Coursera Capstone Final Assignment

Placing a sucessful restaurant in Paris

Introduction

In a world where competitiveness is at his highest peak, a simple, modest pastry is trying to fit into Paris. As anyone else, this restaurant wants to be very successful. In order to accomplish this dream they know that the very first step to make it is to decide a location to sell the deliciousness they know they can produce.¶

After some brainstorning they realized they had to come up with certain factors that will impact their decision to choose a location. After a while they come up with security, competition and population. In a city as big as Paris, they wanted to go to a safe location where crime could not affect the costs that they might have. Also, competition must be low so that they could stand out among other restaurants. Finally, without enough clients they wouldn't make it so it's important that they settle in an area with a lot of populations.

So, now that the factors are decided, it's important to analyze data to make the best decision possible.

DataFrame

In order to make the best decision, based on the factors already defined, they had to get the information needed to make the mandatory analysis. After a few research they came up the the dataframe below. The dataframe is consisted in the 20 arrondissements of Paris followed by their location (with latitude and longitude coordinates), population and crime rate.

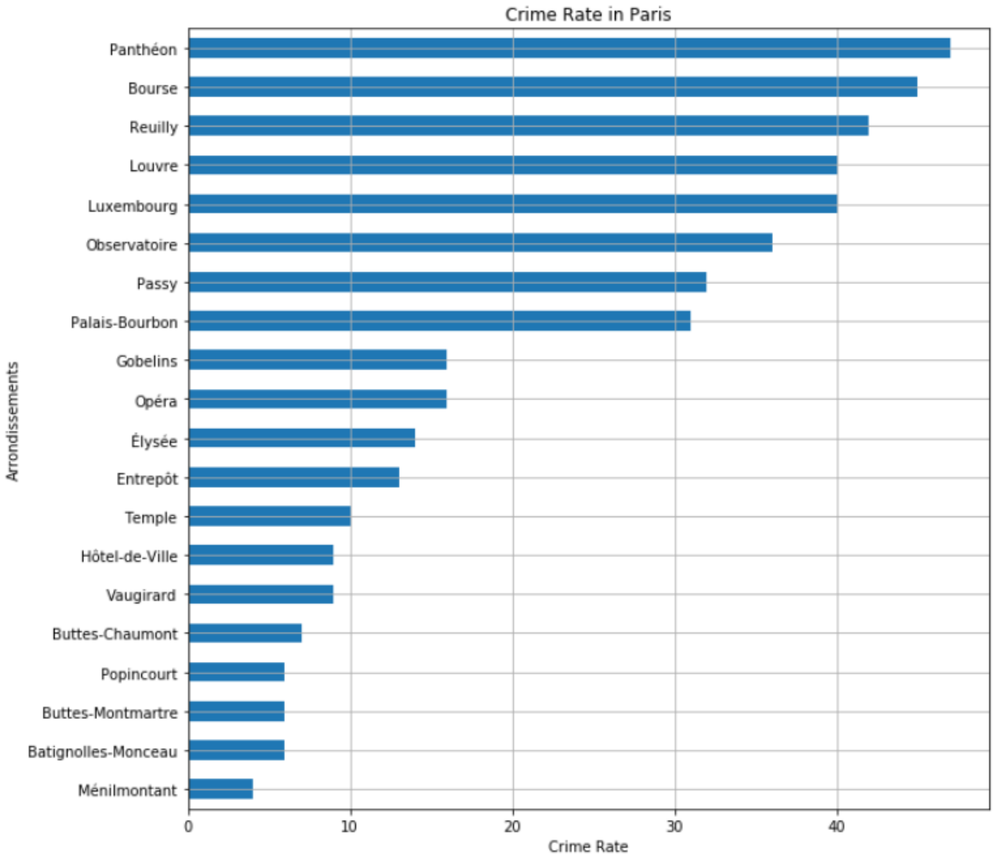
(As a crime rate was not found untill this fate and for the sake of the sucess of this analysis, it was decided to use random crime rates)

