

Similarity of the capital cities in the world

244 Countries around the world

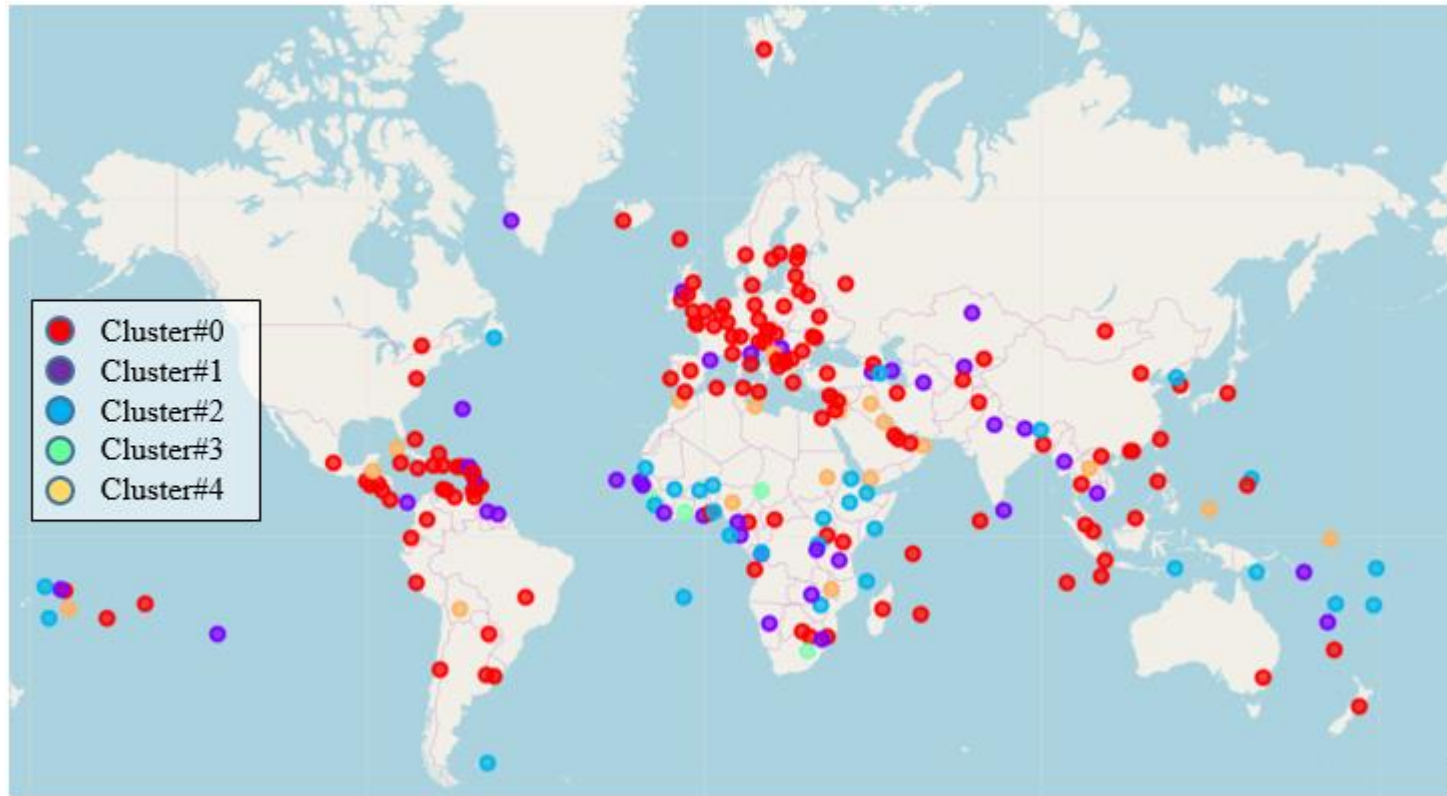
If we can group similar cities to each other, a business model which is success in a certain city might also be success if we implement it in the similar city in the other country



Cluster 244 countries into 5 clusters.

