

# RETO 3: Accidentes de Tránsito en Bogotá

---

## Objetivo

Poner en práctica los conceptos aprendidos en clase acerca de las estructuras de datos que implementan el TAD Mapa Ordenado (árboles) y realizar búsquedas eficientes de información. Específicamente se desea:

1. **Practicar** los conceptos sobre estructuras de tablas ordenadas (árboles binario de búsqueda y balanceados).
2. **Integrar** el uso de las estructuras lineales (tablas de hash, Lista, pilas y colas) a los árboles binarios.
3. **Practicar** los algoritmos de ordenamientos y búsquedas lineales.
4. **Utilizar** adecuadamente el patrón MVC (Modelo-Vista-Controlador).
5. **Aprender** a cargar y procesar en memoria datos en formato CSV.
6. **Utilizar** adecuadamente el administrador de versiones GIT y GitHub.
7. **Aprender** a trabajar en equipo.

## Fecha Límite de Entrega

El reto podrá entregarse hasta el martes **25 de abril** antes de la media noche (11:59 p.m. hora BrightSpace).

## Actualizaciones

A continuación, encontrará un listado de todas las actualizaciones realizadas sobre el enunciado para claridad de los requerimientos y objetivos del Reto.

1. Se **corrigieron** errores de redacción y comprensión sobre los requerimientos del 1 al 7 para facilitar su entendimiento.
2. Se **agregaron** los ejemplos funcionales de los requerimientos de la carga de datos y el bono utilizando el subconjunto de datos -small.

## Contexto

La capital de Colombia, al ser la ciudad más poblada del país, tiene el mayor número de vehículos (incluye especialmente carros, buses, transporte de carga, motocicletas, bicicletas y patinetas). Esto hace que la problemática de movilidad y en consecuencia los accidentes de tránsito con el tiempo se ha vuelto un tema serio de análisis por parte de autoridades distritales, de tránsito, de policía y de justicia.

Gracias a diferentes medidas como el pico y placa, la reducción de velocidad, las ciclovías, las ciclorrutas, el mejoramiento en el servicio de transporte público masivo, la semaforización inteligente y las campañas de concientización están dirigidas a todos los actores en las vías públicas se busca mejorar la movilidad y disminuir la accidentalidad.

Hoy la ciudad está por debajo del índice de mortalidad del país en accidentes, los números de accidentes seguirán creciendo, en especial a medida que la cantidad de motocicletas en circulación aumentan.

## Fuente de Datos

Los datos para este reto fueron tomados del repositorio de Datos abiertos de la Secretaría Distrital de Movilidad de Bogotá, que contiene los registros de todos los accidentes en la ciudad desde el año 2015 hasta el año 2022.

Los archivos **datos\_siniestralidad-<<xxx>>.csv** contienen la información necesaria con todos los siniestros viales de la ciudad, los cuales están preparados y disponibles para los estudiantes en el aula unificada en Bloque Neón (BrightSpace).

Cada uno de estos archivos contiene 15 columnas que permiten establecer las diferentes propiedades de un accidente. Esto incluye su hora, fecha, locación, gravedad, entre otros. Para obtener más detalles puede revisar la [Tabla 1](#) y cada uno de los archivos dispuestos en Bloque Neón.

Nombre de la Columna	Descripción
FORMULARIO	Código del formulario de registro
CODIGO_ACC	Código del accidente
FECHA_OCURRENCIA_ACC	Fecha en formato AAAA/MM/DD del accidente
HORA_OCURRENCIA_ACC	Hora en formato HH:MM:SS de ocurrencia del accidente
ANO_OCURRENCIA_ACC	(Categórica) Año de ocurrencia del accidente
MES_OCURRENCIA_ACC	(Categórica) Mes de ocurrencia del accidente
DIA_OCURRENCIA_ACC	(Categórica) Día de la semana de ocurrencia del accidente (LUNES, MARTES...)
DIRECCION	(Categórica) Dirección en la que ocurrió el accidente
GRAVEDAD	(Categórica) Gravedad definida como: SOLO DANOS, CON HERIDOS o CON MUERTOS
CLASE_ACC	(Categórica) Tipo de accidente definido como: CHOQUE, ATROPELLO; VOLCAMIENTO, INCENDIO, AUTOLESION, CAIDA DE OCUPANTE u OTRO
LOCALIDAD	(Categórica) Localidad en la que ocurrió el accidente
MUNICIPIO	(Categórica) Columna cuyo único valor es Bogotá
FECHA_HORA_ACC	Fecha y hora en formato AAAA/MM/DD HH:MM:SS de ocurrencia del accidente
LATITUD	Latitud geográfica en la que ocurrió el accidente
LONGITUD	Longitud geográfica en la que ocurrió el accidente

Tabla 1. Descripción de campos del archivo CSV *category\_data* con la información de siniestros viales.

