**PROYECTO PARA CLUSTERIZAR LOS BARRIOS DE TORONTO Y NUEVA YORK**

1. **Introducción.**

Los métodos de clusterización o agrupamiento nos sirven de mucho ya que nos permiten agrupar ya sea observaciones o individuos en clases, de modo que los objetos que se encuentren en una misma clase van a tener características similares o siguen un mismo patrón a diferencia de otros objetos que pertenecen a otra clase.

Este proceso o algoritmo de clusterización nos pude ayudar para agrupar a nuestro cliente, productos entre otras cosas.

1. **Problema.**

Se pretende identificar que vecindarios de Toronto son similares o comparten características con vecindarios de Nueva York para así tomar una decisión y ver en que vecindarios es más factible poner una franquicia de pizzerías.

1. **Datos a usar**

Los datos que vamos a usar son datos de vecindarios de NY y Toronto, estos datos están en un dataframe y este tiene las siguientes características

Cantidad de filas y columnas: 409 filas y 5 columnas.

Nombre de columnas:

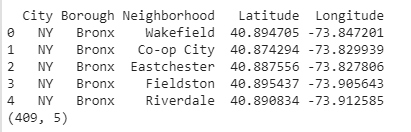
**City:** Nombre de las ciudades (Nueva York y Toronto)

**Borough:** Nombre de los distritos

**Neighborhood:** Nombre de los vecindarios

**Latitude:** Latitud del vecindario

**Longitude:** Longitude del vecindario



1. También haremos uso del API de Forsquare para obtener datos de los lugares que contiene cada vecindario
2. t