Relocation between Neighborhoods in KSA Cities

1. Introduction

Kingdom of Saudi Arabia (KSA) is an Arab Islamic state, the largest country in the Arabian peninsula situated at the strategic crossroads of three continents Europe, Africa and Asia. It is a home to several human civilizations and the cradle of the heavenly messages. Capital of Saudi Arabia is Riyadh.

KSA is renowned for the sacred sites of Islam. The two Holy Mosques' of Makkah Al Mukarramah and Al Madinah Al Munawwarah are the destinations of millions of Muslims annually for Hajj and Umrah additionally they visit other significant sites of Islam.

KSA ranks first in production and reserves of petrol, fifth in natural gas reserves and tenth in natural gas production. Thousands of Muslims heads for the KSA for Hajj and Umrah and for visiting Makkah Al Mukarramah and Al Madinah Al Munawwarah, these visits necessitates provision of housing, catering and health services for the guests of Allah almighty.

part of Saudi Vision 2030 opening Saudi Arabia further for business will boost productivity and smooth journey to become one of the largest economies in the world. They improve business environment, restructure economic cities, create special zones and deregulate the energy market to make it more competitive.

The Saudi Human Resources Gateway is one of the major initiatives of HRDF that aims to create a leap forward in HR Management practice, to contribute to the progression of the HR practitioners and to form a knowledge database where establishments, HR practitioners and researchers can benefit from. The HR Gateway aims to provide specialized, comprehensive, and free of charge Arabic HR content. Also it aims to offer HR practitioners an effective social platform as well as allowing them to have a reach to HR related contents, practicing it and modifying it to suit the company needs and share their experiences.

One of innovative services we can add to Saudi HR Gateway providing advising service for people to move and live in one of KSA cities suitable and nicely and suitable .

a) Problem Description:

Now let me explain the context of coursera Capstone project through a an example. Fahed live on Al Wurud distrect which is a neighborhood in Jeddah city (one of KSA cities). Fahed love this neighborhood, mainly because of all the great surrounding and enviormnent that exist in this neighborhood, such as burger joint, dessert shop, coffee shops,parks, pharmacies, graduate schools and so

on. There are lot of oppertunity and job offer from a great company in Riyadh city (the capital of KSA) with great career prospects. However, given the far distance from his current place he unfortunately must move if he decide to accept the offer.

Wouldn't it be great if we are able to help Fahed or any similar person to determine neighborhoods on the other city that are the similar as the current neighborhood, and if not perhaps similar neighborhoods that are at least closer to his new job?

b) Objective

The aim of this report is to study and analyze the neighborhoods of both Jeddah and Riyadh cities and group them into similar clusters, to analyze those clusters to gather meaningful information. That information can be used to find out Riyadh neighborhoods that are same as current neighborhood in Jeddah or at least similar but .

This will help any person decide to move from one city to another in same country or diffrent country

c) Target Audience

This information provided by this report would be useful for people who are interested in relocating to a different city and are interested in finding new neighborhoods that are highly similar to their existing neighborhood.

2. Data Description & Preparation:

To consider the objective stated above, we can list the below data sources used for the analysis.

- KSA Neighborhood Data (regions, cities, and districts)
- Get all venues in each districts (neighborhood) and its catagory using Fousquare API

```
Libraries imported.
```

a) KSA Neighborhood Data:

The following github project (https://github.com/homaily/Saudi-Arabia-Regions-Cities-and-Districts/tree/master/json (https://github.com/homaily/Saudi-Arabia-Regions-Cities-and-Districts/tree/master/json)) contains Saudi Arabia Regions, Cities and Districts. The data is public data collected from https://maps.address.gov.sa/ (https://maps.address.gov.sa/) as ,all coordinates in (Lat,

Lon) aka (Y, X) format and 8 decimal points. Data points include: Regions, Cities and Districts. Names (Arabic & English). Regions capital city, population & center point. Regions boundaries. Districts boundaries. This data divided into three parts:

- regions.json
- cities.json
- districts.json

Then information obtained will be transformed into a pandas data frame for further analysis.

KSA regions

```
No of regon: 13
```

- KSA regions tabel

Out[4]:

	region_id	name_ar	name_en	population	capital_city_id	Latitude	Longitude
0	1	منطقة الرياض	Riyadh	6777146	3	24.70000	46.73333
1	2	منطقة مكة المكرمة	Makkah	6915006	6	21.42718	39.84349
2	3	منطقة المدينة المنورة	Madinah	1777933	14	24.47058	39.60781
3	4	منطقة القصيم	Qassim	1215858	11	26.33034	43.97436
4	5	المنطقة الشرقية	Eastern Province	4105780	13	26.44199	50.10920

- KSA cities

```
No of cities: 3946
```

- KSA cities tabel

Out[6]:

	city_id	region_id	name_ar	name_en	Latitude	Longitude
0	1	7	تبوك	Tabuk	28.41464	36.53387
1	2	7	نعمي	Na'mi	28.30508	35.74931
2	3	1	الرياض	Riyadh	24.70000	46.73333
3	4	7	حميط	Humayt	28.65152	35.38013
4	5	2	الطائف	At Taif	21.26848	40.41667

- KSA districts

```
No of districts: 2782
```

- Important transformation for data

This file contan each district with boundaries points and there is no centeral location for district then I got all boundaries points Latitude & Longitude and caculate mean of then which give me centeral point of each district

- KSA districts tabel

Out[20]:

	district_id	city_id	region_id	name_ar	Neighborhood	Latitude	Longitude	boundaries
0	10100003001	3	1	حي العمل	Al Amal Dist.	24.644966	46.723598	[[24.64900056, 46.7248791], [24.64742521, 46.7...
1	10100003002	3	1	حي النموذجية	Al Namudhajiyah Dist.	24.655615	46.696523	[[24.65018372, 46.70227584], [24.64939455, 46....
2	10100003003	3	1	حي الجرادية	Al Jarradiyah Dist.	24.618815	46.696110	[[24.61729504, 46.70655201], [24.61727163, 46....
3	10100003004	3	1	حي الصناعية	Al Sinaiyah Dist.	24.646447	46.741675	[[24.64061956, 46.75999871], [24.6342299, 46.7...
4	10100003005	3	1	حي منفوحة الجديدة	Manfuha Al Jadidah Dist.	24.613354	46.716463	[[24.61343234, 46.72718798], [24.61090205, 46....

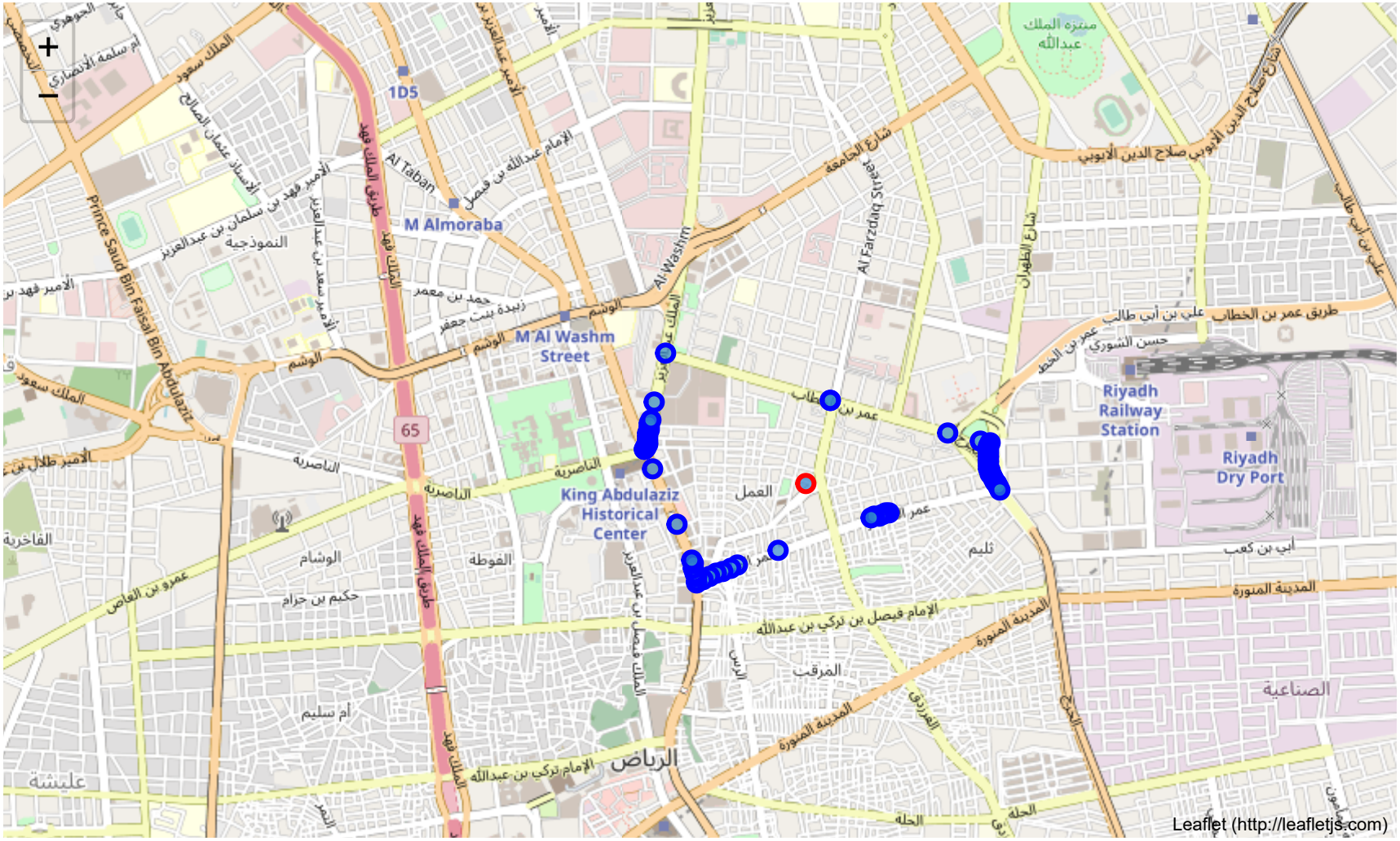
A sample for one district to test center and boundaries on the map shown below

Out[22]:

	district_id	city_id	region_id	name_ar	Neighborhood	Latitude	Longitude	boundaries
0	10100003001	3	1	حي العمل	Al Amal Dist.	24.644966	46.723598	[[24.64900056, 46.7248791], [24.64742521, 46.7...

```
Al Amal Dist.
```

Out[23]:



As shown the centeral of the district come in the middle of the boundaries points

b) Coordinate data for each Neighborhood in KSA Cities:

All the venues in each neighborhood in the KSA cities will be collected from Fousquare using Fousquare API By using this api we will get all the venues in each neighborhood. We can filter these venues to get only two cities will be anlyzed.

