Total population and the most popular venues in Zurich city, Switzerland

by Zhanna Bayrychenko on December 22, 2019

Project as part of the Applied Data Science Capstone course on Coursera (Data Science Professional Certificate)

Aim of this project

Use location data (Foursquare API) to explore a geographical location by comparing neighborhoods or cities.

Location : Zurich city

- The largest city in Switzerland
- 430’000 inhabitants in the municipality
- 12 boroughs/districts
- 34 neighborhoods
- Global city

Business problem

- What are the most popular venues in the neighborhoods of Zurich city ?

- Are there differences between more and less populated neighborhoods ?

- Where is the best development potential for stakeholders ?

Data

- Names of the boroughs and neighborhoods of Zurich city, their total population and population density (https://wikipedia.org)

- Latitude and longitude of each neighborhood

(https://www.geonames.org)

- Geographical data (.json file) with the neighborhoods of Zurich city

(https://opendata.swiss)

- Foursquare API for type, location of the venues

Dataframe

	Borough	Neighborhood	Latitude	Longitude	Population (2018)	Population density / km²
0	Kreis 1 – Altstadt	Rathaus	47.37161	8.54501	3267	3194
1	Kreis 1 – Altstadt	Hochschulen	47.36601	8.54594	664	3194
2	Kreis 1 – Altstadt	Lindenhof	47.37188	8.54036	990	3194
3	Kreis 1 – Altstadt	City	47.37269	8.53576	829	3194
4	Kreis 2	Wollishofen	47.34010	8.53134	18923	3151

Methodology

- Generate choropleth maps of Zurich city with its neighborhoods by using the .json file and showing the total population / pop. density

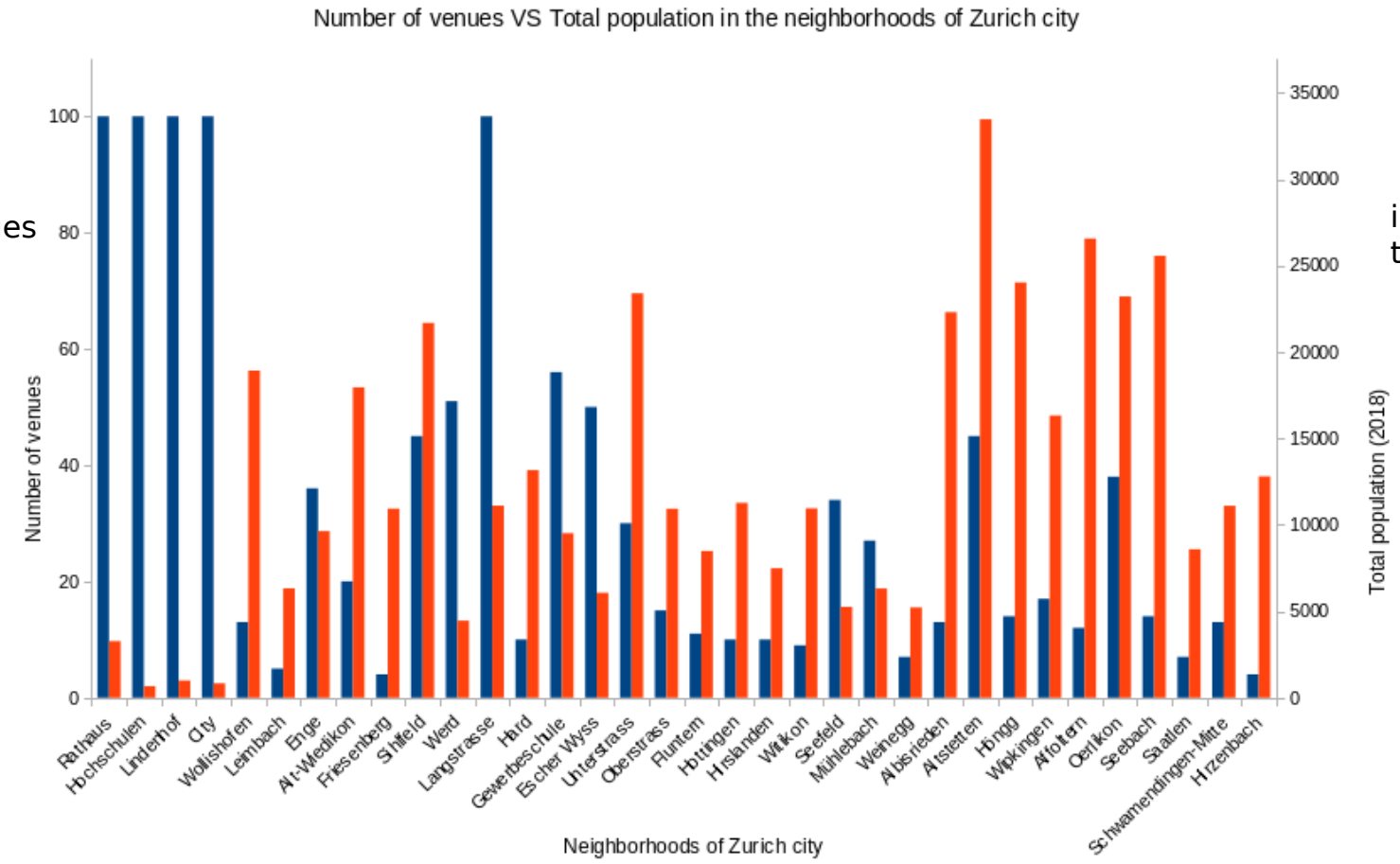
- Explore the most popular venues with Foursquare API

