Capstone Project
Battle of neighborhoods

FIND THE RIGHT AREA TO OPEN A NEW FRENCH RESTAURANT

IN **MONTREAL**, CANADA

Find the right area to open a new French Restaurant in Montreal

- Start a new business in the best conditions thanks to Data Science

- Use of Open Data and Foursquare API

- **How to limit competition ?**

Data acquisition and cleaning

Montréal | Données ouvertes

DESC_LIEU	Longitude	Latitude
POSTE DE QUARTIER 4	-73.807304	45.494380
POSTE DE QUARTIER 30	-73.617433	45.577744
POSTE DE QUARTIER 23	-73.545915	45.556866
POSTE DE QUARTIER 10	-73.701029	45.528543
POSTE DE QUARTIER 24	-73.615612	45.520110
POSTE DE QUARTIER 39	-73.615625	45.607724

■Open Data from https://donnees.montreal.ca
➔Coordinates of the Police posts in Montreal
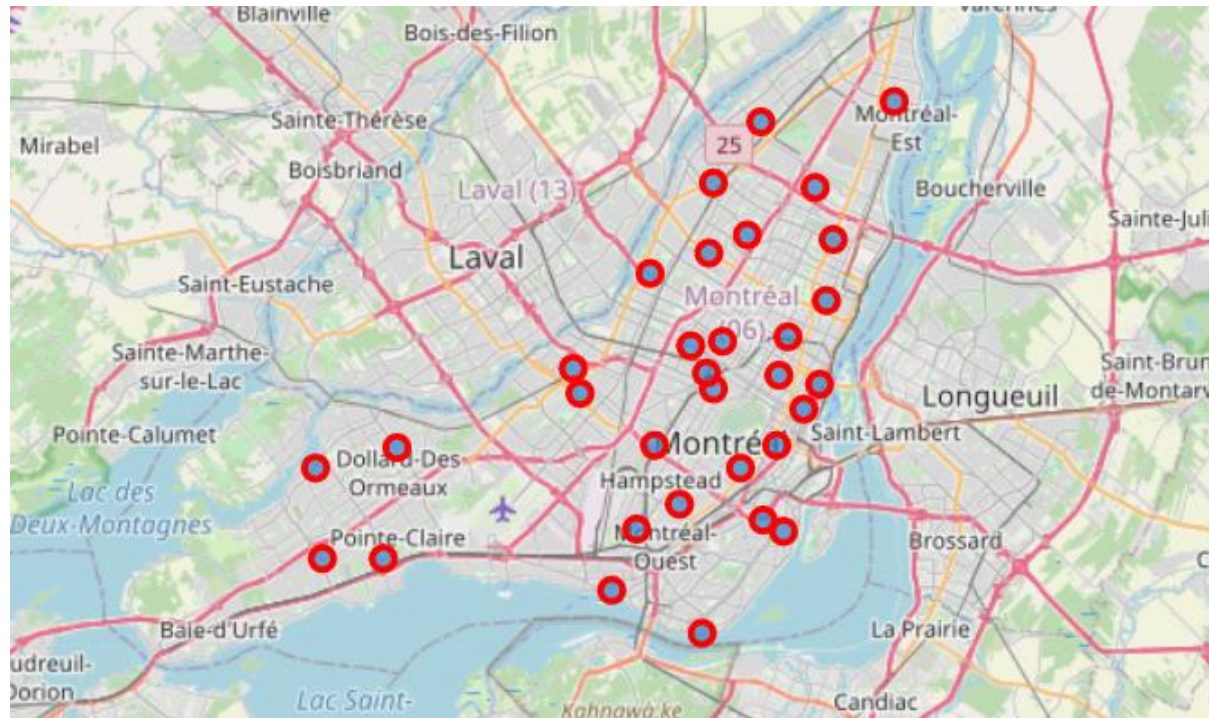➔Allow us to identify the main geographical areas of this city

■ Foursquare API
➔obtain the venues around these areas
➔focus on the existing French restaurants

FOURSQUARE

Data Visualization from analyzed areas

- A representative view of the split of Montreal in term of human activities:

Classification model

- Use of **k-means** to clusterize these areas in Montreal per presence of French restaurants

- Divide the geographical areas into non-overlapping subsets (cluster), without any **cluster-internal structure**

- Within each cluster, **a strong similarity will be present:** we need to find a cluster with a little number of French Restaurants

Clusters found

3 different clusters have been identified thanks to **k-means clustering ➔**
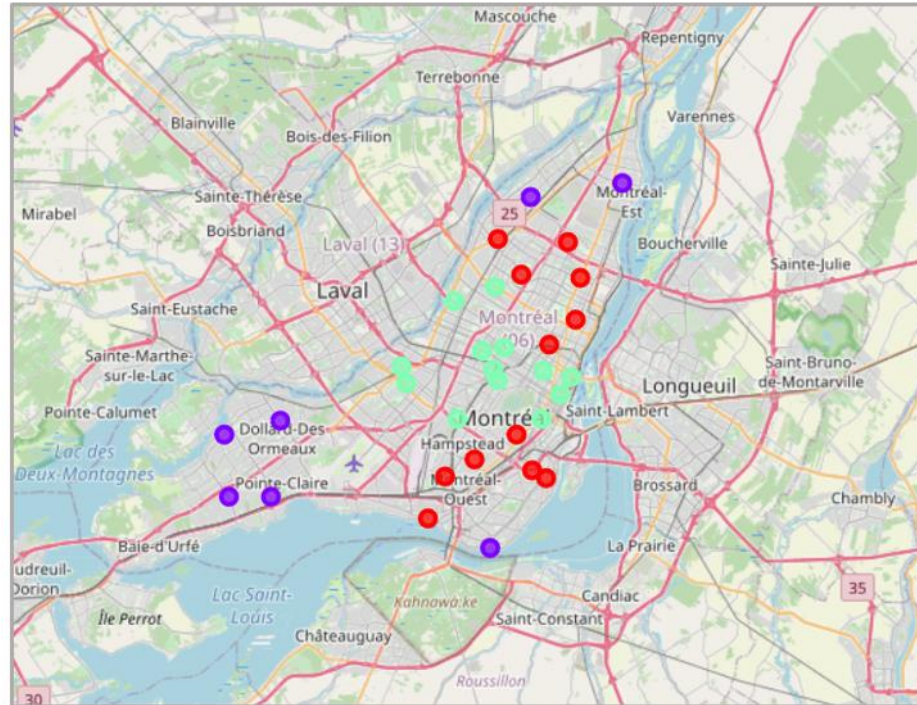
High concentration of French Restaurants

Medium concentration of French Restaurants

Low concentration of French Restaurants

Best areas to open a new French Restaurant

Conclusion and future directions

- **The areas in the 3rd cluster in purple, with a low concentration of French restaurants is clearly the more appropriate** to launch such a business
  - 6 areas are concerned
  - **The competitors will be fewer than in the other areas**

- This project could be completed with other Data:
  - **Demographical Data** (habitant per areas)
  - **Economic Data** (companies actives in these areas)
  - **Sociological Data** (habits for lunch, dinner…)