Introduction

Being motivated by the projects of exploring Toronto and New York neighborhoods in this course, I decided to leverage Foursquare location data to explore my city, Amsterdam in the Netherlands. In Amsterdam, we can basically classify all neighborhoods by two regional types, Buurt and Wijk in Dutch. The former means a typical neighborhood or a localized community within a city, and the latter stands for a more residential area with a lot of old and/or historical family houses. In this project, I would like to investigate whether these two types of region have a similar distribution of venues, or backwardly, can we cluster all neighborhoods based on their venues in order to identify the region types of each of them.

Data

Statistics Netherlands CBS, a Dutch governmental institution that gathers statistical information about the Netherlands, published a dataset of some basic figures for the districts and neighborhoods of Amsterdam in 2016. Those figures included important data that I need in this project, such as Neighborhood, Region Type, Population, Latitude and Longitude.

Given this Amsterdam dataset, I am going to clean it up first, only focus on those neighborhoods with a relatively large number of inhabitants. Then, go over the coordinate of those neighborhoods, to retrieve the venues information by applying foursquare API.