Para evitar problemas de buffer en la lectura de los archivos se recomienda aumentar el tamaño de los campos de lectura de la librería **Python CSV** al máximo posible para el sistema con el siguiente comando en la librería CSV en el **controler.py** del Reto.

```
import csv
...
csv.field_size_limit(2147483647)
```

En algunos casos experimentales puede que Python y el IDE declaren que se alcanzó el límite de recursión con un mensaje **“RecursionError: maximum recursion depth exceeded in comparison”**, en este caso se recomienda actualizar en el **view.py** este límite con las siguientes líneas de código:

```
import sys
...
default_limit = 1000
sys.setrecursionlimit(default_limit*10)
```

## Trabajo Propuesto

### Parte 1: Configuración Repositorio

Complete los siguientes pasos para configurar su repositorio de trabajo:

- p.1. [Cree](#) en GitHub un repositorio basado en la plantilla propuesta para el reto, el cual se encuentra en el URL: <https://github.com/ISIS1225DEVs/Reto3-Template>
- p.2. [Renombre](#) el repositorio de su reto con el formato **Reto3-G<<Número del grupo>>** ej.: **Reto3-G01** para el grupo 1 de la sección.
- p.3. [Edite](#) el **README** del repositorio e incluya los nombres completos, correo Uniandes y códigos de los miembros del equipo de trabajo.
- p.4. [Realice](#) el procedimiento según lo aprendido en clase para clonar el repositorio en su máquina local y sincronizarlo con su repositorio en GitHub.
- p.5. [Descargue](#) los datos desde la sección unificada del curso y cópielos en la carpeta **data** del repositorio local.

### Parte 2: Carga de Datos

En la sección unificada en la sección de reto 3 encontrarán los datos oficiales del proyecto. el ZIP contiene varios archivos con los sufijos -small, -large, -5pct, -10pct, -20pct, -30pct, -50pct y -80pct. Son archivos con diferente número de registros. Esto facilita la implementación y pruebas en computadores con memoria RAM y procesadores reducidos.

Para responder a los requerimientos presentados deberán cargar la información del archivo entregado; recuerde que solo se permite leer una vez la información de este y que las pruebas finales sobre sus algoritmos serán sobre el archivo `_large.csv`.

Al final de la carga de datos debe reportar los siguientes datos:

- El total de los registros cargados desde el archivo.
- Mostrar los primeros 3 y últimos 3 registros de accidentes cargados con las siguientes características:
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).

## Recomendaciones:

- Se recomienda utilizar librerías por extensión de Python como `tabulate`<sup>1</sup> para imprimir adecuadamente los resultados como se muestran en los ejemplos.
- Se recomienda utilizar de la librería **`datetime`** nativa de Python para facilitar el procesamiento de las fechas y horas<sup>2</sup> dentro del proyecto.

## EJEMPLO: carga de datos archivos con sufijo **-small**.

```
-----
Información de los accidentes cargados:
Total de accidentes: 2358
Total de columnas cargadas: 15
-----

Los primeros tres registros de accidentes cargados fueron:
```

FECHA_OCURRENCIA_A CC	FECHA_HORA_ACC	LOCALIDAD	DIRECCION	GRAVEDAD	CLASE_ACC	LATITUD	LONGITUD
2022/09/21	2022/09/21 15:14:00+00	PUENTE ARANDA	KR 65 -CL 13	SOLO DANOS	CHOQUE	4.635	-74.115
2017/01/25	2017/01/25 20:00:00+00	BARRIOS UNIDOS	CL 79-KR 29	SOLO DANOS	CHOQUE	4.67163	-74.0674
2022/02/13	2022/02/13 23:50:00+00	CIUDAD BOLIVAR	KR 27 B -CL 71 H S	SOLO DANOS	CHOQUE	4.55	-74.159

```
-----

Los últimos tres registros de accidentes cargados fueron:
```

FECHA_OCURRENCIA_A CC	FECHA_HORA_ACC	LOCALIDAD	DIRECCION	GRAVEDAD	CLASE_ACC	LATITUD	LONGITUD
2022/07/06	2022/07/06 18:40:00+00	KENNEDY	AV AVENIDA CIUDAD DE CALI-KR 13	SOLO DANOS	CHOQUE	4.653	-74.143
2016/03/12	2016/03/12 12:00:00+00	LOS MARTIRES	CL 22-KR 29	SOLO DANOS	CHOQUE	4.6191	-74.0849
2020/06/10	2020/06/10 09:20:00+00	PUENTE ARANDA	KR 34-CL	SOLO DANOS	CHOQUE	4.60501	-74.1062

<sup>1</sup> Python tabulate, PyPi URL: <https://pypi.org/project/tabulate/>

<sup>2</sup> datetime — Basic date and time types, URL: <https://docs.python.org/3/library/datetime.html>

### Parte 3: Desarrollo de los Requerimientos

Para este reto se han identificado ocho (8) requerimientos, siete (7) obligatorios y un (1) bono. Divididos de acuerdo con su dificultad en básico, intermedio y avanzado. Adicionalmente, es importante resaltar que tres (3) de estos requerimientos se deben desarrollar de forma individual y los cinco (5) restantes son grupales.

El resumen de los requerimientos se muestra en la siguiente tabla y se explican detalladamente en la siguiente sección.



**NOTA:** En caso de que el equipo este conformado solamente por **dos (2)** integrantes deberán resolver solamente **cinco (5)** requerimientos obligatorios: uno de los requerimientos **básicos**, dos de los requerimientos **intermedios**, y los dos requerimientos **avanzados**.