Using nearby venues around the center of the capital city of each country

Using Kmeans method



Result of the 1st clustering

9,623 venues out of 11,346 venues of total venues nearby all of the capital cities are in the 142 countries of cluster#0

Country Capital Population Latitude Longitude

Cluster Labels

0	142	142	142	142	142
1	38	38	38	38	38
2	31	31	31	31	31
3	4	4	4	4	4
4	19	19	19	19	19

The 10 most common venues: most likely tourist attraction, restaurants, and hotels

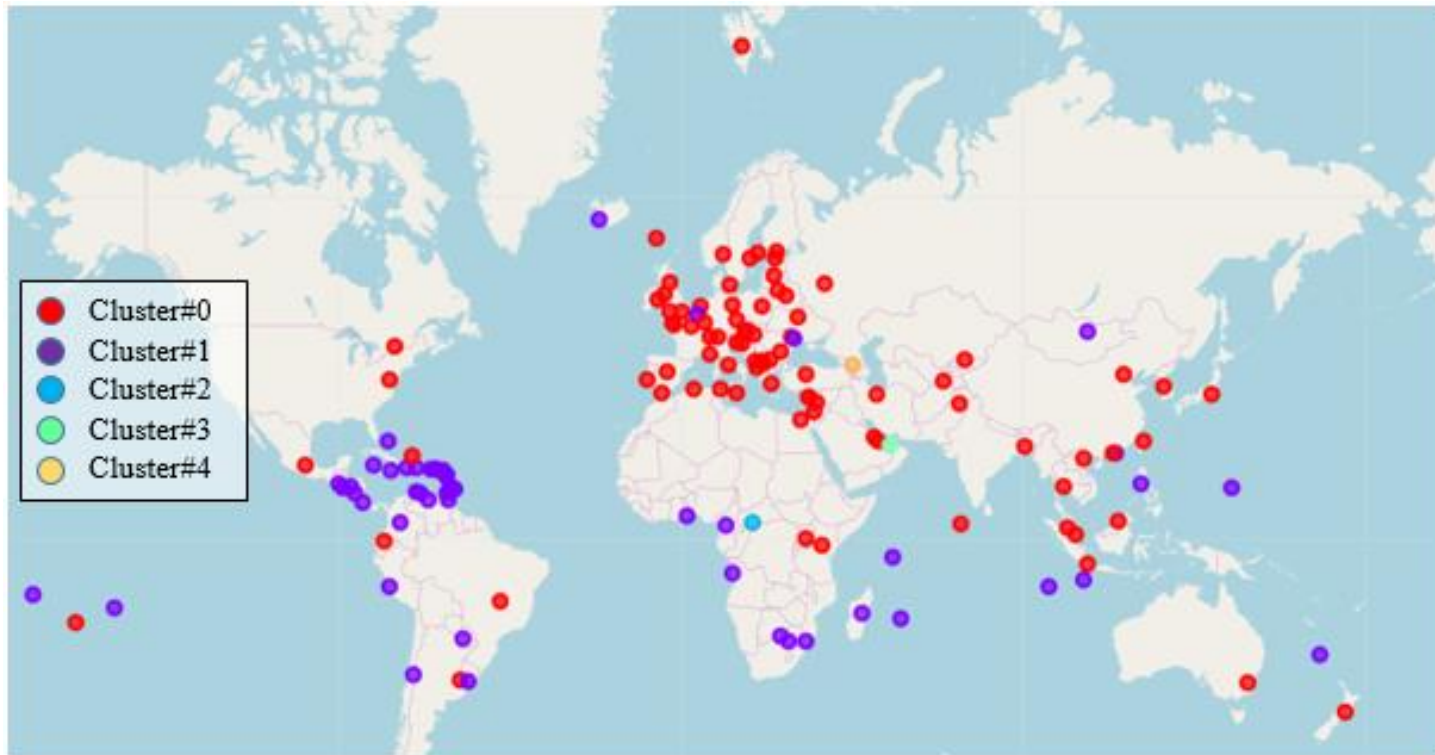
Most of the big country, developed and developing countries are clustered into cluster#0

Country	Capital	Cluster Labels	Country	Capital	Cluster Labels	Country	Capital	Cluster Labels			
0	China	Beijing	0	6	Egypt	Cairo	0	15	Hong Kong	Hong Kong	0
1	Japan	Tokyo	0	7	Mexico	Mexico City	0	16	Colombia	Bogotá	0
2	Russia	Moscow	0	8	United Kingdom	London	0	18	Chile	Santiago	0
4	Indonesia	Jakarta	0	9	Bangladesh	Dhaka	0	19	Singapore	Singapore	0
5	South Korea	Seoul	0	10	Peru	Lima	0	20	Turkey	Ankara	0
				11	Iran	Tehran	0	21	Germany	Berlin	0
				12	Thailand	Bangkok	0	23	Algeria	Algiers	0
								24	Spain	Madrid	0

2nd Clustering for 142 countries into 5 clusters.

