Peer-graded Assignment: Capstone Project - The Battle of Neighborhoods (Week 2)

TABLE OF CONTENT

1 - Introduction

Business Problem

I would like to open a new italian restaurant in one of biggest city in Europe. To identify a proper location I will evaluate the lease rate and the presence of other italian restaurant. I would like to evaluate systematically a proper solution because the budget is fixed thus I would like to be quite sure to be able to set up the restaurant and having enough visibility in the selected city.

It is also important to identify the district of the city because we're evaluating big cities having many different diversity inside so the district is the minimun geographic level for our evaluation.

Who would be interested in this project.

The audience interested could be someone who want to invest his money in the Restaurants having already culinary experience and want to minimize the risks.

2 - Data

Data needed to analyze the problem are

- a **list of cities** to be analyized and all relevant **district** because this is our minimun geographic level of evaluation.

For our purpose we're evaluating two cities : Rome and London

- for each **district** having the **polygon** representation as json file

a json file with the polygon coordinates have been retrived from European-Neighborhoods-Json-Coords-master.

- ... for each distric- city to determine the **number of italian restaurant** and the **lease rete**

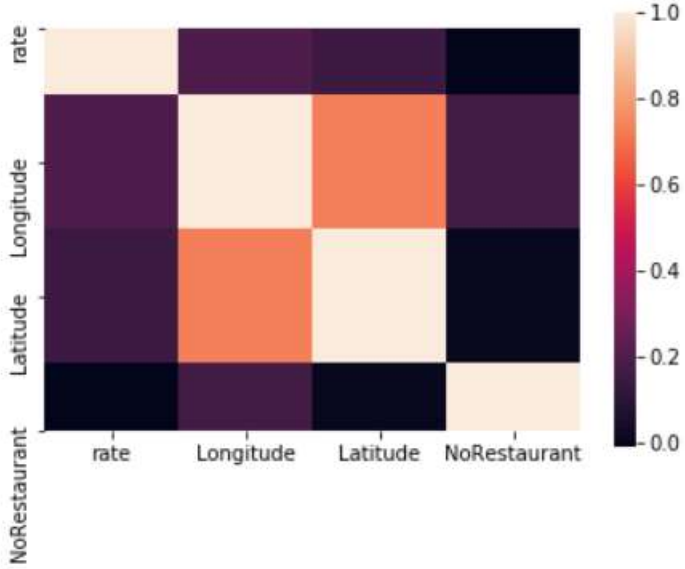
The number of italian restaurant have been retrived using Foursquare and quering "italian Restaurant" by distric coordinates

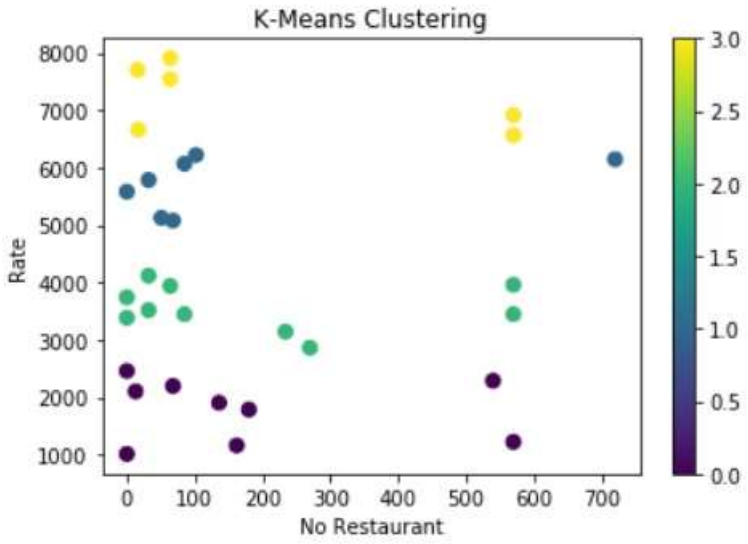
3 - Metodology

Out target is determine districts having similarity in the same city and districts similar between Rome and London. The similarity is based on two numeric indicator :