## Requerimiento No. 1 (Grupal): Reportar todos los accidentes dado un rango de fechas

Como analista de la secretaría de movilidad **deseo** conocer los accidentes que ocurrieron en la ciudad durante un intervalo de fechas específico.

Los **parámetros de entrada** de este requerimiento son:

- Fecha inicial del intervalo (FECHA\_HORA\_ACC).
- Fecha final del intervalo (FECHA\_HORA\_ACC).

La **respuesta esperada** debe contener:

- El número total de accidentes ocurridos en el rango de fechas.
- Todos los accidentes ocurridos, ordenados de más reciente a menos reciente por su fecha de ocurrencia (FECHA\_HORA\_ACC).
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).

### Recomendación:

- Muestre los accidentes en orden del más reciente al menos reciente, teniendo en cuenta año, mes, día, hora y minuto del accidente.

**EJEMPLO:** Deseo un reporte de los accidentes ocurridos entre el **01/11/2016** y el **08/11/2016**:

Hay 9 accidentes registrados entre 2016/11/1 y 2016/11/8

CODIGO_ACCIDENTE	DIA_OCURRENCIA_ACC	DIRECCION	GRAVEDAD	CLASE_ACC	LOCALIDAD	FECHA_HORA_ACC	LATITUD	LONGITUD
4463272	MARTES	CL 37-KR 23 S	CON HERIDOS	CHOQUE	RAFAEL URIBE URIBE	2016/11/08 07:15:00+00	4.57958	-74.1189
4463898	SABADO	CL 72-KR 54	CON HERIDOS	CHOQUE	BARRIOS UNIDOS	2016/11/05 20:50:00+00	4.66986	-74.076
4463113	SABADO	AC 63-KR 57	CON HERIDOS	ATROPELLO	TEUSAQUILLO	2016/11/05 13:30:00+00	4.65527	-74.0858
4463111	SABADO	AV AVENIDA CARACAS-CL 67	CON HERIDOS	ATROPELLO	CHAPINERO	2016/11/05 08:00:00+00	4.654	-74.0634
4463553	VIERNES	AV AVENIDA SUBA-CL 115	SOLO DANOS	CHOQUE	SUBA	2016/11/04 21:40:00+00	4.69808	-74.0696
4463183	VIERNES	AV AVENIDA CIUDAD DE CALI- CL 88	SOLO DANOS	CHOQUE	ENGATIVA	2016/11/04 13:30:00+00	4.70818	-74.0963
4462881	JUEVES	AV AVENIDA DE LAS AMERICAS-KR 80G	CON HERIDOS	CHOQUE	KENNEDY	2016/11/03 17:39:59+00	4.63595	-74.1533
4462914	JUEVES	KR 96I-CL 15B	SOLO DANOS	CHOQUE	FONTIBON	2016/11/03 12:40:00+00	4.66433	-74.149
4462898	MARTES	CL 100-KR 8A	SOLO DANOS	CHOQUE	CHAPINERO	2016/11/01 10:40:00+00	4.68154	-74.0416

## Requerimiento No. 2 (Grupal): Reportar todos los accidentes en un intervalo de horas del día para un mes y año dados

Como analista de la secretaría de movilidad **deseo** conocer los accidentes ocurridos en un intervalo de tiempo del día (desde una hora y minutos iniciales hasta una hora y minutos finales) para un año y mes dados.

Los **parámetros de entrada** de este requerimiento son:

- Hora y minutos iniciales del intervalo de tiempo.
- Hora y minutos finales del intervalo de tiempo.
- Año de consulta.
- Mes de consulta.

La **respuesta esperada** debe contener:

- El número total de accidentes ocurridos en el intervalo de tiempo, año y mes dados.
- Todos los accidentes ocurridos, ordenados de menos reciente a más reciente por su hora y minutos de ocurrencia.
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Hora del accidente (HORA\_OCURRENCIA\_ACC).
  - Fecha del accidente (FECHA\_OCURRENCIA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).

- Longitud (LONGITUD).

#### Recomendación:

- Muestre los accidentes en orden del más reciente más antiguo.
- En caso de que haya varios accidentes ocurridos en una misma hora-minutos, estos accidentes se ordenan por fecha de la menos reciente a la más reciente.

**EJEMPLO:** Deseo un reporte de los accidentes ocurridos para el mes de noviembre de 2022 entre 18:00 y 21:00:

```

===== Req No. 2 Inputs =====
Reportar todos los accidentes para el año '2022' para el mes de 'NOVIEMBRE' en el intervalo de horas '18:00:00' y '21:00:00'

===== Req No. 2 Answer =====
Hay 3 accidentes en el intervalo de horas dado '18:00:00' y '21:00:00' en el año '2022' y en el mes 'NOVIEMBRE'
  