Using nearby venues around the center of the capital city of each country

Using Kmeans method



Result of the 2nd clustering

The 10 most common venues of cluster#0 and cluster#1, It seems like nearby the center of the cities of cluster#1 are more like residential area. On the other hand, for cluster#0, common venues are tourist attractions, restaurants, business areas, and hotels.

	Country	Capital	Population	Latitude	Longitude
Cluster Labels					
0	86	86	86	86	86
1	53	53	53	53	53
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1

tourist attractions,
restaurants, business
areas, and hotels.

parks, restaurants, and
public places

Since this observation aims to find out cities' similarity and seek business opportunity so we decided to look more into the cities in cluster#0 of the 2nd clustering. So before going further, this is where the cities that we are considering, are located.



Final, 3rd Clustering into 5 clusters.

Using nearby venues around the center of the capital city of each country together with country population and GDP per capita
Using Kmeans method



Country population

	Country	Country Population
0	China	1433783686
1	India	1366417754
2	United States	329064917
3	Indonesia	270625568
4	Pakistan	216565318

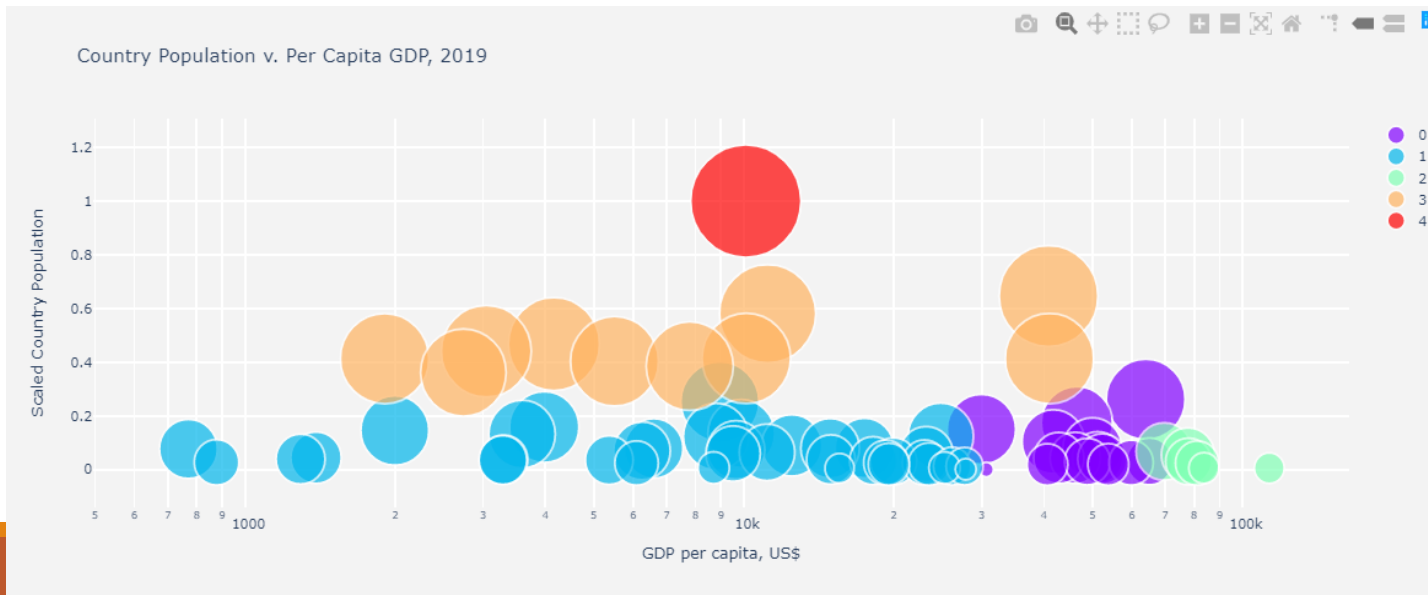
GDP per capita

	Rank	Country	GDP US\$ per capita
0	1	Luxembourg	113196
1	2	Switzerland	83716
2	—	Macau	81151
3	3	Norway	77975
4	4	Ireland	77771

