

UNIVERSIDAD NACIONAL DE LA MATANZA

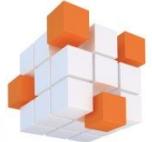
INTELIGENCIA DE NEGOCIOS

Introducción a la Inteligencia de Negocios

Conceptos Fundamentales

Docente: ING. LORENA R. MATTEO

Autores ppt orig.: Lic. HUGO M. CASTRO / MG. DIEGO BASSO



EL CASO FOODSERVICE

- FoodService es una empresa de comidas rápidas con 20 locales que se distribuyen en tres zonas:
 - Capital
 - Provincia de Buenos Aires
 - Interior
- Sus menús comprenden platos de carne y pollo, papas fritas y una variedad de ensaladas y postres.
- Por supuesto ofrecen bebidas frías (gaseosas y agua mineral) y también bebidas calientes (té, café) y pastelería, principalmente a la mañana.



EL CASO FOODSERVICE

- José, el Presidente, dice:
Tenemos que aumentar las ventas



- Mario, Director de Marketing, se pregunta:
¿Qué podemos ofrecer a nuestros clientes?





EL CASO FOODSERVICE

- Para responder a esa pregunta, Mario necesita saber:
 - ¿Qué productos se venden más?
 - ¿Qué sucursales venden mejor?
 - ¿En qué horas hay más clientes?
 - ¿Qué días de la semana son más flojos en ventas?
 - ¿Cómo evolucionaron las ventas respecto del año anterior?





EL CASO FOODSERVICE

- ¿Quién puede darle esa información?
 - Mario sabe que el área de Sistemas procesa los tickets de las sucursales en los procesos de stock de mercadería y contabilidad.
 - Por eso se dirige a Carlos, Responsable de Sistemas.



EL CASO FOODSERVICE



- Carlos recibe el pedido y estima que tendrá lista la información en un plazo de veinte días a un mes.



- Mario pregunta:

¿Cómo un mes? ¿Acaso la información no está dentro de las bases de datos de la empresa?



- Carlos responde:

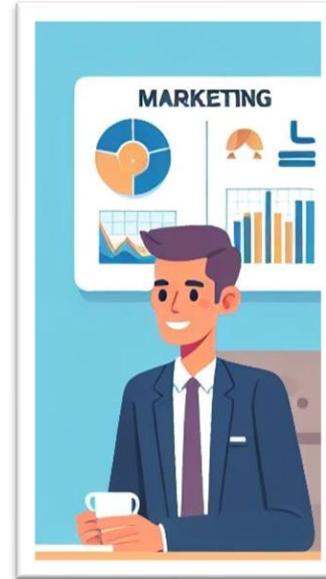
Sí, los datos están, pero no tienen la estructura adecuada para contestar esas preguntas

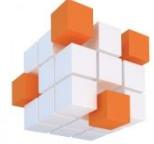


EL CASO FOODSERVICE



- Mario queda convencido de que, si los datos están guardados, tan difícil no debe ser obtener las respuestas que busca.
- Carlos queda convencido de que Mario siempre pide cosas diferentes y todo lo quiere para ayer.

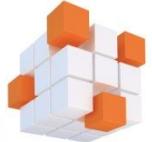




EL CASO FOODSERVICE

- Este es el momento en que recurren a nosotros para que los ayudemos.





EL CASO FOODSERVICE

Analicemos la situación: *¿Dónde está el problema?*

- Mario tiene razón en que los datos están en las bases de datos.
- Y Carlos tiene razón en que no es fácil darles la forma que Mario necesita.

Ambos hacen uso de datos, pero...

- Mario necesita analizar la información para saber qué ocurre y tomar decisiones estratégicas.
- Carlos necesita llevar a cabo los procesos que requiere la operatoria diaria de la empresa.



EL CASO FOODSERVICE

Hay una diferencia fundamental...

- A Carlos le basta con manejar *datos*.
- Mario necesita extraer *información* de esos datos.





INTELIGENCIA DE NEGOCIOS. INTRODUCCIÓN

- La información es el activo más valioso.
 - **Problema:** Cada área funcional de la empresa maneja sus propias fuentes de datos.
- Necesidad de integrar todos los sistemas de información de una organización.
 - Procesar la información eficientemente.
 - Crear inteligencia empresarial que sirva para todas las actividades que se realizan.



INTELIGENCIA DE NEGOCIOS. INTRODUCCIÓN

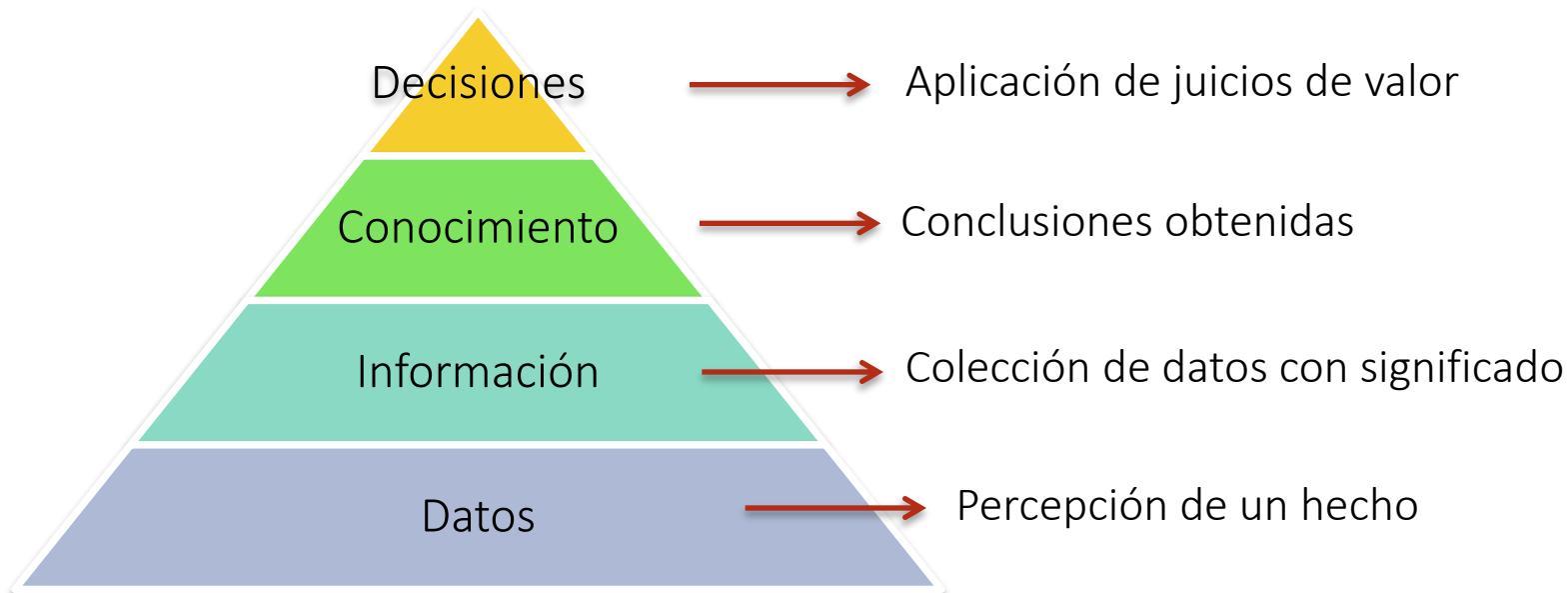
- Se necesita entender no solo QUÉ está pasando, sino CUÁNDO, DÓNDE, QUIÉN y POR QUÉ.
- Solución a los requerimientos de información con OPORTUNIDAD.
- Escalar y compartir la información a todos los tipos de usuarios en la organización.

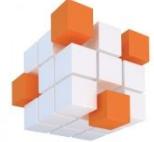
*Se necesita información y conocimiento
a partir de los datos de la empresa.*



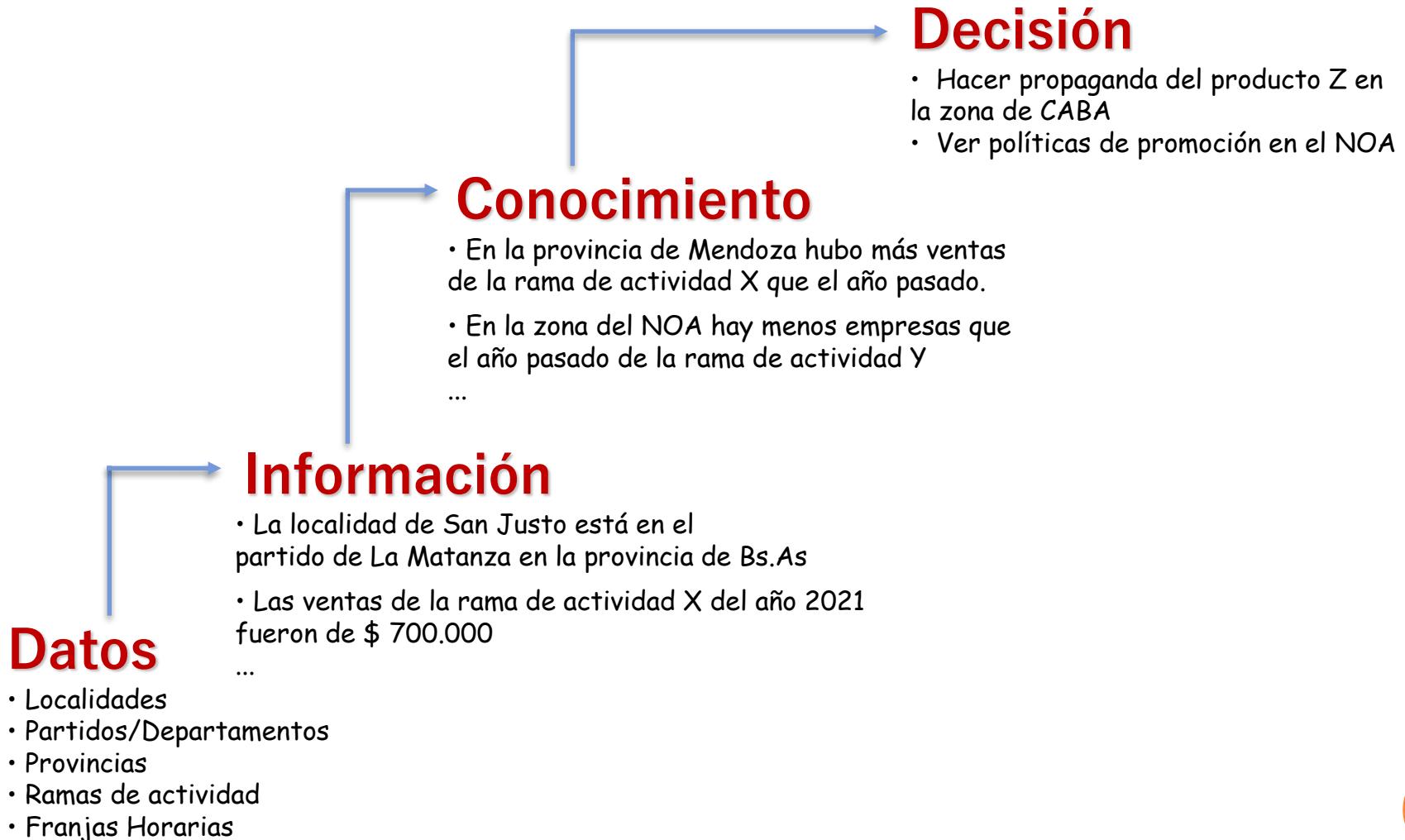
SISTEMA DE INFORMACIÓN

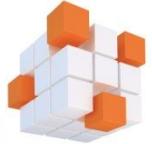
- Conjunto de subsistemas relacionados entre sí, encargados de recolectar, almacenar, procesar y distribuir información, dándole tipo y forma para las operaciones y toma de decisiones de una organización.





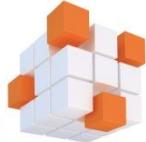
PIRÁMIDE DE INTELIGENCIA DE NEGOCIOS





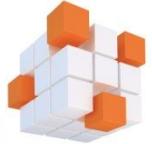
DEFINICIÓN DE BI

- Son las iniciales de Business Intelligence.
 - Inteligencia de Negocios
 - Inteligencia Empresarial
- *Es el conjunto de procesos y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de los datos existentes, para ayudar a la toma de decisiones en una organización o empresa.*
- *Procesos, tecnologías y herramientas necesarias para transformar datos en información, información en conocimiento, y conocimiento en planes que nos lleven a tomar una acción de negocio rentable. [TDWI*]*



¿CUÁL ES EL PROBLEMA PARA BI?

- Los datos de que dispone la empresa:
 - Son reunidos con sistemas desconectados entre sí.
 - Están en diversas plataformas.
 - Están con distintas codificaciones.
 - Datos con igual nombre significan conceptos distintos.
 - Datos con distinto nombre son el mismo concepto.



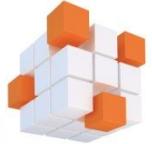
¿QUÉ HACEN LOS USUARIOS?

- Las distintas áreas de negocios encaran sistemas propios sin relación con el resto.
- Los usuarios manejan copias de la información en planillas y muchas veces las corrigen y/o completan.
- *Consecuencia:* Nunca se sabe cuál es la versión verdadera de los datos.



ESTO CAUSA QUE...

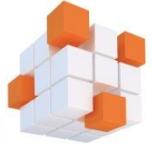
- Cada vez que hay que hacer un trabajo que utilice datos de distintos sistemas hay que:
 - Localizar a el/los dueños de esos datos.
 - Pedir copia de esos datos.
 - Ver las conversiones a efectuar.
 - Ver la correspondencia entre los nombres utilizados.



CONCLUSIÓN

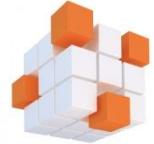
- NO es fácil encontrar los datos que los usuarios necesitan y prepararlos para responder a los pedidos de información.
- NO es fácil que los usuarios los utilicen.

ESTO OCASIONA DEMORAS INACEPTABLES



ADEMÁS ...

- Los usuarios de negocios ven el problema en forma mucho más simple y más abstracta.
- No ven la complejidad de localizar los datos, cargarlos, transformarlos y darles la forma que necesita el usuario.
- Se producen roces entre la gente de negocios y la de informática.



Los DESAFÍOS DE BI

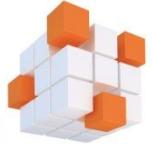
- Tener una única verdad con datos confiables.



- Lograr una visión integral del negocio.



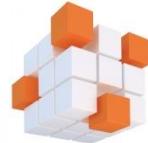
- Poder acceder a la información que necesito en el momento adecuado, sin requerir permanentemente de la intervención del área de sistemas.