- Number of Italian Restaurant
- Lease rate of the district

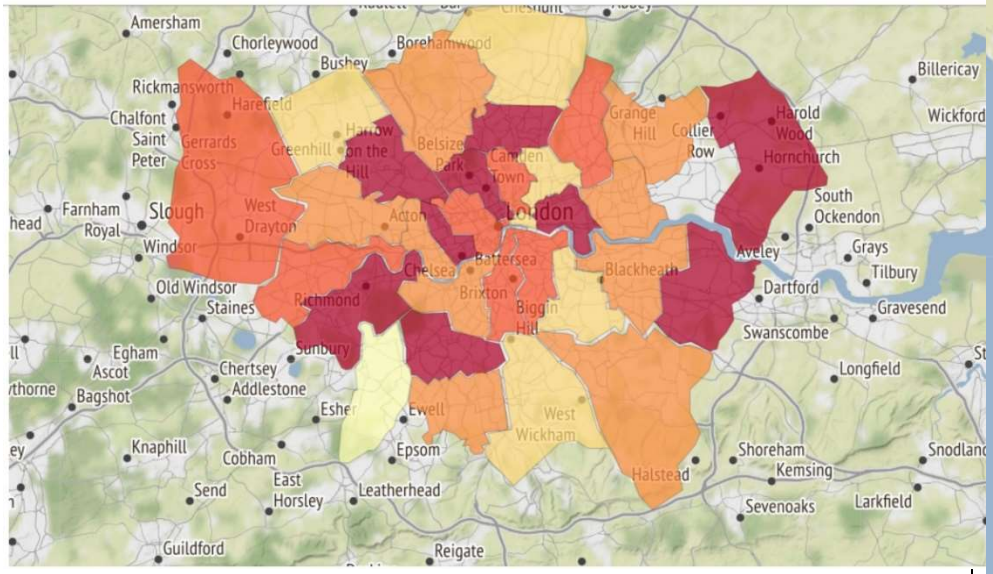
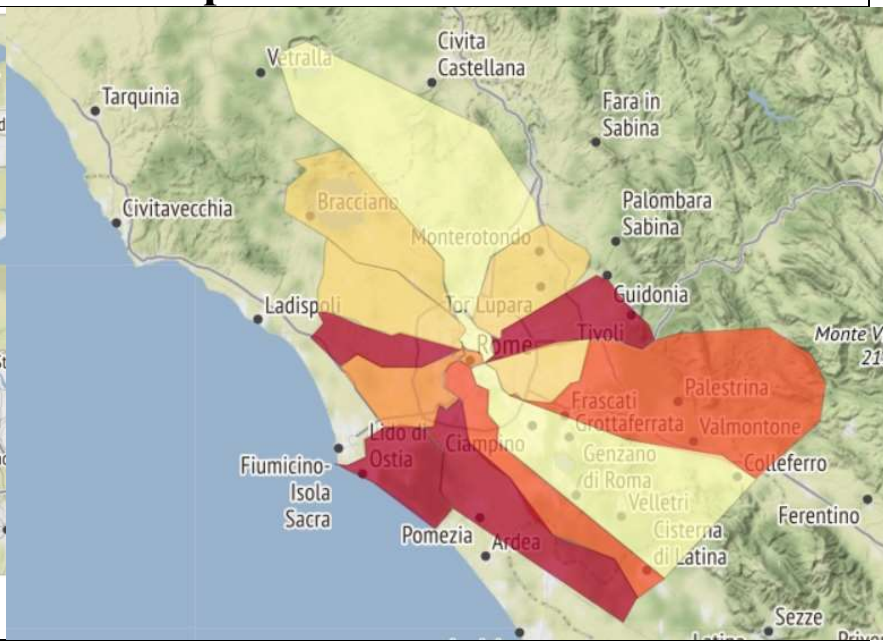
Methodology used is clustering the districts of Rome and London using k-Mean algorithm and representing the result by a clorophel map and the list of districts of both cities by cluster.
On the left the correlation among the attributes available. The couple No Restaurant and rate (=lease rate) are strong correlated. On the right the k-mean calculation on Number of restaurant and rate showing 4 different cluster.





