República Bolivariana de Venezuela Universidad Nacional Experimental de Guayana Ingeniería informática

Asignatura: BASE DE DATOS II

Sección: 1

PROYECTO Nº 1 BASES DE DATOS BASADAS EN COLUMNAS APACHE CASSANDRA

Docente: Integrantes:

CLINIA CORDERO MANSOUR ORTEGA

HECTOR VASQUEZ

LUIS ZAMORA

Puerto Ordaz 06 de junio del 2025

Introducción

Este informe explica cómo se ideó, se armó y se puso en marcha un sistema sencillo para dar sugerencias de música. Se usaron cosas modernas y se pensó en enseñar cómo funciona. La idea junta una base de datos diferente (Apache Cassandra), un motor hecho con Node. js y una página web fácil de usar.

Este plan es como una forma fácil de entender cómo funciona el negocio, usando preguntas juntas que parecen funciones OLAP (Procesamiento Analítico en Línea). Estas se enfocan en cosas importantes como el tipo de música y el tiempo. Para ello, se usó una forma de armar el sistema que es como piezas, de tal forma sería más fácil de cambiar y crecer.

El trabajo se partió en tres momentos claros: pensar cómo sería y cómo funcionaría, hacerlo de verdad y, al final, revisarlo y escribir cómo se hizo. Se hicieron cosas como dibujar cómo serían los datos, preparar el lugar para trabajar con Docker Compose, meter datos desde archivos CSV, construir partes para dar recomendaciones y hacer preguntas OLAP básicas.

Este informe incluye cada una de las partes del sistema, dando una explicación al porqué se eligieron, cómo están organizadas, cómo funcionan en general y cómo se ven los resultados.

Objetivo General

Diseñar e implementar un sistema de recomendación de música con un análisis OLAP simplificado, utilizando una base de datos NoSQL (Cassandra).

Objetivos Específicos

Modelado de Datos NoSQL Simplificado

- Diseñar un esquema NoSQL para almacenar información básica sobre usuarios, canciones y escuchas.
- o Elegir Cassandra como base de datos NoSQL.
- Justificar la elección basada en la escalabilidad y eficiencia para manejar grandes volúmenes de datos.

Implementación de Recomendación Básica

 Implementar un algoritmo simple basado en "Canciones más escuchadas por género" o "Canciones más escuchadas en la misma ciudad del usuario".

Análisis OLAP Muy Simplificado

- o Identificar dos dimensiones clave: usuario y tiempo.
- o Definir una jerarquía simple para la dimensión tiempo: día → mes.
- Definir medidas como número de escuchas.
- o Diseñar un cubo OLAP conceptual para análisis.

Implementación de Consultas OLAP Básicas

- Implementar consultas como "Número total de escuchas por género y mes".
- Mostrar resultados mediante visualizaciones simples (tablas o gráficos).

Dataset Utilizado

El sistema se alimenta inicialmente de tres archivos CSV proporcionados por el docente, con datos estructurados en torno a canciones, usuarios y escuchas. Se recomienda ampliar cada archivo a 100 registros para análisis más robustos.

Canciones (canciones.csv)

cancion_id,titulo,artista,genero
1,Bohemian Rhapsody,Queen,Rock
2,Like a Prayer,Madonna,Pop
3,Smells Like Teen Spirit,Nirvana,Grunge
4,Hey Jude,The Beatles,Rock
5,Billie Jean,Michael Jackson,Pop

Usuarios (usuarios.csv)

usuario_id,nombre,ciudad 101,Ana Perez,Caracas 102,Carlos Gomez,Bogota 103,Laura Torres,Mexico DF 104,Javier Ruiz,Caracas 105,Sofia Martinez,Bogota

Escuchas (escuchas.csv)

usuario_id,cancion_id,fecha_escucha 101,1,2024-01-05 102,2,2024-01-05 101,3,2024-01-10 103,1,2024-01-12 104,4,2024-01-15 105,5,2024-01-20 102,1,2024-01-25 103,4,2024-02-01

Estos datos iniciales permiten probar y validar la funcionalidad básica del sistema. La expansión a 100 registros por archivo facilitará evaluaciones más detalladas y significativas en la etapa de pruebas.

Justificación Técnica

Elección de Cassandra como base de datos NoSQL

Cassandra fue seleccionada como el núcleo de datos, gracias a su capacidad sobresaliente para administrar cantidades enormes de información de manera descentralizada y con la posibilidad de crecer sin límites. Dentro del marco del sistema para recomendar música, se requiere una estructura capaz de aguantar una gran cantidad de adiciones y extracciones de datos, algo común debido al registro incesante de reproducciones y las indagaciones analíticas. Además, Cassandra facilita la réplica automática y asegura una disponibilidad continua, cualidades esenciales para una aplicación actual.

Uso de Docker Compose

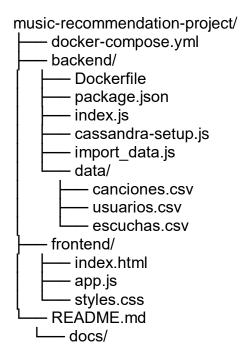
Para organizar el ambiente de desarrollo, nos apoyamos en Docker Compose. Con él, fue sencillo activar y vincular los contenedores que requeríamos, como el servidor de Cassandra y el 'backend' de Node. js. Esto nos dio la tranquilidad de tener una configuración que se puede replicar y llevar a cualquier parte, simplificando el desarrollo y las implementaciones que vengan. Aparte, nos da la certeza de que los ambientes locales y los de producción serán iguales, disminuyendo los problemas de configuración.

Uso de Node.js

Se optó por Node. js para el backend porque maneja de maravilla las operaciones de entrada/salida, además de que se vincula sin problemas con Cassandra gracias a librerías como cassandra-driver. Gracias a esta unión de tecnologías, el proyecto puede presumir de un diseño que es modular, eficiente y, sobre todo, muy escalable.

Estructura del Proyecto

El proyecto se organiza en una estructura clara que separa los componentes principales del sistema en backend, frontend y configuración de entorno.



- **docker-compose.yml**: Orquesta los servicios necesarios para la aplicación (Cassandra y el backend de Node.js).
- backend/: Contiene toda la lógica del servidor, conexión a Cassandra y carga de datos.
 - o index.js: Punto de entrada del servidor backend.
 - cassandra-setup.js: Encargado de crear las tablas necesarias en Cassandra.
 - import_data.js: Automatiza la carga de datos desde los archivos CSV a la base de datos.
 - data/: Contiene los archivos CSV con datos iniciales de usuarios, canciones y escuchas.
- frontend/: Interfaz básica de usuario para visualizar consultas OLAP.
 - o index.html: Estructura básica del sitio web.
 - app.js: Lógica de cliente que realiza peticiones al backend y renderiza resultados.
 - o styles.css: Estilos simples para la visualización.
- README.md y docs/: Documentación del proyecto para facilitar instalación y uso.

