Amsterdam Pizzeria

Introduction

Imagine that you want to open a new pizzeria in Amsterdam. The big question is: What is a good location?
This notebook is trying to answer this question by examining the amount of pizzerias and population per neighborhood.

Data

Three different data sources will be used in this project:

- The geodata about the neighborhoods is retrieved from https://data.overheid.nl/dataset/mea3qdtnvln9ca. This data is a so called shapefile. To read and modify this data, the geopandas package is used.

- The population data about the neighborhoods is retrieved from https://data.amsterdam.nl/datasets/DMknRs8hEH-CtA/bevolking-wijken/. The file with expected population between 2020 and 2050 is used. This contains data from 2020, 2025, 2030, 2040 and 2050 on neighborhood level.

- The amount of pizzerias and food places in Amsterdam is retrieved by using the Foursquare API. At first, only the data of all the pizzeria's is retrieved by specifying it in the query. After this, all food places in the area of the centroid of the neighborhood are retrieved.

Methodology

The starting point was retrieving and transforming the data from the shapefile about the defined areas in Amsterdam. This resulted in 386 defined areas and the following map:

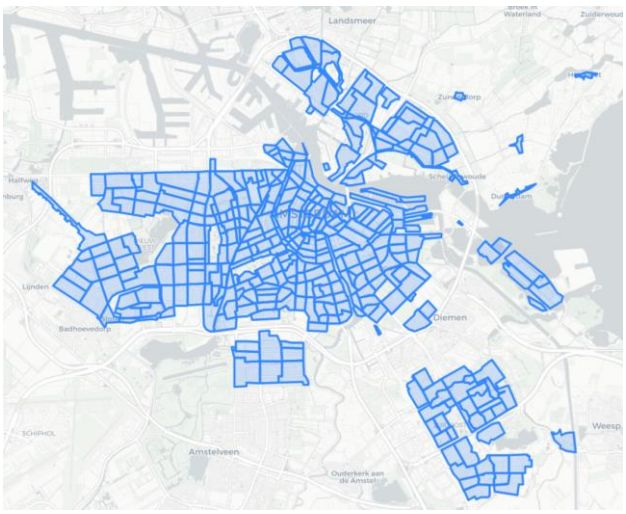

Figure 1: Defined areas in Amsterdam

After this, population data is retrieved and examined. In this file, there are only 101 defined neighborhoods with the following summary statistics:

	2020	2025	2030	2040	2050
count	91.000000	91.000000	91.000000	91.000000	91.000000
mean	9296.901099	9755.604396	10071.912088	10728.692308	11163.670330
std	5571.873208	5818.566328	5962.490286	6435.329907	6935.461238
min	0.000000	1121.000000	1098.000000	1063.000000	1047.000000
25%	5187.500000	5566.500000	5489.000000	5906.500000	5886.500000
50%	8537.000000	8831.000000	8964.000000	9387.000000	9592.000000
75%	12576.000000	13829.000000	14181.500000	15011.000000	15360.000000
max	29788.000000	29999.000000	31944.000000	31619.000000	35390.000000

Figure 2: Summary statistics of population data

So in 2020, the most populated neighborhood had 29788 inhabitants, while the least one had zero inhabitants.

Given that the area do not match, the geodata had to be aggregated to neighborhood level. After this, the two dataframes are merged and give the following choropleth map of the population in 2020:

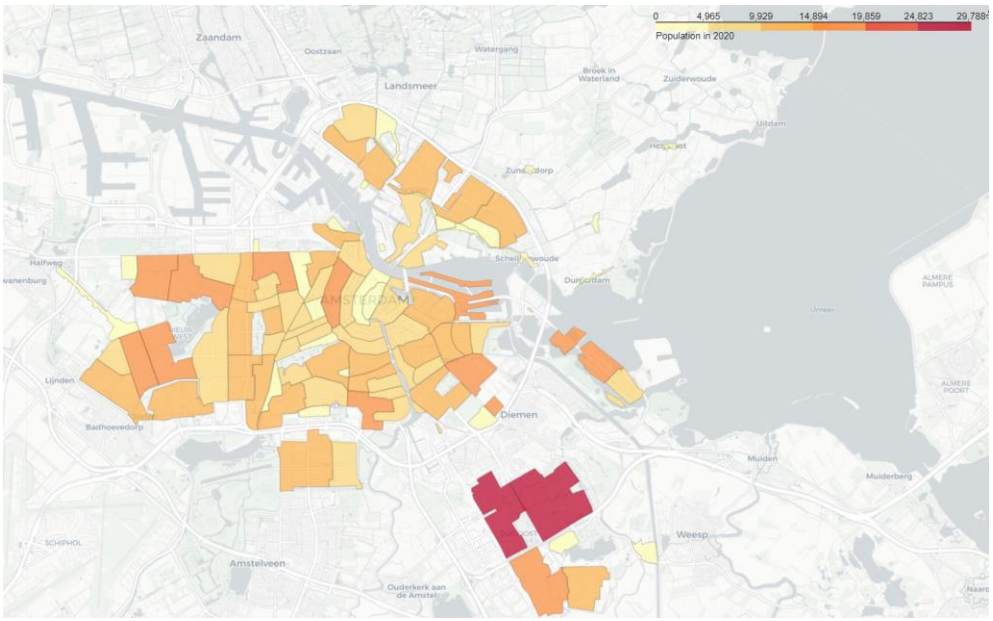

Figure 3: Choropleth map of Amsterdam population 2020

After this, I moved on to the Foursquare data. At first, I was only interested in the existing pizza places in Amsterdam, so this was defined in the query. There were only 41 pizza places in Amsterdam:

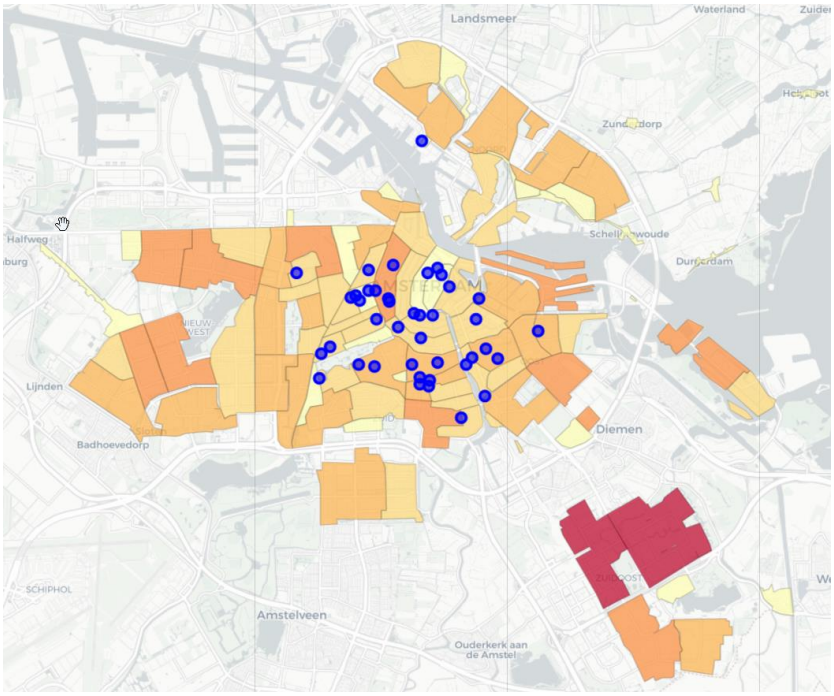

Figure 4: Location of pizza places in Amsterdam according to Foursquare

So there are a lot of areas without a pizza place. Therefore, further analysis is necessary to find the optimal neighborhood to start a new pizzeria. This is done by retrieving all the food places from Foursquare API based on the centroid of the neighborhood and a 500 meter radius. This was the result:

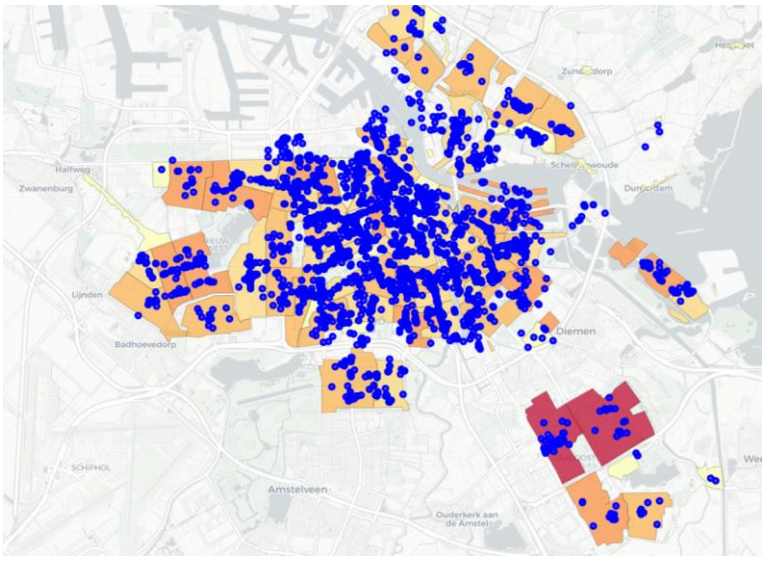

Figure 5: All food places in Amsterdam according to Foursquare

If we transform figure 5 to a choropleth map, this is the result:

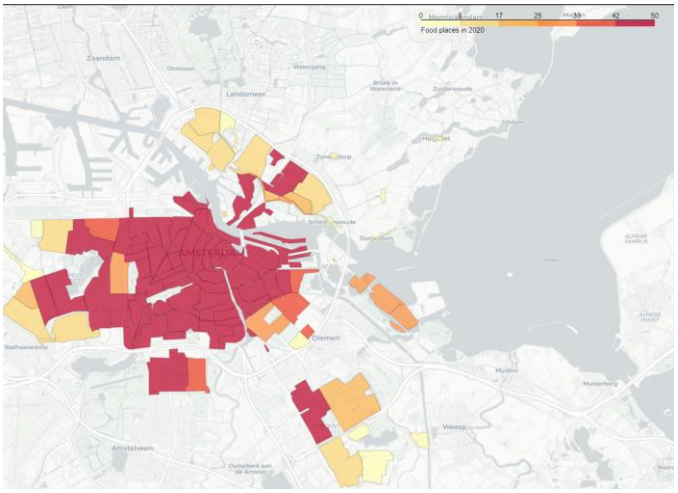

Figure 6: Amount of food places per neighborhood

Results & Discussion

To determine the best location, the population of 2020 and 2025 is used. These numbers are divided by the amount of food places according to Foursquare. The next figure shows a summary of the ratio of 2020:

count	91.000000
mean	361.174931
std	357.542519
min	0.000000
25%	143.950204
50%	236.354167
75%	408.318182
max	1518.500000

Figure 7: Summary statistics of ratio 2020

So there is one food place for every 361 inhabitants in Amsterdam. The top 5 least covered neighborhoods are the following ones:

	Name	Venue	2020	Ratio 2020
87	Nellestein	2	3037	1518.500000
89	Gein	8	11327	1415.875000
67	Banne Buiksloot	11	14781	1343.727273
86	Bijlmer-Oost (E,G,K)	23	29788	1295.130435
72	Geuzenveld	13	16535	1271.923077

Figure 8: Top 5 least covered areas in 2020

If we do the same analysis based on population data of 2025, the results are the same:

	Name	Venue	2025	Ratio 2025
87	Nellestein	2	2863	1431.500000
89	Gein	8	11297	1412.125000
67	Banne Buiksloot	11	14787	1344.272727
86	Bijlmer-Oost (E,G,K)	23	29999	1304.304348
72	Geuzenveld	13	16407	1262.076923

Figure 9: Top 5 least covered areas in 2025

Therefore, the results seem consistent.