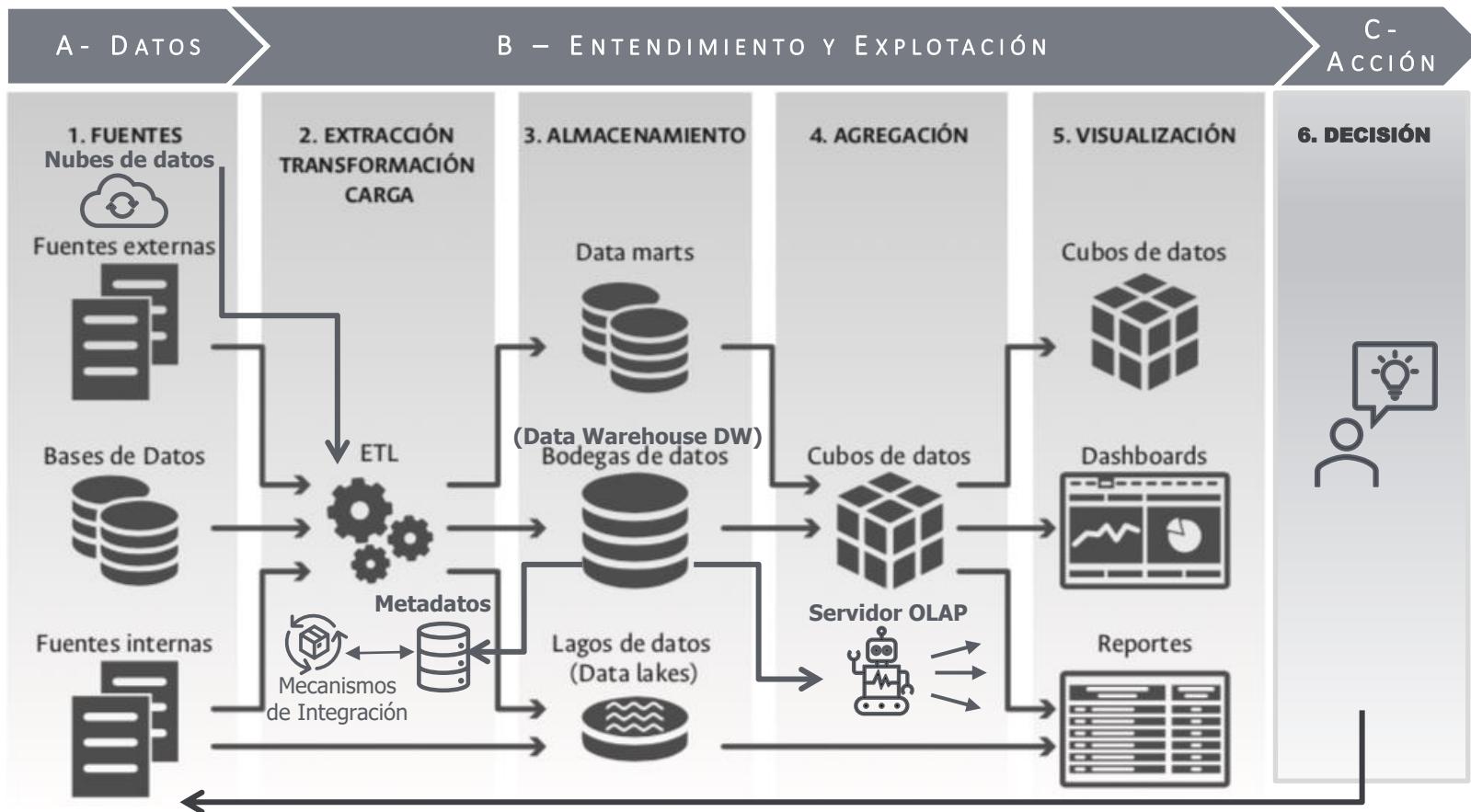
NUEVA ESTRUCTURA DE BASE DE DATOS

- Si los datos están en las bases de datos que tiene la empresa, pero la forma no es la adecuada, la solución es: *crear una base de datos con la forma adecuada.*
- La base de datos con la estructura adecuada es lo que se denomina **DATA WAREHOUSE**.
- Es una base de datos separada de los sistemas transaccionales e independiente de ellos.





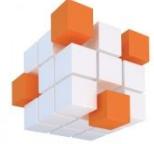
ARQUITECTURA BASE DE BI



A - Datos: Centralizar información de múltiples fuentes en un DW.

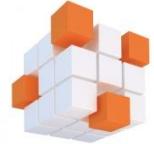
B - Entendimiento y Explotación: Herramientas de BI y DM para analizar y mejorar el entendimiento del negocio.

C - Acción: Actuar sobre los hallazgos realizados en el análisis.



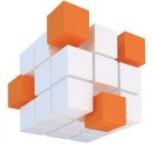
FUENTES DE DATOS

- **OLTP:** sistemas transaccionales utilizados para la operatoria del negocio (ERP, CRM, etc.).
- **Fuentes Internas o Externas:** archivos como salida de otros sistemas, o archivos creados a mano, logs, CSV / TXT/ JSON, XML, PDF, APIs, etc.
- **Otros DW:** porque se encuentra en otro lugar, o porque aprovecho la información ya limpia, BD, DM o cubos.
- **Web 2.0:** datos no estructurados, de e-commerce, de redes sociales, de CMS (sistemas de gestión de contenidos), blogs, wikis, servicios de creación multimedia, Data Clouds, Data Lakes, etc.



+ FUENTES DE DATOS – TENDENCIAS ACTUALES

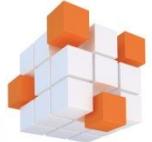
- Mayor uso de **APIs en tiempo real** para conectar fuentes OLTP y **Web 2.0** directamente al BI. **Web Services, Big Data, NLP.**
- Auge de **plataformas en la nube** (Snowflake, Databricks) como DW y almacenamiento de **datos no estructurados**: streaming data, imágenes, audio y video.
- Incremento en datos de **IoT** y **Datos en Tiempo Real** (ej. sensores en retail, salud, fábricas, etc.) integrados en BI. **Open Data** (información pública: censos, clima, salud), **Blockchain** (auditorías, contratos inteligentes, registros descentralizados), **Datos de IA y ML**, etc.



DATA WAREHOUSE

Un Data Warehouse es un repositorio de datos con una estructura **orientada al negocio, integrada, no volátil y variable en el tiempo**, organizada de forma tal que facilita el análisis de grandes volúmenes de datos para la toma de decisiones. [Bill Inmon]





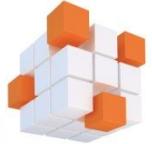
DATA WAREHOUSE

Un DW contiene:

- Información histórica
 - Para visualizar tendencias y efectuar comparaciones
- Información consolidada
 - Para acelerar la respuesta a las consultas

Las bases de datos más voluminosas son Data Warehouses





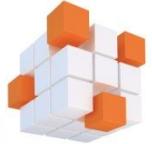
ORIENTADA AL NEGOCIO

- Organiza y presenta los datos desde la perspectiva de los conceptos que maneja la empresa (*fecha, cliente, producto, sucursal, ventas*).
- Los datos tienen el nivel de detalle y la estructura que necesitan los que toman decisiones.



INTEGRADA

- Se construye a partir de fuentes de datos heterogéneas.
 - Bases de datos relacionales, archivos planos, hojas de cálculo, documentos impresos, etc.
- Se unifican denominaciones, codificaciones, formatos.
 - Limpieza de información
 - Integración de datos



No VOLÁTIL

- En el DW los datos **no se modifican**, es decir no hay aplicaciones de usuarios que modifiquen datos.
- El DW se renueva.
- Los datos permanecen intactos entre renovaciones.
- Sólo existen dos operaciones:
 - Carga (inserts)
 - Acceso (select)



VARIABLE EN EL TIEMPO

- El horizonte temporal del DW es más amplio que el de los sistemas transaccionales.
 - La vida útil de los datos es mayor
 - Maneja datos históricos
- La fecha es un dato fundamental.
 - Posibilidad de analizar en el tiempo
 - Integración de datos



METADATOS

- Son datos que describen objetos del DW
 - Estructura del DW
 - Esquema, visiones, dimensiones, datos derivados, ubicación y contenido de los data mart
 - Datos sobre los datos
 - Origen de los datos
 - Validez de los datos (activo, histórico)
 - Información de control (estadísticas de uso, errores, información de auditoría)
 - Algoritmos que se usan para la consolidación
 - Correspondencia entre datos operativos y los del DW
 - Datos de Negocios
 - Definiciones de términos del negocio, dueños de los datos



ALTERNATIVAS

○ Data Warehouse (Almacén de Datos)

- Es una base de datos centralizada optimizada para consultas y análisis de toda la empresa, diseñada para integrar datos de múltiples fuentes en un esquema estructurado.

○ Data Mart

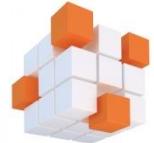
- Sirve a un grupo específico de usuarios.
- Su alcance se reduce a un área en particular:
 - Marketing, Ventas, Compras, etc.

○ Data Warehouse virtual

- Se forma a partir de distintos Data Marts.

○ Data Lake

- Permite almacenar de manera centralizada datos estructurados y no estructurados en bruto a cualquier escala, sin necesidad de un esquema predefinido.



ALTERNATIVAS (EVOLUCIÓN ARQUITECTURA BI)

○ Data Lakehouse

- Combinación de Data Warehouse y Data Lake.
- Usa almacenamiento escalable, datos estructurados y semi-estructurados en un mismo entorno.
 - Ejemplo: *Empresa de e-commerce que almacena logs de usuario en bruto y datos de ventas procesados en un mismo repositorio.*

○ Data Vault (Bóveda de Datos)

- Modelo diseñado para trazabilidad, auditoría y flexibilidad.
- Separación de los datos en Hubs (entidades principales), Links (relaciones) y Satellites (atributos y cambios en el tiempo).

○ Data Mesh

- Modelo distribuido y descentralizado.
- Cada dominio de negocio maneja sus propios datos como un producto y bajo estándares compartidos.



ALTERNATIVAS (EVOLUCIÓN ARQUITECTURA BI)

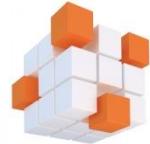
○ Data Fabric

- Es un enfoque unificado que automatiza el movimiento y transformación de datos en entornos híbridos y multi-cloud.
- Se sustenta en una arquitectura unificada con servicios (o tecnologías) que corren por encima de ella y que ayudan a las empresas en la tarea de la gestión de datos.
- Se busca maximizar el valor de los datos y acelerar la transformación digital, así como también simplificar e integrar la gestión de datos en entornos en la nube.



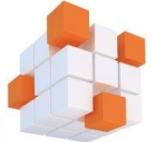
PROCESO DE EXPLOTACIÓN DE DATOS

- Guardar y estructurar los datos en un DW es sólo parte de la tarea.
- Necesitamos analizar los datos para la toma de decisiones.
- Independizar al usuario de negocios del área informática mediante las herramientas adecuadas.
- Herramientas de redacción de informes orientada a las necesidades de los distintos usuarios.
- Empezamos por producir los mismos informes que el personal de negocios estaba recibiendo hasta ahora.
 - Pero ahora él mismo puede armarlos en forma interactiva.



REPORTES - INFORMES

- Para obtener los informes *no es necesario escribir ningún programa.*
- Lo puede definir el mismo profesional de negocios.
- La herramienta que produce los informes está preparada para que él la utilice.
- El profesional de negocios decide:
 - Criterios de selección a usar sobre la base de datos
 - Datos que se van a incluir en los informes
 - Cómo se van a ordenar los datos
 - Cómo se van a agrupar los datos
- Con estas herramientas queda resuelto el problema de la producción de informes



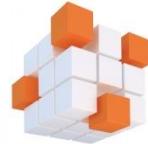
REPORTES - INFORMES

- A partir de los informes obtenidos, Mario y su gente encuentran que:

Las respuestas generan nuevas preguntas...

- ¿Cómo se vende en cada sucursal por franja horaria?
- ¿Qué productos se venden más por la tarde?
- ¿Venden más las sucursales que tienen servicio en el auto?
- ¿Y las que tienen estacionamiento?





Los 5 ESTILOS DE BI

- Estadios de madurez o fases de implantación del análisis de negocios en una organización. Es ideal tener una plataforma de BI que pueda integrar los 5 estilos.





Los 5 ESTILOS DE BI

○ Alertas y Notificaciones Proactivas

Radar Personalizado de Información

Revisar los datos
En la Base de Datos

Las Alertas se disparan
Basadas en Agendas, Eventos,
Reglas de Negocio

Reportes y Alertas son enviados
Al dispositivo preferido
De cada usuario

Los usuarios pueden conocer más
Detalle abriendo los adjuntos o
Clickeando los Web links





Los 5 ESTILOS DE BI

- Tableros de Comando y Scorecards: Brindan información instantánea y general sobre la performance del negocio.



Tableros

- Consumo facil de información con tacómetros
- Vista integrada de los datos de toda la empresa

Scorecards

- Usando Metodologías formales
- Drill Down para ver Performance Scorecards de Grupos/Empleados

Department Operational Metrics For Week Ending 11/21/03 11/24/2003 8:57 AM						
Status	Trend	Metric	Actual	Target	% Difference	
●	▲	Revenue	\$750,456	\$875,000	-14%	
■	▲	Operating Expense	\$286,893	\$200,000	+43%	
○	—	Gross Profit Margin	41%	85%	-37%	
●	▲	# Unique Prospective Buyer Visits	227	250	-9%	
●	▲	Lead-To-Order Conversion Rate	15%	25%	-40%	
○	—	% Sales To Total Region Sales	30%	40%	-25%	
■	▲	# Open Sales & Service Positions	5	3	+67%	
■	▲	Average Employee Turnover	34%	25%	+36%	
●	▲	Average Employee Performance Score	3	5	-40%	
●	▲	Sales Of Existing Models	\$385,794	\$400,000	-4%	
●	▲	Sales Of New Models	\$228,938	\$350,000	-34%	
●	—	Sales Of Pre-Owned Models	\$143,727	\$80,000	+80%	
○	▲	# New Customers	47	50	-6%	
○	—	% Customers Returning For Service Within 6 Months	72%	90%	-20%	
●	▲	Average Customer Satisfaction Score	5	5	-20%	
●	▲	Regional Ad Campaign Expense	\$760,000	\$750,000	0%	
●	—	Quality Level Associated With Brand	5	5	-20%	
○	▲	% Closed Deals With Incentives Applied	81%	50%	+62%	

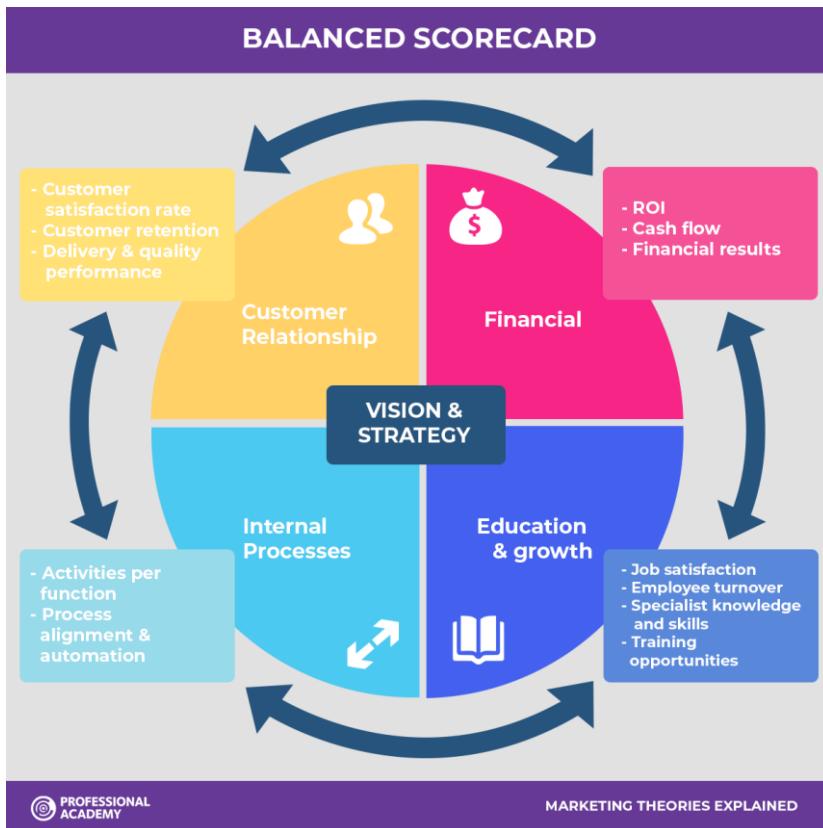
Managed Metrics

- Permite a todos monitorear las métricas que a cada uno le interesa
- Indicación Inmediata de cuando se requiere accionar

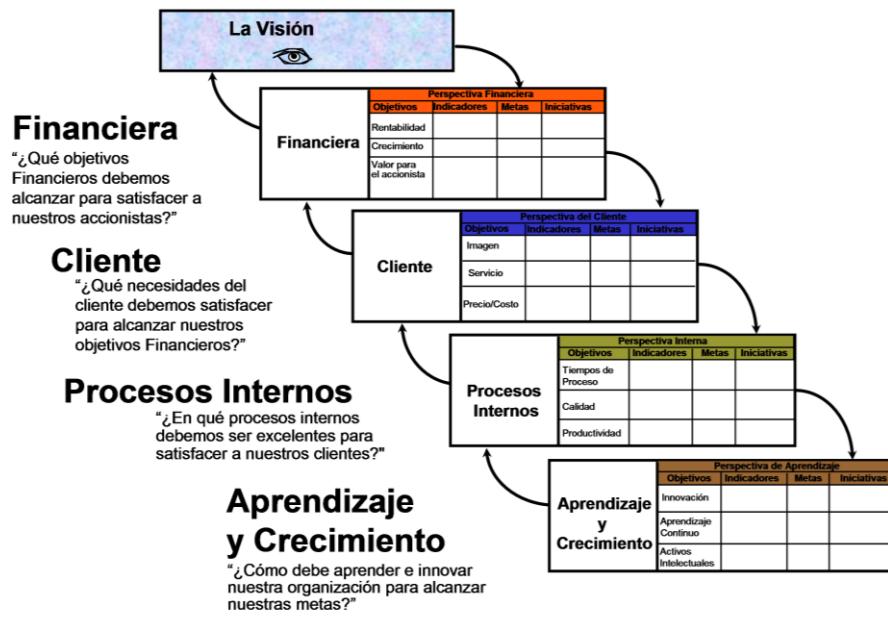


Los 5 ESTILOS DE BI

- Balance Scorecards (BSC): un cuadro de mando integral (CMI), concepto creado por Kaplan y Norton en 1992, es tipo de tablero que proporciona una visión completa del rendimiento general de una compañía, incluyendo la actividad de todos los departamentos y unidades de negocio.



