



MÓDULO 2: EXTRAER INFORMACIÓN DE LA WEB Y BASES DE DATOS >  
LECCIONES

Copy ▾

# Proyecto MusicStream

## Proyecto: Análisis de Popularidad de Canciones en la Era Digital

- Recuerda que este proyecto es un ejercicio académico, ante la duda pregunta a tus profesores.

### 1. Resumen

La plataforma de streaming **MusicStream** busca entender mejor las tendencias musicales y mejorar la experiencia de sus usuarios. Este proyecto se enfocará en analizar la popularidad de canciones y álbumes desde el año 2000 hasta la fecha. El análisis se basará en criterios como las calificaciones, el número de reproducciones, las reseñas de los usuarios, etc.

El objetivo es identificar las canciones y álbumes más populares en **MusicStream**, utilizando técnicas de extracción de datos desde diferentes fuentes, como la API de Spotify, y last.fm. Posteriormente, almacenaremos la información en una base de datos y realizaremos consultas para obtener insights clave.

## 1.1 Cómo Enfrentar el Proyecto

- **Dividan el trabajo:** Asignen tareas según los intereses y habilidades de cada una.
- **Lean la documentación de las APIs:** Revisen cómo funcionan las APIs de Spotify y Last.fm antes de programar.
- **Usen ejemplos de código:** Adaptar ejemplos de la documentación puede facilitar el trabajo.
- **Prueben con datos pequeños:** Extraigan una muestra antes de obtener grandes volúmenes de datos.
- **Verifiquen la información:** Asegúrense de que los datos obtenidos sean correctos.
- **Código limpio y organizado:** Usen comentarios y nombres de variables descriptivos.
- **Hagan copias de seguridad:** Guarden su trabajo en GitHub con frecuencia.
- **Practiquen consultas SQL:** Antes de escribir las consultas finales, prueben con conjuntos pequeños de datos.
- **Ensayan la presentación:** Hagan una prueba antes del día final para ganar confianza.

## 2. Objetivos

- Consolidar los conocimientos de Python y SQL.
- Implementar control de versiones en equipo, usando Git y GitHub.
- Usar Scrum como marco de referencia, aplicando principios ágiles.
- Mejorar la comunicación y presentación de resultados en equipo.

## 3. Fases del Proyecto

### Fase 1: Extracción de Datos

Trabajarás con una muestra de datos en lugar de realizar la extracción completa de toda la información (esto se aclara en el siguiente apartado).

#### 1.1. Extracción de Datos desde la API de Spotify.

En esta fase, usarás la API de Spotify para obtener información sobre canciones, álbumes, géneros y artistas. La documentación de la API está disponible aquí: [Spotify Web API ↗](#).

- Ejemplo de endpoint: <https://api.spotify.com/v1/albums/{id}> ↗

A modo de ejemplo, este endpoint proporciona detalles sobre un álbum específico, incluyendo nombre, año de lanzamiento, y artistas involucrados.

Tareas:

1. Extraer canciones y álbumes lanzados entre 2000 y 2024.
2. Elegir un rango de 5 años para extraer datos (por ejemplo de 2002 a 2007).
3. Seleccionar 4 géneros musicales para el desarrollo de este proyecto.
4. Extraer la siguiente información:
  - Artista.
  - Género musical.
  - Tipo (canción o álbum).
  - Nombre.
  - Año de lanzamiento.
  - ID.

#### 1.2. Extracción de Detalles de Canciones/Álbumes con la API de last.fm.

Por ultimo, usaremos la API de last.fm, que nos permite completar aun mas información. Recuerda que habíamos dicho que íbamos a centrarnos en la popularidad, por lo que aquí es relevante el uso de esta API que puedes encontrarla disponible aquí: [last.fm ↗](https://last.fm/api).

Tareas:

1. Extraer detalles como:
  - Biografía del Artista.
  - Popularidad y Estadísticas de Reproducción.
  - Artistas Similares.

## Fase 2: Organización y Almacenamiento de Datos

Una vez que hayas reunido toda la información necesaria, es hora de pensar en la organización. Usando SQL debes pensar en la estructura que debe tener las bases de datos para almacenar toda la información recopilada y crear todas las tablas y conexiones entre ellas.

