



# Inteligencia de Negocio: Conceptos y Aplicaciones con PENTAHO

## Tema 3

- Elementos de la Arquitectura de Inteligencia de Negocio -



Junio 2008

V 1.0



# TEMA 3

## Necesidad

### Arquitectura

### Infraestruct.

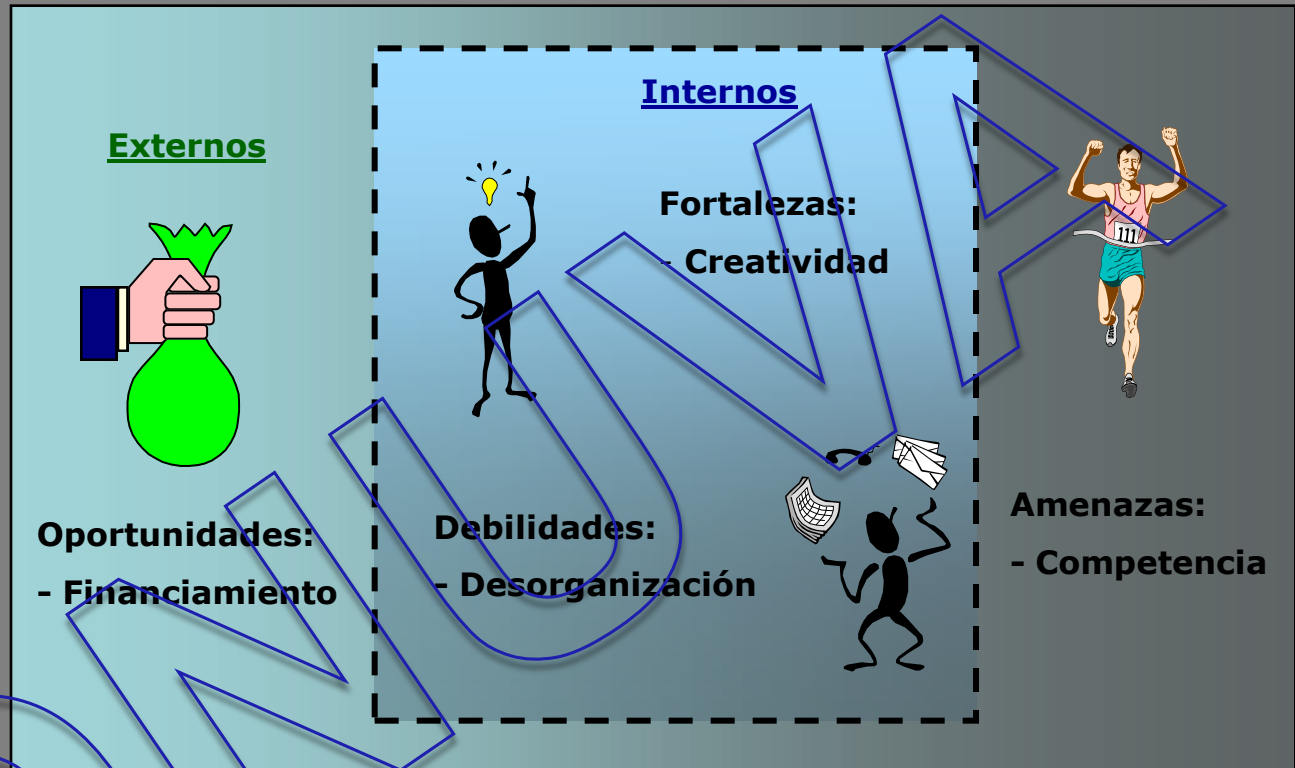
- No-Técnica
- Técnica

### Data Warehouse

### OLAP

### Tableros de Comandos

### Minería de Datos



# Necesidad del Negocio



# TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

Oportunidades / Problemas

Requerimientos de Información

Internos		Externos
<ul style="list-style-type: none"><li>- Capacidad de producción diaria.</li><li>- Capacidad de distribución.</li></ul>	<b>Oportunidad</b> Aumento del consumo de petróleo	<ul style="list-style-type: none"><li>- Capacidad de producción diaria de la competencia.</li><li>- Demanda de los clientes.</li></ul>
<ul style="list-style-type: none"><li>- Flujo de caja.</li><li>- Costos de producción.</li></ul>	<b>Problema</b> Altos intereses de las fuentes de financiamiento	<ul style="list-style-type: none"><li>- Tasas de interés.</li><li>- Regulaciones monetarias.</li></ul>

## Necesidad del Negocio



# TEMA 3

Necesidad

Arquitectura

Infraestruct.

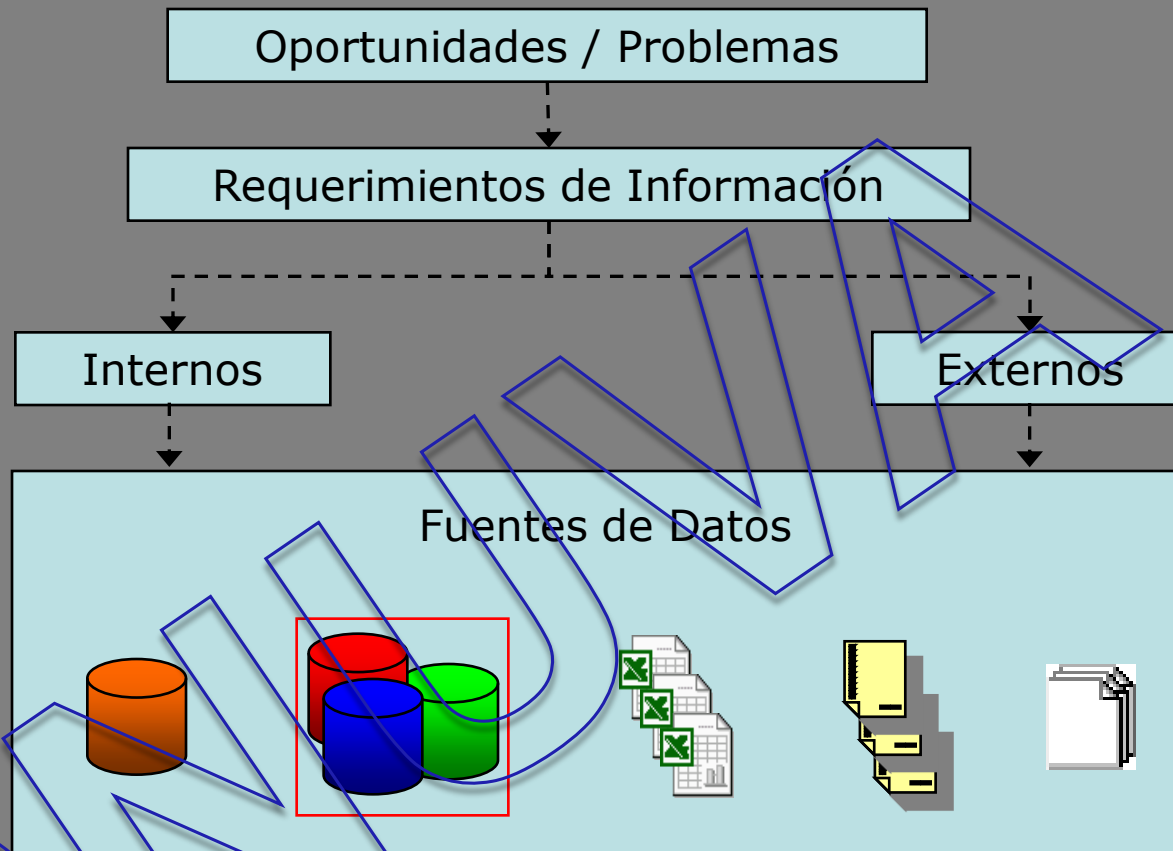
- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos



## Arquitectura de Inteligencia de Negocio



# TEMA 3

Necesidad

Arquitectura

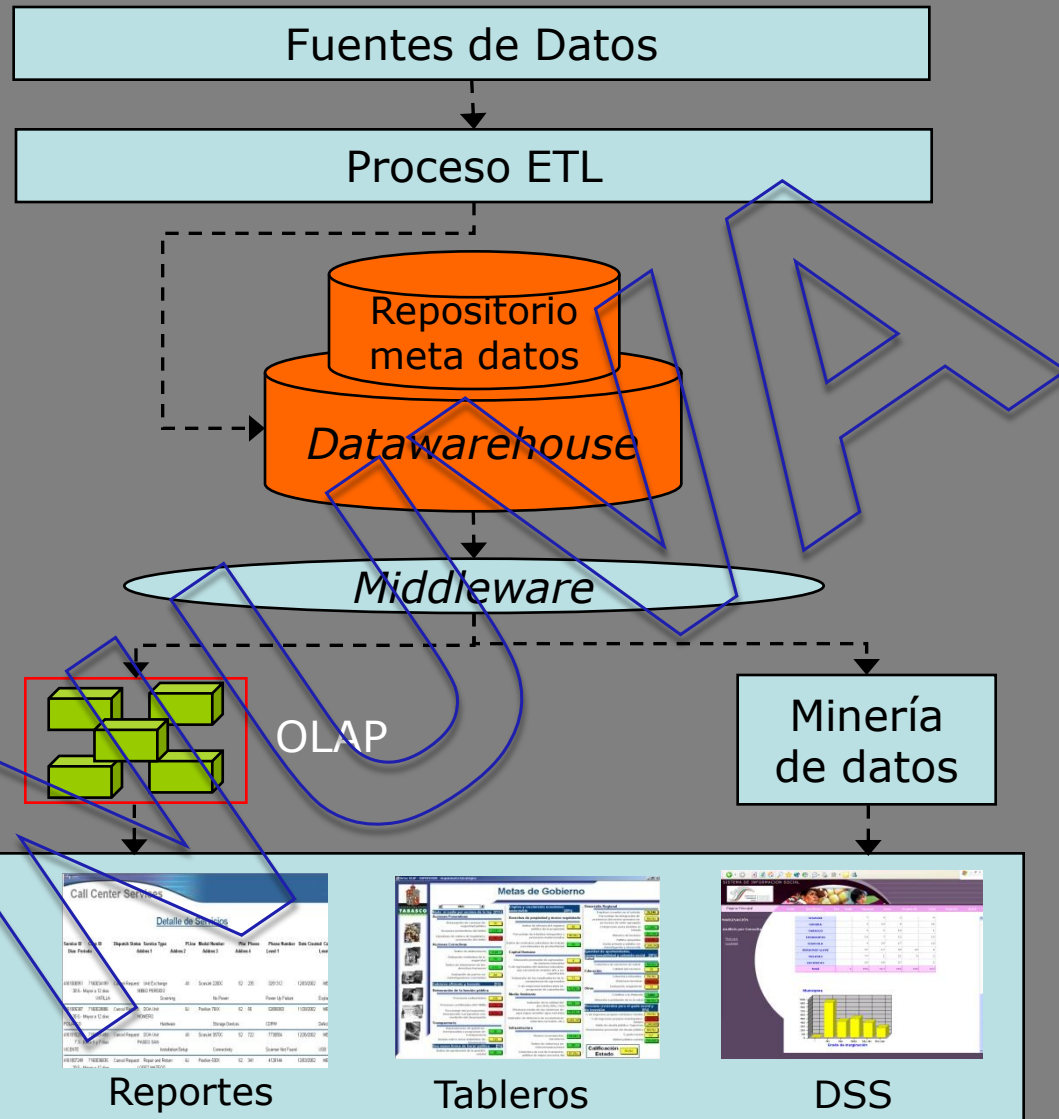
Infraestruct.  
- No-Técnica  
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos



## Arquitectura de Inteligencia de Negocio



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

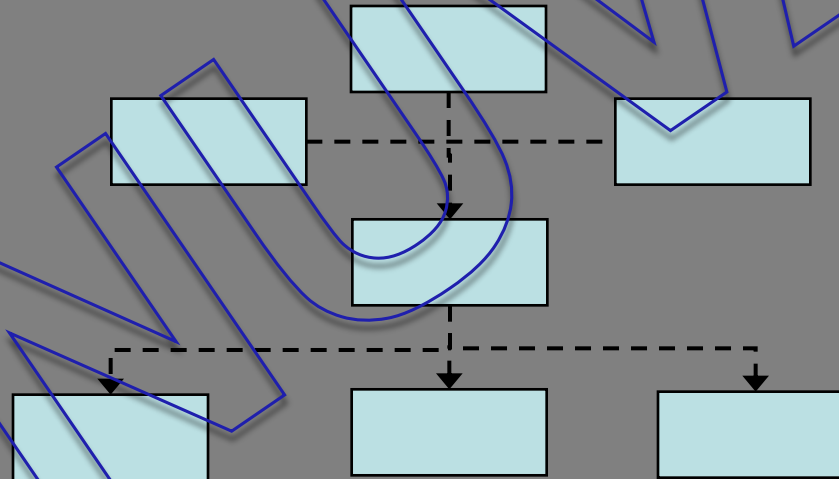
Tableros de Comandos

Minería de Datos

## Infraestructura No Técnica

Permitirá a la organización evitar que el ambiente de IN de soporte a la decisión se fragmente, lo cual proporcionará una integración entre las aplicaciones de IN y el negocio. Algunos de sus componentes son los siguientes:

✓ **Modelo de la función del negocio:**



## Infraestructura No-Técnica



Pág: 6



### TEMA 3

## ✓ Modelo de los procesos de negocio:

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

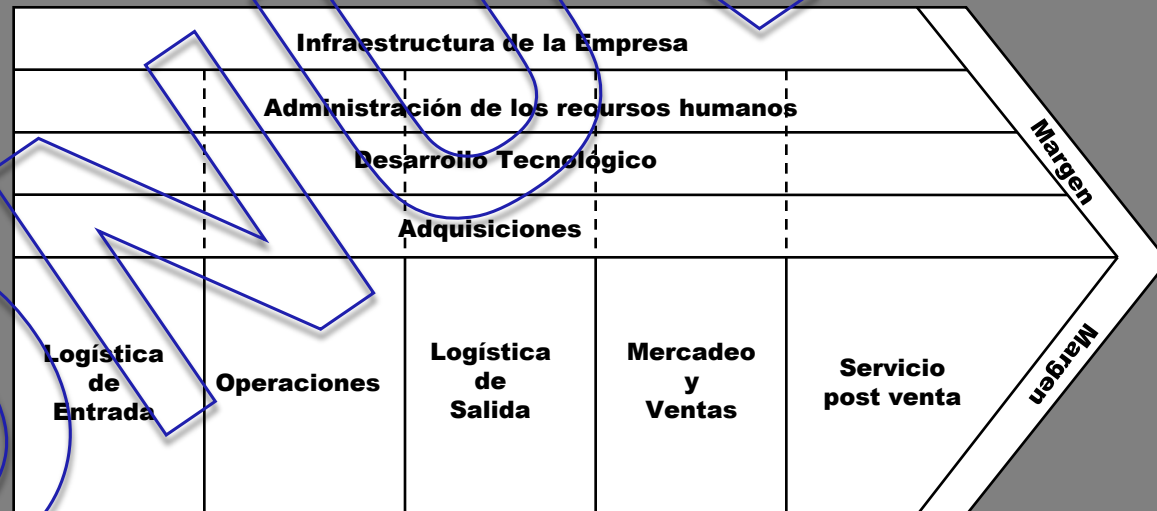
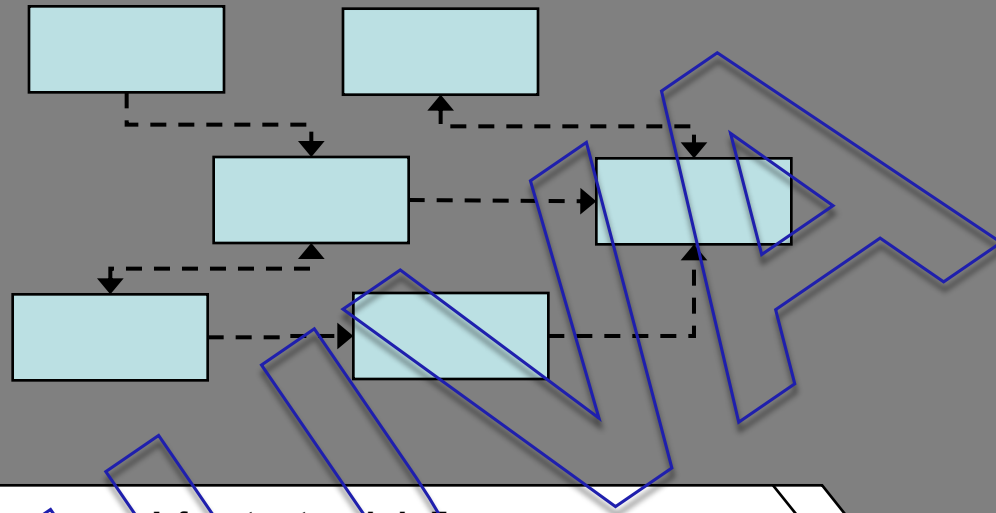
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos



## Infraestructura No-Técnica



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

Data Warehouse

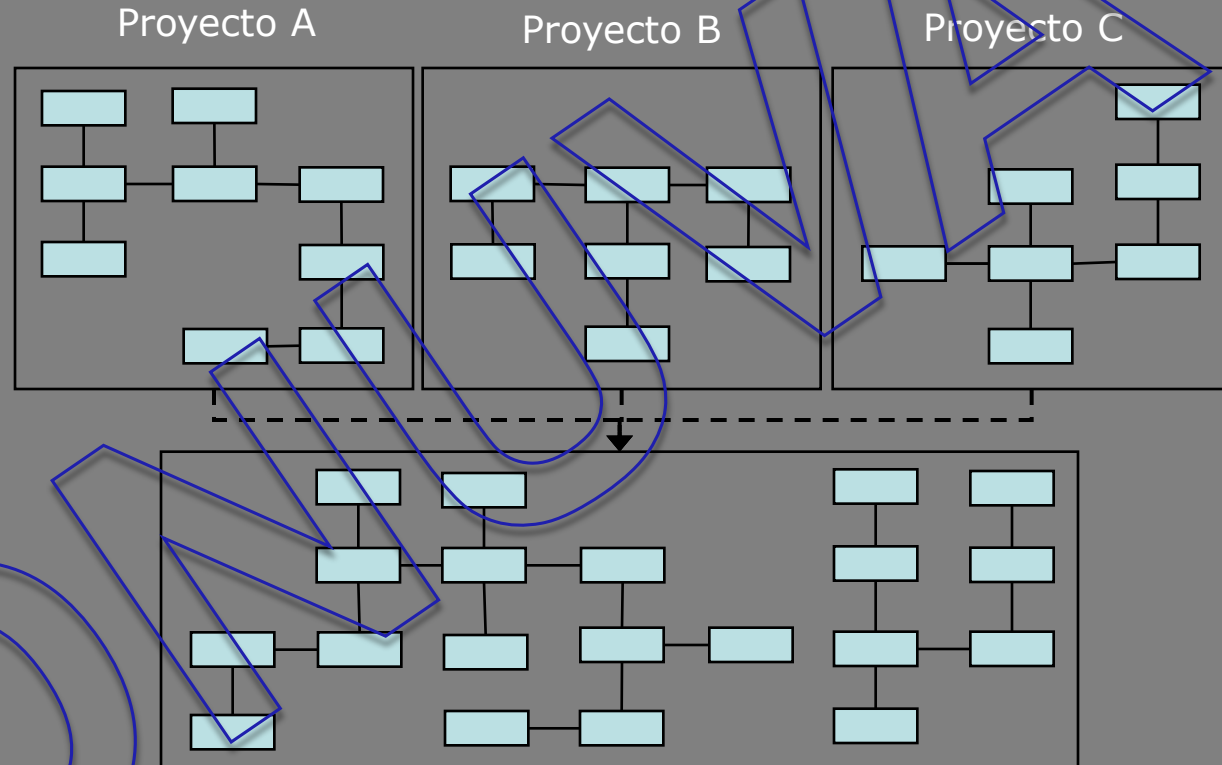
OLAP

Tableros de Comandos

Minería de Datos

### ✓ Modelo (lógico) de datos del negocio:

Su diseño representa el primer paso para explotar la información. Incorpora descripción de los objetos de negocio y usuarios, así como las diferentes interacciones.



Modelo lógico de datos del negocio

## Infraestructura No-Técnica



Pág: 8



## TEMA 3

Necesidad

Arquitectura

**Infraestruct.**

- No-  
Técnica

- Técnica

Data  
Warehouse

OLAP

Tableros de  
Comandos

Minería de  
Datos

### ✓ **Modelo de gestión del negocio:**

Define todo el proceso de conversión de los datos del negocio en información de interés para la toma de decisión.

Los principales objetivos de este modelo son:

1. Evaluar el desempeño de los diferentes procesos del negocio.
2. Contar con información histórica estadística.
3. Tener la información normalizada para el reporte de gestión mensual (o trimestral).
4. Generar escenarios para simular cambios en el comportamiento de variables de interés.

# Infraestructura No-Técnica



Pág: 9



### TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

Los elementos de este modelo son:

Datos de entrada

Finanzas

Clientes

Procesos internos

Formación y desarrollo

Proceso de conversión

Modelo de Gestión del Negocio

Cuentas

Fórmulas

Escenarios

Reportes de salida

Estados Financieros

Indicadores Estratégicos

Indicadores Operativos

## Infraestructura No-Técnica



## TEMA 3

Necesidad

Arquitectura

**Infraestruct.**

- No-Técnica
- **Técnica**

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

# Infraestructura Técnica

Una completa solución de Inteligencia de Negocios suele ejecutarse sobre una plataforma cliente-servidor. Debido al gran volumen de datos que maneja demanda grandes capacidades de almacenamiento, bastante poder de procesamiento y una estructura de red sólida. El software sobre el cual se implemente debe ser eficiente, robusto, soportable y adaptable a los cambios.

## ✓ La plataforma de hardware

Características:

- Poder
- Escalabilidad



# Infraestructura Técnica



Pág: 11



## TEMA 3

Necesidad

Arquitectura

**Infraestruct.**

- No-Técnica
- Técnica

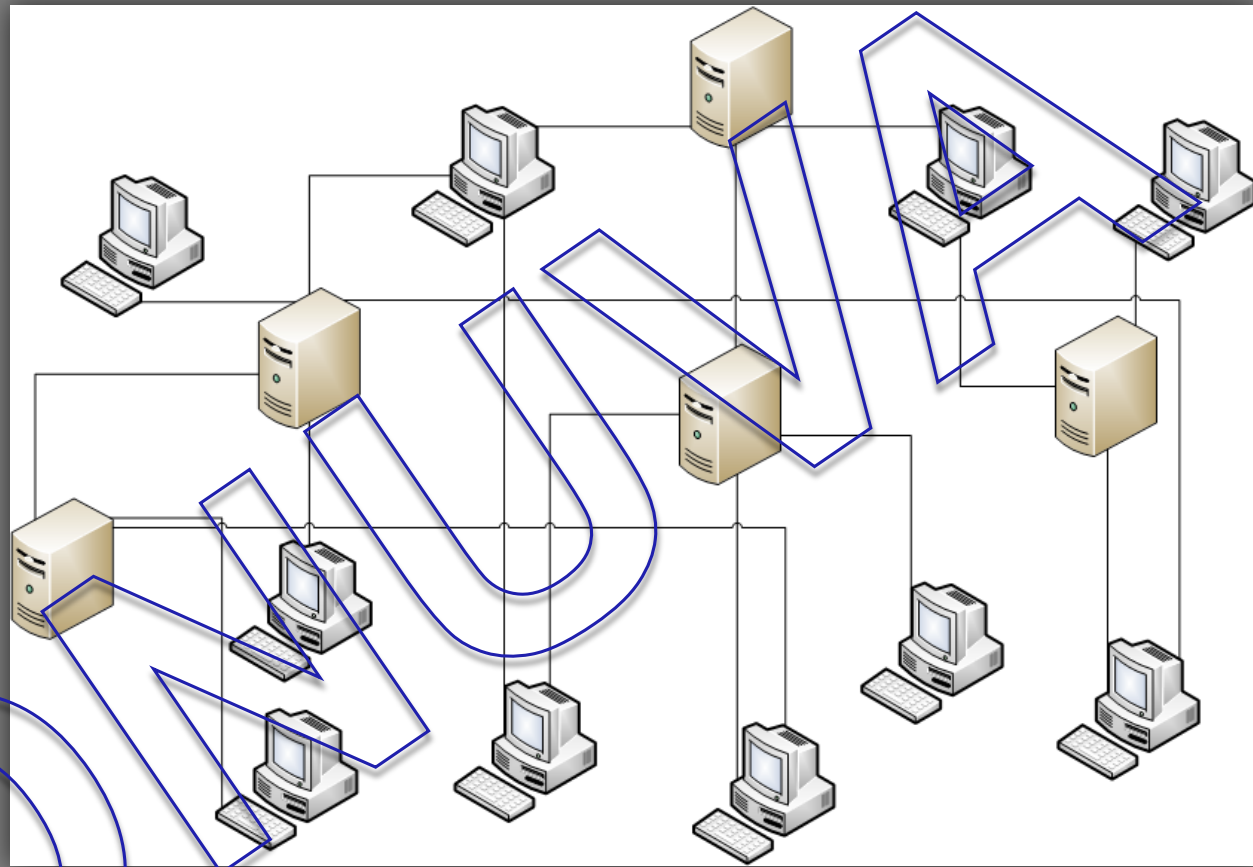
Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

## Caos Controlado



## Infraestructura Técnica



## TEMA 3

**Necesidad**

**Arquitectura**

**Infraestruct.**

- No-  
Técnica

- **Técnica**

**Data  
Warehouse**

**OLAP**

**Tableros de  
Comandos**

**Minería de  
Datos**

# Plataforma de Operación Óptima

Para minimizar el caos en una nueva plataforma de red se debe considerar:

- ✓ Las nuevas plataformas de base de datos deben ser compatibles con la configuración del hardware.
- ✓ El Sistema Gestor de Base de Datos (SGDB) debe poder crecer tanto en capacidad de almacenamiento como en resistencia al uso intensivo en la plataforma de hardware elegida.
- ✓ La escalabilidad es el principal objetivo a alcanzar.
- ✓ La elección de la plataforma debe estar dirigida por la necesidad de conseguir interoperatividad entre las distintas plataformas de hardware.
- ✓ El costo y el Retorno de la Inversión (ROI) son calificadores a controlar.

