Market Research for Opening New italian Restaurant in Paris, France

HOW TO FIND THE BEST PLACE USING DATA-DRIVEN APPROACH?
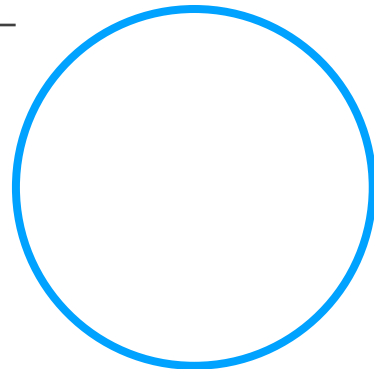
R. LECLERCQ, 2021

TABLE OF CONTENT

1. Introduction

2. Data description

3. Methodology

4. Results

5. Conclusion
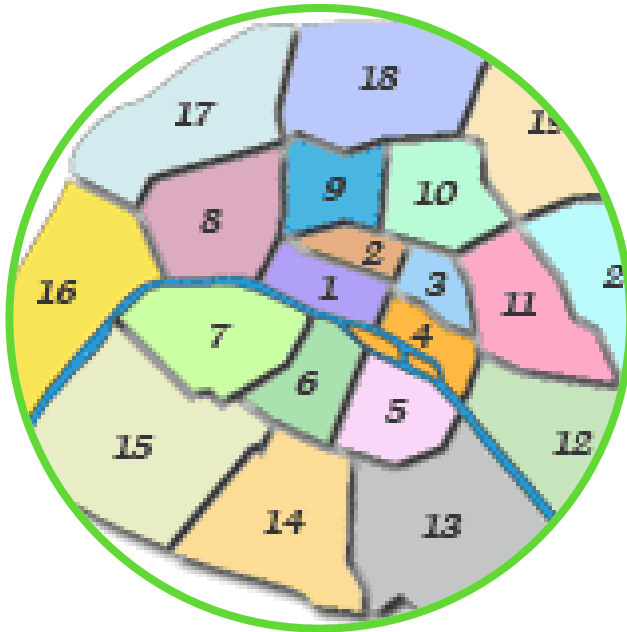
6. References

INTRODUCTION

- Paris is the French capital

- 2 millions of residents in 105 km^2

- > 160 000 italian people in France

- Business Question :

HOW TO FIND THE BEST PLACE TO OPEN A NEW ITALIAN RESTAURANT IN PARIS ?

DATA DESCRIPTION

NEIGHBORHOOD DATA OF 20
DISTRICTS IN PARIS, FRANCE

> Paris Arrondissements
[10 – 11]

COORDINATES (LATITUDE,
LONGITUDE) OF EACH DISTRICT

> Geopy

FOURSQUARE API

> Most famous venues and
location (< 500m radius)

METHODOLOGY

1. Building Paris neighborhood data (Postcode, Neighborhood)

2. Building the coordinates of all districts in Paris, France using tool geopy.

3. Using Foursquare API to generate to the most famous venues and the relative locations.

4. Analyzing, exploring the data and finding the optimal number of clusters for K-Means clustering using the **Elbow**

5. Using K-Means to cluster the neighborhoods in the city of Paris based on the top 10 venues for each neighborhood district.

6. Visualizing geographic details of Paris and its district neighborhoods which are superimposed on top using tool folium

7. Analyzing the clustering result and then proposing some suggestions about the good districts to open Vietnamese Restaurant in Paris

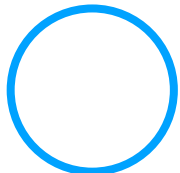8. Giving some perspectives to enhance the performance

