

CAPSTONE PROJECT

The Battle of the Neighborhoods

Introduction: Business Problem

Find a house or an apartment in Paris with the best possible value for money

- At least one subway station.
- Dynamic place with a lot of activities and transactions around it
- The lowest possible price

Data

- Opendata Paris: the shape of Paris and its boroughs
- Data RATP: the positions of all existing subway stations
- Cadastre Data: the transactions in Paris since 2014
- Foursquare: the venues close to potential locations
- Nominatim: the addresses corresponding to our final potential locations

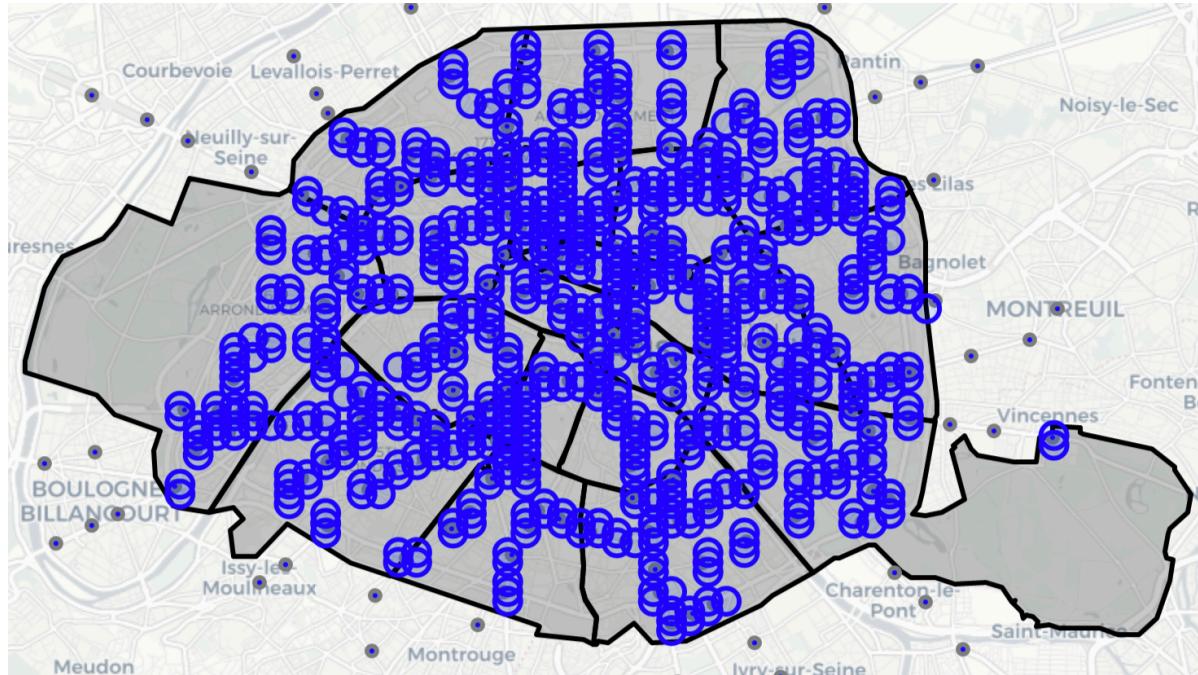
Methodology: Initial Potential Locations

Keep only the 2487 circles of 200 meters whose center is in Paris.



