



Instituto Politécnico Nacional  
Escuela Superior de Cómputo



# Bases de almacén y fuentes de datos

De Luna Ocampo Yanina  
Medina Barreras Daniel Ivan



# Contenidos

01

Definición bases de  
almacén

02

Tipos de bases de  
almacén

03

Bases de almacenamiento  
VS  
Bases transaccionales

04

Arquitectura y  
ventajas

# Contenidos

05

Definición Fuentes de  
datos

06

Tipos Fuentes de datos

07

Procesamiento de  
datos

08

Beneficios y  
herramientas



“Predicts that the  
Global Datasphere will grow from 33  
Zettabytes (ZB) in 2018 to 175 ZB by 2025.”

—IDC (International Data Corporation)



# 01

## Definición bases de almacén

¿Qué son las bases de almacén?

# Bases de almacén



## Definición

Una base de almacén es una base de datos que organiza y analiza datos de diferentes fuentes para tomar decisiones empresariales.

- Organización y análisis de datos de múltiples fuentes.
- Uso en la toma de decisiones empresariales.
- Consolidación y estructuración de grandes volúmenes de datos.



## Función

Almacenar y recuperar grandes volúmenes de datos para consultas y análisis complejos

- Capacidad de almacenamiento de grandes cantidades de datos
- Optimización para consultas y análisis complejos
- Mejora la eficiencia y velocidad de los procesos de análisis de datos.



# 02

## Tipos de bases de almacén

Definición y características



# Características clave



	Almacenamiento	Transaccionales
Definicion	Diseñadas para almacenar grandes volúmenes de datos	Diseñadas para la gestión de datos en tiempo real y mantener la integridad de los datos.
Objetivo	proporcionar un acceso fácil y rápido a grandes volúmenes de datos históricos para su análisis y reporting.	Para apoyar la toma de decisiones empresariales en tiempo real
Optimizacion	Están optimizadas para realizar consultas y análisis complejos de datos.	Están optimizadas para realizar operaciones en línea.
Integridad de los datos	Se mantiene a través de procesos de ETL, ELT, etc. técnicas de modelado e indexación.	A través de reglas de negocio, restricciones de integridad referencial y validación de datos.



# 03

## Bases de almacén **vs** Bases transaccionales

Definición y Características



# Generales



## SQL

SQL es un tipo de base de datos relacional que utiliza tablas para almacenar datos que pueden ser consultados y manipulados mediante el uso del lenguaje SQL.

- Esquema fijo
- Lenguaje SQL estructurado
- Alta consistencia



## NoSQL

NoSQL es un tipo de base de datos no relacional que utiliza modelos de datos flexibles para almacenar y manejar grandes volúmenes de datos no estructurados o semi-estructurados.

- Sin esquema fijo
- Modelos de datos flexibles
- Alta disponibilidad y escalabilidad horizontal.



# Storage



## Data Warehouses

Almacenamiento centralizado de datos.



## Data Lakes

Repositorio centralizado de datos sin estructura

# Transaccionales



## Data Marts

Almacén de datos especializado.



## In-Memory Databases

Base de datos ultrarrápida en memoria.