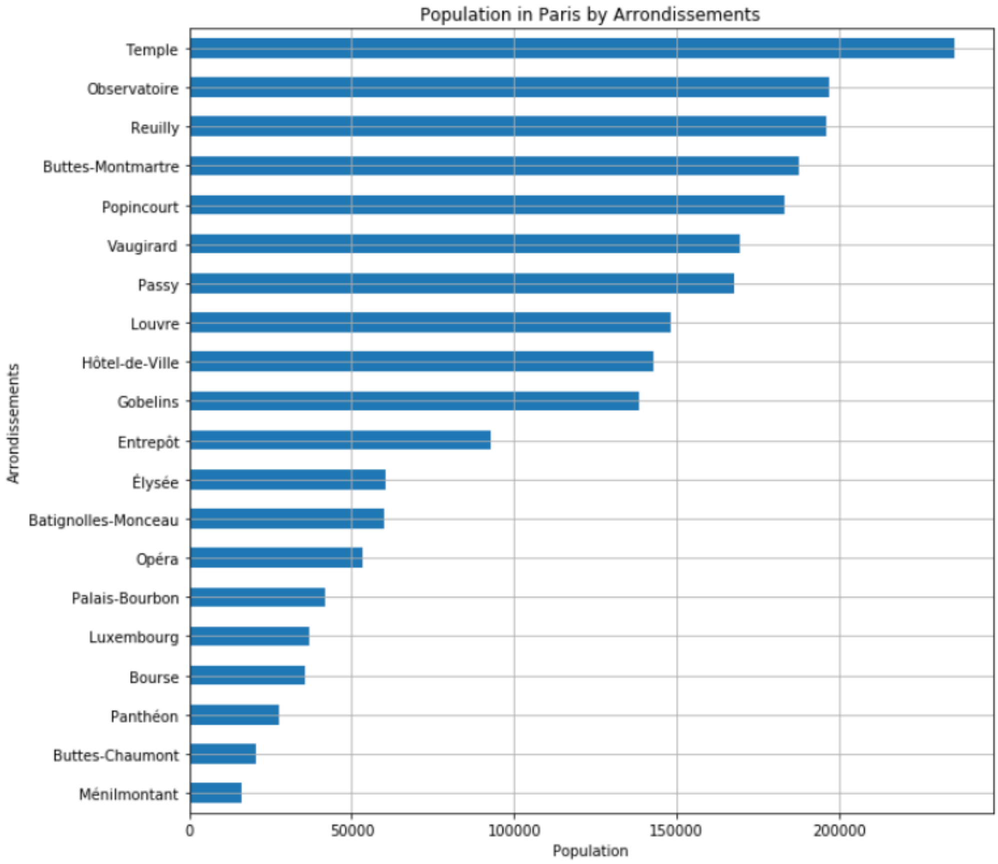
	L_AROFF	Latitude	Longitude	Population	CrimeRate
0	Temple	48.86287238	2.3600009859	235469	10
1	Bourse	48.8682792225	2.34280254689	35469	45
2	Batignolles-Monceau	48.887326522	2.30677699057	59947	6
3	Palais-Bourbon	48.8561744288	2.31218769148	41831	31
4	Hôtel-de-Ville	48.8543414263	2.35762962032	142661	9
5	Élysée	48.8727208374	2.3125540224	60235	14
6	Buttes-Montmartre	48.892569268	2.34816051956	187760	6
7	Louvre	48.8625627018	2.33644336205	148339	40
8	Popincourt	48.8590592213	2.3800583082	183117	6
9	Gobelins	48.8283880317	2.36227244042	138218	16
10	Observatoire	48.8292445005	2.3265420442	196884	36
11	Ménilmontant	48.8634605789	2.40118812928	16338	4
12	Opéra	48.8771635173	2.33745754348	53415	16
13	Buttes-Chaumont	48.8870759966	2.38482096015	20410	7
14	Vaugirard	48.8400853759	2.29282582242	169375	9
15	Entrepôt	48.8761300365	2.36072848785	92660	13
16	Passy	48.8603921054	2.26197078836	167706	32
17	Panthéon	48.8444431505	2.35071460958	27795	47
18	Luxembourg	48.8491303586	2.33289799905	37053	40
19	Reuilly	48.8349743815	2.42132490078	196143	42

Data Analysis

EDA

After cleaning the data, some analysis is mandatory to come up with a decision. They started with some EDA (Explanatory Data Analysis) and realized the most safe places to set up a business as well as the most populated places where there was more potencial for a higher number of clients.




