

Recommendations on Arrondissement / Neighbourhoods to open a new restaurant in Paris

Capstone project

Naveen B
December 19, 2020

1. Introduction

As a part of this project, a problem statement needs to be identified that can take leverage of the Foursquare location data to come up with insights on solving the problem. Based on the problem statement identified, analysis need to be performed on the data to come up with insights on the best approach towards solving the problem.

2. Business Problem

A contractor is looking towards opening a restaurant in the city of Paris. The contractor has not yet decided on the neighbourhood where he is going to open the restaurant. He has approached a data scientist to provide him some guidance based on the current trends in the available data and enable him to decide on the neighbourhood for his restaurant with valid justification. Additionally, the contractor is also interested in knowing the cuisine that would attract the most customers in that area. As a data scientist, I would be using the Foursquare location data and other available datasets to analyse the current trends using graphs and other data visualization methods for narrowing down on the neighbourhood best suited for the restaurant.

3. Data Section

As a part of this analysis, FourSquare location data would be used to identify the restaurants and it's types across various neighbourhoods in Paris. This data would be plotted on a map using Folium and density of restaurants in each neighbourhood would be identified.

An additional dataset would be used called food_coded.csv. This dataset is an open dataset found in Kaggle. For this project purpose, this dataset would be considered as the data collected from the various people in Paris. This dataset would contain details on their occupations, their food preferences and their willingness to eat at a restaurant. In cases where they are eating outside, their comfort foods which they would look out for are also recorded in this dataset. Based on this information, it would be possible to obtain an insight on the need of the customers and neighbourhood which is in most need of a restaurant.

In the following paragraph a brief insight into this Food_coded dataset is provided:

Snippet of food_coded csv:

Gender	breakfast	calories_chicken	calories_day	calories_scone	coffee	comfort_food	comfort_food_reasons	comfort_food_reasons_coded	...	soup	sports
2	1	430	NaN	315.0	1	none	we dont have comfort	9.0	...	1.0	1.0
1	1	610	3.0	420.0	2	chocolate, chips, ice cream	Stress, bored, anger	1.0	...	1.0	1.0
1	1	720	4.0	420.0	2	frozen yogurt, pizza, fast food	stress, sadness	1.0	...	1.0	2.0
1	1	430	3.0	420.0	2	Pizza, Mac and cheese, ice cream	Boredom	2.0	...	1.0	2.0
1	1	720	2.0	420.0	2	Ice cream, chocolate, chips	Stress, boredom, cravings	1.0	...	1.0	1.0

Additional information on the dataset

Please note that this dataset would be processed to remove the columns that are not required for his analysis.

Some interesting information in the food_coded dataset that we might be using on this analysis are the following:

- Gender (1 – Female and 2 – Male)
- cook – how often do you cook?
- eating_out - frequency of eating out in a typical week
- fav_cuisine - What is your favourite cuisine?
- fav_food - was your favourite food cooked at home or store bought?
- marital_status
- on_off_campus – living situation
- parents_cook - Approximately how many days a week did your parents cook?
- pay_meal_out - How much would you pay for meal out?

Based on above information, we would be able to identify the target audience for our restaurant. This would mean, we would be able to decide would be the customers that have a higher chance of choosing this restaurant. Hence, we could make the restaurant more appealing to these customers during the initial stage.

The above information would allow us to choose a speciality cuisine as well for the restaurant, provide us the pricing expectations.

Once we have the above information sorted out, we would be able to target the neighbourhoods having higher percentage of these target customer using the venues in that we can retrieve using the FourSquare API.

This would be my approach towards this problem.

Population Density

Another dataset that we would be using is population density dataset. Paris is divided into 20 Arrondissements. The population density of each arrondissements is provided in this dataset. A peak in to the dataset is shown below:

A snippet of the population density dataset:

	Arrondissement	Area (km2)	Population	Population per km2
0	1	1.826	17,268	9,457
1	2	0.992	22,558	22,740
2	3	1.171	36,727	31,364
3	4	1.601	28,068	17,532
4	5	2.541	61,080	24,038

4. Methodology Section

Based on above information described, we would be able to identify the target audience for our restaurant. This would mean, we would be able to decide would be the customers that have a higher chance of choosing this restaurant. Hence, we could make the restaurant more appealing to these customers during the initial stage.