```

CODIGO_ACCIDENTE	HORA_OCURRENCIA_AC C	FECHA_OCURRENCIA_A CC	DIA_OCURRENCIA_ACC	LOCALIDAD	DIRECCION	GRAVEDAD	CLASE_ACC	LATITUD	LONGITUD
10584622	19:23:00	2022-11-30	MIERCOLES	KENNEDY	KR 89 C-CL 42 S	CON HERIDOS	ATROPELLO	4.63724	-74.1731
10583987	19:37:00	2022-11-13	DOMINGO	SUBA	CL 145 -KR 58	CON HERIDOS	CHOQUE	4.73069	-74.0635
10583959	20:02:00	2022-11-10	JUEVES	ANTONIO NARINO	AV AVENIDA CARACAS-CL 13 A S	SOLO DAÑOS	VOLCAMIENTO	4.58482	-74.096

### Requerimiento No. 3 (Individual): Reportar los 3 accidentes más recientes de una clase particular ocurridos a lo largo de una vía

**Como** analista de la secretaría de movilidad **deseo** un reporte de los 3 accidentes de cierta clase ocurridos en una de las vías de la ciudad

Los **parámetros de entrada** de este requerimiento son:

- Clase del accidente.
- Nombre de la vía de la ciudad.

La **respuesta esperada** debe contener:

- El número total de accidentes ocurridos en la vía en cuestión que correspondan a la clase.
- Los 3 accidentes más recientes ocurridos en la vía en cuestión que correspondan a la clase. Por cada accidente se debe mostrar:
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).



Recomendación:

- Muestre los accidentes en orden descendente del más reciente al más antiguo.

**EJEMPLO:** Deseo un reporte de los accidentes con choques ocurridos a lo largo de la Avenida de las Américas:

===== Req No. 3 Inputs =====  
Hay 31 accidentes de la clase 'CHOQUE' ocurridos a lo largo de la vía 'AV AVENIDA DE LAS AMERICAS' y los tres mas recientes son:  
===== Req No. 3 Answer =====

CODIGO_ACCIDENTE	FECHA_HORA_ACC	DIA_OCURRENCIA_ACC	LOCALIDAD	DIRECCION	GRAVEDAD	CLASE_ACC	LATITUD	LONGITUD
10580796	2022-09-16 16:40:00	VIERNES	PUENTE ARANDA	AV AVENIDA DE LAS AMERICAS-KR 50	SOLO DANOS	CHOQUE	4.6276	-74.1066
10573004	2022-06-06 18:59:59	LUNES	KENNEDY	AV AVENIDA DE LAS AMERICAS-KR 68 D	SOLO DANOS	CHOQUE	4.629	-74.126
10557872	2021-12-07 07:40:00	MARTES	KENNEDY	AV AVENIDA DE LAS AMERICAS-KR 78 K	SOLO DANOS	CHOQUE	4.632	-74.15

## Requerimiento No. 4 (Individual): Reportar los 5 accidentes más recientes dada una gravedad y un rango de fechas

Como analista de la secretaría de movilidad **deseo** un reporte de los 5 accidentes más recientes para un intervalo de fechas y una gravedad dadas.

Los **parámetros de entrada** de este requerimiento son:

- Fecha inicial del intervalo.
- Fecha final del intervalo.
- Gravedad del accidente.

La **respuesta esperada** debe contener:

- El número de accidentes para el intervalo de fechas especificado.
- Los 5 registros de accidentes de la gravedad dada más antiguos. Por cada accidente se debe presentar la siguiente información:
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).

### Recomendaciones:

- Muestre los accidentes en orden descendente del más reciente al más antiguo.

**EJEMPLO:** Deseo un reporte de los accidentes con muertos que hayan ocurrido entre el 01/10/2016 y el 01/10/2018.

```
===== Req No. 4 Inputs =====
Reporte de los 5 accidentes 'CON MUERTOS' más recientes entre '2016-10-01 00:00:00' y '2018-10-01 23:59:59'

===== Req No. 4 Answer =====
Hay 9 accidentes entre las fechas '2016-10-01 00:00:00' y '2018-10-01 23:59:59'
```

CODIGO_ACCIDENTE	FECHA_HORA_ACC	DIA_OCURRENCIA_ACC	LOCALIDAD	DIRECCION	CLASE_ACC	LATITUD	LONGITUD
10452028	2018-06-23 19:20:00	SABADO	KENNEDY	AV AVENIDA PRIMERO DE MAYO-CL 35 S	VOLCAMIENTO	4.61919	-74.1433
10448670	2018-05-23 20:45:00	MIERCOLES	TEUSAQUILLO	AK 30-CL 4	ATROPELLO	4.63482	-74.0793
4505097	2018-01-09 18:35:00	MARTES	USME	KR 14-CL 75A S	CHOQUE	4.5137	-74.1157
4502060	2017-12-08 05:00:00	VIERNES	USAQUEN	AC 9-KR 166	CHOQUE	4.74438	-74.0317
4494461	2017-09-22 21:14:59	VIERNES	BOSA	CL 71A-KR 82	CHOQUE	4.61244	-74.1949

## Requerimiento No. 5 (Individual): Reportar los 10 accidentes menos recientes ocurridos en un mes y año para una localidad de la ciudad

Como analista de la secretaría de movilidad **deseo** un reporte de los 10 accidentes más recientes ocurridos en un mes y un año en una localidad de la ciudad.