- Define Person requirement to move from current city to new city

We will define from and to city which defined by the person this can be the entry point for application which will advice person with best neighborhood/district in the new city match his current neighborhood/district in the current city:

```
Person need to move from:  Jeddah city to :  Riyadh  city
```

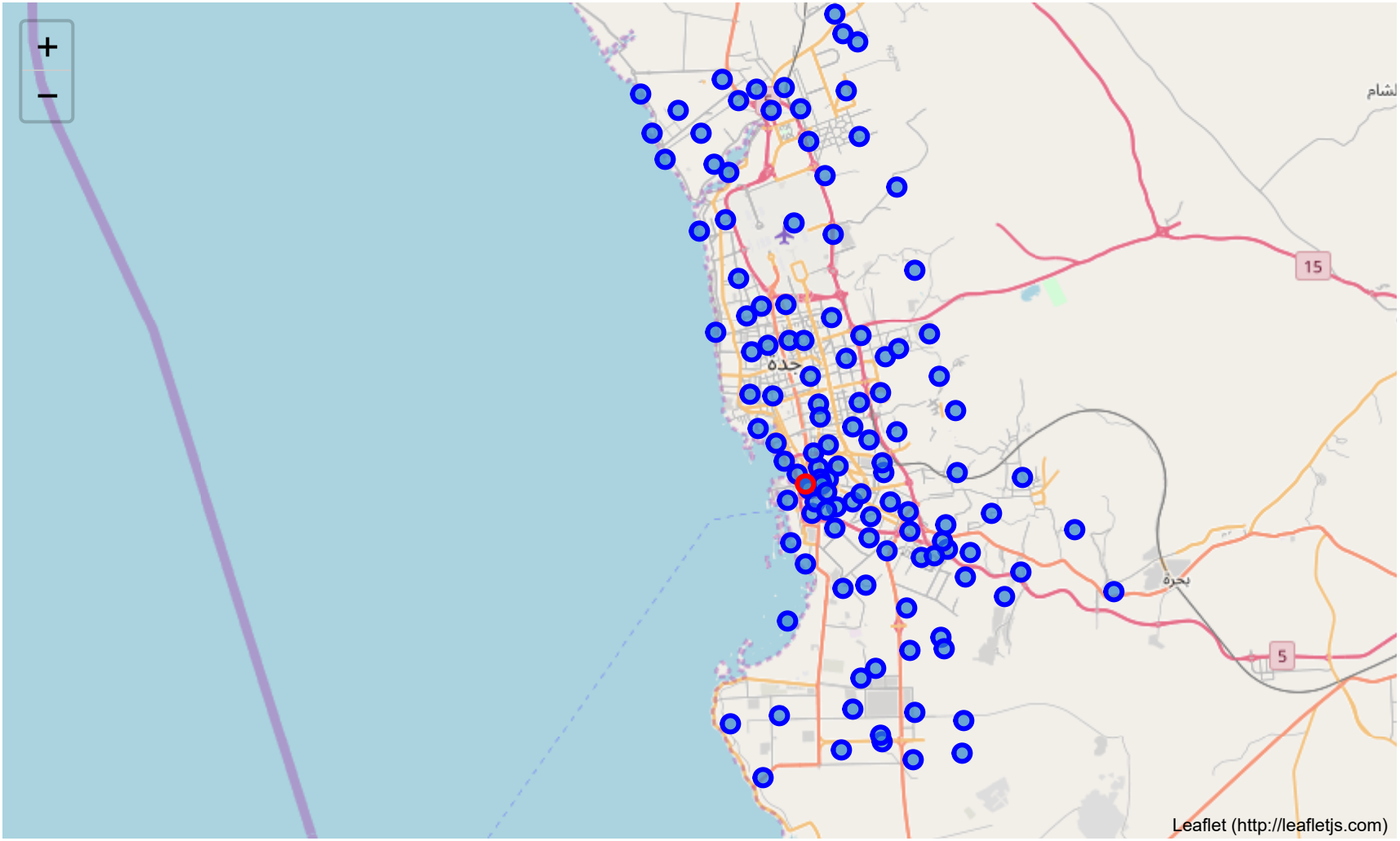
- Now will filter neighborhood/district of current and new city from the whole country_districts dataframe

```
Jeddah 119  districts
Riyadh 187  districts
```

Draw from city

Jeddah

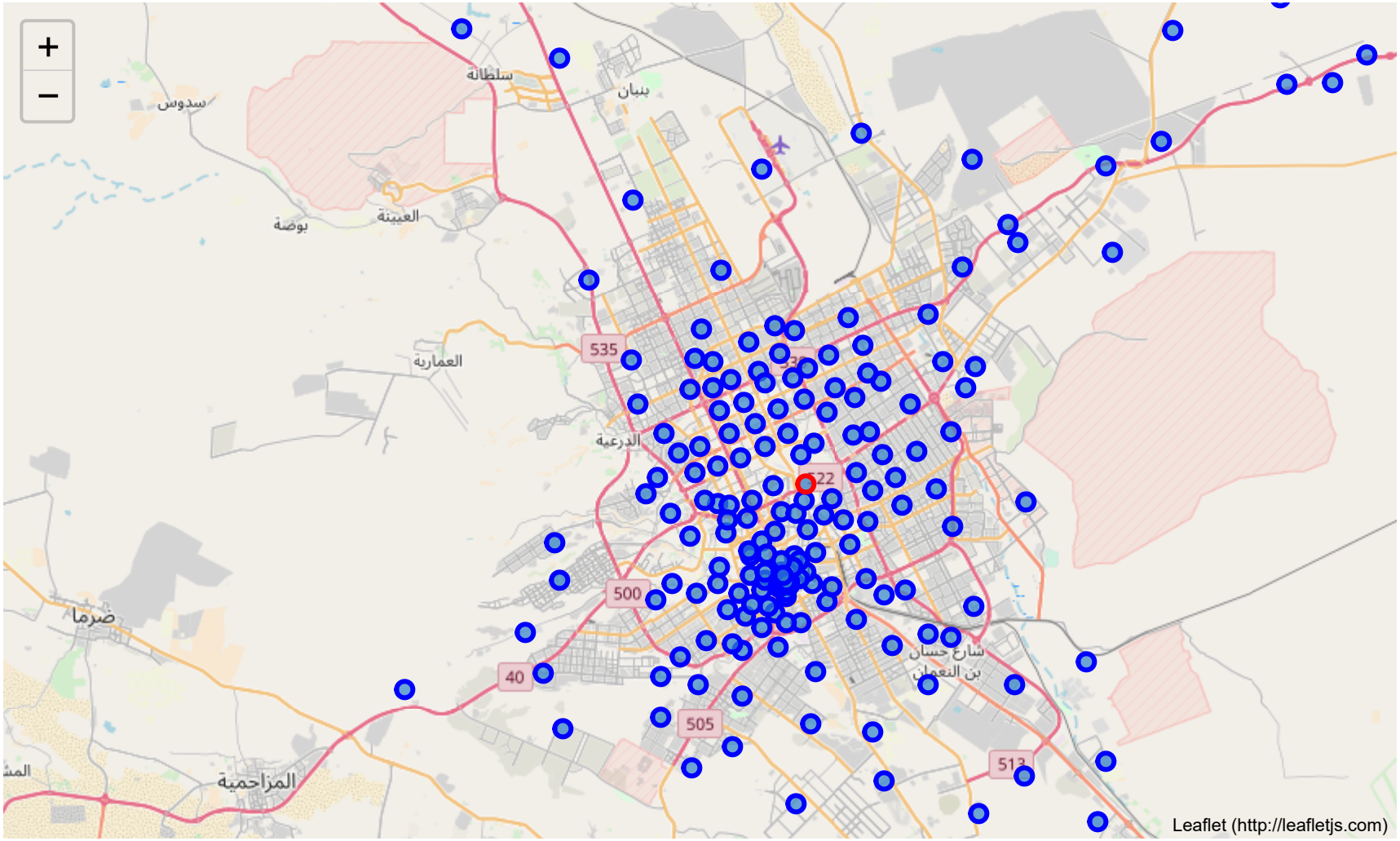
Out[116]:



Draw to city

```
Riyadh
```

Out[115]:



- Now combine two city to get all NearbyVenues for each neighborhood/district using foursquare API

Out[137]:

	district_id	city_id	region_id	name_ar	Neighborhood	Latitude	Longitude	boundaries
627	10200018001	18	2	حي الزمرد	18_Az Zomorod Dist.	21.792042	39.042490	[[21.80927021, 39.03429563], [21.81053914, 39....
628	10200018002	18	2	حي اللؤلؤ	18_Al Loaloa Dist.	21.762660	39.052579	[[21.74946321, 39.05707049], [21.74951489, 39....
629	10200018003	18	2	حي الياقوت	18_Al Yaqoot Dist.	21.779845	39.074368	[[21.7637526, 39.07722576], [21.7815796, 39.06...
630	10200018004	18	2	حي الصوارى	18_As Swaryee Dist.	21.804336	39.110714	[[21.79647589, 39.1222053], [21.78069932, 39.1...
631	10200018005	18	2	حي الامواج	18_Al Amwaj Dist.	21.741037	39.063722	[[21.73475579, 39.06916738], [21.73475841, 39....