Cuatro Perspectivas de Negocio



Fuentes: <https://balancedscorecard.org/> / <https://blog.bismart.com/tipos-de-cuadros-de-mando/> / <https://gestion.pensemso.com/cuadro-de-mando-integral-ejemplo-definitivo-6-plantillas/> / <https://bscdesigner.com/es/bsc-plantillas-y-ejemplos.htm> / <https://margaritaberdugo.wordpress.com/2015/10/29/el-balanced-scorecard/>



Los 5 ESTILOS DE BI

- **Enterprise Reporting:** Permite que el BI llegue al público poniendo a su disposición información con alto nivel de detalle, lo que impacta a los encargados de la toma de decisiones de las organizaciones.

Category Sales and Profit Performance by Region								
Category: Electronics			Subcategory: Audio Equipment					
Region: Northeast			Total Audio Equipment Sales = \$408,120					
Qtr '03	\$6,420	26.0%	\$16,048	\$4,011	\$4,011	\$4,011	\$4,011	\$4,011
Qtr '03	\$12,598	29.0%	\$31,563	\$6,385	\$6,385	\$6,385	\$6,385	\$6,385
Half '03	\$17,598	54.94%	\$46,543	\$8,126	\$2,435	\$4,253	\$1,188	\$3,438
Half '03	\$12,937	26.41%	\$31,094	\$6,056	\$6,056	\$3,142	\$5,071	\$3,142
Dec '03	\$14,438	85.784	\$37,088	\$4,843	\$4,843	\$4,843	\$4,843	\$4,843
Dec '03	\$14,438	26.0%	\$37,088	\$4,843	\$4,843	\$4,843	\$4,843	\$4,843

Category Sales and Profit Performance by Region								
Category: Electronics			Subcategory: Audio Equipment					
Region: Southeast			Total Audio Equipment Sales = \$20,400					
Qtr '03	\$14,438	50.78	\$4,011	\$1,269	\$6	\$6	\$6	\$6
Qtr '03	\$14,438	26.0%	\$37,088	\$4,843	\$4,843	\$4,843	\$4,843	\$4,843
Half '03	\$13,879	53.66%	\$34,448	\$3,743	\$3,743	\$3,743	\$3,743	\$3,743
Half '03	\$13,215	28.05%	\$31,048	\$2,627	\$2,627	\$2,627	\$2,627	\$2,627
Dec '03	\$16,895	84.94%	\$39,576	\$5,495	\$5,495	\$5,495	\$5,495	\$5,495
Dec '03	\$13,349	26.1%	\$31,218	\$2,627	\$2,627	\$2,627	\$2,627	\$2,627

Category Sales and Profit Performance by Region								
Category: Electronics			Subcategory: Cameras					
Region: Northeast			Total Camera Sales = \$23,770					
Qtr '03	\$17,198	84.81%	\$18,630	\$3,236	\$3,236	\$3,236	\$3,236	\$3,236
Qtr '03	\$12,076	26.4%	\$14,164	\$2,486	\$2,486	\$2,486	\$2,486	\$2,486
Half '03	\$26,768	89.77%	\$31,186	\$4,810	\$4,810	\$4,810	\$4,810	\$4,810
Half '03	\$24,424	27.0%	\$30,978	\$4,927	\$4,927	\$4,927	\$4,927	\$4,927
Dec '03	\$26,865	84.94%	\$33,786	\$5,770	\$5,770	\$5,770	\$5,770	\$5,770
Dec '03	\$26,865	27.0%	\$32,918	\$5,770	\$5,770	\$5,770	\$5,770	\$5,770

Category Sales and Profit Performance by Region								
Category: Electronics			Subcategory: Cameras					
Region: Southeast			Total Camera Sales = \$14,000					
Qtr '03	\$16,378	60.381	\$16,638	\$4,484	\$4,484	\$4,484	\$4,484	\$4,484
Qtr '03	\$16,378	29.54%	\$16,067	\$1,098	\$1,098	\$1,098	\$1,098	\$1,098
Half '03	\$16,408	86.388	\$20,278	\$4,981	\$4,981	\$4,981	\$4,981	\$4,981
Half '03	\$16,211	27.0%	\$19,069	\$2,045	\$2,045	\$2,045	\$2,045	\$2,045
Dec '03	\$17,298	84.94%	\$20,438	\$5,389	\$5,389	\$5,389	\$5,389	\$5,389
Dec '03	\$17,298	27.0%	\$19,211	\$5,389	\$5,389	\$5,389	\$5,389	\$5,389

- Reportes Operacionales Print-perfect
- Via Web e Impresión
 - Facil navegación a través de cientos de páginas de reportes
 - Prompts que permiten a los usuarios definir el contenido

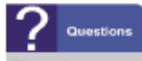
Period	Header	Detail
Qtr '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Qtr '03	Total Orders: 30,838	Total Revenue: \$2,930,080
Half '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Half '03	Total Orders: 30,838	Total Revenue: \$2,930,080
Dec '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Dec '03	Total Orders: 30,838	Total Revenue: \$2,930,080

Period	Header	Detail
Qtr '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Qtr '03	Total Orders: 30,838	Total Revenue: \$2,930,080
Half '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Half '03	Total Orders: 30,838	Total Revenue: \$2,930,080
Dec '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Dec '03	Total Orders: 30,838	Total Revenue: \$2,930,080

Period	Header	Detail
Qtr '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Qtr '03	Total Orders: 30,838	Total Revenue: \$2,930,080
Half '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Half '03	Total Orders: 30,838	Total Revenue: \$2,930,080
Dec '03	Total Orders: 40,144	Total Revenue: \$3,012,380
Dec '03	Total Orders: 30,838	Total Revenue: \$2,930,080

- Reportes de Negocios Pixel-perfect
- Creado por usuarios de negocio, no por el departamento de sistemas
 - Tablas y gráficos integrados

- Page-perfect Invoices and Statements
- On-line Billing Applications
 - Statements
 - Other Page Forms

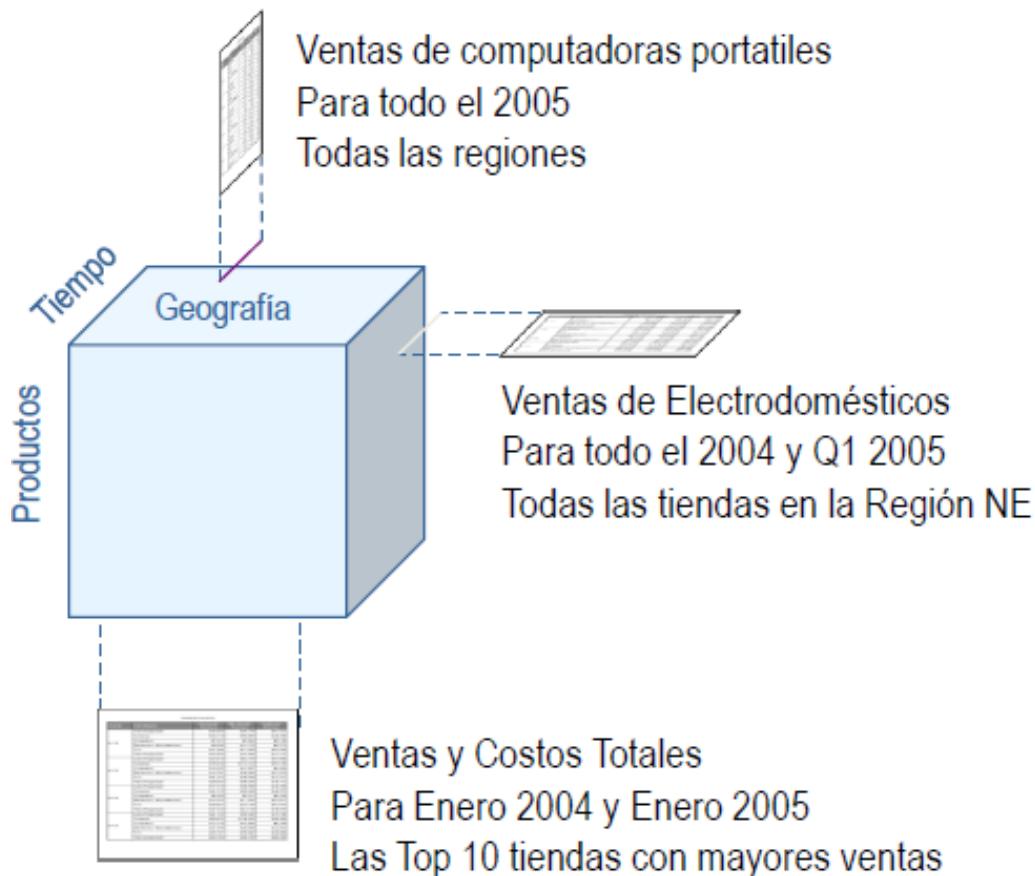


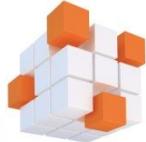


Los 5 ESTILOS DE BI

○ Análisis Multidimensional OLAP

Cortar y extraer
con un análisis OLAP





Los 5 ESTILOS DE BI

○ Análisis Predictivo y Avanzado

Análisis Predictivo Básico
Basado en Técnicas de Regresión

On-Line Sales Forecast



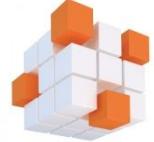
Notar ambas Predicción Lineal y Lineas de Predicción Estacional

Análisis Predictivo Avanzado
Basado en Técnicas de Data Mining

DETERMINAR
QUIEN ES
CAPAZ DE ...

Alcanzar las Ventas
Mantenerse en el presupuesto
Responder
Comprar
Defraudar
Ser Rentable
Puntual

Redes Neuronales, Clustering Algoritmos
Arbol de Algoritmos, Regresiones Multi-
Variables



EVOLUCIÓN DE LOS 5 ESTILOS DE BI

A raíz de la transformación digital y de la integración de tecnologías emergentes podemos considerar una visión actualizada de estos estilos:

- **BI en la Nube (Cloud BI):** Implementación flexible, accesible desde cualquier dispositivo, escalabilidad.
- **Mobile BI:** Acceso a reportes y dashboards desde dispositivos móviles.
- **Analítica Embebida (Embedded BI):** Integra capacidades analíticas directamente en aplicaciones empresariales o procesos operativos, permitiendo que la información llegue en el contexto de la acción.
- **BI en Tiempo Real (Real-Time o Streaming BI):** Se enfoca en el análisis y visualización de datos en tiempo real, esencial en entornos donde la rapidez en la toma de decisiones es crítica (como en e-commerce o logística).



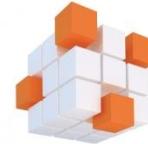
EVOLUCIÓN DE LOS 5 ESTILOS DE BI (CONT)

- **BI Auto-Servicio (Self-Service BI):** Interfaces intuitivas, análisis sin intervención de TI, dashboards personalizados.
- **Análisis Argumentativo(Augmented Analytics):** Utiliza inteligencia artificial y machine learning para automatizar el descubrimiento de insights, recomendaciones y predicciones, haciendo más accesible el análisis avanzado.
- **BI Colaborativo (Collaborative BI):** Integración de herramientas de BI con plataformas de colaboración empresarial.
- **Data Storytelling & NLP en BI:** Visualización narrativa de datos, consultas en lenguaje natural.
- **Blockchain & BI Descentralizado:** Auditoría segura, registros inmutables, contratos inteligentes.



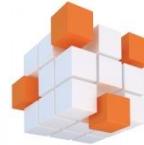
RESUMEN 5 ESTILOS DE BI

Estilo de BI	Características Clave	Público Objetivo	Beneficios Principales
Análisis Predictivos y Avanzados	Identificación de patrones, predicción, técnicas estadísticas y de ML (clasificación, regresión, clustering).	Científicos de datos, analistas estadísticos, analistas de negocio avanzados.	Descubrimiento de insights ocultos, predicción de tendencias, análisis profundo de relaciones complejas.
Análisis Multidimensional (Cubos OLAP)	OLAP, análisis multidimensional, "slice and dice", cubos predefinidos.	Gerentes, analistas de negocio.	Ánálisis rápido de datos agregados, exploración multidimensional.
Informes Empresariales (Consultas y Análisis Ad Hoc)	Auto-servicio, flexible, impulsado por el usuario, sin necesidad de conocimientos técnicos profundos.	Usuarios de negocio, analistas.	Respuestas rápidas a preguntas específicas, exploración de datos sin restricciones, reducción de la carga de TI.
Tableros de Control y Scorecards	Formato preciso, distribución amplia, programado o bajo demanda, personalización	Ejecutivos, gerentes, consumidores de información.	Visión consolidada del rendimiento, informes estandarizados, gobernanza de datos.
Entrega de Informes y Alertas	Distribución proactiva, basada en horarios o eventos, personalización, notificaciones.	Amplia gama de usuarios, tanto internos como externos.	Monitorización en tiempo real, acción oportuna ante eventos críticos, reducción de la necesidad de monitorización constante.



EVOLUCIÓN DE LOS 5 ESTILOS DE BI

Estilo de BI	Características Clave	Público Objetivo	Beneficios Principales
BI en la Nube (Cloud BI)	Implementación flexible, accesible desde cualquier dispositivo, escalabilidad.	Empresas de todos los tamaños.	Reducción de costos, mayor accesibilidad y colaboración.
Mobile BI	Acceso a reportes y dashboards desde dispositivos móviles.	Ejecutivos, equipos en campo.	Toma de decisiones desde cualquier ubicación.
Analítica Embebida (Embedded Analytics)	Integración de dashboards e informes dentro de otras aplicaciones empresariales.	Usuarios de sistemas CRM, ERP, e-commerce.	Acceso directo a datos en el flujo de trabajo, decisiones más rápidas.
BI en Tiempo Real (Real-Time BI)	Procesamiento y visualización de datos en streaming, dashboards dinámicos.	Equipos operativos, analistas.	Reacción inmediata a eventos, optimización de operaciones.



EVOLUCIÓN DE LOS 5 ESTILOS DE BI

Estilo de BI	Características Clave	Público Objetivo	Beneficios Principales
BI de Auto-Servicio (Self-Service BI)	Interfaces intuitivas, análisis sin intervención de TI, dashboards personalizados.	Usuarios de negocio, gerentes.	Democratización del acceso a datos, autonomía en el análisis.
Análisis Argumentativo (Augmented Analytics)	IA y Machine Learning en BI: Análisis predictivo, detección de anomalías, insights automatizados.	Analistas avanzados, tomadores de decisiones.	Insights más precisos, automatización del análisis.
Collaborative BI	Integración de herramientas de BI con plataformas de colaboración empresarial.	Equipos de negocio y analistas.	Mayor participación y discusión de insights en equipo.
Data Storytelling & NLP en BI	Visualización narrativa de datos, consultas en lenguaje natural.	Usuarios sin conocimientos técnicos.	Comunicación efectiva de insights, accesibilidad mejorada.
Blockchain & BI Descentralizado	Auditoría segura, registros inmutables, contratos inteligentes.	Empresas de sectores regulados, finanzas.	Seguridad y transparencia en los datos.