Los **parámetros de entrada** de este requerimiento son:

- Localidad de Bogotá en la que ocurrieron accidentes.
- Un mes.
- Año entre 2015 y 2022.

La **respuesta esperada** debe contener:

- El número total de accidentes ocurridos en la localidad durante la fecha indicada.
- Los 10 registros de accidentes más recientes en la localidad. Cada accidente debe mostrar la siguiente información:
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).

### Recomendaciones:

- Muestre los accidentes en orden descendente del más reciente al más antiguo.

**EJEMPLO:** Deseo un reporte de los accidentes menos recientes ocurridos en la localidad de Fontibón para noviembre del 2016.

Hay 4 accidentes ocurridos en la localidad de FONTIBON en el mes NOVIEMBRE del año 2016  
Estos son los 10 accidentes menos recientes:

CODIGO_ACCIDENTE	DIA_OCURRENCIA_ACC	DIRECCION	GRAVEDAD	CLASE_ACC	FECHA_HORA_ACC	LATITUD	LONGITUD
4465469	VIERNES	AV AVENIDA BOYACA-CL 18	SOLO DANOS	CHOQUE	2016/11/25 00:30:00+00	4.65037	-74.1257
4464648	JUEVES	CL 26-KR 69	SOLO DANOS	CHOQUE	2016/11/17 21:59:59+00	4.65786	-74.1068
4464646	MIERCOLES	AV AVENIDA CENTENARIO-KR 87	SOLO DANOS	CHOQUE	2016/11/16 21:00:00+00	4.66053	-74.137
4462914	JUEVES	KR 96I-CL 15B	SOLO DANOS	CHOQUE	2016/11/03 12:40:00+00	4.66433	-74.149

## Requerimiento No. 6 (Grupal): Mostrar los N accidentes ocurridos dentro de una zona específica para un mes y un año

Como analista de la secretaría de movilidad **deseo** un reporte que muestre un número particular de accidentes que ocurrieron dentro de una zona circular específica de la ciudad para un mes y un año.

Los **parámetros de entrada** de este requerimiento son:

- Mes
- Año entre 2015 y 2022
- Coordenadas del centro del área (Latitud y Longitud)
- Radio del área en km
- Número de accidentes

La **respuesta esperada** debe contener:

- Los N accidentes ocurridos, organizados por su cercanía al centro de la zona.
- Los registros de todos los N accidentes ocurridos en la zona. Cada accidente debe mostrar la siguiente información:
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).

### INSTRUCCIONES ADICIONALES

Dado que la tierra no es plana, sino que tiene la forma aproximada de una esfera, el cálculo de la distancia entre dos puntos de la superficie debe realizarse usando la **fórmula de Haversine**. Esta fórmula permite aproximar dicha distancia a partir de la latitud y longitud de dos puntos sobre una esfera.

En términos generales, la fórmula de Haversine se define así:

$$D = 2 \cdot \sin^{-1} \left( \sqrt{\sin^2 \left( \frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left( \frac{\lambda_2 - \lambda_1}{2} \right)} \right) \cdot \mu_E$$

En donde:

- D es la distancia entre los dos puntos en kilómetros.

- $\phi_1$  es la latitud del primer punto en radianes.
- $\phi_2$  es la latitud del segundo punto en radianes.
- $\lambda_1$  es la longitud del segundo punto en radianes.
- $\lambda_2$  es la longitud del segundo punto en radianes.
- $\mu_E$  es el radio de la tierra (6371 km).

Para más información, puede consultar los siguientes recursos:

- [Haversine - Distance \(vcalc.com\)](https://vcalc.com/haversine-distance/)
- [Finding distances based on Latitude and Longitude \(hedges.name\)](https://hedges.name/finding-distances-based-on-latitude-and-longitude/)

## Recomendaciones

- Para calcular la distancia entre dos ubicaciones geográficas con datos de latitud y longitud se sugiere utilizar la **Fórmula del semiverseno (Haversine formula)**<sup>3,4,5,6</sup>.

**EJEMPLO:** Deseo un reporte de los 3 accidentes más cercanos al punto con latitud 4.674, longitud -74.068 dentro de un radio de 5.0 km para el mes de enero de 2022.

Los 3 accidentes más cercanos al punto (4.674, -74.068) dentro de un radio de 5 km para el mes de ENERO de 2022							
CODIGO_ACCIDENTE	DIA_OCURRENCIA_ACC	DIRECCION	GRAVEDAD	CLASE_ACC	FECHA_HORA_ACC	LATITUD	LONGITUD
10562094	VIERNES	KR 14 -CL 78	SOLO DANOS	CHOQUE	2022/01/28 16:57:59+00	4.66458	-74.0572
10561695	MIERCOLES	CL 63 -KR 28	CON HERIDOS	ATROPELLO	2022/01/26 12:10:00+00	4.651	-74.076
10561926	VIERNES	CL 64 -KR 71 A BIS	CON HERIDOS	CHOQUE	2022/01/28 17:19:59+00	4.675	-74.102

<sup>3</sup> Haversine formula, URL: [https://en.wikipedia.org/wiki/Haversine\\_formula](https://en.wikipedia.org/wiki/Haversine_formula)