## Infraestructura Técnica



Pág: 13



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

# Plataforma de Middleware

El Middleware es un software de conectividad que ofrece un conjunto de servicios que hacen posible el funcionamiento de aplicaciones distribuidas sobre plataformas heterogéneas. Funciona como una capa de abstracción de software distribuida, que se sitúa entre las capas de aplicaciones y las capas inferiores (sistema operativo y red).



## Infraestructura Técnica



## TEMA 3

**Necesidad**

**Arquitectura**

**Infraestruct.**

- No-Técnica
- Técnica

**Data Warehouse**

**OLAP**

**Tableros de Comandos**

**Minería de Datos**

## Plataforma de SGBD

La infraestructura de la base de datos cambia con el tamaño y complejidad del ambiente de inteligencia de negocios.

- ✓ Negocios pequeños y medianos demandan bases de datos relacionales.
- ✓ Negocios medianos y grandes demandan bases de datos relacionales y multidimensionales.

**Data Warehouse**

**Infraestructura Técnica**



Pág: 15



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

# Características de la Plataforma de SGBD

Una plataforma de base de datos para ser utilizada en IN debe cumplir los siguientes requerimientos:

- ✓ Soporte de red
- ✓ Manejo de *triggers* y procedimientos.
- ✓ Soporte y monitoreo administrativo.
- ✓ Utilización de caché y buffer de datos.
- ✓ Manejo óptimo de procesos concurrentes.
- ✓ Compatibilidad con discos, buses, memorias y procesadores múltiples y de gran capacidad.
- ✓ Tolerancia a fallos, soporte para checkpoints, recuperación y reinicio rápido.
- ✓ Optimizadores de consultas.
- ✓ Indexación avanzada.
- ✓ Sistema de seguridad robusto y de acceso restringido.

## Infraestructura Técnica



Pág: 16



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

## Situaciones comunes en un entorno sin Data Warehousing

- ✓ No puedo encontrar la data que necesito.
- ✓ No puedo obtener la data que necesito.
- ✓ No entiendo la data que necesito.
- ✓ No puedo usar la data que necesito.



### Causas

- ✓ La data está esparcida por toda la red interna.
- ✓ Existen muchas versiones con muchas diferencias.
- ✓ No existe documentación de la data.
- ✓ La documentación existente es deficiente.
- ✓ La data se encuentra encriptada o necesita ser transformada.

# Data Warehouse



Pág: 17



## TEMA 3

**Necesidad**

**Arquitectura**

**Infraestruct.**

- No-Técnica
- Técnica

**Data Warehouse**

**OLAP**

**Tableros de Comandos**

**Minería de Datos**

## Data Warehouse

Es un almacén único, completo y consistente de datos obtenidos desde una amplia variedad de orígenes o fuentes de datos y disponibles a usuarios finales de manera tal que puedan entenderlos o utilizarlos en un contexto de negocio.



## Data Warehousing

Es el proceso de transformar la data en información y colocarla a la disposición de los usuarios en el menor tiempo posible.

# Data Warehouse



Pág: 18



**TEMA 3**

Necesidad

Arquitectura

Infraestruct.

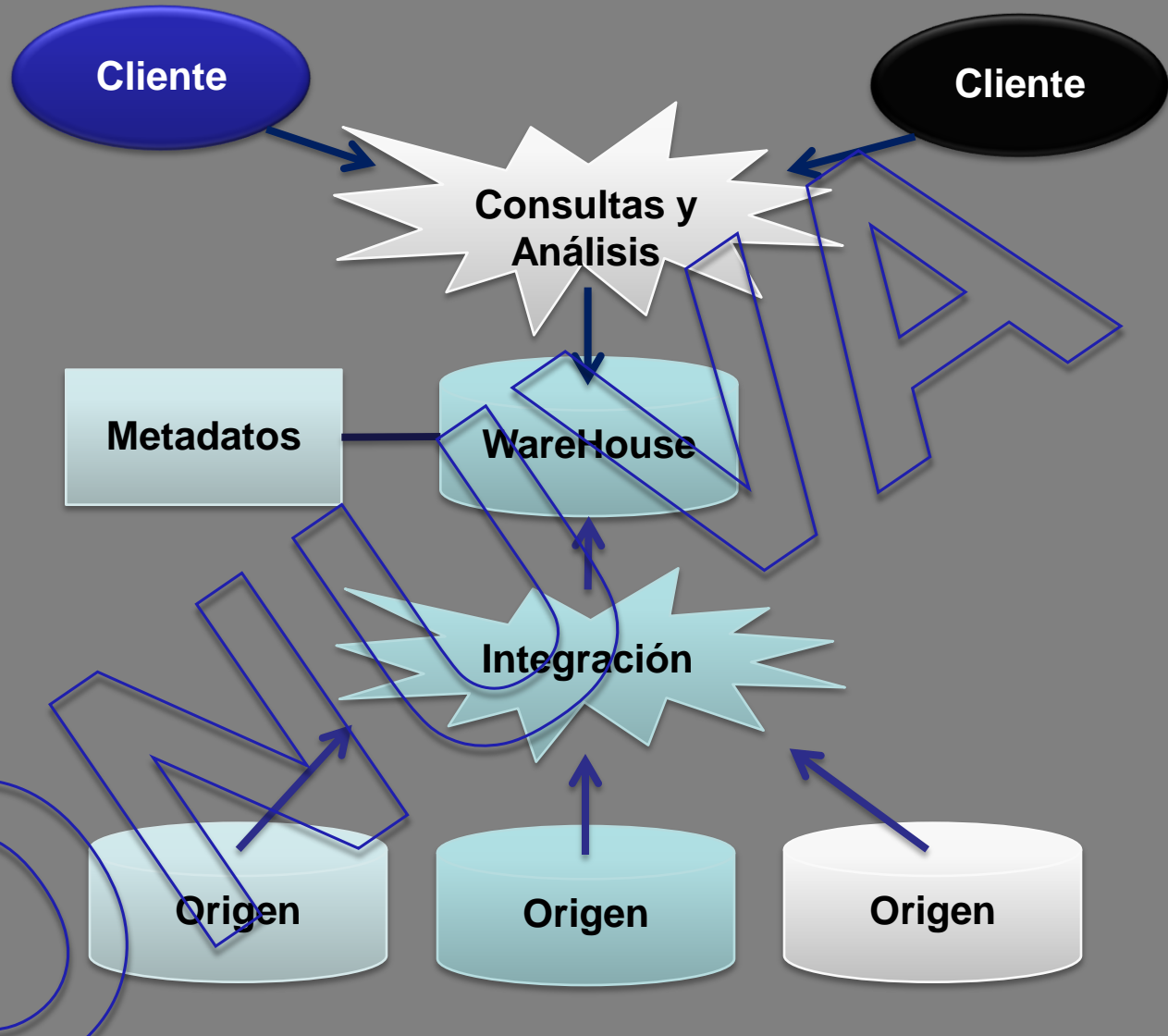
- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos



# Data Warehouse



## TEMA 3

**Necesidad**

**Arquitectura**

**Infraestruct.**

- No-Técnica
- Técnica

**Data Warehouse**

**OLAP**

**Tableros de Comandos**

**Minería de Datos**

## Justificación de hacer Data Warehousing

- ✓ La data debería estar integrada en toda la organización.
- ✓ La información resumida tiene un gran valor para la organización.
- ✓ La data histórica tiene la clave para estudiar el pasado, entender el presente y mejorar el futuro de la organización.

