

Capstone Project – The Battle of Neighborhoods Restaurants at Brussels Capital Region Region

**Final Project for
Applied Data Science Capstone by IBM @ Coursera**

Michael Fan

Introduction

- Restaurants are local business. It primarily depends on
 - Population and demographics
 - Economic condition
- Expats are people who live in foreign country.
 - Question: Will there be more restaurants resembling to the originating country if there is larger population in the area?
- Auderghem is a municipality at Brussels Capital Region. It is reported to have majority of Japanese expatriates living in Brussels.
 - Question: Are there more Japanese restaurants in Auderghem than any other municipalities in Brussels Capital Region?

Data

- Data largely defines the scope of this work, from which analysis can be performed to draw conclusion.
- Sources of Data in this work:
 - Foursquare API: information of local venues, e.g. name, location, category, and others.
 - Google Map API: A good source to obtain address from longitudes and latitudes, and vice versa
 - Others from internet: Brussels municipality information and GeoJson municipality boundary information for creating maps.

Data – Foursquare API

- Abundant local venue information can be queried upon.
- Restaurant's unique venue ID, name, location, categories are used for this work.
- Venue category ID database:
 - Provided by Foursquare to query for venues of given category ID
 - Hierarchical database of json document available for query
 - Multiple category ID's can be supplied for each query.
 - However, single category ID may provide more consistent result, which is used in this work.
- Limited results per query, at least for free or low cost query.

Data – Brussels Municipality Information

- Municipality name, post codes, population, and others:
 - https://en.wikipedia.org/wiki/List_of_municipalities_of_the_Brussels-Capital_Region
- Municipality boundary for maps:
 - Link: <http://opendatastore.brussels/dataset/region-bruxelloise>
 - Data: <http://geoservices-urbis.irisnet.be/geoserver/Urbis/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=Urbis:Mu&outputFormat=JSON&srsName=EPSG:4326&maxFeatures=100>

Data – Google Map API

- Foursquare API provides detailed local business information but address information is not in consistent format for big data analysis.
- Google Map API provides formatted address from longitude-latitude pair. It is quite consistent, though not perfect
- To determine the municipality of a restaurant venue:
 - Longitude and latitude are queried from Foursquare.
 - Formatted address queried from Google Map API from longitude and latitude
 - Post code from formatted address is used to determine municipality of the given venue.
 - A consistent approach for a large datasets

Data Collection and Preparation

- Python programming is used to
 - Scrape data from internet
 - Query from database API
 - Collect data from sources
 - Prepare data for analysis
- Tools were developed to utilize Foursquare venue category ID database to connect research question for query such as:
 - What are venue category ID's that include certain keyword?
 - For a given category, what are its subcategories? What are their venue category ID's?

Prepared Restaurant Data

```
df_restaurants_brussels.head()
```

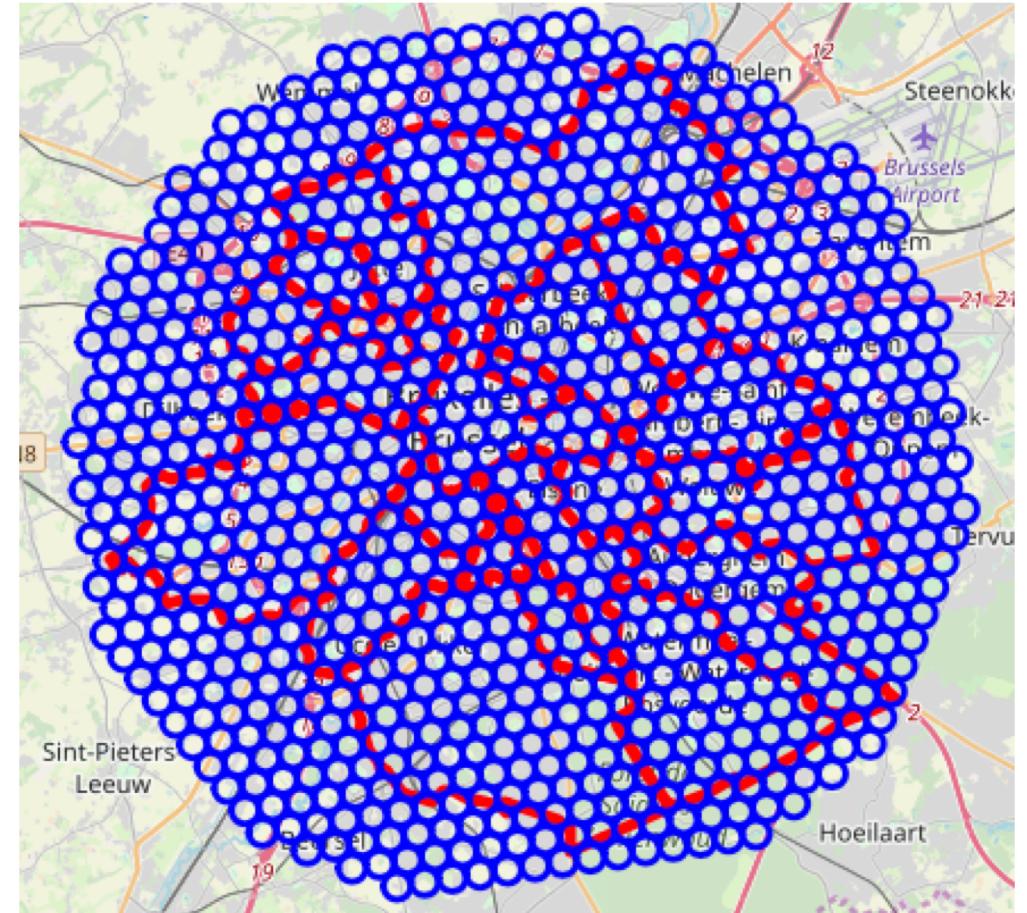
	vid	name	latitude	longitude	postcode	municipality	venueCategory2
47	5718fb6b498e930b29d77485	glacier lamy	50.775501	4.354569	1180	Uccle	Restaurant
49	56abd8c6498eaae8293f874f	À table 🍔🍜	50.774065	4.362790	1180	Uccle	Fast Food Restaurant
50	4f5fd2c3e4b00cb4dfcaf272	Burger	50.776336	4.372664	1180	Uccle	Burger Joint
61	533be9b3498e7c582a2e2d22	Pile poele	50.786713	4.417077	1170	Watermael-Boitsfort	Sandwich Place
74	4def7a001fc7e81f6da505c4	L' Entre-Pots (Uccle)	50.780449	4.325775	1180	Uccle	Italian Restaurant