<sup>4</sup> Calculate distance, bearing and more between Latitude/Longitude points, URL: <https://www.movable-type.co.uk/scripts/latlong.html>

<sup>5</sup> Haversine Formula in Python (Bearing and Distance between two GPS points), URL: <https://stackoverflow.com/questions/4913349/haversine-formula-in-python-bearing-and-distance-between-two-gps-points>

<sup>6</sup> haversine 2.5.1, URL: <https://pypi.org/project/haversine/>

## Requerimiento No. 7 (Grupal): Reportar los accidentes más temprano y más tarde para cada día de un mes y año dado, y graficar el histograma de frecuencias de accidentes por hora para ese mismo mes y año

Como analista de la secretaría de movilidad **deseo** conocer el accidente más temprano y más tarde para cada día de un año y mes dados. Adicionalmente, generar el histograma con el número de accidentes por cada hora (entera) del día para los accidentes del mismo año y mes dados.

Los **parámetros de entrada** de este requerimiento son:

- Mes
- Año entre 2015 y 2022

La **respuesta esperada** debe contener:

- Para cada día del mes en el año dados, dar la información del accidente ocurrido más temprano y más tardío. Cada accidente debe mostrar la siguiente información:
  - Código del accidente (CODIGO\_ACCIDENTE).
  - Fecha y hora del accidente (FECHA\_HORA\_ACC).
  - Día del accidente (DIA\_OCURRENCIA\_ACC).
  - Localidad (LOCALIDAD).
  - Dirección (DIRECCION).
  - Gravedad (GRAVEDAD).
  - Clase de accidente (CLASE\_ACC).
  - Latitud (LATITUD).
  - Longitud (LONGITUD).
- La gráfica de barras con una representación (simplificada) de los accidentes ocurridos en cada franja horaria durante el mes y año dados, entendiendo una franja horaria como el tiempo entre una hora y justo antes de la hora anterior (p.ej de 3:00 a 3:59).

**IMPORTANTE:** todas las operaciones de conteo deben realizarse en el modelo (model.py) con la librería DISCLib, mientras que el histograma solo debe configurarse y ejecutarse en la vista (view.py).

### Recomendaciones

- Muestre la respuesta agrupada del primer día al último día del mes, mostrando primero el accidente más temprano y luego el accidente más tardío.
- Si un día no tiene accidentes, no debe mostrarse, y si sólo hay un accidente, debe aparecer en ambas categorías.
- Si hay dos accidentes a la misma hora, debe mostrar el que tenga el menor código (CODIGO\_ACCIDENTE).
- Apóyese de los requerimientos anteriores para resolver este ejercicio.

- Para realizar el diagrama de barras, utilice la librería [matplotlib](#) de Python con listas de Python. No debe usar la librería Pandas o similares.

**EJEMPLO:** Deseo un reporte de los accidentes menos recientes ocurridos en el mes de diciembre de 2019.

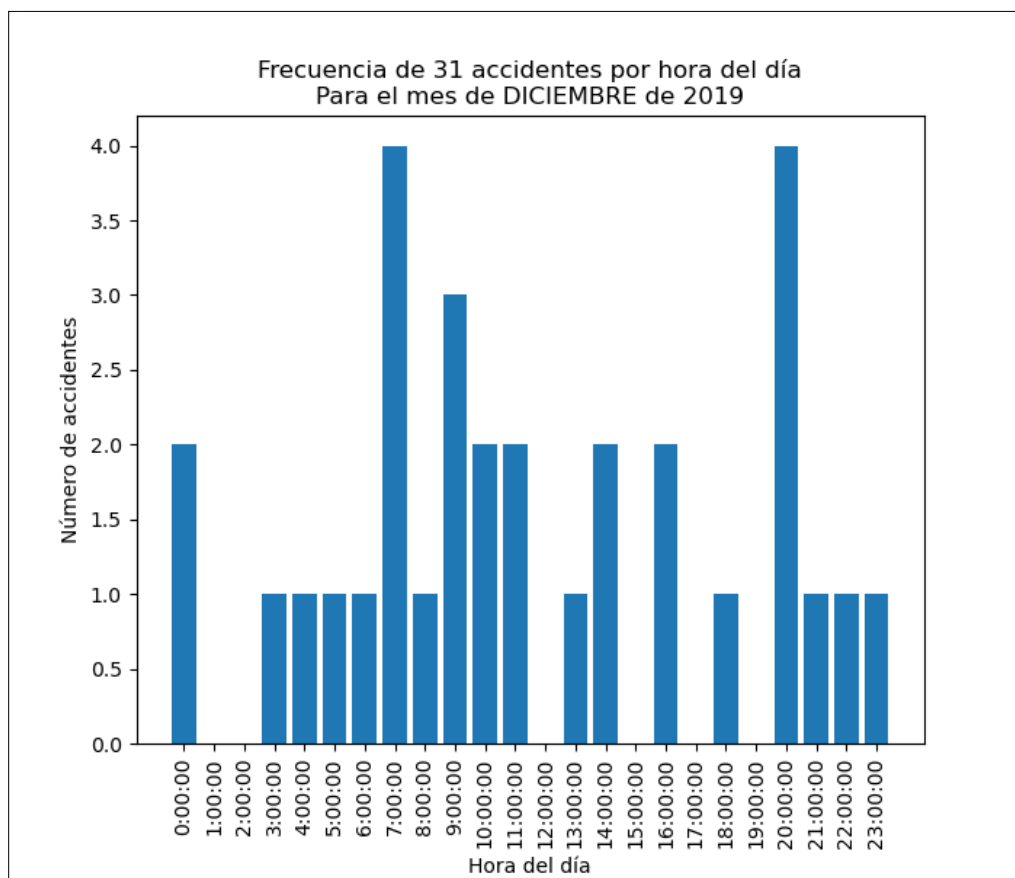
Accidentes más temprano y tardíos para el mes de DICIEMBRE de 2019