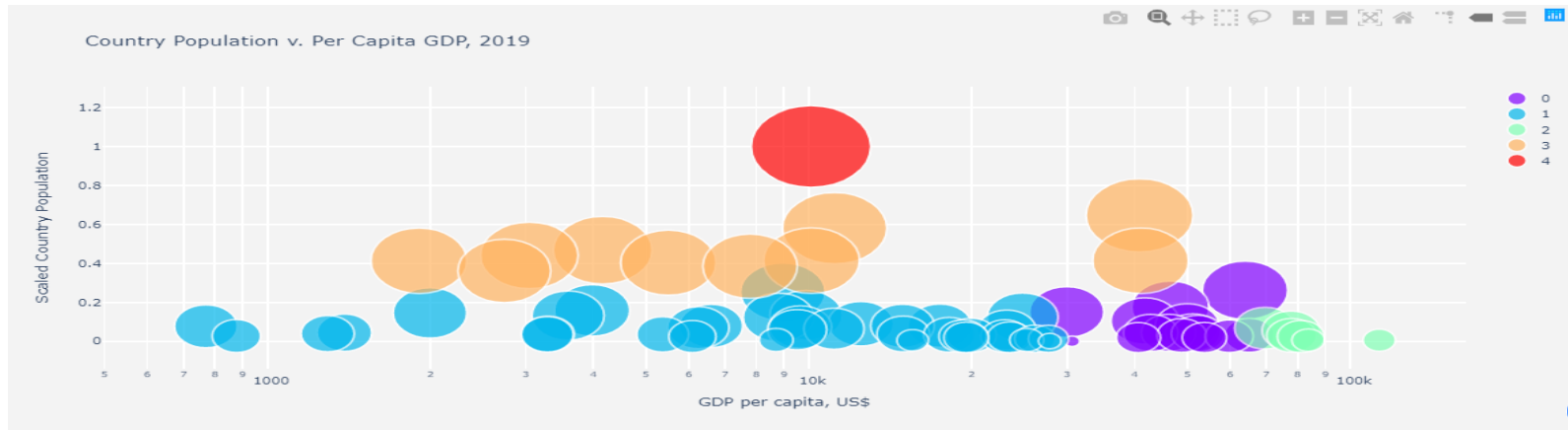
Result of the 3rd clustering



	Country	Capital	Population	Latitude	Longitude	Country Population	GDP US\$ per capita
Cluster Labels							
0	15	15	15	15	15	15	15
1	37	37	37	37	37	37	37
2	6	6	6	6	6	6	6
3	10	10	10	10	10	10	10
4	1	1	1	1	1	1	1



Observation conclusion



Cluster#0: High GDP per capita with fewer number of country population while their capital city has the number of population in the range of low to medium.

- Singapore, Germany, Spain, France, Austria, Sweden, Canada, Israel, Netherlands, United States, Finland, Denmark, Australia, New Zealand, Malta

Cluster#1: Majority of the result, low to medium GDP per capita with less number of country population and their capital city also has the number of population in the range of low to medium.

- Turkey, Algeria, Kenya, Argentina, Ukraine, Taiwan, Brazil, Romania, Hungary, Poland, Belarus, Uganda, Lebanon, Ecuador, Malaysia, Bulgaria, Czech Republic, Pakistan, Kyrgyzstan, Croatia, Moldova, Tunisia, Albania, Latvia, Greece, Tajikistan, Portugal, Lithuania, North Macedonia, Estonia, Slovakia, Slovenia, Cyprus, Montenegro, Bahrain, Maldives, Brunei

Cluster#2: Highest GDP per capita with quite less number of country population and their capital city also has a small population.

- Qatar, Ireland, Norway, Macao, Switzerland, Luxembourg

Cluster#3: Medium to high GDP per capita with a high number of country population and their capital city also have a high number of population.

- Japan, Russia, Indonesia, Egypt, Mexico, United Kingdom, Bangladesh, Iran, Thailand, Vietnam

Cluster#4: This cluster contains only one country which is China. This city has its own characteristic, medium GDP per capita with a very high number of country populations and a very high population for the capital city which is Beijing.

- China