RESULTS

	postcode	neighbourhood
0	75001	75002,75003,75004,75005,75006,75007,75008,75009
1	75002	75001,75003,75009,75010
2	75003	75001,75002,75004,75010,75011
3	75004	75001,75003,75005,75006,75011,75012
4	75005	75001,75004,75006,75012,75013,75014
5	75006	75001,75004,75005,75007,75014,75015
6	75007	75001,75006,75008,75015,75016
7	75008	75001,75007,75009,75016,75017,75018
8	75009	75001,75002,75008,75010,75017,75018
9	75010	75002,75003,75009,75011,75018,75019,75020
10	75011	75003,75004,75010,75012,75019,75020
11	75012	75004,75005,75011,75013,75020
12	75013	75005,75012,75014
13	75014	75005,75006,75013,75015
14	75015	75006,75007,75014,75016
15	75016	75007,75008,75015,75017
16	75017	75008,75009,75016,75018
17	75018	75008,75009,75010,75017,75019
18	75019	75010,75011,75018,75020
19	75020	75010,75011,75012,75019

Table 1. Dataframe of Neighborhood Data in Paris, France

RESULTS

	postcode	latitude	longitude
0	75001	48.863554	2.338856
1	75002	48.867418	2.344256
2	75003	48.862607	2.360211
3	75004	48.856004	2.357028
4	75005	48.852752	2.346343
5	75006	48.853537	2.343370
6	75007	48.855913	2.313839
7	75008	48.872385	2.312707
8	75009	48.877355	2.336856
9	75010	48.879201	2.354391
10	75011	48.855630	2.370806
11	75012	48.839734	2.380054
12	75013	48.826997	2.353396
13	75014	48.828590	2.307541
14	75015	48.838461	2.315728
15	75016	48.855031	2.273958
16	75017	48.883508	2.304923
17	75018	48.893074	2.343881
18	75019	48.878076	2.376198
19	75020	48.857126	2.409257

Table 2. The coordinates
of 20-district in Paris, France
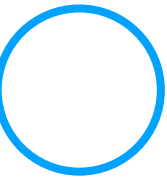
RESULTS
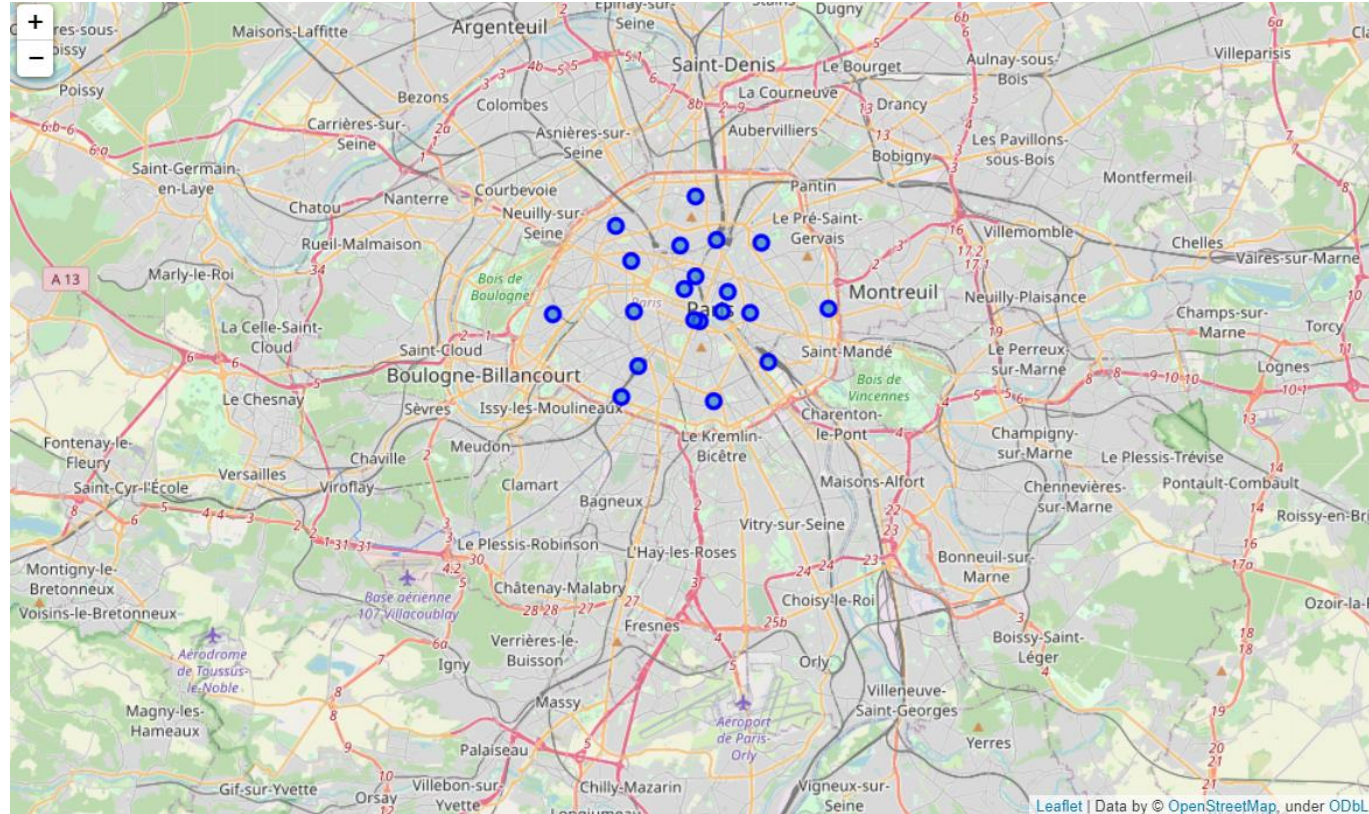
Visualization of geographic details of Paris and its district neighborhood which are superimposed on top

