

Applied Data Science Capstone Project

Opening a new bakery/cafe business in Hamburg

...

Mai 17, 2020

Overview

- The project consists of supporting the stakeholders of a small bakery franchise to extend their business in Hamburg
- By leveraging the Foursquare location data, Hamburg neighborhood will be investigated based on their similarities, specifically in terms of availability of bakeries and cafes.
- In addition, the population distribution per Hamburg neighborhood, as well as, its population density will support the selection of the most promising region for the new business.

Project objective:

Identify the most promising neighborhood to open a new bakery/cafe establishment

Project steps

Data extraction

Two sets of data were used in this project. The first set of data consists of information about Hamburg neighborhoods and their geographical location.

These data will be combined with the population density per Hamburg neighborhood and number of establishment extracted from Foursquare API

Data scraping

Python 3 and jupyter notebook will be used to scrap the data, process it and plot graphics

Data analysis and decision

Comparison between neighborhood population density and number of establishment will provide valuable information for stakeholders to decide where to open a new bakery/cafe business

Data extraction

(Hamburg geographical information)

| | Neighborhood | Postal Code | Latitude | Longitude |
|---|--------------|-------------|----------|-----------|
| 0 | Harburg | 21073 | 53.459 | 9.981 |
| 1 | Tonndorf | 22045 | 53.587 | 10.117 |
| 2 | Uhlenhorst | 22085 | 53.571 | 10.019 |
| 3 | Billstedt | 22115 | 53.528 | 10.147 |
| 4 | Langenhorn | 22417 | 53.667 | 10.034 |

Fig. 2: Hamburg's neighborhood data. Head of extracted dataframe



Fig 3. Neighborhood superimposed on top of Hamburg map

Data extraction

(Distribution of establishment per neighborhood)



| | Neighborhood | Postal Code | Latitude | Longitude | Bakery | Café | Total |
|----|-----------------|-------------|----------|-----------|--------|------|-------|
| 38 | Hoheluft-Ost | 20251 | 53.583 | 9.981 | 1.0 | 14.0 | 15 |
| 33 | Altona-Altstadt | 22767 | 53.55 | 9.935 | 6.0 | 6.0 | 12 |
| 9 | Neustadt | 20354 | 53.55 | 9.979 | 1.0 | 9.0 | 10 |
| 39 | Hoheluft-West | 20253 | 53.581 | 9.968 | 5.0 | 4.0 | 9 |
| 62 | Sternschanze | 20357 | 53.561 | 9.962 | 0.0 | 5.0 | 5 |

Fig 4. Number bakeries and cafes in Hamburg, extracted from Foursquare AIP.

| | Neighborhood | Area (km ²) | Population | Population/km ² |
|---|------------------|-------------------------|------------|----------------------------|
| 0 | Hamburg-Altstadt | 24 | 2350.000 | 979 |
| 1 | HafenCity | 22 | 4925.000 | 2239 |
| 2 | Neustadt | 23 | 12.762 | 5549 |
| 3 | St. Pauli | 25 | 22.097 | 8839 |
| 4 | St. Georg | 24 | 11.358 | 4733 |

Fig.5: Head of the data frame containing the population data per neighborhood

Results and discussion

The data bakeries and cafes information extracted from Foursquare API per Hamburg neighborhood were overlaid on the Hamburg map, as shown in Fig. 6. Each neighborhood was grouped according to the total quantity of establishments. Thus an interactive tool is provided for a easy visualization of establishments per neighborhood