Methodology

- Categorical statistical data analysis
 - Categories: venue category, municipality
- Forms:
 - Question and Answer: A research question followed by analysis and discussion

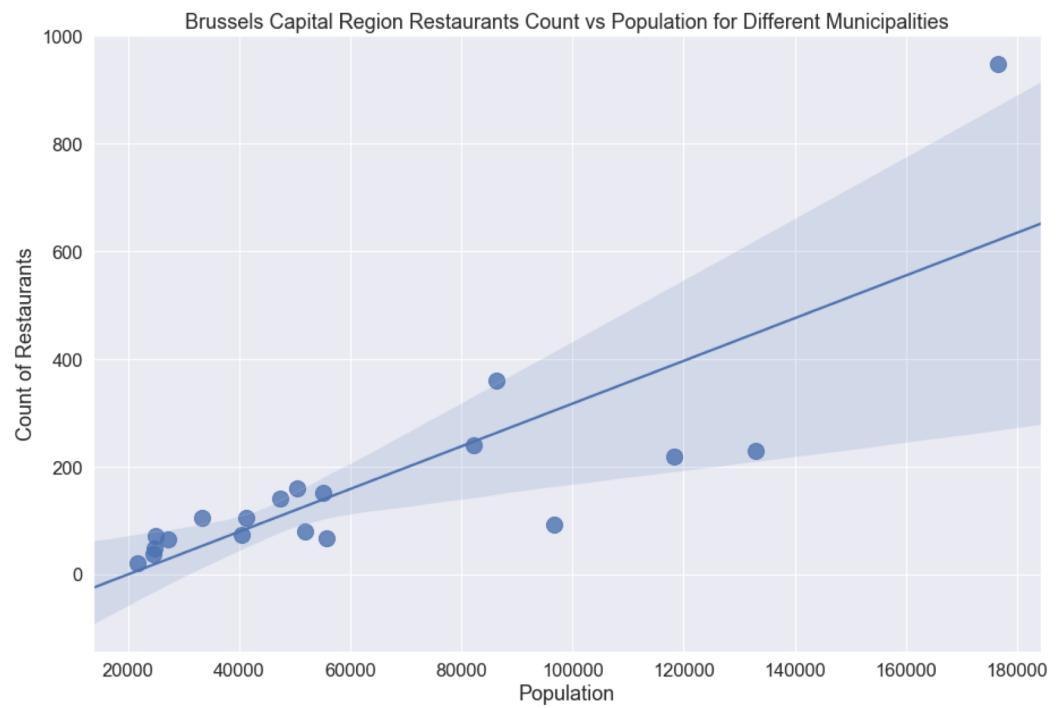
Brussels Capital Region Map

- Municipality boundary in red.
- The region is divided into small blue blocks.
- API query is performed for each block. This is required because limited results per query.



