On the Preferred Location of a Multifaceted Bar Business in Paris

IBM Data Science Professional Certificate

March 2020

Outline

- Background and business problem
- Data
- Methods
- Results and Discussion
- Conclusion

Background and business problem

- Paris: the most visited city in the world
  - 16 million foreign visitors a year
  - 450 million nights spent for accommodation
- Paris: the most expensive city in the world

- Business problem: **what is the best location(s) for opening a bar/coffee shop business in Paris that would tap both tourist and local populations, during day and nighttime?**

Data

- 16 variables covering **location**, **socioeconomic**, **demographic**, and **tourism** data
- 5 **location variables** sourced in Foursquare
  - Bars
  - Restaurants
  - Theaters and comedy clubs
  - Hotels
  - Museum and monuments

Data (cont'd)

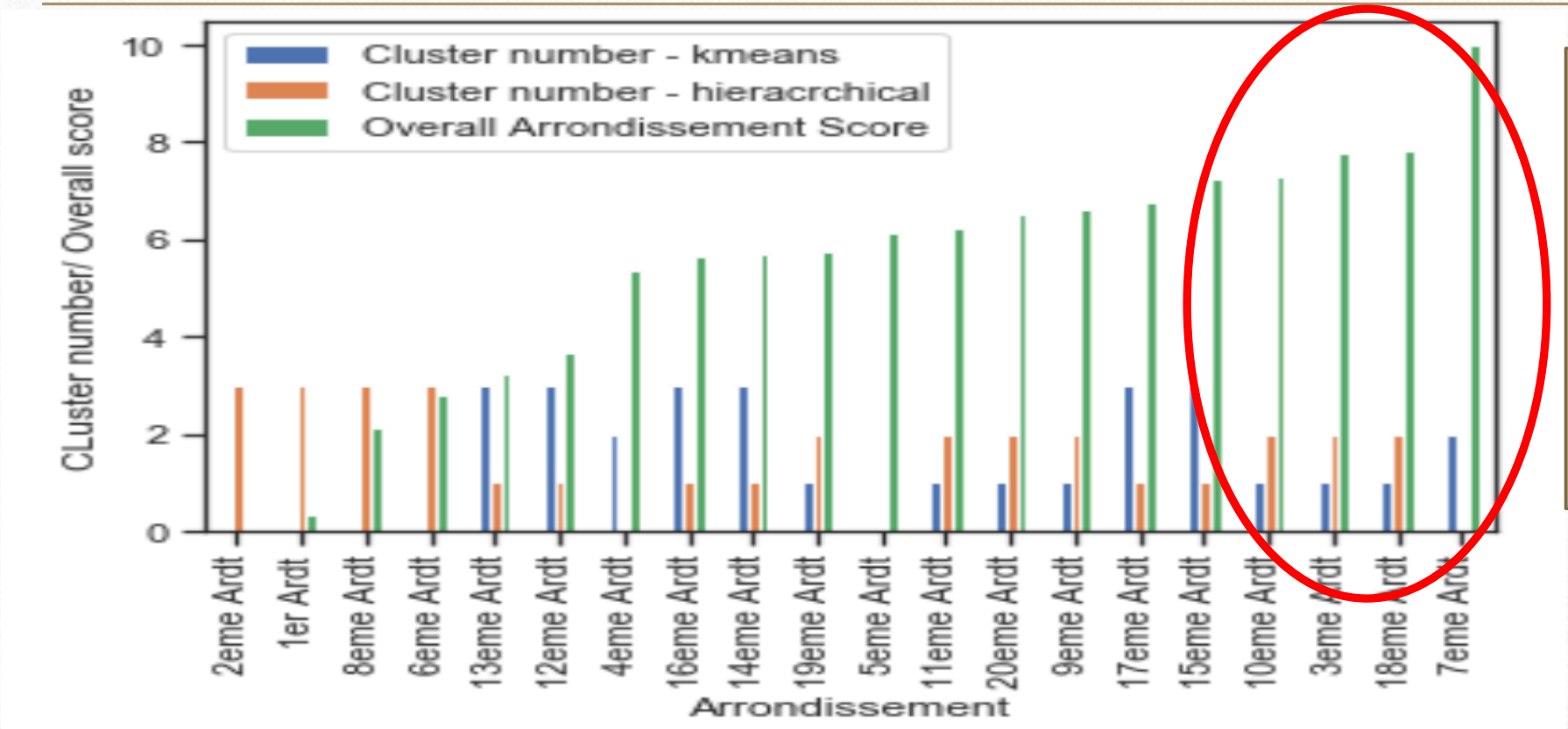
- **Wealth variables** (poverty rate, median annual income, percentage of taxable income households)
- **Population variables (**total population, population density, and percentage variation of population over 2013-2016)
- **Real-estate prices data** (2018 square meter price in euro, percentage change of that price over 2018–2019)
- **15–64 age category population density**, as a proxy fo "youth" clientele
- **Tourism data** of 15 most visited sites