## Ventajas de utilizar un Data Warehouse

- ✓ Realización rápida de consultas.
- ✓ Las consultas son privadas y no están disponibles fuera del data warehouse.
- ✓ Los orígenes de datos no son afectados.
- ✓ Pueden realizarse consultas cuando los orígenes de datos no se encuentran disponibles.
- ✓ Realizar reportes, OLAP y minería de datos.

# Data Warehouse



Pág: 20



## TEMA 3

**Necesidad**

**Arquitectura**

**Infraestruct.**

- No-Técnica
- Técnica

**Data Warehouse**

**OLAP**

**Tableros de Comandos**

**Minería de Datos**

## OLAP

OLAP es el acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing). Y su objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras multidimensionales (o Cubos OLAP) que contienen datos resumidos de grandes Bases de Datos (un data warehouse, por ejemplo).

### Cubo OLAP

- ✓ Un cubo OLAP es una estructura de datos multidimensional que permite analizar con rapidez la data.
- ✓ Su contenido son valores numérico medibles y sus ejes se denominan dimensiones.
- ✓ Un cubo de más de tres dimensiones se define como hipercubo.

# OLAP



Pág: 21



# TEMA 3

## Cubo OLAP

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

tienda_id	nombre_ci	producto_id	nombre_pr	precio
001	Caracas	001	Arroz	4
002	Valencia	002	Maiz	1,8

venta_id	tienda_id	producto_id	fecha	cantidad
111	002	001	2008-03-12	214
222	001	002	2008-02-15	732

### Tabla de Dimensiones

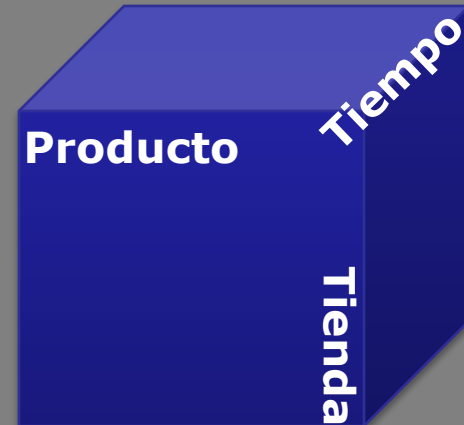
- ✓Tienda
- ✓Producto

### Tabla de Hechos

- ✓Venta

### Medidas

- ✓precio x cantidad



# OLAP



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

## Ventajas de OLAP

- ✓ Es un poderoso paradigma de visualización.
- ✓ Proporciona tiempos de respuesta rápidos.
- ✓ Permite realizar consultas de manera interactiva.
- ✓ Es bueno para analizar grupos de datos organizados por series de tiempo.
- ✓ Permite identificar semejanzas y diferencias en la data.
- ✓ Permite realizar Pivot.
- ✓ Permite realizar Slice & Dice.
- ✓ Permite realizar Drill & Down.

# OLAP



Pág: 23



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

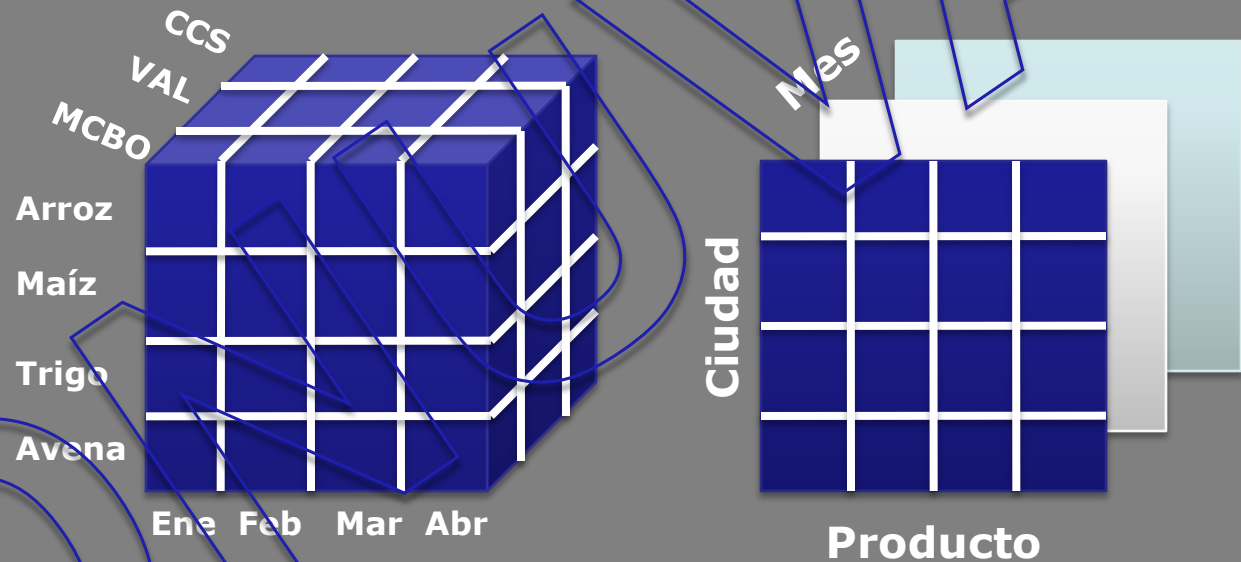
OLAP

Tableros de Comandos

Minería de Datos

## Pivot (Rotar)

Permite reordenar los ejes de la estructura multidimensional con la finalidad de obtener una mejor visión de los datos y reducir el tiempo de análisis.



# OLAP



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

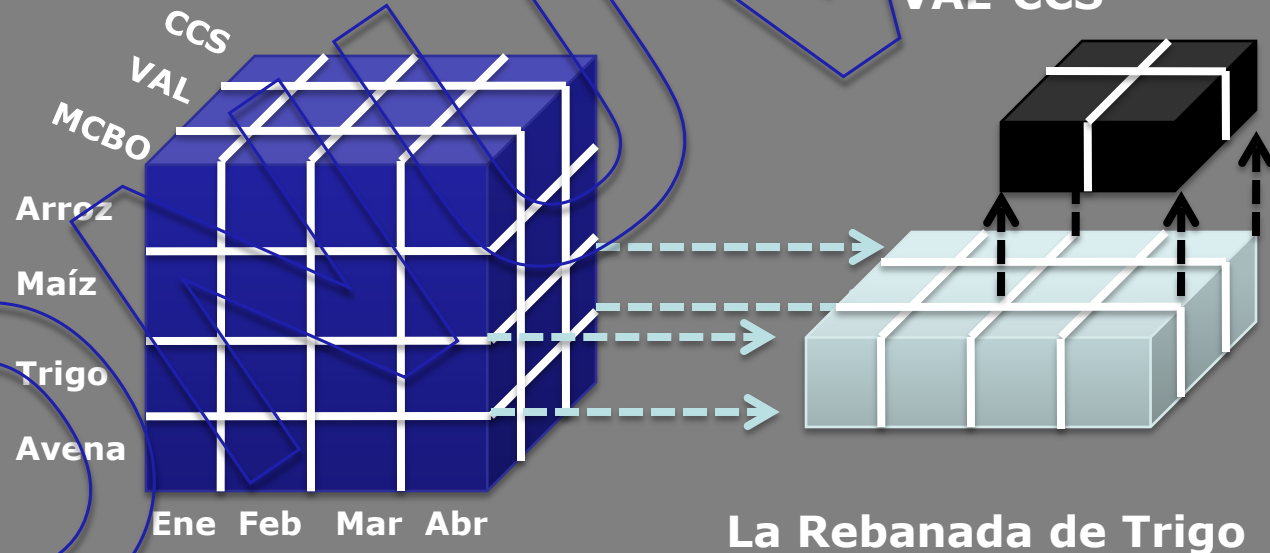
Minería de Datos

## Slice & Dice

**Rebanar (*Slice*):** Una rebanada es un subconjunto de un arreglo multidimensional que ha sido extraído seleccionando uno o mas valores de un único eje o dimensión.