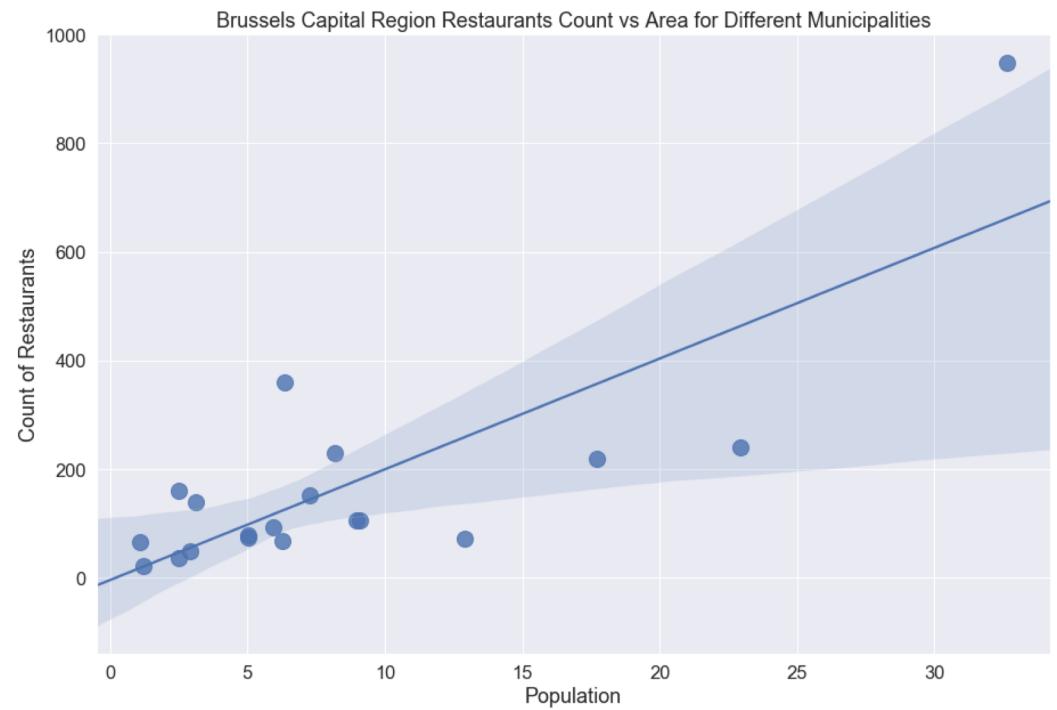
Restaurants vs Population

- Number of restaurants vs municipality population.
- More restaurants are in municipality with larger population.
- Some municipality show below- or above- trend restaurant counts