# 04

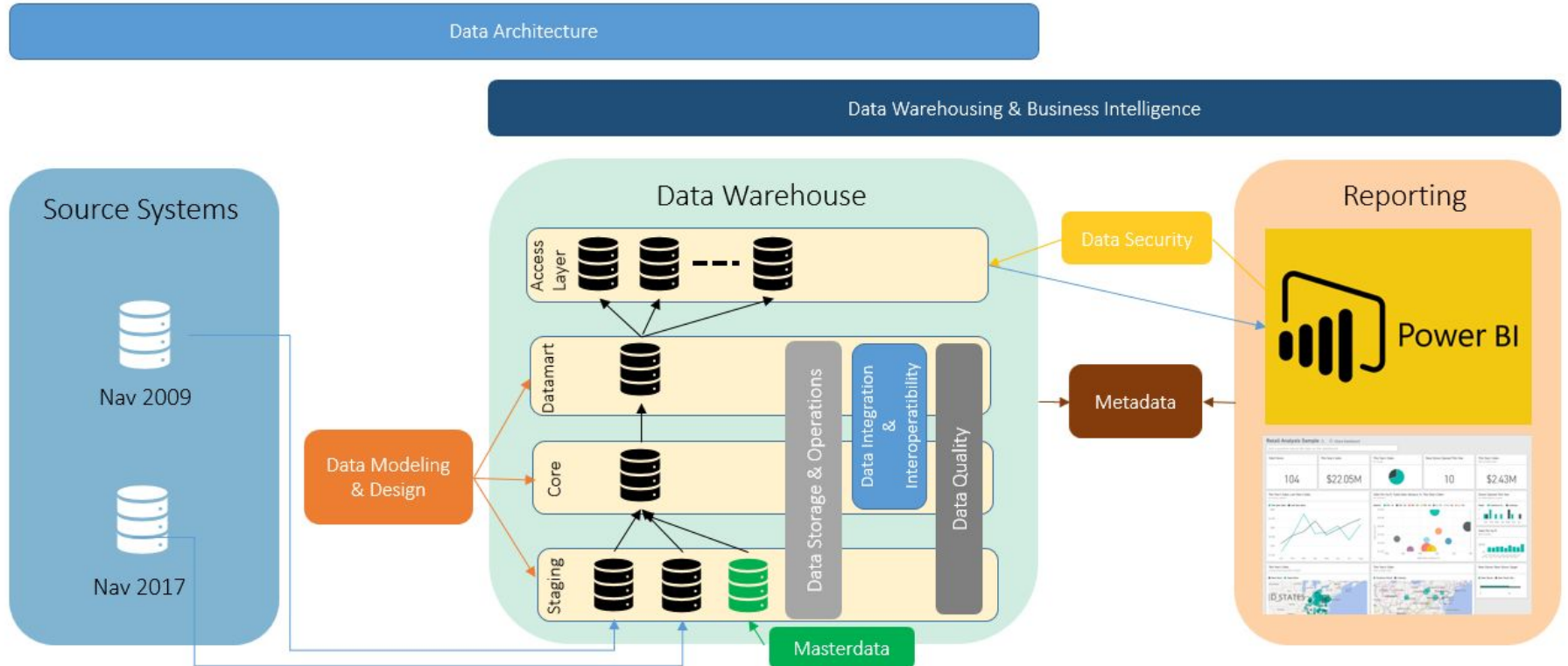
## Arquitectura Y Ventajas

¿Cómo puede ayudar a su empresa?



# Data Governance

Esquema de la aplicación de los puntos de Data Governance



## Integridad

Uniformidad entre tipos de datos y fuentes dispares.

## Seguridad y recuperación

Se duplican y hacen copias de seguridad automáticas.

## No volátiles

Cuando la información entra, no cambia y es estable.

I

O

N

A

C

C

## VENTAJAS

❖❖❖❖❖

## Analíticas de negocio

Mejores tomas de decisiones

## Consultas más rápidas

Podemos consultar grandes cantidades de datos de forma rápida.

## Calidad de datos

Integridad en los datos

# 05

## Definición de fuentes de datos

Entendiendo las fuentes de datos



# Definición.

Las fuentes de datos son los lugares o sistemas de los cuales se recopilan los datos. Pueden ser internas o externas a la empresa y pueden incluir bases de datos transaccionales, registros de clientes, registros de ventas, redes sociales, entre otros.

**En resumen:** Las fuentes de datos son los lugares o sistemas de dónde se obtienen los datos.





# 06

## Tipos de fuentes de datos

Categorías



# Fuentes de **datos**

1

Internas

2

Estructuradas

3

Semi-estructuradas

Tipos

Externas

4

No estructuradas

5

Transaccionales

6



07

# Procesamiento de datos

Procesos

# ETL

Conjunto de técnicas y herramientas utilizadas para el ETL.

1. Extracción: se extraen de diferentes fuentes, como bases de datos, servicios web, archivos, etc.
2. Transformación: se cambian a un formato común y se limpian para eliminar datos duplicados, incompletos o incorrectos.
3. Carga: los datos se cargan en un almacén de datos centralizado.
4. Análisis: se utilizan herramientas de análisis de datos y se convierten en conocimiento útil para la toma de decisiones empresariales.



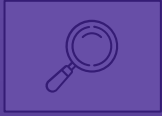
# 08

## Beneficios y herramientas

Por qué usarlo



# Herramientas



## ETL

Extract, Transform, Load.

Ej: Talend, DataStage, SSIS

Spark,, kafka



## Bases de Datos

Necesitan de una base de datos que permita la consulta y manipulación de datos.

Ej: Oracle, SQL Server, MySQL  
Hadoop



## Visualización

Presentar datos de manera visual y accesible.

Ej: Tableau, PowerBI



## Seguridad

Asegurar la integridad y seguridad de los datos almacenados.

Ej: Symantec Data Loss Prevention



## Planificación y Monitoreo

Planificar y monitorear la carga y transformación de los datos.

Ej: BMC Control-M



## Modelado de Datos

Diseñar y modelar el esquema del almacén de datos.

Ej: ERwin, Visio, SQL Power Architect



## Minería de Datos

Analizar grandes conjuntos de datos para encontrar patrones y tendencias.

Ej: RapidMiner, KNIME, SAS



## Gestión de Metadatos

Gestionar la información sobre los datos almacenados en el almacén de datos.

Ej: Informatica Metadata Manager, IBM InfoSphere



# Gracias!

¿Preguntas?



CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon and infographics & images by Freepik

# Fuentes

- [https://docs.aws.amazon.com/es\\_es/machine-learning/latest/dg/creating-and-using-datasources.html](https://docs.aws.amazon.com/es_es/machine-learning/latest/dg/creating-and-using-datasources.html)
- <https://support.microsoft.com/es-es/office/introducci%C3%B3n-a-los-tipos-de-datos-y-las-propiedades-de-los-campos-30ad644f-946c-442e-8bd2-be067361987c>
- <https://learn.microsoft.com/es-es/sql/reporting-services/report-builder/generating-data-feeds-from-reports-report-builder-and-ssrs?view=sql-server-ver16>

