

# 1 Enunciado

Para este proyecto se requiere implementar en lenguaje C++ un programa que soporte la ejecución de un conjunto de consultas espaciales sobre un conjunto de datos abierto. El conjunto de datos principal con el que se debe trabajar está compuesto por cientos de miles de registros que representan viajes de taxi en la ciudad de New York correspondientes al año 2015 <sup>1</sup>. Estos datos son ofrecidos por la organización llamada NYC Taxi and Limousine Commission. El segundo conjunto de datos es la demarcación de los barrios de la ciudad de New York en formato GeoJson <sup>2</sup>.

## 1.1 Consultas

Las consultas que su solución debe soportar son:

- ¿Qué viajes tuvieron como origen y destino el mismo barrio?
- ¿Cuáles son los barrios con mayor cantidad de viajes? Retornar los top 5. (Considerando apenas las coordenadas de partida).
- Dados dos puntos P1 y P2 que representan una región rectangular, retornar el número de viajes que comenzaron en tal región.
- Determinar por cuáles vecindarios pasó un taxi en un determinado viaje, considerando la ruta como una línea recta.
- Dado un punto P y una distancia D, encontrar los viajes que empezaron o terminaron a una distancia D como máximo a partir del punto P.

## 2 Consideraciones

- El conjunto de datos principal a utilizar está compuesto por 36 archivos CSV, que son los registros de 3 proveedores distintos para cada mes del año. Dado que el total del tamaño de los datos está en el orden de los GBs, se debe considerar esta escala desde el comienzo para el diseño de su propuesta.
- A pesar que el conjunto de datos completo se conoce a priori, el objetivo es que su programa utilice estructuras de datos dinámicas que además puedan permitir las consultas de manera eficiente.
- Se sugiere se utilice alguna herramienta de terceros para validar el resultado de algunos experimentos tales como PostGIS, MongoDB u Oracle.

---

<sup>1</sup>[https://github.com/toddschneider/nyc-taxi-data/blob/master/setup\\_files/raw\\_data\\_urls.txt](https://github.com/toddschneider/nyc-taxi-data/blob/master/setup_files/raw_data_urls.txt)

<sup>2</sup><https://data.beta.nyc/dataset/pediacities-nyc-neighborhoods/resource/35dd04fb-81b3-479b-a074-a27a37888ce7>