

# Title: Diversity in food culture: comparing availability of international cuisines in selected metropolitan cities

Author: Aleksandras Gutmanas

Date: Feb-Mar 2020

## Introduction

### Background

Most large cities are diverse places, whose populations include large numbers of migrants hailing from nearly everywhere on the planet and enriching their new environments. Perhaps the most tangible expression of such an enrichment is our ability to access a large variety of cuisines in a global metropolis like London or Paris, and even in smaller centres such as Stockholm and Vilnius. Having said that, diversity can often be limited by segregation, usually caused by a very natural desire for newcomers to settle in close proximity to their former compatriots, thus creating a familiar milieu in a new place of residence. The evidence of this natural segregation is ample in neighbourhood names, such as the ubiquitous *Chinatowns* and *Little Italies* in many North American cities, mostly established about a century ago. It is thus also possible for these diverse segments of population to co-exist in parallel spaces, without really interacting with each other. The questions of diversity and segregation have of course been discussed and debated in academic and political contexts, and I like to think of them as linked to the ideas of Mono-, Multi-, Inter- and Transculturality. For the purposes of this project, a simplified description of these theories will suffice, and a curious reader is referred to Benessaieh (2010) and Welsch (1999) for a more thorough discussion. Monoculturality refers to the idea that cultures can be likened to islands, completely self-sufficient and separate from each other, with only sporadic contacts between them coloured by indifference or suspicion if not outright hostility. In the context of food, one can think of a 100% authentic traditional dish consumed by people who can also prepare it. Multiculturality describes a condition where cultures co-exist in close physical proximity (e.g., in the same city), but the contacts are limited to a bare minimum of mostly business-like transactions: that same authentic dish can now be occasionally consumed by “the others”. Interculturality still treats cultures as islands, distinct and separate from each other, but the contacts are numerous and varied, with frequent borrowing of ideas (or food ingredients in our case). And finally, Transculturality rejects the idea of cultures as islands, instead treating them as constantly changing streams with blurred boundaries. From the culinary aspect, this could be likened to a fusion dish.

### The problem

A lot of the discussion on matters of cultural diversity does not go beyond opinions and anecdotal data. For example, a respected chef claims that Britain is more open to food from the rest of the world than other European countries (Slater, 2015). This may be true, but can we quantify it? If yes, how? In this project, I will leverage the venue information from the Foursquare API and cluster neighbourhoods in four European cities: London, Paris, Stockholm and Vilnius, with the intent to quantify the degree of general diversity (a proxy for Multiculturality), physical proximity of restaurants offering different cuisines (a proxy for Interculturality) and will also count the number of fusion restaurants (a proxy for Transculturality). The particular choice of the cities is purely arbitrary at this stage, and is inspired by a personal curiosity: I know these cities fairly well, having visited them many times over the years.

## Stakeholders and potential interest

Beyond my personal curiosity and the purely academic interest in quantifying the degree to which cities are multi-, inter- or “transculinary”, if I may be allowed to create a new word, there are other stakeholders for this analysis. For example, city councils and tourist agencies may replicate this work and use the findings to advertise their cities; a business person wishing to open a new restaurant could use the data to select a suitable location (e.g., gaining an appreciation of what food the local residents like), and tourists could learn at a glance what to expect when visiting a new city.

## Data

### Data sources

The project relies on two main types of data. The first is the geographic data on each city, which I have obtained from the following sources:

- 33 London boroughs from <http://insideairbnb.com/get-the-data.html>. Surprisingly, I found no immediate way to download this data from a Government source.
- 20 Paris boroughs (“arrondissements”) (available from the French Government’s open data portal at <https://www.data.gouv.fr/en/datasets>) are very compact, and do not compare easily with the Greater London data set above. The same is true for the AirBnB dataset (URL above). I have therefore obtained the data for the “Metropole du Grand Paris”, or the Greater Paris area, which consists of 12 areas and besides the city itself, also includes a ring of suburbs around Paris, linked with the city centre, and to an extent with each other by an extensive transport network. The data is at [https://opendata.apur.org/datasets/4c41f3fc9d96418c829855041bff5970\\_0?geometry=0.286%2C48.512%2C4.477%2C49.144](https://opendata.apur.org/datasets/4c41f3fc9d96418c829855041bff5970_0?geometry=0.286%2C48.512%2C4.477%2C49.144);
- Stockholm neighbourhoods from <http://insideairbnb.com/get-the-data.html> suffer from a similar problem as Paris – only the compact inner city is made available. The Greater Stockholm area is nowadays defined as the entire Stockholm county (see <https://sv.wikipedia.org/wiki/Storstockholm> in Swedish). Fortunately, <https://www.thenmap.net/> offers lower resolution, but acceptable borders for Swedish counties (“län”) and municipalities (“kommun”). Not only that, the team even produced a PyPi package to access their API;
- In the case of Vilnius, the only comprehensive dataset I came across includes regions (counties), municipalities and subdistricts for the entire country (Lithuania). These datasets were prepared by Julius Seporaitis (<https://github.com/seporaitis>) and made available on his GitHub account at <https://github.com/seporaitis/lt-geojson>. I have cloned this data, but will only require the municipality for Vilnius city itself, and for the suburban region around it.

