2. Data

2.1 Data needed

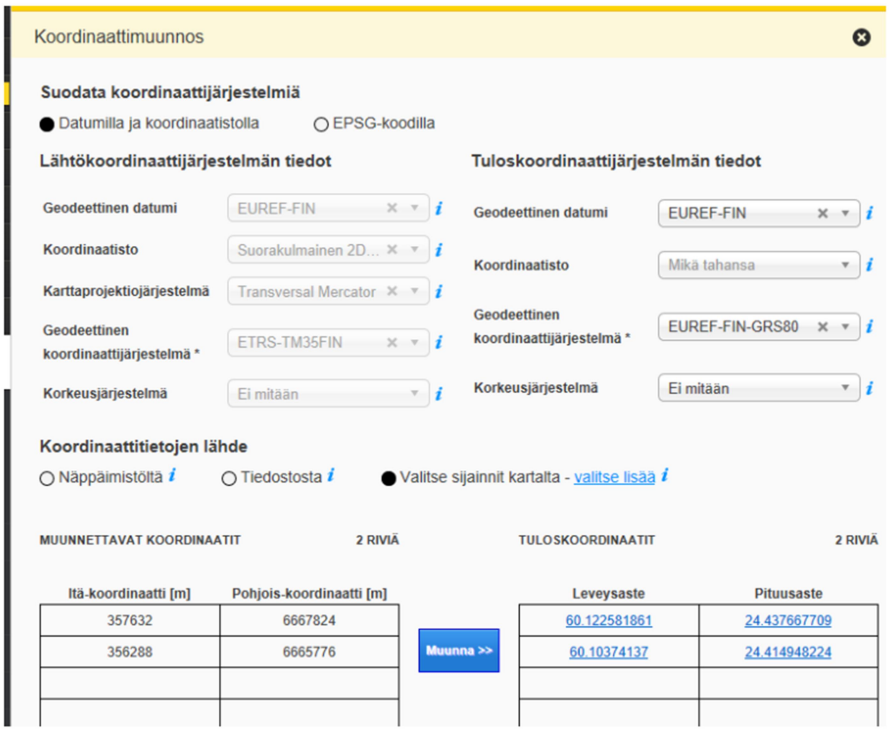
Based on definition of our problem, factors that will influence our decission are:

- number of existing kebab restaurants/pizzerias/fast food places or other restaurants in the neighborhood (any type of restaurant)
- number of people per kebab restaurants/pizzerias/fast food places or other restaurants in the neighborhood or borough
- number of competitors in the near vicinity, if any.
- number of shops around the kebab restaurant/pizzeria/fast food place/other restaurant
- number of bars around the kebab restaurant/pizzeria/fast food place/other restaurant.

2.2 Data sources

We get the following information from https://fi.wikipedia.org/wiki/Luettelo_Tampereen_tilastoalueista: neighbours, boroughs, planning area, the number of inhabitants in borough.

The location of the neigbours were mainly calculated using the Nominatim library of geopy.geocoders. However, it did not give all the results, and the missing results were varying, and in some cases the result was not accurate. Therefore, in addition to geopy, I also used the open data available in https://geodata.tampere.fi/geoserver/kiinteistot/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=kiinteistot:NIMISTO_MVIEW&outputFormat=json. The coordinates are not in the form of latitudes and longitudes, but in the format: ETRS-GK24 (EPSG:3878). Therefore I changed those coordintes using a converter on the page https://kartta.paikkatietoikkuna.fi/?lang=fi .



I also calculated the distances from neighbourhoods to the Tampere city center (Keskustori) by Google Maps, using either distance by car, walk or bike. At the moment there are big tram constructions on the way to Tampere center, and the easiest way to get to the center is either by walk or by bus. Some traffic is only allowed for public traffic or taxis. Using your own car mean longer routes to get to the center. Finally, I figured out Radius information for each neighbourhood according to the size of the neighbourhood. This is used as an information how far we are going to search places around the neighbouhood center. All the data mentioned above which was not available in https://fi.wikipedia.org/wiki/Luettelo_Tampereen_tilastoalueista I collected and stored into a local file.

The number of kebab restaurants/pizzerias/fast food places and other restaurants and their type and location in every neighborhood will be obtained using **Foursquare API.** We also use that same api to get the number of competitors, shops and bars in the near vicinity of the kebab restaurants/pizzerias/fast food places and other restaurants.

2.3 Data cleaning

First I read the data from https://fi.wikipedia.org/wiki/Luettelo_Tampereen_tilastoalueista , and put it into a pandas dataframe. Then I dropped the not needed column 'Kaupunginosan numero'. Then I renamed columns and borouhs into English for better readibility. The ’NumberOfPeople’ data includes a special charachter ’\xa0’ between numbers which had to be removed. This peculiar space like character was noticed when type conversion from object type to int type gave an error.

Next I read from my locally saved data/csv file the following information into a different dataframe

- Neighbourhood
- Latitude
- Longitude
- Distance from center (by car or public transportation or if close by walk or bycycle)
- Radius used to seek places nearby

Then I Repaired the data so that scandinavic characters are included (ae -> ä and eo -> ö) and drop the column ’Unnamed: 0’. After that I joined data read from https://fi.wikipedia.org/wiki/Luettelo_Tampereen_tilastoalueista and from my local file. When checking the data I could see that there were neighbourhoods with the same Latitude, Longitude. I joined these neighbourhoods together (Amuri A and Amuri B to Amuri, Kyttälä A and Kyttälä B to Kyttälä, and Tammela A and Tammela B to Tammela).

Then I used Foursquare API to search kebab restaurants/pizzerias/fast food places or other restaurants in the neighborhood using a specific radius for each neighborhood. After that I removed multiple venues leaving only the food place closest to the neighbourhood. When visualizing and investigating the map thorougly near Tampere boarders I could find that there are still two venues outside Tampere. One is 'Jagin Kööki' which is clearly in Ylöjärvi, and the other is 'Kokkinurkkaus' which is just close to the border, and it locates in Kangasala. I removed these venues. Finally I counted the number of shops and bars and other competitors 100 meters around each kebab/pizza/fast food place or other restaurant.