On the Preferred Location of a Multifaceted Bar Business in Paris

I - Introduction and business problem description

Paris has been one of the most touristic cities in the world for a while now. It actually tops tourists' destination since at least the beginning of the 2010s, alternating this first place with London (see AFP, 2014; AFP, 2019a). [1]

Roughly 16 to 17 million foreign visitors visit Paris every year, with around 450 million nights spent by tourists for accommodation (as of 2018), including hotels, campsites, and youth hotels— yet excluding home-sharing platforms like Airbnb, (see AFP; 2014) [2]. The main 60 touristic sites or so total around 80 million visits yearly, with —according to the *Office du Tourisme et des Congrès de Paris*[3]—roughly 14 million visitors for the *Cathédrale Notre-Dame de Paris*, 11 million for the *Basilique du Sacré-Cœur de Montmartre,* or 9 million for the *Musée du Louvre*.

Meanwhile, as of 2015, Paris is one the most expensive—and as a corollary, on average, wealthy—city in the world (see AFP, 2019b quoting a 2019 Economist Intelligence Unit report),[4] trailed by Zurich, Geneva and Osaka.[5] Even more, Paris was the only eurozone capital in the top 10 most expensive cities in the world, with a woman's haircut and a two-piece business suite respectively quoted around 150 and 1700 euros in 2018.

Against this background, I was approached by a foreign investor who envisages to open a *bar/pub* in Paris. The investor seeks advice on the location of such a business, a location that would tap both the tourists' population—acknowledging the fact that numerous, if not a fair majority of, tourists are as interested in the city's nightlife, as they are in its cultural sites—as well as the local, resident clientele which will support smoothing the business's revenues in view of the seasonality of tourism influxes. The business would also function as a coffee shop during daytime. Tapping both foreign and resident clienteles, as well as mainly night (and secondarily day) consumers needs is intended to ensure a solid, long run, financial viability of the project.

---

[1] See respectively https://www.thelocal.fr/20140311/paris-still-worlds-top-tourist-destination, and https://www.thelocal.fr/20190410/france-retains-crown-as-most-visited-country-on-earth.

[2] See https://www.thelocal.fr/20190410/peak-year-for-tourists-in-france-despite-yellow-vests-and-rail-strikes.

[3] See https://pro.parisinfo.com/etudes-et-chiffres/enquetes-et-dossiers.

[4] https://www.thelocal.fr/20190319/paris-tops-ranking-of-worlds-most-expensive-cities-to-live-for-the-first-time.

[5] https://www.france24.com/en/20190319-paris-france-world-most-expensive-cities-singapore-hong-kong; see also https://www.novinite.com/articles/195890/Paris+is+Now+the+Most+Expensive+City+in+the+World.

The preceding paragraphs have laid down some of the main components of the business problem which can be simplistically stated as follows: **what is the best location(s) for starting the aforementioned, multifaceted business in Paris?**

The area of the city exceeds 100 km2 indeed. It counts 2.2 million inhabitants with wide spatial variation in income levels and other sociodemographic characteristics, and is organized in 20 *arrondissements* (let us view this as a borough), and 80 administrative neighborhoods. In this preliminary report, I identify such a location at the *arrondissement* level.[6]
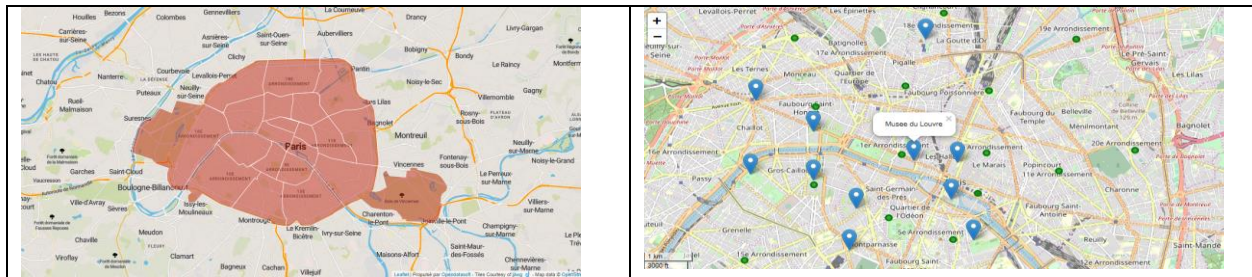
Note: In the final report, I provide a link to the notebook and code that will allow the reader to interact with the map and identify the main touristic sites in Paris along with the *arrondissement* they belong to.

