

"UNIVERSIDAD AUTÓNOMA GABRIEL RENE MORENO"

**FACULTAD DE INGENIERÍA EN
CIENCIAS DE LA COMPUTACIÓN Y
TELECOMUNICACIONES
CARRERA EN INGENIERÍA EN SISTEMAS**



DATA WEREHOUSE

Estudiante: Vanesa Vino Apaza

Registro: 220053243

Materia: Sistemas para el Soporte y la Toma de Decisiones

Docente: Ing. Miguel Peinado Pereira

Semestre: 2-2024

Sigla: INF-432

Grupo: SA

SANTA CRUZ – BOLIVIA

DATA WEREHOUSE

Un data werehouse o deposito de datos, es un sistema que agrega y combina informacion desde diferentes fuentes. Todos estos datos están integrados en un único almacén centralizado y permite a las empresas mejorar los procesos de toma de decisiones.

Es una base de datos diseñada para almacenar, filtrar, extraer y analizar grandes colecciones de datos, están específicamente desarrollados para trabajar con Big Data.

Estructura de un Data Werehouse

- **Básica:** Funciona con sistemas operativos básicos y archivos planos, proporcionando datos en bruto almacenados junto con metadatos. Los usuarios pueden acceder a estos datos para análisis y generación de informes.
- **Básica con área de ensayo:** Incluye un área de ensayo entre las fuentes de datos y el almacén para limpiar y personalizar los datos antes de su almacenamiento, permitiendo la adaptación a diferentes grupos dentro de la organización.
- **Básica con área de ensayo y data marts:** Diseñada para líneas de negocio específicas, permite incluir data marts para diferentes departamentos (ventas, inventario, compras), facilitando el acceso a datos específicos por parte de los usuarios finales.

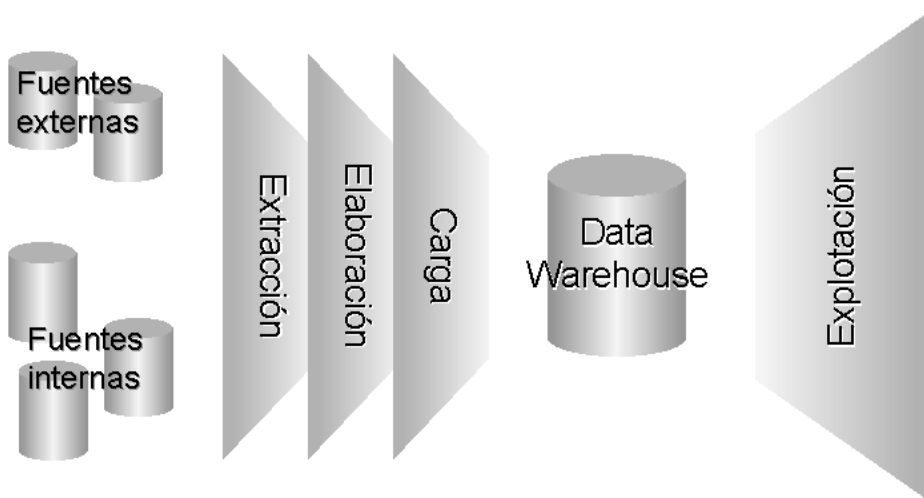
Características del Data Warehouse

- **Integrado:** Los datos en el DWH deben estar completamente integrados y libres de inconsistencias, con posibles niveles específicos ajustados a las necesidades de los usuarios.
- **Temático:** Los datos están organizados por temas, facilitando el acceso y entendimiento, como agrupar información de clientes en una única tabla.
- **Histórico:** El DWH incluye la variable temporal, reflejando el estado del negocio y permitiendo el análisis de tendencias a lo largo del tiempo.
- **No volátil:** La información es permanente, solo se lee y no se modifica; las actualizaciones se realizan agregando nuevos valores, sin alterar los existentes.

Componentes clave del Data Warehouse

- **Cubos de información:** Estas estructuras multidimensionales contienen datos agregados y organizados para facilitar el análisis desde diferentes perspectivas, como tiempo, ubicación y productos.
- **Variables cuantitativas:** Permiten desglosar la información y obtener datos más detallados para un análisis más preciso.
- **Metadatos:** Describen los datos almacenados en el Data Warehouse, proporcionando información relevante como el origen, formato y ubicación de los datos.
- **Funciones ETL:** Estas funciones, encargadas de extraer, transformar y cargar los datos, garantizan la calidad y consistencia de la información almacenada en el Data Warehouse.
- **Middleware:** Esta capa intermedia facilita la comunicación entre las aplicaciones y el Data Warehouse, asegurando una correcta transmisión y procesamiento de los datos.

Procesos clave en la gestión del Data Warehouse



- **Extracción:** obtención de información de las distintas fuentes tanto internas como externas.
- **Elaboración:** filtrado, limpieza, depuración, homogeneización y agrupación de la información.
- **Carga:** organización y actualización de los datos y los metadatos en la base de datos.
- **Explotación:** extracción y análisis de la información en los distintos niveles de agrupación.