ALCANCE DENTRO DE LA ORGANIZACIÓN



CEO
Directores

Tableros
Scorecards

Gerentes
de Negocio

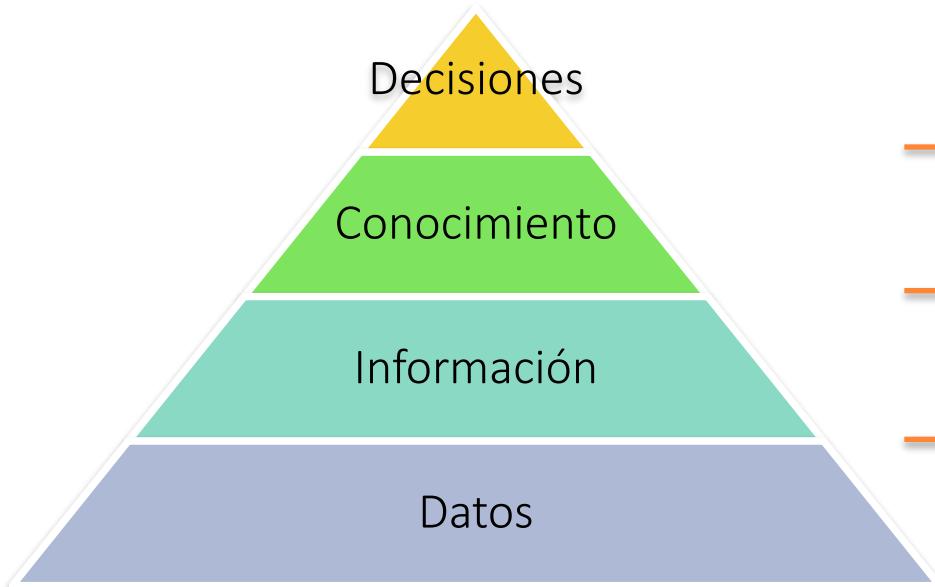
Reportes
Tableros

Analistas
Consumidores
de Información

Reportes
Alertas



ALCANCE DENTRO DE LA ORGANIZACIÓN



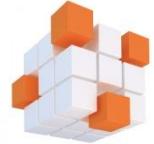
*Definición de estrategias
de negocio a implementar*

Análisis avanzado (Data Mining)
Tableros / Scorecards

Reportes / Tableros
Análisis predictivo (estadístico)

Bases de datos
DW/ Data Mart

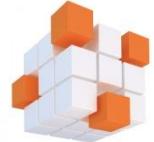
*La cantidad de uso que se le de en las áreas
dependerá de la cultura de la organización*



SISTEMAS OLTP vs OLAP

Sistemas OLTP (On Line Transactional Process)

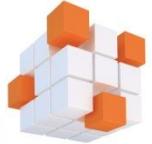
- Sistemas preparados para realizar transacciones y procesos instantáneos.
 - Alineados por módulos o funcionalidad dentro de la aplicación. No integrados.
 - Tiene un objetivo operacional. Se debe garantizar la consistencia de los datos.
 - Actualización online.
 - Disponibilidad de datos recientes o de períodos cortos.
 - Acceso a datos para altas, bajas, modificaciones.
 - Información detallada y no redundante orientada a favorecer la operación transaccional.



SISTEMAS OLTP vs OLAP

Sistemas OLAP (On Line Analytical Process)

- Sistemas preparados para atender consultas complejas y de grandes volúmenes de datos.
 - Integrados y alineados en dimensiones que tienen sentido para el análisis que requiere el negocio.
 - Consolida datos ya validados y los adecua a las necesidades propias de la toma de decisiones.
 - Actualización batch (ETL).
 - Disponibilidad de datos históricos.
 - El acceso a los datos es sólo de lectura y consulta.
 - Información detallada, agregada y redundante para favorecer el análisis.



PROYECTO DE BI - CONSIDERACIONES

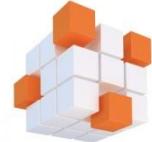
- Un proyecto de BI debe basarse en satisfacer las necesidades de negocio.
- Los datos se deben presentar de acuerdo con las variables del negocio.
- Un DW está en permanente evolución.
- Cada proyecto de BI debe tener un principio y fin.



PROYECTO DE BI - CONSIDERACIONES



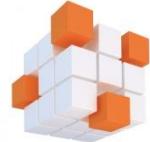
Fuentes: Año 2023: <https://www.tech-wonders.com/2022/12/top-31-business-intelligence-tools-to-go-for-in-2023.html>
<https://www.selecthub.com/c/business-intelligence-tools/>
<https://www.effectivesoft.com/blog/best-business-intelligence-tools-comparison.html>



NECESIDAD DE UNA SOLUCIÓN BI

- ¿Quién necesita soluciones BI?
 - Existen empresas que piensan que no necesitan contar con una solución BI.
- Obtención caótica de la información
 - Problemas cuando se necesita consolidar información o realizar tareas de análisis.
 - ¿Dónde está almacenado cada dato, con qué formato y qué nivel de consistencia tiene?





NECESIDAD DE UNA SOLUCIÓN BI

- ¿Quién necesita analizar la información?
 - El éxito de una organización y de la gestión de la empresa se centra en el uso que se hace de la información.
 - No se puede gestionar lo que no se controla.
 - No se puede controlar lo que no se mide, si no se tiene información para controlar los procesos ocurrirá el caos.

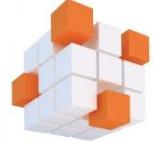


- La información reduce la incertidumbre y facilita la toma de mejores decisiones



NECESIDAD DE UNA SOLUCIÓN BI

- Si una organización tiene un servicio o producto que comercializa, si existen *objetivos* a corto y largo plazo que *deben ser alcanzados* y si existe, por sobre todas las cosas, ideales de *competencia* y *crecimiento*, debe existir también dentro de la empresa un sistema BI.
- Tomar decisiones sin la información adecuada es un riesgo que ninguna empresa debería correr.



NECESIDAD DE UNA SOLUCIÓN BI

- ¿Cuándo *necesita* y *puede* la organización hacer uso de la información?

ES NECESARIO *DECIDIR AHORA* Y SE DEBE *TENER LA INFORMACIÓN AHORA.*

- Suponer que desarrollar un sistema BI es algo lujoso, de costos muy elevados o que es un elemento de Marketing es una concepción errónea de la idea.



Tendencia actual

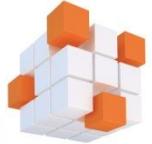


ORGANIZACIÓN IMPULSADA POR DATOS

- El viaje hacia la innovación comienza con los datos.
- Existe una necesidad urgente de capacitar a toda la organización para tomar decisiones rápidas y efectivas. Convertirse en una organización impulsada por datos requiere una estrategia integral de datos.

¡Datos impulsan ideas y agilidad empresarial!





DESAFÍOS PARA CONVERTIRSE EN UNA ORGANIZACIÓN BASADA EN DATOS

- Unificación de datos para una visión completa.
- Capacidad de almacenamiento escalable (servicios en la nube).
- Arquitecturas de seguridad, protegiendo datos sin sacrificar velocidad.
- Analítica ágil mediante herramientas eficientes para decisiones rápidas.
- Capacitación en Machine Learning.
- Automatización de gobernanza, liberando tiempo para la innovación.
- Colaboración multidisciplinaria lo cual impulsa la innovación a través de la diversidad.



CUADRANTE MÁGICO DE BI

- Fabricantes de software BI para empresas (Jun 2024)
Analytics and Business Intelligence (ABI)



Fuente: <https://www.gartner.com/doc/reprints?id=1-2HWG6WTS&ct=240621&st=sb>



CUADRANTE MÁGICO DE BI

○ Fabricantes de software BI para empresas

El Cuadrante Mágico de Gartner es una **publicación anual que mide la situación de mercado de un producto tecnológico**. Contempla, por un lado, la **innovación del proveedor** y, por otro, la **habilidad en el desarrollo** de los productos. Su impulsora, Gartner, es la consultora de IT más importante del mundo, y se dedica a analizar las tendencias de mercado.

Este Cuadrante Mágico de Gartner se ha erigido como **herramienta líder a nivel mundial en la industria IT**, al permitir a las empresas que quieran contratar servicios y soluciones IT tener una **visión global de un área de productos**. Algo que, indudablemente, repercute después en su elección de proveedor.

No es de extrañar que las empresas que figuran en este cuadrante **saquen pecho por ello**: este indicador es una de las principales referencias de IT de la industria. Además de **reconocer la labor** de los proveedores, también les **aporta imagen e influencia** en el sector.

Los estudios del Grupo Gartner son una referencia internacional para **más de 200 productos y servicios**. En el Cuadrante Mágico de Gartner, el **eje X** indica la “**integridad de visión**”, y clasifica la habilidad de los proveedores para aprovechar el momento que vive el mercado para generar valor. El **eje Y** indica la “**capacidad de ejecutar**” y clasifica la habilidad de los proveedores para ejecutar con éxito su visión de mercado.

Los ejes dividen **cuatro zonas de calificación**, que son las siguientes:

- Líderes:** Destacan normalmente por tener una gran cuota de mercado. Desarrollan bien su negocio en función de las características del mercado y están bien posicionados para el futuro.
- Visionarios:** Son capaces de ofrecer productos innovadores. Saben hacia dónde va el mercado, pero no tienen todavía la capacidad de realizar implantaciones por su tamaño u otras circunstancias. Sería es caso de las Startups.
- Aspirantes:** Tienen buena ejecución del negocio y son capaces de dominar un gran segmento del mercado, pero no demuestran un entendimiento real de hacia dónde va éste.
- Jugadores de nichos específicos:** Se enfocan con éxito en un nicho determinado, pero no adquieren una visión global ni se caracterizan por grandes innovaciones.

Microsoft vuelve a coronarse como líder demostrando que, además de **trabajar bien sus servicios y soluciones** (en este caso BI y analítica) tiene **conocimiento de hacia dónde va el mercado y avanzará con él** con facilidad.



CUADRANTE MÁGICO DE BI

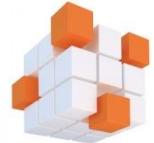
○ 2023-2022-2021



Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms

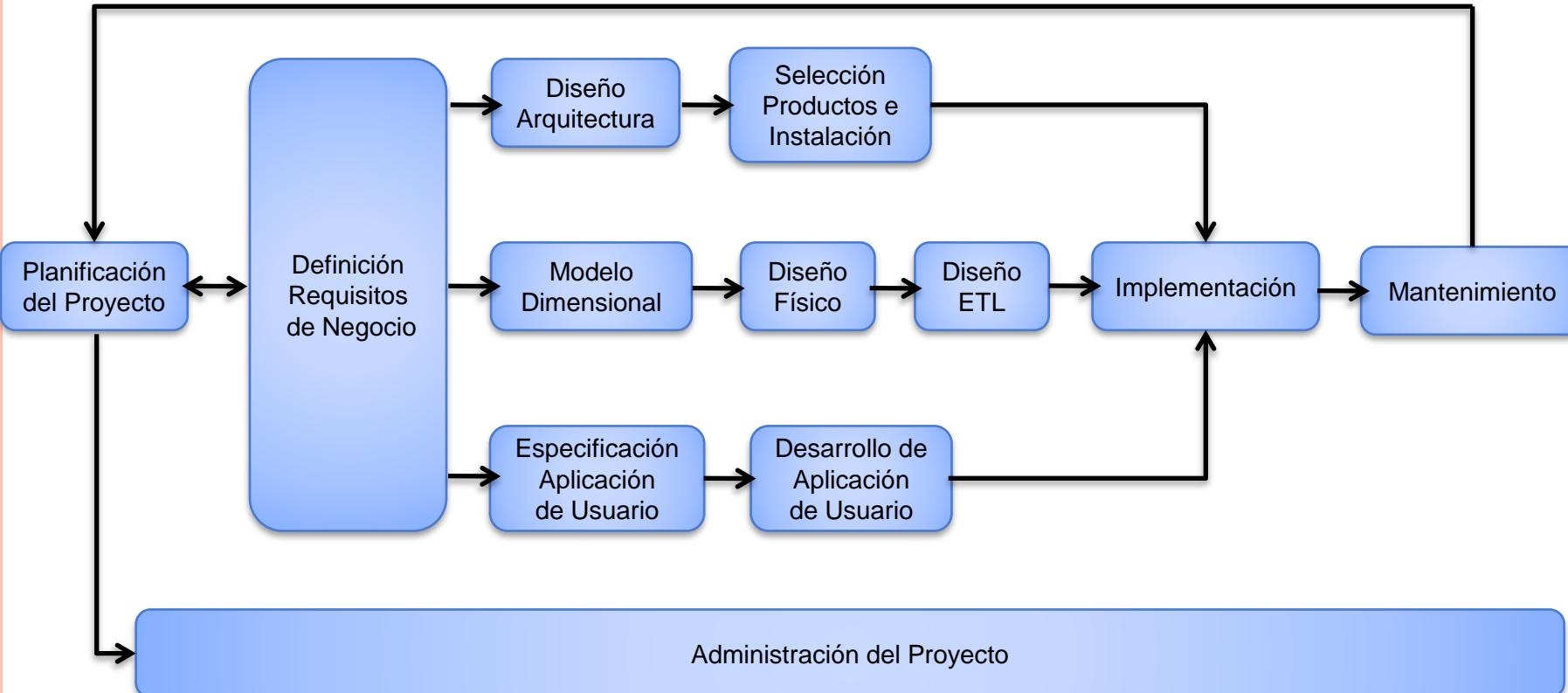


Source: Gartner (February 2021)



CICLO DE VIDA DE UN PROYECTO DE BI

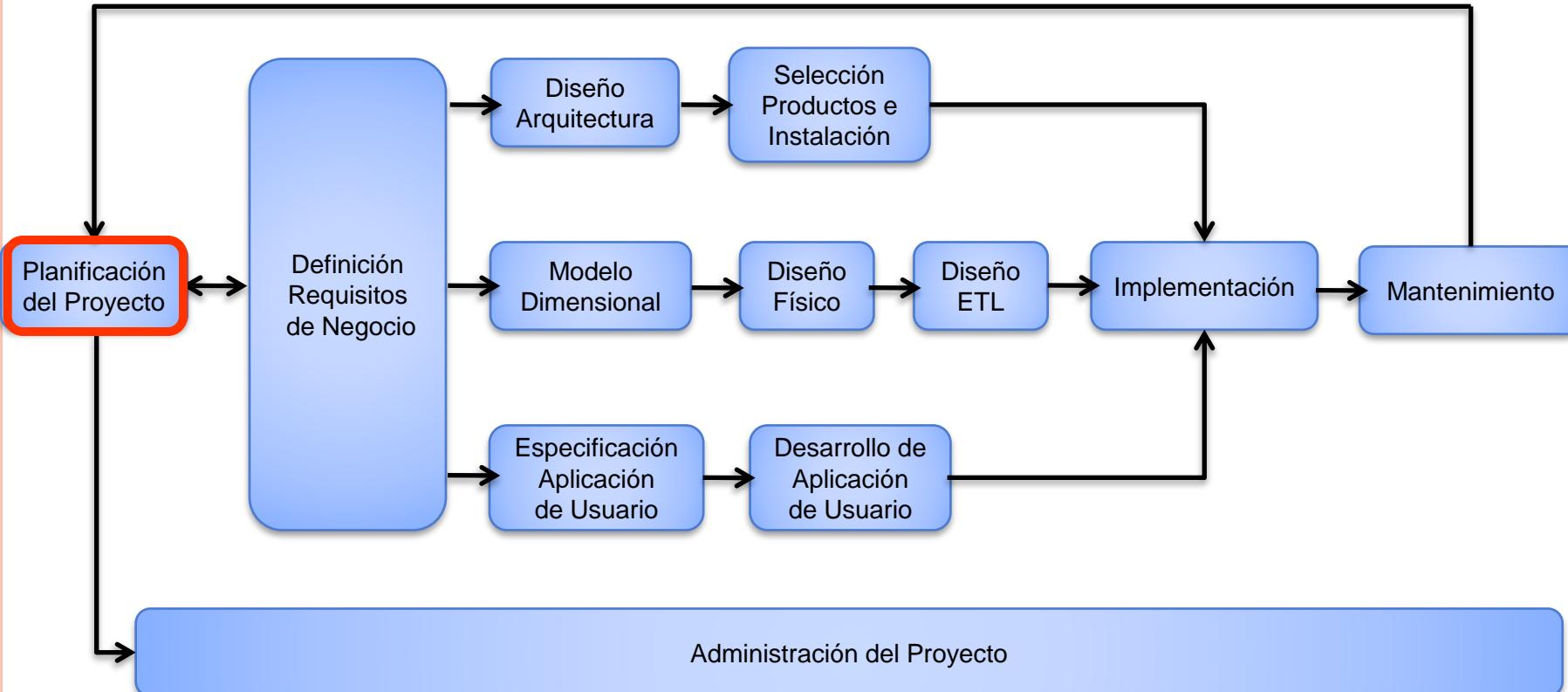
8/4/2025 IN2025





CICLO DE VIDA DE UN PROYECTO DE BI

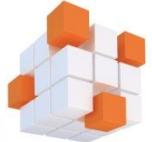
8/4/2025 IN2025





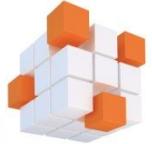
PLANIFICACIÓN DEL PROYECTO

- Ir por área de negocios. Realizar de a un modelo o proceso de negocio.
- Hacer partícipes a usuarios operacionales, técnicos de IT y analistas.
 - Mentalizar a los participantes antes de comenzar, creando expectativas realistas.
 - Expertos y analistas del negocio aseguran la calidad del producto final.
 - Usuarios aseguran la calidad de los datos.
 - Técnicos de TI dan soporte al sistema.