Methods

- Analysis at the ***arrondissement*** **level**
- **Multipronged quantitative approach**
  - **Yields more assurance in final recommendations** based on the comparison of three widely used alternative ranking and clustering methodologies
- **Two classification methods**: kmeans clustering and agglomerative hierarchical clustering, and
- Construction of **a multilevel nested composite index** of the 16 indicators
-

Results and discussion

- Cluster numbers and overall attractiveness score of the 20 arrondissements



Potential *arrondissements* for consideration

Results and discussion (cont'd)

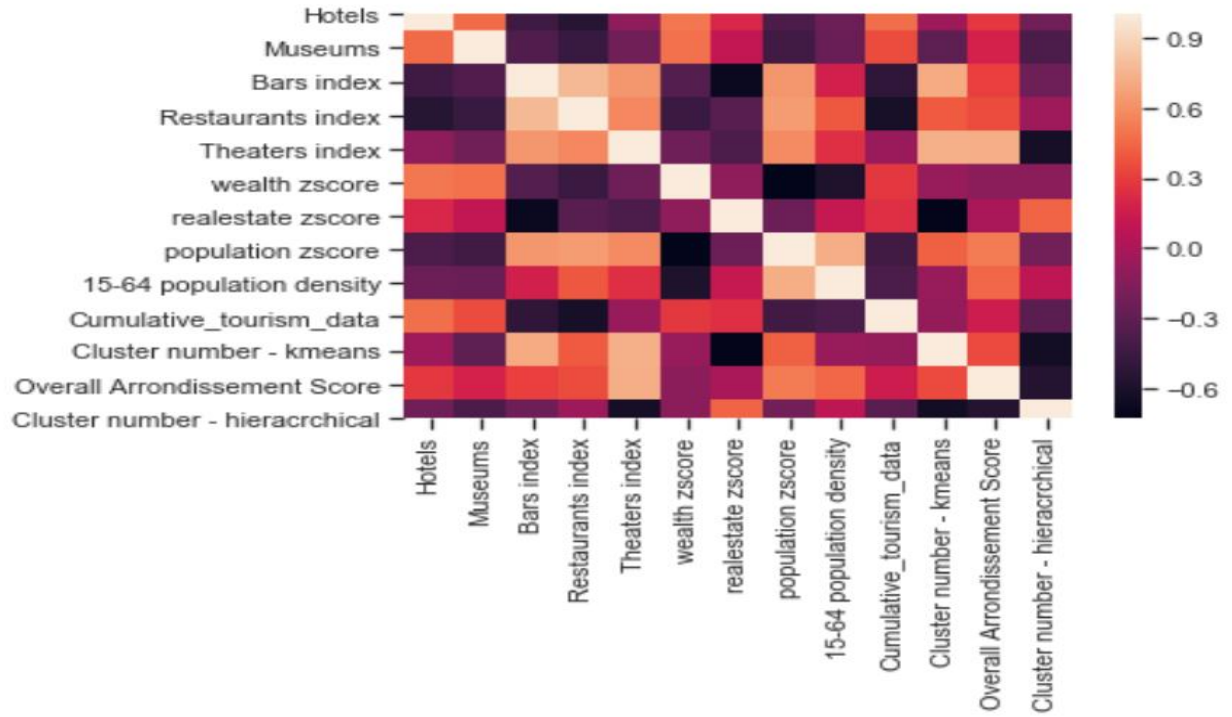
uster mber - ieans	Hotels	Museums	Bars index	Restaurants index	Theaters index	wealth zscore	realestate zscore	population zscore	15-64 population density	Cumulative_tourism_data	Overall Arrondissement Score	Cluster number - hieracrchical
0	0.270000	0.333333	0.277	0.411769	0.388472	0.769790	0.668508	0.178050	0.232780	0.226	0.328500	2.4
1	0.142857	0.095238	0.621549	0.782960	0.876017	0.347663	0.749715	0.786962	0.716990	0.090053	0.685442	2.0
2	0.575	0.388889	0.248114	0.298172	1.000000	0.822139	0.715841	0.269312	0.185426	0.972779	0.768986	0.0
3	0.166667	0.055556	0.898732	0.809	1.000	0.577	0.163	0.689	0.355	0.024	0.537718	1.0