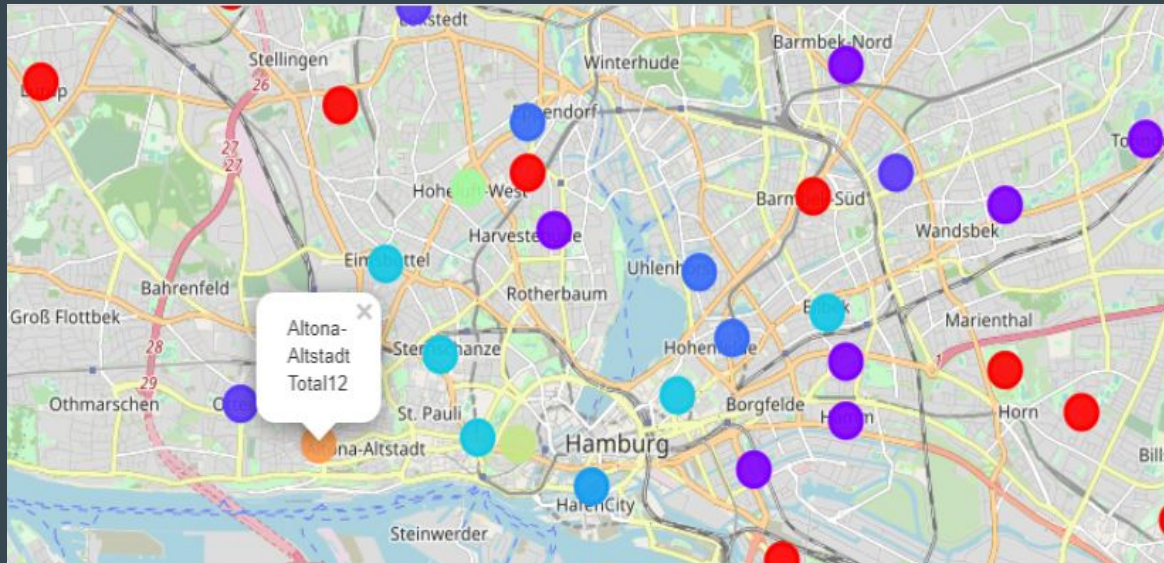


Fig 6. Number of establishment per Neighborhood superimposed on top of Hamburg map

Results and discussion

- Hoheluft-Ost, followed by Altona-Altstadt and Neustadt presented the higher concentration of establishment, with cafes contributing more to the total number of establishments (Fig 7).
- The number of establishments on its own does not provide enough information to decide on the most promising neighborhood to open a new business since it does not consider the population density per neighborhood
- It is expected that the higher the population per km² in a determined neighborhood the higher is the amount of potential clients to the new business
- A graphic of the population density per neighborhood was created to support the assessment of the most promising region to open a new establishment, as shown in Fig.8.

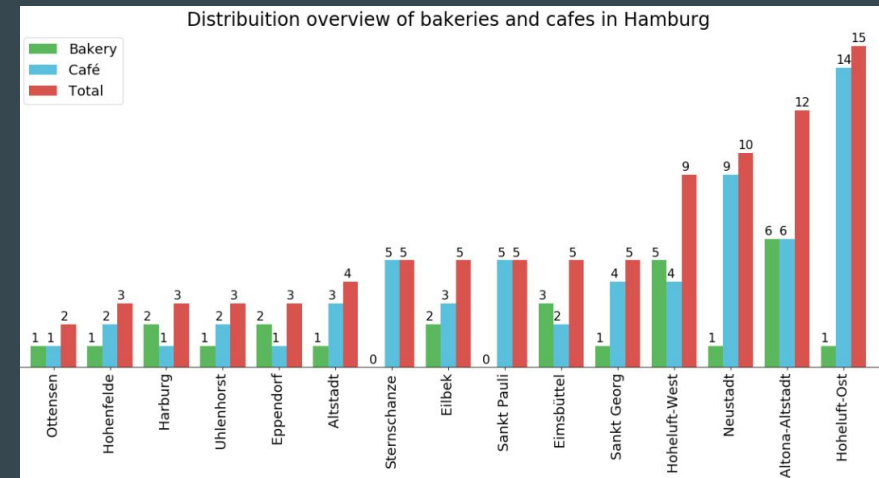


Fig. 7: 15 neighborhoods in Hamburg with more number of establishments

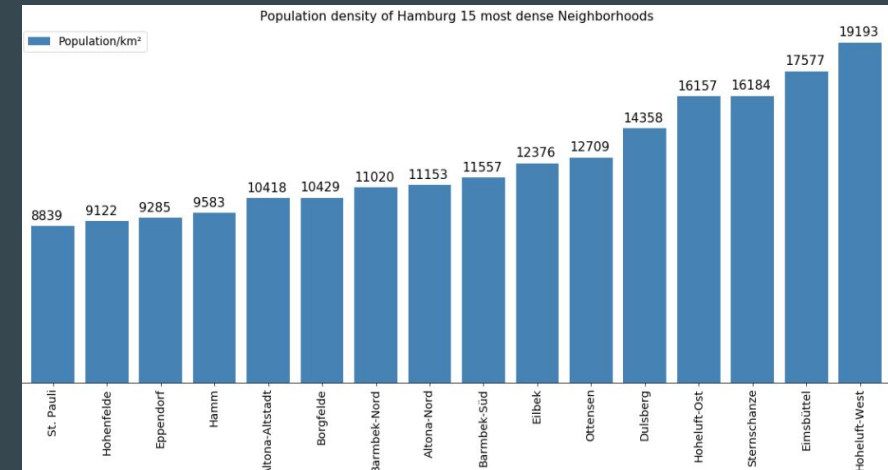


Fig. 8: 15 neighborhoods in Hamburg with higher density (population/km²)