II - Data description

Giving the investor solid scientific insights into the most favorable locations for the business under consideration requires a diverse set of data—a total of 16 variables that required preprocessing and cleaning to a significant degree—to capture various characteristics of the prospective catchment area(s) that will be advised. Here, I describe the data and its sources, along with some visualization for descriptive—rather than analytical—purposes.

II.1 Foursquare sourced data

I make use of Foursquare to obtain a first group of variables that relate to the business's direct and indirect competitors, namely the bars, nightclubs, and restaurants. A high concentration of such businesses in a particular area is assumed to reduce the attractiveness of that location for our investment.

Among the 90 different categories of venues rendered by Foursquare, the category *Bars* are referred to by the following labels: Bar, Pub, Brasserie, Bistro, Brewery, Nightclub, Beer Garden, Gastropub, Lounge. Similarly, the *Restaurants* category is referred to through the labels Restaurant, Diner, Creperie, Fish & Chips Shop, Snack Place or Tea Room; while the labels Theater, Comedy, Club and Movie Theater refer to our category *Theaters*.

---

[6] I will keep using the

I also use Foursquare to extract data related to two more categories that I label *Hotels* and *Museums* (the latter covered by the labels Museum, Art Museum, History Museum, Monument / Landmark, and Science Museum). These are expected to have a favorable impact on my recommendation for a given location, in view of the fact that a high concentration of these increases our business's patronage, by night and day respectively. The following figure gives a snapshot of the five categories' counts.
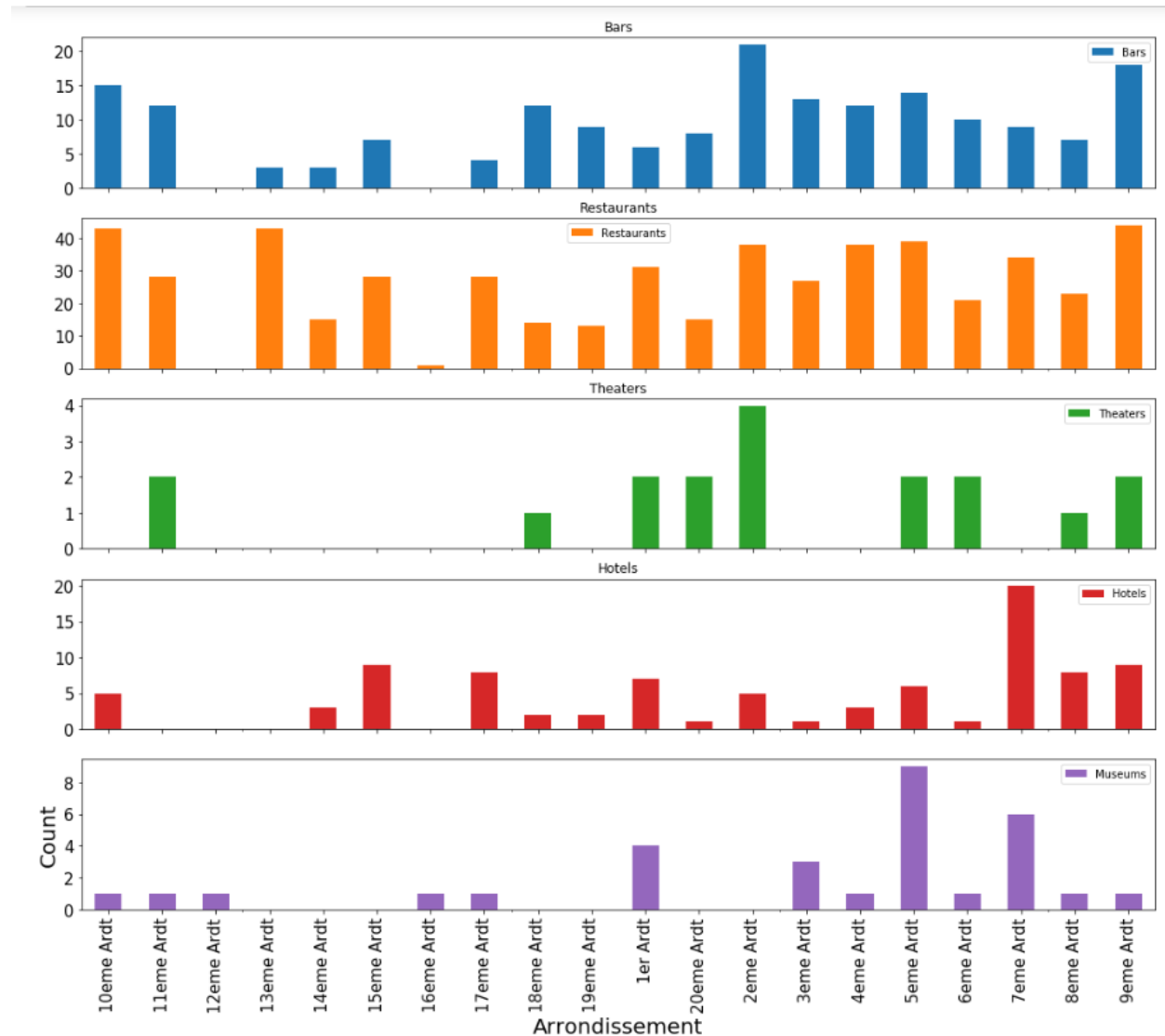