Esta organización modular permite mantener un desarrollo ordenado y facilita futuras extensiones del sistema.

Casos de Uso

A continuación, se presentan tres casos de uso básicos que orientaron el diseño e implementación del sistema:

1. Registrar escucha de una canción

- Actor: Usuario
- Descripción: El usuario reproduce una canción. El sistema almacena la escucha junto con la fecha, el usuario y la canción correspondiente.

2. Recomendar canciones por ciudad

- o Actor: Usuario
- Descripción: El sistema ofrece una lista de las canciones más escuchadas en la ciudad del usuario.

3. Consultar número de escuchas por género y mes

- Actor: Administrador o Analista
- Descripción: El sistema devuelve el total de escuchas agrupadas por género y mes para propósitos analíticos.

Estos casos de uso permitieron mantener el enfoque en funcionalidades concretas, útiles y realistas dentro del alcance del proyecto.

Modelo de Datos NoSQL

El modelo de datos fue diseñado teniendo en cuenta la simplicidad y eficiencia de consulta, estructurando los datos en tres tablas principales:

1. usuarios

- usuario_id (UUID)
- nombre (TEXT)
- ciudad (TEXT)

2. canciones

- cancion_id (UUID)
- titulo (TEXT)
- artista (TEXT)
- genero (TEXT)

3. escuchas

- usuario_id (UUID)
- cancion id (UUID)
- fecha escucha (DATE)

Este modelo permite registrar de manera eficiente las escuchas y realizar consultas relacionadas con la frecuencia de reproducción, géneros preferidos y popularidad de canciones por ubicación o periodo de tiempo. Además, su diseño se alinea con las características de Cassandra, facilitando lecturas rápidas y segmentación de los datos por claves de partición.

Diseño del Cubo OLAP Conceptual

Para el análisis OLAP se diseñó un cubo conceptual basado en las siguientes características:

- Dimensiones:
 - o **Género**: categoría musical de la canción.
 - o **Tiempo**: fecha de la escucha, con jerarquía Día → Mes.
- Medida Principal:
 - Número de escuchas: total de veces que se escuchó una canción, por género y por periodo de tiempo.

Este cubo permite realizar análisis agregados como:

- Total de escuchas por género y mes.
- Comparación de popularidad de géneros entre diferentes meses.

A pesar de no haber implementado una solución OLAP completa con las debidas herramientas tales como Apache Spark e incluso herramientas de BI, nuestro diseño conceptual permitió la ejecución de consultas básicas sobre la base de datos junto a agrupaciones y conteos emulando funcionalidades OLAP en un entorno simple.

Implementación del Algoritmo de Recomendación

El algoritmo implementado en este proyecto tiene un enfoque deliberadamente simple, basado en la ubicación del usuario. La lógica se resume en lo siguiente:

- Dado un usuario, se identifica su ciudad registrada.
- A partir de esta ciudad, se consulta la base de datos para identificar las canciones más escuchadas en dicha localidad.
- El sistema devuelve una lista de canciones populares locales como recomendación.

Este enfoque permite personalizar la experiencia del usuario sin requerir cálculos complejos o grandes volúmenes de datos interrelacionados. La implementación se realiza en el backend utilizando Node.js y consultas CQL (Cassandra Query Language) para agrupar y contar escuchas por ciudad.

Este modelo de recomendación puede ser fácilmente extendido en el futuro para incluir criterios adicionales como género musical preferido, artistas favoritos o interacciones previas del usuario.

Codigo fuente:

Backend:

Cassandra-setup.js

```
const cassandra = require('cassandra-driver');
const keyspace = 'musicrec';
const client = new cassandra.Client({
 contactPoints: [process.env.CASSANDRA_HOST || 'localhost'],
 localDataCenter: 'datacenter1'
});
async function setupKeyspaceAndTables() {
    const createKeyspaceQuery = `
     CREATE KEYSPACE IF NOT EXISTS ${keyspace}
     WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
    await client.execute(createKeyspaceQuery);
    const queryUsuarios = `
     CREATE TABLE IF NOT EXISTS ${keyspace}.usuarios (
       usuario_id int PRIMARY KEY,
       nombre text,
       ciudad text
    await client.execute(queryUsuarios);
```

```
const queryCanciones =
      CREATE TABLE IF NOT EXISTS ${keyspace}.canciones (
        cancion id int PRIMARY KEY,
       artista text,
        genero text
    await client.execute(queryCanciones);
    const queryEscuchas = `
      CREATE TABLE IF NOT EXISTS ${keyspace}.escuchas (
        usuario_id int,
        fecha_escucha date,
        cancion_id int,
        PRIMARY KEY (usuario_id, fecha_escucha, cancion_id)
    await client.execute(queryEscuchas);
    console.log("Keyspace y tablas creadas correctamente.");
  } catch (err) {
    console.error("Error al configurar keyspace y tablas: ", err);
module.exports = { client, setupKeyspaceAndTables, keyspace };
```