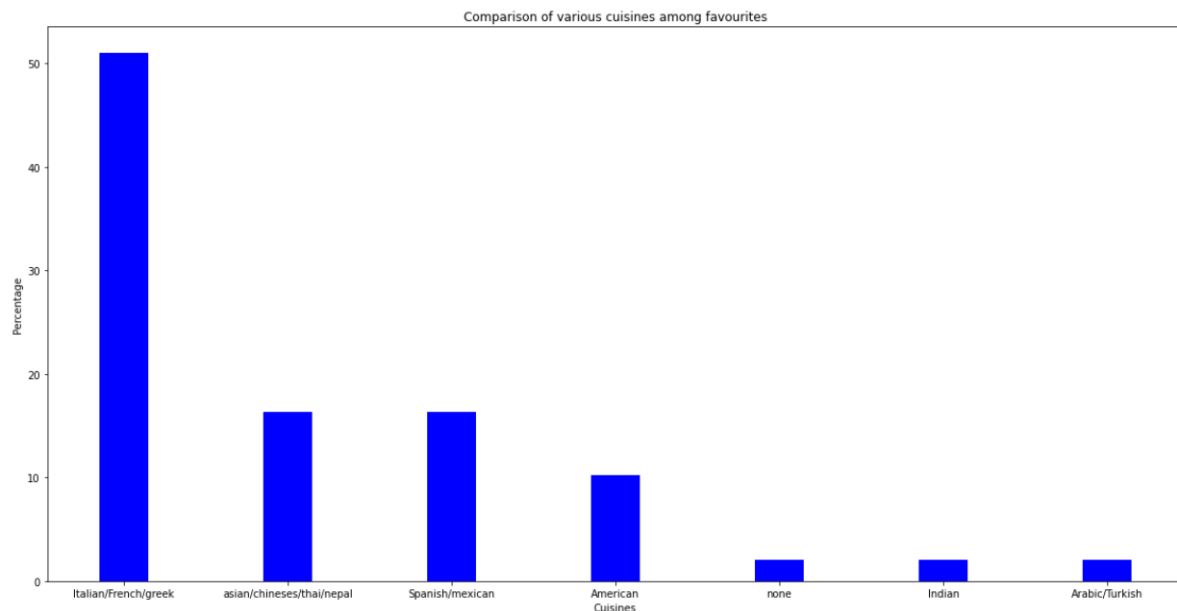
The above information would allow us to choose a speciality cuisine as well for the restaurant, provide us the pricing expectations.

Once we have the above information sorted out, we would be able to target the neighbourhoods having higher percentage of these target customer using the venues in that we can retrieve using the FourSquare API.

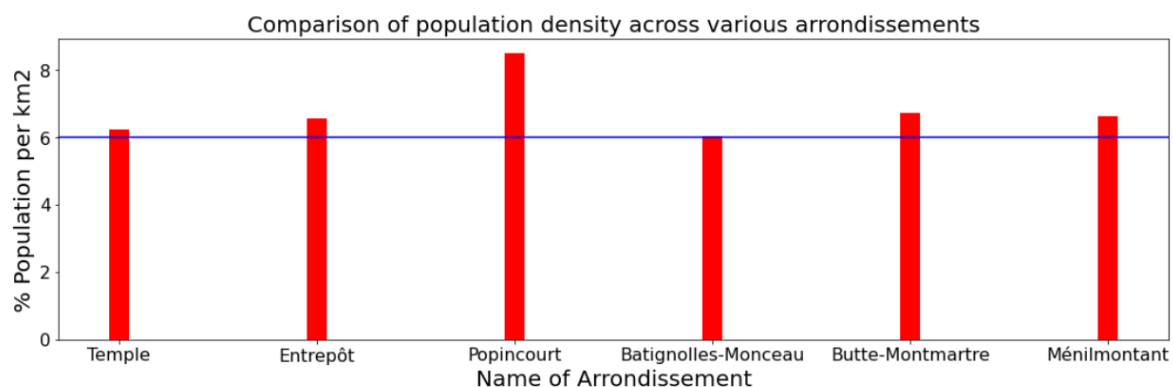
This would be my approach towards this problem.

The data would be pre processed to filter out the unnecessary columns and have only the columns that we require for the analysis. The data was normalised across into percentages to obtain an idea on the factors that significantly affect the recommendations.