It is important to note that this data does not represent the full count of businesses in Paris's 20 *arrondissements*, but only those businesses that appear in Foursquare, which in the case of a non-anglophone city like Paris, reflect tourists' testimonies. This warrants strong skepticism over utilizing such data for decision-making purposes in a real-life problem solving.

Below are a couple of more figures that compare *Bars* and *Restaurants* counts. This shows, for example, that *Restaurants* testimonies are more frequent than Bars testimonies, even though there are differences in the *arrondissements* in this respect.
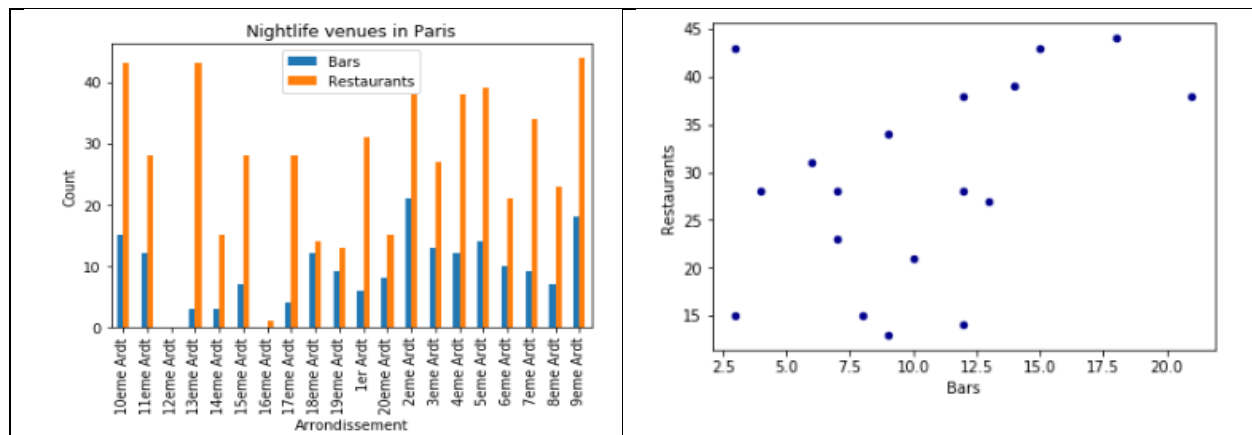
Secondly, and I will get back to this later, while this data gives a sense of the raw numbers of businesses listed in Foursquare, it needs—at least part of it—to be scaled to take into account the differences of the area and population of the *arrondissements* if it is to be used meaningfully in a quantitative analysis. Hence the *Bars* data for the 4th and 7th *arrondissements* for example, are not directly comparable as their areas are 7 and 16 square kilometers, and their populations of 182 000 and 165 000 respectively.

