



UNIVERSIDAD NACIONAL EXPERIMENTAL DE GUAYANA
DEPARTAMENTO DE CIENCIA Y TECNOLOGÍA
COORDINACIÓN DE INGENIERÍA INFORMÁTICA
SISTEMAS DE BASES DE DATOS II
SEMESTRE: 2025-II

PROFA. CLINIA CORDERO

FECHA DE ENTREGA: 05-12-2025

**TEMA VIII SISTEMA DE BASES DE DATOS DOCUMENTALES
PROYECTO N° 1
VALOR 15%**

Proyecto: "GlobalMarket Analytics & Search Engine"

Asignatura: Bases de Datos NoSQL / MongoDB

Plataforma: MongoDB Atlas (M0 Free Tier)

Duración: 15 días

Modalidad: Equipos de 3 personas

Nivel: Intermedio

1. Contexto del Escenario

Su equipo ha sido contratado por "GlobalMarket", una startup de comercio electrónico en expansión. Su base de datos relacional actual está fallando bajo la carga de consultas complejas y la variabilidad de los atributos de sus productos.

Su misión: Migrar y optimizar el catálogo de productos y el registro de ventas a **MongoDB Atlas**, implementando un motor de búsqueda eficiente y un dashboard de análisis de ventas en tiempo real.

2. Fuente de Datos (Dataset)

Para este proyecto, se utilizará un dataset real. Se recomienda el siguiente (o uno similar de E-commerce):

- **Opción recomendada:** [Amazon Sales Dataset](#) o [Global Superstore](#).
- **Volumen:** El dataset debe contener al menos 1,000 documentos para que las pruebas de rendimiento e índices sean significativas.

3. Requerimientos Técnicos (El "Core" del Proyecto)

El proyecto se divide en tres pilares fundamentales que cada integrante del equipo puede liderar, aunque todos deben colaborar:

A. Modelado y Validación (Schema Design)

No se trata solo de importar el JSON/CSV. Los alumnos deben:

1. **Diseñar el Esquema:** Decidir estratégicamente cuándo **embeber** (embedding) y cuándo **referenciar** (referencing). *Ejemplo: ¿Las reseñas van dentro del producto o en una colección aparte?*
2. **Schema Validation:** Implementar reglas de validación JSON en Atlas para asegurar la integridad de los datos (ej. que el precio sea siempre positivo, que el email del usuario tenga formato correcto).
3. **Ingesta:** Cargar los datos a un clúster M0 en Atlas (usando `mongoimport` o MongoDB Compass).

B. Consultas Avanzadas (Aggregation Framework)

Deben construir un "Pipeline de Agregación" para responder preguntas de negocio. Se requieren al menos 3 stages complejos:

1. **Reporte de Ventas:** Calcular el total de ventas por categoría y por mes. (Uso de `$match`, `$group`, `$project`).
2. **Top Clientes/Productos:** Identificar los productos con mejor calificación promedio que tengan más de 50 reseñas.
3. **Bucket Pattern:** Agrupar productos por rangos de precios (Bajo, Medio, Alto) usando `$bucket` o `$bucketAuto`.

C. Búsqueda y Rendimiento (Atlas Search & Indexing)

1. **Atlas Search:** Implementar un índice de búsqueda (Lucene) para permitir búsqueda difusa (*fuzzy search*) por nombre de producto o descripción. *Esto simula la barra de búsqueda de la tienda.*
2. **Optimización:** Crear índices tradicionales (Simples y Compuestos) para acelerar las consultas frecuentes.
3. **Análisis:** Usar la herramienta "Explain Plan" para demostrar la diferencia de tiempo de respuesta antes y después de crear los índices.

4. Cronograma Sugerido (15 Días)

Fase	Días	Actividades Clave
1. Diseño e Ingesta	1-4	Selección del dataset, diseño del modelo (JSON), creación del Clúster en Atlas y carga de datos.
2. Desarrollo	5-10	Creación de Aggregation Pipelines, implementación de Schema Validation y configuración de Atlas Search.
3. Optimización	11-13	Creación de índices, análisis de performance (Explain Plan) y creación de un Dashboard visual.
4. Entrega	14-15	Preparación de la presentación, documentación y script de demostración.

5. Entregables

1. **Repositorio (GitHub/GitLab):**
 - Scripts de creación de validaciones (validation.js).
 - Scripts de las consultas de agregación (queries.js).
 - Archivo README.md explicando cómo levantar el proyecto.
2. **Reporte Técnico (PDF):**
 - Diagrama del modelo de datos (comparativa Relacional vs Documental).
 - Capturas de pantalla del "Explain Plan" mostrando la mejora de rendimiento.
3. **MongoDB Charts (Bonus/Visualización):**
 - Un enlace público o capturas de un Dashboard simple hecho en **MongoDB Atlas Charts** que muestre gráficamente los resultados de las agregaciones (ej. Gráfico de barras de ventas por país).

Nota: El uso de *MongoDB Charts* es una gran ventaja didáctica porque permite ver el resultado de sus datos sin programar un frontend complejo, manteniendo el foco en la base de datos.

6. Rúbrica de Evaluación (Total: 15%)

- **Modelado de Datos:** Decisión correcta de embedding vs referencing y uso correcto de tipos de datos.
Valor 3 %
- **Aggregation Framework:** Complejidad y exactitud de los pipelines (uso de \$lookup, \$unwind, etc.).
Valor 5%
- **Atlas Features:** Implementación exitosa de Atlas Search y Validaciones.
Valor 3 %
- **Performance:** Estrategia de indexación correcta y demostración de mejora.
Valor 2%
- **Trabajo en Equipo y Presentación:** Claridad en la exposición y uso de MongoDB Charts.
Valor 2%