

IBM

Data Science Professional

Find the best location for new Coffee Shop in Riyadh

THE BATTLE OF NEIGHBORHOODS

Eman AlRefaai | Applied Data Science Capstone | Aug-2020

Introduction

Our client want to open new coffee shop in Riyadh city and he is asking about the best location for his business. For sure he can change anything after he open the coffee shop but not the location, so we have to help him choose the right location, which is the one with lowest number of coffee shops for less competitions.

BUSINESS CHANLLENGES

To find the best location for new coffee shop we should focus of what we want to have and what we don't want to have. we want a place with traffic, coffee-drinker people like college, schools, malls..etc. and what we don't want to have is competitors from other coffee shops.

I will use Foursquare database to provide our client the best location by clustering the areas by its distance from the center of the city and find the area with lowest number of coffee shops.

DATA REQUIREMENTS

For this assignment, I will be utilizing the Foursquare API to pull the following location data in Riyadh

- Venue Name
- Venue ID
- Venue Location
- Venue Category
- venue distance

I will cluster the areas according to distance from the center and number of venues category equal to Coffee Shop.

LOAD AND EXPLORE DATA

What we want is a place in Riyadh that has coffee drinkers like the student so I focus on the colleges and universities.

1. Get the list of venues categories in Foursquare

So, I need to get the college information in Riyadh using foursquare but first I need to know the college and universities category id in **Foursquare**, so I load these information using the following URL

<https://api.foursquare.com/v2/venues/categories>

the result of categories is as follows:

	id	name
0	4d4b7104d754a06370d81259	Arts & Entertainment
1	4d4b7105d754a06372d81259	College & University
2	4d4b7105d754a06373d81259	Event
3	4d4b7105d754a06374d81259	Food
4	4d4b7105d754a06376d81259	Nightlife Spot
5	4d4b7105d754a06377d81259	Outdoors & Recreation
6	4d4b7105d754a06375d81259	Professional & Other Places
7	4e67e38e036454776db1fb3a	Residence
8	4d4b7105d754a06378d81259	Shop & Service
9	4d4b7105d754a06379d81259	Travel & Transport

2. Get the colleges in Riyadh with its coordinates

First, I get the Latitude and Longitude of Riyadh city from a repository on GitHub that has all cities coordinates

	city	lat	lng	country	iso2	admin	capital	population	population_proper
0	Riyadh	24.653837	46.715683	Saudi Arabia	SA	Ar Riyāḍ	primary	4465000.0	4205961.0
1	Jeddah	21.542381	39.197969	Saudi Arabia	SA	Makkah al Mukarramah	NaN	3012000.0	2867446.0
2	Ad Dammām	26.434417	50.103263	Saudi Arabia	SA	Ash Sharqīyah	admin	2054710.0	768602.0
3	Medīna	24.468611	39.614167	Saudi Arabia	SA	Al Madīnah al Munawwarah	admin	1010000.0	1010000.0
4	Aṭ Ṭa'if	21.270278	40.415833	Saudi Arabia	SA	Makkah al Mukarramah	NaN	657282.0	530848.0

Then, I called Foursquare explore function to get the colleges in Riyadh using the category id and Riyadh latitude and longitude

