



ESCUELA DE  
INGENIERÍA EN CIENCIAS Y SISTEMAS  
FACULTAD DE INGENIERÍA  
UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



Día, Fecha:

Sábado, 08/03/2025

Hora de inicio:

08:00AM

# Sistemas organizacionales y gerenciales 2[N]

José Luis Reynoso Tiú



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

# Clase 6



# PAREJAS PRÁCTICA



<https://docs.google.com/spreadsheets/d/1DnfKnte8-g7rIGcRaYVSwWscSUKytUCgxZKIQhNrow4/edit?usp=sharing>

# Información de contacto



## FOROS DE UEDI

Foro semanal para la resolución de dudas



**2655896221401@INGENIERIA.USAC.EDU.GT**

[SOG2] Asunto



# DATA WAREHOUSE

Un Data Warehouse es un almacén electrónico donde generalmente una empresa u organización mantiene una gran cantidad de información. Los datos de un data warehouse deben almacenarse de forma segura, fiable, fácil de recuperar y fácil de administrar.

Un data warehouse es un repositorio unificado para todos los datos que recogen los diversos sistemas de una empresa. El repositorio puede ser físico o lógico y hace hincapié en la captura de datos de diversas fuentes sobre todo para fines analíticos y de acceso.

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.

# ¿Cómo funciona un Data Warehouse?

Un data warehouse funciona como un repertorio central. La información proviene de una o varias fuentes de datos, como sistemas transaccionales u otras bases de datos relacionales.

Los datos pueden ser estructurados, semiestructurados o no estructurados. Una vez que están integrados en el warehouse, son tratados y transformados. Los usuarios pueden luego acceder a ellos con la ayuda de herramientas de Business Intelligence, de clientes SQL o de tablas.

Agregando la información en el mismo lugar, una empresa logra tener una visión de conjunto de su clientela o de otros elementos cruciales

# Tipos de Data Warehouse


- **Enterprise Data Warehouse (EDW)** es un almacén centralizado que proporciona servicios de apoyo a la toma de decisiones en toda la empresa.
- **Operational Data Store (ODS)** es simplemente un almacén de datos necesario cuando ni el Data Warehouse ni los sistemas OLTP soportan las necesidades de reporte de las organizaciones.
- **Data Mart** es un subconjunto del data warehouse, diseñado especialmente para una línea de negocio particular, como ventas, finanzas, etc.

365✓

What Is a Data Warehouse?

Share



Watch on  YouTube

# WHAT IS DATA WAREHOUSING?



# ¿Quién utiliza un DWH?

Los Data Warehouse son utilizados por todas las empresas que deben tratar grandes volúmenes de datos o que recolectan datos de múltiples fuentes. Son utilizados también por empresas que desean acceder con mayor facilidad a sus datos.

Para cualquier empresa que busque asistencia en la toma de decisiones, los Data Warehouse pueden resultar pertinentes. Este es también el caso para los usuarios que buscan gestionar informes, gráficos y diagramas a partir de datos. Los Data Warehouses tienen su lugar en todos los sectores de actividad. Sin embargo, son utilizados de formas muy diferentes según la industria.

# OLAP Y OLTP

El procesamiento analítico en línea (OLAP) y el procesamiento de transacciones en línea (OLTP) son sistemas de procesamiento de datos que ayudan a almacenar y analizar datos empresariales. Puede recopilar y almacenar datos de múltiples fuentes, como sitios web, aplicaciones, medidores inteligentes y sistemas internos.

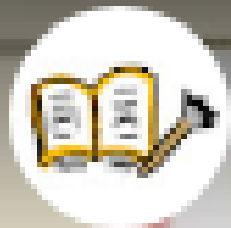
El OLAP combina y agrupa los datos para que pueda analizarlos desde diferentes puntos de vista. Por el contrario, el OLTP almacena y actualiza los datos transaccionales de manera confiable y eficiente en grandes volúmenes. Las bases de datos de OLTP pueden ser uno de los diferentes orígenes de datos de un sistema de OLAP.

El propósito principal del procesamiento analítico en línea (OLAP) es analizar los datos agregados, mientras que el propósito principal del procesamiento de transacciones en línea (OLTP) es procesar las transacciones de bases de datos.

Los sistemas de OLAP se utilizan para generar informes, realizar análisis de datos complejos e identificar tendencias. Por el contrario, los sistemas de OLTP se utilizan para procesar pedidos, actualizar el inventario y administrar las cuentas de los clientes.

Otras diferencias entre OLTP Y OLAP son:

- Formato de datos
- Arquitectura de datos
- Rendimiento
- Requisitos



Explain By Example: OLTP vs OLAP



Explain By Example

Share



WHAT IS

OLTP VS. OLAP?

