

Base de datos de Spotify

Carlos Julián Lucero Tobar, David Hernando Sanchez Lombana

- <https://www.kaggle.com/datasets/asaniczka/top-spotify-songs-in-73-countries-daily-updated>

Contextualización y Planteamiento del Problema

Contextualización

En la actualidad, la industria musical se apoya ampliamente en plataformas de streaming como Spotify, que generan y gestionan grandes volúmenes de datos sobre álbumes, canciones, artistas y su desempeño en diferentes mercados. Estos datos incluyen información detallada sobre atributos musicales, popularidad, duración, contenido explícito y rankings regionales a lo largo del tiempo. Una base de datos bien estructurada es crucial para almacenar, organizar y analizar esta información, permitiendo a sellos discográficos, artistas y analistas tomar decisiones informadas y estratégicas.

Planteamiento del Problema

Las organizaciones enfrentan varios desafíos en la gestión de estos datos:

1. **Integridad y Consistencia:** Garantizar que la información sobre álbumes, canciones y artistas sea precisa y coherente.
2. **Complejidad de Relaciones:** Manejar relaciones complejas, como la asociación muchos a muchos entre canciones y artistas, sin redundancias.
3. **Acceso y Análisis Eficiente:** Facilitar la generación de informes y el análisis de tendencias musicales de manera rápida y efectiva.
4. **Escalabilidad:** Adaptar la base de datos al crecimiento continuo de datos y nuevas métricas sin interrupciones.
5. **Seguridad:** Proteger la información sensible y controlar el acceso para mantener la confidencialidad y cumplir con normativas.

Reglas de negocio Spotify

Identificación Única de Álbumes: Cada álbum debe tener un identificador único (album_id) que lo distinga de otros álbumes en la base de datos.

Nombre Obligatorio de Álbumes: El campo nombre en la tabla Álbumes es obligatorio y no puede estar vacío.

Fecha de Lanzamiento de Álbumes: Cada álbum debe tener una fecha de lanzamiento (fecha_lanzamiento) válida registrada.

Identificación Única de Canciones: Cada canción debe tener un identificador único de Spotify (spotify_id) que la distinga de otras canciones.

Nombre Obligatorio de Canciones: El campo nombre en la tabla Canciones es obligatorio y no puede estar vacío.

Popularidad de Canciones: La popularidad de una canción (popularidad) debe estar entre 0 y 100.

Indicador de Contenido Explícito: Cada canción debe indicar si contiene contenido explícito mediante el campo `es_explicito`.

Duración de Canciones: La duración de una canción (`duracion_ms`) debe registrarse en milisegundos y es obligatoria.

Atributos Musicales: Los atributos como `danceability`, `energy`, `speechiness`, `acousticness`, `instrumentalness`, `liveness` y `valence` deben estar entre 0.00 y 1.00, representando características musicales de la canción.

Relación Álbum-Canción: Cada canción debe estar asociada a un álbum existente mediante el campo `album_id`. Este campo es opcional, permitiendo canciones sin álbum asignado.

Identificación Única de Artistas: Cada artista debe tener un identificador único (`artista_id`) que lo distinga de otros artistas.

Nombre Obligatorio de Artistas: El campo `nombre` en la tabla `Artistas` es obligatorio y no puede estar vacío.

Relación Muchos a Muchos entre Canciones y Artistas: Una canción puede estar asociada a múltiples artistas y un artista puede participar en múltiples canciones, gestionado a través de la tabla `Cancion_Artista`.

Identificación Única de Países: Cada país debe tener un código de país único (`pais_codigo`) de dos caracteres que lo identifique de manera inequívoca.

Nombre Obligatorio de Países: El campo `nombre` en la tabla `Países` es obligatorio y no puede estar vacío.

Gestión de Snapshots: Cada snapshot (`fecha_snapshot`) representa una fecha específica para la captura de rankings y debe ser única.

Ranking Diario de Canciones: Cada ranking diario (`ranking_diario`) debe ser un valor positivo que indica la posición de una canción en un país específico en una fecha determinada.

Movimiento en el Ranking: Se puede registrar el movimiento diario (`movimiento_diario`) y semanal (`movimiento_semanal`) de una canción en el ranking, reflejando cambios en su posición.

Integridad Referencial en Rankings: Cada entrada en la tabla `Rankings` debe referenciar una canción existente, un país existente y un snapshot existente, garantizando la integridad de los datos.

Unicidad de Rankings por Fecha y País: Una misma canción no puede tener múltiples rankings en el mismo país y fecha de snapshot, asegurando que cada combinación de `spotify_id`, `pais_codigo` y `fecha_snapshot` sea única.