In each case the data comes in the standard GeoJSON format, making it compatible with the Folium library for creating maps.

The second source of data will be venue locations and categories provided by the Foursquare API: <https://developer.foursquare.com/docs/api/endpoints>. I will primarily use the “search” endpoint to obtain lists of venues, and will also need to process the category tree available from the “categories” endpoint.

### Data clean-up and preparation

For the geospatial data, I will require the coordinates of each neighbourhood’s centre to use as latitude and longitude in Foursquare API queries. The main problem I encountered in obtaining definitions of small neighbourhood is that there isn’t really a good open source of this info for any of the four cities. Also, my goal is to cover each city in as complete a fashion as possible, while neighbourhoods could be of arbitrary shape and size. So instead, I decided to cover each metropolitan area in a lattice of hexagons. Hexagons approximate circles well enough. I chose a size of the hexagon by the limitation of the Foursquare API, which returns no more than 50 venues in a result set. A radius

of 0.5 km tends to be small enough to be able to fit within this limit. Table 1 summarises the geographical coverage of each metropolitan area.

For the venues data, the main problem is to filter relevant categories (e.g., I removed fast food chains and coffee houses such as Starbucks). I needed to remap venue categories to a higher node in the category tree: for instance, I reclassified an Alsatian restaurant as French for the purposes of this analysis (even if Alsatians may object to this particular decision). I have also created cultural categories of my own, grouping some of the smaller categories into a larger one. For example, Foursquare has separate categories for Czech, Ukrainian, Russian, Eastern European, Hungarian restaurants. I grouped them all (and a few others) as “East European”. Appendix A shows this manual remapping. Table 1 also gives the number of venues that were returned by the Foursquare API for each city and that were assigned such a cultural category, and the number of areas (hexagons) with such venues.

*Table 1. Summary of geographic coverage and venue data for the four metropolitan areas*

Metropolitan area	Number of boroughs	Number of hexagons	Number of categorised venues	Number of hexagons with venues	Number of hexagons with 50+ venues
London	33	2417	17112	1815	169
Stockholm	26	9435	3751	800	34
Paris	12	1261	10930	1037	107
Vilnius	2	3897	1036	180	14

## Methodology

### Exploratory data analysis

#### *Feature selection and data transformation*

Once the data clean-up is complete, each neighbourhood becomes associated with a filtered list of venues, each labelled with a particular cultural category (e.g., South Asian or French), which are the features I want to cluster the neighbourhoods on. I used the one-hot encoding to transform these lists into a frequency table with each cultural category being a feature. This allows me to cluster the neighbourhoods by their cuisine/cultural profiles, but prior to doing so, I grouped categories with few venues (<100 for London, Paris and Stockholm and <30 for Vilnius) with the “Other” group. This resulted in 12 features for London, 7 for Stockholm, 10 for Paris and 8 for Vilnius (Table 2). I then normalised each profile by the total number of venues in each neighbourhood to account for the fact that residential neighbourhoods typically have fewer venues in total. My question is about diversity, rather than availability, so normalisation had to be done per area/hexagon.

*Table 2. Extracted cultural features for the four metropolitan areas*

Metropolitan area	Number of features	Top 1	Top 2	Top 3
London	12	Mediterranean	Asian	South Asian
Stockholm	7	Mediterranean	Asian	Scandinavian
Paris	10	French	Mediterranean	Asian
Vilnius	8	Mediterranean	Asian	East European

It is an interesting observation that Mediterranean and Asian restaurants are two of the most popular categories in each of these cities. Another interesting point is that in three of the four cities, the cuisine associated with the country or region is also among the top three (top 1 in Paris and top 3 in Stockholm and Vilnius). London is an outlier, as British cuisine is only top 4 there.

### Number of venues per hexagon

As expected, in Vilnius and Stockholm, which are smaller of the four, most hexagons have just 1 or 2 venues, While in London and Paris, the density is higher (Figure 1).