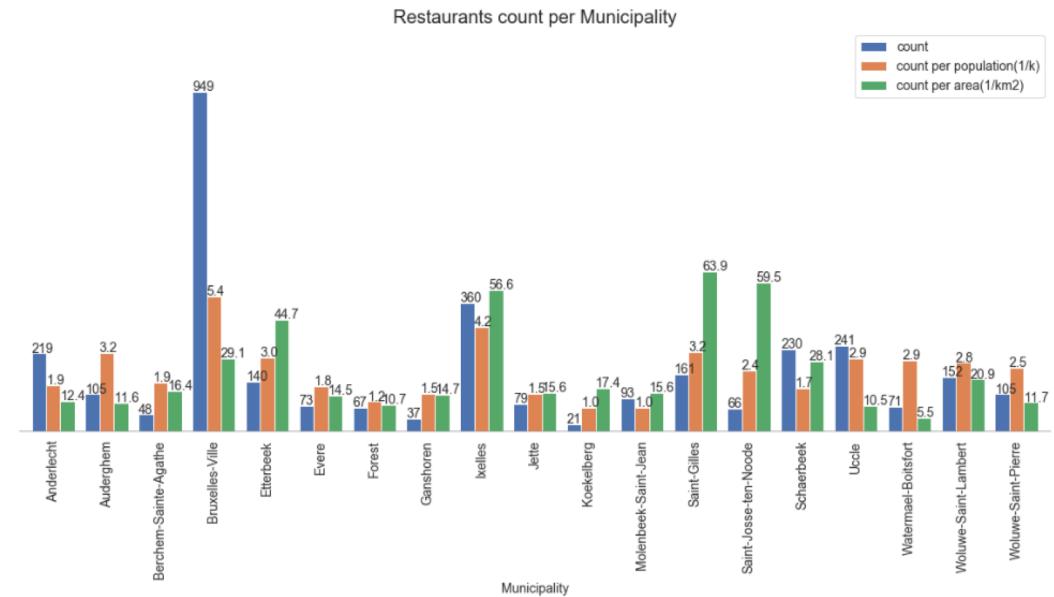
Restaurants vs Area

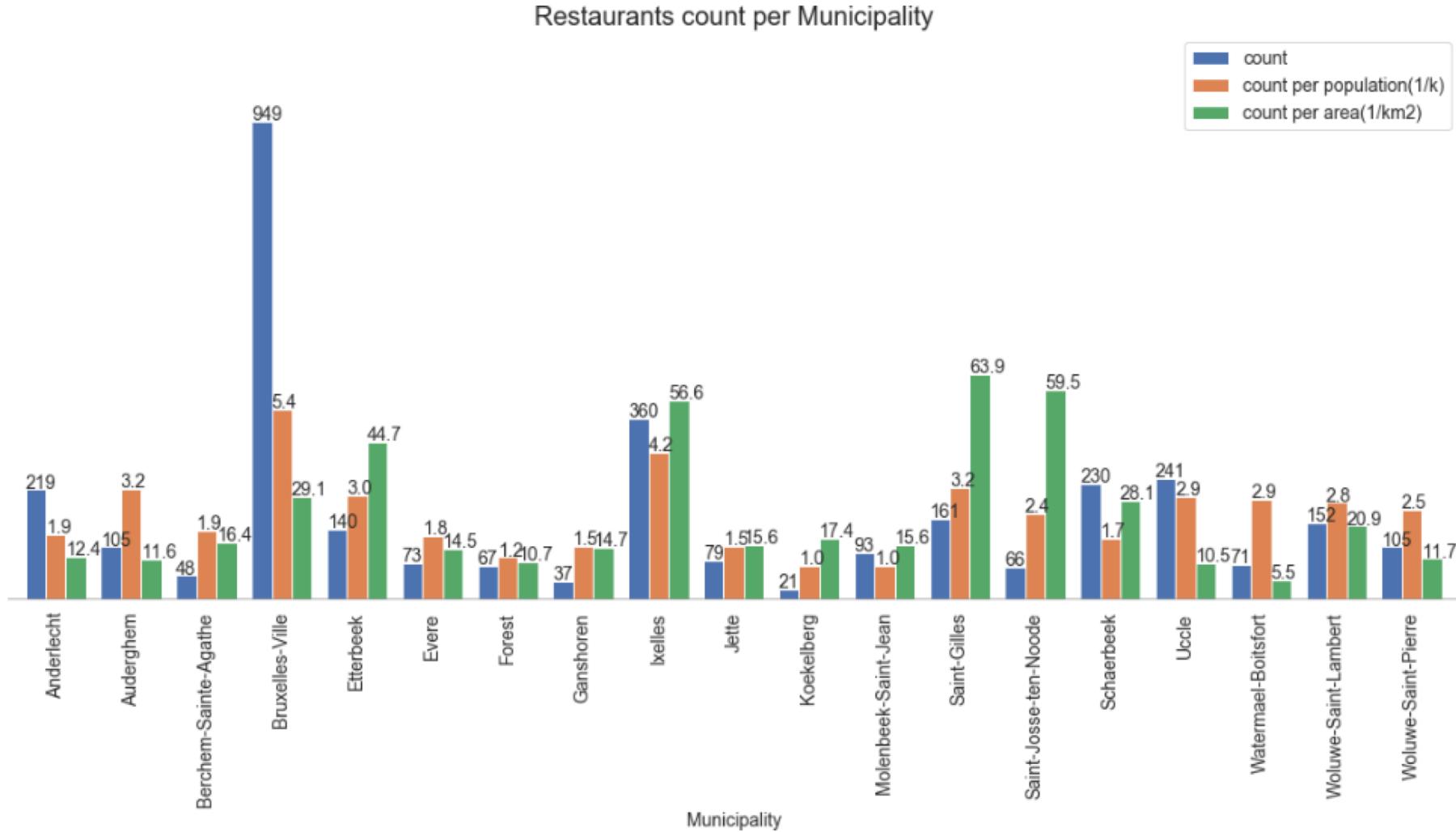
- Number of restaurants vs municipality area.
- More restaurants are in municipality with larger area.
- Some municipality show below- or above- trend restaurant counts



Restaurant count per Municipality

- Below-trend count: Koekelberg and Molenbeek-Saint-Jean
- Above-trend count: Bruxelles-Ville and Ixelles
- Possible reason:
 - Income level per capita.
 - Over-crowded business
 - Under-served area