Define Foursquare Credentials and Version

```
Your credentails:
CLIENT_ID: 5AHI1B14PNWWA3ZKPOQPLKSXRZ0P3OXLSUU1EMYDLVWWH1VY
CLIENT_SECRET:WSPO43X5JSJXEKH2F2X3MOUPFUQCBYLSY0WQYZL3FADJJAOR
```

Now call the foursquareAPI on each neighborhood to get all nearby venues

then store new dataframe locally on file to deal with limited number of API call of foursquare daily

then drop null and explore data

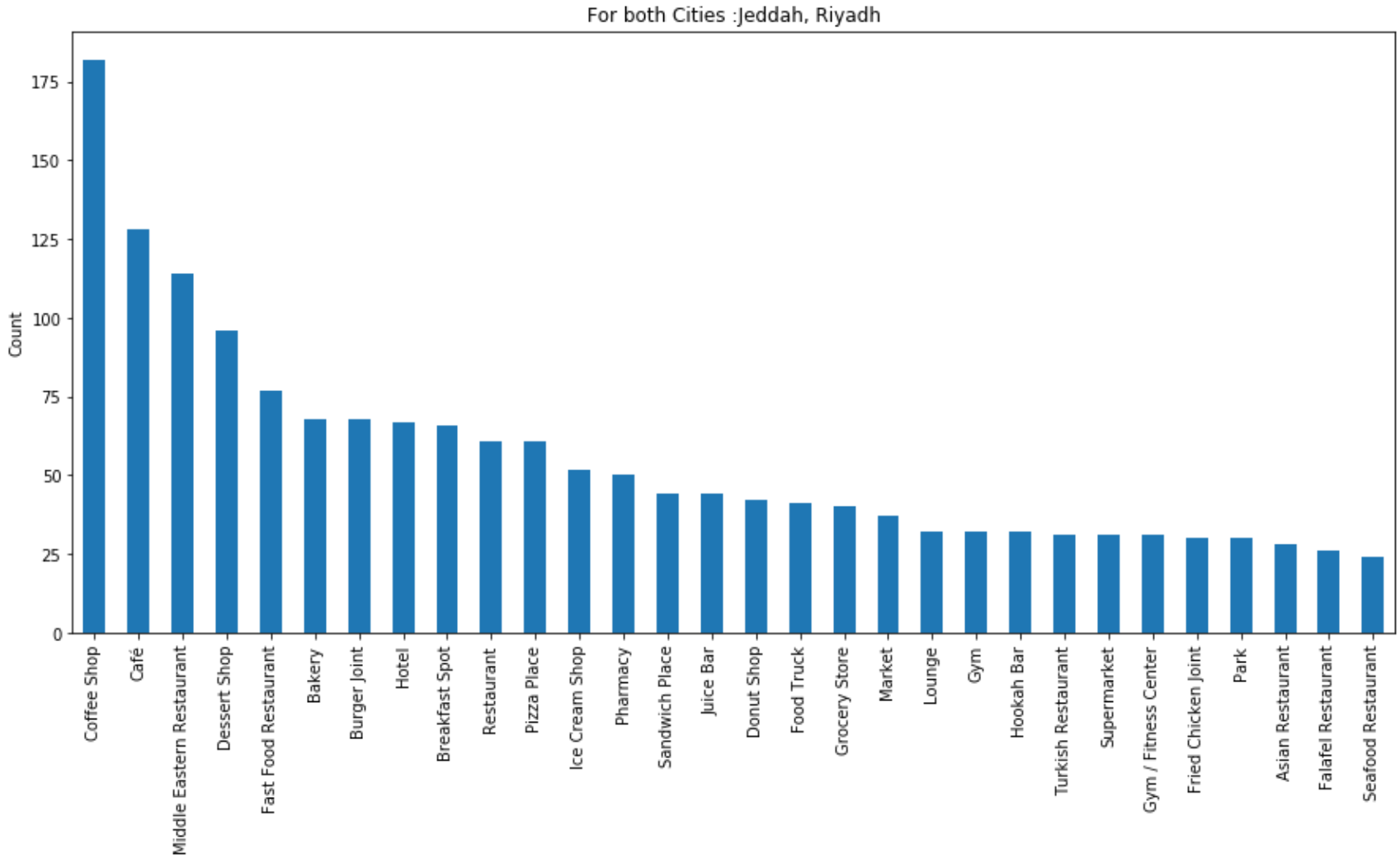
Out[28]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue_id	Venue	Latitude	Longitude	Category	city_id
0	18_Az Zomorod Dist.	21.792042	39.04249	5a1d4f79c0af5707fbf8723a	Obhur	21.790107	39.042135	Beach	18
1	18_Az Zomorod Dist.	21.792042	39.04249	52c808e1498e4d3e96049621	Private Beach	21.793894	39.044243	Beach	18
2	18_Az Zomorod Dist.	21.792042	39.04249	5c053acad1a402002c4d683a	Boho The Beach	21.795712	39.043003	Beach	18
3	18_Az Zomorod Dist.	21.792042	39.04249	50c1f787e4b0757ad3b06704	Indigo	21.788826	39.044985	Resort	18
4	18_Az Zomorod Dist.	21.792042	39.04249	54134a91498e038f80469a59	Amaya	21.789050	39.045123	Restaurant	18

Now analyse current and new cities by display common venues catagories as shown below

- Data Analysis

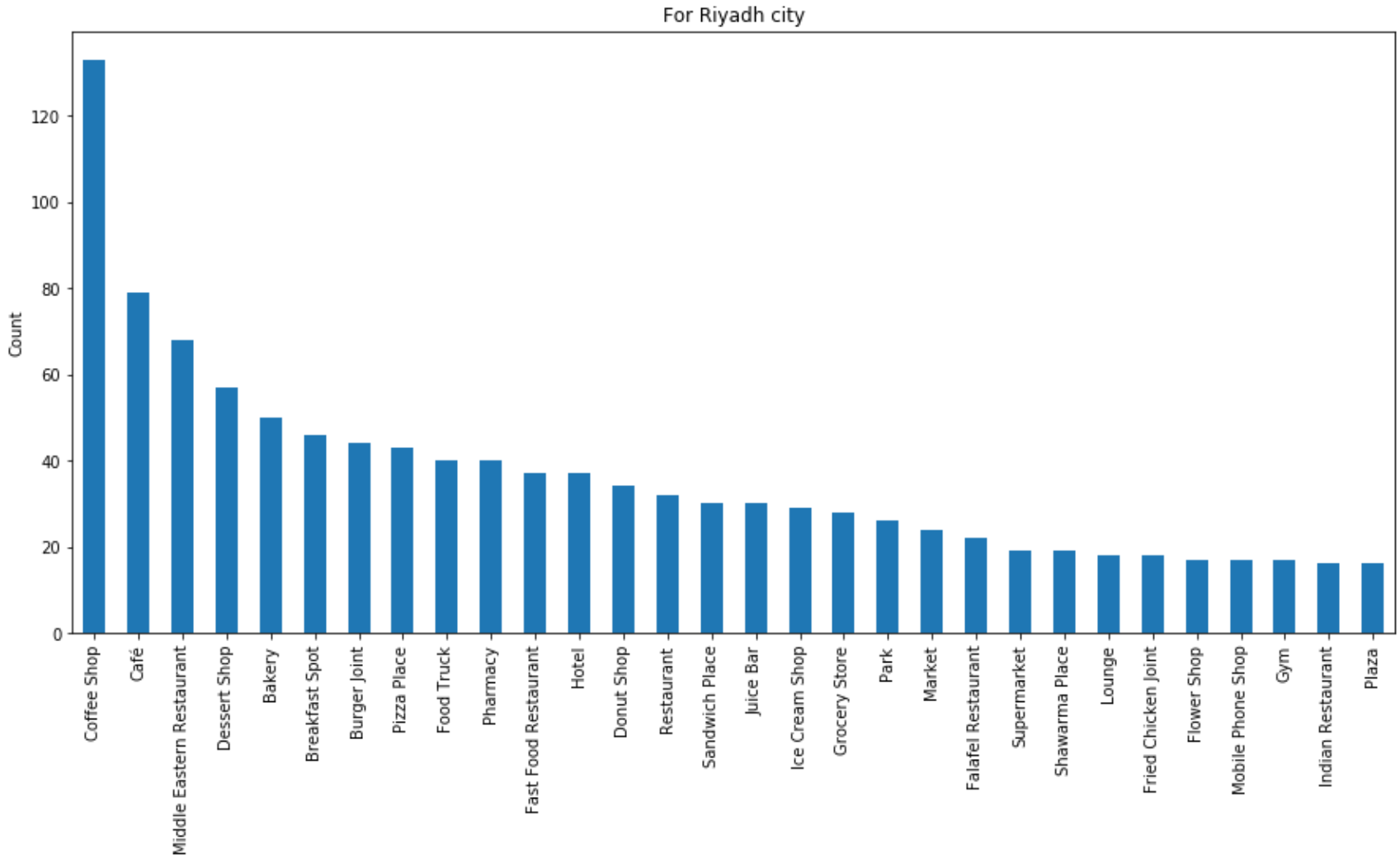
- Common categories for both cities


For both Cities :Jeddah, Riyadh
Count
0
25
50
75
100
125
150
175
Coffee Shop
Café
Middle Eastern Restaurant
Dessert Shop
Fast Food Restaurant
Bakery
Burger Joint
Hotel
Breakfast Spot
Restaurant
Pizza Place
Ice Cream Shop
Pharmacy
Sandwich Place
Juice Bar
Donut Shop
Food Truck
Grocery Store
Market
Lounge
Gym
Hookah Bar
Turkish Restaurant
Supermarket
Gym / Fitness Center
Fried Chicken Joint
Park
Asian Restaurant
Falafel Restaurant
Seafood Restaurant

As shown most common categories in both cities:

- Coffee Shop
- Café
- Middle Eastern Restaurant
- Dessert Shop
- Bakery

- Common categories for new city

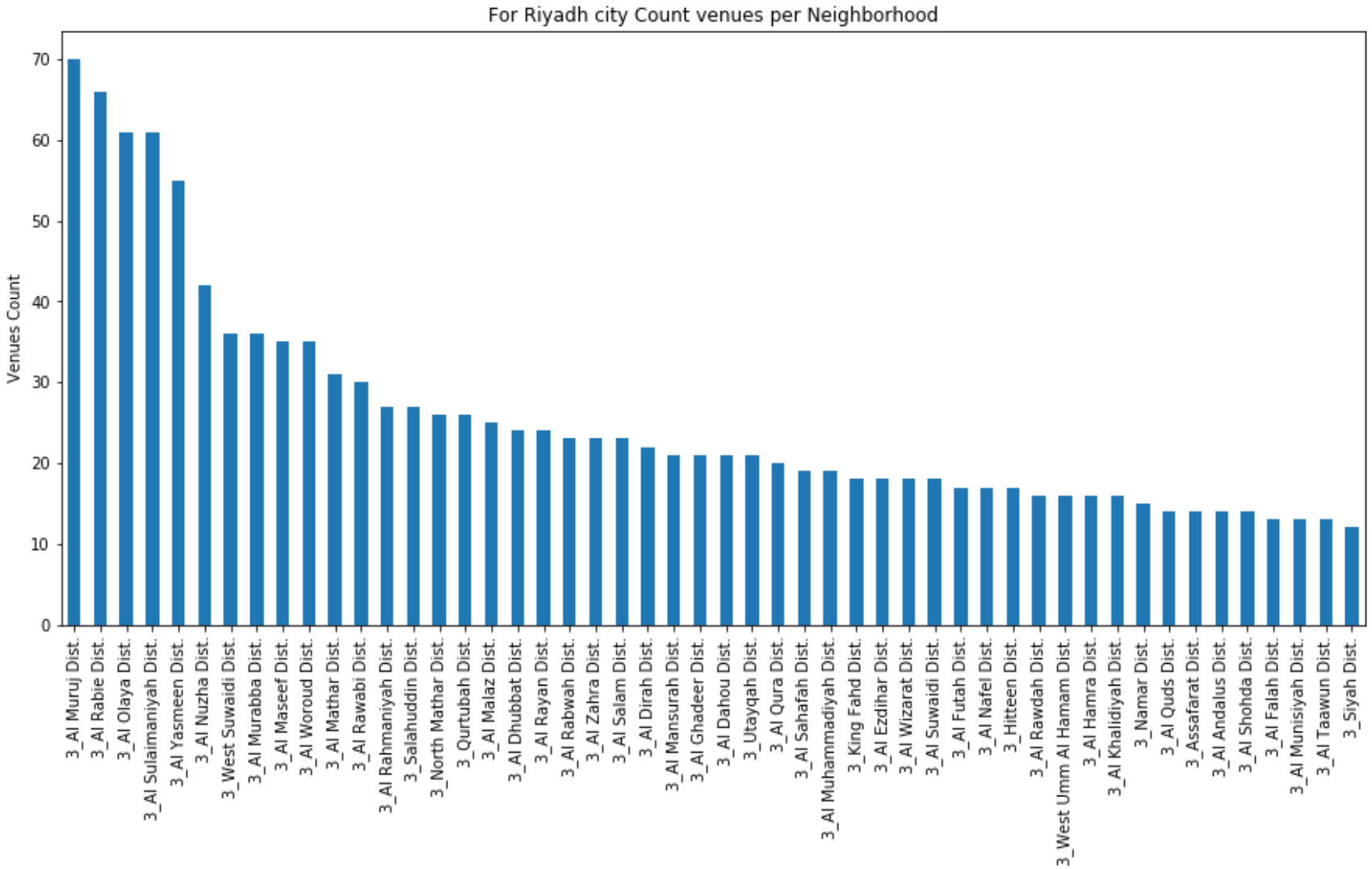

For Riyadh city
Count
0
20
40
60
80
100
120
Coffee Shop
Café
Middle Eastern Restaurant
Dessert Shop
Bakery
Breakfast Spot
Burger Joint
Pizza Place
Food Truck
Pharmacy
Fast Food Restaurant
Hotel
Donut Shop
Restaurant
Sandwich Place
Juice Bar
Ice Cream Shop
Grocery Store
Park
Market
Falafel Restaurant
Supermarket
Shawarma Place
Lounge
Fried Chicken Joint
Flower Shop
Mobile Phone Shop
Gym
Indian Restaurant
Plaza

As shown most common categories in new city:

- Coffee Shop
- Café
- Middle Eastern Restaurant
- Dessert Shop
- Bakery

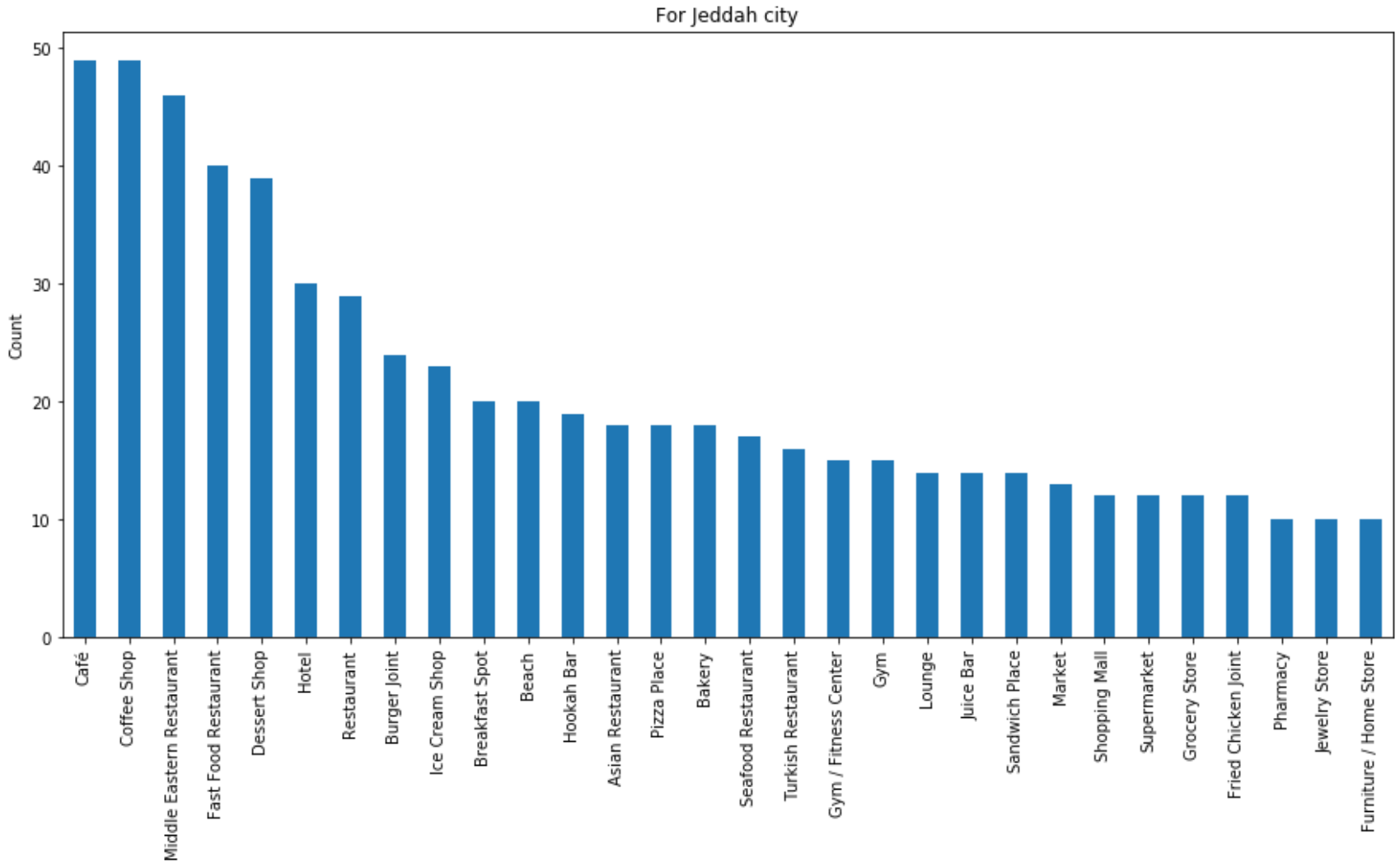
and as notes number of venues near from double of current catagory and this is becuse new city is the captial of KSA

- Then display number of venues per Neighborhood for new city


For Riyadh city Count venues per Neighborhood
Venues Count
0
10
20
30
40
50
60
70
3_Al Muruj Dist.
3_Al Rabie Dist.
3_Al Olaya Dist.
3_Al Sulaimaniyah Dist.
3_Al Yasmeen Dist.
3_Al Nuzha Dist.
3_West Suwaidi Dist.
3_Al Murabba Dist.
3_Al Maseef Dist.
3_Al Woroud Dist.
3_Al Mathar Dist.
3_Al Rawabi Dist.
3_Al Rahmaniyah Dist.
3_Salahuddin Dist.
3_North Mathar Dist.
3_Qurtubah Dist.
3_Al Malaz Dist.
3_Al Dhubbat Dist.
3_Al Rayan Dist.
3_Al Rabwah Dist.
3_Al Zahra Dist.
3_Al Salam Dist.
3_Al Dirah Dist.
3_Al Mansurah Dist.
3_Al Ghadeer Dist.
3_Al Dahou Dist.
3_Utayqah Dist.
3_Al Qura Dist.
3_Al Sahafah Dist.
3_Al Muhammadiyah Dist.
3_King Fahd Dist.
3_Al Ezdihar Dist.
3_Al Wizarat Dist.
3_Al Suwaidi Dist.
3_Al Futah Dist.
3_Al Nafel Dist.
3_Hitteen Dist.
3_Al Rawdah Dist.
3_West Umm Al Hamam Dist.
3_Al Hamra Dist.
3_Al Khalidiyah Dist.
3_Namar Dist.
3_Al Quds Dist.
3_Assafarat Dist.
3_Al Andalus Dist.
3_Al Shohda Dist.
3_Al Falah Dist.
3_Al Munisiyah Dist.
3_Al Taawun Dist.
3_Siyah Dist.

