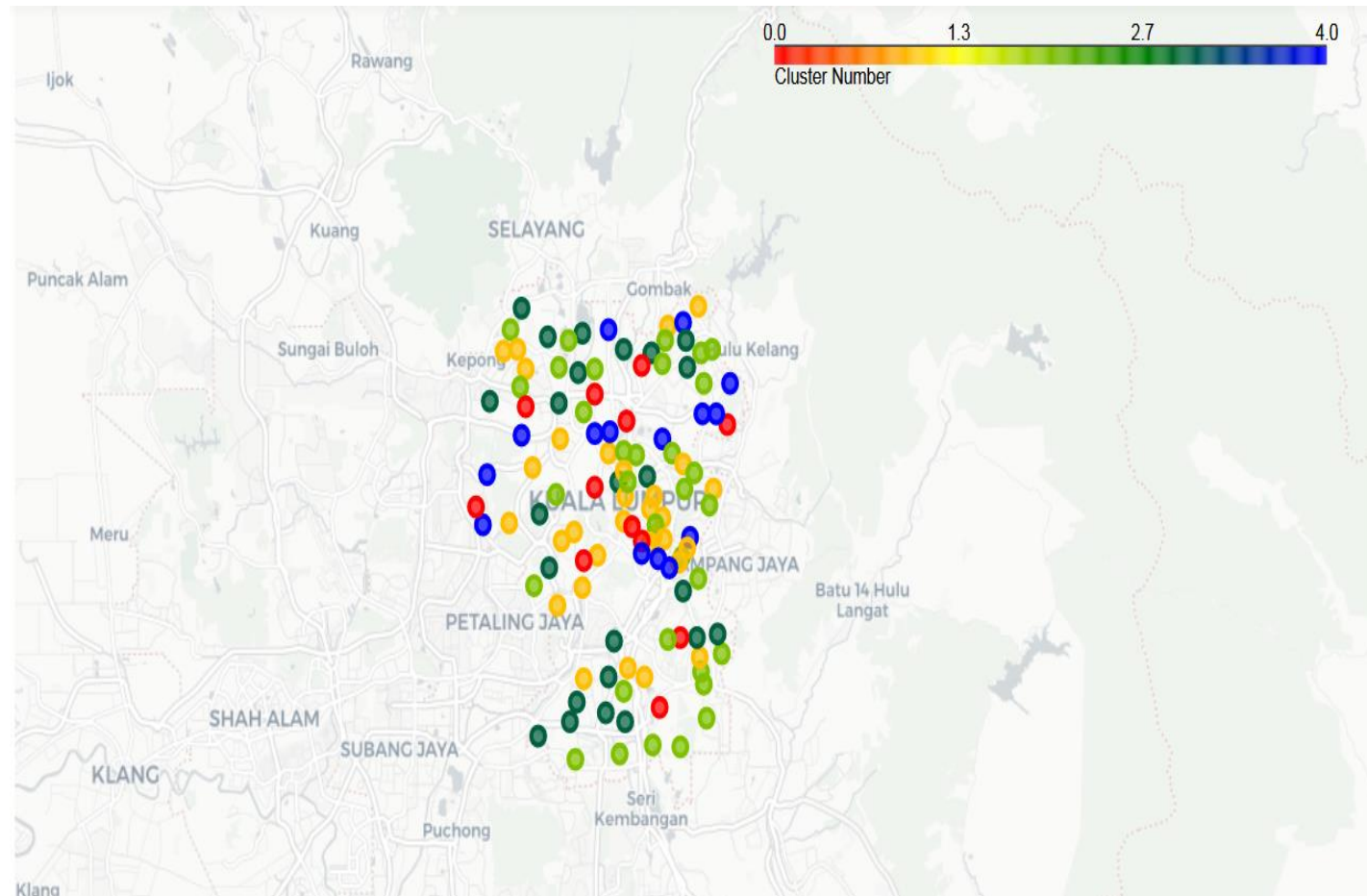| Suburb	Venue Category	Venue
TIONG NAM	Shopping Mall	53
	School	7
	Shopping Plaza	7
	University	7
	Bookstore	3
	High School	3
	General College & University	1
	Language School	1
	Supermarket	1
TAYNTON VIEW	School	17