Una vez que tengas la estructura de la base de datos lista, debes insertar todos los datos en las bases de datos .

## Fase 3: Análisis y Consultas de Datos

Realiza consultas SQL para responder a las siguientes preguntas:

- ¿Cuál es el artista con más albums?
- ¿Qué género es el mejor valorado?
- ¿En qué año se lanzaron más álbumes?
- ¿Cuál es la canción mejor valorada?
- ¿Cuál es el artista con más valoración?
- ¿Cuál es el album más valorado de los años pares de mi selección?
- ¿Qué país tiene más artistas? (ordenar por popularidad)
- ¿Qué artista estuvo más tiempo y cuántos albums tiene?

## 4. Planificación del Proyecto

Para la realización de este proyecto trabajaremos en sprints (Consulta tu calendario de Google para saber cuántos sprints tienes). Siguiendo los principios ágiles, estableceremos pequeños ciclos iterativos de forma que al final de cada uno generaremos valor perceptible por nuestros usuarios.

Dedicaremos el primer día a la planificación del sprint (sprint planning) y el resto a trabajar en el desarrollo del proyecto. Al final de cada sprint haremos un Sprint Review del proyecto para presentar los resultados conseguidos y recoger feedback.

También haremos una retro corta revisando los working agreements que hemos acordado al inicio del proyecto y añadiendo cualquier otro feedback que nos permita mejorar el proyecto.

Al final del proyecto (que coincidirá con el final del módulo), haremos una sesión de presentación más completa, más allá de lo que sería un Sprint Review.

### Planificación de sesiones

Para optimizar el tiempo, podríamos considerar dividir las tareas o formar subgrupos que trabajen simultáneamente en distintas actividades.

- **Sesión 1:** Planificación del Proyecto
  1. Elegir a la scrum master.
  2. Crear el repositorio en GitHub: repositorio en GitHub y con acceso para todos los miembros del equipo..
  3. Establecer los horarios de trabajo y reuniones.
  4. Crear el tablero de tareas del proyecto y las responsabilidades.
  5. Investigar las APIs a utilizar y analizar qué información es relevante para resolver el proyecto.
- **Sesión 2, Sesión 3, Sesión 4, Sesión 5:** Extracción de Datos

1. Extraer datos de la API de Spotify.
  2. Extraer datos de la API de last.fm.
  3. Almacenar los datos extraídos en una estructura temporal antes de insertarlos en la base de datos.
  4. Documentar el proceso de extracción de datos y cualquier desafío encontrado para el Sprint Review
- **Sesión 6, Sesión 7, Sesión 8:** Creación y Almacenamiento de Datos
    1. Diseñar la Base de Datos: Definir las tablas necesarias y sus relaciones (por ejemplo, género, artistas, países, premios, etc según corresponda).
    2. Escribir scripts de Python/SQL para crear la base de datos y las tablas e insertar datos en la base de datos.
    3. Insertar los datos extraídos en la base de datos.
    4. Validar que los datos se han insertado correctamente y que no hay duplicados ni errores.
  - **Sesión 7, Sesión 8, Sesión 9:** Análisis de Resultados, Presentación y Repo Documentado
    1. Escribir y ejecutar consultas SQL para responder a preguntas específicas (por ejemplo, ¿qué género tiene más artistas?).
    2. Analizar los resultados de las consultas y verificar su exactitud.
    3. Documentar las consultas realizadas y los hallazgos obtenidos.
  - 1. Documentar el proceso y cualquier desafío encontrado para el Sprint Review
- **Durante las entrevistas técnicas:** Presentación Final
    1. Preparar una presentación que incluya una visión general del proyecto, metodologías utilizadas, y resultados clave.
    2. Ensayar la presentación y recibir retroalimentación para mejorarla antes de la sesión final.
    3. Tener en el repositorio de GitHub todo el código del desarrollo del proyecto.

## Criterios de Aceptación Generales

- Crear la infraestructura necesaria: repositorio en GitHub y con acceso para todos los miembros del equipo.
- Extraer datos de distintas fuentes de datos y creación de una Base de Datos.
- Tener la extracción de datos, creación de la Base de Datos e inserción de datos automatizada en funciones (obligatorio) y clases (optativo).
- Tener en el repositorio de GitHub todo el código del desarrollo del proyecto.