DETERMINING THE OPTIMAL VALUE K

- The charts of the Elbow Method using the distortion or the inertia

- Difficult to determine the Elbow point using the inertia.

- Applying the distortion with the number of clusters, we choose the optimal number of clusters for K-Means is 8.

The Elbow Method using Distortion

The Elbow Method using Inertia
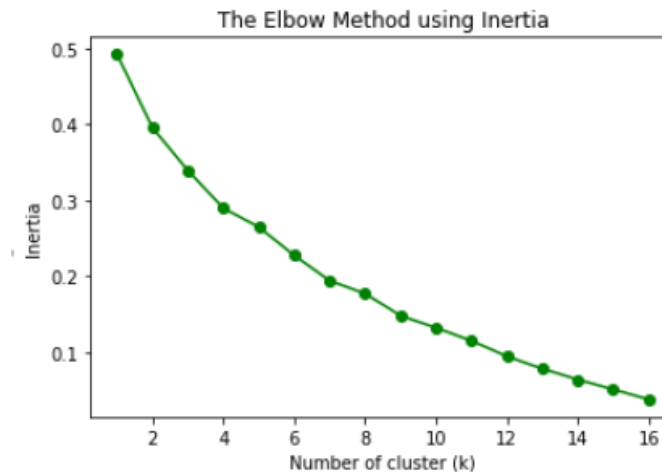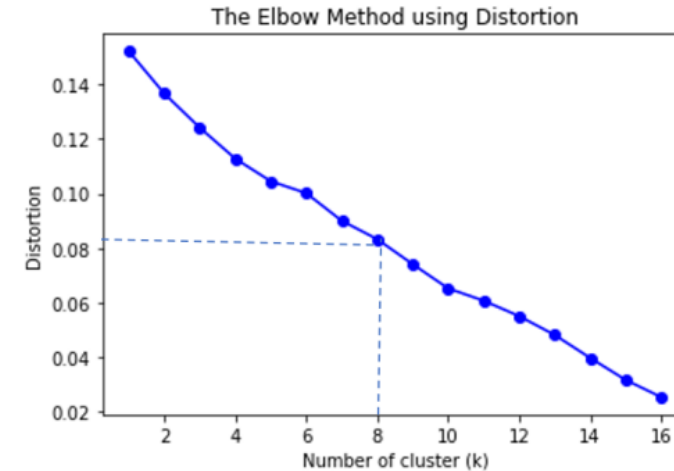
CLUSTERING NEIGHBORHOODS OF PARIS
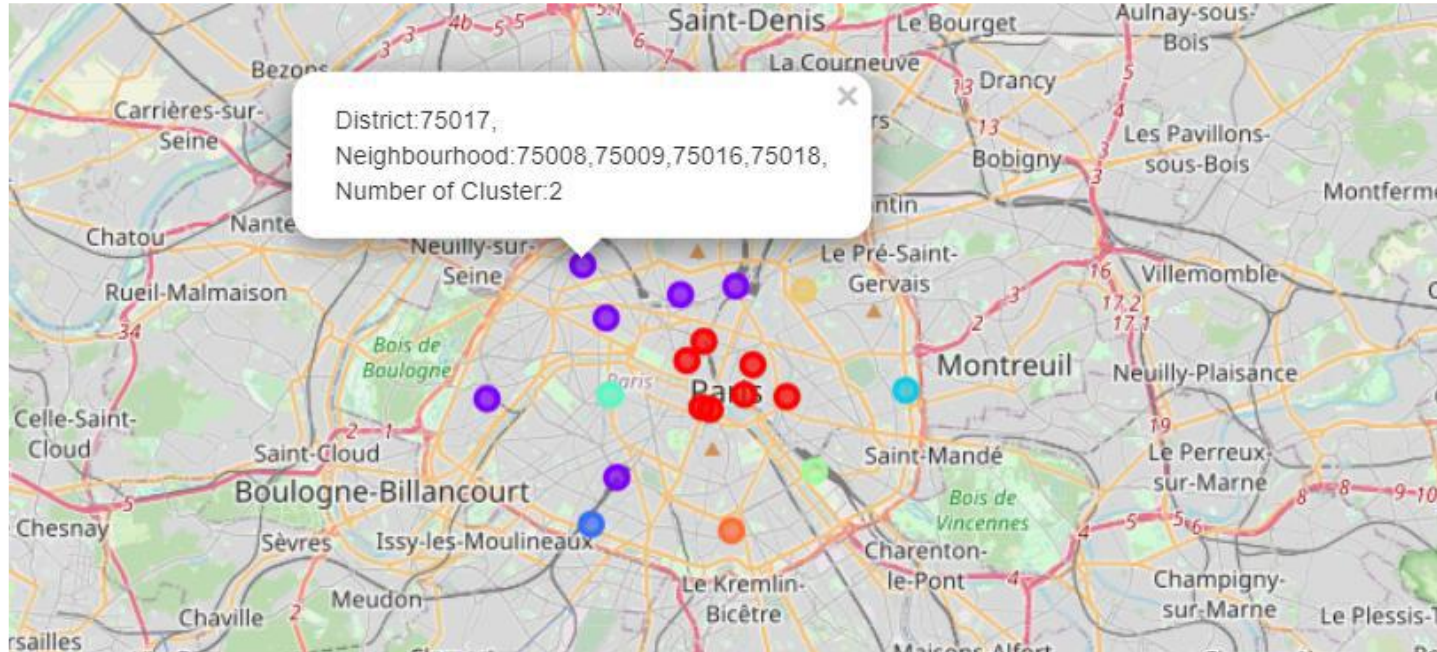
Using the optimal value k, we launch K-Means technique to cluster the neighborhood into 8 clusters

CONCLUSION

- Italian restaurant is present in several cluster: 1,2, 5 and 8.

- Rated second in cluster 2

- Best choice is **District 15.**

Perspectives : new variables

-the transport info (public transport, parking, etc.),

-the information of italian communities,

-the information of major tourist venues

-the number of families as italian restaurant are a typical menu for french family