Conclusion

Based on the population per food place, the conclusion can be made that Nellestein, Gein, Banne Buiksloot, Bijlmer-Oost, and Geuzenveld are the best options to start a new pizzeria. This results seems to be consistent over the next 5 years. Furthermore, there are no pizza places in those areas yet:

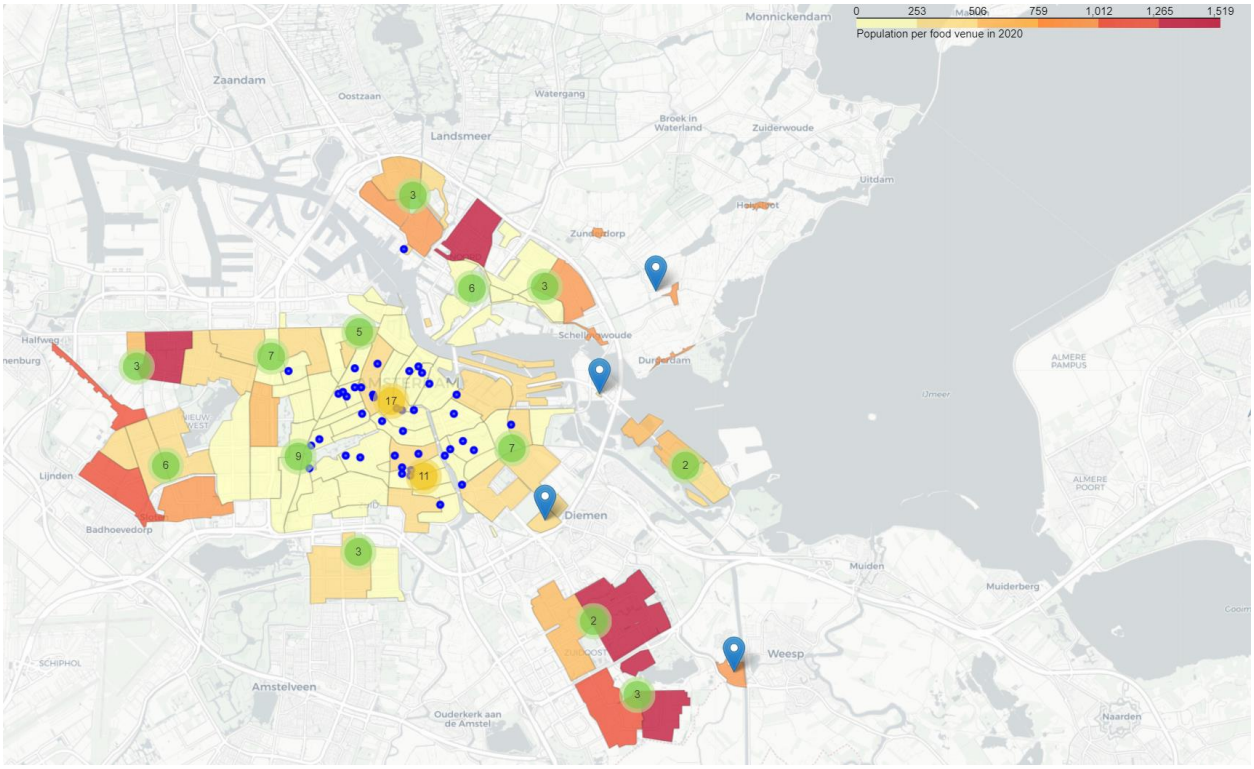

Figure 10: Population per food venue in 2020