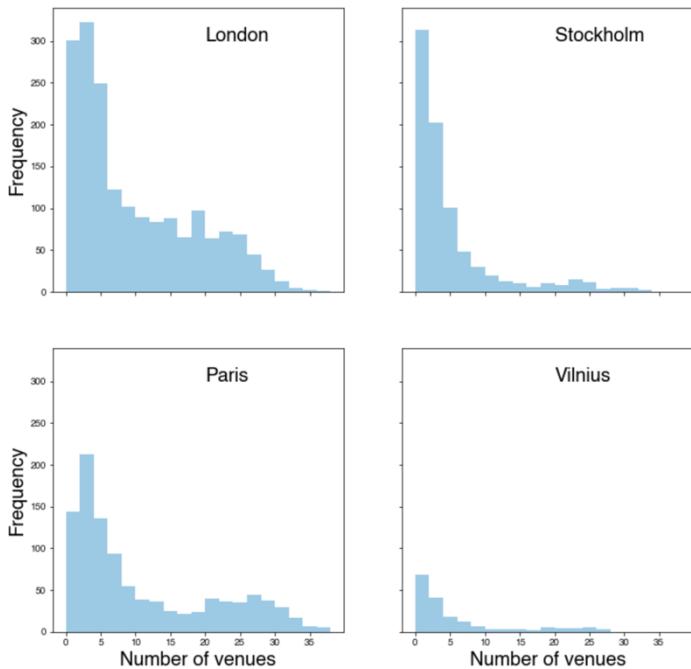


Figure 1. Density of venues in the four cities.

### K-means clustering

The goal of the project is cluster neighbourhood in each city based on their cultural profiles. It is appropriate to use non-hierarchical clustering algorithms, as at this point I am not interested in detailed relationships between different areas. K-means is a suitable approach. K-means clustering was performed with the help of the scikit learn package with the following parameters: k-means++ method, and 12 rounds of cluster seeding, leaving other parameters at defaults. I have iterated over the number of expected clusters and then evaluated each result with in-model scoring methods, as there is no “ground truth” to score against. In the case of London and Paris, the number of clusters can be determined with the help of the Calinski-Harabasz score (Calinski and Harabasz, 1974), which aims to maximise the ratio between inter- and intra-cluster variability. It is implemented in the scikit learn library. I tested two other scores (Silhouette and Davies-Bouldin) but they were inconclusive (Figure 2). I chose 6 clusters for London and 4 for Paris for the final mapping.

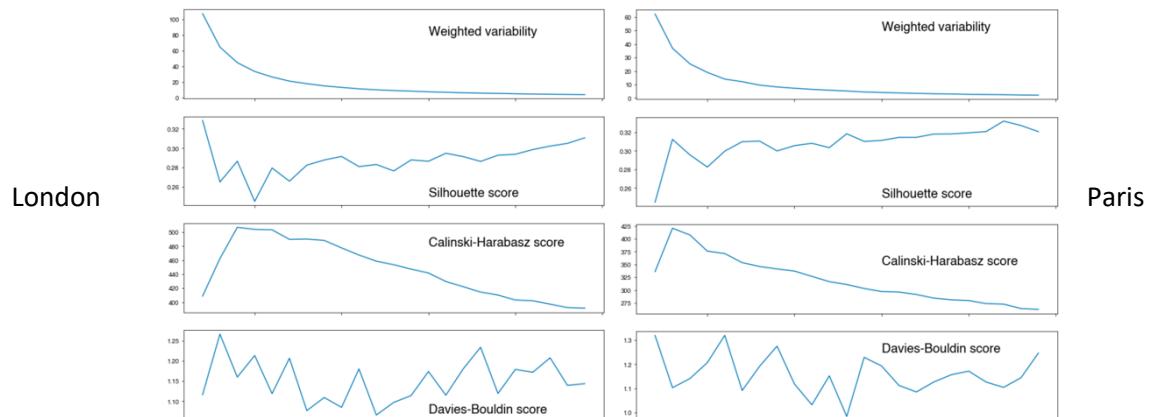


Figure 2. Evaluating goodness of clustering for different numbers of clusters. Left: London. Right: Paris

Similar analysis for Stockholm and Vilnius is completely inconclusive (Figure 3).