import data.js

```
// Importa los módulos necesarios para manejo de archivos, rutas, CSV y Cassandra
const fs = require('fs'); // Para leer archivos
const path = require('path'); // Para manejar rutas de archivos
const csv = require('csv-parser'); // Para parsear archivos CSV
const { client, setupKeyspaceAndTables } = require('./cassandra-setup'); // Cliente y setup de
function importCSV(filePath, insertRow) {
 return new Promise((resolve, reject) => {
    const promises = [];
    fs.createReadStream(filePath)
      .pipe(csv())
      .on('data', (row) => {
        promises.push(insertRow(row)); // Inserta cada fila en la base de datos
      })
      .on('end', async () => {
          await Promise.all(promises); // Espera a que todas las inserciones terminen
          resolve();
        } catch (err) {
          reject(err);
      })
      .on('error', reject); // Maneja errores de lectura o parseo
```

```
});
async function importData() {
 await setupKeyspaceAndTables(); // Asegura que el keyspace y las tablas existen
 // Importar datos de Usuarios
 const usuariosFile = path.join(__dirname, 'data', 'usuarios.csv'); // Ruta al CSV de
 await importCSV(usuariosFile, async (row) => {
    const query = 'INSERT INTO musicrec.usuarios (usuario_id, nombre, ciudad) VALUES (?, ?,
    await client.execute(query, [parseInt(row.usuario_id), row.nombre, row.ciudad], { prepare:
true });
  });
 console.log("Usuarios importados.");
 // Importar datos de Canciones
 const cancionesFile = path.join(__dirname, 'data', 'canciones.csv'); // Ruta al CSV de
 await importCSV(cancionesFile, async (row) => {
    const query = 'INSERT INTO musicrec.canciones (cancion_id, titulo, artista, genero) VALUES
(?, ?, ?, ?)';
    await client.execute(query, [parseInt(row.cancion_id), row.titulo, row.artista,
row.genero], { prepare: true });
 });
 console.log("Canciones importadas.");
 // Importar datos de Escuchas
 const escuchasFile = path.join(_dirname, 'data', 'escuchas.csv'); // Ruta al CSV de
  await importCSV(escuchasFile, async (row) => {
    const query = 'INSERT INTO musicrec.escuchas (usuario_id, fecha_escucha, cancion_id)
VALUES (?, ?, ?)';
    const fecha = new Date(row.fecha escucha);
    await client.execute(query, [parseInt(row.usuario id), fecha, parseInt(row.cancion id)], {
prepare: true });
  });
 console.log("Escuchas importadas.");
 // Termina el proceso cuando todo está listo
 process.exit(0);
importData();
```

```
const express = require('express'); // Framework web para Node.js
const bodyParser = require('body-parser'); // Middleware para parsear JSON
const { client, setupKeyspaceAndTables } = require('./cassandra-setup'); // Cliente y setup de
const cors = require('cors'); // Middleware para CORS
const multer = require('multer'); // Middleware para manejo de archivos
const csv = require('csv-parser'); // Parser de archivos CSV
const fs = require('fs'); // Módulo de sistema de archivos
const path = require('path'); // Módulo para rutas de archivos
const upload = multer({ dest: 'uploads/' }); // Configura multer para subir archivos a la
const app = express(); // Crea la app de Express
let open;
(async () => {
 open = (await import('open')).default;
})();
app.use(cors({
 origin: '*'
})); // Permite peticiones desde cualquier origen (CORS)
app.use(bodyParser.json()); // Permite recibir JSON en las peticiones
app.use(bodyParser.urlencoded({ extended: true })); // Permite recibir datos de formularios
const PORT = process.env.PORT || 3000;
const frontendPath = path.join(__dirname, '../frontend');
app.use(express.static(frontendPath));
app.get('/', (req, res) => {
 res.sendFile(path.join(frontendPath, 'index.html'));
});
async function startServer() {
  await setupKeyspaceAndTables(); // Crea keyspace y tablas si no existen
  app.listen(PORT, async () => {
    console.log(`Servidor escuchando en http://localhost:${PORT}`);
    if (open) {
      await open(`http://localhost:${PORT}`);
  });
startServer();
app.get('/recommendation', async (req, res) => {
    const usuarioId = parseInt(req.query.usuario_id); // ID del usuario
    const modo = req.query.modo || 'ciudad'; // Modo de recomendación
    if (isNaN(usuarioId)) {
```

```
return res.status(400).json({ error: 'usuario_id debe ser un número' });
    if (modo === 'ciudad') {
      // --- RECOMENDACIÓN POR CIUDAD ---
      const userQuery = 'SELECT ciudad FROM musicrec.usuarios WHERE usuario_id = ?';
      const userResult = await client.execute(userQuery, [usuarioId], { prepare: true });
      if (userResult.rowLength === 0) {
        return res.status(404).json({ error: 'Usuario no encontrado' });
      const ciudad = userResult.rows[0].ciudad;
      const allUsersQuery = 'SELECT usuario_id, ciudad FROM musicrec.usuarios';
      const allUsersResult = await client.execute(allUsersQuery);
      const matchingUsers = allUsersResult.rows
        .filter(u => u.ciudad === ciudad)
        .map(u => u.usuario_id);
      let songCounts = {};
      for (let uid of matchingUsers) {
        const escuchaQuery = 'SELECT cancion_id FROM musicrec.escuchas WHERE usuario_id = ?';
        const escuchaResult = await client.execute(escuchaQuery, [uid], { prepare: true });
        escuchaResult.rows.forEach(row => {
          songCounts[row.cancion_id] = (songCounts[row.cancion_id] || 0) + 1;
        });
      const sortedSongs = Object.entries(songCounts)
        .sort((a, b) \Rightarrow b[1] - a[1])
        .slice(0, 5);
      let recommendations = [];
      for (let [cancion_id, count] of sortedSongs) {
        const songQuery = 'SELECT titulo, artista, genero FROM musicrec.canciones WHERE
cancion_id = ?';
        const songResult = await client.execute(songQuery, [parseInt(cancion_id)], { prepare:
true });
        if (songResult.rowLength > 0) {
          recommendations.push({
            cancion id: cancion id,
            titulo: songResult.rows[0].titulo,
            artista: songResult.rows[0].artista,
            genero: songResult.rows[0].genero,
            listens: count
          });
      return res.json({ recommendations }); // Devuelve las recomendaciones por ciudad
    } else if (modo === 'genero') {
      // --- RECOMENDACIÓN POR GÉNERO FAVORITO ---
      // 1. Obtiene todas las escuchas del usuario
      const escuchasQuery = 'SELECT cancion_id FROM musicrec.escuchas WHERE usuario_id = ?';
      const escuchasResult = await client.execute(escuchasQuery, [usuarioId], { prepare: true
});
      if (escuchasResult.rowLength === 0) {
        return res.status(404).json({ error: 'El usuario no tiene escuchas registradas' });
      // 2. Cuenta cuántas veces escuchó cada género
      let genreCounts = {};
      for (let row of escuchasResult.rows) {
```

```
const cancionId = row.cancion id;
        const songQuery = 'SELECT genero FROM musicrec.canciones WHERE cancion_id = ?';
        const songResult = await client.execute(songQuery, [cancionId], { prepare: true });
        if (songResult.rowLength > 0) {
          const genero = songResult.rows[0].genero;
          genreCounts[genero] = (genreCounts[genero] | | 0) + 1;
      const favoriteGenre = Object.entries(genreCounts).sort((a, b) => b[1] - a[1])[0]?.[0];
      if (!favoriteGenre) {
        return res.status(404).json({ error: 'No se pudo determinar el género favorito' });
      // 4. Busca todas las escuchas de ese género (de todos los usuarios)
      const allEscuchasQuery = 'SELECT cancion_id FROM musicrec.escuchas';
      const allEscuchasResult = await client.execute(allEscuchasQuery);
      let songCounts = {};
      for (let row of allEscuchasResult.rows) {
        const cancionId = row.cancion_id;
        const songQuery = 'SELECT genero FROM musicrec.canciones WHERE cancion_id = ?';
        const songResult = await client.execute(songQuery, [cancionId], { prepare: true });
        if (songResult.rowLength > 0 && songResult.rows[0].genero === favoriteGenre) {
          songCounts[cancionId] = (songCounts[cancionId] || 0) + 1;
      const sortedSongs = Object.entries(songCounts)
        .sort((a, b) \Rightarrow b[1] - a[1])
        .slice(0, 5);
      let recommendations = [];
      for (let [cancion_id, count] of sortedSongs) {
        const songQuery = 'SELECT titulo, artista, genero FROM musicrec.canciones WHERE
cancion id = ?';
        const songResult = await client.execute(songQuery, [parseInt(cancion_id)], { prepare:
true });
        if (songResult.rowLength > 0) {
          recommendations.push({
            cancion_id: cancion_id,
            titulo: songResult.rows[0].titulo,
            artista: songResult.rows[0].artista,
            genero: songResult.rows[0].genero,
            listens: count
          });
      return res.json({ recommendations, genero_favorito: favoriteGenre }); // Devuelve las
    } else {
      return res.status(400).json({ error: 'Modo de recomendación no válido' });
  } catch (err) {
    console.error("Error en /recommendation: ", err);
    res.status(500).json({ error: 'Error interno del servidor' });
});
// --- Endpoint OLAP: escuchas por género y mes ---
// Devuelve un análisis de escuchas agrupadas por género y mes
app.get('/olap/genre-month', async (req, res) => {
```

```
try {
    const escuchasQuery = 'SELECT usuario_id, cancion_id, fecha_escucha FROM
musicrec.escuchas';
    const escuchasResult = await client.execute(escuchasQuery, [], { prepare: true });
    // Mapea cancion_id a género
    const songsQuery = 'SELECT cancion_id, genero FROM musicrec.canciones';
    const songsResult = await client.execute(songsQuery, [], { prepare: true });
    let songGenreMap = {};
    songsResult.rows.forEach(row => {
      songGenreMap[row.cancion_id] = row.genero;
    });
    let aggregation = {};
    // Agrupa escuchas por género y mes
    escuchasResult.rows.forEach(row => {
      const fecha = new Date(row.fecha_escucha);
      const month = `${fecha.getFullYear()}-${("0" + (fecha.getMonth() + 1)).slice(-2)}`;
      const genero = songGenreMap[row.cancion_id] || 'Desconocido';
      const key = `${genero}_${month}`;
      aggregation[key] = (aggregation[key] || 0) + 1;
    });
    let resultArray = [];
    for (let key in aggregation) {
      const [genero, month] = key.split('_');
      resultArray.push({ genero, month, listens: aggregation[key] });
    res.json({ olap: resultArray }); // Devuelve el análisis OLAP
  } catch (err) {
    console.error("Error en /olap/genre-month: ", err);
    res.status(500).json({ error: 'Error interno del servidor' });
});
// Devuelve un análisis de escuchas agrupadas por género, mes y ciudad
app.get('/olap/genre-month-city', async (req, res) => {
 try {
    // 1. Obtener escuchas con usuario_id, cancion_id, fecha_escucha
    const escuchasQuery = 'SELECT usuario_id, cancion_id, fecha_escucha FROM
musicrec.escuchas';
    const escuchasResult = await client.execute(escuchasQuery, [], { prepare: true });
    const songsQuery = 'SELECT cancion_id, genero FROM musicrec.canciones';
    const songsResult = await client.execute(songsQuery, [], { prepare: true });
    let songGenreMap = {};
    songsResult.rows.forEach(row => {
      songGenreMap[row.cancion_id] = row.genero;
    // 3. Mapear usuario_id a ciudad
    const usersQuery = 'SELECT usuario_id, ciudad FROM musicrec.usuarios';
    const usersResult = await client.execute(usersQuery, [], { prepare: true });
    let userCityMap = {};
    usersResult.rows.forEach(row => {
      userCityMap[row.usuario_id] = row.ciudad;
    // 4. Agrupar escuchas por género, mes y ciudad
    let aggregation = {};
    escuchasResult.rows.forEach(row => {
      const fecha = new Date(row.fecha_escucha);
      const month = `${fecha.getFullYear()}-${("0" + (fecha.getMonth() + 1)).slice(-2)}`;
```

```
const genero = songGenreMap[row.cancion_id] || 'Desconocido';
      const ciudad = userCityMap[row.usuario_id] || 'Desconocido';
      const key = `${genero}_${month}_${ciudad}`;
      aggregation[key] = (aggregation[key] || 0) + 1;
    });
    let resultArray = [];
    for (let key in aggregation) {
      const [genero, month, ciudad] = key.split('_');
      resultArray.push({ genero, month, ciudad, listens: aggregation[key] });
    res.json({ olap: resultArray });
  } catch (err) {
    console.error("Error en /olap/genre-month-city: ", err);
    res.status(500).json({ error: 'Error interno del servidor' });
});
app.get('/health', (req, res) => {
 res.send('OK');
});
// --- Scripts de utilidad ---
const scripts = {
 start: "node index.js",
 import: "node import_data.js"
};
// Permite agregar un usuario desde el frontend o por API
app.post('/usuario', async (req, res) => {
    const { usuario_id, nombre, ciudad } = req.body;
    if (!usuario_id || !nombre || !ciudad) {
      return res.status(400).json({ error: 'Faltan datos para el usuario' });
    const query = 'INSERT INTO musicrec.usuarios (usuario_id, nombre, ciudad) VALUES (?, ?,
    await client.execute(query, [parseInt(usuario_id), nombre, ciudad], { prepare: true });
    res.json({ mensaje: 'Usuario insertado correctamente' });
 } catch (err) {
    console.error('Error al insertar usuario:', err);
    res.status(500).json({ error: 'Error al insertar usuario' });
});
// --- Endpoint para insertar una canción ---
// Permite agregar una canción desde el frontend o por API
app.post('/cancion', async (req, res) => {
 try {
    const { cancion_id, titulo, artista, genero } = req.body;
    if (!cancion_id || !titulo || !artista || !genero) {
      return res.status(400).json({ error: 'Faltan datos para la canción' });
    const query = 'INSERT INTO musicrec.canciones (cancion_id, titulo, artista, genero) VALUES
(?, ?, ?, ?)';
```

```
await client.execute(query, [parseInt(cancion_id), titulo, artista, genero], { prepare:
true });
    res.json({ mensaje: 'Canción insertada correctamente' });
  } catch (err) {
    console.error('Error al insertar canción:', err);
    res.status(500).json({ error: 'Error al insertar canción' });
});
// Permite agregar una escucha desde el frontend o por API
app.post('/escucha', async (req, res) => {
 try {
    const { usuario_id, cancion_id, fecha_escucha } = req.body;
    if (!usuario_id || !cancion_id || !fecha_escucha) {
     return res.status(400).json({ error: 'Faltan datos para la escucha' });
    const fecha = new Date(fecha_escucha);
    const query = 'INSERT INTO musicrec.escuchas (usuario_id, fecha_escucha, cancion_id)
VALUES (?, ?, ?)';
    await client.execute(query, [parseInt(usuario_id), fecha, parseInt(cancion_id)], {
prepare: true });
    res.json({ mensaje: 'Escucha insertada correctamente' });
  } catch (err) {
    console.error('Error al insertar escucha:', err);
    res.status(500).json({ error: 'Error al insertar escucha' });
});
// --- Endpoint para carga masiva de CSV ---
// Permite cargar usuarios, canciones o escuchas desde un archivo CSV
app.post('/cargar-csv/:tipo', upload.single('csv'), async (req, res) => {
  const tipo = req.params.tipo; // Tipo de datos a cargar
  const filePath = req.file.path; // Ruta temporal del archivo
 let insertQuery, insertFn;
 try {
    if (tipo === 'usuarios') {
      insertQuery = 'INSERT INTO musicrec.usuarios (usuario_id, nombre, ciudad) VALUES (?, ?,
      insertFn = row => client.execute(insertQuery, [parseInt(row.usuario_id), row.nombre,
row.ciudad], { prepare: true });
    } else if (tipo === 'canciones') {
      insertQuery = 'INSERT INTO musicrec.canciones (cancion_id, titulo, artista, genero)
VALUES (?, ?, ?, ?)';
      insertFn = row => client.execute(insertQuery, [parseInt(row.cancion_id), row.titulo,
row.artista, row.genero], { prepare: true });
    } else if (tipo === 'escuchas') {
      insertQuery = 'INSERT INTO musicrec.escuchas (usuario_id, fecha_escucha, cancion_id)
VALUES (?, ?, ?)';
      insertFn = row => client.execute(insertQuery, [parseInt(row.usuario_id), new
Date(row.fecha_escucha), parseInt(row.cancion_id)], { prepare: true });
    } else {
      fs.unlinkSync(filePath);
      return res.status(400).json({ error: 'Tipo de datos no válido' });
    const promises = [];
    fs.createReadStream(filePath)
      .pipe(csv())
```

```
.on('data', row => promises.push(insertFn(row)))
      .on('end', async () => {
        try {
          await Promise.all(promises);
          fs.unlinkSync(filePath);
          res.json({ mensaje: 'Datos cargados correctamente.' });
        } catch (err) {
          fs.unlinkSync(filePath);
          res.status(500).json({ error: 'Error al insertar datos.' });
      })
      .on('error', err => {
        fs.unlinkSync(filePath);
        res.status(500).json({ error: 'Error al procesar el archivo.' });
      });
  } catch (err) {
    if (fs.existsSync(filePath)) fs.unlinkSync(filePath);
    res.status(500).json({ error: 'Error interno del servidor.' });
});
```