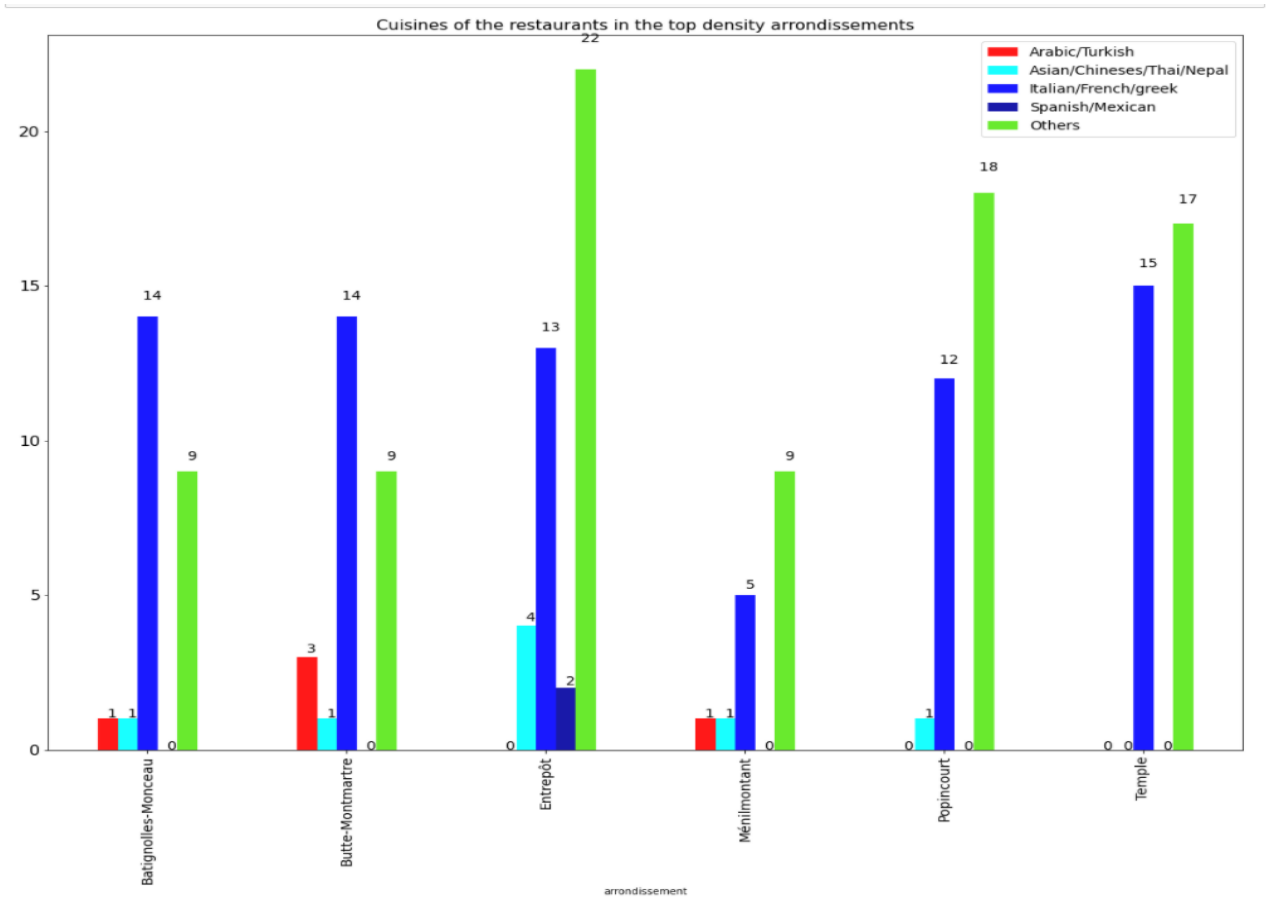
The food_coded dataset was analysed to get an idea on the preferences of the people living in Paris. From the dataset the subset of people who would frequently visit restaurants were filtered out. Once this was done, we then started focussing on the cuisines which were most popular among this set of people. To do this , a new dataframe was created with the subset of people who would frequently dine outside and their favourite cuisines was analysed.



Each cuisine was plotted against the percentages of people who had chosen this cuisine as the favourite. This helped us arrive at an conclusion as which cuisine could be chosen for our nre restaurant.

