**ACTIVIDAD 2**

**Tipo actividad: Obtención de datos y análisis con SQLite**

Para el presente taller se pueden trabajar con diversos conjuntos de datos de Kaggle, preferiblemente que estén en formato SQLite, como es el caso de la base de datos de vuelos (<https://www.kaggle.com/code/mpwolke/airlines-sqlite> ) o de futbolistas (<https://www.kaggle.com/datasets/hugomathien/soccer/code> ). Se debe orientar a los estudiantes a realizar un análisis exploratorio de los datos utilizando comandos de SQLite en DB Browser. (también es posible emplear el conector de Python para SQLite).

Para realizar el taller, agrupe los estudiantes y permítales escoger con qué conjunto de datos quieren trabajar. Si ellos quieren utilizar una alternativa diferente, otorgue 10 minutos para que busquen (debe ser un conjunto de datos en CSV o SQlite preferiblemente) Una vez todos tengan una base de datos resuelva con los estudiantes las siguientes preguntas: (las preguntas aplican para el set de datos de aviones y de futbolistas, en caso de tener uno diferente se deben adecuar).

La idea con las preguntas a responder es que los estudiantes trabajen diferentes operaciones con SQL. Puede complementar las preguntas dependiendo del dataset que hayan escogido los estudiantes. Algunas sugerencias son:

1. ¿Cuántas tablas tiene la base de datos? (uso de consultas select)
2. ¿Qué información hay en cada tabla? (uso de select en una tabla)

**ii**. Para el dataset de aviones responder:

1. ¿En el dataset de aviones: que tipos de asientos y tiquetes hay?
2. ¿Qué modelos de aviones hay?
3. ¿Cuántas reservas hay?
4. ¿Cuál es el precio medio de las reservas? ¿el más caro? ¿El más barato?
5. ¿Qué rangos de aeronaves hay? (ordenarlos descendentemente)

**iii.** Para el dataset de fútbol responder:

1. ¿Cuántos jugadores hay en la base de datos?
2. ¿Cuáles son los 10 primeros jugadores en la tabla player?
3. ¿Cuántos países hay? ¿Cuántos jugadores hay en cada país?
4. ¿Cuántas ligas hay?

**10.** ¿Cuántas ligas hay por país?

(puede tomar como referencia las consultas en el libro: <https://www.kaggle.com/code/dimarudov/data-analysis-using-sql/notebook> )