

FACULTAD DE INGENIERÍA
CIENCIAS DE LA COMPUTACIÓN Y TELECOMUNICACIONES



Tarea de investigación Data Warehouse

Nombre y Apellido: Robert Lorenzo Navarro Valle

Registro: 219036489

Materia: Sistemas para el soporte de la toma de decisiones

Santa Cruz Bolivia – 2024

¿QUE ES UN DATA WAREHOUSE?

Un Datawarehouse es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinitud de perspectivas y con grandes velocidades de respuesta. La creación de un datawarehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de Business Intelligence.

Normalmente, un data warehouse se aloja en un servidor corporativo o cada vez más, en la nube. Los datos de diferentes aplicaciones de procesamiento de transacciones Online (OLTP) y otras fuentes se extraen selectivamente para su uso por aplicaciones analíticas y de consultas por usuarios.

Data Warehouse es una arquitectura de almacenamiento de datos que permite a los ejecutivos de negocios organizar, comprender y utilizar sus datos para tomar decisiones estratégicas. Un data warehouse es una arquitectura conocida ya en muchas empresas modernas.

La ventaja principal de este tipo de bases de datos radica en las estructuras en las que se almacena la información (modelos de tablas en estrella, en copo de nieve, cubos relacionales... etc.). Este tipo de persistencia de la información es homogénea y fiable, y permite la consulta y el tratamiento jerarquizado de la misma (siempre en un entorno diferente a los sistemas operacionales).

Es una forma de consolidar y administrar datos, provenientes de diversas fuentes, con el propósito de responder a necesidades de información de negocio y tomar decisiones, que antes no eran posibles de realizar.

CARACTERÍSTICAS DE UN DATA WAREHOUSE

Integrado:

- Un Data Warehouse recopila datos de múltiples fuentes heterogéneas, como bases de datos transaccionales, archivos planos, aplicaciones externas, etc., y los integra en un formato coherente y uniforme.
- **Variable en el tiempo:**
- Los datos en un Data Warehouse están etiquetados con un período de tiempo específico, lo que permite análisis históricos. Se almacenan datos históricos, lo cual es esencial para identificar tendencias a lo largo del tiempo.
- **No volátil:**

- Una vez que los datos han sido almacenados en el Data Warehouse, no se modifican ni se eliminan. Esto asegura la estabilidad y consistencia de los datos para análisis y reportes.

Componentes de un Data Warehouse:

1. Fuentes de datos externos:

- Las fuentes de datos son los sistemas de origen desde los cuales se extraen los datos para cargarlos en el Data Warehouse. Pueden incluir bases de datos operativas, archivos de texto, sistemas ERP, aplicaciones de CRM, entre otros.
- **Área de almacenamiento:**
- El área de almacenamiento es donde se almacena toda la información integrada y transformada. Está organizada de manera que facilita el acceso rápido a los datos para su análisis.
- **Meta datos (Metadata):**
- Los metadatos describen el contenido, estructura y uso de los datos en el Data Warehouse. Incluyen descripciones de las fuentes de datos, procesos ETL, esquemas de la base de datos y las relaciones entre los diferentes datos.
- **Data Marts:**
- Son subconjuntos de un Data Warehouse que se centran en un área o departamento específico, como ventas o finanzas. Permiten un análisis más rápido y específico.
- Una colección de datos organizado para un área temática o para un grupo de usuarios • Mejora el desempeño de las consultas para un departamento o área de desempeño • Depende del Data Warehouse o directamente de los sistemas operacionales
- D.M : Son repositorios de un área específica. Ej. Comercial, finanzas, etc. D.W : Se construye con todos los D.M

Aplicaciones del Data Warehouse

- Administración de relaciones con clientes (CMR)
- Administración del conocimiento (KM)
- Minería de datos (Data Mining)
- Alertas (Balanced Scorecard)
- Detección de fraudes • Evaluación de créditos • Etc