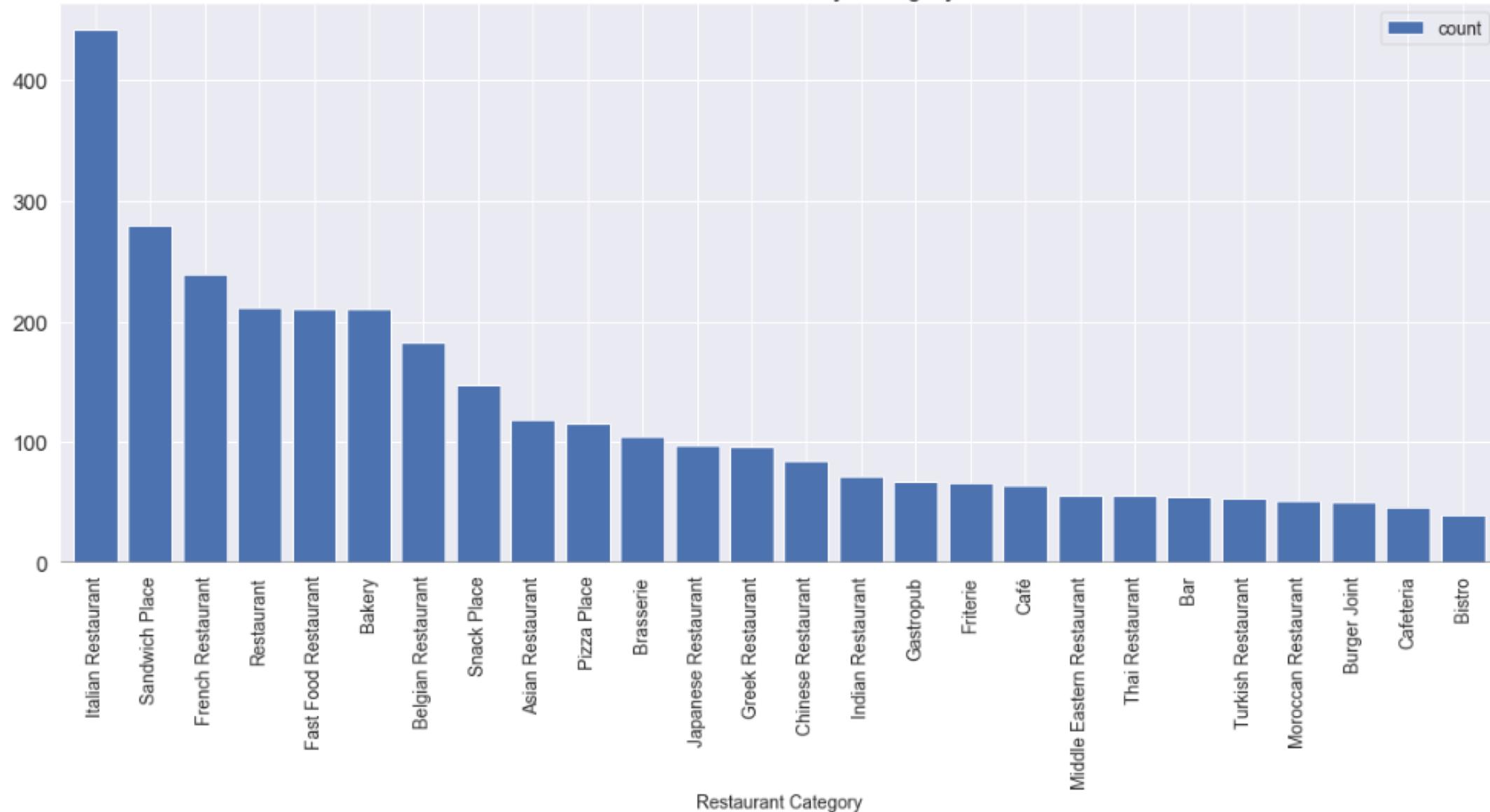


Below-trend count: Koekelberg and Molenbeek-Saint-Jean

Above-trend count: Bruxelles-Ville and Ixelles

Possible reason: Income level per capita, over-crowded business, under-served area