CLUSTERING COLLEGES WITH NEARBY COFFE SHOPS

I did clustering by the following steps

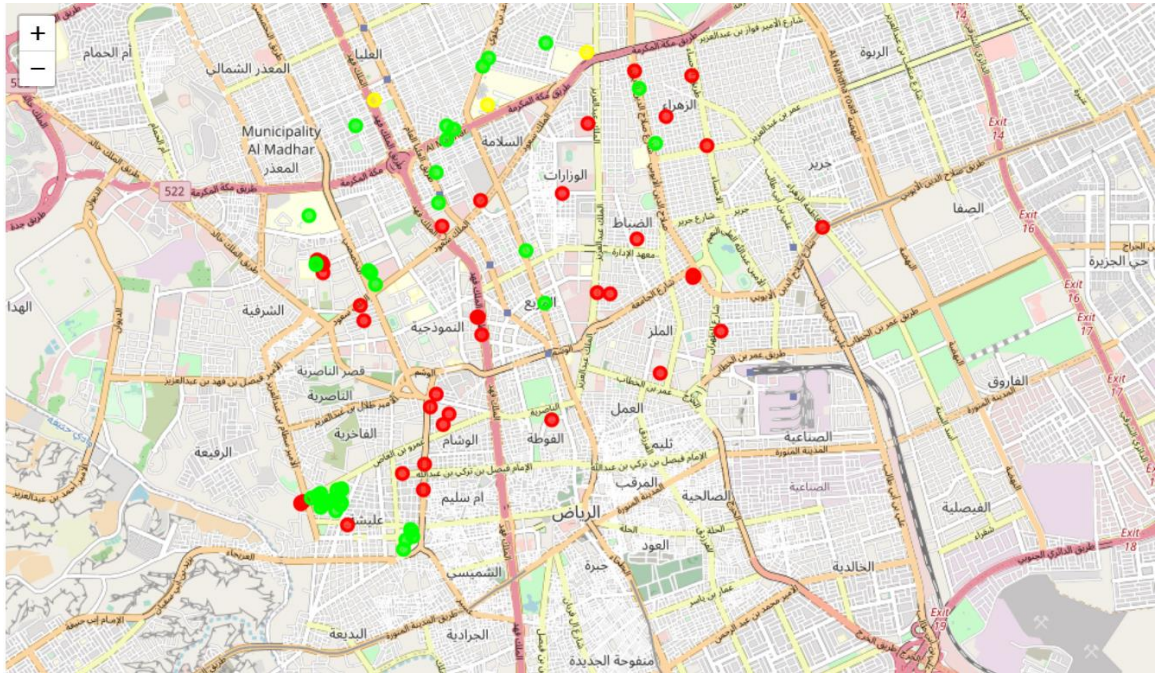
- 1- First, I grouped each college by the number of nearby coffee shops

	College	CollegeLatitude	CollegeLongitude	CF Count
0	02GA010	24.633398	46.679511	9
1	AU Library (مكتبة جامعة الفيصل)	24.665466	46.676914	2
2	Academic Tower KSMC Calss 2	24.629217	46.690792	6
3	Al Thaqafi Center Higher Cultural Institute fo...	24.682908	46.697048	8
4	Alfaisal University (جامعة الفيصل)	24.664931	46.676748	8
5	Building 6 (مبنى ٦)	24.663244	46.732449	4
6	Building No.3 KSU	24.631871	46.679717	8
7	Bulding 5 - King Abdulaziz University Hospital	24.660911	46.720102	4
8	College Of Applied Medical Sciences (CAMS)	24.632779	46.674634	5
9	College Of Arts - English Literature	24.632958	46.680605	8
10	College Of Dentistry (كلية طب الأسنان)	24.657858	46.700616	2
11	College Of Law And Political Science	24.633602	46.676186	8
12	College of Business - Alfaisal University	24.664911	46.677711	3
13	College of Business Administration , Building ...	24.633404	46.679296	9
14	College of Engineering - Alfaisal University	24.663690	46.677825	4
15	College of Medicine - Alfaisal University	24.664734	46.677896	4
16	College of Science - Alfaisal University	24.665140	46.677643	3
17	Der AlZahraa Bookstore (مكتبة دار الزهراء)	24.681508	46.696095	8

- 2- Then I sorted them by the coffee shops count.
- 3- K-mean is not directly applicable to categorical variables so I dropped the “College Name” from the data frame.
- 4- Run k-mean to cluster the data into 3 clusters
- 5- Add cluster label to the data frame, data will be as follows:

	Cluster Labels	College	CollegeLatitude	CollegeLongitude	CF Count
29	0	Institute of Saudi Electricity Company	24.650254	46.727350	1
82	0	مدارس التضامن الأهلية	24.638183	46.692760	1
39	0	Prince Saud Al-Faysal Institute of Diplomatic ...	24.645778	46.693649	1
72	0	جلسة مجلس الامن الدولي	24.634670	46.692614	1
80	0	كلية امن الشبكات جامعة الملك سعود	24.647422	46.694448	1

6- Finally, I visualize the resulted cluster



CONCLUSION

from the result (green spots), south west Riyadh is the best place for new coffee shop as it has lots of coffee-drinker people with less competitors.

