



## Trabajo Práctico Aéreos

### Introducción

Los bases de datos orientadas a grafos (GDBMS's) están diseñadas para resolver en forma eficiente consultas que involucran la *navegación* de un grafo en que los nodos representan entidades y los arcos representan interrelaciones entre dichas entidades. Por ejemplo, estas bases de datos se utilizan para resolver consultas de camino mínimo, de búsqueda de patrones, y en sistemas de recomendación, entre otras aplicaciones. Un ejemplo de base de datos que puede estructurarse como un grafo es la red internacional de vuelos y aeropuertos. En esta red los nodos representan aeropuertos del mundo (identificados con su código IATA de 3 letras) mientras que los arcos representan vuelos ofrecidos por distintas compañías. Las consultas típicas sobre este tipo de bases de datos están relacionadas con encontrar vuelos entre dos puntos sujeto a una serie de restricciones (fecha, duración, presupuesto, lugares disponibles, etc.).

### Objetivos del Trabajo Práctico

El objetivo principal es diseñar una base de datos en Neo4J que permita buscar itinerarios de vuelo en forma eficiente.

### Desarrollo del Trabajo Práctico

Se identifican las siguientes tareas que se deben cumplir:

1. Diseñar un esquema de base de datos que permita responder de forma eficiente los requerimientos del sistema. Se deberá utilizar Neo4J como SGBD.
2. Descargar los datos de vuelos y aeropuertos en formato CSV del Campus de la materia, e ingresarlos en Neo4J (Nota: Los datos provienen de <https://openflights.org/data.html>).
3. Instanciar los vuelos en forma diaria durante un mes, asignándoles atributos como: cantidad de asientos disponibles, horario de salida, duración, costo. Para extraer los horarios de vuelo, utilizar la API de Skyscanner (<https://rapidapi.com/skyscanner/api/skyscanner-flight-search>) o bien los resultados de Google Flights.

4. Desarrollar las consultas de búsqueda de itinerarios indicadas en este enunciado. Se deberá considerar que un itinerario puede constar de uno o más vuelos escalonados en distintas etapas.

Se deberá entregar un informe del trabajo práctico que contenga:

- Para cada integrante del grupo, rol asignado y división de tareas según el mismo.
- Detalle del modo de importar los datos a Neo4J que se utilizó: parámetros con los que se haya invocado a programas de conversión de datos o scripts en algún otro lenguaje de programación que se hayan invocado; cualquier tipo de post procesamiento que se haya realizado a los datos.
- Consultas en lenguaje CYPHER solicitadas: Se deberá mostrar un ejemplo de la consulta y el resultado de ejecutarla en la base de datos, indicando qué parte de la consulta es parametrizada y qué valores deben ser modificados.
- Detalle de análisis de performance realizados: Indicar qué se intentó mejorar, cómo se intentó hacerlo y qué resultados (positivos o negativos) se tuvo.

## **Requerimientos del sistema**

El sistema deberá cumplir con los siguientes requerimientos:

1- Para cada vuelo existente, se lo deberá instanciar durante 10 días consecutivos (es decir, cada arco del grafo deberá representar una salida concreta del vuelo en un día específico), con los horarios extraídos de Skyscanner o Google Flights para un día en particular. Para aquellos vuelos que no sean encontrados, se podrá asignar un horario de salida cualquiera durante el día, y una duración de vuelo proporcional a la distancia entre los aeropuertos, y suponiendo una velocidad media de 800km/h.

2- Para cada instancia de vuelo se deberá asignar una cantidad de asientos disponibles aleatoria entre 0 y 5, y un costo proporcional a la distancia recorrida.

## **Consultas a resolver:**

Garfield quiere irse de vacaciones a la paradisíaca isla de Zanzíbar en el Océano Índico.

- a) Encuentre un itinerario de modo que Garfield llegue lo antes posible a Zanzibar saliendo desde su hogar en “Indiana, IL” mañana:
- b) Encuentre un segundo itinerario que resuelva los siguientes problemas:
  - Asegúrese de que todos los vuelos cuenten con dos lugares disponibles, ya que Garfield quiere llevar también a Odie.

- Asegúrese de que el costo no supere el presupuesto de Garfield, que es de 2000 pesos por individuo (4000 pesos en total).
- Asegúrese de que el vuelo encontrado no pase por Gibraltar, ya que su pista de aterrizaje es una de las más peligrosas del mundo, y Garfield definitivamente no quiere pasar por allí.