## 5. Historias de Usuario

### 1. Selección de Herramientas para el proyecto.

**Historia de Usuario:** Como desarrollador de **MusicStream**, quiero seleccionar las herramientas adecuadas para realizar el proyecto efectivamente, de modo que pueda asegurarme de que la información recopilada sea precisa y completa.

Criterios de Aceptación:

- Identificar y evaluar las potenciales APIs para mi proyecto.
- Seleccionar herramientas que cumplan con los requisitos del proyecto (como por ejemplo, Python, MySQL u otras herramientas que se consideren necesarias para abordar el proyecto).
- Documentar el proceso de selección y justificación de las herramientas elegidas.

## 2. Extracción de Datos de APIs

**Historia de Usuario:** Como analista de datos, quiero extraer información relevante de APIs para asegurarme de que tengo todos los datos necesarios para el análisis posterior.

Criterios de Aceptación:

- Configurar y realizar solicitudes a la API de Spotify, MusicBrainz y last.fm para obtener datos iniciales.
- Almacenar los datos extraídos en una estructura temporal antes de insertarlos en la base de datos (por ejemplo en CSV).
- Documentar el proceso de extracción de datos y cualquier desafío encontrado.

## 3. Decisión de la Estructura de la Base de Datos

**Historia de Usuario:** Como arquitecto de bases de datos, quiero diseñar una estructura de base de datos que pueda almacenar de manera eficiente toda la información recolectada, para que los datos puedan ser fácilmente consultados y analizados.

Criterios de Aceptación:

- Crear un diagrama ER (Entidad-Relación) detallado de la base de datos.
- Definir las tablas necesarias y sus relaciones (por ejemplo, género, artistas, países, premios, etc según corresponda).
- Documentar la estructura de la base de datos y justificar las decisiones de diseño.

## 4. Inserción de Datos en la Base de Datos

**Historia de Usuario:** Como desarrollador de bases de datos, quiero insertar todos los datos recolectados en la base de datos diseñada, para que la información esté organizada y sea accesible para consultas.

Criterios de Aceptación:

- Escribir scripts de Python/SQL para insertar datos en la base de datos.
- Validar que los datos se han insertado correctamente y que no hay duplicados ni errores.
- Documentar el proceso de inserción de datos y cualquier ajuste necesario en la estructura de la base de datos.

## 5. Realización de Consultas para Análisis de Datos

**Historia de Usuario:** Como analista de datos, quiero realizar consultas SQL para extraer información específica de la base de datos, para responder a preguntas clave, como por ejemplo las tendencias y popularidad de las canciones o artistas.

Criterios de Aceptación:

- Escribir y ejecutar consultas SQL para responder a preguntas específicas (por ejemplo, ¿qué género tiene mas artistas?).
- Analizar los resultados de las consultas y verificar su exactitud.
- Documentar las consultas realizadas y los hallazgos obtenidos.

## 6. Presentación de Resultados

**Historia de Usuario:** Como miembro del equipo de **MusicStream**, quiero presentar los resultados del análisis de datos de manera clara y efectiva, para que las partes interesadas puedan entender y utilizar la información para la toma de decisiones.

Criterios de Aceptación:

- Preparar una presentación que incluya una visión general del proyecto, metodologías utilizadas, y resultados clave.
- Ensayar la presentación y recibir retroalimentación para mejorarla antes de la sesión final.

## 6. Entrega

El formato de entrega de este proyecto será mediante la carga del mismo a la plataforma de GitHub. Se creará un repositorio. El nombre del repositorio deberá estar compuesto de las siguientes partes, todo ello separado por guiones:

- La palabra: proyecto-da.
- Número de la promoción: promo-numero.
- Número del módulo: modulo-2.
- Número del equipo: team-numero.

Por ejemplo:

- da-project-promo-123-modulo-2-team-123
- da-project-promo-47-modulo-2-team-3

En lo relacionado con las fechas de los sprints y las entregas, revisar vuestro calendario de clase. En las sprint review se revisará que se hayan solucionado todas las tareas técnicas asociadas a la entrega de esos sprints; si algo quedara pendiente, se arrastraría al siguiente sprint.

