

ALMACENES DE DATOS (DATA WAREHOUSE)



¿QUE SON?

Es un repositorio digital que almacena y protege la información de los sistemas informáticos. Un almacén de datos puede ser un almacenamiento conectado a la red, un almacenamiento distribuido en la nube, un disco duro físico o un almacenamiento virtual

¿PARA QUE SIRVEN?

Centraliza y fusiona grandes cantidades de datos de múltiples fuentes. Las organizaciones pueden obtener información empresarial valiosa a partir de los datos y mejorar las decisiones. Con el tiempo, se construye un registro histórico de gran valor. Gracias a estas funciones, un almacén de datos puede considerarse la "fuente única de datos" de una organización



¿COMO EMPEZO?

Cuando aparecieron por primera vez los data warehouses a finales de los 80, su objetivo era que los datos pasaran de los sistemas operativos a los sistemas de apoyo a las decisiones (DSS). Aquellos primeros data warehouses necesitaban un volumen enorme de redundancias. La mayoría de empresas tenían múltiples entornos DSS para sus diferentes usuarios. Aunque los entornos DSS utilizaban, en gran medida, los mismos datos, la recopilación, la limpieza y la integración de estos se solía replicar para cada entorno.

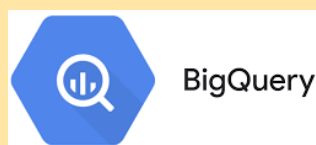
No es sino hasta fines de los años 80 que asistimos a la emergencia de la primera Data Warehouse empresarial, desarrollada por Paul Murphy y Barry Devlin de IBM.



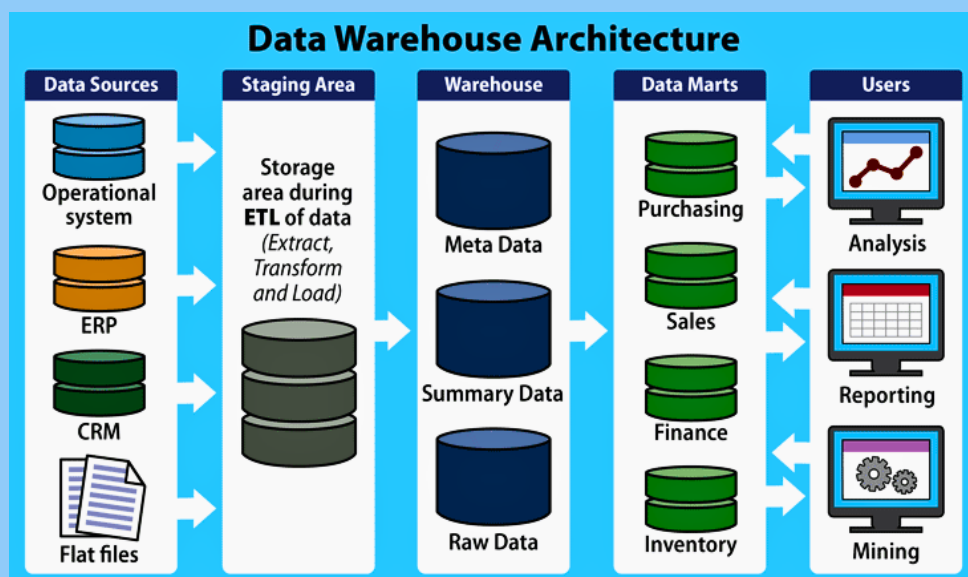
PROVEEDORES

Como principales proveedores de data warehouse tenemos a :

- Amazon Redshift
- Google BigQuery
- Snowflake
- Microsoft Azure Synapse Analytics
- IBM Db2 Warehouse
- Oracle Exadata
- Teradata
- SAP Data Warehouse Cloud
- Vertica (Micro Focus)
- Couchbase



ARQUITECTURA



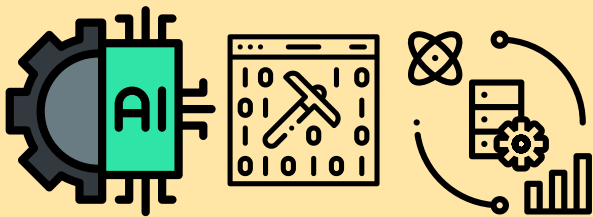
La arquitectura de un data warehouse depende de las necesidades específicas de la organización. Algunas de las arquitecturas más comunes son:

