Coursera Capstone Project – SF restaurants

1. Introduction/Business Problem

A client wants to open a restaurant in the area of San Francisco.
The client is new to the city and is not bound to a specific kind of cuisine.
He needs to know:

1. what kind of cuisine would attract a lot of customers
2. which area is suitable for that kind of cuisine
3. a possible additional competitor analysis

Data science will be used to weigh between the location and cuisine of restaurants, using data gathered from the website foursquare.com.

2. Data

Three kinds of data are necessary to answer the business problem:

1. The location of a restaurant (geo location, district)
2. The type of restaurant (e.g. Chinese or Mexican)
3. The business metrics of the restaurant (likes)

In the Case of the Foursquare API, the relevant data will be extracted from the 'search' and 'likes' endpoint:

1. The geo data from the previous Coursera assignments will mark the city border.
   Within the border arbitrary exploration points will be defined by Longitude and latitude.
2. The *'categoryId'* tag for food ('4d4b7105d754a06374d81259') will be used to extract the id, category, location and name of a restaurant around an exploration point.
3. The *'categoryId'* from 2. will be used for an additional API query.
   The result contains the number of likes per restaurant and a list of a few users who liked the restaurant.

The number of premium calls for the *'details'* endpoints is very limited.
 The rating of the restaurant, which is part of the response of calling the *'details'* endpoint would be a better option, but again, premium calls are limited.

Therefore, I will use the like count as a metric for the business performance of a restaurant, which allows to collect more than 50 data points for the large area of SF.
I will assume, that many likes correspond to many and happy customers and therefore large revenue for the restaurant's owners.