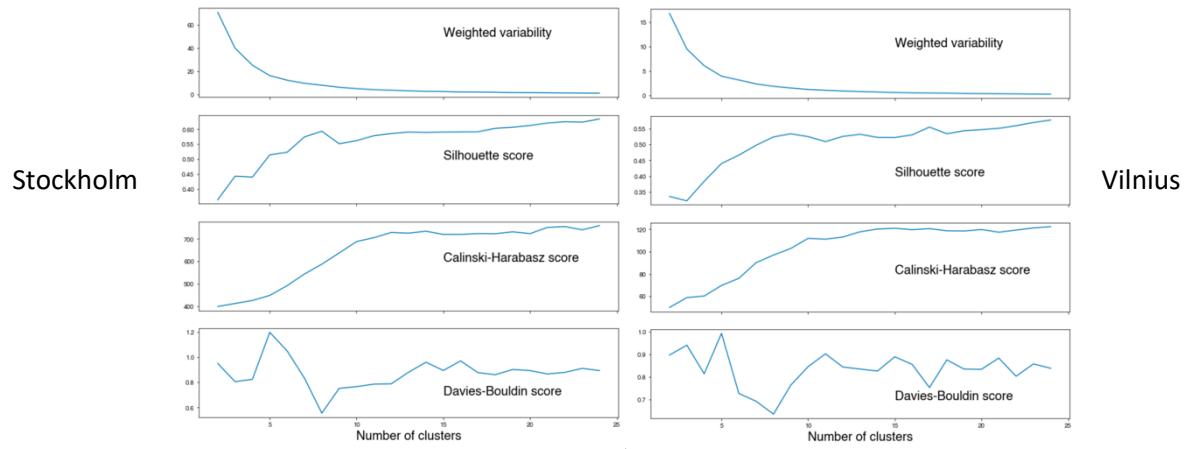


Figure 3. Evaluating goodness of clustering. Left: Stockholm. Right: Vilnius

Given that there were only 7 and 8 features in these two datasets, the maximum number of clusters should not really go beyond this number. I have therefore chosen 8 clusters.

## Results

### London clusters

Table 3 shows the resulting 6 London cultural clusters with frequencies of each cuisine's appearance in each cluster.

Table 3. Cultural profiling (in %) of London areas, and size of each cluster

Cluster ID	0	1	2	3	4	5
African	1	2	1	0	0	0
American	2	5	1	2	2	0
Asian	12	25	12	3	12	91
British and Irish	5	14	6	87	7	2
Caribbean	2	3	1	1	1	0
East European	1	1	1	1	0	0
French	0	2	1	1	2	0
Latin American	0	2	0	0	1	0
Mediterranean	7	20	3	0	59	2
Middle Eastern	61	9	3	1	5	1
South Asian	7	15	70	2	9	3
Rest of the world	1	3	0	1	1	0
Total number of hexagons	105	978	218	133	271	110

There are clear patterns – the largest cluster is also the most mixed, while smaller clusters have stronger signal for a particular cuisine. Geographically (Figure 4), these clusters follow a pattern. The mixed cluster (1, red) takes nearly the entire central London with some smaller pockets of Mediterranean (4, pale blue) and Middle Eastern (0, green) cuisines. The South Asian (2, purple) is concentrated on the West side of the Greater London area.

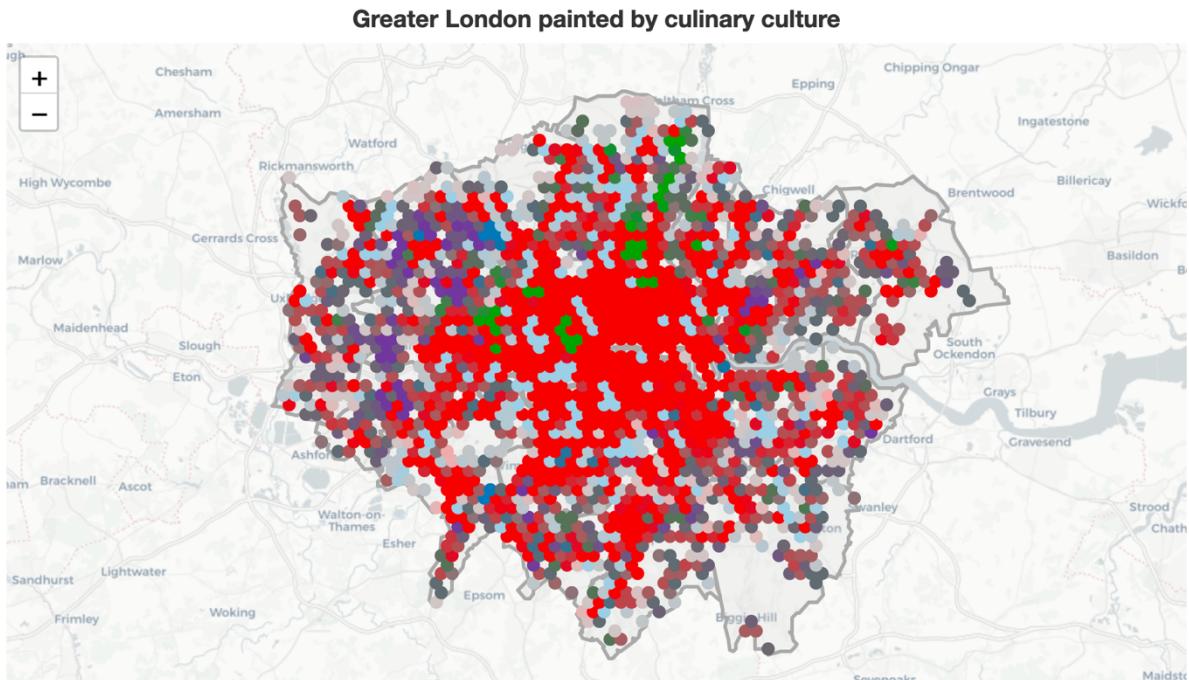


Figure 4. Clustering of neighbourhoods in the Greater London. Red – mixed cluster, purple – South Asian, pale blue – Mediterranean, deep blue – Asian, green – Middle Eastern, pink – British and Irish. Areas with fewer than 10 venues have their colors faded.