ENTIDAD-RELACIÓN (ACERCAR CON ZOOM)

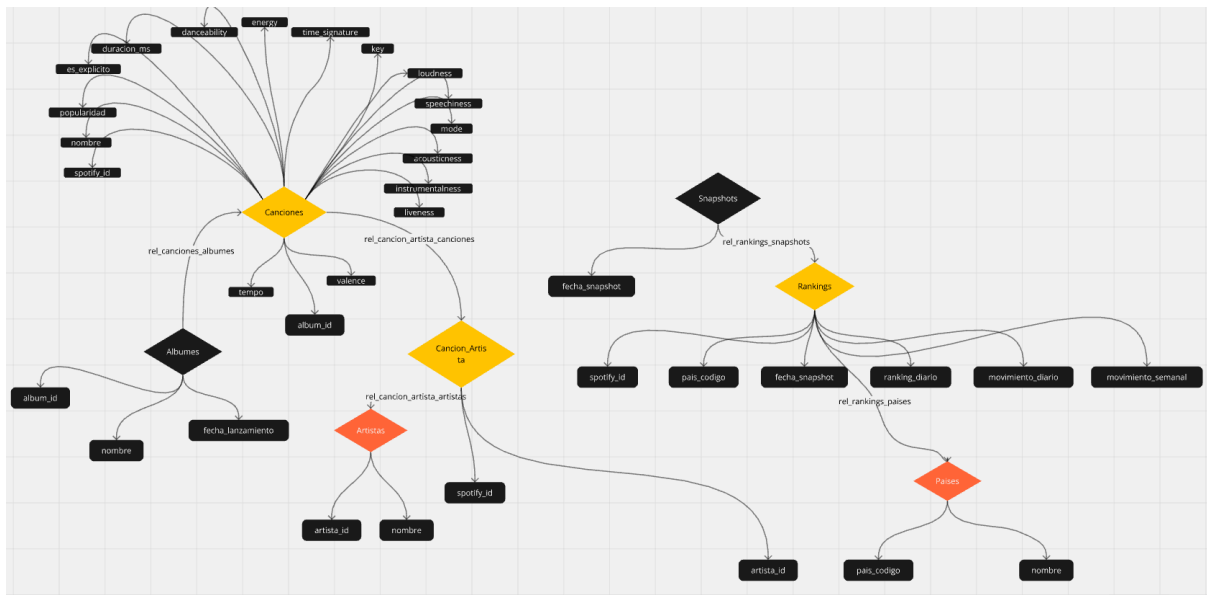
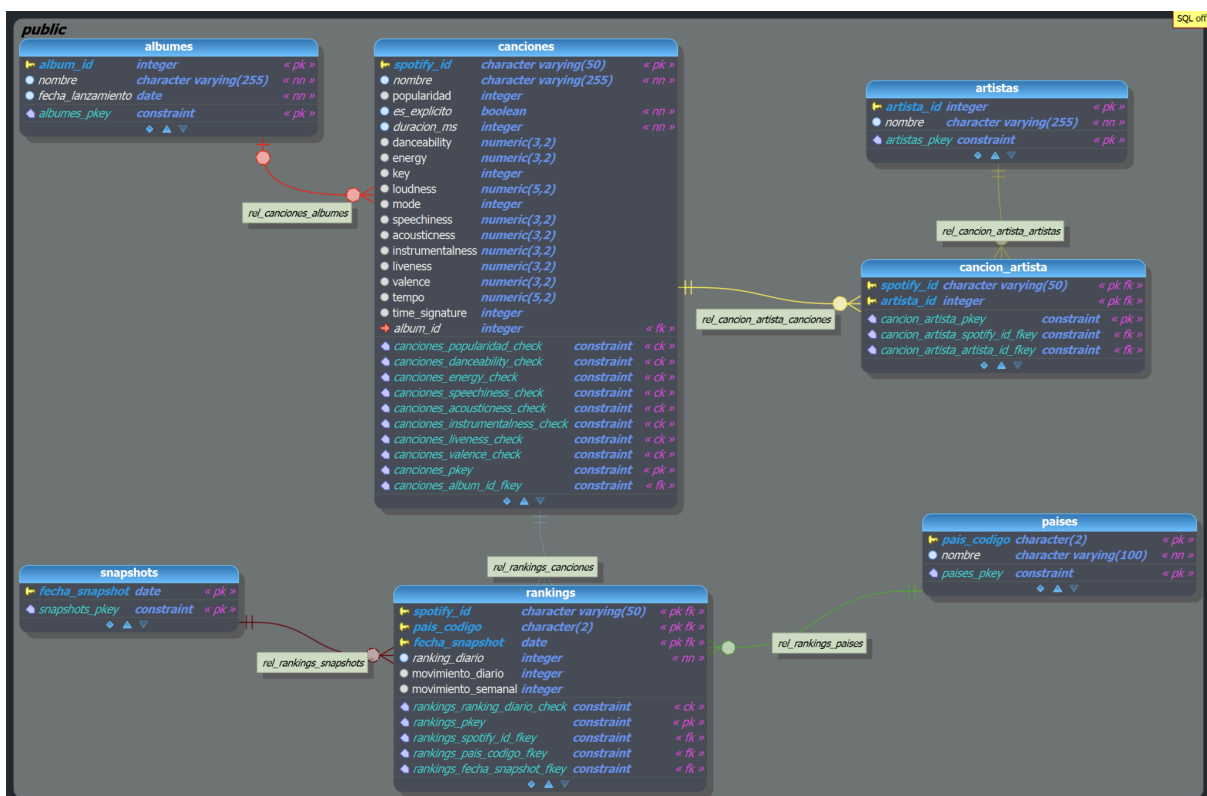