**Dice (*Cortar*):** Es la acción de realizar un slice en mas de un eje.

**El Dado Trigo, Mar-Abr, VAL-CCS**



# OLAP



Pág: 25



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

Data Warehouse

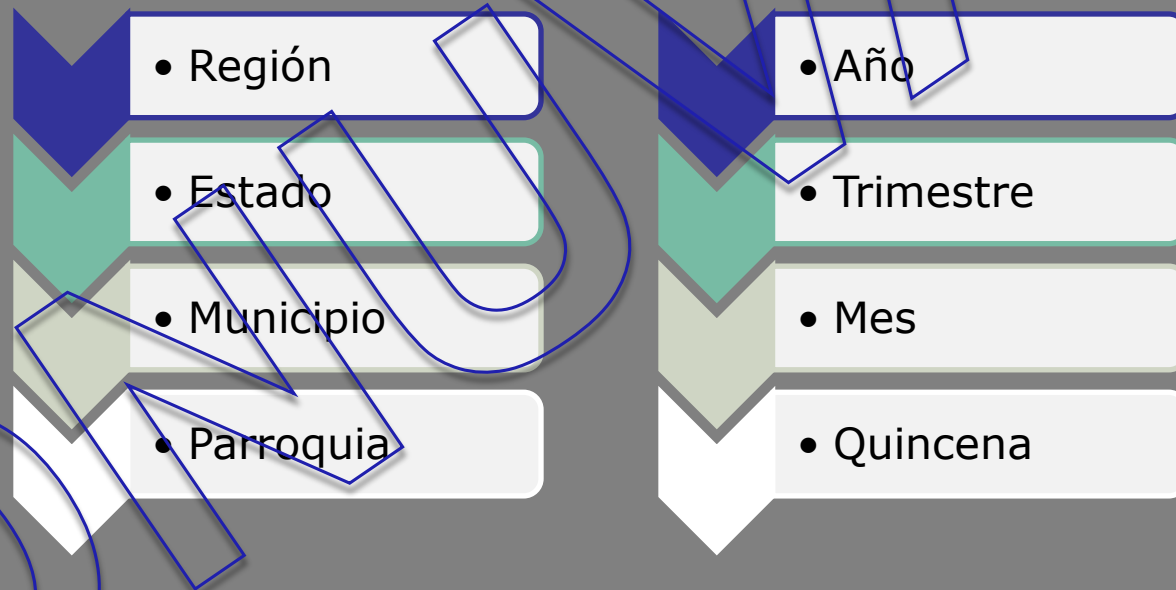
OLAP

Tableros de Comandos

Minería de Datos

## Drill & Down

**Taladrar y Excavar (*Drill & Down*):** Es una técnica analítica donde el usuario navega entre los rangos de datos des una vista resumida a una mas detallada.



# OLAP



Pág: 26



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

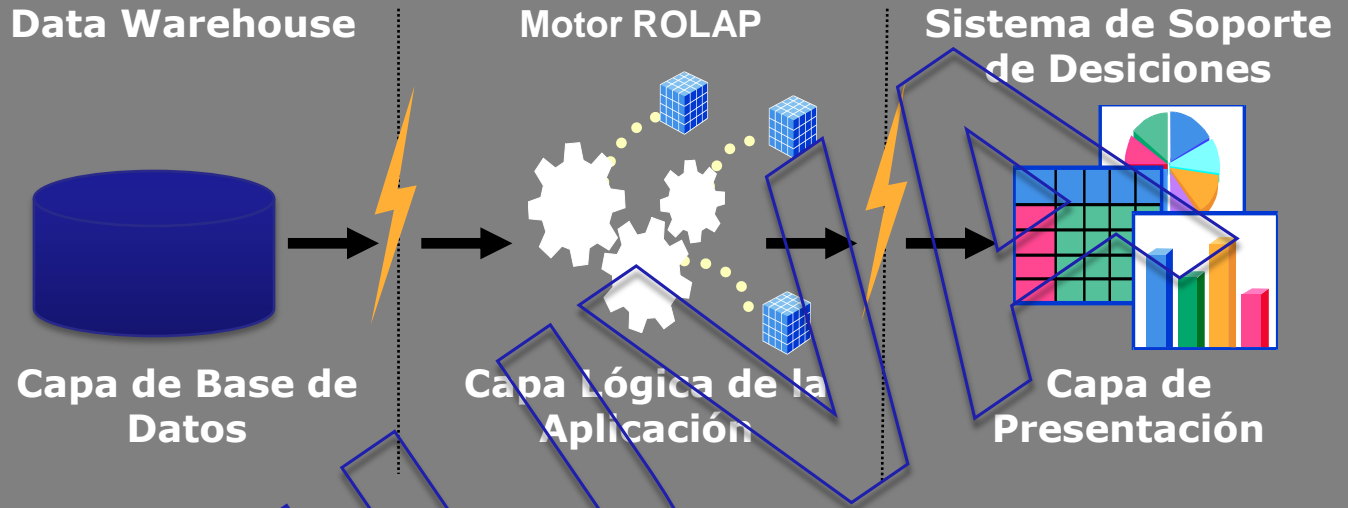
Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

# Arquitectura de un Sistema OLAP Relacional



- ✓ **Capa de Base de Datos:** Repositorio de datos y metadatos relacional.
- ✓ **Capa Lógica de la Aplicación:** Transforma las consultas multidimensionales OLAP en consultas SQL que puedan ejecutarse sobre el Data Warehouse relacional.
- ✓ **Capa de Presentación:** Interfaz de usuario que modela el resultado de la consulta en gráficos, tablas, y elementos interactivos.

# OLAP



Pág: 27



## TEMA 3

**Necesidad**

**Arquitectura**

**Infraestruct.**

- No-Técnica
- Técnica

**Data Warehouse**

**OLAP**

**Tableros de Comandos**

**Minería de Datos**

# Tableros de Comandos

Un tablero de comando es una interfaz de usuario desarrollada principalmente utilizando tecnologías web donde se refleja información en tiempo real y resumida de varios aspectos (indicadores) de un negocio.

Los tableros de comandos gerenciales están compuestos principalmente por Indicadores de Rendimiento Clave (Key Performance Indicators, KPI's en inglés) y permiten monitorear y analizar el desempeño del negocio.

## Tableros de Comando



## TEMA 3

# Diseño de Tableros Comandos

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

Característica	Recompensa
Intuitivo	• Mayor claridad
Personalizable	• Reduce la carga de trabajo innecesaria del departamento de IT.
Interactivo y Accionable	• Localizar origen de la causa. • Tiempo de Drill & Down reducido. • Análisis en Serie. • Análisis y alertas de tendencias. • Análisis predictivo.
Parte de una estrategia de inteligencia de negocios	• Una sola versión de la verdad. • Eficiencia y responsabilidad.

## Tableros de Comando



Pág: 29



### TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica

- Técnica

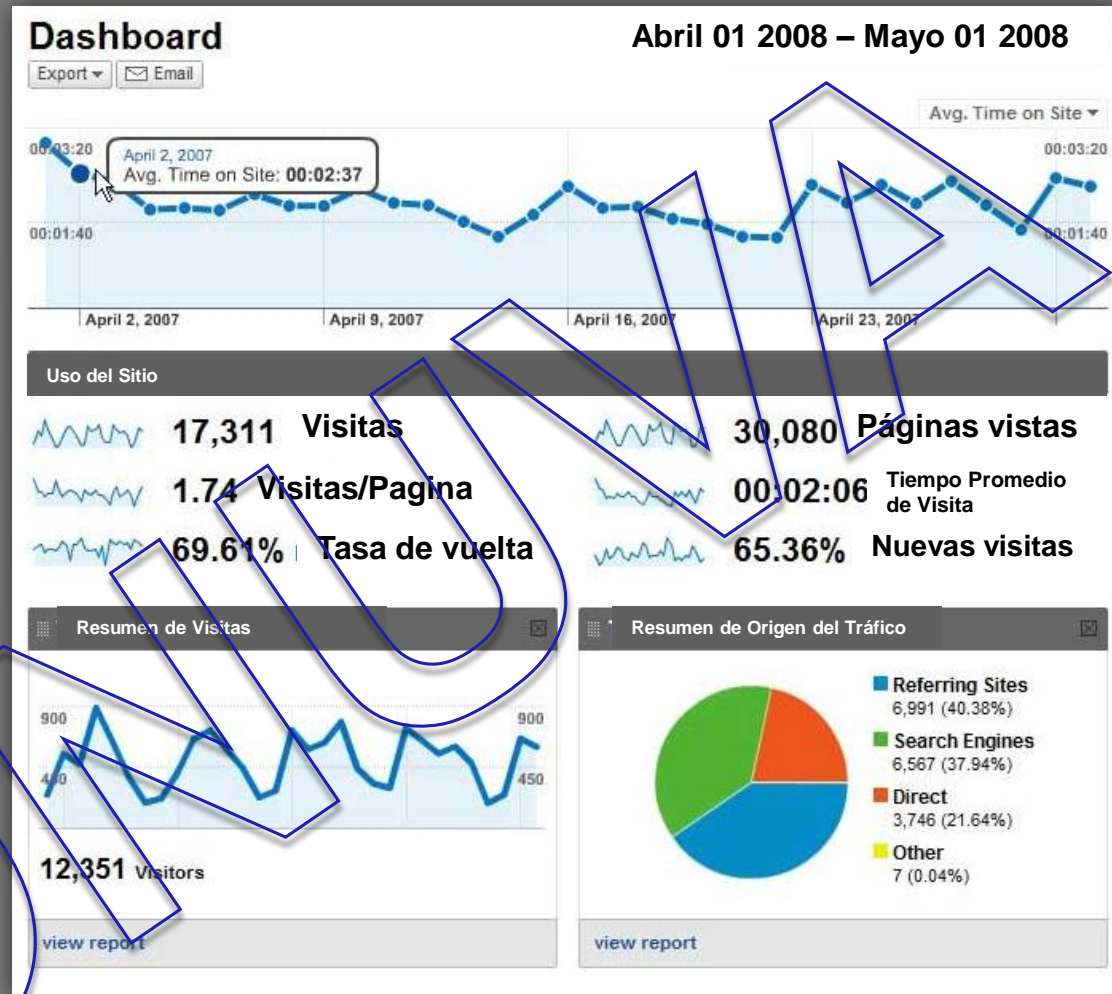
Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