As shown in the above Figure most of Neighborhood with more than 30 venues are downtone Neighborhoods

- Common categories for current city

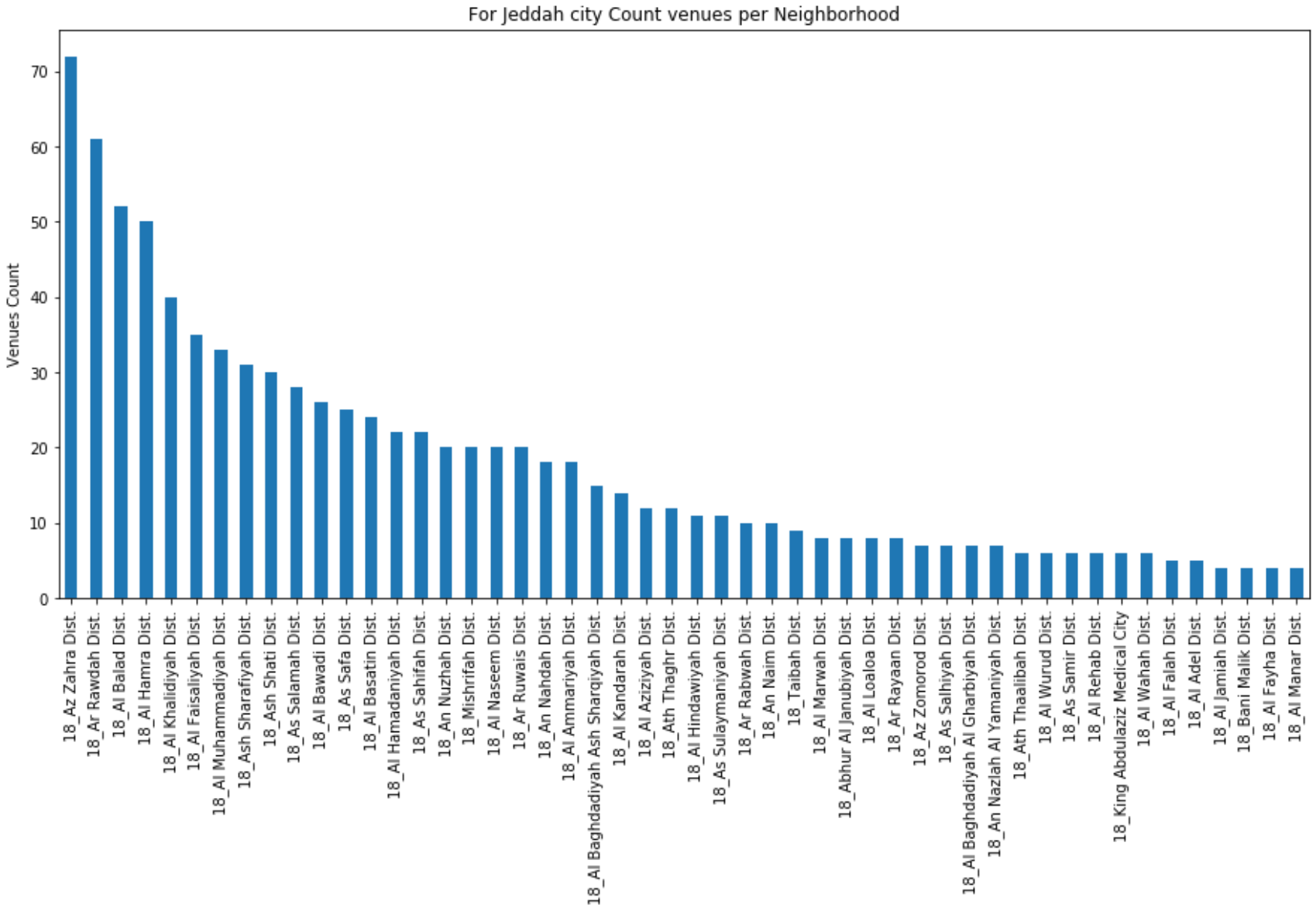


As shown most common categories in current city:

- Coffee Shop
- Café
- Middle Eastern Restaurant
- Fast Food
- Dessert Shop
- Hotel

and as notes there is ***Beach*** category appear in this city as its on Red Sea

Now let us list the current Neighborhood catagories

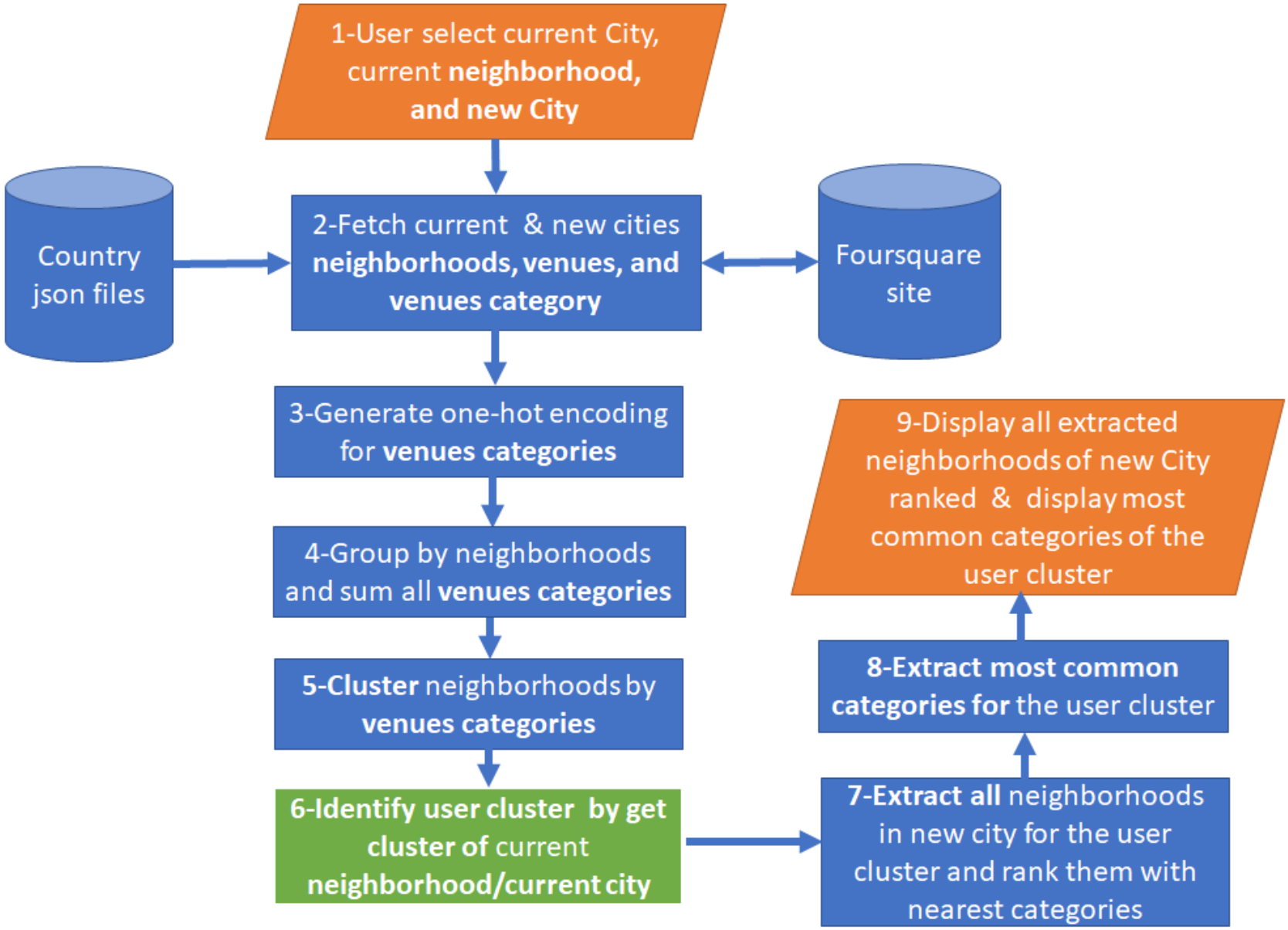

For Jeddah city Count venues per Neighborhood
Venues Count
0
10
20
30
40
50
60
70
18_Az Zahra Dist.
18_Ar Rawdah Dist.
18_Al Balad Dist.
18_Al Hamra Dist.
18_Al Khalidiyah Dist.
18_Al Faisaliyah Dist.
18_Al Muhammadiyah Dist.
18_Ash Sharafiyah Dist.
18_Ash Shati Dist.
18_As Salamah Dist.
18_Al Bawadi Dist.
18_As Safa Dist.
18_Al Basatin Dist.
18_Al Hamadaniyah Dist.
18_As Sahifah Dist.
18_An Nuzhah Dist.
18_Mishrifah Dist.
18_Al Naseem Dist.
18_Ar Ruwais Dist.
18_An Nahdah Dist.
18_Al Ammariyah Dist.
18_Al Baghdadiyah Ash Sharqiyah Dist.
18_Al Kandarah Dist.
18_Al Aziziyah Dist.
18_Ath Thaghr Dist.
18_Al Hindawiyah Dist.
18_As Sulaymaniyah Dist.
18_Ar Rabwah Dist.
18_An Naim Dist.
18_Taibah Dist.
18_Al Marwah Dist.
18_Abhur Al Janubiyah Dist.
18_Al Loaloa Dist.
18_Ar Rayaan Dist.
18_Az Zomorod Dist.
18_As Salhiyah Dist.
18_Al Baghdadiyah Al Gharbiyah Dist.
18_An Nazlah Al Yamaniyah Dist.
18_Ath Thaalibah Dist.
18_Al Wurud Dist.
18_As Samir Dist.
18_Al Rehab Dist.
18_King Abdulaziz Medical City
18_Al Wahah Dist.
18_Al Falah Dist.
18_Al Adel Dist.
18_Al Jamiah Dist.
18_Bani Malik Dist.
18_Al Fayha Dist.
18_Al Manar Dist.

