



The Battle of Neighborhoods

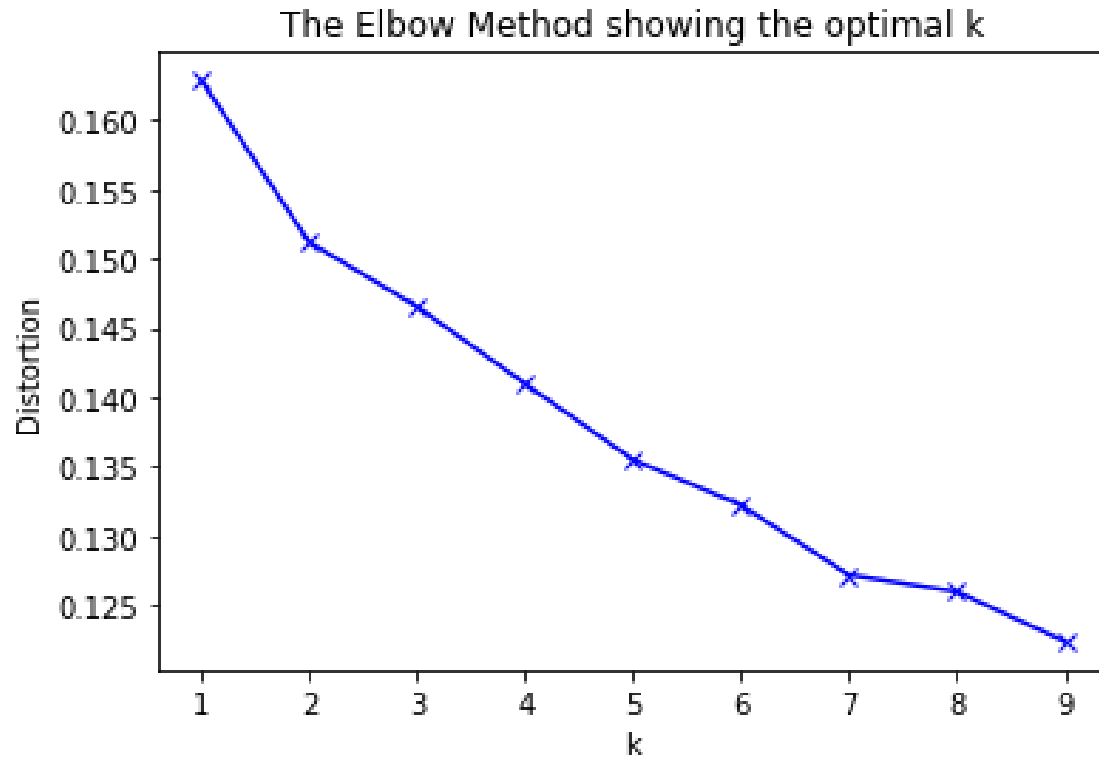
Clustering locations in Paris based on Restaurant Types