## 8. Presentación

El último día del módulo presentaréis la versión final de este proyecto a vuestras compañeras y el docente. Cada equipo realizará una presentación de 10 minutos y posteriormente habrá 5 minutos de feedback por parte del público.

El objetivo es que practiquéis la realización de las demos de los proyectos que habéis desarrollado, explicándolo desde un punto de vista técnico y también desde la perspectiva de producto, mejorando además vuestras habilidades de exposición, objetivo de desarrollo profesional del curso.

Para que la presentación salga bien es imprescindible una buena preparación. Por ello, durante el segundo sprint del módulo tendréis que asignar responsabilidades dentro del equipo relacionadas con la preparación de ésta. Algunos tips para prepararlos este demo son:

- Todas las participantes del equipo deben hablar en la presentación (sin práctica no hay mejora).
- Identificar los objetivos de la presentación: Debemos tener claro qué es lo que queremos lograr con la presentación. ¿Queremos demostrar la funcionalidad de un producto? ¿Queremos mostrar los resultados de un experimento? ¿Queremos atraer inversores? Dependiendo del objetivo, deberemos enfocar la presentación de manera distinta.
- Conoce a tu audiencia: La presentación debe estar adaptada al tipo de audiencia que se espera. Si se presentará ante posibles inversores, la presentación debe estar enfocada en los beneficios y la rentabilidad del producto. Si es una presentación para usuarios, deberá enfocarse en la usabilidad y la facilidad de uso.
- Debéis ser claras y concisas: La presentación debe ser fácil de entender y no debe ser demasiado larga. Es importante presentar la información de manera clara y concisa. Debemos tener en cuenta que los detalles técnicos pueden ser interesantes, pero no deberían opacar la idea principal.
- Demostrad la funcionalidad del producto: Si la presentación es para demostrar un producto, es importante que demostremos su funcionalidad. Podemos hacerlo mediante un video o demostrando el producto en vivo. Es importante asegurarse de que el producto funciona correctamente antes de la presentación.
- Resaltad los aspectos más importantes: En cualquier presentación, hay aspectos que son más importantes que otros. Debéis resaltar los aspectos que sean más relevantes para su objetivo. Por ejemplo, si queréis atraer inversores, deberéis resaltar los beneficios y la rentabilidad del producto.
- Practicad la presentación: Es importante que practiquéis la presentación varias veces antes del evento. Debéis aseguraos de que la presentación esté bien estructurada y que os sentís cómodas hablando frente a la audiencia.

- Preparad respuestas a preguntas frecuentes: Es probable que la audiencia tenga preguntas después de la presentación. Debéis preparaos para responder preguntas frecuentes y tener la información necesaria a mano.

Además de esto, para mejorar vuestras habilidades de exposición en público y hacer la presentación más rica, podréis incorporar otros elementos adicionales (son solo ideas, sentíos libres de innovar y ser creativas):

- Dejar muy claro quién ha sido vuestro cliente y qué fue lo que os pidió.
- Explicar qué necesidades cubre o qué problemas soluciona el producto, cuál es el beneficio principal que aporta y qué lo hace único comparado con otros productos parecidos del mercado.
- Aportaciones "únicas y diferenciadoras" de cada equipo al proyecto.
- Cómo ha sido la organización del equipo, el reparto de tareas y la coordinación a la hora de trabajar todas en el mismo código.
- Cuál de las tareas o los puntos ha sido el que más esfuerzo ha requerido.
- Cuál de las tareas o partes del producto es la que hace que el equipo esté más orgulloso.
- Las tecnologías qué habéis utilizado y para qué sirven, y algunas partes del código que habéis desarrollado que merezca la pena resaltar.

La presentación debe tener un "buen inicio y un buen cierre" que nos haga a todos estar atentos y aplaudir... ahí os dejamos que echéis a volar vuestra imaginación.

- No habléis en primera persona de lo que habéis hecho, hablad del equipo.
- No mencionéis problemas, sino "retos" que os han hecho aprender y crecer.

Previous  
Ejercicios

Next  
Repasos

Last updated 12 months ago