

# RETO 3: UFO Sightings

---

## 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

3 de noviembre antes de la media noche (11:59 p.m. hora Brightspace)

## Contexto

Por mucho tiempo se ha hablado de vida extraterrestre. Y hay muchos testimonios alrededor del mundo que hablan de avistamientos de Ovnis (Objetos Voladores No Identificados, UFO por sus siglas en inglés) e incluso de encuentros con extraterrestres. En este reto analizaremos un conjunto registros de avistamientos de Ovnis.

## Fuente de Datos

Hay dos conjuntos de datos públicos en formato CSV que están disponibles en el sitio Web UFO Sightings (<https://www.kaggle.com/NUFORC/ufo-sightings>). Sin embargo, vamos a proveer dos conjuntos de datos "limpios" que en principio tienen menos datos inconsistentes que los datos originales.

Estos datos están contenidos en los archivos **UFOS-utf8-small.csv**, **UFOS-utf8-large.csv** y en sus correspondientes subconjuntos del 5.0%, 10.0%, 20.0%, 30.0%, 50.0% y 80.0% localizados en el material del reto en el aula Bloque Neón Unificada. Cada registro de avistamiento cuenta con la siguiente información:

Propiedad	Descripción
<b>datetime</b>	Fecha y hora del avistamiento en formato DD/MM/AAAA HH: MM. Por ejemplo, 15/10/2008 21:00
<b>city</b>	Ciudad
<b>state</b>	Estado (aplica para USA y Canadá)
<b>country</b>	País (este campo puede estar vacío para avistamientos en países diferentes a USA y Canadá)
<b>shape</b>	Forma reportada del OVNI. Las posibles formas son: cylinder, light, circle, sphere, disk, fireball, unknown, oval, other, cigar, rectangle, chevron, triangle, formation, delta, changing, egg, diamond, flash, teardrop, cone, cross, pyramid, round, crescent, flare, hexagon, dome, changed
<b>duration (seconds)</b>	Duración en segundos
<b>duration (hours/min)</b>	Duración en formato variable: horas, minutos o un rango de los anteriores
<b>comments</b>	Breve descripción de lo visto por los testigos
<b>date posted</b>	Fecha y hora en formato MM/DD/AAAA de cuando se reportó este avistamiento en el conjunto de datos. Por ejemplo, 5/15/2013.
<b>latitude</b>	Coordenada geográfica de latitud
<b>longitude</b>	Coordenada geográfica de longitud

Tabla 1. Propiedades de información de un avistamiento.

## Trabajo Propuesto

### Parte 1: Configuración Repositorio

Complete los siguientes pasos para configurar su repositorio de trabajo:

- 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>
- Renombre el repositorio de su grupo con el esquema **Reto2-G<<Número del grupo>>** ej.: **Reto2-G01** para el grupo 1 de la sección 2.
- Edite el **README** del repositorio e incluya los nombres completos, correo Uniandes y códigos de los miembros del equipo de trabajo.
- Realice el procedimiento según lo aprendido en clase para clonar el repositorio en su máquina local y sincronizarlo con su repositorio en GitHub.
- Descargue los datos desde la sección unificada del curso y cópielos en la carpeta **data** del repositorio local.

## Parte 2: Carga de Datos

Para responder a los requerimientos presentados deberán cargar la información de uno de los archivos entregados; recuerde que solo se permite leer una vez la información de cada archivo.

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

- El total de avistamientos cargados.
- Mostrar los primeros 5 y últimos 5 avistamientos cargados con sus características.

**Nota:** Los ejemplos dados en el documento están hechos basados en el subconjunto de datos más pequeño ("\_small.csv").

## Parte 3: Desarrollo de los Requerimientos

### Requerimiento 1 (Grupal): Contar los avistamientos en una ciudad

Primero, se desea conocer el total de ciudades donde se han reportado avistamientos y segundo listar los avistamientos en una ciudad específica.

En caso de encontrar múltiples avistamientos en la misma ciudad mostrarlos ordenados cronológicamente por su fecha y hora (en formato AAAA-MM-DD HH:MM).

La entrada de este requerimiento es el nombre de la ciudad a consultar.

Y como respuesta debe presentar en consola la siguiente información:

- El total de avistamientos en la ciudad con la ciudad con más avistamientos.
- La información de los primeros 3 avistamientos y de los últimos 3 avistamientos obtenidos en la consulta (respetando el orden cronológico) incluyendo la siguiente información en cada uno de los registros:
  - La fecha y hora.
  - Ciudad, País.
  - Duración en segundos.
  - Forma del objeto.

