


# LABBOOK - DESARROLLO DEL DATA WAREHOUSE PARA HISTORIAL DE SPOTIFY

FRANCISCO JAVIER DELGADO RODRIGUEZ Y ALBERTO PALENCIA QUIÑONES  
BASES DE DATOS AVANZADAS  
ESCUELA SUPERIOR DE INFORMATICA



### **Día 1: Diseño del Data Warehouse**

- Se definieron los requerimientos del sistema.
- Se diseñó el esquema dimensional con sus respectivas tablas y relaciones.
- Se redactó el script SQL para la creación de la base de datos.

### **Día 2: Desarrollo del Esquema de Base de Datos**

- Se implementó la base de datos en MySQL utilizando DWH.sql.
- Se verificaron las claves primarias y foráneas.
- Se realizaron pruebas de inserción de datos manuales para verificar la integridad de las tablas.

### **Día 3: Desarrollo de la Aplicación Java**

- Se creó la conexión con MySQL utilizando JDBC.
- Se desarrolló la funcionalidad de lectura de archivos CSV.
- Se implementó la lógica para volcar datos a la base de datos.
- Se probaron las inserciones y se corrigieron errores de formato.

### **Día 4: Implementación de Consultas**

- Se diseñó el menú interactivo para ejecutar consultas sobre la base de datos.
- Se implementaron consultas para obtener artistas, álbumes, canciones, fechas y plataformas.
- Se realizaron pruebas de ejecución y optimización de consultas.

### **Día 5: Pruebas y Corrección de Errores**

- Se ejecutaron pruebas con datos reales.
- Se corrigieron errores en la inserción de datos y en las consultas.
- Se optimizó el código SQL y Java para mejorar el rendimiento.
- Se documentó el código y el esquema de la base de datos.

### **Anexo: Resumen del Tiempo Dedicado y Dificultades Encontradas**

#### **Tiempo Dedicado:**

- Diseño del esquema: 6 horas.
- Implementación de la base de datos: 5 horas.
- Desarrollo en Java: 8 horas.
- Implementación de consultas: 4 horas.
- Pruebas y corrección de errores: 6 horas.
- Documentación: 3 horas.

**Dificultades Encontradas:**

- Problemas con la conversión de formatos de fecha y hora al insertar datos.
- Recorte del archivo de líneas de datos del CSV por gran demora en el tiempo de compilación
- Diferencias en los nombres de las columnas entre el CSV y la base de datos.
- Manejo de claves foráneas en la inserción de datos.
- Optimización de consultas para mejorar la velocidad de respuesta.
- Debugging de la conexión JDBC debido a errores de credenciales.