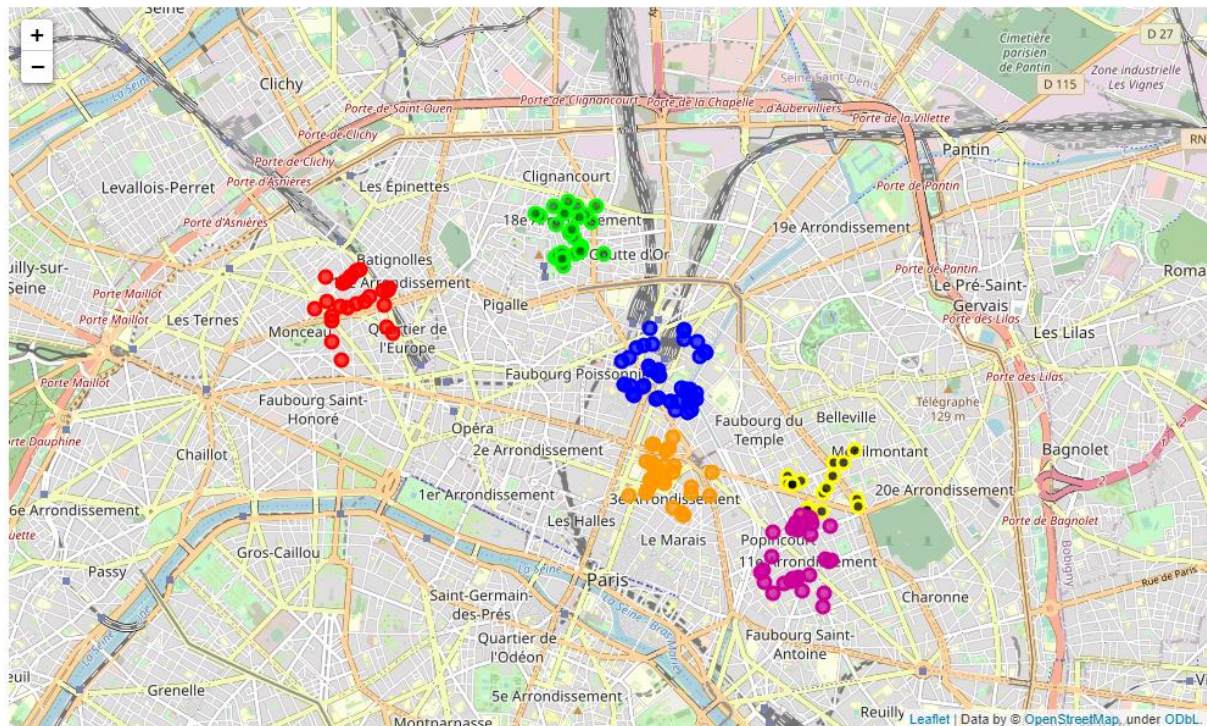


The population density dataset was taken to find the most populated arrondissements / neighbourhoods in Paris. Since the population density plays a huge role in the number of customers visiting the restaurant, the neighbourhoods with more than 6% of the total population density of the city of Paris were chosen. The blue line indicates 6% density of population.



The next step in our approach was to look at the restaurants in those chosen neighbourhoods. Using the FourSquare API, we were able to retrieve the restaurants in and around those highly populated arrondissements of Paris. Based on that analysis, we were able to narrow down on one arrondissement which contained a smaller number of Italian/Greek/French cuisine restaurants in that neighbourhood.

Folium was used to plot the clusters of the restaurants across the highly populated arrondissements as shown below:



The above map indicates the proximity of the restaurants with each other in each neighbourhood:

- Batignolles-Monceau - Red Color spots
- Butte-Montmartre - Lime Color spots
- Entrepôt - blue color spots
- Ménilmontant - yellow with black fill spots
- Popincourt - purple color spots
- Temple - orange color spots

5. Results Section:

From the bar plot, it can be seen that Italian/French/Greek cuisines had a significant higher value compared to the other cuisines like Asian, Chinese, Spanish, Mexican, Arabic etc. In order to find population density across various arrondissements or neighbourhoods in Paris. This can be observed in the bar plot 'Comparison of population density across various arrondissements'. From this bar plot, the top arrondissements with high population density can be inferred. The plot 'Cuisines of the restaurants in the top density arrondissements' provides information on the cuisine of the restaurants present in the high-density arrondissements or neighbourhoods of

Paris and we can see that the Ménilmontant has the least number of Italian/French/Greek cuisine restaurants. A map was also plotted which shows the proximity of the restaurants between each other in each neighbourhood.

6. Discussion Section

Based on the plots discussed in the results section, the following observations were made:

1) Observations from the graph 'Comparison of various cuisines among favourites' This plot provides insight on the preferences among the people who frequently dine out. Based on the plot it can be clearly seen that the Italian/Greek/French cuisine is the cuisine that this section of people who dine out frequently prefer.

2) Observations from the graph 'Comparison of population density across various arrondissements': This plot provides an insight across the various high densely populated arrondissements in the city of Paris. From the plot, it can be understood that the most people live per square kilometre in the following neighbourhoods:

- Batignolles-Monceau
- Butte-Montmartre
- Entrepôt
- Ménilmontant
- Popincourt
- Temple

3) Observations from the plot "Cuisines of the restaurants in the top density arrondissements" This plot shows the numbers of restaurants on the various cuisines in neighbourhood selected from the second graph. This shows that Ménilmontant has the least amount of Italian/Greek/French cuisine restaurant. Hence it is most probable that the people might need a new Italian/French/Greek cuisine restaurant in this area.

7. Conclusion Section

From the above shown analysis and insights obtained from data visualization, it can be understood that there is a section of people in Paris who eat frequently in restaurants. A major portion of this section of people tend to like Italian/Greek/French cuisines. We can see that some of the neighbourhoods are densely populated and hence has a chance of better customer incoming. Out of these neighbourhoods/arrondissements we could see that Ménilmontant has

very few Italian/Greek/French cuisine restaurants. Based on these points, we conclude that we would recommend opening an Italian, Greek or a French restaurant in the arrondissement / neighbourhood of Ménilmontant, Paris for the best outcome.