**EJEMPLO:** Se quiere conocer el total de ciudades donde hay reportados avistamientos y los avistamientos en la ciudad de “las vegas”.

```
===== Req No. 1 Inputs =====
UFO Sightings in the city of: las vegas

===== Req No. 1 Answer =====
There are 679 different cities with UFO sightings...
The city with most UFO sightings is:

There are 7 sightings at the: las vegas city.
The first 3 and last 3 UFO sightings in the city are:
```

datetime	city	state	country	shape	duration (seconds)
1998-12-05 20:30:00	las vegas	nv	us	unknown	10
1999-07-19 21:30:00	las vegas	nv	us	oval	600
2000-01-07 16:20:00	las vegas	nv	us	unknown	600
2006-10-15 08:30:00	las vegas	nv	us	circle	1200
2007-03-04 11:00:00	las vegas	nv	us	circle	660
2013-04-05 08:30:00	las vegas	nv	us	formation	420

## Requerimiento 2 (Individual): Contar los avistamientos por duración

Primero, se desea conocer el número de avistamientos con la duración en segundos más larga que se tenga(n) registrado(s).

Segundo, dado un rango de duración en segundos (segundos-min, segundos-máx.). Mostrar los avistamientos ordenados ascendentemente por su duración y en caso de múltiples avistamientos de la misma duración, mostrarlos ordenados alfabéticamente por su combinación ciudad y país (country-city).

Las entradas de este requerimiento son los límites del rango en segundos

- Límite inferior en segundos (máximo).
- Límite superior en segundos (mínimo).

Y como respuesta debe presentar en consola la siguiente información:

- El total de avistamientos registrados con duración máxima
- La información de los tres primeros y tres últimos avistamientos dentro del rango, ordenados por su duración e incluyendo la siguiente información:
  - La fecha y hora.
  - Ciudad, País.
  - Duración en segundos.
  - Forma del objeto.

**EJEMPLO:** Se quiere conocer el/los avistamientos(s) con la duración en segundos más larga que se tenga registrados y los avistamientos con una duración entre 30.0 y 150.0 segundos.

===== Req No. 2 Inputs =====

UFO sightings between 30.0 and 150.0

===== Req No. 2 Answer =====

There are 58 different durations of UFO sightings...

The longest UFO sightings are:

duration (seconds)	count
172800	1

There are 193 sightings between: 30.0 and 150.0 duration.

The first 3 and last 3 UFO sightings in the duration time are:

datetime	city	state	country	shape	duration (seconds)
1997-08-22 23:30:00	atlantic city	wy	us	light	30
2006-12-05 22:00:00	bloomfield	nj	us	disk	30
2007-08-04 21:39:00	brockport	ny	us	light	30
1977-09-26 19:00:00	willimas lake (canada)	bc	ca	unknown	120
2008-10-28 17:55:00	iowa city	ia	us	unknown	150
2006-02-18 17:16:00	silverdale	wa	us	unknown	150

### Requerimiento 3 (Individual): Contar avistamientos por Hora/Minutos del día

Primero, se desea conocer los avistamientos con la hora y minuto del día (formato HH:MM) más tardíos que se tengan registrados. Y segundo, listar los avistamientos el rango de tiempo en el día [HH:MM inicial, HH:MM final]. Mostrar los avistamientos en orden cronológico por Hora/Minutos. En caso de múltiples avistamientos en una misma Hora/Minutos, mostrarlos ordenados cronológicamente por su fecha.

La entrada de este requerimiento son los límites de tiempo del rango

- Límite inferior en formato HH: MM.
- Límite superior en formato HH: MM.

Como respuesta debe presentar en consola la siguiente información:

- El total de avistamientos en el rango.
- La información de los tres primeros y tres últimos avistamientos dentro del rango, ordenados cronológicamente e incluyendo la siguiente información:
  - La fecha y hora.
  - Ciudad, País.
  - Duración en segundos.
  - Forma del objeto.

**EJEMPLO:** Se quiere conocer los avistamientos más tardíos que se tengan registrados y los avistamientos entre las 20:45 y 23:15.

===== Req No. 3 Inputs =====

UFO sightings between 20:45:00 and 23:15:00

===== Req No. 3 Answer =====

There are 249 UFO sightings with different times [hh:mm:ss]...