## Paris clusters

A different pattern emerges for Paris (Table 4 and Figure 5). The centre of Paris is dominated by French cuisine. This French cluster is surrounded by a ring of the mixed cluster, and only further out does one see a fainter fragmented ring with pockets of Mediterranean, French, and Asian cuisines.

Table 4. Cultural profiling (in %) of Parisian areas, and size of each cluster.

Cluster ID	0	1	2	3
African	1	0	0	1
American	3	0	1	2
Asian	22	3	8	64
British and Irish	2	1	1	1
French	19	2	67	3
Latin American	2	0	1	1
Mediterranean	28	92	14	20
Middle Eastern	11	2	3	3
South Asian	6	1	2	2
Rest of the world	6	0	2	3
Total number of hexagons	544	119	211	163

Greater Paris painted by culinary culture

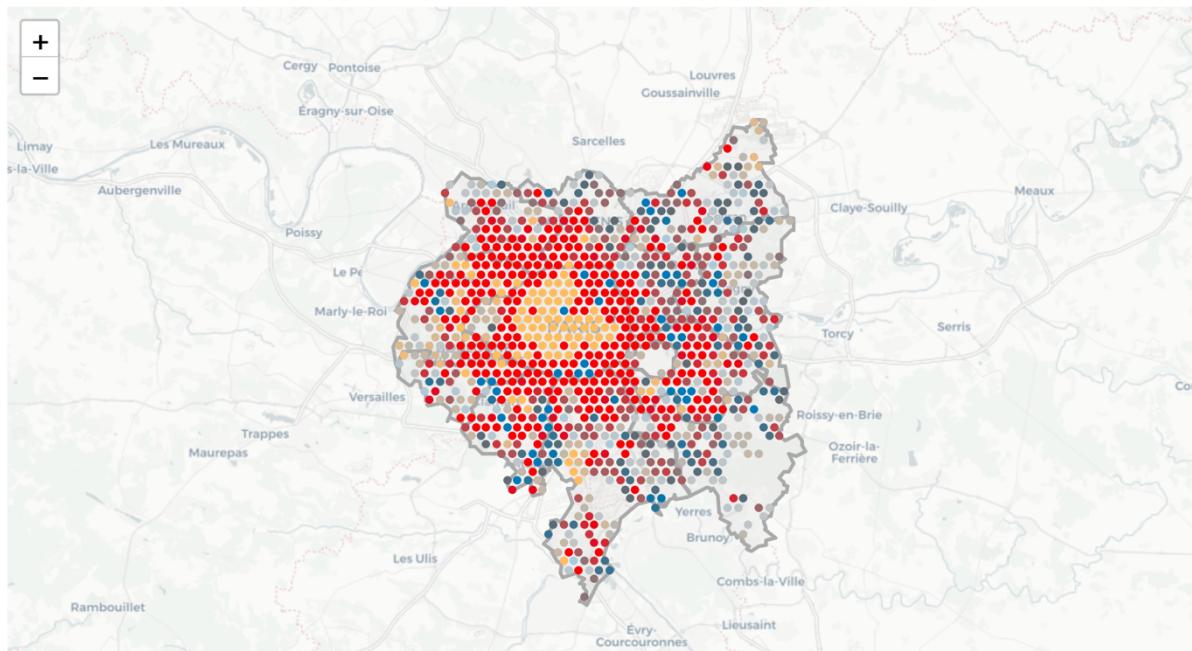


Figure 5. Clustering of neighbourhoods in the Greater Paris. Red – mixed cluster, yellow – French, pale blue – Mediterranean, deep blue – Asian. Areas with fewer than 10 venues have their colors faded.

