**Clase M5 AE2: Enfoques para el almacenamiento y gestión de los datos - Parte 1**

| **Aspecto** | **Data Lake** | **Data Warehouse** | **Data Mart** |
| --- | --- | --- | --- |
| **Tipo de datos** | Estructurados, semiestructurados y no estructurados | Estructurados | Estructurados |
| **Aplicación del esquema** | *Schema-on-read* (al leer los datos) | *Schema-on-write* (al cargar los datos) | *Schema-on-write* |
| **Casos de uso ideales** | Ciencia de datos, machine learning, almacenamiento bruto | Reportes de negocio, inteligencia de negocio, análisis histórico | Análisis específicos por área (marketing, finanzas, ventas) |
| **Nivel técnico necesario** | Alto (requiere conocimiento en big data, ETL/ELT, y herramientas analíticas) | Medio-alto (ETL, modelado dimensional, SQL avanzado) | Medio (conocimiento básico de análisis y BI) |
| **Costos y tiempo** | Bajo costo de almacenamiento, pero alto costo en procesamiento | Costoso en implementación y licencias, pero más optimizado | Menor costo y tiempo de implementación que un DW completo |
| **Usuarios típicos** | Científicos de datos, analistas avanzados | Analistas de negocio, usuarios de BI | Analistas de área, supervisores, personal no técnico |

| **Tecnología / Caso** | **Uso** |
| --- | --- |
| **Data Lake: Amazon S3 + AWS Glue + Athena** | Empresa que almacena logs de sensores IoT, imágenes y datos crudos para análisis predictivo. |
| **Data Warehouse: Google BigQuery / Snowflake** | Retail que necesita generar reportes de ventas históricos, inventario y márgenes. |
| **Data Mart: Power BI + SQL Server** | Departamento de marketing que analiza campañas, ROI y conversión de leads. |