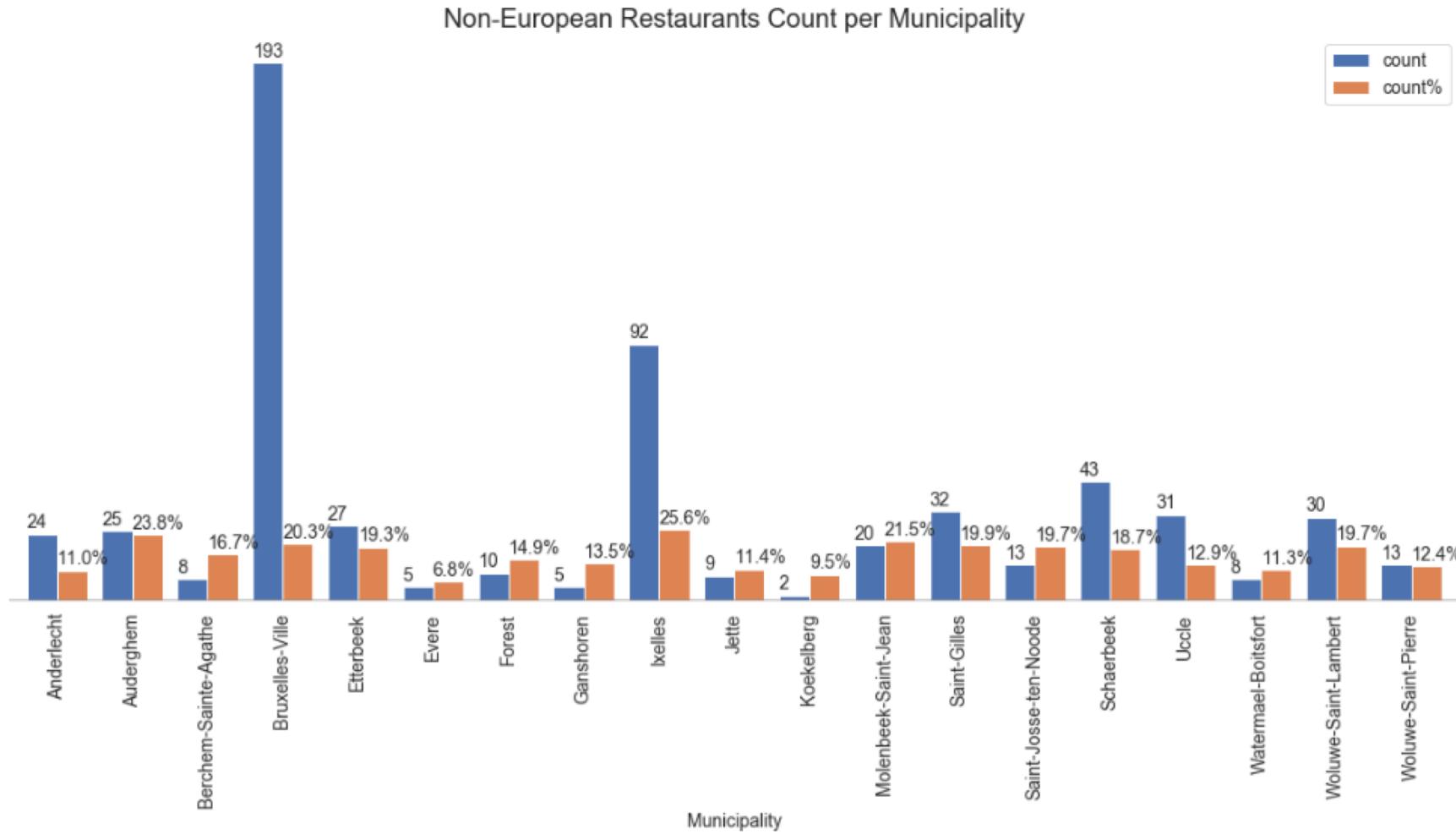
Restaurants count By Category



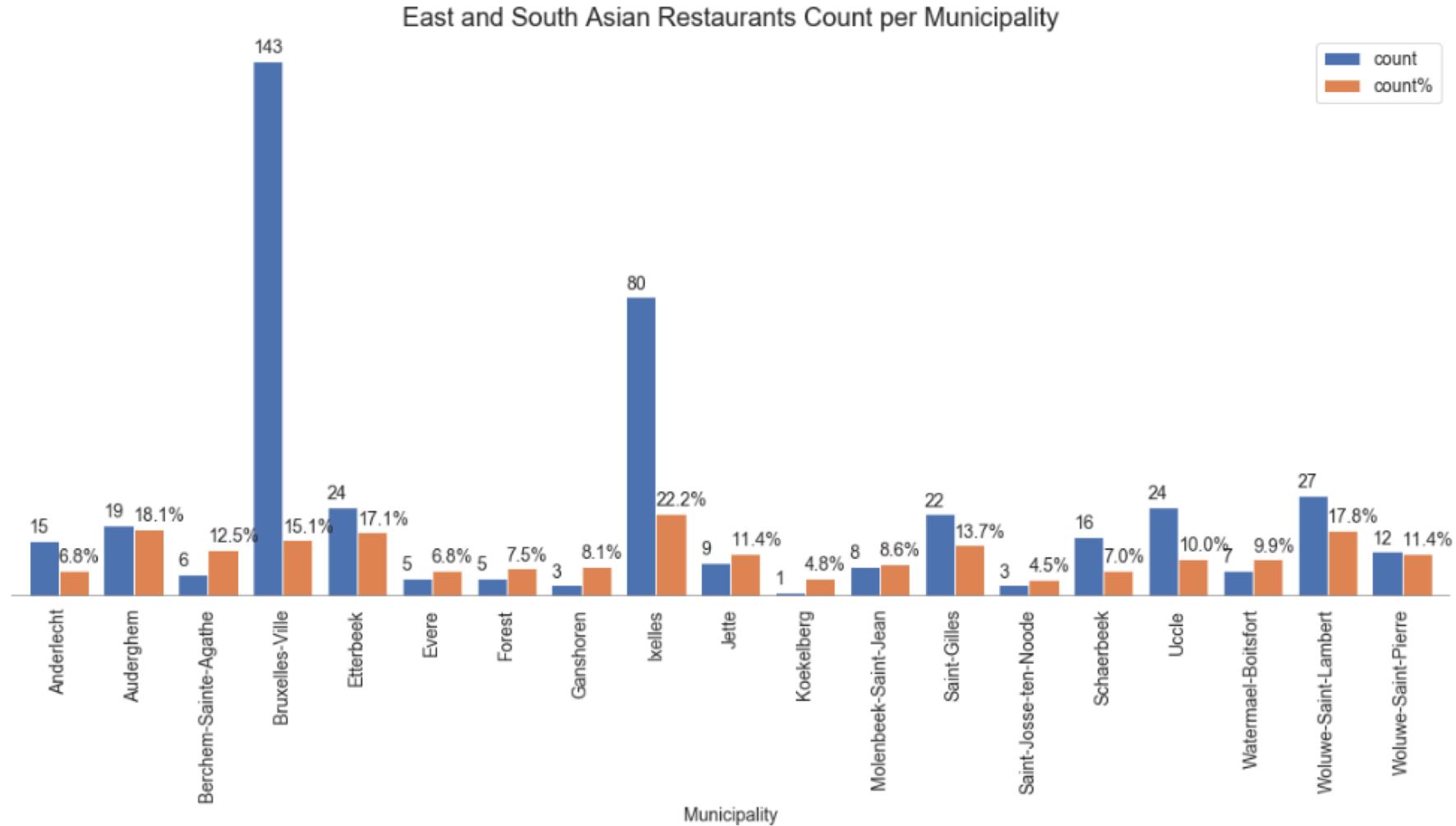
Most popular: Italian restaurants

Top Restaurant Category Per Municipality

	1	2	3	4	5
Municipality	Bakery	Italian Restaurant	Snack Place	Restaurant	Sandwich Place
Anderlecht	Bakery	Italian Restaurant	Snack Place	Restaurant	Sandwich Place
Auderghem	Italian Restaurant	Bakery	Sandwich Place	Asian Restaurant	Fast Food Restaurant
Berchem-Sainte-Agathe	Snack Place	Bakery	Greek Restaurant	Italian Restaurant	Japanese Restaurant
Bruxelles-Ville	Italian Restaurant	Sandwich Place	Belgian Restaurant	French Restaurant	Restaurant
Etterbeek	Italian Restaurant	Sandwich Place	Bakery	Fast Food Restaurant	Chinese Restaurant
Evere	Friterie	Snack Place	French Restaurant	Sandwich Place	Bakery
Forest	Sandwich Place	Pizza Place	Italian Restaurant	Belgian Restaurant	Snack Place
Ganshoren	Italian Restaurant	Bakery	Belgian Restaurant	Friterie	Restaurant