Accidentes del día 2019/12/2

CODIGO_ACCIDENTE	DIA_OCURRENCIA_ACC	DIRECCION	GRAVEDAD	CLASE_ACC	LOCALIDAD	FECHA_HORA_ACC	LATITUD	LONGITUD
10504223	LUNES	KR 69-CL 64	CON HERIDOS	CHOQUE	ENGATIVA	2019/12/02 09:49:59+00	4.67257	-74.0947
10504223	LUNES	KR 69-CL 64	CON HERIDOS	CHOQUE	ENGATIVA	2019/12/02 09:49:59+00	4.67257	-74.0947

Accidentes del día 2019/12/3

CODIGO_ACCIDENTE	DIA_OCURRENCIA_ACC	DIRECCION	GRAVEDAD	CLASE_ACC	LOCALIDAD	FECHA_HORA_ACC	LATITUD	LONGITUD
10504352	MARTES	CL 100-KR 9	CON MUERTOS	CHOQUE	USAQUEN	2019/12/03 06:40:00+00	4.68498	-74.048
10504346	MARTES	CL 132-KR 50	CON HERIDOS	CHOQUE	SUBA	2019/12/03 22:09:59+00	4.71849	-74.0564





## Requerimiento No. 8 (Bono Grupal): Visualizar todos los accidentes de una clase particular para un rango de fechas en el mapa de Bogotá

Como analista de la secretaría de movilidad **deseo** visualizar gráficamente en un mapa los puntos en los que ocurrieron accidentes por tipo para un rango de fechas especificado.

Los **parámetros de entrada** de este requerimiento son:

- Fecha inicial del intervalo.
- Fecha final del intervalo.
- Clase de accidente

La **respuesta esperada** debe contener:

- El número total de accidentes ocurridos en el rango de fecha.
- Un mapa interactivo de clústeres que muestre todos los accidentes según su tipo en el mapa de Bogotá, en donde cada marcador tenga asignado un color de acuerdo con su gravedad.

Para completar este requerimiento recomendamos utilizar la librería por extensión de Python llamada “folium” que se puede instalar en su ambiente por medio del comando “`pip install folium`”.

Para más información sobre esta librería dirigirse a los siguientes enlaces:

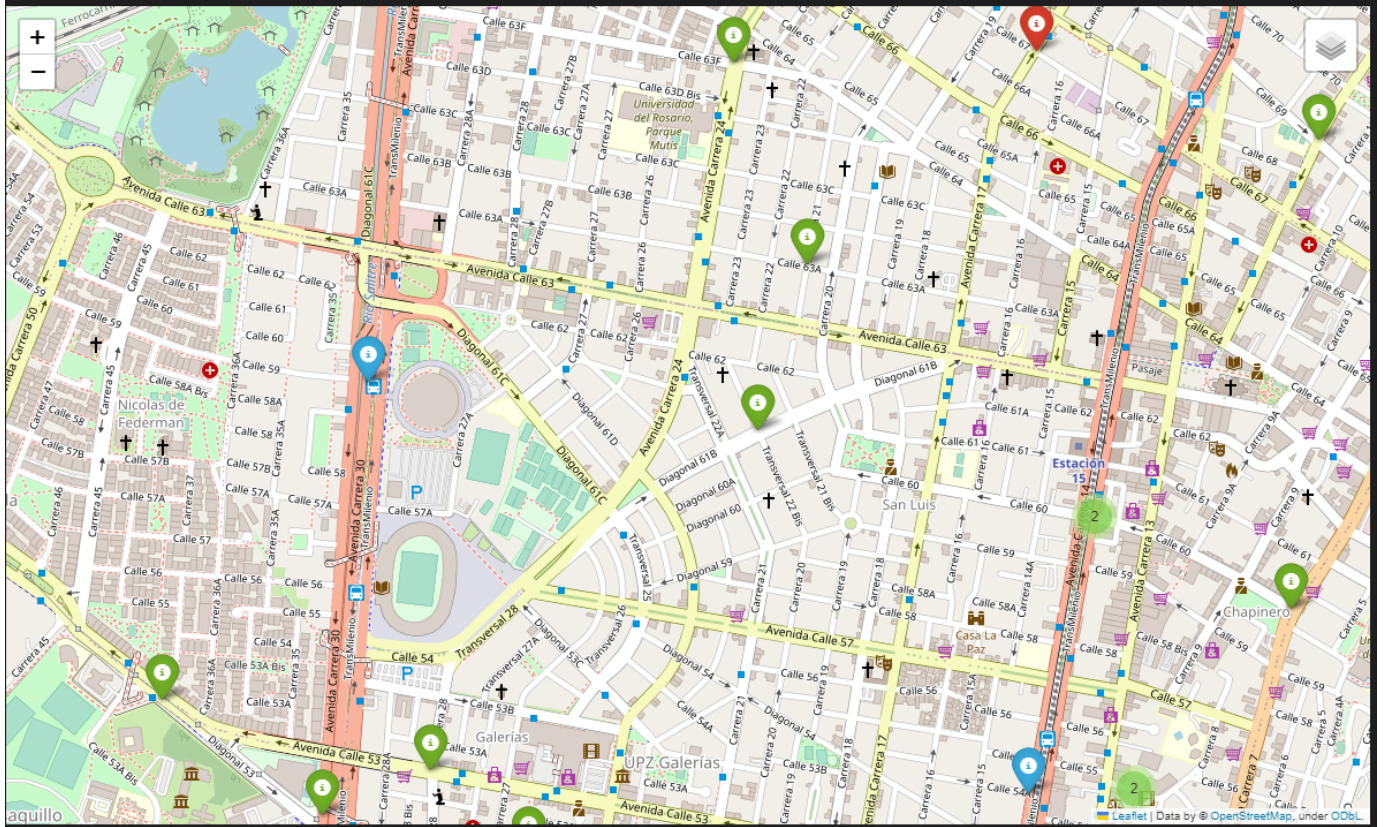
- Enlace oficial de PYPI, URL: <https://pypi.org/project/folium/>
- Enlace oficial de la librería, URL: <https://github.com/python-visualization/folium>
- Enlace al tutorial de la Librería. URL: <https://python-visualization.github.io/folium/quickstart.html>