Consequently, I scale the *Bars*, *Restaurants*, and *Theaters* counts by dividing these by the *arrondissements'* population densities (number of inhabitants per square kilometers). Further making use of a variable that classifies the arrondissements three categories which I label *Least*, *Moderately*, and *High touristic* according to their attractiveness to tourists (see paragraph II.2.4), I am able to produce a matrix scatter plot that gives interesting insights into this extracted Foursquare data. For instance, one can see that high touristic sites have received more assessments for their bars and restaurants which lends support to the fact that those businesses that are established in high touristic locations rely proportionately more for on tourists' consumption for their revenues and profits—assuming the behavior of tourists, as it relates their decision to assess bars and restaurants in Foursquare, is not skewed accross touristic attractiveness.

Touristic_sites_dummy
- Least touristic
- Moderately touristic
- Very touristic
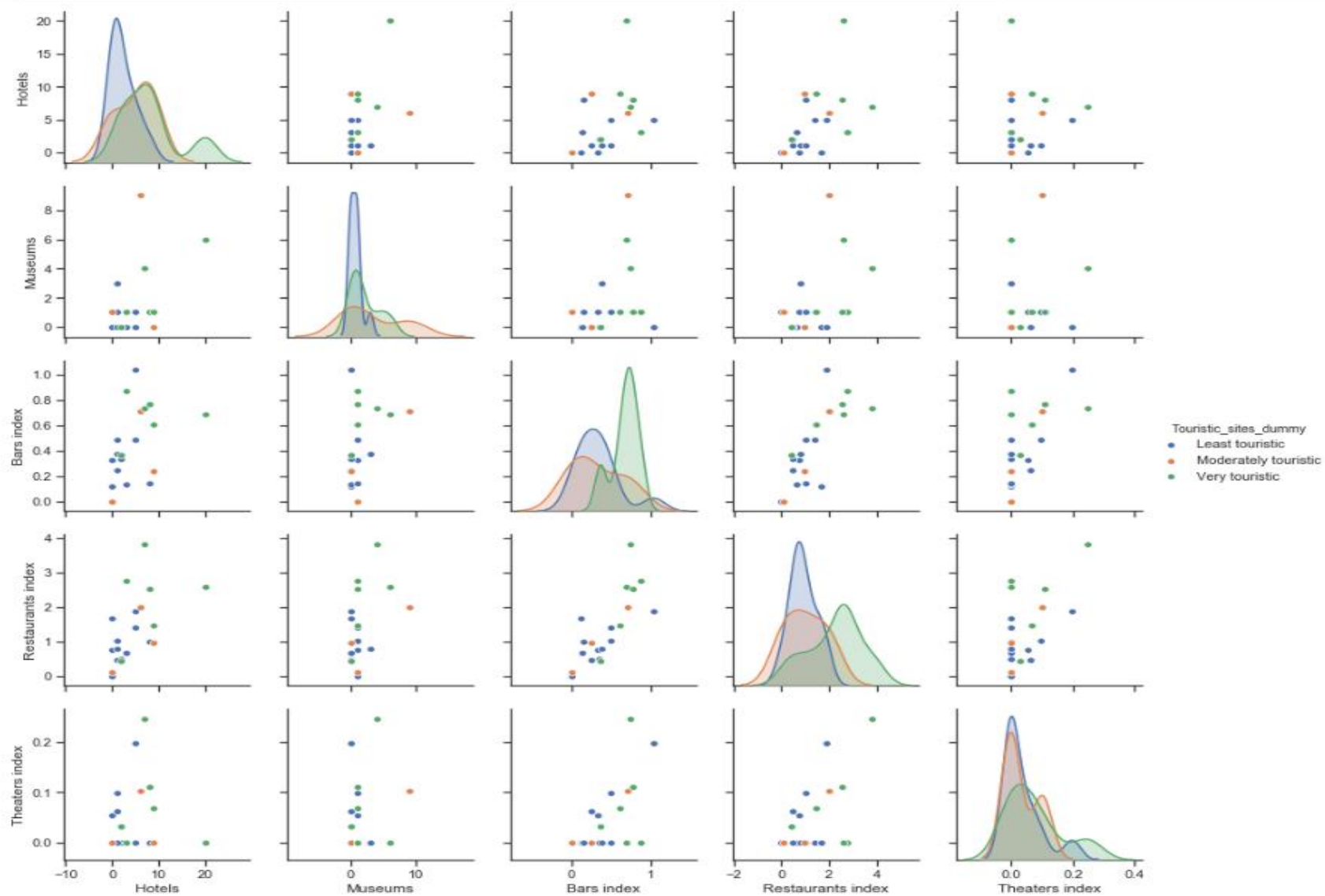
II.2 Socioeconomic, demographic, and touristic data