Ixelles	Italian Restaurant	French Restaurant	Sandwich Place	Fast Food Restaurant	Japanese Restaurant
Jette	Bakery	Sandwich Place	Snack Place	Italian Restaurant	Belgian Restaurant
Koekelberg	French Restaurant	Fast Food Restaurant	Restaurant	Bakery	Italian Restaurant
Molenbeek-Saint-Jean	Italian Restaurant	Fast Food Restaurant	Bakery	Sandwich Place	Turkish Restaurant
Saint-Gilles	Fast Food Restaurant	Italian Restaurant	French Restaurant	Sandwich Place	Greek Restaurant
Saint-Josse-ten-Noode	Italian Restaurant	Fast Food Restaurant	Sandwich Place	Middle Eastern Restaurant	Snack Place
Schaerbeek	Fast Food Restaurant	Italian Restaurant	Bakery	Sandwich Place	Snack Place
Uccle	Bakery	Italian Restaurant	French Restaurant	Restaurant	Sandwich Place
Watermael-Boitsfort	French Restaurant	Restaurant	Italian Restaurant	Cafeteria	Sandwich Place
Woluwe-Saint-Lambert	Italian Restaurant	Sandwich Place	Bakery	Fast Food Restaurant	French Restaurant
Woluwe-Saint-Pierre	Bakery	French Restaurant	Italian Restaurant	Restaurant	Sandwich Place