Recomendación:

- **SOLAMENTE** puede utilizar librerías para la representación gráfica en el mapa de clústeres, todas las consultas al dataset deben realizarse utilizando **DISClib** y las **ADTs** disponibles en el mismo.

**EJEMPLO:** Deseo visualizar gráficamente los accidentes de clase choque que hayan ocurrido entre el 20/10/2016 y el 01/10/2018.

Req No. 8 Answer  
Hay 570 accidentes entre las fechas '20/10/2016' y '1/10/2018'



## Parte 4: Análisis de resultados

Dentro del proyecto debe incluir, en la carpeta **Docs**, un documento en formato **PDF** donde se evidencie el análisis de complejidad y las pruebas de tiempos de ejecución para cada requerimiento. Se sugiere que el documento tenga la siguiente distribución del contenido:

- Nombres, código y correo Uniandes de los integrantes del grupo.
- Para los requerimientos individuales se debe indicar que estudiante del equipo lo realizó.
- Un diagrama de las estructuras de datos que usó para resolver los diferentes requerimientos y que se crean al realizar la carga de datos.
- Descripción de las funciones y código implementado en la carga de datos y en cada uno de los requerimientos (incluye el bono si aplica).
- Análisis de complejidad temporal de cada uno de los requerimientos en **Notación O** (incluye el bono si aplica).
- Pruebas de tiempos de ejecución para cada uno de los requerimientos. En estas pruebas se deben incluir:
  - las tablas de tiempos de ejecución registrados para los requerimientos del 1 al 7 (incluye el bono si aplica).
  - Graficas de los tiempos de ejecución registrados para los requerimientos del 1 al 7 (incluye el bono si aplica).

- Un análisis de resultados comparándolo con la estimación en **Notación O**.

#### Recomendaciones:

- Tomen como guía las herramientas, metodología y análisis realizados en los laboratorios; en especial los laboratorios 4, 5 y 7.
- Ejecute las pruebas de los requerimientos siempre con los mismos parámetros de entrada y con los archivos más grande que pueda procesar su computador.
- Ejecute las pruebas de los requerimientos con la configuración optima del catálogo (LINKED\_LIST, ARRAY\_LIST, PROBING, CHAINING) y los algoritmos de ordenamiento (Shell, Insertion, Merge, etc.).

## Entrega

Para realizar la entrega del reto deben:

1. Agregar los usuarios de los monitores y profesores del curso a su organización de GitHub para hacer la entrega adecuada de la actividad,
2. Dar permisos adecuados repositorio a los monitores y al profesor, de lo contrario el taller **NO** podrá ser calificado,
3. Asegurarse que la visibilidad del repositorio entregado sea **privada** y que solo pueda accederse con los permisos configurados para los integrantes del grupo,
4. Enviar el enlace de GitHub en la actividad correspondiente dentro de Bloque Neón Uniandes (BrightSpace).
5. Incluir en el repositorio GIT todo el material, código y documentos solicitados durante la actividad.

**IMPORTANTE:** Recuerde que solo se calificará el material hasta el último **COMMIT** realizado previo a la fecha límite **(11:59 PM) del 26 de abril 2023**.