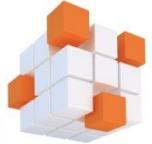
PLANIFICACIÓN DEL PROYECTO

- Seleccionar una aplicación piloto con alta probabilidad de éxito.
 - Estudiar la viabilidad
 - Estudiar los orígenes de los datos
 - Estudiar el hardware y software disponibles
 - Planificar el entrenamiento del usuario
- Construir prototipos rápida y frecuentemente.
- Reportar activamente y publicar los casos de éxito.
- Herramientas para visualizar los datos fáciles de usar.



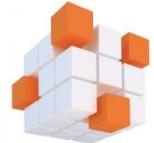
FACTORES DE ÉXITO DEL PROYECTO

- Patrocinador o Sponsor
 - Visión del potencial impacto del DW en la organización.
 - Capacidad de convencer a sus pares para que apoyen el proyecto.
 - Un solo patrocinador podría hacer estancar o demorar el proyecto si éste decide dejar la empresa o atender otros asuntos.
- Apoyo de la alta gerencia de la organización.
- Motivación para la construcción del DW
 - Facilitar el acceso a los datos.
 - Resolver requerimientos específicos y críticos del negocio.
 - Alinearse con los aspectos estratégicos de la empresa. *Ejemplo:* ganar mercado, mejorar la competitividad, etc.



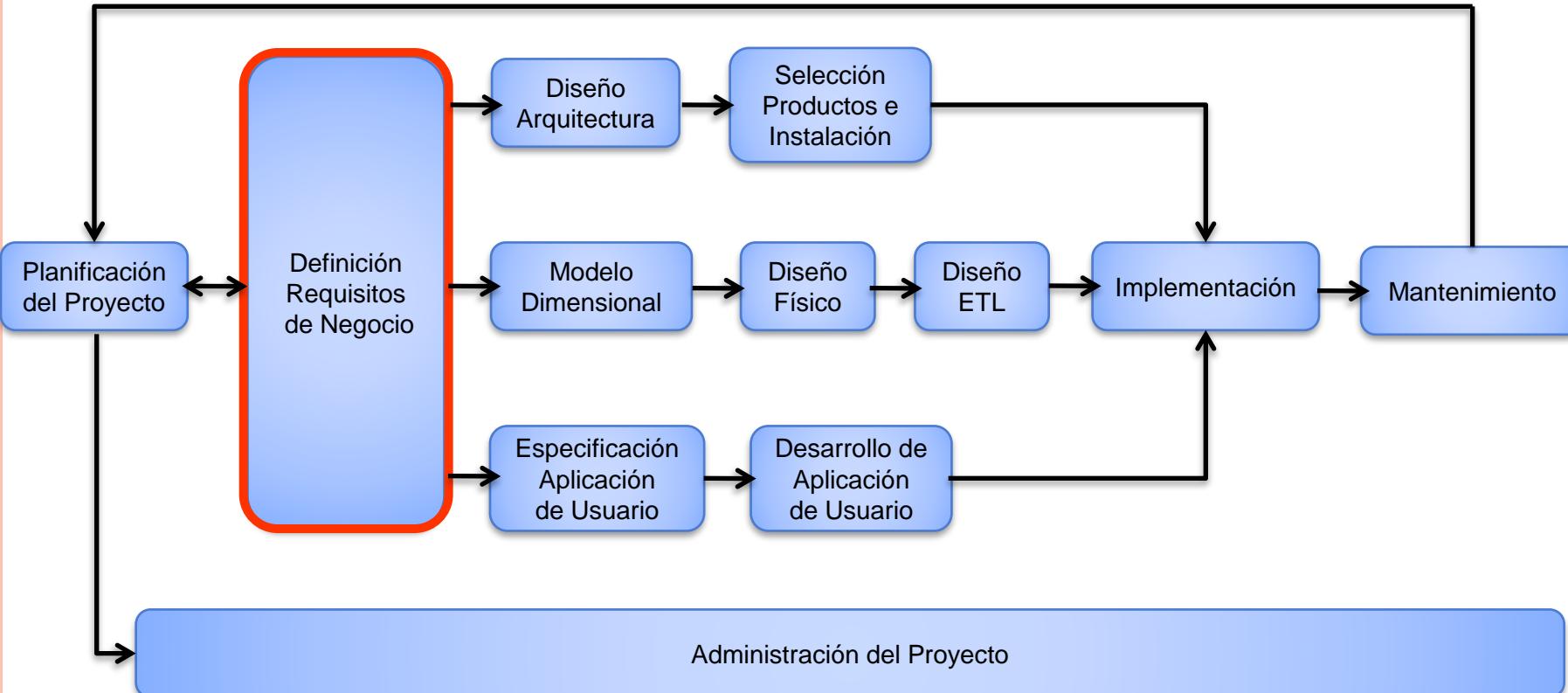
FACTORES DE ÉXITO DEL PROYECTO

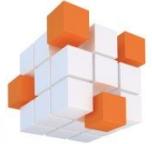
- Viabilidad
 - Tecnología, recursos, etc.
 - Disponibilidad de los datos en los sistemas operacionales.
- Participación de la gente de negocios y TI
 - El proyecto debe propiciar una buena oportunidad para que ambos avancen en la misma sintonía.
- Cultura de análisis de la información
 - ¿Los analistas toman decisiones basadas en hechos reales o basadas en la intuición o hechos anecdóticos?
 - Si no hay uso actual de información es probable que la empresa no necesite un DW.
 - Se debe invertir un esfuerzo considerable en cambiar la cultura de la organización.



CICLO DE VIDA DE UN PROYECTO DE BI

8/4/2025 IN2025

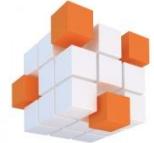




DEFINICIÓN DE REQUISITOS DEL NEGOCIO

- Foco en el proceso de negocio e indicadores.
 - Sino el proyecto se demora o se gasta más.
 - No alcanza los objetivos previstos.
 - ¿Qué hacen los usuarios, cómo lo hacen, qué resultados buscan, etc.?
- Identificar el ciclo del proceso de negocio.
 - Seleccionar el proceso que queremos modelar.
 - *Ventas*
 - Establecer el inicio del proceso, qué eventos lo desencadenan.
 - *Pedidos*
 - Establecer el fin del proceso, qué salidas genera.
 - *Entregas*

Esto ayudará a establecer el alcance del modelo a realizar



DEFINICIÓN DE REQUISITOS DEL NEGOCIO

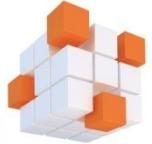
- Establecer reuniones y entrevistas con los usuarios
 - Prioridades y objetivos de la entrevista y del proyecto.
 - Buscar que el usuario sea el que hable.
 - Buscar indicadores y datos que nos ayuden a bosquejar la dimensionalidad del negocio.
 - ¿Qué considera el usuario como proyecto exitoso?
- Documentar las reuniones y las fases del proyecto.

La solución BI a implementar deberá ir de la mano con la visión de la empresa y los objetivos de negocio.



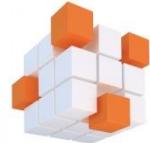
DEFINICIÓN DE REQUISITOS DEL NEGOCIO

- Dar una definición del ciclo del proceso, su misión y objetivos principales, los atributos de cada objetivo así como el cliente interno o externo del proceso.
 - El cliente realiza los pedidos de los productos que desea. Los pedidos son procesados en casa central, y en caso que haya stock del producto solicitado, se concreta el pedido generando una factura (venta concretada). Los productos vendidos son entregados por el equipo de reparto en la fecha acordada. Se trata que la cantidad de clientes a los que se les vende crezca el 1% mensual.
 - Existen dos modalidades de venta: directa e indirecta. Se trata que la mayor cantidad de ventas sea directa.
 - Los objetivos de las ventas están establecidos a principio de cada mes con un volumen mínimo por cumplir que está estipulado por área de venta.



DEFINICIÓN DE REQUISITOS DEL NEGOCIO

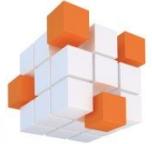
- Transformar los objetivos y sus atributos operacionales en mediciones e indicadores.
 - El cliente realiza los pedidos de los productos que desea. Los pedidos son procesados en casa central, y en caso que haya stock del producto solicitado, se concreta el pedido generando una factura (venta concretada). Los productos vendidos son entregados por el equipo de reparto en la fecha acordada. **Se trata que la cantidad de clientes a los que se les vende crezca el 1% mensual.**
 - Existen dos modalidades de venta: directa e indirecta. **Se trata que la mayor cantidad de ventas sea directa.**
 - Los objetivos de las ventas están establecidos a **principio de cada mes con un volumen mínimo por cumplir** que está **estipulado por área de venta.**



DEFINICIÓN DE REQUISITOS DEL NEGOCIO

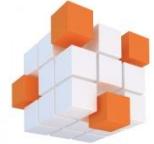
○ Identificar sustantivos.

- El **cliente** realiza los **pedidos** de los **productos** que desea. Los pedidos son procesados en casa central, y en caso que haya stock del producto solicitado, se concreta el pedido generando una **factura** (venta concretada). Los productos vendidos son entregados por el equipo de reparto en la **fecha** acordada. Se trata que la **cantidad de clientes** a los que se les vende crezca el 1% mensual.
- Existen dos **modalidades de venta**: directa e indirecta. Se trata que la mayor cantidad de ventas sea directa.
- Los **objetivos de las ventas** están establecidos a principio de cada **mes** con un **volumen mínimo** por cumplir que está estipulado por **área de venta**.

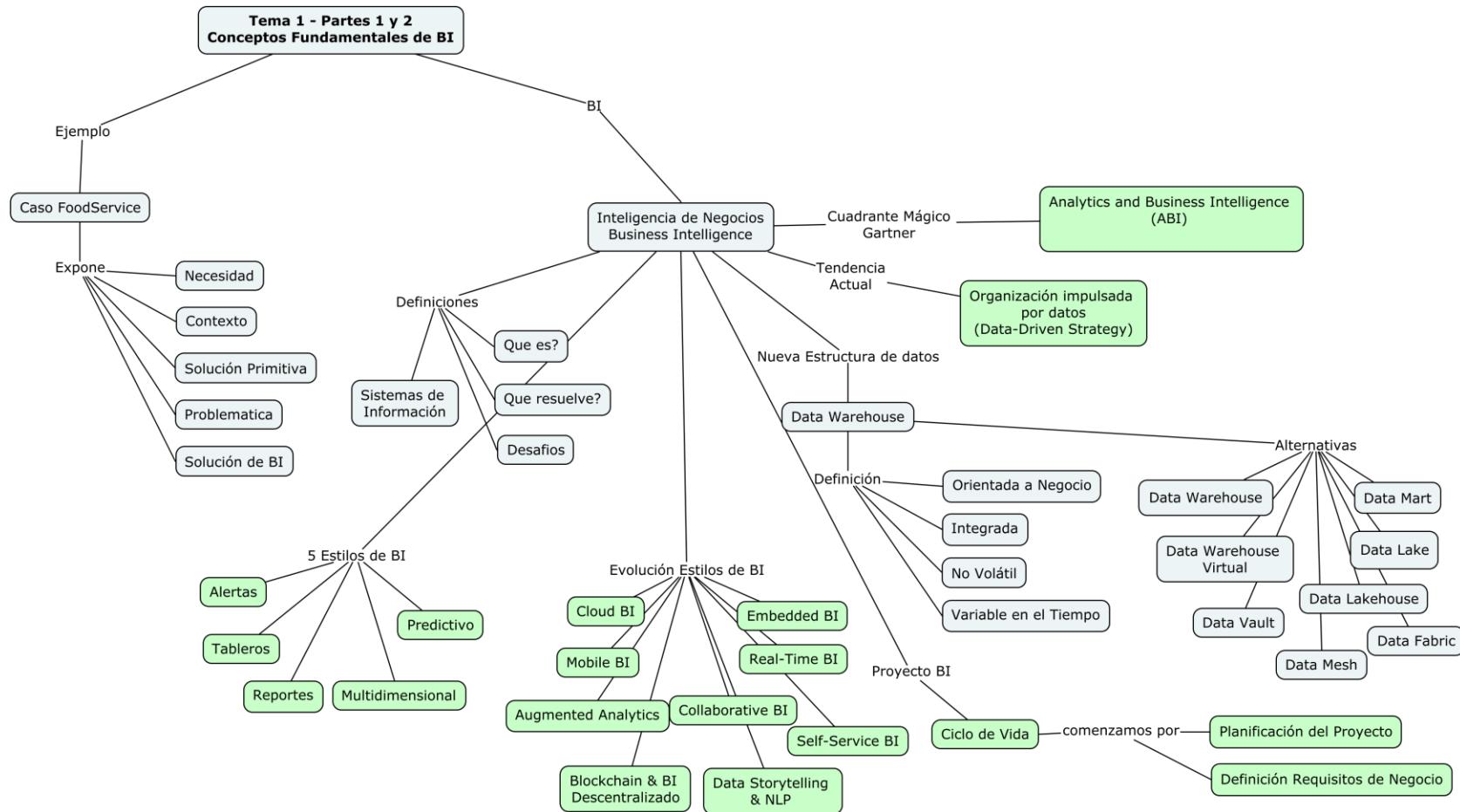


DEFINICIÓN DE REQUISITOS DEL NEGOCIO

- Consultar con los stakeholders como miden el proceso de negocio y entender cómo se calcula cada métrica.
 - ✓ Volumen vendido 5% mayor al objetivo.
 - ✓ Incremento en las marcas premium respecto al mes anterior.
 - ✓ Crecimiento del 1% en la cantidad de clientes a los que se les vendió.



RESUMEN TEMA 1 – PARTES 1 Y 2



TP CASOS ESTUDIO: PFIZER / KODAK

ENTREGA OPCIONAL (*)

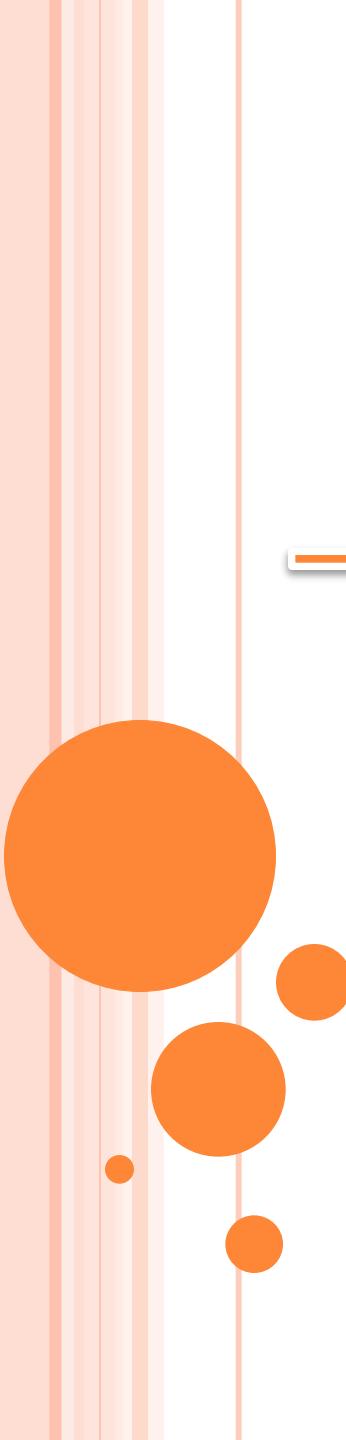


- Analizar los casos y responder a los interrogantes que surgen a partir de las necesidades de información y decisión enunciadas.
- Utilizando la plantilla del Test de Disponibilidad de Litmus, completar cada uno de los factores tenidos en cuenta los casos, en función de la disposición de la empresa para la implementación de una solución de BI.