Frontend:

App.js

```
const backendUrl = 'http://localhost:3000';
document.getElementById('recommendation-form').addEventListener('submit', async (e) => {
 e.preventDefault(); // Evita que el formulario recargue la página
 const usuarioId = document.getElementById('usuario id').value; // Obtiene el ID de usuario
 const modo = document.getElementById('modo_recomendacion').value; // Obtiene el modo de
recomendación
  const resultDiv = document.getElementById('recommendation-result'); // Div donde se muestran
los resultados
 resultDiv.innerHTML = 'Cargando recomendaciones...'; // Mensaje de carga
  try {
    const response = await
fetch(`${backendUrl}/recommendation?usuario id=${usuarioId}&modo=${modo}`);
    const data = await response.json(); // Convierte la respuesta a JSON
    if (data.error) {
     resultDiv.innerHTML = `Error: ${data.error}`;
    } else {
      // Si hay recomendaciones, las muestra en una lista
      let html = '<h3>Recomendaciones:</h3>';
      data.recommendations.forEach(rec => {
        if (modo === 'ciudad') {
         html += `${rec.titulo} - ${rec.artista} (Género: ${rec.genero}, Escuchas en tu
ciudad: ${rec.listens})`;
        } else {
         html += `${rec.titulo} - ${rec.artista} (Género: ${rec.genero}, Escuchas:
${rec.listens})`;
```

```
});
      html += '';
      resultDiv.innerHTML = html;
  } catch (err) {
    resultDiv.innerHTML = `Error al conectar al backend.`;
});
// --- Visualización OLAP 2D: Género/Mes/Ciudad ---
document.getElementById('fetch-olap').addEventListener('click', async () => {
  const olapDiv = document.getElementById('olap-result');
  olapDiv.innerHTML = 'Cargando análisis OLAP...';
  try {
    // Solicita el análisis OLAP 2D al backend
    const response = await fetch(`${backendUrl}/olap/genre-month-city`);
    const data = await response.json();
    if (data.error) {
      olapDiv.innerHTML = `Error: ${data.error}`;
    } else {
      // Agrupar por ciudad para mostrar un gráfico por ciudad
      const grouped = {};
      data.olap.forEach(item => {
        if (!grouped[item.ciudad]) grouped[item.ciudad] = [];
        grouped[item.ciudad].push(item);
      });
      Object.keys(grouped).forEach(ciudad => {
        grouped[ciudad].sort((a, b) => {
          if (a.month < b.month) return -1;</pre>
          if (a.month > b.month) return 1;
          // Si el mes es igual, por género alfabético
          if (a.genero < b.genero) return -1;</pre>
          if (a.genero > b.genero) return 1;
          return 0;
        });
      });
      // Crear un gráfico simple por ciudad usando canvas
      let html = '';
      Object.keys(grouped).forEach(ciudad => {
        html += `<h4>${ciudad}</h4><canvas id="chart-${ciudad}" width="500"</pre>
height="250"></canvas>`;
      });
      olapDiv.innerHTML = html;
      // Dibujar los gráficos
      Object.keys(grouped).forEach(ciudad => {
        const canvas = document.getElementById(`chart-${ciudad}`);
        if (!canvas) return;
        const ctx = canvas.getContext('2d');
        const labels = grouped[ciudad].map(item => `${item.month}\n${item.genero}`);
        const values = grouped[ciudad].map(item => item.listens);
        const max = Math.max(...values, 1);
        const barWidth = Math.max(20, 400 / values.length);
        ctx.clearRect(0, 0, 500, 250);
```

```
// Dibujar barras
        values.forEach((v, i) => {
          ctx.fillStyle = '#4a90e2';
          ctx.fillRect(40 + i * barWidth, 200 - (v / max) * 180, barWidth - 4, (v / max) *
180);
          // Mostrar valor encima de la barra
          ctx.fillStyle = '#333';
          ctx.font = '10px Arial';
          ctx.fillText(v, 40 + i * barWidth, 195 - (v / max) * 180);
        // Dibujar etiquetas de género y mes debajo de cada barra, sin rotar y con salto de
        labels.forEach((label, i) => {
          const [mes, genero] = label.split('\n');
          ctx.save();
          ctx.font = '10px Arial';
          ctx.textAlign = 'center';
          ctx.fillStyle = '#222';
          ctx.fillText(mes, 40 + i * barWidth + (barWidth - 4) / 2, 215);
          ctx.fillStyle = '#e26a4a';
          ctx.font = 'bold 11px Arial';
          ctx.fillText(genero, 40 + i * barWidth + (barWidth - 4) / 2, 230);
          ctx.restore();
        });
        // Ejes
        ctx.beginPath();
        ctx.moveTo(35, 200);
        ctx.lineTo(480, 200);
        ctx.moveTo(40, 200);
        ctx.lineTo(40, 20);
        ctx.strokeStyle = '#000';
        ctx.stroke();
        ctx.font = '12px Arial';
        ctx.fillText('Escuchas', 5, 30);
      });
  } catch (err) {
   olapDiv.innerHTML = `Error al conectar al backend.`;
});
// Evento para el formulario de nuevo usuario
document.getElementById('insert-usuario-form').addEventListener('submit', async (e) => {
  e.preventDefault(); // Evita recarga
  const usuario_id = document.getElementById('nuevo_usuario_id').value; // ID usuario
  const nombre = document.getElementById('nuevo_usuario_nombre').value; // Nombre
  const ciudad = document.getElementById('nuevo_usuario_ciudad').value; // Ciudad
  const resultDiv = document.getElementById('insert-usuario-result'); // Div resultado
  resultDiv.innerHTML = 'Insertando...'; // Mensaje de carga
    // Envía los datos al backend por POST
    const response = await fetch(`${backendUrl}/usuario`, {
      method: 'POST',
      headers: { 'Content-Type': 'application/json' },
      body: JSON.stringify({ usuario_id, nombre, ciudad })
    });
    const data = await response.json(); // Respuesta JSON
    if (data.error) {
```

```
// Muestra error si ocurre
     resultDiv.innerHTML = `Error: ${data.error}`;
     resultDiv.innerHTML = `${data.mensaje}`;
  } catch (err) {
   resultDiv.innerHTML = `Error al conectar al backend.`;
});
document.getElementById('insert-cancion-form').addEventListener('submit', async (e) => {
 e.preventDefault(); // Evita recarga
 const cancion_id = document.getElementById('nueva_cancion_id').value; // ID canción
 const titulo = document.getElementById('nueva_cancion_titulo').value; // Título
  const artista = document.getElementById('nueva_cancion_artista').value; // Artista
  const genero = document.getElementById('nueva_cancion_genero').value; // Género
  const resultDiv = document.getElementById('insert-cancion-result'); // Div resultado
  resultDiv.innerHTML = 'Insertando...'; // Mensaje de carga
 try {
   const response = await fetch(`${backendUrl}/cancion`, {
     method: 'POST',
     headers: { 'Content-Type': 'application/json' },
     body: JSON.stringify({ cancion_id, titulo, artista, genero })
   });
   const data = await response.json(); // Respuesta JSON
   if (data.error) {
     resultDiv.innerHTML = `Error: ${data.error}`;
   } else {
     // Muestra éxito
     resultDiv.innerHTML = `${data.mensaje}`;
  } catch (err) {
   resultDiv.innerHTML = `Error al conectar al backend.`;
});
// --- Insertar Escucha ---
document.getElementById('insert-escucha-form').addEventListener('submit', async (e) => {
  e.preventDefault(); // Evita recarga
  const usuario_id = document.getElementById('nueva_escucha_usuario_id').value; // ID usuario
 const cancion_id = document.getElementById('nueva_escucha_cancion_id').value; // ID canción
  const fecha escucha = document.getElementById('nueva escucha fecha').value; // Fecha
  const resultDiv = document.getElementById('insert-escucha-result'); // Div resultado
  resultDiv.innerHTML = 'Insertando...'; // Mensaje de carga
 try {
   const response = await fetch(`${backendUrl}/escucha`, {
     method: 'POST',
     headers: { 'Content-Type': 'application/json' },
     body: JSON.stringify({ usuario_id, cancion_id, fecha_escucha })
```

```
const data = await response.json(); // Respuesta JSON
   if (data.error) {
     // Muestra error si ocurre
     resultDiv.innerHTML = `Error: ${data.error}`;
   } else {
     resultDiv.innerHTML = `${data.mensaje}`;
  } catch (err) {
   resultDiv.innerHTML = `Error al conectar al backend.`;
});
// --- Carga masiva de CSV ---
// Evento para el formulario de carga de CSV
document.getElementById('csv-upload-form').addEventListener('submit', async (e) => {
 e.preventDefault(); // Evita recarga
 const tipo = document.getElementById('csv-type').value; // Tipo de datos a cargar
 const fileInput = document.getElementById('csv-file'); // Input de archivo
 const resultDiv = document.getElementById('csv-upload-result'); // Div resultado
 if (!fileInput.files.length) {
   // Si no hay archivo seleccionado, muestra error
   resultDiv.innerHTML = 'Selecciona un archivo CSV.';
   return;
  const file = fileInput.files[0]; // Archivo seleccionado
  const formData = new FormData();
  formData.append('csv', file); // Agrega el archivo al formData
  try {
   const response = await fetch(`${backendUrl}/cargar-csv/${tipo}`, {
     method: 'POST',
     body: formData
   });
   const data = await response.json(); // Respuesta JSON
   if (data.error) {
     // Muestra error si ocurre
     resultDiv.innerHTML = `Error: ${data.error}`;
   } else {
     // Muestra éxito
     resultDiv.innerHTML = `${data.mensaje}`;
  } catch (err) {
   resultDiv.innerHTML = `Error al conectar al backend.`;
});
```