## Stockholm clusters

As discussed in the methodology section, clustering for Stockholm and Vilnius was not as robust. These two cities had fewer venues over a larger territory and thus lower venue density. Thus many of the areas (hexagons) contained just 1 or 2 venue (Figure 1) and their profiles were thus very “pure”. As a result, the clustering algorithm tends to group such areas together. Table 5 and Figure 6 summarise the results for Stockholm. As for London and Paris, the “Mixed” cluster (3) is the largest,

*Table 5. Cultural profiling (in %) of Stockholm areas, and size of each cluster.*

Cluster ID	0	1	2	3	4	5	6	7
American	0	2	2	3	100	0	0	0
Asian	0	6	0	38	0	12	94	5
Mediterranean	97	14	2	38	0	9	2	19
Middle Eastern	1	3	0	4	0	0	1	68
Scandinavian	1	5	95	8	0	0	1	3
South Asian	0	1	0	4	0	79	1	4
Rest of the world	0	70	1	5	0	0	1	1
Total number of hexagons	206	33	61	343	22	17	80	38



*Figure 6. Clustering of neighbourhoods in the Stockholm county. Red – mixed cluster, yellow – Rest of the world, pale blue – Mediterranean, deep blue – Asian, orange – Scandinavian, green – Middle Eastern, brown – American, purple – South Asian. Areas with fewer than 5 venues have their colours faded towards grey. Inset map shows clusters without colour fading, emphasizing hexagons with few venues.*

## Vilnius clusters

And a similar picture with a different combination of exact profiles emerges for Vilnius, summarised by Table 6 and Figure 7. The largest cluster is again the mixed one!

Table 6. Cultural profiling (in %) of Vilnius areas, and size of each cluster

Cluster ID	0	1	2	3	4	5	6	7
American	2	12	75	5	0	0	0	0
Asian	93	0	0	25	0	1	0	0
British and Irish	0	0	0	4	0	1	4	3
East European	0	10	18	12	0	0	96	3
Latin American	0	0	0	5	0	0	0	55
Mediterranean	0	5	0	33	6	97	0	30
Middle Eastern	4	0	7	11	94	1	0	5
Rest of the world	2	73	0	6	0	0	0	3
Total number of hexagons	9	13	14	77	10	27	19	11

Vilnius painted by culinary culture

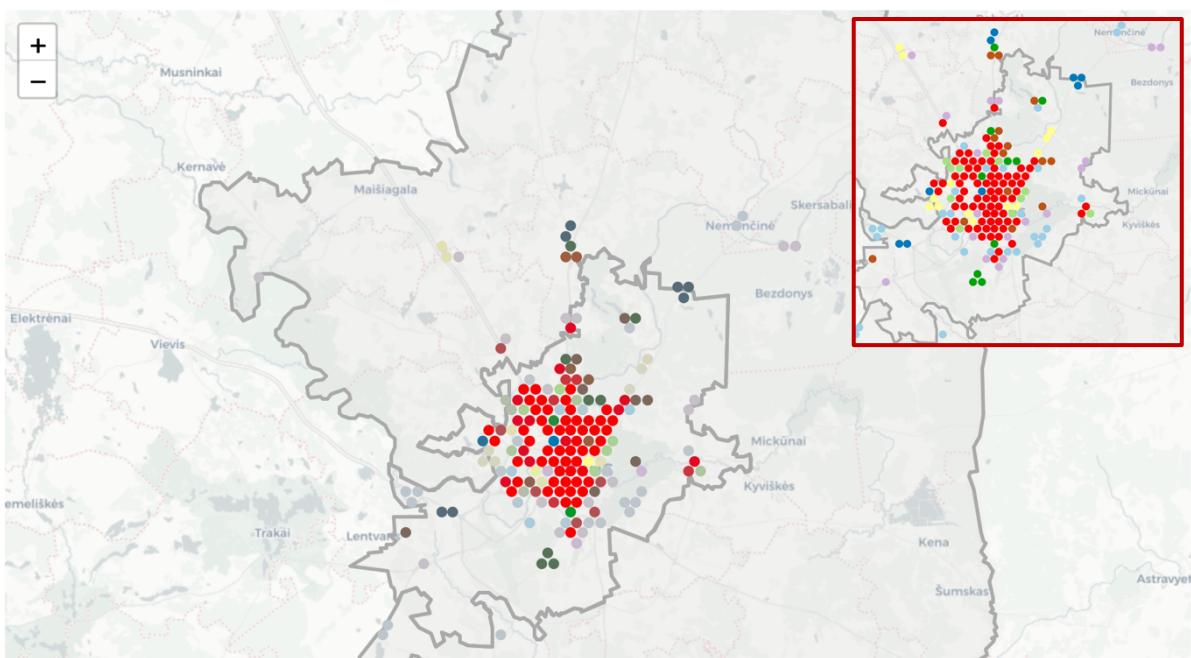


Figure 7. Clustering of neighbourhoods in Vilnius and its suburban area. Red – mixed cluster, yellow – Rest of the world, pale blue – Mediterranean, deep blue – Asian, green – Middle Eastern, brown – American, pale purple – East European, pale green – Latin American. Areas with fewer than 5 venues have their colours faded towards grey. Inset map shows clusters without colour fading, emphasizing hexagons with few venues.