(*)

- *Los TPs Opcionales sirven para afianzar conceptos necesarios para realizar los TPs de Aplicación, serán corregidos en clase y/o mediante Autoevaluación.*
- *Los TPs de Aplicación tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción, serán corregidos en detalle por los docentes.*
- *Ver documentos: "Condiciones de Cursada en MIeL - sección: Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio" + "Círculo Entrega TPs Teams/MIeL"*





UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

Modelo Dimensional – Parte 1

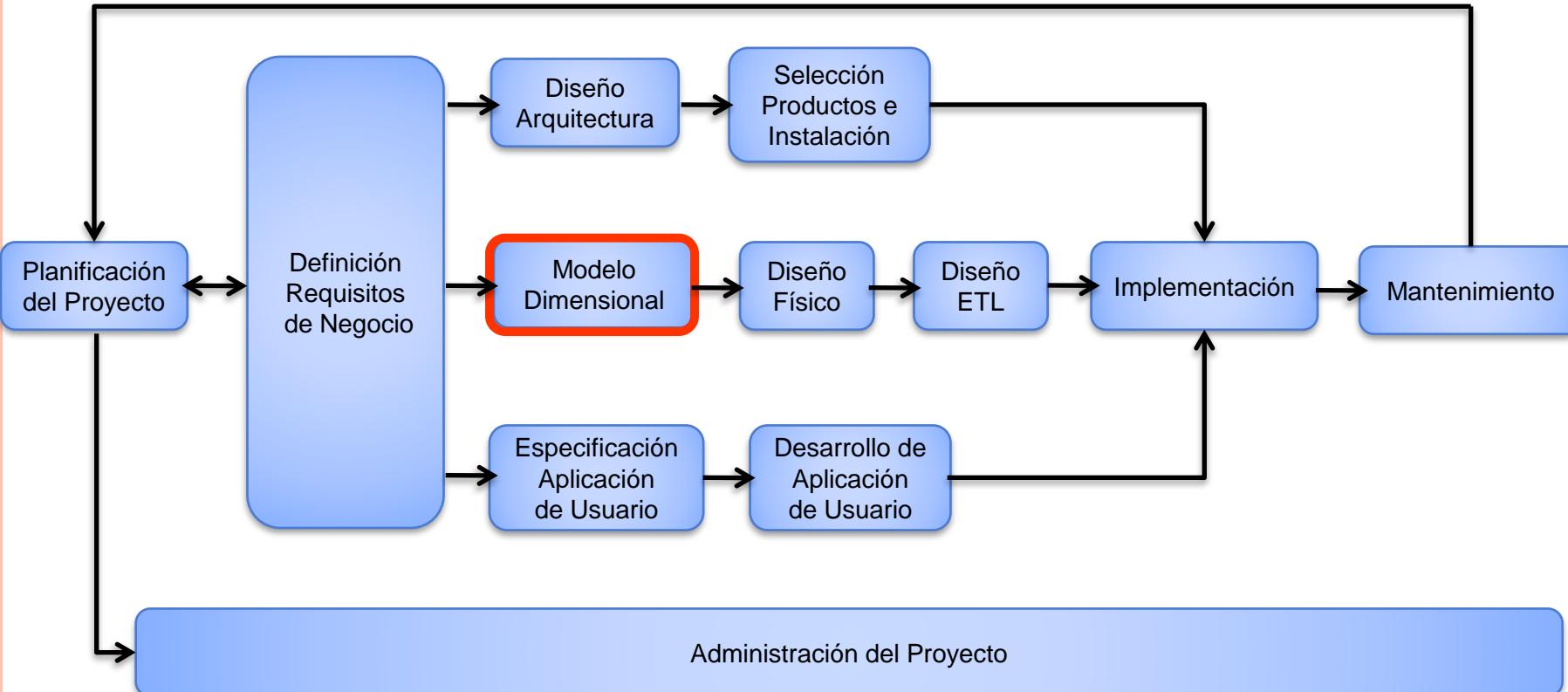
Docentes: ING. LORENA R. MATTEO

Autores ppt orig.: Lic. HUGO M. CASTRO / MG. DIEGO BASSO



CICLO DE VIDA DE UN PROYECTO DE BI

15/4/2025 IN2025





MODELO DE DATOS

○ Modelo de Entidad-Relación (OLTP)

- Orientado a la implementación de los procesos transaccionales
- Tareas operacionales
- Entidades, Atributos y Relaciones

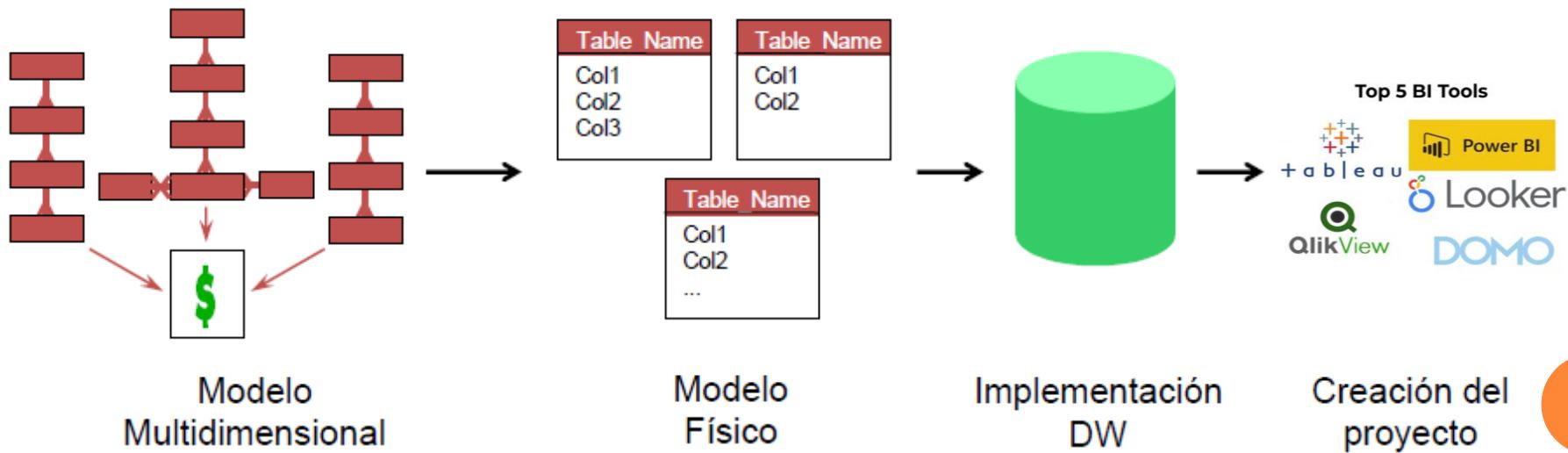
○ Modelo Dimensional (OLAP)

- Orientado a las características del negocio (variables del negocio)
- Ofrece una visión de los datos orientada hacia el análisis y una rápida y flexible navegación por estos.



MODELO DIMENSIONAL

- “El modelo dimensional es una técnica de diseño que busca presentar los datos en un framework estándar, intuitivo y escalable, que permite un acceso a los datos altamente performante, basándose en el modelado relacional, pero con algunas restricciones de diseño importantes”.





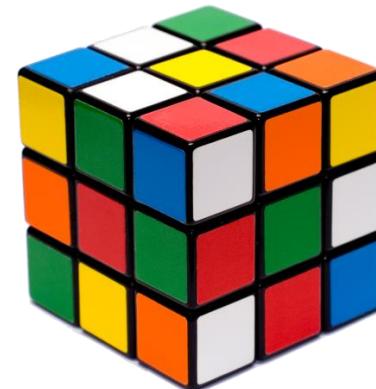
MODELO DIMENSIONAL

- Busca acercar los datos a la manera en que estos serán convertidos en información útil para los usuarios del negocio.
- Componentes
 - **Hechos** (Fact) / Medidas / Indicadores de negocio
 - **Dimensiones**
 - **Atributos**
 - Elementos
 - Jerarquías
 - **Relaciones**



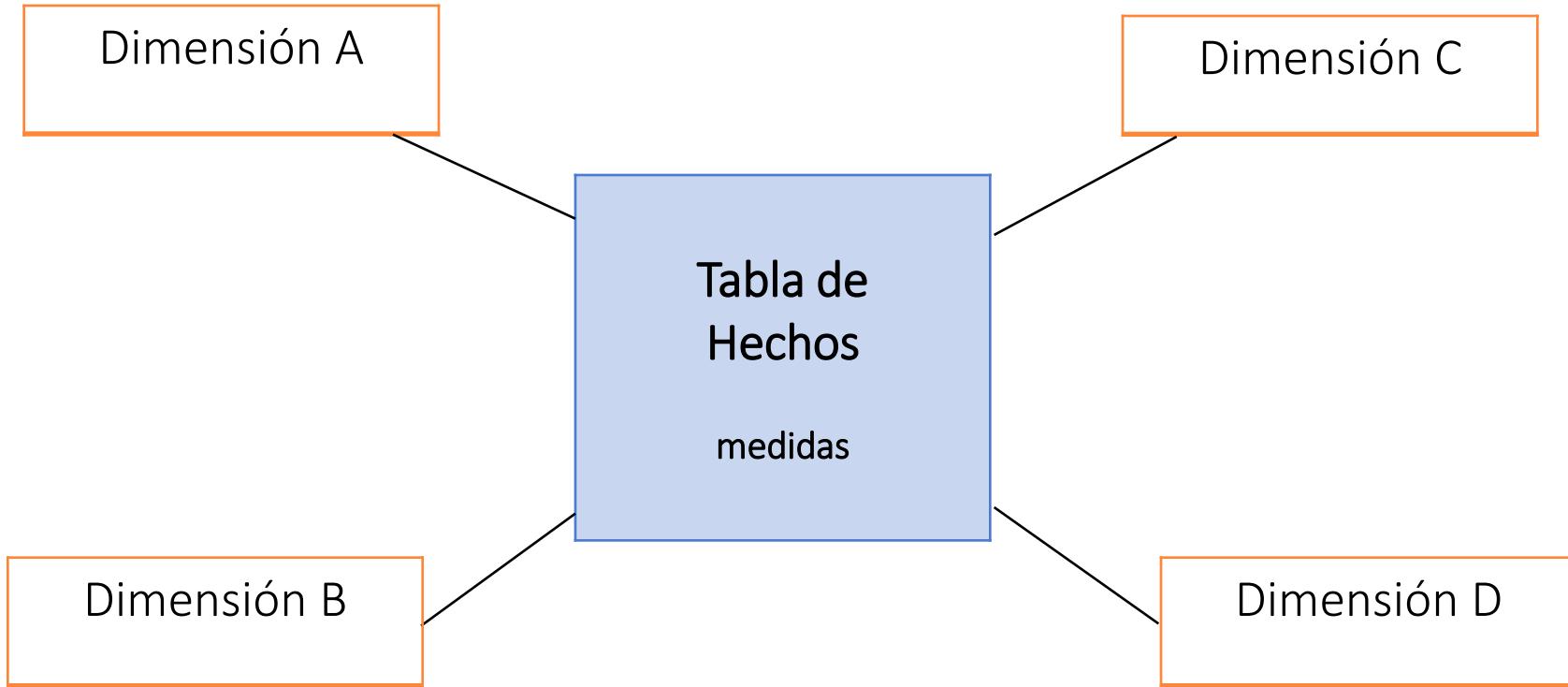
MODELO DIMENSIONAL

- Representación de la estructura que va a tener el DW.
- Define un nivel mínimo de detalle (**Granularidad**).
- Se puede revisar con el profesional de negocios.
- El modelo de datos dimensional es lo que hace que un DW sea una base de datos **orientada al negocio**.
- La forma de representar la organización de los datos en un modelo dimensional es a través de un **cubo**.





ESQUEMA DE UN MODELO DIMENSIONAL





HECHOS O MEDIDAS (FACTS)

○ Hecho

- Es algo que ocurre en el tiempo (operación o actividad).

○ Medidas / Indicadores

- Son valores **numéricos** que describen el **hecho** analizado.
- Las medidas más usadas son aditivas.
- Ayudan a medir el desempeño del negocio.
- Tipos
 - **Básicas**: existen físicamente en el DW junto a los atributos que las caracterizan. Provienen de las fuentes de datos OLTP.
 - **Calculadas o derivadas**: construidas a partir de medidas básicas y pueden o no estar almacenadas en el DW.



HECHOS O MEDIDAS (FACTS)

○ Medidas / Indicadores

- Columnas agregables del DW
- Admiten sumas, promedios, consolidaciones, etc.
- Dan lugar a las métricas.

○ Ejemplos

- ✓ \$ Venta (en un supermercado)
- ✓ Unidades vendidas (en un supermercado)
- ✓ Minutos hablados (en una telefónica)
- ✓ Duración de la llamada (en una telefónica)
- ✓ Saldo \$ (en un banco)

○ Todos los hechos, medidas o indicadores de negocio se guardan en la **tabla de hechos** y deben tener el **mismo nivel de granularidad**.



DIMENSIONES

- Categorías para describir el contexto (variables del negocio) en el que se pueden analizar las medidas.
- Contienen uno o más **atributos** de tipo texto agrupados lógicamente. Las dimensiones no comparten atributos.
- Sirven para **mostrar, agrupar** y **filtrar** información.
- Ejemplos

- ✓ TIPO DE LLAMADA
- ✓ SUCURSAL
- ✓ CLIENTE
- ✓ PRODUCTO
- ✓ EMPLEADO
- ✓ GEOGRAFÍA
- ✓ VENDEDOR
- ✓ AÑO

- Si hay valores numéricos deben ser categorizados.
 - *Ejemplos:* Rangos de edades, niveles de precios



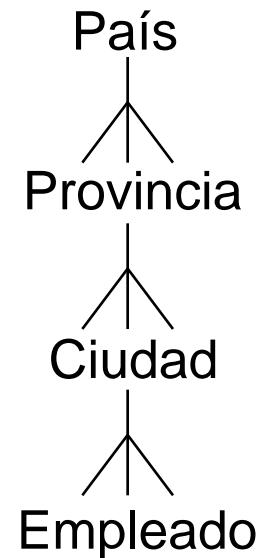
DIMENSIONES CON ATRIBUTOS

- Las dimensiones definen los niveles de análisis (**jerarquías**). Esto permite obtener summarización o agregación.

DIM Tiempo

TIEMPO	
■■■	AÑO
■■■	SEMESTRE
■■■	BIMESTRE
■■■	MES
■■■	FECHA
<new level>	

DIM Geografía

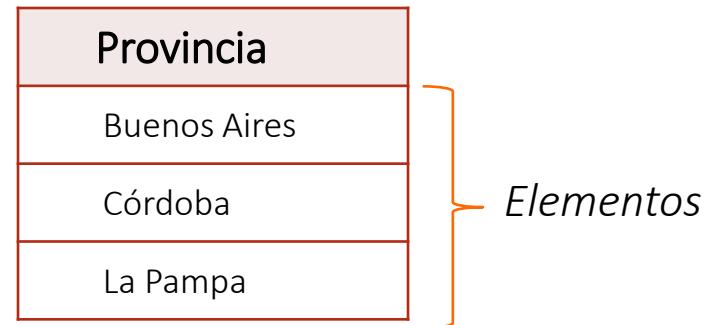




ATRIBUTOS

- Representan categorías o clases de elementos que tienen el mismo nivel lógico dentro de una dimensión.
- Valores cualitativos de una transacción.
- Visualizar la información de las dimensiones a distintos **niveles de detalle** y **agrupar** datos para analizar las métricas.
- Son una agrupación de **elementos** (valores que toma).
- Son parecidos a las entidades del modelo E-R.
- Ejemplos:

- | | |
|-------------|------------|
| ✓ Provincia | ✓ Empleado |
| ✓ Ciudad | ✓ Edad |
| ✓ Mes | ✓ Cliente |
| ✓ Ítem | ✓ Año |





ELEMENTOS

- Son los valores o instancias que toma un **atributo**.
- Ejemplos:

- | | |
|-------------|------------|
| ✓ Provincia | ✓ Empleado |
| ✓ Ciudad | ✓ Edad |
| ✓ Mes | ✓ Cliente |
| ✓ Ítem | ✓ Año |

Provincia
Buenos Aires
Córdoba
La Pampa

Elementos

Mes
Enero 2024
...
Diciembre 2024

Elementos

Edad
0-15 años
16-30 años
31-45 años

Elementos



JERARQUÍAS DE ATRIBUTOS

- Ordenamiento lógico de atributos dentro de la dimensión.
- Contiene relaciones entre los atributos de una misma dimensión.
 - 1:1
 - 1:N
 - N:N
- Las **jerarquías** definidas permiten bajar el nivel de complejidad de los datos y ayudar al usuario a ver los datos en el nivel de detalle que necesita.



JERARQUÍAS DE ATRIBUTOS

- **Importante:** si en el modelo de la dimensión aparece un sólo nivel, prestar atención a la cardinalidad.
 - *Cardinalidad alta:* Hace necesaria la definición de jerarquías.
 - Ejemplo: Si hay 20.000 productos, hace falta definir niveles jerárquicos para poder analizar.
 - *Cardinalidad baja:* Aceptable no tener jerarquías.
 - Ejemplo: Dimensión Mercado tiene dos posibles valores (Local o Extranjero).