Data Exploration

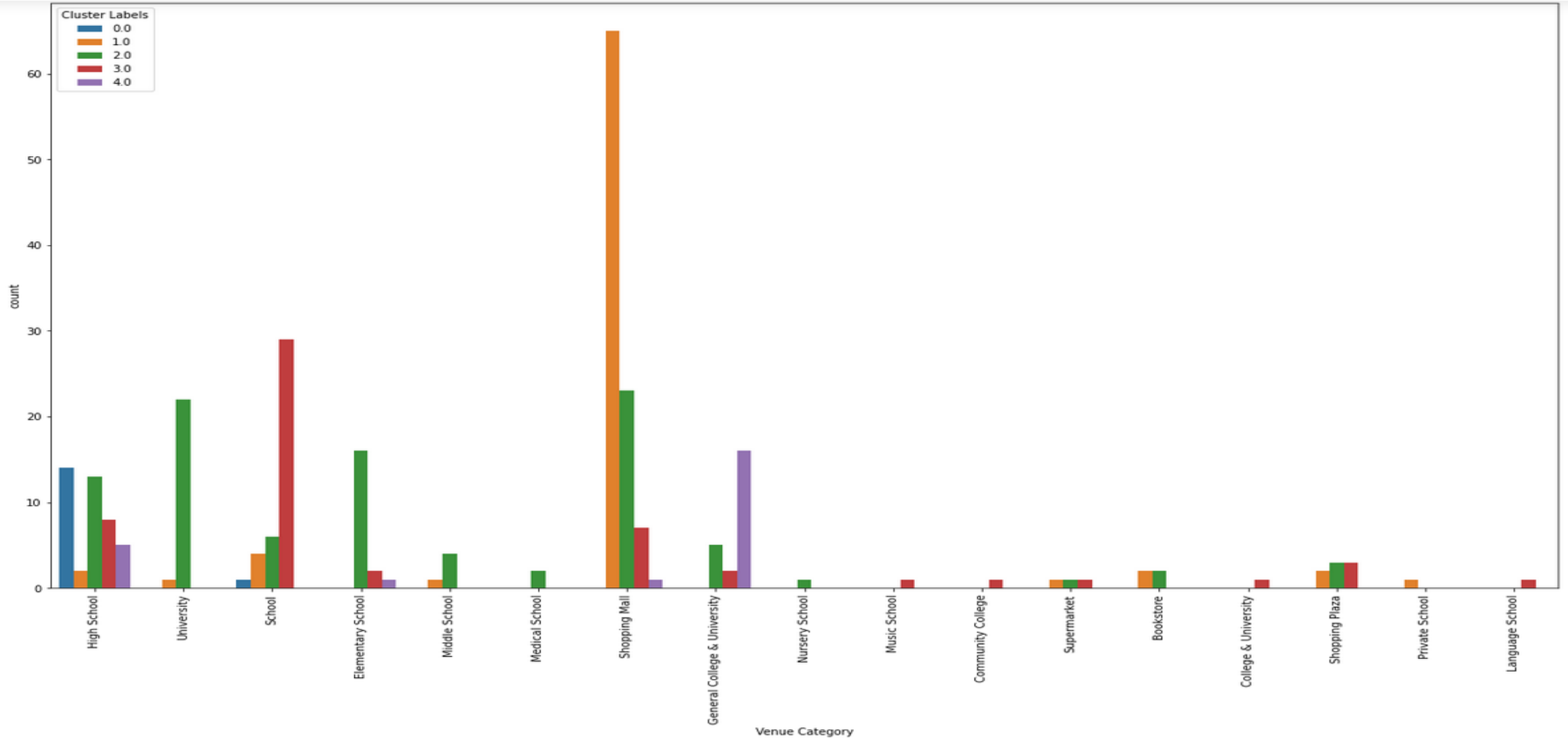- Find top 10 common venues of interest

	Suburb	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	1st Most Common Venue Count	2nd Most Common Venue Count	3rd Most Common Venue Count
0	BANDAR MANJALARA	High School	School	Shopping Plaza	Shopping Mall	Supermarket	University	College & University	Community College	Elementary School	General College & University	2	2	1
1	BANDAR SRI PETALING	School	Music School	University	Language School	College & University	Community College	Elementary School	General College & University	High School	Middle School	1	1	0
2	BANDAR TASIK SELATAN	High School	School	University	Medical School	College & University	Community College	Elementary School	General College & University	Language School	Middle School	2	1	0
3	BANGSAR BARU	Shopping Mall	University	Medical School	College & University	Community College	Elementary School	General College & University	High School	Language School	Middle School	3	0	0
4	BATU 3 - 4 JALAN CHERAS	School	University	Medical School	College & University	Community College	Elementary School	General College & University	High School	Language School	Middle School	1	0	0

Data Exploration

- KMeans Clustering to cluster neighbourhood

Analyzing the Results

Analyzing the Results …

- Analysing Clusters

Cluster	Colour	Top venue category
Cluster 0	Blue	High School
Cluster 1	Brown	Shopping Malls
Cluster 2	Green	University
Cluster 3	Red	School
Cluster 4	Violet	General College & Universtiy

Analyzing the Results (Competitions)...

- Competition Landscape

Cluster Labels	Educational Institutions	Bookstores	Educational Institutions per Bookstores
0.0	15	NaN	NaN
1.0	9	2.0	4.5
2.0	69	2.0	34.5
3.0	45	NaN	NaN
4.0	22	NaN	NaN

Analyzing the Results (Competitions)...

- Competition Landscape

Cluster Labels	Educational Institutions	Bookstores	Educational Institutions per Bookstores
0.0	15	NaN	NaN
1.0	9	2.0	4.5
2.0	69	2.0	34.5
3.0	45	NaN	NaN
4.0	22	NaN	NaN

