Applied Data Science - Capstone Project

Author: Ramajeyam Gopalraj

Introduction / Business Problem

I am a business man who is looking to setup a bar in Toronto area. I want to pick a place where there are not many existing bars. Hence, I want to apply data science using FourSquare data on Toronto area and locate area where there are other amenities such as Pizza Place and Diners, but not many Bars. We will use K-Means clustering to group the data by number of bars and pick the area with least number of bars and many Pizza and Diner places.

Data

We will use the following data attributes to work with.

1. Neighborhood
2. Venue
3. Venue Category
4. Latitude
5. Longitude