index.html

```
<!DOCTYPE html>
<html lang="es">
 <meta charset="UTF-8">
  <title>Sistema de Recomendación de Música</title>
  <link rel="stylesheet" href="styles.css">
  <div class="container">
    <h1>Sistema de Recomendación de Música</h1>
    <section id="recommendation-section">
      <h2>Recomendación Personalizada</h2>
      <form id="recommendation-form">
        <label for="usuario id">Ingresa tu ID de usuario:</label>
        <input type="number" id="usuario id" name="usuario id" required>
        <label for="modo recomendacion">Modo de recomendación:</label>
        <select id="modo recomendacion" name="modo recomendacion">
          <option value="ciudad">Por ciudad</option>
          <option value="genero">Por género favorito</option>
        </selects
        <button type="submit">Obtener Recomendaciones/button>
      <div id="recommendation-result"></div>
    </section>
    <section id="olap-section">
      <h2>Análisis OLAP: Escuchas por Género, Mes y Ciudad</h2>
      <button id="fetch-olap">Ver OLAP 2D</button>
      <div id="olap-result"></div>
    </section>
    <section id="insert-section">
      <h2>Insertar Usuario, Canción o Escucha</h2>
      <div style="display: flex; gap: 30px; flex-wrap: wrap; justify-content: center;">
        <form id="insert-usuario-form">
          <h3>Nuevo Usuario</h3>
          <input type="number" id="nuevo_usuario_id" placeholder="ID Usuario" required><br>
          <input type="text" id="nuevo_usuario_nombre" placeholder="Nombre" required><br>
          <input type="text" id="nuevo usuario ciudad" placeholder="Ciudad" required><br>
          <button type="submit">Insertar Usuario</button>
          <div id="insert-usuario-result"></div>
        </form>
        <form id="insert-cancion-form">
          <h3>Nueva Canción</h3>
          <input type="number" id="nueva_cancion_id" placeholder="ID Canción" required><br>
          <input type="text" id="nueva cancion titulo" placeholder="Título" required><br>
          <input type="text" id="nueva cancion artista" placeholder="Artista" required><br>
          <input type="text" id="nueva cancion genero" placeholder="Género" required><br>
          <button type="submit">Insertar Canción</button>
          <div id="insert-cancion-result"></div>
        </form>
        <form id="insert-escucha-form">
          <h3>Nueva Escucha</h3>
          <input type="number" id="nueva escucha usuario id" placeholder="ID Usuario"</pre>
required><br>
```

```
<input type="number" id="nueva_escucha_cancion_id" placeholder="ID Canción"</pre>
required><br>
          <input type="date" id="nueva escucha fecha" required><br>
          <button type="submit">Insertar Escucha</button>
          <div id="insert-escucha-result"></div>
        </form>
    </section>
    <section id="csv-upload-section">
     <h2>Cargar datos masivos desde CSV</h2>
     <form id="csv-upload-form">
       <label for="csv-type">Tipo de datos:</label>
       <select id="csv-type" required>
         <option value="usuarios">Usuarios</option>
         <option value="canciones">Canciones</option>
         <option value="escuchas">Escuchas</option>
        </select>
        <input type="file" id="csv-file" accept=".csv" required>
        <button type="submit">Cargar CSV</button>
      <div id="csv-upload-result"></div>
 <script src="app.js"></script>
</body>
```