Analyzing the Results …

- Clusters 1 and 2 there are bookstores in their vicinity, whereas for other clusters no bookstore nearby
- **Cluster 3, this cluster has next higher number but no bookstore in the vicinity.**

Analyzing the Results (Cluster 3)

- Understand 3 more

	Suburb	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
23	TAMAN BATU PERMAI	School	High School	General College & University	University	Medical School	College & University	Community College	Elementary School	Language School	Middle School
77	BANDAR MANJALARA	High School	School	Shopping Plaza	Shopping Mall	Supermarket	University	College & University	Community College	Elementary School	General College & University
51	SEKSYEN 1 WANGSA MAJU	School	Shopping Mall	General College & University	High School	University	Language School	College & University	Community College	Elementary School	Middle School
160	KUCHAI	School	Shopping Mall	High School	University	Language School	College & University	Community College	Elementary School	General College & University	Middle School
164	TAMAN YARL	School	Shopping Mall	High School	University	Language School	College & University	Community College	Elementary School	General College & University	Middle School

Choices of Location for Bookstore

- Shopping Malls near the top 5 neighbourhood in Cluster 3

	Suburb	Venue
16442	TAMAN YARL	Plaza OUG
16011	KUCHAI	The Scott Garden
7714	BANDAR MANJALARA	Kepong Village Mall
5105	SEKSYEN 1 WANGSA MAJU	AEON Alpha Angle Shopping Centre

Conclusion

- Location based search and exploration using services such FourSquare API, enabled me in analysing and understand a wealth of information about the neighbourhoods of Kuala lumper

- KMeans clustering algorithm was used cluster neighbourhood

- Top 5 locations were selected from the cluster that is close to business understanding and requirements

- The recommended locations were based on vicinity to education institutions, shopping malls and based on competition landscape.