DIAGRAMA RELACIONAL



Entidades y Atributos

1. Albums

- **album_id**: Identificador único del álbum (Clave primaria).

- **nombre:** Nombre del álbum (Obligatorio).
- **fecha_lanzamiento:** Fecha de lanzamiento del álbum (Obligatorio).
- 2. **Canciones**
 - **spotify_id:** Identificador único de la canción en Spotify (Clave primaria).
 - **nombre:** Nombre de la canción (Obligatorio).
 - **popularidad:** Nivel de popularidad (0 a 100).
 - **es_explicito:** Indica si la canción contiene contenido explícito (Booleano).
 - **duracion_ms:** Duración de la canción en milisegundos (Obligatorio).
 - **danceability, energy, speechiness, acousticness, instrumentalness, liveness, valence:** Atributos musicales que varían entre 0.00 y 1.00.
 - **key, loudness, mode, tempo, time_signature:** Otros atributos musicales.
 - **album_id:** Referencia al álbum al que pertenece (Clave foránea).
- 3. **Artistas**
 - **artista_id:** Identificador único del artista (Clave primaria).
 - **nombre:** Nombre del artista (Obligatorio).
- 4. **Cancion_Artista**
 - **spotify_id:** Referencia a la canción (Clave foránea).
 - **artista_id:** Referencia al artista (Clave foránea).
 - **Clave primaria:** Combinación de **spotify_id** y **artista_id**.
- 5. **Países**
 - **pais_codigo:** Código único de país de dos caracteres (Clave primaria).
 - **nombre:** Nombre del país (Obligatorio).
- 6. **Snapshots**
 - **fecha_snapshot:** Fecha del snapshot (Clave primaria).
- 7. **Rankings**
 - **spotify_id:** Referencia a la canción (Clave foránea).
 - **pais_codigo:** Referencia al país (Clave foránea).
 - **fecha_snapshot:** Referencia al snapshot (Clave foránea).
 - **ranking_diario:** Posición diaria en el ranking (Obligatorio).
 - **movimiento_diario, movimiento_semanal:** Cambios en la posición del ranking.
 - **Clave primaria:** Combinación de **spotify_id**, **pais_codigo** y **fecha_snapshot**.

Relaciones

1. **Albúmenes y Canciones**
 - **Tipo:** Uno a Muchos.
 - **Descripción:** Un álbum puede contener múltiples canciones. Cada canción pertenece opcionalmente a un álbum específico.
2. **Canciones y Artistas**
 - **Tipo:** Muchos a Muchos.
 - **Descripción:** Una canción puede ser interpretada por varios artistas y un artista puede participar en múltiples canciones. Esta relación se gestiona a través de la tabla intermedia **Cancion_Artista**.
3. **Rankings, Canciones, Países y Snapshots**
 - **Tipo:** Relaciones múltiples.

- **Descripción:** La tabla Rankings relaciona una canción específica (spotify_id) con un país (pais_codigo) en una fecha determinada (fecha_snapshot), registrando su posición en el ranking diario y los movimientos correspondientes.

4. **Países y Rankings**

- **Tipo:** Uno a Muchos.
- **Descripción:** Un país puede tener múltiples entradas de rankings para diferentes canciones y fechas.

5. **Snapshots y Rankings**

- **Tipo:** Uno a Muchos.
- **Descripción:** Un snapshot (fecha específica) puede incluir múltiples rankings de diferentes canciones y países.