## Tablero de Comando del Tráfico de un Sitio Web



## Tableros de Comando



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

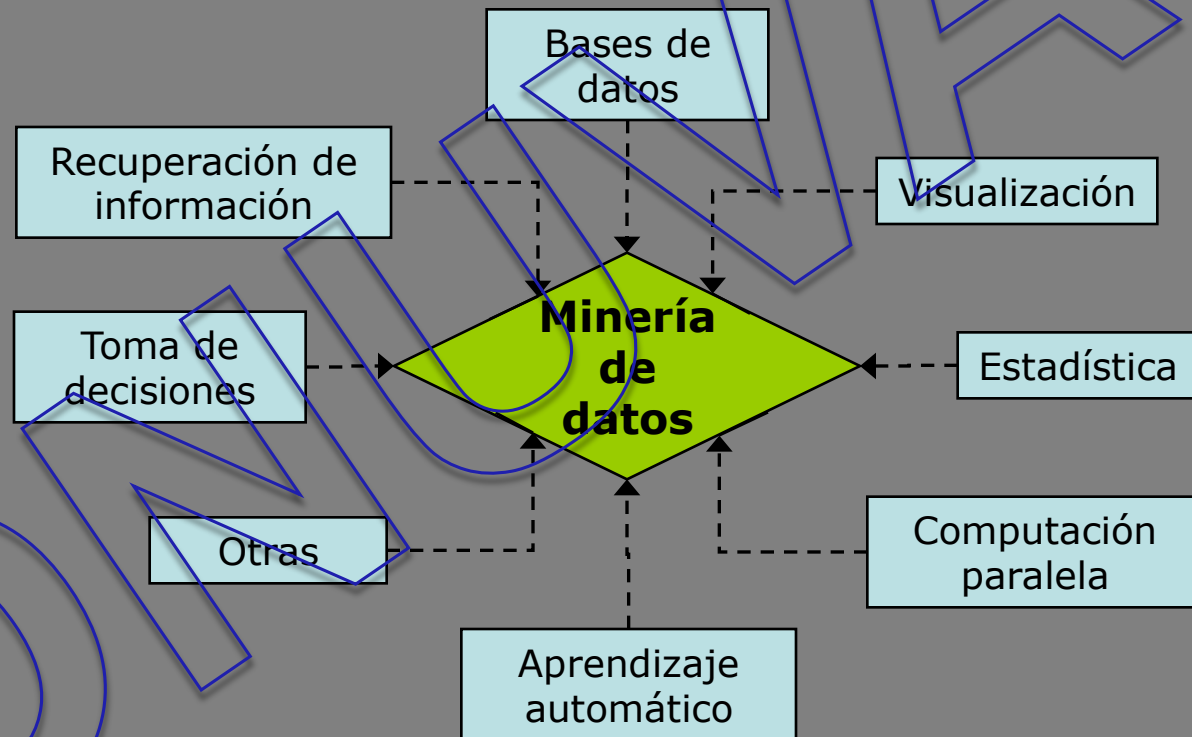
OLAP

Tableros de Comandos

Minería de Datos

# Minería de Datos

Integración de numerosas técnicas de análisis de datos y extracción de modelos. Surge por la necesidad de **explotar el potencial de los datos (volumen y variedad) almacenados en los sistemas de información.**



# Minería de Datos



Pág: 31



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

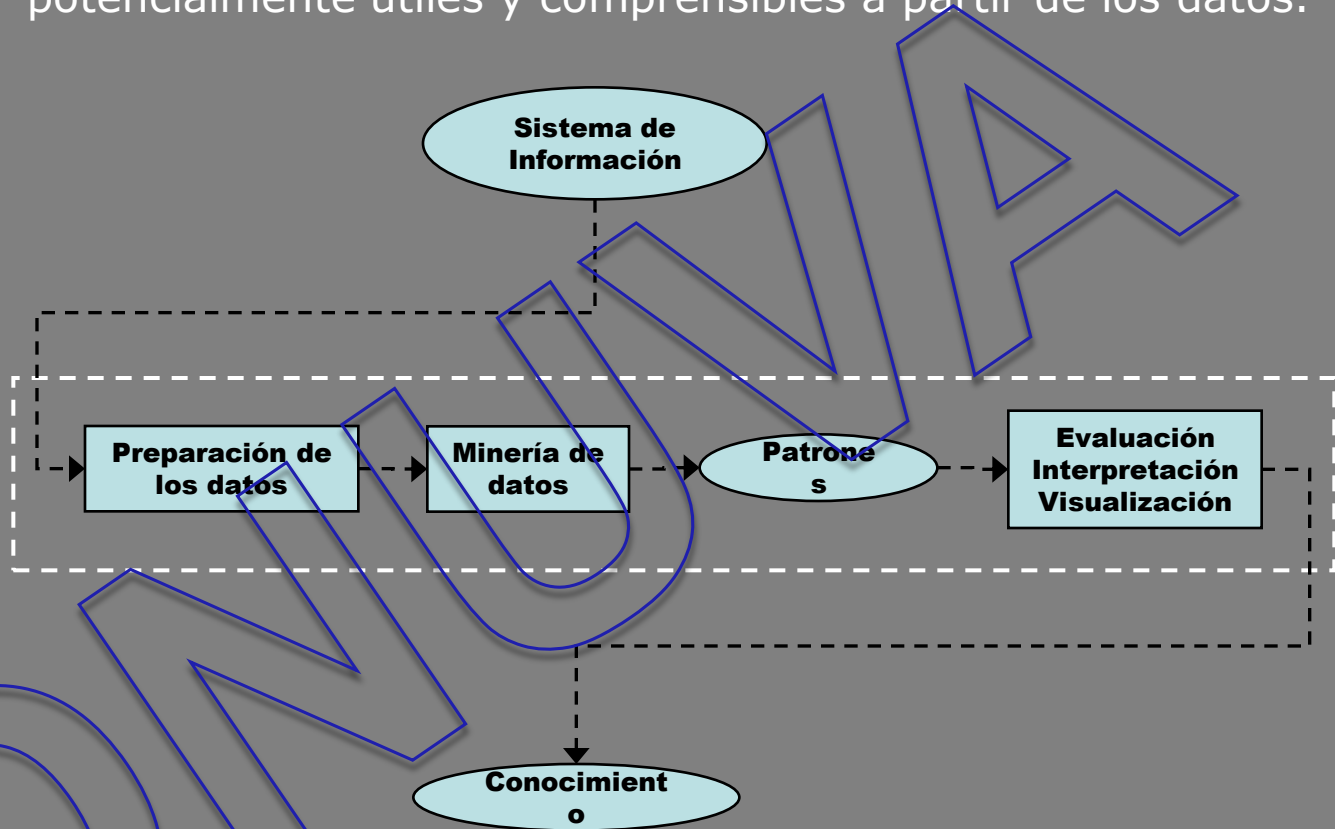
OLAP

Tableros de Comandos

Minería de Datos

# Descubrimiento de Conocimiento

Proceso de identificar patrones válidos, novedosos, potencialmente útiles y comprensibles a partir de los datos.



# Minería de Datos



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos

# Minería de datos

## Tareas

Problemas a resolver

**Predictivas:** predecir uno o más valores para uno o más ejemplos.

- Clasificación.
- Categorización.
- Priorización.
- Regresión.