4 – Result

Here below the geographical clusterization between London and Rome where the same colour means a cluster having the same behaver.

London map	Rome map
	

and the list table of cluster where for each cluster are merged together the districts of both cities.

Cluster	City	District	Latitude	Longitude	NoRestaurant	Lease rate
1000	Roma	Parioli	49.42911	11.01737	0	2,865
		Monte Sacro	37.60482	-0.98189	24	2,254
		Prenestino Centocelle	41.88451	12.56718	510	2,491
		Aurelia	42.13718	11.78496	0	2,667
		Monte Mario	41.92440	12.45213	510	1,726
		Flaminio	41.91283	12.47714	540	2,122
	London	City	51.51562	-0.09200	540	4,378
		Croydon	51.37130	-0.10196	51	3,693
		Enfield	50.71645	-106.66213	0	4,239
		Hackney	51.54324	-0.04936	180	4,893
		Harrow	51.59677	-0.33728	64	4,020
		Lewisham	51.46243	-0.01013	234	4,803
2000	Roma	Center Prati	42.29278	13.94995	0	5,735
		Arvalia Portuense	41.84248	12.40774	170	5,229
	London	Barnet	44.29732	-72.04971	0	7,993
		Bromley	51.40280	0.01481	32	7,622
		Ealing	51.51266	-0.30520	162	7,653
		Greenwich	52.03673	1.16893	32	7,625
		Hammersmith and Fullham	51.48487	-0.20188	570	7,490
		Newham	51.53000	0.02932	68	7,063
		Redbridge	51.57632	0.04541	32	7,602
		Sutton	49.91727	-106.20154	0	7,059
		Wandsworth	51.45703	-0.19326	270	7,196
3000	Roma	Roma Delle Torri	41.85592	12.65625	128	7,813
		Cinecitta	41.85259	12.56917	510	7,942
		Appia Antica	41.12257	13.91131	12	7,568
	London	Wesminster	37.71192	-90.42218	0	2,076
		Hillingdon	51.54252	-0.44833	16	2,611
		Hounslow	51.46861	-0.36135	68	1,591
		Islington	51.53843	-0.09991	570	2,430
		Lambeth	51.50130	-0.11729	570	1,676
		Southwark	51.50292	-0.10346	570	1,164

		Waltham Forest	51.55700	-0.00584	85	2,220
4000	Roma	Tiburtina	25.66592	-100.37696	255	4,342
		EUR	50.06664	10.12590	0	4,481
		Ostia	-31.45507	-64.20182	34	3,324
		Monte Verde	-47.00873	-67.10830	0	4,165
	London	Bexley	51.44168	0.15049	64	5,227
		Brent	39.05478	-84.43355	13	6,141
		Camden	39.94484	-75.11989	720	6,124
		Haringey	51.58793	-0.10541	136	5,325
		Havering	51.03586	-2.84175	0	6,151
		Kensington and Chelsea	51.48754	-0.16822	570	6,028
		kingston upon Thames	51.40963	-0.30626	64	6,255
		Merton	51.41080	-0.18810	102	6,318
		Richmond upon Thames	51.44035	-0.30815	85	6,588
		Tower Hamlets	51.12886	1.29867	17	5,619

5 – Discussion

Applying the process of analysis described we can show that London and Rome even they are far city from each other, are quite similar if we evaluate the distribution of Italian restaurant and the lease rent.
Both cities are in Europe and are adopting the same commercial rules on the same market. If an Investor would open a Italian restaurant chain in both cities can encounter the same condition.

It becomes strategic for the investor to identify the proper cluster and based on that open the restaurant having the same identity and branding.

6 - Conclusion

The model described can be used in a similar manner to reach many different purpose. Can be possibile to elevate the number of attribute to use, , not only the number of restaurant and lease rate and to extend the analysis toward a bigger set of cities.
This model can support an Investor to identify new opportunities minimizing the risk.