II.2.1 Wealth data

Wealthier locations are more attractive for any business. I use three socioeconomic indicators to capture that wealth effect, namely the poverty rate, the median annual income, and the percentage of taxable income households (sourced in French National Institute of Statistics and Economic Studies, INSEE). [7] Based on these, I construct a composite indicator, after normalizing these variables using the min-max method. This allows the ranking of the *arrondissements* according to their wealth level.

II.2.2 Population data

Similarly, I create a composite indicator based on three variables related to the *arrondissements'* population, namely the total population, the population density, and the percentage variation of population over 2013-2016. A higher value on that indicator increases the attractiveness of a given location.

II.2.3 Real-estate price data

Locations with high real-estate prices negatively impact my recommendation for a given location, even though they indirectly correlate to the wealth indicator. I retain two variables, the 2018 square meter price in euro, and the percentage change of that price over 2018–2019, which I use to create the real-estate composite along the same lines described in the preceding to paragraphs.

II.2.4 Other demographic and tourism indicators

The final set of indicators first includes one demographic variable, the 15–64 age category population density—our target local clientele— which is assumed to positively impact the assessment of an *arrondissement* location. The last two indicators are related tourism. The first of these adds up the tourism data of the touristic sites—among the 15 most visited sites in 2018— belonging to the various *arrondissements*, while the last variable classifies the arrondissements into three categories according to the degree to which they were preferred destinations for tourists.

---

[7] All socioeconomic and demographic data is official statistics provided by the French National Institute of Statistics and Economic Studies. See https://www.insee.fr/en/accueil.

Arrondissement Number	15-64 population density	Cumulative tourism data	Tourism patronage
10eme Ardt	23328		Least touristic
11eme Ardt	29496		Least touristic
12eme Ardt	5988		Least touristic
13eme Ardt	17385		Least touristic
14eme Ardt	16703		Least touristic
15eme Ardt	18903	1.17	Moderately touristic
16eme Ardt	6369	1.4	Moderately touristic
17eme Ardt	20720		Least touristic
18eme Ardt	23558	11	Very touristic
19eme Ardt	18724		Least touristic
1er Ardt	6266	9.13	Very touristic
20eme Ardt	22745		Least touristic
2eme Ardt	15868		Least touristic
3eme Ardt	22173		Least touristic
4eme Ardt	12231	17.45	Very touristic
5eme Ardt	16076	3.62	Moderately touristic
6eme Ardt	12633		Least touristic
7eme Ardt	8463	16.5	Very touristic
8eme Ardt	6458	7	Very touristic
9eme Ardt	19877		Very touristic