**Descriptivas:** describir los datos de los ejemplos.

- Agrupamiento.
- Correlación.
- Reglas de asociación.
- Detección.

## Métodos

Cómo resolver los problemas

Incluye las técnicas siguientes:

- Algebraicas y estadísticas.
- Bayesianas.
- Conteos de frecuencia.
- Árboles de decisión.
- Aprendizaje de reglas.
- Redes neuronales.
- Procesos estocásticos.

# Minería de Datos



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-Técnica
- Técnica

Data Warehouse

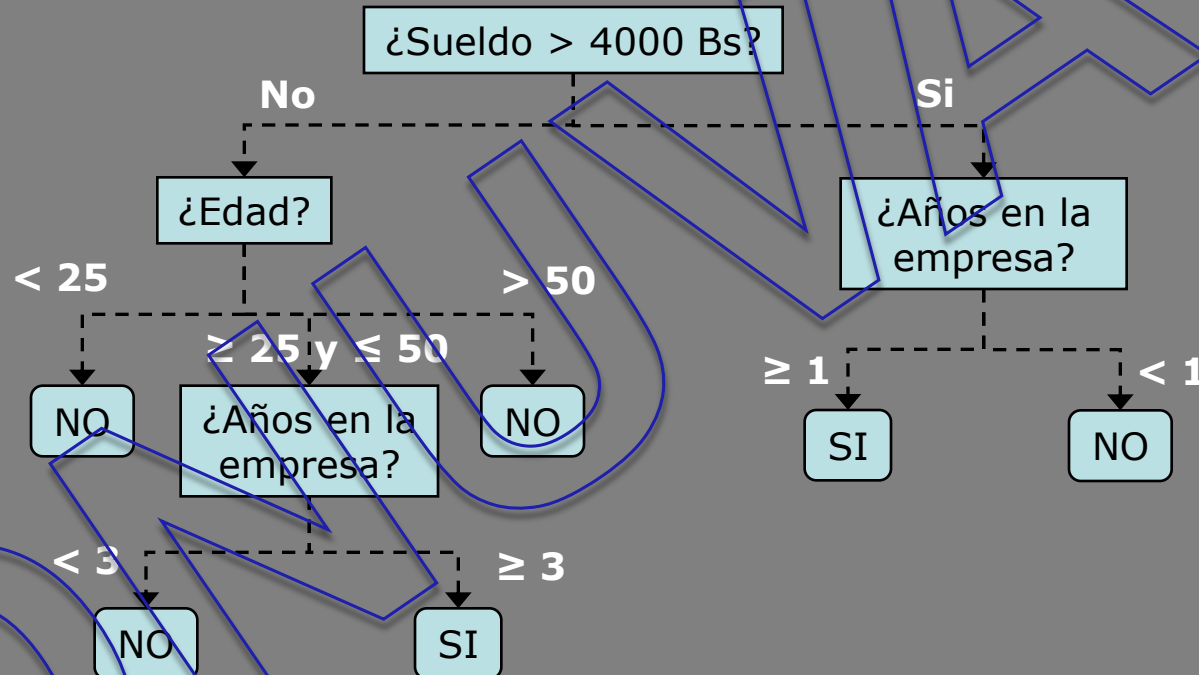
OLAP

Tableros de Comandos

Minería de Datos

# Árboles de Decisión

Conjunto de condiciones organizadas en un árbol, de tal manera que la decisión final se puede determinar siguiendo las condiciones que se cumplen desde la raíz hasta alguna de las hojas.



SI (NO) se concede el crédito

# Minería de Datos



Pág: 34



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-  
Técnica

- Técnica

Data  
Warehouse

OLAP

Tableros de  
Comandos

Minería de  
Datos

## Reglas de Asociación

Expresan las combinaciones de los valores de los atributos (nominales) que suceden más frecuentemente. Muy ligadas al mercadeo directo.

Cliente	Items comprados
1	Jugo de naranja, refresco
2	Leche, jugo de naranja, limpiador de ventanas,
3	Jugo de naranja, detergente
4	Jugo de naranja, detergente, refresco
5	Limpiador de ventanas, refresco

	Jugo naranja	Limpiador de ventanas	Leche	Refrescos	Detergente
Jugo de naranja	4	1	1	2	2
Limpiador de ventanas	1	2	1	1	0
Leche	1	1	1	0	0
refrescos	2	1	0	3	1
Detergents	2	0	0	1	2

# Minería de Datos



Pág: 35



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

- No-  
Técnica

- Técnica

Data  
Warehouse

OLAP

Tableros de  
Comandos

Minería de  
Datos

### Patrones detectados:

- ✓ Jugo de naranja y refresco son los productos más frecuentes que se compran juntos.
- ✓ El detergente nunca se compra con leche o limpiador de ventanas.
- ✓ La leche nunca se compra con refresco o detergente.

Reglas de asociación:

- ✓ **SI** jugo de naranja **ENTONCES** refresco;  
Confianza = 2 de 4 compras de jugo de naranja incluyen refresco (50%)
- ✓ **SI** refresco **ENTONCES** jugo de naranja;  
Confianza = 2 de 3 compras de refresco incluyen jugo de naranja (67%)

# Minería de Datos



Pág: 36



## TEMA 3

Necesidad

Arquitectura

Infraestruct.

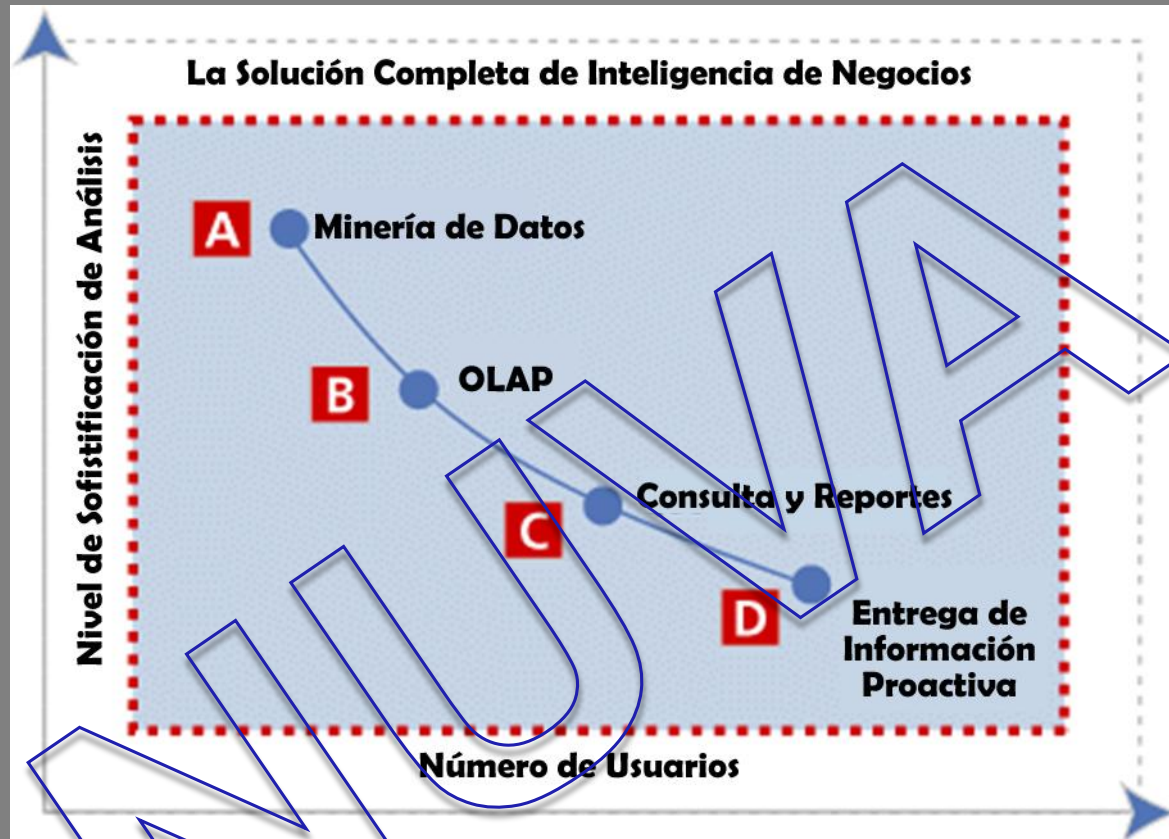
- No-Técnica
- Técnica

Data Warehouse

OLAP

Tableros de Comandos

Minería de Datos



**Entrega proactiva de información** (*Proactive Information Delivery*). Recepción de información, programada o por eventos, a través de canales como web, dispositivos inalámbricos o dispositivos de voz.

# Minería de Datos

