

# Capstone Project - The Battle of the Neighborhoods

Giang Nguyen

## 1. Introduction

### 1.1. Overview

Düsseldorf is the capital and second-largest city of the most populous German federal state of North Rhine-Westphalia after Cologne, as well as the seventh-largest city in Germany with a population of 617,280. Mercer's 2012 Quality of Living survey ranked Düsseldorf the sixth most livable city in the world.

Düsseldorf is an international business and financial centre, renowned for its fashion and trade fairs and is headquarters to one Fortune Global 500 and two DAX companies. Messe Düsseldorf organises nearly one fifth of premier trade shows.

The Rhine Metropolis is one of the most diverse areas in terms of culinary diversity. Düsseldorf, with the third largest Japanese community in Europe, not only provides a wide range of culinary cuisine but also has a solid foundation of Authentic Asian food in the city. Düsseldorf's exceptional culinary cuisine has been recognized and visited by the Worldwide leading travel guide of Lonely Planet. Along with a broad range of diverse cultural cuisine, Düsseldorf is also home to various Michelin starred restaurants that are world renowned

### 1.2. Opportunities and challenges

Our stakeholder are really into opening a restaurant in Duesseldorf because of the current and potential opportunities:

- this is one of the biggest and the most liveable cities in Germany.
- it is also the central of cuisine, and the duesseldorf's locals and suburban, as well as the visitors are keen on the diverse cultural cuisine here.
- the locals' income is also high. it is a favourable condition for establishing a business relating to service area as restaurant.

Of course, choosing a location for the restaurant is one of the stressful and controversial tasks, since there are a lot of challenges relating to

- the number of potential competitors
- the number of of alternatives
- the population
- the parking lot numbers

### 1.3. Objectives

This project conduct an analysis in order to find out the optimal place to build up a cuisine restaurant based on the mentioned attributes. In fact there are many determinants but because of the limitation of time and the course requirements, thus the project just focuses on the main ones.

## 2. Data

### 2.1. Data description

Based on criteria listed above the following data will be utilized in our analysis:

- the number of restaurants within the certain radius of each borough (Foresquare API)
- the number of parks within the certain radius of each borough (Foresquare API)
- the population density in each borough.

Data sources:

- the population and the population density of the borough.  
<https://www.citypopulation.de/de/germany/dusseldorf/admin/>  
Source: citypopulation.de
- the square of Duesseldorf's Districts  
<https://www.duesseldorf.de/bv/?L=0>  
Source: Official Website of Dusseldorf city
- the coordinates of the borough.  
<https://latitude.to/articles-by-country/de/germany/210179/dusseldorf>  
Source: latitude to

### 2.2 Data Preparation

Table 1. The Data Frame of the information of Duesseldorf's District

	Borough	Population	Population Density	Latitude	Longitude
0	Altstadt	2404	7603.008850	51.223999	6.770830
1	Angermund	6626	678.516775	51.324665	6.774664
2	Bilk	40651	4838.669355	51.201499	6.774497
3	Carlstadt	2285	7603.008850	51.220332	6.769330
4	Derendorf	20610	7603.008850	51.239666	6.788830
5	Düsseltal	28177	8899.580420	51.237832	6.801997
6	Eller	31701	2894.265403	51.199999	6.838330
7	Flehe	2726	4838.669355	51.188833	6.768164
8	Friedrichstadt	19883	4838.669355	51.207333	6.775164
9	Garath	19109	4470.661896	51.141333	6.890163

Table 2. The detail information of data frame

	Population	Population Density	Latitude	Longitude
count	39.000000	39.000000	39.000000	39.000000
mean	<a href="#">13138.153846</a>	3860.467043	51.220563	6.804561
std	10430.690926	2226.325238	0.042230	0.051867
min	1612.000000	678.516775	51.125999	6.709497
25%	5049.500000	2560.856519	51.200749	6.770497
50%	10598.000000	3415.128205	51.221332	6.788830
75%	19626.500000	4838.669355	51.247166	6.848413
max	40651.000000	8899.580420	51.324665	6.906496

### 3. Methodology and Analysis

After cleaning and preparing the data, let us identify the steps, that have to be performed in order to find the most optimal boroughs. Firstly, we will apply some basic exploratory analysis to our data. For that let's find the location of each borough on the map. Then we can visually inspect some values in our data with the help of bar charts. Secondly, we have the possibility to reduce the number features in data frame by replacing them with more reasonable data. Finally, we will perform cluster analysis to find the best cluster of boroughs with meaningful features.

#### 3.2 Cluster Analysis

To identify groups (clusters) with similar characteristics, the unsupervised learning method to our data, namely K-Means algorithm, was applied to our data.

But before that, to reduce dimensionality of the problem the columns "Population", "Number of restaurants" and "Number of beer restaurants" were removed. These three columns were replaced with two new ones, namely, "Number of restaurants per thousand people" and "Number of beer restaurants per thousand people".

### 4. Results and discussion

During the analysis, two clusters were defined. Two groups were clustered according to the Population Density. It is obvious, that the cluster with highest population density has the highest priority for us (Cluster 1).

Duesseldorf Pempelfort and Duesseldorf Flehe are the most attractive options in terms of low competitiveness (low number of current restaurants) and high customer number (high population density).

What could be done better?

Foursquare doesn't represent the full picture, since many venues are not on the list. For that reason, another maps could be utilized such as Google map or Openstreet map.

Boroughs have too complex geometry, thus defining the closest venues within the certain radius brings additional error to our analysis.

## **5. Conclusion**

To conclude, the basic data analysis was performed to identify the most optimal boroughs for the placement of the restaurant in the city of Duesseldorf. During the analysis, several important statistical features of the boroughs were explored and visualized. Furthermore, clustering helped to highlight the group of optimal areas. Finally, Duesseldorf Pempelfort and Duesseldorf Flehe were chosen as the most attractive options for the further analysis.