As shown in the above Figure most of Neighborhood with more than 30 venues are downtone Neighborhoods

Now we are ready to move to next phase of project : ***Methodology*** to anlysis data and build suitable model

3-Methodology

The following Figure define the overall steps from getting user input to display output to user for moving from current new city to new city


1-User select current City, current neighborhood, and new City
Country json files
2-Fetch current & new cities neighborhoods, venues, and venues category
Foursquare site
3-Generate one-hot encoding for venues categories
4-Group by neighborhoods and sum all venues categories
5-Cluster neighborhoods by venues categories
6-Identify user cluster by get cluster of current neighborhood/current city
7-Extract all neighborhoods in new city for the user cluster and rank them with nearest categories
8-Extract most common categories for the user cluster
9-Display all extracted neighborhoods of new City ranked & display most common categories of the user cluster

*** Steps 1,2 already performed in Data Description & Preperatio phase***

a) Generate one-hot encoding for venues categories

We use One Hot Encoding for venues categories, allow us group neighborhoods of both cities, and then find out the top ten venues categories present in each neighborhood.

Out[35]:

	Neighborhood	city_id	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport Gate	Airport Lounge	American Restaurant	Antique Shop	Arcade	Arepa Restaurant	Art Gallery
0	18_Az Zomorod Dist.	18	0	0	0	0	0	0	0	0	0	0	0
1	18_Az Zomorod Dist.	18	0	0	0	0	0	0	0	0	0	0	0
2	18_Az Zomorod Dist.	18	0	0	0	0	0	0	0	0	0	0	0
3	18_Az Zomorod Dist.	18	0	0	0	0	0	0	0	0	0	0	0
4	18_Az Zomorod Dist.	18	0	0	0	0	0	0	0	0	0	0	0

As shown the catagories become a columns 276 (be execluse Neighborhood and city_id) and each venues will be assigned each 1 or 0 for each catagory

b) Group by neighborhoods and sum all venues categories

as shown in above one-hot encodong tabel, there are number of recoreds for each Neighborhood (venues), now we will group data by Neighborhood to have one value between Neighborhood and catagory as shown below:

Out[36]:

	Neighborhood	city_id	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport Gate	Airport Lounge	American Restaurant	Antique Shop	Arcade	Arepa Restaurant	Art Gallery
0	18_Abhur Al Janubiyah Dist.	18	0	0	0	0	0	0	0	0	0	0	0
1	18_Al Adel Dist.	18	0	0	0	0	0	0	0	0	0	0	0
2	18_Al Ajwad Dist.	18	0	0	0	0	0	0	0	0	0	0	0
3	18_Al Amir Fawaz Al Janouby Dist.	18	0	0	0	0	0	0	0	0	0	0	0
4	18_Al Amir Fawaz Ash Shamaly Dist.	18	0	0	0	0	0	0	0	0	0	0	0

As shown the rows become 221, which are number of Neighborhoods in both cities and each Neighborhood row contains number for each catagory represent number of venues belong to this catagory in this area

- Calculating most common venues catagories for each Neighborhood

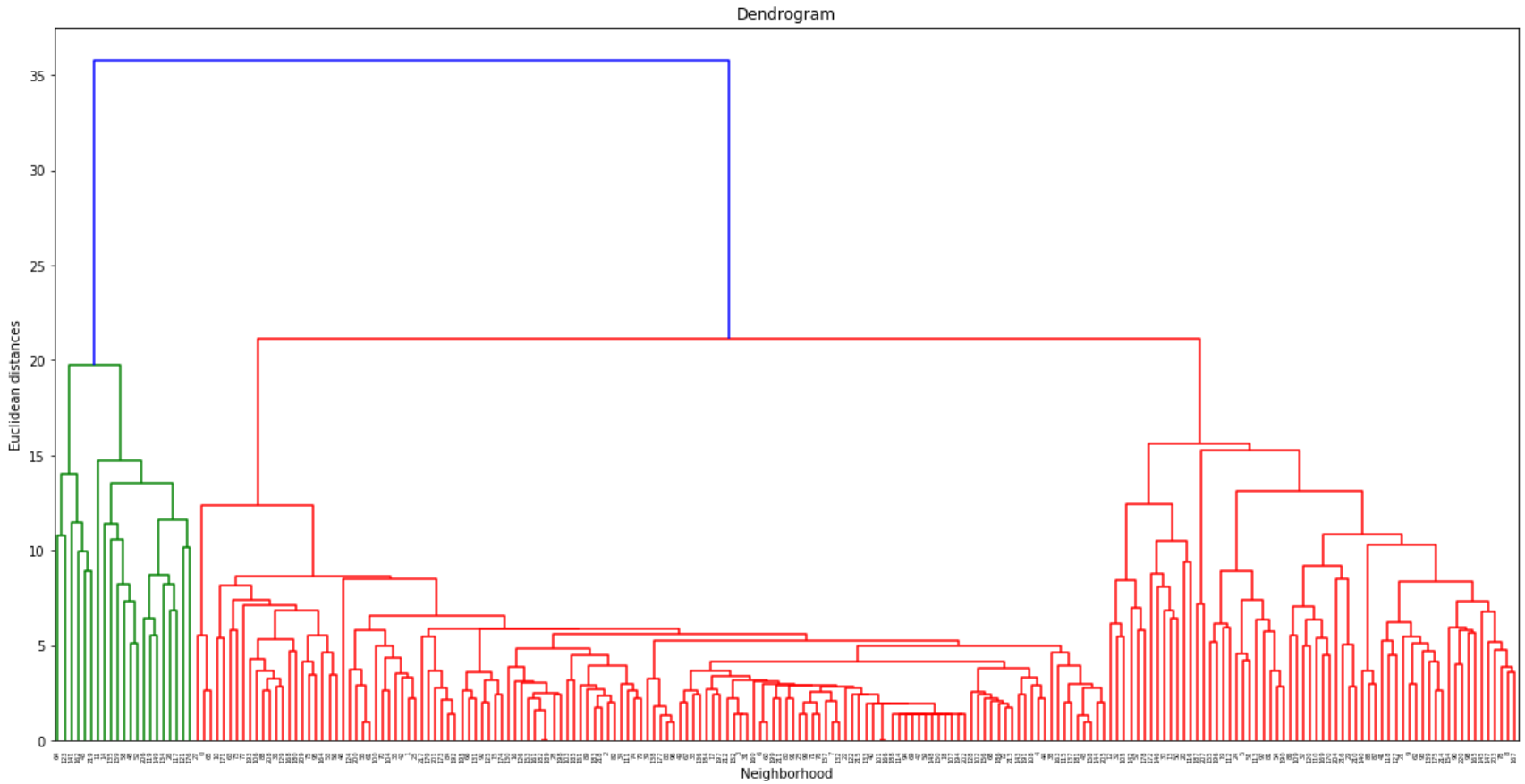
as shonw Neighborhood appears associated with most common categories sorted with most catagories first

Out[38]:

	city_id	Neighborhood	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	18	18_Abhur Al Janubiyah Dist.	Beach	Seafood Restaurant	Boat Rental	Fountain	Restaurant	Surf Spot	Yoga Studio	Filipino Restaurant	Farm	Farmers Market
1	18	18_Al Adel Dist.	Gym	Dessert Shop	Gym / Fitness Center	Arts & Crafts Store	Donut Shop	Food & Drink Shop	Food	Flower Shop	Flea Market	Fish Market
2	18	18_Al Ajwad Dist.	Market	Hookah Bar	Burger Joint	Food Truck	Department Store	Dessert Shop	Food Service	Food Court	Food & Drink Shop	Food
3	18	18_Al Amir Fawaz Al Janouby Dist.	Trail	Mini Golf	Yoga Studio	Fish & Chips Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish Market
4	18	18_Al Amir Fawaz Ash Shamaly Dist.	Turkish Restaurant	Hospital	Seafood Restaurant	Yoga Studio	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant

c) Cluster neighborhoods by venues categories

will use hierarchical clustering instead of k-mean, as hierarchical clustering has fewer assumptions about the distribution of your data - the only requirement (which k-means also shares) is that a distance can be calculated each pair of data points. Hierarchical clustering typically 'joins' nearby points into a cluster, and then successively adds nearby points to the nearest group. You end up with a 'dendrogram', or a sort of connectivity plot. You can use that plot to decide after the fact of how many clusters your data has, by cutting the dendrogram at different heights. Of course, if you need to pre-decide how many clusters you want (based on some sort of business need) you can do that too. Hierarchical clustering can be more computationally expensive but usually produces more intuitive results.

