

Problem

In Toronto, there are many people aspiring to become business owners. Toronto is a profitable and prime region for businesses to be built.

Nevertheless, it is pertinent that based on the target products, the businesses establish their outlets in the most viable locations.

A significant percentage of them are unsure of where to start their business in. In fact, a considerable percentage of startups fail within their first few years in Toronto.

Data employed

To help these startup business owners to be better informed about better locations to build their business, we will analyze the data provided in

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada: M

Data will also be made use of from https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/lab_s_v1/Geospatial_Coordinates.csv

Data consists primarily of geospatial information, including PostalCode, Borough, Neighborhood, Latitude and Longitude. Using additional information from FourSquare API, the surrounding amenities and shops will also be accounted for.

Approach

By analyzing these information, we can provide reasonable suggestions for locations for business to establish their outlets.

We will analyze by employing machine learning. Locations will be clustered using the renowned K-means clustering algorithm, an unsupervised machine learning algorithm.