

Título de la Tarea: Diseño de un Sistema Big Data para una Empresa Minorista

Descripción del Caso Práctico:

La empresa **RetailNow**, dedicada a la venta de productos tecnológicos, ha contactado con tu consultora para diseñar un sistema Big Data que les permita gestionar el enorme volumen de datos que generan a diario. RetailNow tiene varios puntos de venta físicos y una plataforma de e-commerce. Sus necesidades principales son las siguientes:

1. **Fuentes de datos internas:** Los puntos de venta físicos y el e-commerce generan datos transaccionales de manera continua, incluyendo ventas, inventarios, comportamiento de los clientes, entre otros. Además, los sensores de las tiendas físicas generan datos en tiempo real sobre la afluencia de personas y la interacción de los clientes con los productos.
2. **Fuentes de datos externas:** La empresa está interesada en recopilar datos de redes sociales, tendencias de búsqueda en internet, y estadísticas de productos en otros marketplaces para poder comparar su rendimiento y conocer mejor los intereses de los consumidores.
3. **Volumen de datos:** Actualmente almacenan 200 TB de datos y calculan que cada año se suman otros 50 TB de datos nuevos. Estos datos incluyen tanto información estructurada como no estructurada, como imágenes y registros de actividad en línea.
4. **Acceso a la información:** La junta directiva necesita informes analíticos mensuales. Además, les gustaría tener acceso a un cuadro de mandos interactivo que permita filtrar datos y ver información detallada sobre el rendimiento de productos, clientes y ventas.
5. **Mantenimiento y escalabilidad:** La empresa espera duplicar su volumen de ventas en los próximos cinco años, por lo que necesitarán un sistema escalable que pueda adaptarse al crecimiento de los datos y seguir ofreciendo análisis en tiempo real.

Requisitos:

Tu tarea consiste en diseñar el **prediseño del sistema Big Data** que permitirá a RetailNow gestionar y analizar de manera efectiva toda esta información. Debes detallar cómo el sistema gestionará las distintas fuentes de datos y cómo responderá a las necesidades analíticas de la empresa. A continuación se indican los apartados que debes abordar:

Apartado 1: Prediseño de un sistema Big Data

1. **Arquitectura del sistema:**

- Describe la arquitectura Big Data que utilizarás para RetailNow. Explica qué capas serán necesarias para garantizar el almacenamiento, procesamiento y análisis de los datos de manera eficiente.
- Indica cómo gestionará el sistema tanto datos estructurados (ventas, inventarios) como no estructurados (imágenes, redes sociales).

2. Escalabilidad y almacenamiento:

- Explica cómo se irá ampliando la capacidad del **clúster** de almacenamiento y procesamiento a medida que se incrementa el volumen de datos. ¿Qué tipo de almacenamiento elegirías y por qué (distribuido, en la nube, etc.)?

3. Procesamiento de datos:

- Detalla las estrategias de procesamiento de datos que implementarán. ¿Utilizarás procesamiento **en tiempo real** o **por lotes**? Justifica tu elección.
- Explica si será necesario cumplir con las características **ACID** para ciertos tipos de datos, y si es necesario implementar subsistemas **OLTP** para las transacciones y **OLAP** para el análisis de datos.

4. Analíticas y cuadros de mando:

- Proporciona una estrategia para crear los **cuadros de mando** interactivos que la junta directiva requiere. Explica cómo serán capaces de filtrar la información y realizar análisis dinámicos sobre los datos.

5. Modelos predictivos:

- Considerando la cantidad de datos históricos, indica si será necesario crear **modelos predictivos** para ayudar a RetailNow a identificar tendencias de ventas, comportamiento de los clientes y optimización de inventarios. Si es así, describe brevemente cómo sería la implementación de estos modelos.

Recursos necesarios y recomendaciones

1. Recursos necesarios:

- Ordenador con capacidad para ejecutar software de análisis de datos.
- Acceso a software para simulaciones de procesamiento de datos en clústeres (ej. Apache Spark, Hadoop).

2. Recomendaciones:

- Lee detenidamente la unidad didáctica y consulta el material proporcionado por el docente antes de abordar la tarea.
- Aclara tus dudas con el profesor antes de empezar el desarrollo del prediseño.
- Asegúrate de que el documento sea claro, coherente y con un lenguaje técnico apropiado.

Indicaciones de entrega

- El trabajo debe presentarse en un **documento de texto** nombrado de la siguiente forma: Apellido1_Apellido2_Nombre_SBD01_Tarea
- El archivo debe subirse a la plataforma antes de la fecha límite. Evita el uso de caracteres especiales como la ñ o acentos en el nombre del archivo.

Criterios de evaluación RA1:

- e. Se han establecido objetivos y prioridades, secuenciación y organización del tiempo de realización.
- f. Se han seleccionado e integrado sistemas de información que satisfacen las necesidades del problema.
- g. Se han determinado criterios de coste y calidad necesarios para la eficacia y eficiencia de la implementación de un sistema Big Data

Criterios de evaluación (sobre 10,00 puntos):

Cada apartado tendrá un valor específico, y la evaluación total será de **10,00 puntos**. Se valorará tanto el contenido técnico como la claridad y estructura del trabajo.

1. **Descripción del sistema Big Data propuesto** (2,00 puntos):
Explica correctamente cómo será el sistema Big Data para RetailNow, cubriendo las fuentes de datos y las capas necesarias para el análisis.
2. **Explicación de la arquitectura y escalabilidad** (2,00 puntos):
Detalla la arquitectura Big Data y cómo se garantizará la escalabilidad del sistema frente al crecimiento de los datos.
3. **Estrategias de procesamiento de datos** (2,00 puntos):
Justifica el uso de procesamiento en tiempo real o por lotes, y explica el uso de sistemas **ACID**, **OLTP** y **OLAP** según sea necesario.

4. **Cuadros de mando y análisis dinámicos** (2,00 puntos):
Proporciona una solución para que la junta directiva pueda analizar los datos mediante cuadros de mando interactivos.
 5. **Modelos predictivos y análisis de tendencias** (2,00 puntos):
Discute si será necesaria la creación de modelos predictivos y cómo se implementarán para optimizar las ventas y el inventario.
-

Puntuación total:

El total de la tarea es de **10,00 puntos**. Cada apartado será valorado sobre **2,00 puntos**, sumando un total de 10,00. Se restarán puntos por faltas de ortografía o errores gramaticales (-0,1 por cada error).

Recomendaciones:

- **Lee cuidadosamente** el material proporcionado antes de comenzar con el desarrollo de la tarea.
- **Aclara tus dudas** con el profesor antes de comenzar.
- **Revisa tu trabajo** para asegurar la coherencia técnica, la correcta justificación y un lenguaje adecuado.