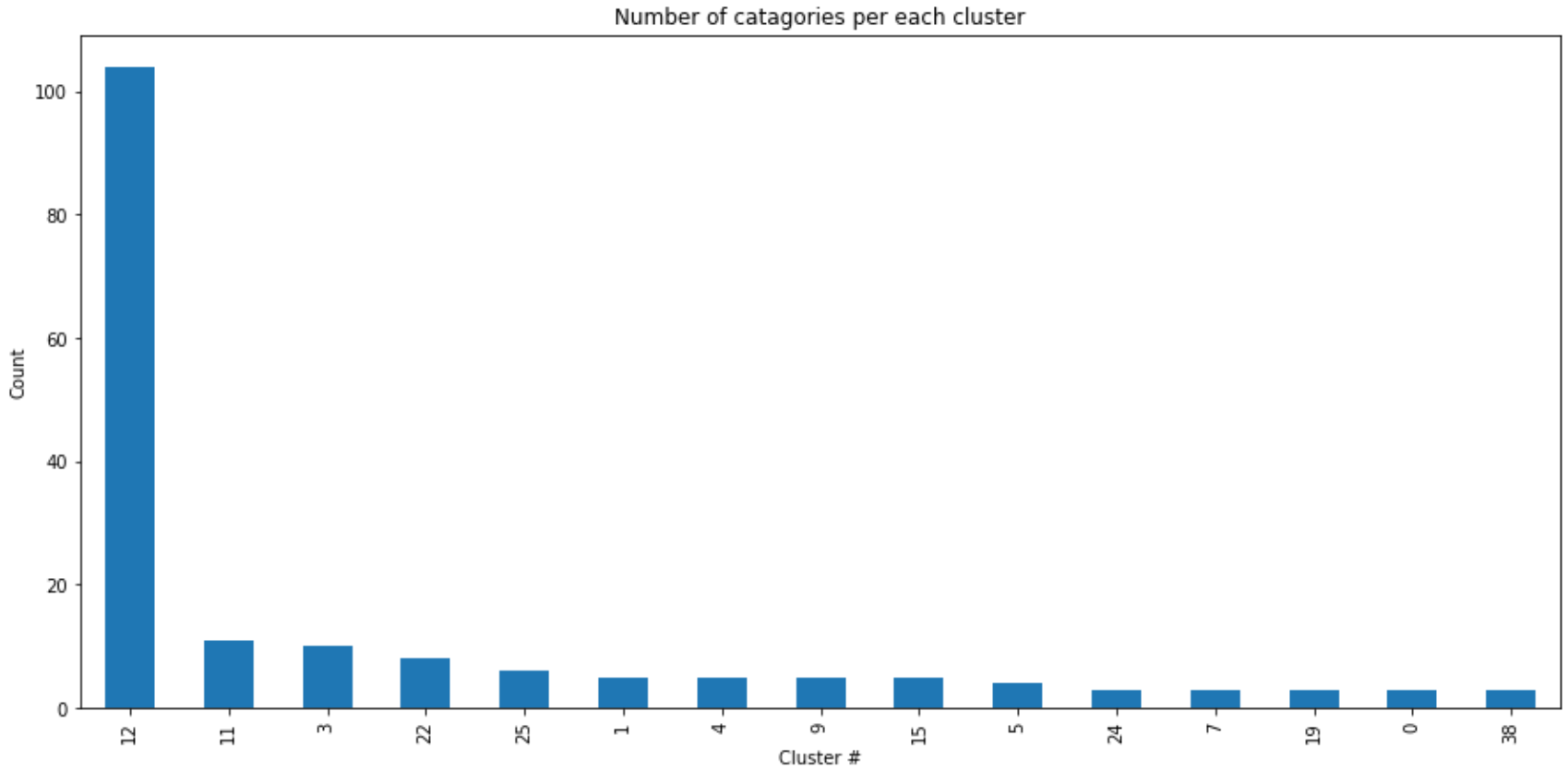


As shown from the above dendrogram cutting at 50 clusters is more suitable which will make most of clusters Neighborhoods around 10 Neighborhoods

As shown below table depicts the clustered data along with the top 10 most common venues catagoires in that cluster.

Out[40]:

	city_id	Neighborhood	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	18	18_Abhur Al Janubiyah Dist.	Beach	Seafood Restaurant	Boat Rental	Fountain	Restaurant	Surf Spot	Yoga Studio	Filipino Restaurant	Farm	Farmers Market
1	18	18_Al Adel Dist.	Gym	Dessert Shop	Gym / Fitness Center	Arts & Crafts Store	Donut Shop	Food & Drink Shop	Food	Flower Shop	Flea Market	Fish Market
2	18	18_Al Ajwad Dist.	Market	Hookah Bar	Burger Joint	Food Truck	Department Store	Dessert Shop	Food Service	Food Court	Food & Drink Shop	Food
3	18	18_Al Amir Fawaz Al Janouby Dist.	Trail	Mini Golf	Yoga Studio	Fish & Chips Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish Market
4	18	18_Al Amir Fawaz Ash Shamaly Dist.	Turkish Restaurant	Hospital	Seafood Restaurant	Yoga Studio	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant


Number of catagories per each cluster
Count
0
20
40
60
80
100
12
11
3
22
25
1
4
9
15
5
24
7
19
0
38
Cluster #

Display most common categories in cluster

Cluster:12

- First 5 Neighborhoods in this cluster

	city_id	Neighborhood	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
2	18	18_Al Ajwad Dist.	Market	Hookah Bar	Burger Joint	Food Truck	Department Store	Dessert Shop	Food Service	Food Court	Food & Drink Shop	Food
3	18	18_Al Amir Fawaz Al Janouby Dist.	Trail	Mini Golf	Yoga Studio	Fish & Chips Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish Market
4	18	18_Al Amir Fawaz Ash Shamaly Dist.	Turkish Restaurant	Hospital	Seafood Restaurant	Yoga Studio	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant
6	18	18_Al Amwaj Dist.	Beach	Yoga Studio	Flea Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flower Shop
7	18	18_Al Andalus Dist.	Spa	Pizza Place	Yoga Studio	Fish & Chips Shop	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant

- The most five common categories in this cluster

	Catagory	Count
0	Fast Food Restaurant	84
1	Farm	82
2	Farmers Market	81
3	Falafel Restaurant	75
4	Filipino Restaurant	70

- The number of Neighborhoods per each city in this cluster

	City	Count
0	3	73
1	18	31

Cluster:11

- First 5 Neighborhoods in this cluster

	city_id	Neighborhood	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 Mos Commor Venue
1	18	18_Al Adel Dist.	Gym	Dessert Shop	Gym / Fitness Center	Arts & Crafts Store	Donut Shop	Food & Drink Shop	Food	Flower Shop	Flea Market	Fish Marke
25	18	18_Al Kawthar Dist.	Gym / Fitness Center	Café	Dessert Shop	Fish & Chips Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish Marke
35	18	18_Al Wahah Dist.	Burger Joint	Turkish Restaurant	Gym / Fitness Center	Bathing Area	Bar	Yoga Studio	Farm	Farmers Market	Fast Food Restaurant	Filipino Restauran
42	18	18_Ar Rabwah Dist.	Gym / Fitness Center	Market	Café	Pharmacy	African Restaurant	Hookah Bar	American Restaurant	Gym	Burger Joint	Fast Food Restauran
55	18	18_As Sororyah Dist.	Furniture / Home Store	Fish Market	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Yoga Studio	Exhibi

- The most five common categories in this cluster

	Catagory	Count
0	Fast Food Restaurant	7
1	Fish & Chips Shop	6

	Catagory	Count
2	Farm	6
3	Fish Market	6
4	Farmers Market	5

- The number of Neighborhoods per each city in this cluster

	City	Count
0	18	7
1	3	4

Cluster:3

- First 5 Neighborhoods in this cluster

	city_id	Neighborhood	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 Com V
9	18	18_Al Baghdadiyah Al Gharbiyah Dist.	Hotel	Seafood Restaurant	Ice Cream Shop	Supermarket	Café	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Resta
21	18	18_Al Hindawiyah Dist.	African Restaurant	Food	Coffee Shop	Hotel	Afghan Restaurant	Snack Place	American Restaurant	Breakfast Spot	Fast Food Restaurant	
41	18	18_An Nuzhah Dist.	Hotel	Pizza Place	Moving Target	Intersection	Bridal Shop	Grocery Store	Clothing Store	Trail	Coffee Shop	Disc
62	18	18_Ath Thaalibah Dist.	Coffee Shop	Seafood Restaurant	Hotel	Restaurant	Filipino Restaurant	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Resta
93	3	3_Al Falah Dist.	Coffee Shop	Hotel	Grocery Store	Donut Shop	Theater	Convenience Store	Food Court	Heliport	Salon / Barbershop	Fast Resta

- The most five common categories in this cluster

	Catagory	Count
0	Coffee Shop	9
1	Hotel	6
2	Fast Food Restaurant	6
3	Farm	5
4	Farmers Market	4

- The number of Neighborhoods per each city in this cluster

	City	Count
0	3	6
1	18	4

Cluster:22

- First 5 Neighborhoods in this cluster

	city_id	Neighborhood	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 Co
36	18	18_Al Wurud Dist.	Hotel	Department Store	Bed & Breakfast	Park	Asian Restaurant	Middle Eastern Restaurant	Yoga Studio	Fish Market	Farmers Market	Fas Rest
88	3	3_Al Dubiyah Dist.	Coffee Shop	Flower Shop	Park	Pet Store	Pakistani Restaurant	Flea Market	Food	Fish Market	Fish & Chips Shop	
106	3	3_Al Khuzama Dist.	Park	Stables	Farm	Vacation Rental	Fast Food Restaurant	Yoga Studio	Filipino Restaurant	Fabric Shop	Falafel Restaurant	Fa

	city_id	Neighborhood	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 Co
129	3	3_Al Nakheel Dist.	Park	Middle Eastern Restaurant	Grocery Store	Pharmacy	Burger Joint	Yoga Studio	Falafel Restaurant	Farm	Farmers Market	Fas Rest
168	3	3_Al Wahah Dist.	Lounge	Hookah Bar	Restaurant	Park	Trail	Gym / Fitness Center	Yoga Studio	Farm	Exhibit	

- The most five common categories in this cluster

	Catagory	Count
0	Park	8
1	Yoga Studio	5
2	Fast Food Restaurant	4
3	Farmers Market	4
4	Grocery Store	3