Methodology: Close to a subway station

Keep only the 681 potential locations that are close to a subway station



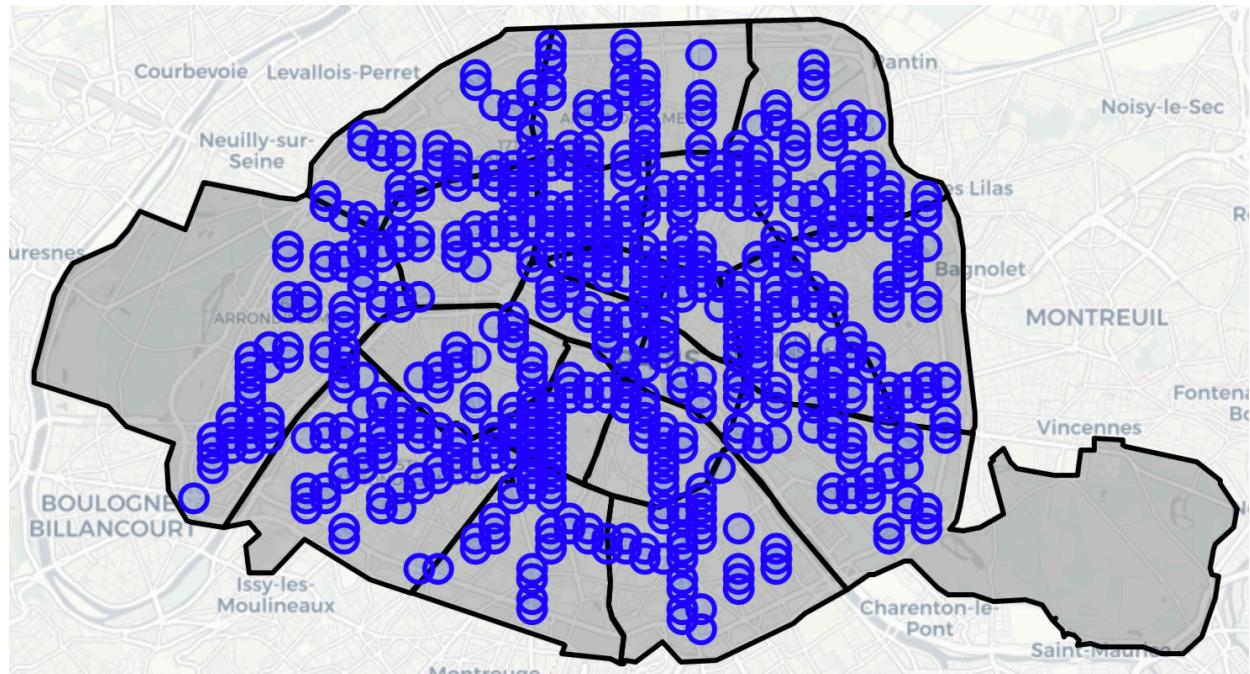
Methodology: Transactions in Paris

All the transactions
in Paris for the past
5 years



Methodology: With TXN and venues

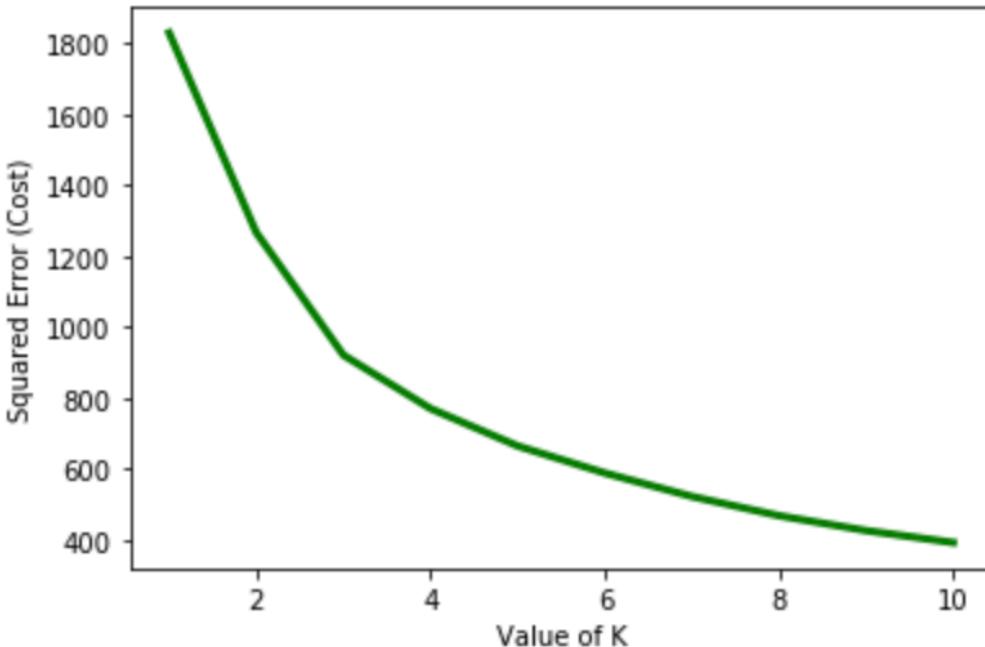
Keep only the 610 potential locations with at least 10 transactions and with venues



