

Choosing place for café in Paris

Coursera Capstone Project

Sviatoslav Shapkin

2020

Problem

Location is one of the most important decisions that will determine whether the cafe will be a success or a failure

Objective: Analyse and select the best locations in Paris to open a new cafe

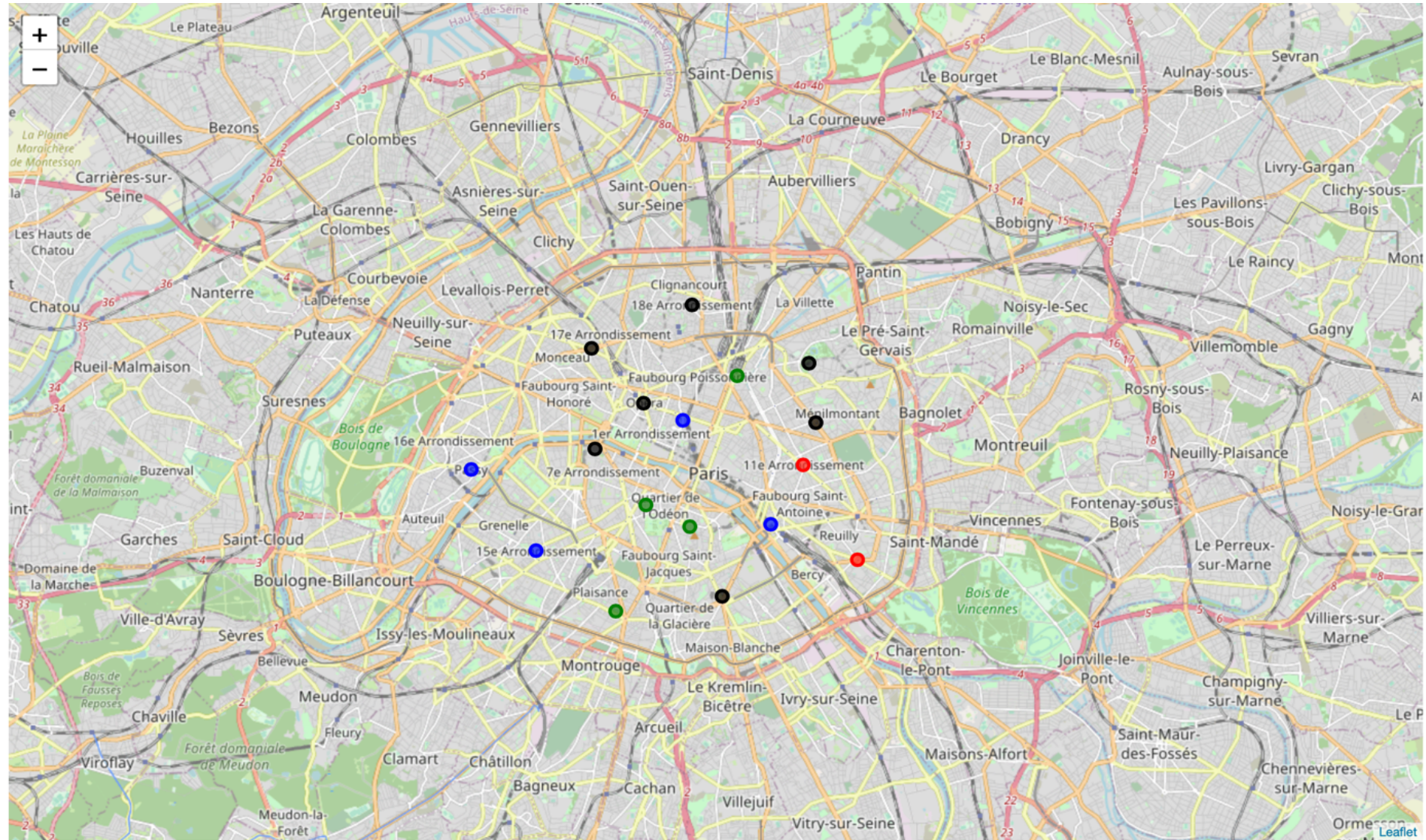
Data

- Table with Paris districts from Wikipedia (https://en.wikipedia.org/wiki/Arrondissements_of_Paris)
- Geo Data to identify the centers of districts for further applications
- Foursquare API to get info about different organizations in the districts.

Methodology

- Web scraping Wikipedia page for district list.
- Get latitude and longitude coordinates using Geocoder.
- Use Foursquare API to get venue data.
- Filter venue category by Café field.
- Perform clustering by using k-means clustering.
- Visualize the clusters using Folium framework.

Results



Discussion

- Only one metric has been used for KMeans algorithm
- Only central districts have been used (excluding suburbs like Saint-Denis)

Conclusion

We have gone through the process of:

- identifying the problem,
- specifying and preprocessing the data,
- performing machine learning by clustering the data,
- providing recommendations to the relevant stakeholders.