- The number of Neighborhoods per each city in this cluster

	City	Count
0	3	7
1	18	1

Cluster:25

- First 5 Neighborhoods in this cluster

	city_id	Neighborhood	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	10 C
53	18	18_As Salhiyah Dist.	Fast Food Restaurant	Music Venue	Dessert Shop	Campground	Bus Station	Tram Station	Yoga Studio	Filipino Restaurant	Farm	F
56	18	18_As Sulaymaniyah Dist.	Fast Food Restaurant	Ice Cream Shop	Grocery Store	Falafel Restaurant	Burger Joint	Trail	Middle Eastern Restaurant	Dessert Shop	Pharmacy	
75	18	18_Taibah Dist.	Pizza Place	Fast Food Restaurant	Turkish Restaurant	Juice Bar	Pharmacy	Dessert Shop	Doner Restaurant	Falafel Restaurant	Farm	F
95	3	3_Al Fayha Dist.	Juice Bar	Pizza Place	Middle Eastern Restaurant	Burger Joint	Fast Food Restaurant	Supermarket	Donut Shop	Turkish Restaurant	Breakfast Spot	Res
164	3	3_Al Suwaidi Dist.	Fast Food Restaurant	Burger Joint	Ice Cream Shop	Middle Eastern Restaurant	Pet Store	Convenience Store	Park	Candy Store	Breakfast Spot	

- The most five common categories in this cluster

	Catagory	Count
0	Fast Food Restaurant	5
1	Middle Eastern Restaurant	4
2	Juice Bar	3
3	Burger Joint	3
4	Farmers Market	3

- The number of Neighborhoods per each city in this cluster

	City	Count
0	3	3
1	18	3

The summary of most common catagories for first 5 clusters

Out[541]:

	cluster:12	cluster:11	cluster:3	cluster:22	cluster:25
0	Fast Food Restaurant	Fast Food Restaurant	Coffee Shop	Park	Fast Food Restaurant
1	Farm	Fish Market	Hotel	Yoga Studio	Middle Eastern Restaurant
2	Farmers Market	Fish & Chips Shop	Fast Food Restaurant	Fast Food Restaurant	Burger Joint
3	Falafel Restaurant	Farm	Farm	Farmers Market	Dessert Shop
4	Filipino Restaurant	Farmers Market	Grocery Store	Farm	Juice Bar

4-Results

a) Identify user cluster by get cluster of current neighborhood current city

After get the current neighborhood in current city, will display this neighborhood associted with Most Common catagory and cluster number

Out[47]:

	city_id	Neighborhood	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	Cl
75	18	18_Taibah Dist.	Pizza Place	Fast Food Restaurant	Turkish Restaurant	Juice Bar	Pharmacy	Dessert Shop	Doner Restaurant	Falafel Restaurant	Farm	Farmers Market	

b) Extract all neighborhoods in new city for the user cluster and rank them with nearest categories

25

Out[48]:

	city_id	Neighborhood	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
95	3	3_Al Fayha Dist.	Juice Bar	Pizza Place	Middle Eastern Restaurant	Burger Joint	Fast Food Restaurant	Supermarket	Donut Shop	Turkish Restaurant	Breakfast Spot
164	3	3_Al Suwaidi Dist.	Fast Food Restaurant	Burger Joint	Ice Cream Shop	Middle Eastern Restaurant	Pet Store	Convenience Store	Park	Candy Store	Breakfast Spot
209	3	3_Shubra Dist.	Food Truck	Shawarma Place	Bakery	Restaurant	Gym	Gym / Fitness Center	Supermarket	Juice Bar	Farmers Market

d) Display all extracted neighborhoods of new City ranked

as shown list all candidate Neighborhoods name in the new city ranked with most similar Neighborhoods

```
Nesrest Neighborhoods for  18_Taibah Dist.
```

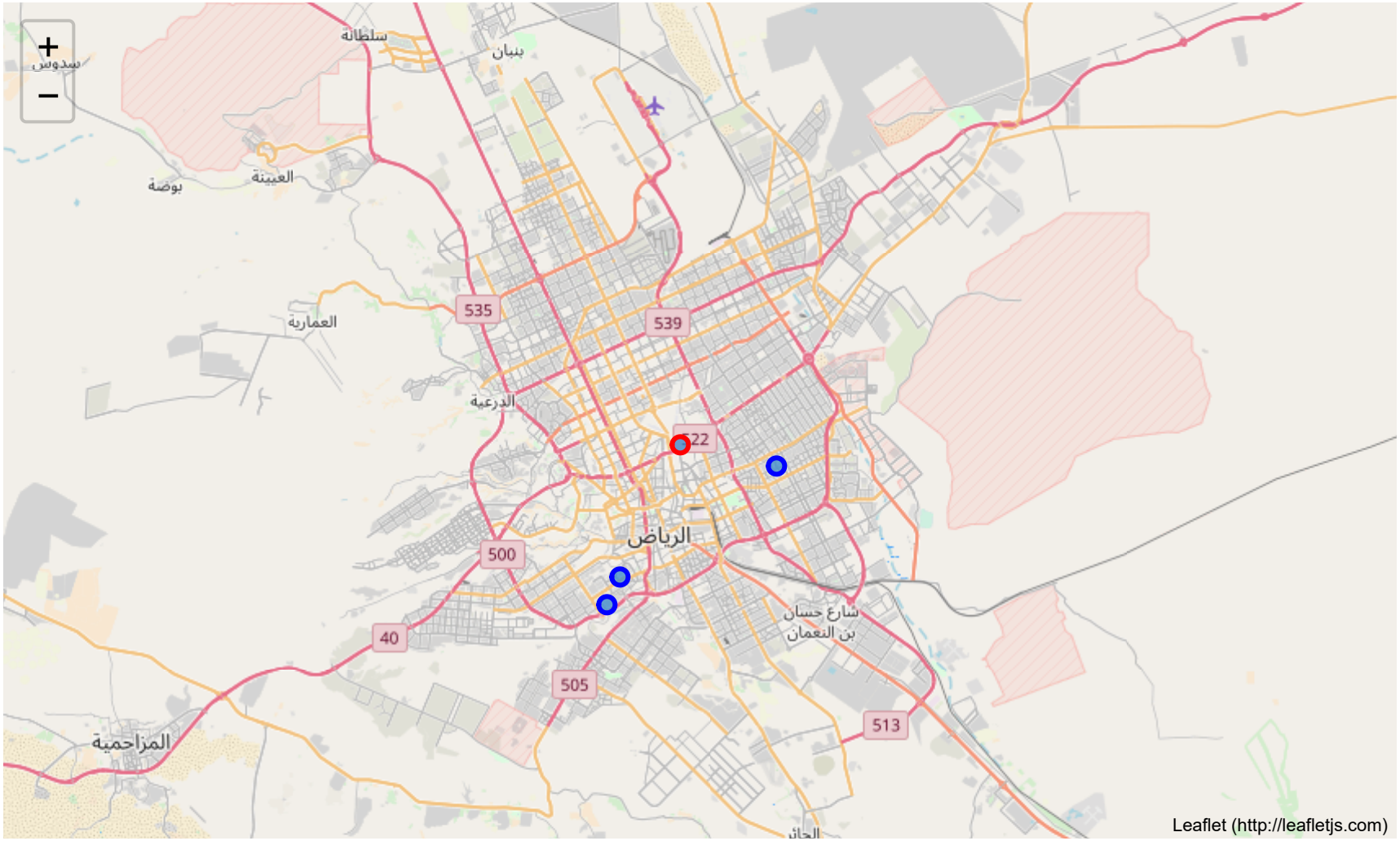
Out[50]:

	candidate_Neighborhoods	similarity
0	3_Al Fayha Dist.	0.288675
2	3_Shubra Dist.	0.235702
1	3_Al Suwaidi Dist.	0.200000

The Customer cam now see the candidate Neighborhoods on the map as following

```
Riyadh
```

Out[59]:



c) most common categories of the user cluster

Extract most 10 common categories for the user cluster & display it

Out[52]:
```
Fast Food Restaurant         5
Middle Eastern Restaurant    4
Juice Bar                    3
Burger Joint                 3
Farmers Market               3
Dessert Shop                 3
Falafel Restaurant           3
Supermarket                  2
Ice Cream Shop               2
Pizza Place                  2
Breakfast Spot               2
Turkish Restaurant           2
Farm                         2
Pharmacy                     2
Bakery                       1
Name: Venue, dtype: int64
```

5-Discussion

The intent with which analysis was carried out was to find out similar neighborhoods for a person relocating from current city to new city.

As we analyze the results section, we can analyze the clusters and see similar neighborhoods in the new city. as the example we have, if we compare the Current Neighborhood with the best candidate Neighborhood in the same cluster .

List 1: the Current Neighborhood: "18_Taibah Dist." and list of nearest venues catagories

```
                     catagory
Fast Food Restaurant         2
Pizza Place                  2
Dessert Shop                 1
Juice Bar                    1
Pharmacy                     1
Doner Restaurant             1
Turkish Restaurant           1
```

List 2: the best candidate Neighborhood ('3_Al Fayha Dist.') and list of nearest venues catagories

The Neighborhood '3_Al Fayha Dist.' selected as it has the best similarity score 0.288675 as shown above

	catagory
Juice Bar	2
Turkish Restaurant	1
Donut Shop	1
Middle Eastern Restaurant	1
Falafel Restaurant	1
Supermarket	1
Burger Joint	1
Fast Food Restaurant	1
Pizza Place	1
Breakfast Spot	1

As seen in the above List 1,2 current & the best candidate Neighborhoods have a lot of similar venues catagories:

- Fast Food Restaurant
- Pizza Place
- Juice Bar
- Turkish Restaurant

If a person’s current location were in the Neighborhood of 18_Taibah District in in Jedda city, which has venues like Fast Food, Pizza, Juice Bar and Turkish Restaurant nearby, the person, would like to relocate to a neighborhood like Fayha District in Riyadh city which also has venues like Fast Food, Pizza, Juice Bar and Turkish Restaurant nearby. This is just one example of how our data analysis can help people relocate from one city to another which similar to their current localities.

6-Conclusion

In a fast moving world, there are many real life problems or scenarios where data can be used to find solutions to those problems. Like seen in the example above, data was used to cluster neighborhoods in KAS cities based on the most common venues in those neighborhoods. Similarly, data can also be used to solve other problems, which most people face in metropolitan cities.