Cluster number - racrchical	Hotels	Museums	Bars index	F	_tourism_data	Cluster number - kmeans	Overa Arrondisseme Sco
0	0.483333	0.592593	0.270234		0.717670	1.333333	0.7178
1	0.166667	0.055556	0.898732		0.024546	3.000000	0.5377
2	0.142857	0.095238	0.621549		0.090053	1.000000	0.6854
3	0.262500	0.166667	0.268755		0.231089	0.000000	0.1317

High attractiveness can build on different constituents structures:
Tourist/local populations
Daytime/nighttime revenues
Higher/lower real-estate prices
Etc.

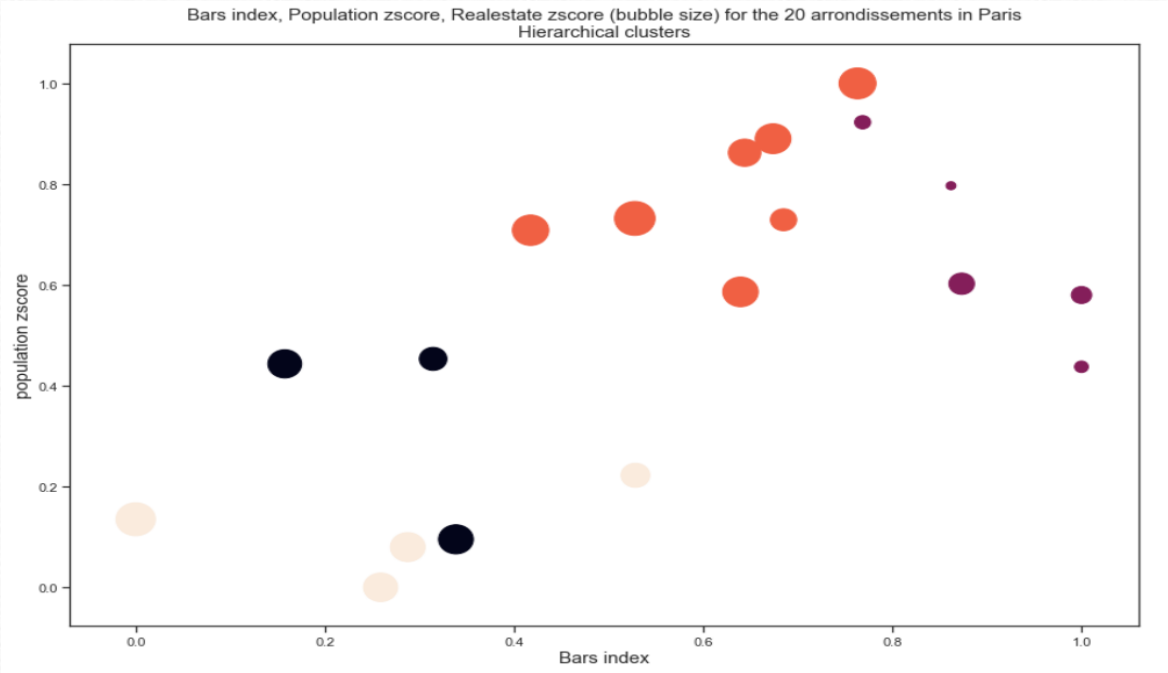
Results and discussion (cont'd)

- Higher attractiveness tends to have:
  - ✓ less competition,
  - ✓ more favorable local population dynamics,
  - ✓ yet less favorable real-estate prices and local residents' income levels.



Results and discussion (cont'd)

- Again, important arbitrages to consider, e.g.,
  - ✓ more favorable businesses competition
  - ✓ highly associated with favorable population dynamics
  - ✓ yet higher real-estate prices



Results and discussion (cont'd)

- Final recommendation boils down to the:
- **7th *arrondissement***: relatively higher share of revenues coming from tourists during daytime, taping a wealthier local albeit less important resident population, but faces a stronger business competition
- **18th *arrondissement***: much less pronounced competition; business relies more on a younger, more important, but less wealthy resident population.

Conclusion

- More refined analysis needed:
  - ✓ consider location at a lower spatial level, i.e., administrative neighborhood level (*quartiers administratifs*)
  - ✓ take into account official records regarding the numbers businesses actually established in those *arrondissements*
  - ✓ weighting of Foursqaure data used so far with clients' evaluations of the businesses.

THANK YOU FOR YOUR ATTENTION