## Discussion

### Quality of clustering

As indicated by Figures 1 and 2, the clustering for London and Paris is reasonably robust. That can't be said about Stockholm and Vilnius (Figure 3), and any conclusions for these two cities should be taken with caution. This is mainly due to the fact that these two cities are much smaller and contain fewer venues. They also attract fewer tourists, and thus lower population and lower tourism traffic perhaps reduce the amount of data contributed to the Foursquare API.

## City demographics and the overall view of the results

Despite the limitations of the data sources on the available venues, and some clear deficiencies of the manual remapping of categories from the Foursquare API results to, arguably, arbitrary “cultural categories”, there is a clear picture emerging that allows one to look into the degree of cultural mixing in different cities. The four chosen cities are very unequal in their size and demographics (Table 7).

*Table 7. Comparing cultural clusters to demographics of the cities. Sources of demographic data: [https://en.wikipedia.org/wiki/Greater\\_London](https://en.wikipedia.org/wiki/Greater_London) (London); <http://worldpopulationreview.com/world-cities/stockholm-population/> (Stockholm; <https://www.metropolegrandparis.fr/fr/une-metropole-mondiale-142> and [Working Together for Local Integration of Migrants and Refugees in Paris](#) (OECD Publishing, 2018) (Paris); <http://worldpopulationreview.com/world-cities/vilnius-population/> (Vilnius)*

Metropolitan area	Total population (million)	Population density (per sq km)	Residents identifying with non-dominant ethnicity	Size of the Mixed cluster	Size of cluster representing dominant culture
London	8.8	4,761 <sup>b</sup>	40 %	978	133
Stockholm	2.2	4,800 <sup>a</sup>	27 %	363	61
Paris	7.2	8,598 <sup>b</sup>	24 % <sup>d</sup>	544	211
Vilnius	0.6	1,392 <sup>a</sup>	37 % <sup>c</sup>	77	19

*Table footnotes: <sup>a</sup>Population density for Stockholm and Vilnius is given for the core municipality; <sup>b</sup>Population density for London and Paris is given over the entire metropolitan area; <sup>c</sup>Ethnic identification of Vilnius residents as not Lithuanians includes a large number of Poles (16.5%), Russians (12%) and Belorussians and Ukrainians (data not available), as these communities lived in Vilnius for centuries. They are classified as East Europeans for the purposes of this analysis anyway; <sup>d</sup>ethnic composition for Paris is inferred from the number of naturalised citizens and number of foreign citizens residing in the area.*

Yet, in each case the largest cluster of food cultures is the “Mixed” one. The exact composition of each mix is different in different cities, but they all show a strong presence of Asian and Mediterranean cuisines (Tables 3-6), which as we may recall, are consistently in the top 3 most popular cuisines in each city. They account for 58% of the “mix” in Vilnius, 76% in Stockholm, 50% in Paris and 45% in London. The numbers for Vilnius and Stockholm are somewhat overestimated, as there are additional small clusters called “Rest of the world”, which could arguably be folded into the “Mixed” cluster. Unsurprisingly, the further away from the city centre, the fewer venues are available from the Foursquare API (colours in Figures 4-7 are faded towards the edges of each city). The mixed cluster in London, Stockholm and Vilnius dominates the city centre (Figures 4, 6, 7), while the centre of Paris is dominated by the French cuisine – a clear outlier in this dataset (Figure 5).

## What does the data tell about each city?

### *London*

The western boroughs of London show a strong presence of South Asian cuisine. This is nearly certainly due to the large migration of people from the Indian subcontinent to Britain. What is interesting is that there is a clear signal in the restaurant data, indicating that perhaps these populations did indeed settle (and perhaps continue to settle) in close proximity to their compatriot migrants. North London and pockets of West London show a perceptible signal of Middle Eastern cuisine, pointing to a similar pattern of migration, while the Mediterranean cluster is spread quite evenly, and I would speculate that this has more to do with Londoners’ tastes than with migration patterns – few people don’t like a pizza!

### *Paris*

As discussed above, the centre of Paris is dominated by French cuisine. Unlike in the other three cities, the “home” food is the most popular option. The curious concentration in the city centre

though makes me speculate that perhaps tourism is partly at play. French cuisine is advertised world over, and it is not inconceivable that tourists expect to find it in the city. One would need to investigate which customers frequent these venues, but at this point, one can at least state that the Parisians are fond of their own food to a degree unmatched by other cities with respect to their “own” food. Yet Parisians still enjoy food from the rest of the world, with a universally shared penchant for Asian and Mediterranean cuisines. There is only a slightly discernible culinary segregation with a few pockets of the Asian food cluster in the outer suburbs, but like in other cities, they sport fewer venues (faded colours) (Figure 5). This is surprising, as Paris does have large immigrant communities. Perhaps the density of population forces more of the mixing?

