

	<p>UNIVERSIDAD NACIONAL EXPERIMENTAL DE GUAYANA COORDINACION DE INGENIERIA INFORMÁTICA UC SISTEMAS DE BASES DE DATOS II SEMESTRE: 2025.I PROFA. CLINIA CORDERO FECHA: 17-05-2025</p>	<p>FECHA DE ENTREGA DE PROYECTO: Semana 8</p>
---	--	--

**PROYECTO Nº 1
BASES DE DATOS BASADAS EN COLUMNAS
APACHE CASSANDRA
VALOR 12%**

Proyecto: Sistema Básico de Recomendación de Música con Análisis OLAP Simplificado en NoSQL

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:

1. Modelado de Datos NoSQL Simplificado:

- Diseñar un esquema NoSQL para almacenar información básica sobre usuarios, canciones y escuchas.
- Por ejemplo:
 - Usuarios: {usuario_id, nombre, ciudad}
 - Canciones: {cancion_id, titulo, artista, genero}
 - Escuchas: {usuario_id, cancion_id, fecha_escucha}
- Elegir *una* base de datos NoSQL (Cassandra).
- Justificación breve y enfocada en la *razón principal* de la elección de NoSQL (por ejemplo, escalabilidad para grandes volúmenes de datos).

2. Implementación de Recomendación Básica:

- Implementar *un único y muy simple* algoritmo de recomendación.
- Por ejemplo: "Canciones más escuchadas por género" o "Canciones más escuchadas en la misma ciudad del usuario".
- No se busca la complejidad de Collaborative Filtering o híbridos en esta versión.

3. Análisis OLAP Muy Simplificado:

- Identificar *2 dimensiones clave* (por ejemplo, usuario y tiempo o género y tiempo).
- Definir *una jerarquía simple* para *una* dimensión (por ejemplo, tiempo: día -> mes).
- Definir *1 o 2 hechos/medidas* (por ejemplo, número de escuchas).

- Diseñar un *cubo OLAP conceptual simple* (sin implementación completa de todas las operaciones).
- 4. Implementación de Consultas OLAP Básicas:**
 - Implementar 2-3 *consultas OLAP básicas* (por ejemplo, "Número total de escuchas por género y mes").
 - Generar *visualizaciones simples* (tablas o gráficos básicos) para mostrar los resultados.

Etapas del Proyecto (3 Semanas):

- **Semana 1: Especificación y Diseño (Intensivo)**
 - Modelo de Casos de Uso (muy básico, 2-3 casos).
 - Modelo de Datos NoSQL (el esquema simple definido arriba).
 - Requisitos OLAP (dimensiones, jerarquías, hechos/medidas simplificados).
 - Diseño del Cubo OLAP (conceptual).
 - Diseño general del sistema (capas simplificadas).
- **Semana 2: Implementación (Enfocado)**
 - Implementación de la base de datos NoSQL.
 - Implementación del algoritmo de recomendación.
 - Implementación de las *consultas OLAP básicas*.
- **Semana 3: Pruebas, Evaluación y Finalización (Combinado)**
 - Pruebas básicas del sistema.
 - Evaluación *cualitativa* de las recomendaciones (¿tienen sentido?).
 - Evaluación de la *ejecución de las consultas OLAP*.
 - Documentación del proyecto.
 - Preparación de la presentación.

Dataset proporcionado por el Docente:

A continuación, se presenta un dataset pequeño. Se recomienda el incremento de cada archivo .csv a 100 registro.

1. 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

2. 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

3. 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

Entregables:

- Documento de Especificación/Diseño.
- Código fuente.
- Base de datos NoSQL implementada.
- Informe de pruebas y evaluación.
- Defensa ó Presentación.

Claves para el Éxito en 3 Semanas:

- Enfoque en la simplicidad: Cada componente (recomendación, OLAP) debe ser lo más simple posible.
- Claridad en los objetivos: Se participa a los estudiantes que el objetivo es la *comprensión básica* de NoSQL y OLAP, no un sistema de recomendación completo, debido al tiempo limitado que se tiene.
- Gestión del tiempo: Los estudiantes deben trabajar de forma intensiva y organizada.