Results and discussion

- Neighborhoods such as Barmbek-Süd and Barmbek-Nor are between the densest neighborhood in Hamburg, however, they did not feature between the top 15 with most establishments in Hamburg (Fig. 7)
- Such a simple assessment suggests that these neighborhoods might present a great opportunity for the extension of a bakery business
- Eppendorf one of the most preferred neighborhoods in the Hamburg feature between the 15 with higher density and presents only 3 establishments with only 1 cafe. Therefore, it might also accommodate more establishments

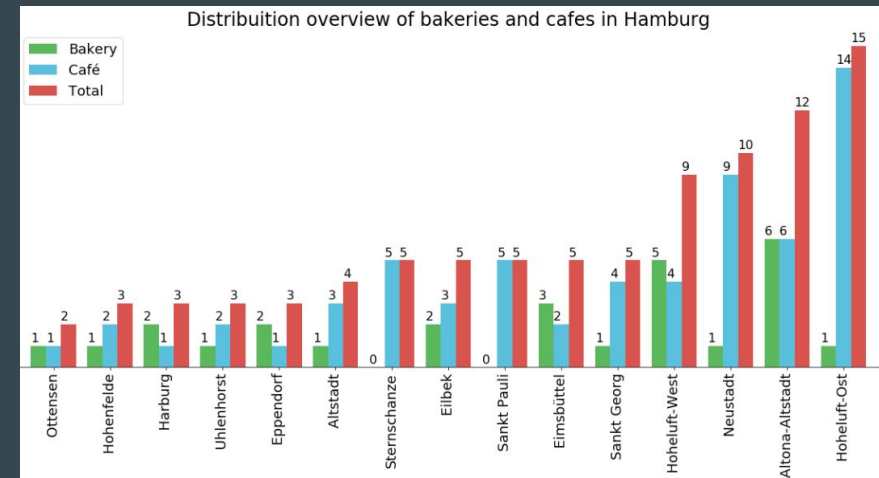


Fig. 7: 15 neighborhoods in Hamburg with more number of establishments

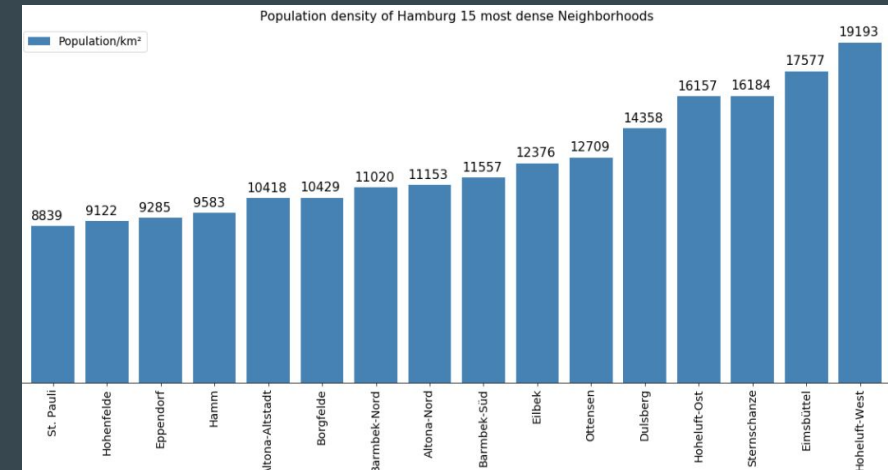


Fig. 8: 15 neighborhoods in Hamburg with higher density (population/km2)

Results and discussion

- This project successfully explored the Hamburg neighborhood and classified them based on the total quantity of available bakeries and cafes
- By comparing this data with the population density of each neighborhood, regions with a higher number of potential clients can be identified and support the decision on where to open a new establishment.
- Such an approach was applied first to bakeries and cafes, however, the available code provides information for 100 different venues per neighborhood.
- The analysis showed that Barmbek-Süd and Barmbek-Nor are between the densest neighborhood in Hamburg, however, they did not feature between the top 15 with most establishments in Hamburg, suggesting that these neighborhoods might present a great opportunity for the extension of a bakery business.
- Eppendorf one of the most preferred neighborhoods in the Hamburg feature between the 15 with higher density and presents only 3 establishments with only 1 cafe, which might indicate that it can accommodate more establishments.
- It is important to mention, however, that such analysis did not consider the price involved on open a new establishment on each neighborhood, which is affected by different variables not investigated within this project.