Watch on  YouTube

# Bases de datos, DWH y Data Lake





# Diferencias entre DWH y Data Lake

## Data Lake:

- Conserva **todos los datos**, incluso los que no se usan actualmente.
- Almacena datos en **bruto**, sin procesar.
- Es **escalable** y **económico** para almacenar grandes cantidades de datos.
- Soporta **todos los tipos de datos**, incluyendo datos no tradicionales.
- Usa el enfoque "**Schema on Read**".

## Data Warehouse:

- Solo conserva los datos **necesarios** para informes específicos.
- Almacena datos **procesados** y **estructurados**.
- Puede ser **costoso** y **difícil de escalar**.
- No suele soportar **datos no tradicionales**.
- Usa el enfoque "**Schema on Write**".

# Schema-on-read y Schema-on-write

Tradicionalmente se crea una base de datos relacional con tablas especificando el esquema. Luego de configurar los esquemas y creadas las tablas, se puede comenzar a cargar los datos. Se pueden cargar datos en masa o en pequeños lotes, pero siempre se ingieren basados en el esquema configurado. Y, una vez que los datos se cargan en la tabla, se puede comenzar a ejecutar consultas analíticas. Esto es el **schema-on-write (esquema en escritura)**.

Si se tienen datos y no se conoce el esquema. Para crear la tabla y analizar los datos, se necesita conocer el esquema de antemano, por lo que el schema-on-write falla en esta situación.

¿Qué pasa si se cargan los archivos/datos y luego se analiza su estructura y características? Esto se denomina **schema-on-read (esquema en lectura)**.

# DATA MINING

La Minería de Datos es un proceso de encontrar patrones potencialmente útiles en conjuntos de datos enormes. Es una habilidad multidisciplinaria que utiliza aprendizaje automático, estadísticas e inteligencia artificial para extraer información y evaluar la probabilidad de eventos futuros. Los conocimientos derivados de la Minería de Datos se utilizan para marketing, detección de fraudes, descubrimiento científico, etc.

La Minería de Datos se trata de descubrir relaciones ocultas, insospechadas y previamente desconocidas pero válidas entre los datos. También se le llama Descubrimiento de Conocimiento en Datos (KDD), extracción de conocimiento, análisis de datos/patrones, recolección de información, etc.

# ¿Por qué es importante la minería de datos?

La minería de datos constituye un elemento crucial para cualquier iniciativa de análisis exitosa.

Las empresas pueden utilizar el proceso de descubrimiento de conocimientos para aumentar la confianza de los clientes, encontrar nuevas fuentes de ingresos y lograr que los clientes regresen. **Una minería de datos eficaz resulta de gran utilidad para diversos aspectos de la planificación empresarial y la administración de las operaciones**

# PROCESO DE MINERÍA DE DATOS





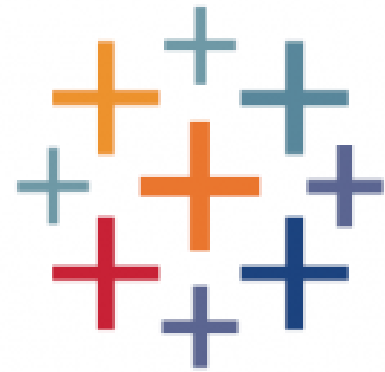
# BUSINESS INTELLIGENCE

Business Intelligence, o inteligencia empresarial en español, es un término que abarca un conjunto de estrategias, tecnologías y aplicaciones enfocadas en recolectar, analizar y transformar datos de negocio en información útil y accionable. Su objetivo es brindar a las empresas la capacidad de tomar decisiones mejor informadas y basadas en datos.

## Beneficios de BI:

- Aumento de la eficiencia
- Mayor rentabilidad
- Reducción de riesgos
- Mejora de la satisfacción del cliente