Methodology: Find the best value of K

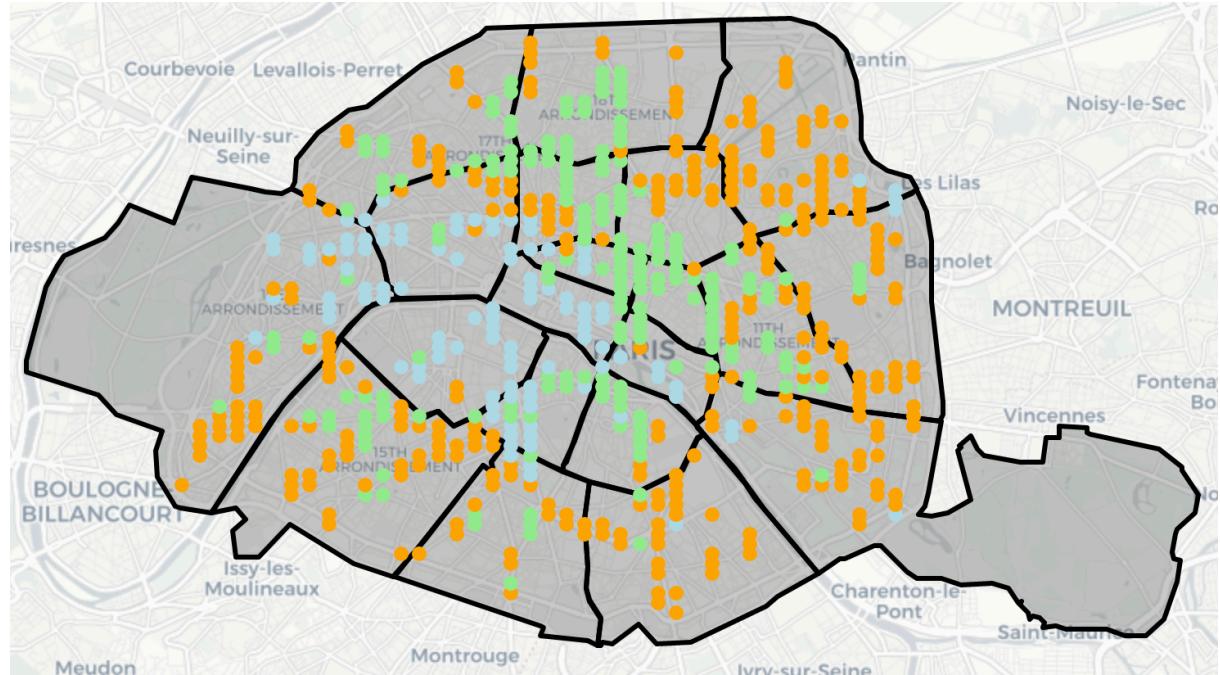
Use the elbow approach to find K.

Here $K = 3$



Methodology: K-Means Clustering

The remaining
locations distributed
in 3 clusters



Methodology: K-Means Clustering

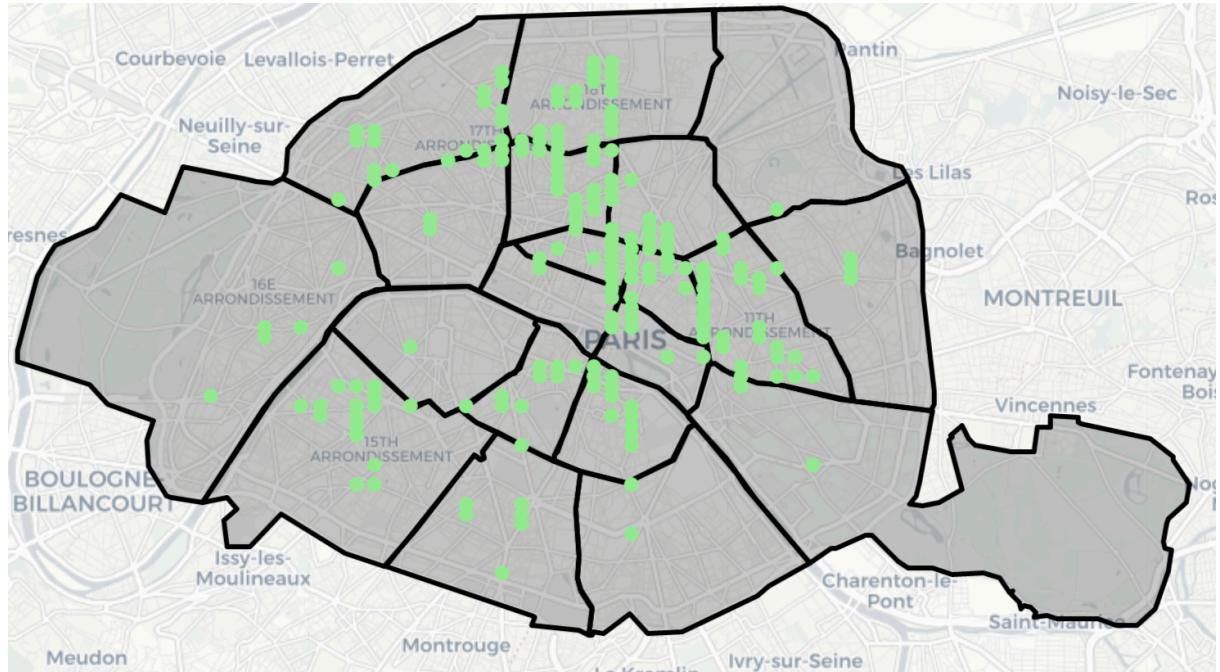
The averages for each cluster.