### *Stockholm and Vilnius*

Unlike the other cities, patterns for Stockholm are much more difficult to discern because of the sparser venue data and, as a consequence, less robust clustering. The only conclusion I venture drawing is that even a smaller city in Europe sports a varied food culture, not dissimilar to other cities in Europe in its favourites. Unlike other cities, Vilnius has a low proportion of population with origins from outside Eastern Europe, so it is difficult to draw cultural conclusions from the available data.

## Conclusion

My original questions pertained to the theories of culture. None of the cities can be described as monocultural – that much is clear. In each case, the Mixed cluster is the largest and contains most venues. Unfortunately, the Foursquare API data does not offer a “Fusion” category for food, hence my last question on the degree of transculturality (in food) must remain unanswered. This leaves the remaining discussion to focus on deciding between multiculturality (i.e., parallel existences) and interculturality (shared existences). It appears that both Paris and London are somewhere between these two concepts. They each contain a large cluster with a dominant cuisine, which is also geographically compact (the South Asian cluster in London and the French cluster in Paris), suggesting that there is a degree of parallel existence in these areas. Yet, as has been mentioned a few times, the Mixed cluster is the largest, leading me to believe that Londonians and Parisians go beyond occasional sampling of an “exotic” food and are in fact embracing the world cuisines. The unique aspect about Parisians is that they do so while simultaneously cherishing their traditional foods, which cannot be said about Londoners or residents of Stockholm or Vilnius for that matter. Is Nigel Slater (Slater, 2015) correct in stating that Britain’s food culture is the most diverse? He may have a point, as the Mixed cluster in London is slightly more diverse than in other cities.

## References

- Benessaieh, A. (2010): ‘Multiculturalism, Interculturality, Transculturality’, in: Benessaieh, A. (ed.): *Transcultural Americas/Amériques transculturelles*. Ottawa: Ottawa University Press, pages 11-38.
- Caliński, T., & Harabasz, J. (1974). “A Dendrite Method for Cluster Analysis”. *Communications in Statistics-theory and Methods* 3: 1-27. doi:10.1080/03610927408827101
- Slater, N. (2015) “Let’s eat together: how immigration made British food great.” The Guardian website: <https://www.theguardian.com/global/2015/may/24/lets-eat-together-cooking-immigration-britain-food>. Accessed on 22 Feb 2020
- Welsch, W. (1999) “Transculturality - The Puzzling Form of Cultures Today”, in Featherstone, Mike; Scott Lash (eds.): *Spaces of Culture: City, Nation, World*. London: Sage, pages 194-213.

Working Together for Local Integration of Migrants and Refugees in Paris. OECD Publishing, Paris, 2018.

## Appendix A

Mapping top-level categories from the Foursquare API to a “Cultural” category.

Foursquare Category	Cultural category
Afghan Restaurant	Middle Eastern
African Restaurant	African
American Restaurant	American
Asian Restaurant	Asian
Australian Restaurant	Other
Austrian Restaurant	Central European
BBQ Joint	American
Bangladeshi Restaurant	South Asian
Belgian Restaurant	Central European
Cajun / Creole Restaurant	Caribbean
Caribbean Restaurant	Caribbean
Caucasian Restaurant	Middle Eastern
Creperie	French
Czech Restaurant	East European
Dutch Restaurant	Central European
Eastern European Restaurant	East European
English Restaurant	British and Irish
Falafel Restaurant	Middle Eastern
Fish & Chips Shop	British and Irish
Fondue Restaurant	French
French Restaurant	French
Gastropub	British and Irish
German Restaurant	Central European
Greek Restaurant	Mediterranean
Halal Restaurant	Middle Eastern
Hawaiian Restaurant	American
Hungarian Restaurant	Central European
Indian Restaurant	South Asian
Irish Pub	British and Irish
Italian Restaurant	Mediterranean
Jewish Restaurant	Middle Eastern
Kebab Restaurant	Middle Eastern
Latin American Restaurant	Latin American

Mediterranean Restaurant	Mediterranean
Mexican Restaurant	Latin American
Middle Eastern Restaurant	Middle Eastern
Modern European Restaurant	Other
Pakistani Restaurant	South Asian
Pizza Place	Mediterranean
Polish Restaurant	East European
Portuguese Restaurant	Mediterranean
Poutine Place	American
Russian Restaurant	East European
Scandinavian Restaurant	Scandinavian
Scottish Restaurant	British and Irish
Slovak Restaurant	East European
Southern / Soul Food Restaurant	American
Spanish Restaurant	Mediterranean
Sri Lankan Restaurant	South Asian
Steakhouse	Other
Swiss Restaurant	Central European
Turkish Restaurant	Middle Eastern
Ukrainian Restaurant	East European
Wings Joint	American
Pub	British and Irish