Data acquisition and cleaning

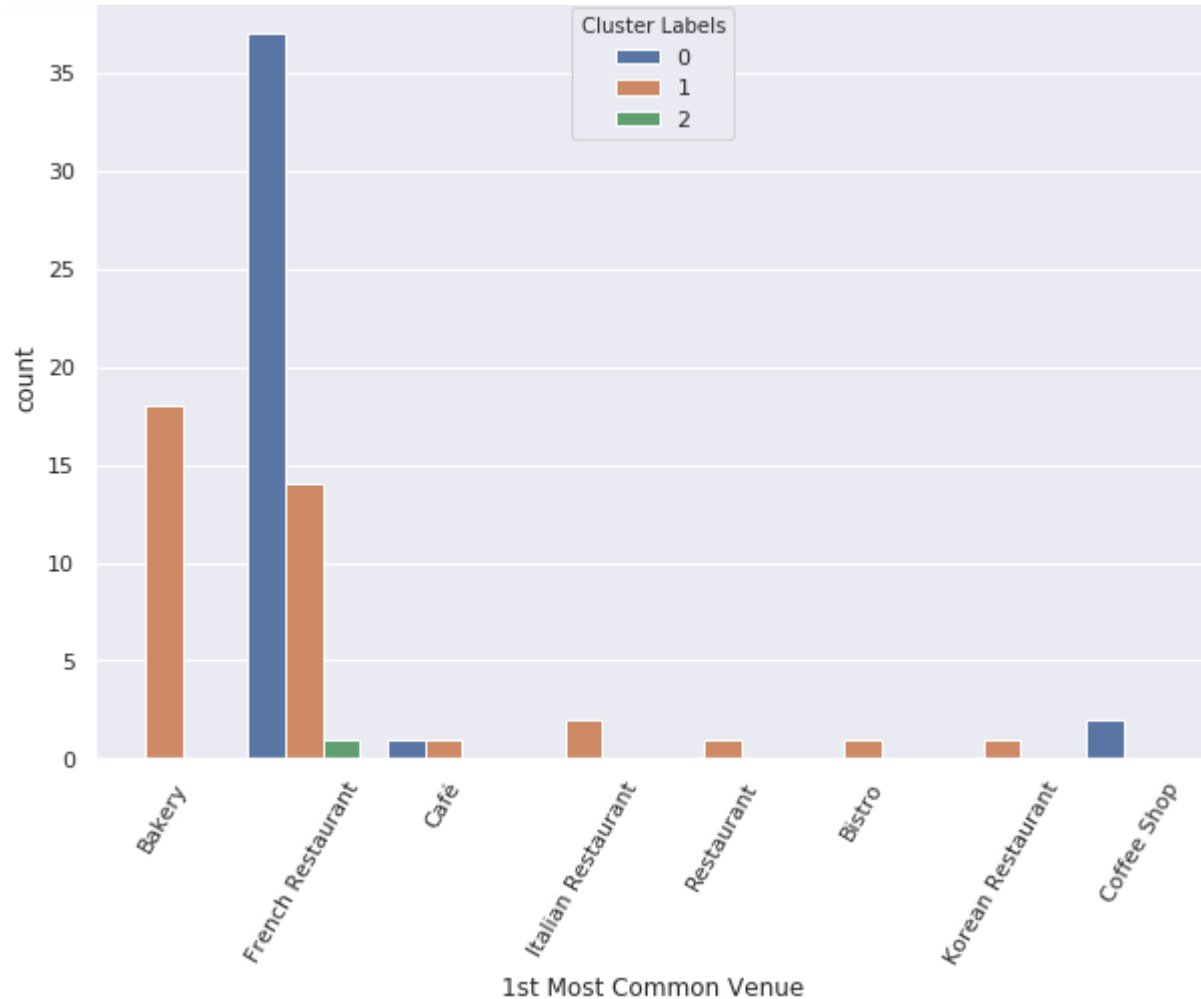
- Paris Neighborhood dataset got from <https://www.data.gouv.fr/fr/datasets/quartiers-administratifs/>
- Foursquare API is used to explore Location Data for each Neighborhood.
- In total, 3,741 rows and 7 columns in the final dataset.
- Cleaned data contains 153 features.

Neighborhoods Restaurants Clustering(1)



- KMeans algorithm used to cluster our dataframe.
- The elbow method leveraged to find the optimal value for number of clusters k.
- The plot looks like an arm with an elbow at $k = 3$, so that point is the optimal value for k.

Neighborhoods Restaurants Clustering(2)



- **Cluster 1 - French Restaurants:**
This cluster is dominated by French Restaurants.
- **Cluster 2- Bakery & Divers Restaurants:**
In this cluster, people prefer Bakery and others food styles(Italian, Korean,..) together with the French food.
- **Cluster 3- Bel-Air:** It is an isolated cluster, it contain only Bel-Air neighborhood.

Conclusion

- The 80 Paris neighborhoods are segmented into 3 clusters according to their similar food categories.
- Cluster 2 neighborhoods are more suitable for opening restaurants with Maghreb foods in Paris.
- Others features can be included to the study for further improvement:
 - Location renting cost,
 - Traffic pattern information,
 - Demographic and lifestyle data.