The Battle of Neighborhoods (Week 1)

1. A description of the problem and a discussion of the background. (**15 marks**)
2. A description of the data and how it will be used to solve the problem. (**15 marks)**

Introduction: Business Problem

The scenario for this project is that we will try to find an optimal location for a restaurant Berlin, Germany. We are working for an Italian restaurant who is looking for an optimal location to attract customers but where the market is not overly saturated. Berlin has a lot of restaurants, so we are interested in areas with no Italian restaurants in vicinity. Ideally, we want the location to be close to the city center.

We will use our data science powers to generate a few most promising neighborhoods based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

Description of Data

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

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to Italian restaurants in the neighborhood
- distance of neighborhood from city center

We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be needed to extract/generate the required information:

- centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using Google Maps API reverse geocoding
- number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API
- coordinate of Berlin center will be obtained using Google Maps API geocoding