Best Choice = 1

Label	Transactions	Price	Venues
0	53.290221	9633.529968	11.851735
1	112.417112	10815.064171	27.609626
2	41.867925	14942.330189	22.632075

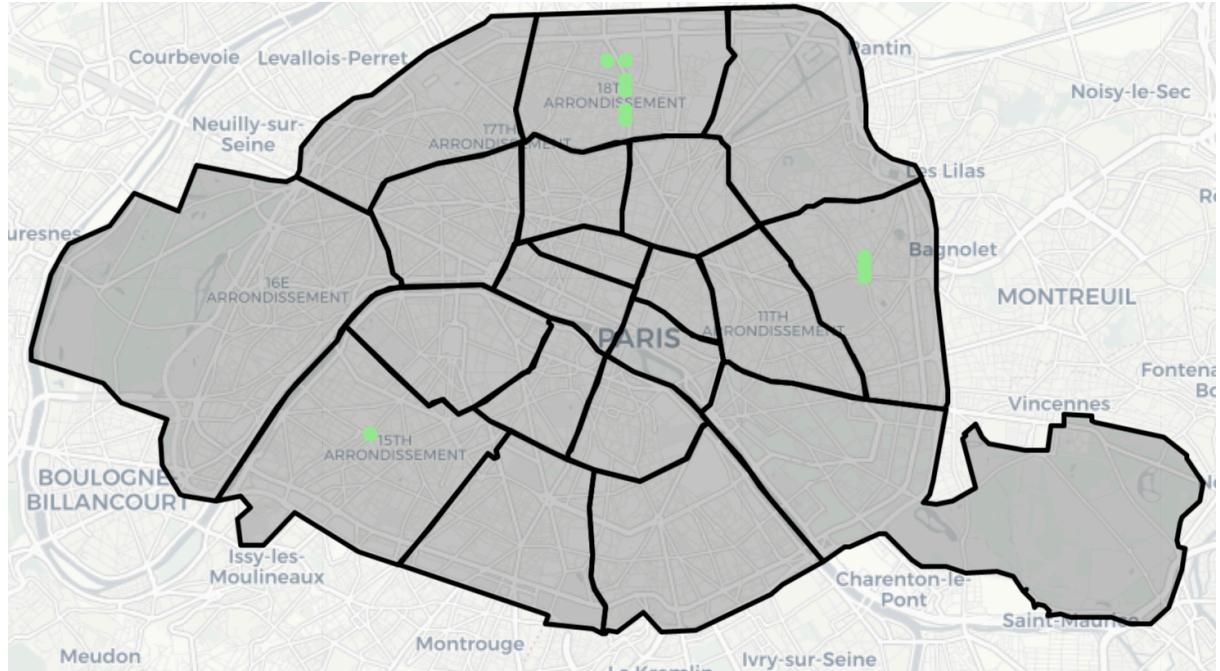
Methodology: Only Cluster 1

Keep only the 187 remaining locations of the cluster with label 1



Methodology: Top 10

The 10 best places
to live



Results and Discussion

- 90% of the proposals in two boroughs known as dynamics and cheap
- Crime rate higher
- Not ideal result due to a lack of data
- Choices made more or less objectives

Conclusion



- Make sure to have all the required data
- Have a clear idea of what is wanted by the stakeholder