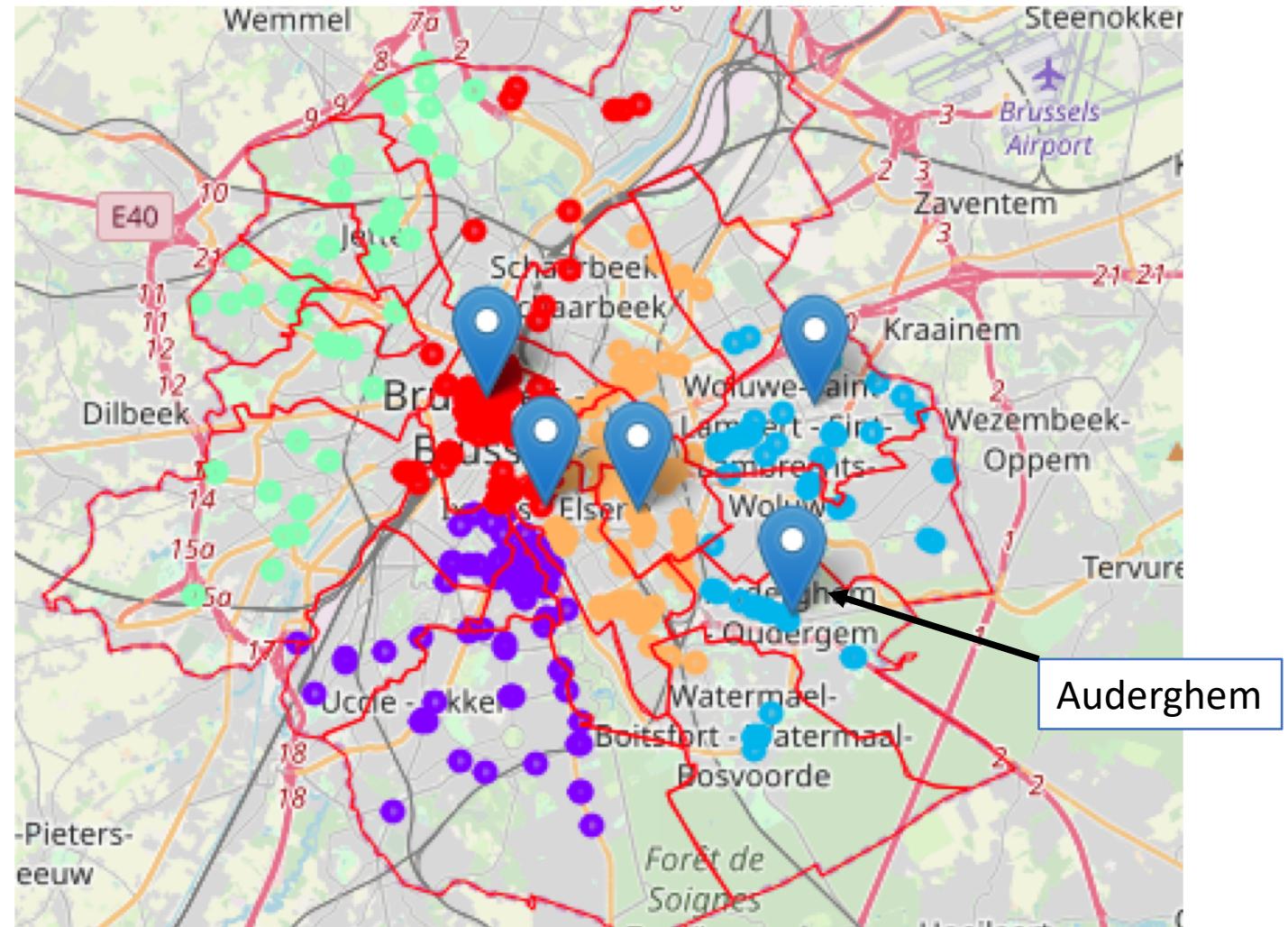
- Brussels-Ville has largest population, thus can be regarded as bench mark (20.3%)
- Auderghem, Ixelles and Molenbeek-Saint-Jean have more (>20.3%) non-european restaurants than other municipalities.



- Brussels-Ville as benchmark: 15.1%
- Auderghem, Ixelles, Etterbeek and Woluwe-Saint-Lambert above benchmark
- Molenbeek-Saint-Jean has many non-european restaurants but not as many east/southern asian restaurants.

Clustering of East/Southern Asian Restaurants

- 5 clusters shown in different colors
- Also shown are location marks for Brussels-Ville, Auderghem, Ixelles, Etterbeek and Woluwe-Saint-Lambert
- Indeed, there are East/Southern Asian Restaurants clustered near Auderghem



What about Japanese Restaurants?

- Some restaurant venues are categorized as Asian Restaurant as primary category, which means it is not specified as one of many subcategories of Asian Restaurants. Foursquare may have more venue details to provide such information. But it is beyond the scope of this work.
- After glancing the data when performing analysis, I found that restaurants with Japanese-related name are categorized as Seafood Restaurant. Foursquare may have more venue details to provide such information. But it is beyond the scope of this work.
- As it can be seen in the analysis performed, Auderghem does not have many restaurant venue information provided by Foursquare: 105 in total restaurant venues excluding those categories with fewer than 50 venues in Brussels Capital Region as whole, 25 for non-european restaurants, 19 for East Southern Asian Restaurants. The statistics population is not very large to draw much meaningful information, especially considering data are obscured by the previous two factors.

Conclusion

- Categorical descriptive statistical analysis has been performed to understand the association between population and restaurants, using data collected from Foursquare.com API, Google MAP API, and wikipedia for Brussels Capital Region. In general, there is obvious association between number of restaurants and population.
- Other factors to affect count of restaurant for a region may be due to economic condition such as income level.
- Data indeed imply Auderghem may have larger population for expatriates from East/Southern Asia. However, not enough information to associate it with larger Japanese expatriates population than other municipalities in Brussels Capital Region.