# Herramientas de BI



+ a b l e a u

qlik ®



# Cuadrante mágico de Gartner

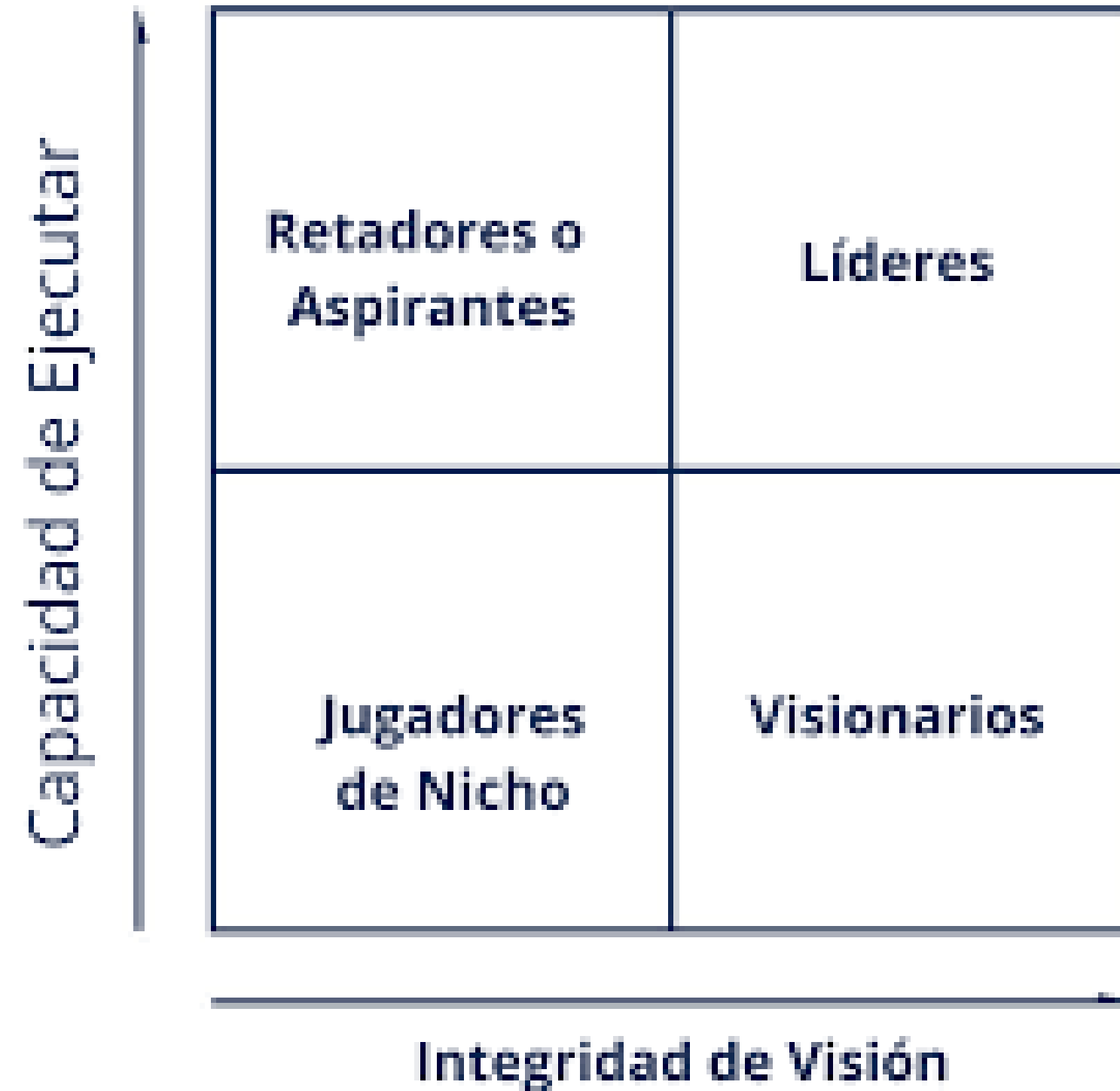
Un Magic Quadrant de Gartner (literalmente, "cuadrante mágico" en español) recoge la culminación de la investigación de un mercado específico, y te proporciona una visión panorámica de las posiciones relativas de sus competidores.

Gracias a su representación gráfica y un conjunto uniforme de criterios de evaluación, cada Magic Quadrant te ayuda a determinar rápidamente cómo los proveedores de tecnología están ejecutando sus ideas y qué tan bien se están desempeñando en comparación con la visión de mercado de Gartner.

# Cuadrante mágico de Gartner

- **Líderes:** se desempeñan bien de acuerdo a la visión actual del mercado y están bien posicionados para el futuro.
- **Visionarios:** entienden hacia dónde se dirige el mercado o tienen una visión para cambiar sus reglas, pero su capacidad de ejecución aún es limitada.
- **Jugadores de nicho:** se centran con éxito en un segmento pequeño, o están dispersos y no innovan por encima de la media ni superan a los demás.
- **Retadores o aspirantes:** se desempeñan bien hoy e incluso pueden dominar un gran segmento, pero no demuestran poseer una buena comprensión de hacia dónde se dirige el mercado.

# Cuadrante mágico de Gartner





# Gartner® Magic Quadrant™ 2024 sobre plataformas de Analítica y Business Intelligence