Docker-compose.yml

Archivos de información:

Funcion-ario: Explicación detallada del proyecto

Descripción General

Este sistema es una aplicación web de recomendación de música que utiliza una base de datos NoSQL (Cassandra) y permite:

- Recomendar canciones personalizadas a los usuarios.
- Analizar escuchas por género y mes (OLAP simplificado).
- Insertar usuarios, canciones y escuchas desde la web.
- Cargar datos masivos desde archivos CSV.

Estructura del Proyecto

- **docker-compose.yml**: Levanta Cassandra en un contenedor Docker.
- **backend/**: Código del servidor Node.js.
- **index.js**: Lógica principal del backend y endpoints.
- **cassandra-setup.js**: Configura el keyspace y las tablas en Cassandra.
- **import data.js**: Importa datos desde CSV a Cassandra.
- **data/**: Archivos CSV de usuarios, canciones y escuchas.
- **frontend/**: Aplicación web (HTML, CSS, JS).
- **index.html**: Interfaz de usuario.
- **app.js**: Lógica de interacción con el backend.
- **styles.css**: Estilos visuales.
- **docs/**: Documentación y diccionario de funciones.

Backend (Node.js + Express)

- **Conexión a Cassandra**: Usa el driver oficial, crea el keyspace y las tablas si no existen.
- **Endpoints principales**:
- `/recommendation`: Devuelve recomendaciones personalizadas según el usuario y el modo (por ciudad o por género favorito).
- `/olap/genre-month`: Devuelve un análisis OLAP de escuchas agrupadas por género y mes.
- '/usuario', '/cancion', '/escucha': Permiten insertar usuarios, canciones y escuchas vía POST.

- `/cargar-csv/:tipo`: Permite cargar datos masivos desde archivos CSV (usuarios, canciones o escuchas).
- **import_data.js**: Script para importar los datos iniciales desde los CSV a Cassandra.

Frontend (HTML + JS + CSS)

- **index.html**: Contiene formularios para:
- Obtener recomendaciones personalizadas (seleccionando modo).
- Ver análisis OLAP.
- Insertar usuarios, canciones y escuchas.
- Cargar archivos CSV masivos.
- **app.js**: Maneja los eventos de los formularios, realiza peticiones al backend y muestra los resultados.
- **styles.css**: Da formato visual moderno y responsivo a la web.

Flujo de Uso

- 1. **Despliegue**: Se levanta Cassandra con Docker y el backend con Node.js.
- 2. **Carga de datos**: Se pueden importar datos iniciales desde CSV o cargar nuevos desde la web.
- 3. **Interacción**: El usuario puede:
 - Consultar recomendaciones personalizadas.
 - Ver análisis OLAP.
 - Insertar nuevos datos manualmente o por CSV.

Lógica de Recomendación

- **Por ciudad**: Recomienda las canciones más escuchadas por usuarios de la misma ciudad.
- **Por género favorito**: Recomienda las canciones más escuchadas del género que más escucha el usuario.