Population in Paris by Arrondissements
Temple
Observatoire
Reuilly
Buttes-Montmartre
Popincourt
Vaugirard
Passy
Louvre
Hôtel-de-Ville
Gobelins
Entrepôt
Élysée
Batignolles-Monceau
Opéra
Palais-Bourbon
Luxembourg
Bourse
Panthéon
Buttes-Chaumont
Ménilmontant
Arrondissements
0
50000
100000
150000
200000
Population

Cluster Analysis

After sorting the data in the EDA, where they got information in safety and client potencial, the last factor analyzed is missing: venues competition. Their objective was to find anarrondissement where there was not a lot of restaurants, so that they could grow and build a community around their restaurant. In order to do it, it was utilized a data from Foursquare to form a dataframe to classify each arrondissement in terms of most common venues. That way we could cluster the arrondissements considering their similarities and therefore choose the best cluster to set up a restaurant.

	Neighborhood	Latitude	Longitude	Population	CrimeRate	Cluster Labels
2	Batignolles-Monceau	48.887326522	2.30677699057	59947	6	0
3	Palais-Bourbon	48.8561744288	2.31218769148	41831	31	0
5	Élysée	48.8727208374	2.3125540224	60235	14	0
10	Observatoire	48.8292445005	2.3265420442	196884	36	0
19	Reuilly	48.8349743815	2.42132490078	196143	42	1
17	Panthéon	48.8444431505	2.35071460958	27795	47	2
15	Entrepôt	48.8761300365	2.36072848785	92660	13	2
14	Vaugirard	48.8400853759	2.29282582242	169375	9	2
13	Buttes-Chaumont	48.8870759966	2.38482096015	20410	7	2
12	Opéra	48.8771635173	2.33745754348	53415	16	2
11	Ménilmontant	48.8634605789	2.40118812928	16338	4	2
0	Temple	48.86287238	2.3600009859	235469	10	2
8	Popincourt	48.8590592213	2.3800583082	183117	6	2
7	Louvre	48.8625627018	2.33644336205	148339	40	2
6	Buttes-Montmartre	48.892569268	2.34816051956	187760	6	2
4	Hôtel-de-Ville	48.8543414263	2.35762962032	142661	9	2
1	Bourse	48.8682792225	2.34280254689	35469	45	2
18	Luxembourg	48.8491303586	2.33289799905	37053	40	2
9	Gobelins	48.8283880317	2.36227244042	138218	16	3
16	Passy	48.8603921054	2.26197078836	167706	32	4

Score Analysis and Conclusions

After the data analysis, for a more quantitive based decision, they ended up creating a score system that you can see in the final database. We can conclude that the restaurant should be in Passy.

	Neighborhood	Latitude	Longitude	Population	CrimeRate	Score
16	Passy	48.8603921054	2.26197078836	167706	32	3.78
2	Batignolles-Monceau	48.887326522	2.30677699057	59947	6	3.66
5	Élysée	48.8727208374	2.3125540224	60235	14	2.97
19	Reuilly	48.8349743815	2.42132490078	196143	42	2.79
3	Palais-Bourbon	48.8561744288	2.31218769148	41831	31	2.76
11	Ménilmontant	48.8634605789	2.40118812928	16338	4	2.55
10	Observatoire	48.8292445005	2.3265420442	196884	36	2.49
9	Gobelins	48.8283880317	2.36227244042	138218	16	2.37
8	Popincourt	48.8590592213	2.3800583082	183117	6	2.28
13	Buttes-Chaumont	48.8870759966	2.38482096015	20410	7	2.25
6	Buttes-Montmartre	48.892569268	2.34816051956	187760	6	2.19
4	Hôtel-de-Ville	48.8543414263	2.35762962032	142661	9	2.07
14	Vaugirard	48.8400853759	2.29282582242	169375	9	1.98
0	Temple	48.86287238	2.3600009859	235469	10	1.89
12	Opéra	48.8771635173	2.33745754348	53415	16	1.86
15	Entrepôt	48.8761300365	2.36072848785	92660	13	1.77
18	Luxembourg	48.8491303586	2.33289799905	37053	40	1.56
7	Louvre	48.8625627018	2.33644336205	148339	40	1.38
1	Bourse	48.8682792225	2.34280254689	35469	45	1.35
17	Panthéon	48.8444431505	2.35071460958	27795	47	1.35