1. A description of the problem and a discussion of the background. (15 marks)

The client is a gastronomy company, serving high-quality cuisines on food trucks. Their business plan is uniquely designed to attract customers that are willing to pay for a relatively pricy meal but are unwilling or too busy to sit in a restaurant or bistro. They mainly operate in neighborhoods that are preferably far from other restaurants or gastronomical competitors. They pay particular attention to the uniqueness of their brand and advertise their company, basically by not advertising for it!

Since the start of the Covid-19 pandemic, the demand for their takeaway meals is soaring, particularly due to the home-office work style of their customers. This has convinced them to expand their business to other cities. Due to logistical reasons, their first choice is the city of Düsseldorf in Germany. The client's wish is to find out the neighborhoods in the city that would be more attractive for their business. More specifically, they would like to find out 2-3 neighborhoods that could fulfill the following criteria:

i) Neighborhoods that offer the highest number of potential customers.

ii) The customers are willing to purchase pricy meals from the food truck.

iii) The neighborhoods are preferably far from other restaurants or gastronomical businesses as potential competitors.

2. A description of the data and how it will be used to solve the problem.

Considering the criteria requested by the client, it would be necessary to collect granular data for different districts of the city of Düsseldorf with representative features like demographic aspects and purchasing power. In order to be able to provide a targeted insight for the client, the focus of data collection was geared toward zip codes in the city. That is because more data was available for the zip codes than for the boroughs.

Specifically, the following data was collected for each zip code:

- The land area of the zip code
- The Population of the zip code
- Average rental price per square meter in the zip code
- Spatial coordinates of the restaurants

Land area and population of the zip code will allow for a quantification of the number density of potential customers in the vicinity of the food truck. The data for these features were collected from this source: www.suche-postleitzahl.org

The average rental price in each zip code is a representative factor for the purchasing power or willingness of the potential customers to pay for a rather expensive meal from the food truck. This was collected from: https://www.miet-check.de/mietpreis-plz/

Spatial coordinates of the restaurants will allow for spotting the denser areas in terms of the potential competitors. The client is more interested in regions that are not highly populated with other restaurants. This data was collected using Foursquare API. It was noted that the zip code data provided by this API for each venue was not always correct. In order to access the true zip code for each restaurant, reverse geo-coding of the coordinates was performed using ArcGIS API.

The data will be used to create a scoring mechanism as well as a clustering algorithm using DBSCAN. It is expected that the regions within the city that are attractive for the business model of the client will more probably be labeled as outliers.