- Sencilla
- Sencilla con zona de preparacion
- Radial
- Entorno de pruebas

AREAS DEL DATA WAREHOUSE

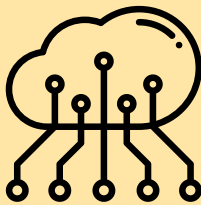
Si tu organización tiene o hace cualquier cosa de las siguientes, es probable que le venga bien un almacén de datos:

- Varias fuentes de datos dispares
- Análisis y visualización de Big Data tanto asíncronos como en tiempo real
- Aprendizaje automático e inteligencia artificial (IA)
- Analíticas de streaming
- Generación de informes personalizados o análisis ad hoc
- Minería de datos
- Ciencia de datos



VENTAJAS DE DATA WAREHOUSE EN LA NUBE

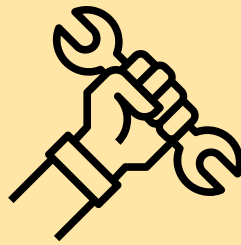
- Totalmente gestionado para ahorrar operaciones
- Mejor tiempo de funcionamiento en comparación con los almacenes de datos on-premise
- Pensados a escala
- Precios flexibles para ser rentables
- Analíticas en tiempo real
- Iniciativas de aprendizaje automático e inteligencia artificial



DESVENTAJAS

Si bien contamos con multiples ventajas, tambien tenemos desventajas como lo puden ser:

- Costos elevados.
- Complejidad en diseño y mantenimiento.
- Tiempo de implementación prolongado.
- Requisitos significativos de recursos.
- Posible latencia en la disponibilidad de datos.
- Rigidez ante cambios en esquema de datos.
- Necesidad de mantenimiento continuo.
- Desafíos en seguridad y privacidad.
- Dependencia de expertos técnicos.
- Complejidad en la escalabilidad.



DATA WAREHOUSE VS BIG DATA VS BUSINESS INTELLIGENCE

Estos 3 apartados comparten un mercado en común, el manejo de grandes cantidades de datos, pero cada uno de ellos tiene su funcionamiento, características y detalles, por lo que es fundamental analizarlos en detalle a cada uno.

DATA WAREHOUSE

Los datos que se almacenan en este sistema son consolidados de diferentes fuentes o sistemas de la empresa. Son datos estructurados, destinados a ser precisos y de alta calidad para realizar aportes en cuanto a la toma de decisiones de la empresa. Los datos pueden ser divididos para diferentes departamentos o sectores, a fin de tomar diferentes estrategias.



BIG DATA

Son datos no estructurados que son almacenados en grandes cantidades con variedad de complejidad y claro, la velocidad con la que estos crecen. No son datos relacionales, por lo que quedan fuera del entorno corporativo. Ofrece un análisis en tiempo real y puede provenir de diferentes tipos de fuentes y formatos



BUSINESS INTELLIGENCE

Hace de cuello de cebolla en los datos que se recogen durante el Data Warehouse, donde tienen información exacta para la toma de decisiones. En este punto ofrece datos que pueden ser transformados en información útil para los negocios y la toma de decisiones de estrategias corporativas.



BIBLIOGRAFIA

- ¿Qué son los almacenes de datos? | Google Cloud. (s. f.). Google Cloud. <https://cloud.google.com/learn/what-is-a-data-warehouse?hl=es>
- ¿Qué es un almacén de datos? (s. f.). <https://www.oracle.com/mx/database/what-is-a-data-warehouse/>
- ¿Qué es un almacén de datos? - Explicación del almacén de datos - AWS. (s. f.). Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/data-store/>
- Emilio. (2020). Comparativa de las mejores soluciones de data warehouse. TodoBI - Business Intelligence, Big Data, ML y AI. <https://todobi.com/comparativa-de-las-15-mejores-soluciones-de-data-warehouse/>
- Roca, C. (2023). Data Warehouse y todo lo que necesitas saber sobre los datos. ThePower Business School. <https://www.thepowermba.com/es/blog/data-warehouse-y-todo-lo-que-necesitas-saber-sobre-los-datos>