Introduction/Business

A bottling company has an established role in the market of Toronto. They are the premier supplier of bottled water with over 1000 clients in Toronto's neighborhoods. His best clients consist of the following types of businesses: restaurants and bars, hotels and coffee shops. The bottled water is currently stored in a central warehouse outside of the city and is distributed to the clientele daily. The main problem with this business plan is that the distribution becomes is inefficient and becomes time consuming and costly.

Our client seeks to reduce the distribution costs by building 5 smaller warehouses in Toronto to serve as smaller hubs and distribute to his clients from those points since it'd be closer, eliminating the time and resource costs caused by the distance of the original warehouse from his clients. In order to accomplish this, our client has asked us to find the 5 best locations in the city of Toronto where the warehouses could be constructed in order to create smaller distribution clusters. The warehouses must form the centers of the clusters in order to minimize their relative distances from each client.

Data

The data required for this task are the locations of the hotels, coffee shops, bars and restaurants in Toronto. In order to gather the required data, we will gather the locations of Toronto neighborhoods from Wikipedia. With this data, we will pinpoint the locations of these neighborhoods on Foursquare. The data will then be filtered in order to acquire the locations of the targeted venues needed for our study. We will utilize Folium to build a map of Toronto and visualize the delivery venues. Following this, the k-means algorithm will be used to define the centroids of the five clusters which will serve as the best possible points at which distribution warehouses can be built.