The latest UFO sightings time is:

```
+-----+-----+
| time   | count |
+=====+=====+
| 23:55:00 | 1 |
+-----+-----+
```

There are 291 sightings between: 20:45:00 and 23:15:00

The first 3 and last 3 UFO sightings in this time are:

```
+-----+-----+-----+-----+-----+-----+-----+
| datetime          | time   | city   | state | country | shape   | duration (seconds) |
+=====+=====+=====+=====+=====+=====+=====+
| 1999-03-04 20:45:00 | 20:45:00 | phoenix | az    | us      | other   | 300 |
+-----+-----+-----+-----+-----+-----+-----+
| 2003-09-26 20:45:00 | 20:45:00 | hudson  | fl    | us      | circle  | 900 |
+-----+-----+-----+-----+-----+-----+-----+
| 2003-10-17 20:45:00 | 20:45:00 | opp     | al    | us      | triangle | 900 |
+-----+-----+-----+-----+-----+-----+-----+
| 2008-07-17 23:15:00 | 23:15:00 | urbandale | ia    | us      | unknown | 30 |
+-----+-----+-----+-----+-----+-----+-----+
| 2010-11-14 23:15:00 | 23:15:00 | rockford | il    | us      | triangle | 4 |
+-----+-----+-----+-----+-----+-----+-----+
| 2011-10-26 23:15:00 | 23:15:00 | new delhi | Not Available | in      | fireball | 3600 |
|                   |         | (india)  |         |         |         |      |
+-----+-----+-----+-----+-----+-----+-----+
```



#### Requerimiento 4 (Grupal): Contar los avistamientos en un rango de fechas

Primero, se desea conocer el número de avistamientos con la fecha más antigua que se tengan registrados.

Segundo, listar los avistamientos registrados entre un rango de fechas del suceso (utilizando un formato AAAA-MM-DD).

La entrada de este requerimiento son los límites de fechas del rango

- Límite inferior en formato AAAA-MM-DD.
- Límite superior en formato AAAA-MM-DD.

Y como respuesta debe presentar en consola la siguiente información:

- El total de los avistamientos en el rango.
- La información de los tres primeros y tres últimos avistamientos dentro del rango, ordenados cronológicamente e incluyendo la siguiente información:
  - La fecha y hora.
  - Ciudad, País.
  - Duración en segundos.
  - Forma del objeto.

**EJEMPLO:** Se desea conocer los avistamientos ocurridos el 6 de agosto de 1945 (1945-08-06) y el 15 de noviembre de 1983 (1984-11-15).

```
===== Req No. 4 Inputs =====
UFO sightings between 1945-08-06 and 1984-11-15

===== Req No. 4 Answer =====
There are 750 UFO sightings with different dates [YYYY-MM-DD]...
The oldest UFO sightings date is:
+-----+-----+
| date      | count |
+-----+-----+
| 1944-07-15 |      1 |
+-----+-----+

There are 45 sightings between: 1945-08-06 and 1984-11-15
The first 3 and last 3 UFO sightings in this time are:
+-----+-----+-----+-----+-----+-----+-----+
| datetime          | date      | city      | state | country | shape  | duration (seconds) |
+-----+-----+-----+-----+-----+-----+-----+
| 1949-05-15 06:30:00 | 1949-05-15 | grays     | sc    | us      | light  | 300 |
+-----+-----+-----+-----+-----+-----+-----+
| 1951-06-15 20:30:00 | 1951-06-15 | greenville | ms    | us      | disk   | 300 |
+-----+-----+-----+-----+-----+-----+-----+
| 1954-07-07 22:00:00 | 1954-07-07 | west columbia | sc    | us      | formation | 600 |
+-----+-----+-----+-----+-----+-----+-----+
| 1983-06-15 19:30:00 | 1983-06-15 | phoenix   | az    | us      | sphere  | 300 |
+-----+-----+-----+-----+-----+-----+-----+
| 1984-07-31 19:00:00 | 1984-07-31 | rosebud   | mo    | us      | other   | 300 |
+-----+-----+-----+-----+-----+-----+-----+
| 1984-10-20 03:00:00 | 1984-10-20 | middle town | ca    | us      | diamond | 180 |
+-----+-----+-----+-----+-----+-----+-----+
```

## Requerimiento 5 (Grupal): Contar los avistamientos de una Zona Geográfica

Se desea conocer el número de avistamientos en una zona geográfica definida por un rango de longitudes [longitud-min, longitud-máx.] y latitudes [latitud-min, latitud-máx.]. Para ello coordenadas geográficas deben aproximarse con dos cifras decimales (ej.: -14.4095083 aproximado a -14.41).