Capturas de Pantalla de la base de datos

A continuación, se muestran algunas capturas de pantalla tablas que componen la base de datos:

1. Tabla de usuarios

cqlsh:musicrec> SELECT * FROM usuarios;					
wario_id	ciudad	nombre			
114	Caracas	Diego Torres			
110	Quito	Andrés Castro			
128	Buenos Aires				
117	Santiago	Daniela Soto			
144					
120	Quito	Esteban Silva			
140	Quito	Lucas Carrasco			
129					
132	Bogota				
185	Bogota				
105	Bogota				
123	Mexico DF	Renata Castro			
153	Mexico DF	Martina Pizarro			
178	Buenos Aires	Ignacio León			
137	Santiago				
118					
194		Matías López			
125		Julieta Gómez			
197	Santiago				
134					
139					
169					
145		Agustina Duarte			
111		Valentina Ríos			
113		Paula Morales			
160		Emilia Castro			
151		Solana Roldán			
1		Mansour			
157					
130	Quito				
122	Bogota				
176	Lima				
177	Santiago				

2. Tabla de escuchas

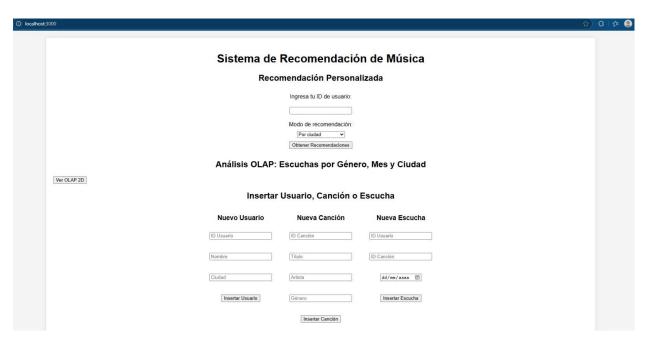
calch-mucicco	c> SELECT * FROM	accuchae-	
cqtsii.riustci e	C> SELECT " FROM	escuciias,	
usuario_id	fecha_escucha	cancion_id	
+			
114	2024-02-09	14	
110	2024-02-05	10	
128	2024-02-23	28	
117	2024-02-12	17	
144	2024-03-11	44	
120	2024-02-15	20	
140	2024-03-07	40	
129	2024-02-24	29	
132	2024-02-27	32	
185	2024-04-21	85	
105	2024-01-19	5	
123	2024-02-18	23	
153	2024-03-20	53	
178	2024-04-14	78	
137	2024-03-04	37	
118	2024-02-13	18	
194	2024-04-30	94	
125	2024-02-20	25	
197	2024-05-03	97	
134	2024-03-01	34	
139	2024-03-06	39	
169	2024-04-05	69	
145	2024-03-12	45	
111	2024-02-06	11	
113	2024-02-08	13	
160	2024-03-27	60	
151	2024-03-18	51	
1	2025-06-04	4	
157	2024-03-24	57	
130	2024-02-25	30	
122	2024-02-17	22	
176	2024-04-12	76	
177	2024-04-13	77	

3. Tabla de canciones

ion_id	artista	genero	titulo
23	Dua Lipa	Pop	Levitating
53	Justin Bieber		
91	Ke\$ha	Pop	
55	Justin Bieber	Pop	Intentions
33	Adele	Pop	Hello
5	Michael Jackson	Pop	Billie Jean
28		Country	Old Town Road
42	Drake	Hip-Hop	In My Feelings
50		Pop	Animals
95	Jennifer Lopez	Pop	On The Floor
88	Clean Bandit	Pop	Rather Be
82		EDM	Lean On
49	Maroon 5 John Lennon	Pop	Sugar
10	John Lennon	Rock	Imagine
83	Major Lazer	EDM	Light it up
60	Coldplay	Rock	Paradise
67		Pop	Apologize
79	Martin Garrix	EDM	Ocean
16			Perfect
63	Coldplay		Fix You
80	Martin Garrix		High on Life
13	Oasis		Wonderwall
30			
72	Avicii	EDM	Levels
99	Jennifer Lopez	Pop	Love Don't Cost a Thing
11	Michael Jackson		Thriller
1		Rock	Bohemian Rhapsody
19			Thinking Out Loud
61			
46	The Chainsmokers	Pop	Paris
43			Hotline Bling

Pruebas de la aplicación:

Pagina inicial



Muestra recomendación por ciudad



Recomendaciones:

- Imagine John Lennon (Género: Rock, Escuchas en tu ciudad: 1)
- One U2 (Género: Rock, Escuchas en tu ciudad: 1)
- Señorita Shawn Mendes (Género: Pop, Escuchas en tu ciudad: 1)
- Lucid Dreams Juice WRLD (Género: Hip-Hop, Escuchas en tu ciudad: 1)
- Animals Maroon 5 (Género: Pop, Escuchas en tu ciudad: 1)

Muestra recomendación por genero

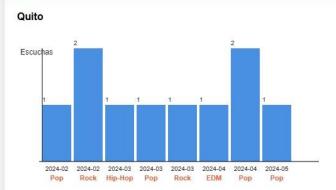
Ingresa tu ID de usuario:
200
Modo de recomendación:
Por género favorito ✔
Obtener Recomendaciones

Recomendaciones:

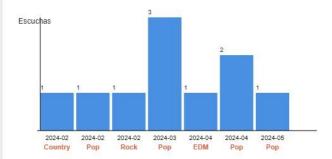
- Like a Prayer Madonna (Género: Pop, Escuchas: 1)
 Billie Jean Michael Jackson (Género: Pop, Escuchas: 1)
- Shape of You Ed Sheeran (Género: Pop, Escuchas: 1)
 Blinding Lights The Weeknd (Género: Pop, Escuchas: 1)
 Thriller Michael Jackson (Género: Pop, Escuchas: 1)

Grafico OLAP

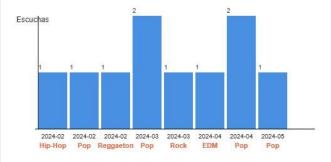




Buenos Aires



Santiago



Zona de inserción de datos:

Insertar Usuario, Canción o Escucha

Nuevo Usuario	Nueva Canción	Nueva Escucha	
ID Usuario	ID Canción	ID Usuario	
Nombre	Título	ID Canción	
Ciudad	Artista	dd/mm/aaaa 🗊	
Insertar Usuario	Género	Insertar Escucha	

Cargar datos masivos desde CSV

Tipo de datos;

[Usuarios ▼]

[Elegir archivo] No se eligió ningún archivo

[Cargar CSV]