

**Prueba Corta #1**

Estudiante: Celina Madrigal Murillo

Carné: 2020059364

1. Explique en que consisten los siguientes conceptos:

**a. Data Warehouse:** almacena datos para crear analytics

**b. Data Lake:** almacenamiento masivo que almacena raw data.

**c. Data Mart:** un data mart es una forma simple de Data Warehouse enfocado en un área funcional o un tema específico.

2. ¿De que forma se benefician las aplicaciones del uso de Columnar Storage? Explique.

Ofrece consultas rápidas y rendimiento de E/S para prácticamente cualquier tamaño de datos. Son adecuadas para cargas de trabajo similares a OLAP. Se usan los datos solo de la columna entonces cuando se hace una consulta no se devuelve toda la fila.

3. ¿En que consiste streaming y batch processing?

El streaming processing es el flujo constante de datos, procesar datos de manera continua y el batch processing consiste en la ejecución de un programa sin el control o supervisión directa del usuario.

4. ¿En que consiste datos estructurados, semi estructurados y no estructurados?

Los datos estructurados son datos que siguen un esquema fijo o tabular, los datos semi estructurados son datos que no necesariamente tienen los mismos atributos y los datos no estructurados son datos que no tienen ningun patrón.

**Referencias:**

Wikipedia Contributors. (2022, June 21). Column-oriented DBMS. Wikipedia; Wikimedia Foundation.  
[https://en.wikipedia.org/wiki/Column-oriented\\_DBMS#Benefits](https://en.wikipedia.org/wiki/Column-oriented_DBMS#Benefits)

oscarfmdc. (2020, June 12). Stream processing - Tecnologías y comparativa. Aprender BIG DATA.  
<https://aprenderbigdata.com/stream-processing/>

de, C. (2006, May 16). Procesamiento por lotes. Wikipedia.org; Wikimedia Foundation, Inc.  
[https://es.wikipedia.org/wiki/Procesamiento\\_por\\_lotes](https://es.wikipedia.org/wiki/Procesamiento_por_lotes)

Data Warehousing on AWS. (2021). <https://d1.awsstatic.com/whitepapers/enterprise-data-warehousing-on-aws.pdf>