Las entradas para este requerimiento son los rangos de longitudes y latitudes aproximadas a 2 cifras decimales determinadas por el usuario

- Límite máximo y mínimo de longitud (longitud-min y longitud-máx.).
- Límite máximo y mínimo de latitud (latitud-min y latitud-máx.).

**Ayuda:** Se propone utilizar una estructura de datos tipo mapa ordenado para representar los avistamientos ordenados por la coordenada longitud (aproximada a dos cifras decimales). Al interior de cada coordenada longitud representar los avistamientos que tengan dicha longitud. Estos avistamientos se pueden ordenar por la coordenada latitud (aproximada a dos cifras decimales).

Y como respuesta debe presentar en consola la siguiente información:

- El total de los avistamientos dentro del área definida.
- La información de los cinco primeros y cinco últimos avistamientos dentro del rango, ordenados por ubicación geográfica (latitud y longitud) e incluyendo la siguiente información:
  - La fecha y hora.
  - Ciudad, País.
  - Duración en segundos.
  - Forma del objeto.
  - Longitud y latitud del avistamiento.

**EJEMPLO:** Se desean conocer los avistamientos reportados en la zona de Nuevo México que se encuentra en una longitud desde los -103.00 a -109.05 y una latitud desde 31.33 a 37.00.

```
===== Req No. 5 Inputs =====
UFO sightings between latitude range of 31.33 and 37.0
plus longitude range of -103.0 and -109.05

===== Req No. 5 Answer =====
There are 6 different UFO sightings in the current area
The first 5 and last 5 UFO sightings in this time are:
+-----+-----+-----+-----+-----+-----+-----+-----+
| datetime | city | state | country | shape | duration (seconds) | latitude | longitude |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2002-12-14 05:00:00 | el paso | tx | us | unknown | 120 | 31.7586 | -106.486 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2004-08-23 21:00:00 | silver city | nm | us | light | 1800 | 32.77 | -108.28 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2004-08-15 00:16:00 | roswell | nm | us | disk | 780 | 33.3942 | -104.522 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 1997-02-05 22:00:00 | albuquerque | nm | us | flash | 180 | 35.0844 | -106.651 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2008-07-04 22:30:00 | albuquerque | nm | us | sphere | 4 | 35.0844 | -106.651 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2012-11-08 23:00:00 | tucumcari | nm | us | flash | 10 | 35.1717 | -103.724 |
+-----+-----+-----+-----+-----+-----+-----+-----+
```

## Requerimiento 6 (BONO Grupal): Visualizar los avistamientos de una zona geográfica.

Se desea visualizar la zona geográfica del requerimiento 5 y sus avistamientos reportados en un mapa.

Las entradas para este requerimiento son los rangos de longitudes y latitudes aproximadas a 2 cifras decimales determinadas por el usuario

- Límite máximo y mínimo de longitud (longitud-min y longitud-máx.).
- Límite máximo y mínimo de latitud (latitud-min y latitud-máx.).

Y como respuesta debe presentar en consola la siguiente información:

- El total de los avistamientos dentro del área definida.
- La información de los cinco primeros y cinco últimos avistamientos dentro del rango, ordenados por ubicación geográfica (latitud y longitud) e incluyendo la siguiente información:
  - La fecha y hora.
  - Ciudad, País.
  - Duración en segundos.
  - Forma del objeto.
  - Longitud y latitud del avistamiento.

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>

**EJEMPLO:** Se desean conocer los avistamientos reportados en la zona de Nuevo México que se encuentra en una longitud desde los -103.00 a -109.05 y una latitud desde 31.33 a 37.00.

## Parte 4: Análisis de resultados

Cree un archivo en formato **PDF** para la entrega y guárdelo en la carpeta **Docs** del repositorio, el documento debe contener las siguientes secciones:

- Nombres, código y correo Uniandes de los integrantes del grupo.
- Análisis de complejidad temporal en **Notación O** para cada uno de los requerimientos. Incluir una breve justificación de la complejidad temporal dada.
- Recuerde indicar quien implemento los requerimientos individuales

## Entrega

Permita el acceso a su **organización y repositorio** a los monitores y profesores del curso.

Envíe el enlace (URL) del repositorio por **BrightSpace** antes de la fecha límite de entrega.

Recuerden que cualquier documento solicitado durante la práctica debe incluirse dentro del repositorio **GIT** y solo se calificarán los entregables hasta el último **COMMIT** realizado previo a la media noche (**11:59 PM**) del **3 de noviembre de 2021**.