- Cluster neighborhoods based on their most common venues with *k*-means model

→ Create choropleth maps showing the population and the clusters of each neighborhood

Results

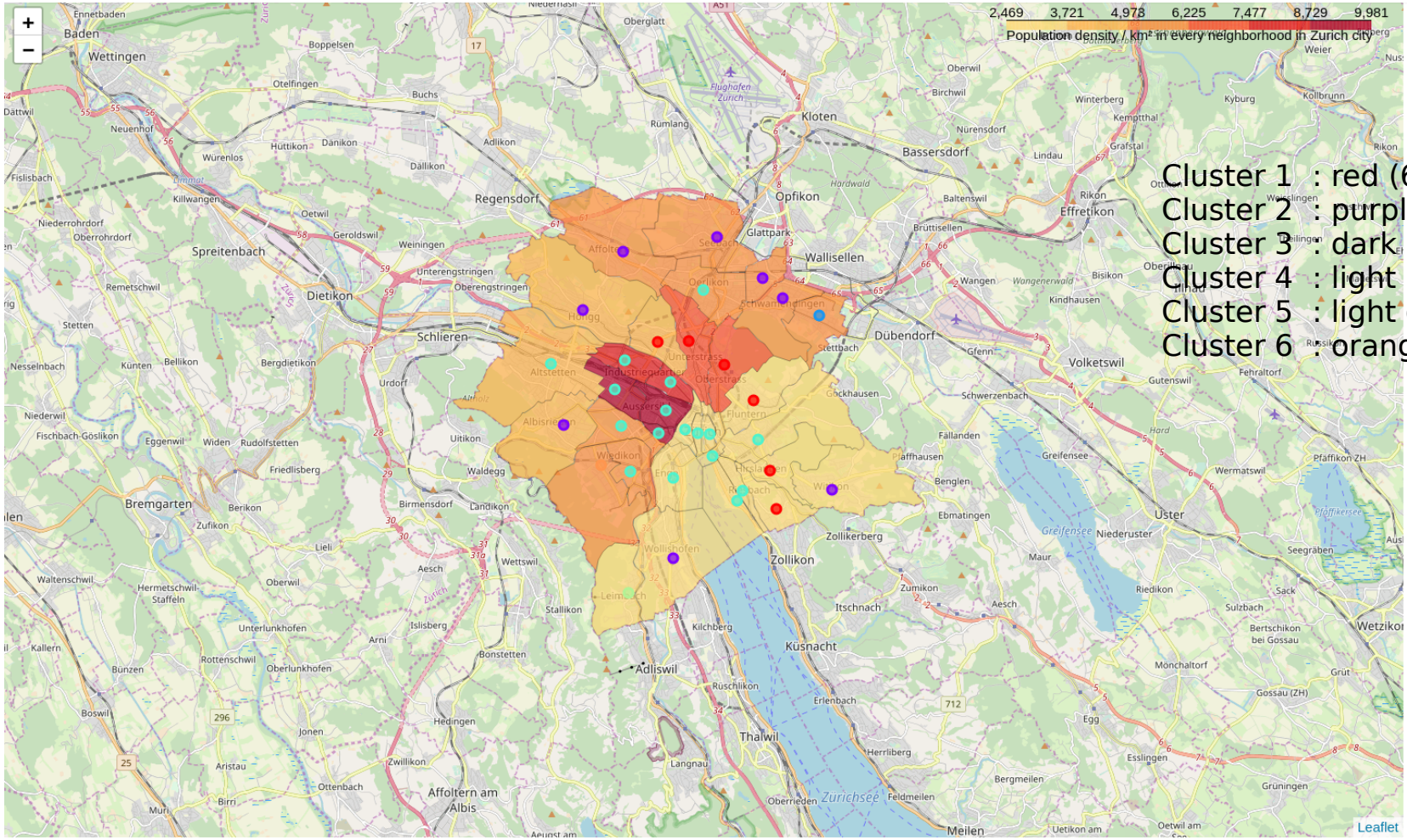
in blue :
number of venues



in red :
total population

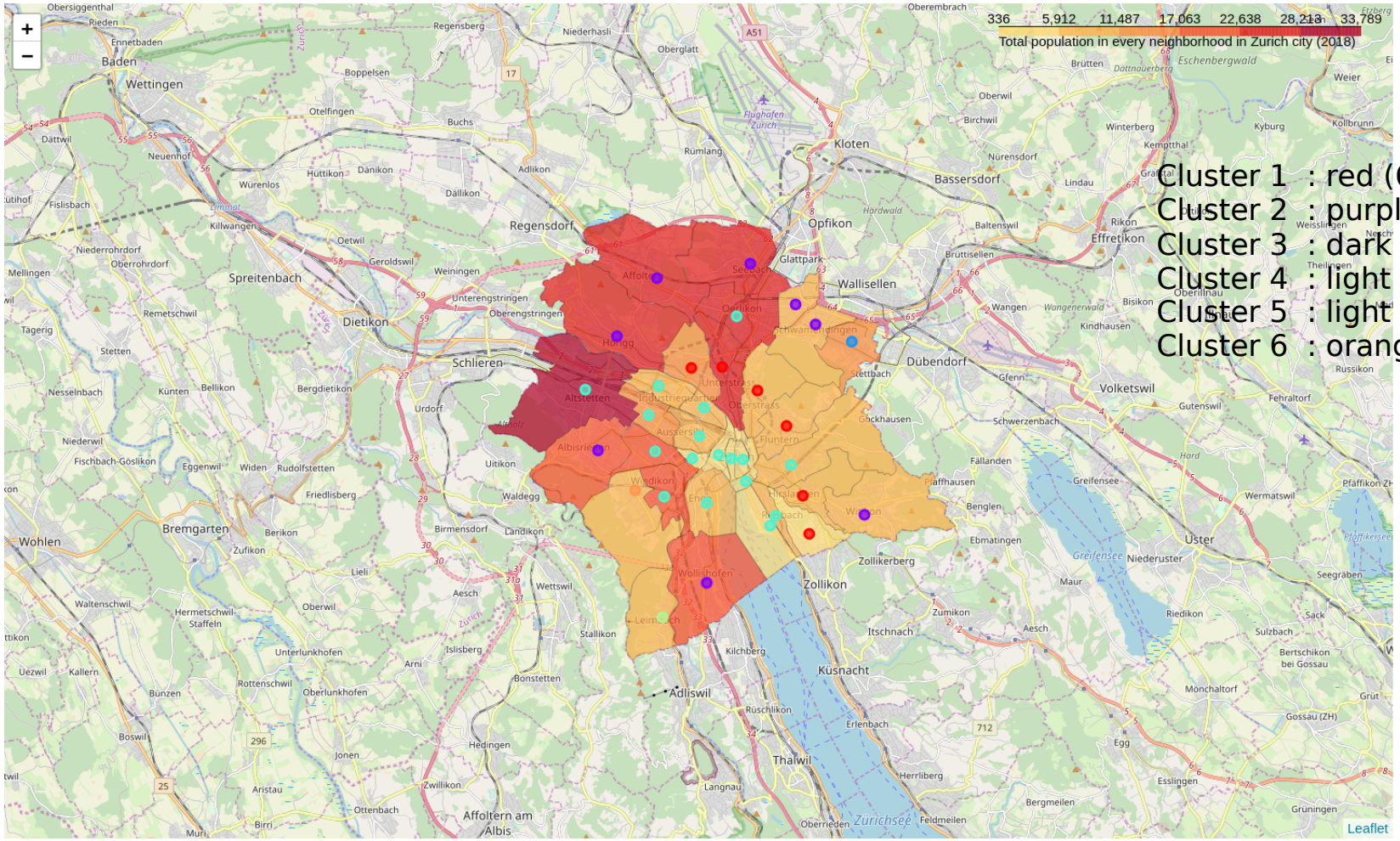
Results : choropleth maps (1/2)

Population density/km² in each neighborhood with the 6 clusters



Results : choropleth maps (2/2)

Total population in each neighborhood with the 6 clusters



Cluster 1 : red (6)
Cluster 2 : purple (8)
Cluster 3 : dark blue (1)
Cluster 4 : light blue (28)
Cluster 5 : light green (1)
Cluster 6 : orange (1)

Discussion : type of neighborhoods

- The neighborhoods with the biggest number of venues have the lowest number of total population and vice versa

→ commercial VS residential neighborhoods

Discussion : Clusters

- Clusters 2, 3, 5 and 6 : peripherical residential neighborhoods

Cluster 2 : Old residential (more developped)

Cluster 3, 5 and 6 : New residential, less developped

Popular venues : bus/tram stations, restaurants, supermarkets/stores, yoga studios, event spaces

High development potential, but should be daytime activities and family-friendly.

Discussion : Clusters

- Cluster 4 : commercial neighborhoods, City Center

Popular venues : cafés, restaurants, bars, shops, etc. Well developped nightlife.

The development potential is lower, as existing venues are already numerous and various, so the competition would be high.

Some neighborhoods are still interesting and some businesses have more potential than others.

Discussion : Clusters

- Cluster 1 : Wealthy Residential-Center, buffer zone

Popular venues : restaurants, stores, hotels, bakeries. Less developped nightlife.

Some neighborhoods have good development potential, for various types of venues. Can also be fancy restaurants, as these neighborhoods seem to be appreciated by wealthy people.

Conclusion

This analysis gave us information on the development potential of the neighborhoods, which can be very useful to different stakeholders.

This project demonstrated again the importance and the interest of data and how this data can bring us answers to all kind of questions.