GRANULARIDAD

- Es el nivel más detallado (mínimo nivel de detalle o nivel atómico) que tiene cada tabla de hechos en el modelo del proceso de negocio. Se expresa como:

<AtribDim1> x ... x < AtribDimN>

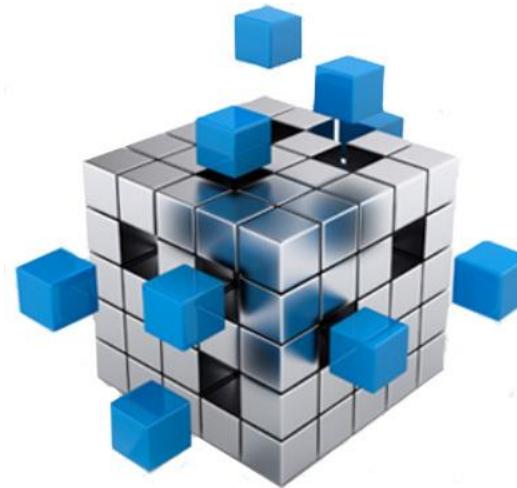
siendo AtribDimx el nivel más detallado de cada Dimensión

- Ejemplos:
 - la granularidad de un modelo dimensional que consta de las dim. de fecha, almacén y producto es producto vendido en el almacén por día, es decir: **Producto x Almacén x Fecha**
 - si en el modelo existe la dimensión Fecha (Tiempo) (con las jerarquías de año y trimestre) la granularidad de esa dimensión en la tabla de hechos estará hasta el nivel trimestral, es decir no se tendrá información para los días o meses individuales: **<AtribDim1>x ... x Trimestre**



CUBO OLAP

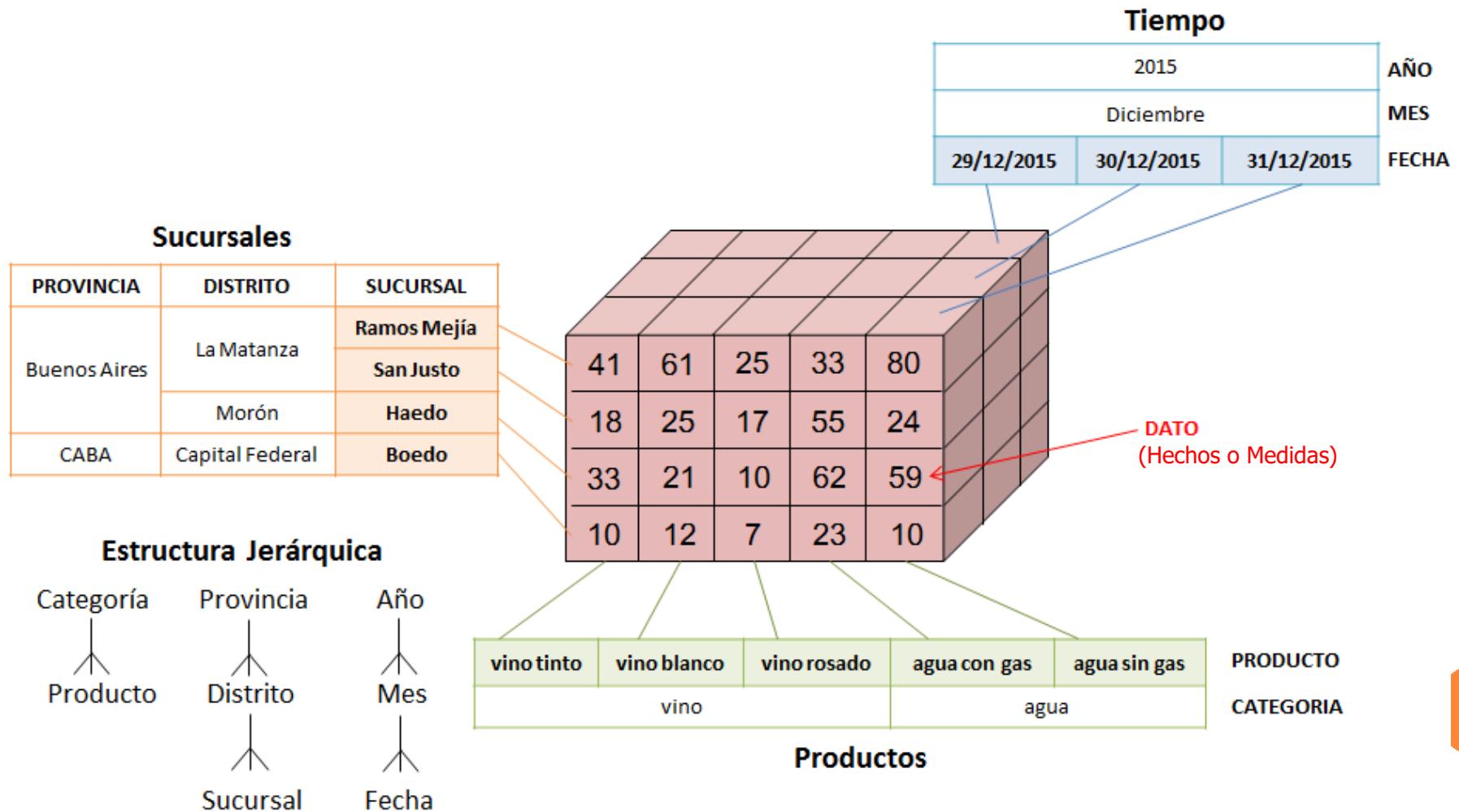
- Es una base de datos que posee diversas dimensiones.
- Cada dimensión es una “arista” del cubo.
- En las celdas del cubo están las medidas (hechos).
- Es independiente de la forma en que realmente se almacenan los datos.





EJEMPLO DE CUBO OLAP

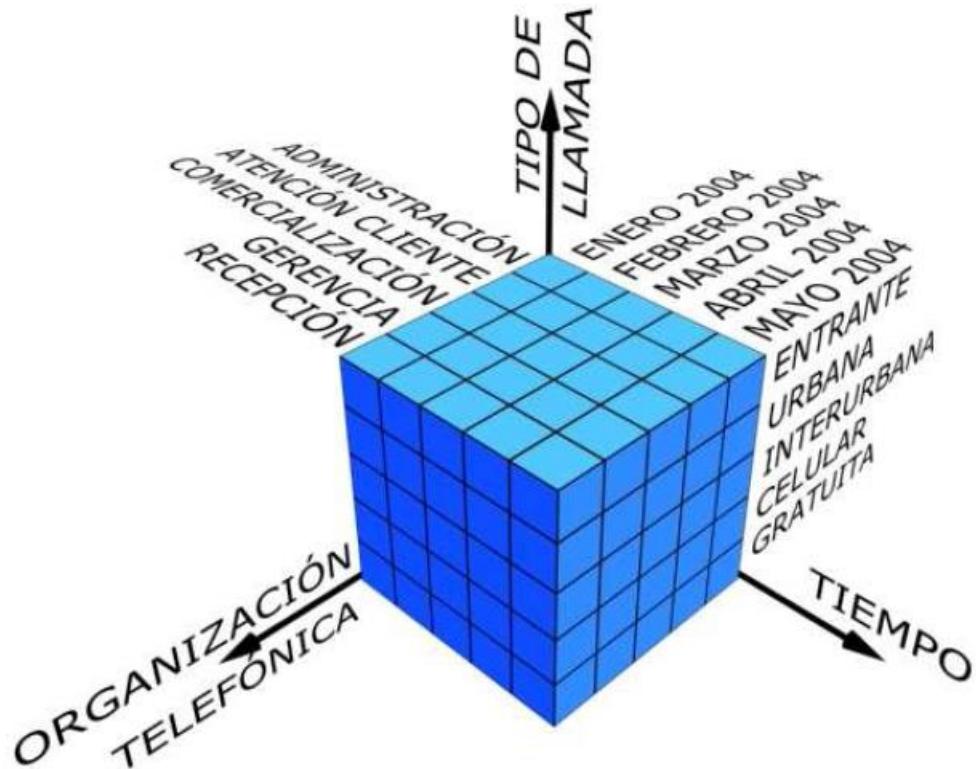
- Cubo que posee como medida la “Cantidad de pedidos” por Fecha, Producto y Sucursal.





EJEMPLO DE CUBO OLAP

- Cubo que posee como medida la “*Duración de las Llamadas*” y como dimensiones Tiempo, Tipo de Llamada y Organización Telefónica.





¿POR DÓNDE EMPEZAMOS?

- La variedad y disparidad de datos dentro de una empresa requiere un enfoque incremental.
- No se puede cargar todo y ver después quién lo va a usar.
- Este enfoque puede llevar al fracaso.
- Se pierde tiempo y esfuerzo sin obtener resultados.





¿POR DÓNDE EMPEZAMOS?

- Por las necesidades de uso de información de los usuarios de negocios.
 - Es decir, por los *requerimientos*.
- Priorizar
 - Comenzar por lo más importante para la empresa.
 - ¿Qué valores numéricos necesitan analizar?
 - ¿Cuáles son el/los procesos de negocios que generan esos valores?





PROCESO DE DISEÑO DIMENSIONAL

Pasos

1. Elegir el **proceso de negocios** a modelar (ventas diarias, manejo de stock, etc.).
2. Elegir las **dimensiones** que van a intervenir (tener en claro los **atributos, jerarquías** y elementos de atributos).
3. Elegir la **granularidad** (nivel de detalle) del proceso de negocios con que se van a guardar los datos en el DW.
4. Elegir los **hechos** y **medidas** que se van a utilizar en la tabla de hechos (básicas y calculadas).

*Metodología de Ralph Kimball para
el armado incremental del DW*



CASO DE ESTUDIO

○ Relevamiento

- Un supermercado posee distintas sucursales en distintos lugares del país en las que comercializa distintos productos.
- Cada producto pertenece a una familia y cada familia pertenece a un tipo de producto. Ejemplo: el producto Yogur A x 170gr pertenece a la familia yogur y al tipo lácteo.
- Registra los tickets de las ventas de sus sucursales en una BBDD.
- La información que se registra de cada producto vendido es :

• Fecha	• Costo de cada producto
• Hora	• Cantidad de unidades
• Sucursal	• Precio unitario
• Nro. de ticket	• Importe
• Producto	



CASO DE ESTUDIO

○ Relevamiento

- Quiere aprovechar esta información del año actual y los dos anteriores para ver, entre otras cosas.
1. ¿Cuánto se vendió (en unidades e importes) por sucursal y tipo de producto en los distintos trimestres, del mismo año o año actual y alguno de los anteriores.
 2. ¿Cómo fueron las ventas en las distintas provincias en las que están las sucursales?
 3. ¿Cuál fue la ganancia obtenida en la venta de cada producto?
 4. ¿Cómo fueron las ventas mensuales con tarjetas (crédito y débito) respecto al pago en efectivo (por sucursal, por producto, por mes, o trimestre o año)?
 5. ¿Cómo evolucionaron las ventas en los últimos 12 meses?



CASO DE ESTUDIO

ANÁLISIS DE LOS REQUERIMIENTOS

1. Ventas por sucursal y tipo de producto en los distintos trimestres, del mismo año o año actual y alguno de los anteriores.
 - ❑ Medidas básicas: cantidad de unidades, importe de venta
 - ❑ Ver el detalle de las medidas por **sucursal**. El supermercado posee distintas sucursales.
 - ❑ Las medidas están referidas a un **producto** y a su vez se desea verlos agrupados por **tipo de producto**.
 - ❑ A su vez cada producto pertenece a una **familia de producto** y cada familia pertenece a un tipo de producto.
 - ❑ Ver la evolución de las ventas en el transcurso del **tiempo**, detallados por **trimestre** o **año**.
- ✓ Para satisfacer estas necesidades debemos crear las dimensiones **SUCURSAL**, **PRODUCTO** y **TIEMPO**.



CASO DE ESTUDIO

ANÁLISIS DE LOS REQUERIMIENTOS

2. Ventas en las distintas provincias en las que están las sucursales.
 - ❑ La información de las ventas por **sucursal** requiere que también pueda visualizarse por **provincia** asociada.
3. Ganancia obtenida en la venta de cada producto.
 - ❑ Medidas básicas: **costo del producto**
 - ❑ Medidas derivadas o calculadas: **ganancia de venta**
4. Ventas mensuales con tarjetas (crédito y débito) y pago en efectivo (por sucursal, por producto, por mes, o trimestre o año)
 - ❑ Analizar las ventas por diferentes **formas de pago**.
5. Ver evolución de las ventas en el transcurso del **tiempo**, detallados por **trimestre**, **año** o **mes**. La información del ticket permite conocer hasta la **fecha** y **hora**.



CASO DE ESTUDIO

ANÁLISIS DE LOS REQUERIMIENTOS

- Resumen de los requerimientos, agrupar por apertura de las dimensiones y comenzar a definir los cubos a crear.

Hecho a medir: Venta de Productos			
	Dimensiones		
Medidas	Tiempo	Sucursal	Producto
Cantidad	X	X	X
Importe	X	X	X
Costo	X	X	X
Ganancia	X	X	X

- Proceso de negocio a modelar: **Ventas**

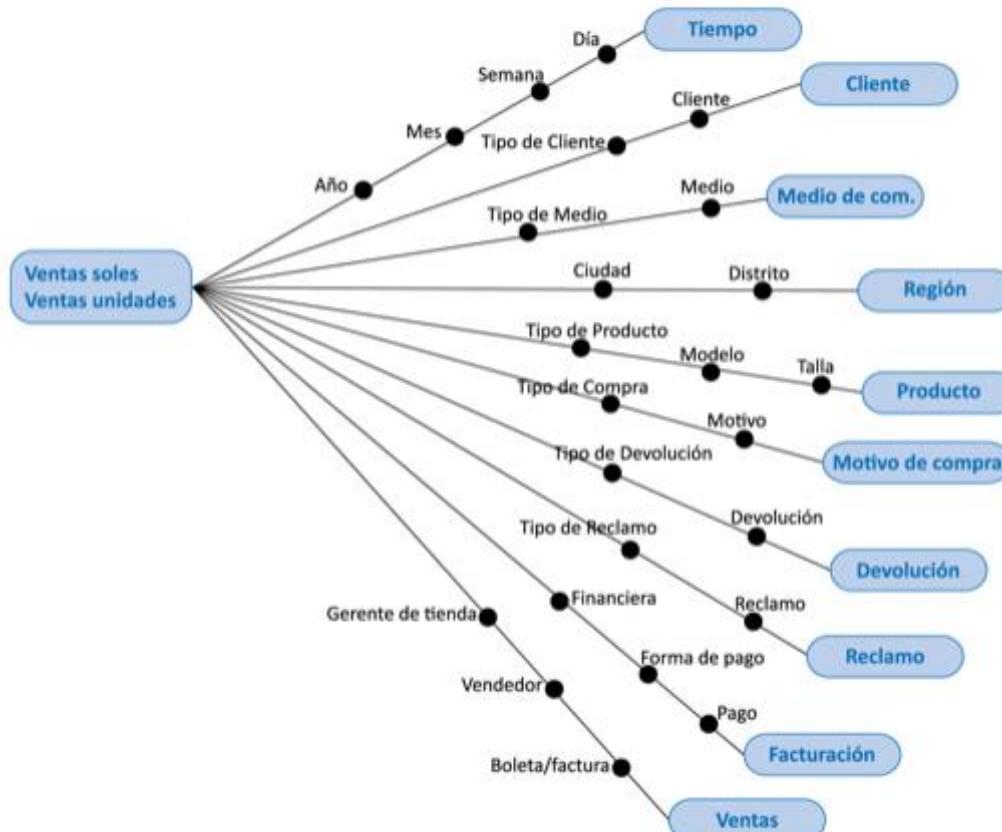


CASO DE ESTUDIO

ANÁLISIS DE LOS REQUERIMIENTOS STARNET

- Otra forma de mostrar los requerimientos gráficamente si se cuenta con ese nivel de detalle:

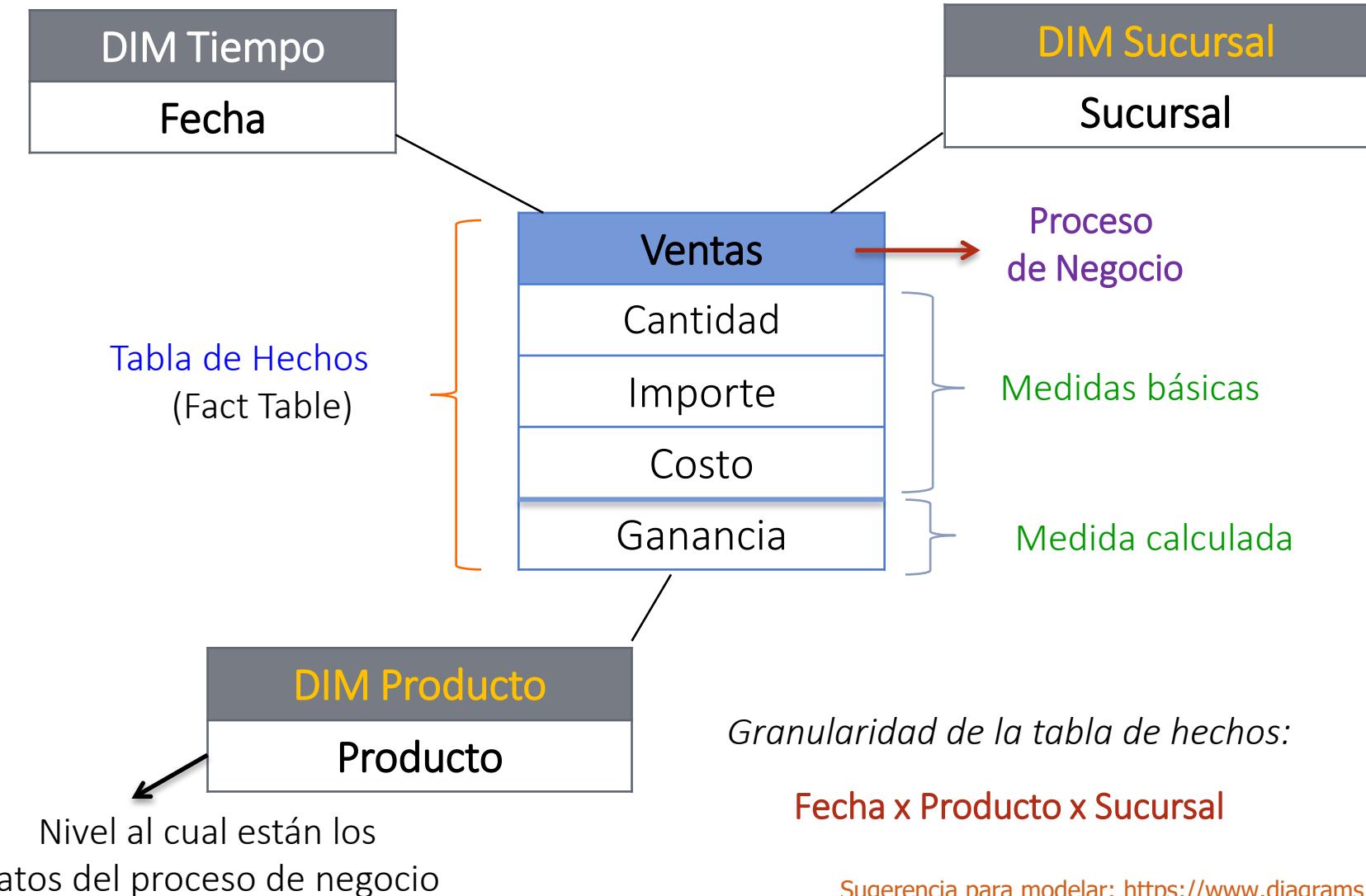
Diagrama Starnet de requerimientos para un sistema de Business Intelligence



Fuente: <https://docplayer.es/81892156-Rediseno-de-las-areas-de-marketing-ventas-y-recursos-humanos-del-retail-exxa-en-lima.html>



MODELO DIMENSIONAL “CONCEPTUAL” (GRUPO DE HECHOS Y GRANULARIDAD)





RELACIONES ENTRE ATRIBUTOS

1:1 (uno-a-uno)

- A cada código de artículo le corresponde una descripción.
- A cada descripción le corresponde un código de artículo.
- Ambos atributos forman parte de la misma dimensión.

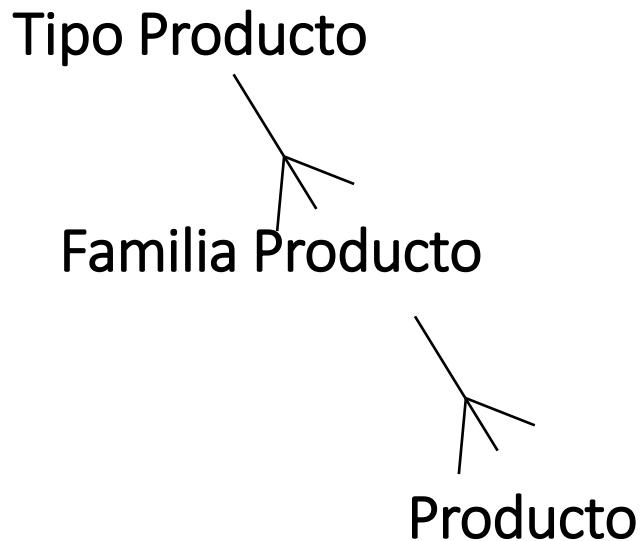
1:N (uno-a-muchos)

- Analicemos la dimensión *Producto*
- Cada **producto** pertenece a una **familia de producto** y cada familia pertenece a un **tipo de producto**.
 - Una familia de productos comprende varios productos
 - Un tipo de producto comprende varias familias



DIMENSIÓN PRODUCTO

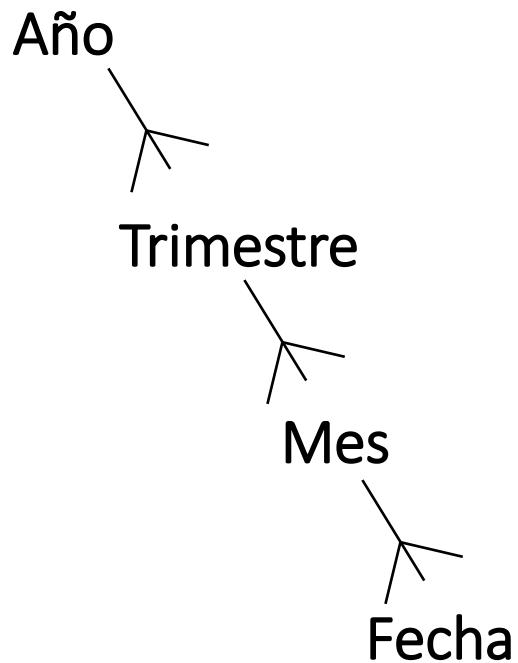
- La relación jerárquica es:





DIMENSIÓN TIEMPO

- ¿De qué manera se quiere analizar la información?
 - Monto de ventas por sucursal y tipo de producto en distintos **trimestres**, del mismo **año** o año actual.
 - Evolución de ventas en los últimos 12 **meses**.
- La relación **jerárquica** es:

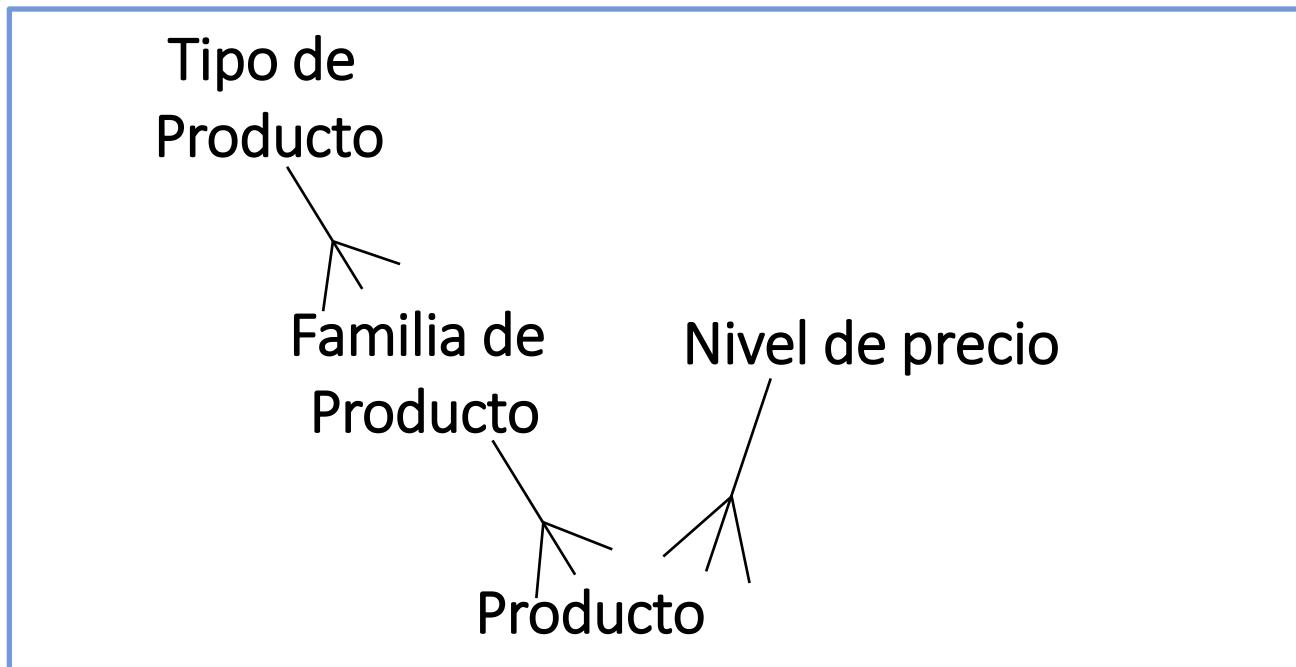




DIMENSIONES CON JERARQUÍAS MÚLTIPLES

- Puede ocurrir que además de agrupar los productos por **tipo de producto** y **familia de producto** sea útil agruparlos por **nivel de precio** (caro, mediano, barato)
- Esto se representa con una doble jerarquía:

DIM PRODUCTO





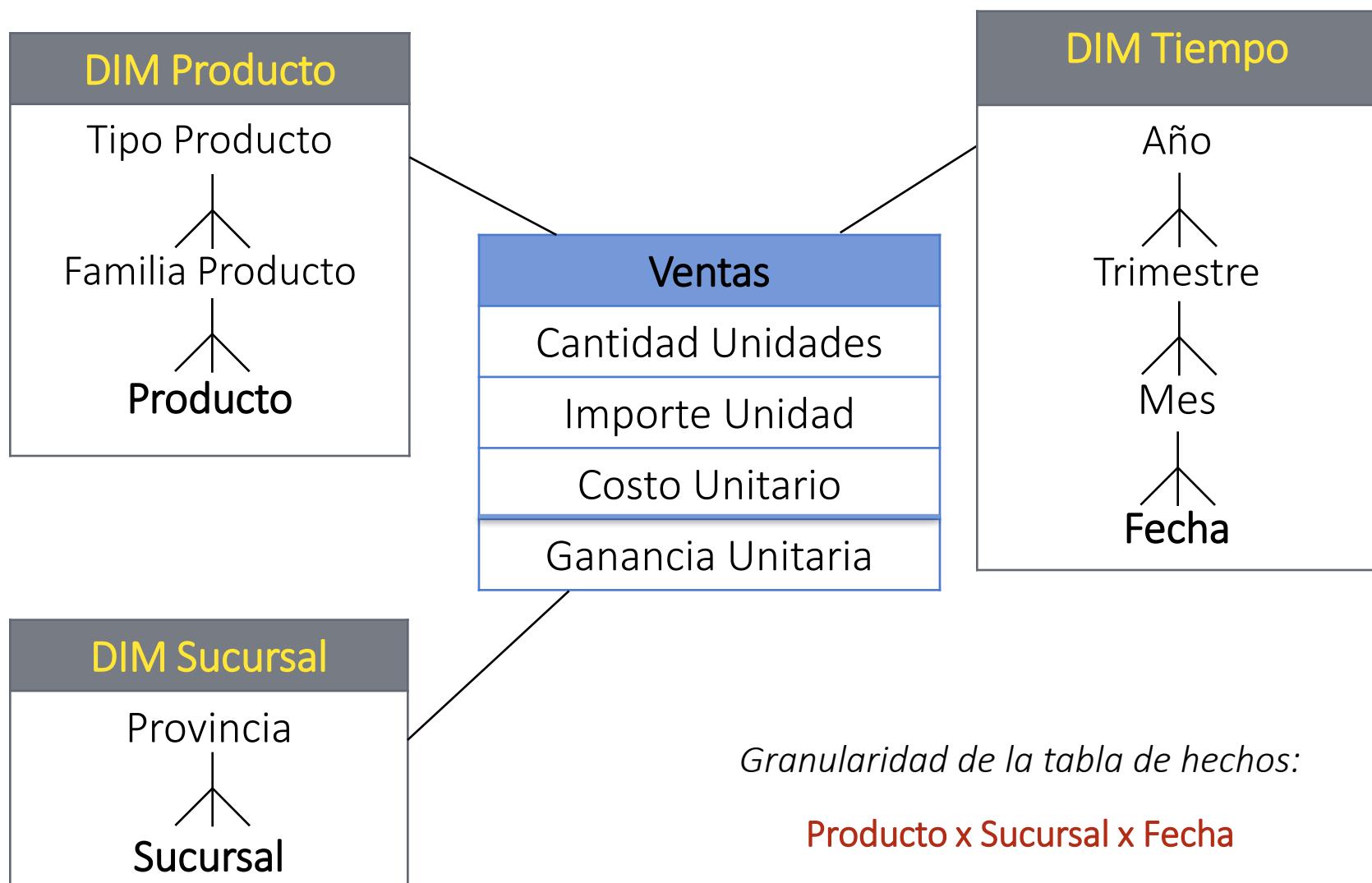
RELACIONES ENTRE ATRIBUTOS

N:N (muchos-a-muchos)

- En el caso de estudio: “El supermercado posee distintas **sucursales** en distintos lugares del país en las que comercializa distintos **productos**.
 - Cada producto se vende en varias sucursales
 - En cada sucursal se venden varios productos
- Los atributos *Sucursal* y *Producto* corresponden a dimensiones diferentes.
- Se relacionan a través la **tabla de hechos** y representan la relación muchos a muchos que existe entre las dimensiones.



MODELO DIMENSIONAL “LÓGICO” (MUESTRA JERARQUÍAS)





CASO DE ESTUDIO - CONCLUSIONES

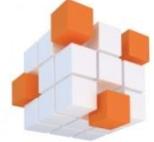
- Al consultar los datos y usar los metadatos con las estructuras jerárquicas de las dimensiones se pueden obtener totales por varias cosas:
 - ✓ Importes totales por tipo de producto, provincia, mes.
 - ✓ Importes totales mensuales por Familia de productos de la provincia de Córdoba o Buenos Aires.
 - ✓ Comparar el total de unidades vendidas de un mismo producto en diferentes sucursales o provincias.
- Recordar que las **dimensiones** sirven para:
 - Mostrar
 - Filtrar
 - Agrupar



CASO DE ESTUDIO - CONCLUSIONES

- Si un usuario nos pregunta:

- ¿Voy a poder obtener los totales trimestrales de los importes de todas las provincias? **SI**
- ¿Voy a poder obtener los totales trimestrales de los impuestos de todas las provincias? **NO**
 - No hay una medida para el valor del impuesto
- ¿Voy a poder obtener los totales mensuales de los descuentos de todas las provincias? **NO**
 - No hay una medida para el descuento
- ¿Voy a poder obtener los totales anuales de los importes de los productos “Yogur A” vendidos en cada sucursal de CABA? **SI**



CASO DE ESTUDIO - CONCLUSIONES

- Si un usuario nos pregunta:

- ¿Voy a poder obtener los totales mensuales de los importes de todas las provincias por forma de pago? **NO**
 - El ticket no contiene información sobre la Forma de Pago.
 - No hay una dimensión Forma de Pago.
- ¿Voy a poder obtener los totales mensuales de los importes de todas las sucursales grandes de todas las provincias? **NO**
 - En este modelo no está la información sobre las sucursales para saber si son grandes o chicas.
 - Se puede resolver fácilmente agregando otra jerarquía para la dimensión Sucursal y teniendo una fuente de datos que traiga esa información.



VENTAJAS DEL MODELO DIMENSIONAL

- Presenta la información de una manera estándar, sencilla y sobre todo intuitiva para los usuarios.
- Resiste cambios inesperados en la conducta del usuario.
- Todas las dimensiones son equivalentes y pueden ser pensadas como puntos de acceso a la tabla de hechos.
- Facilidad para adaptarse a cambios.
 - ❑ Agregar hechos o medidas a la tabla de hechos, siempre que sean consistentes con el mayor nivel de detalle de las dimensiones.
 - ❑ Agregar atributos a las dimensiones.
 - ❑ Agregar nuevas dimensiones.



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

1. Agregar un nuevo hecho no planeado, siempre y cuando sea compatible con la granularidad definida.

Fecha	SKU	Volumen
05/04/2011	Agua 1.5 L	10
05/04/2011	Agua 2.25 L	35
05/04/2011	Cola 1 L	15
06/04/2011	Cola 1.5 L	10



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

- Supongamos que quiero agregar el *importe*:
 - Agrego una columna a la tabla de hechos.
 - Completo los valores para los registros existentes.
 - Con algún valor por defecto.
 - Voy al transaccional a buscar el valor correcto (muy costoso).

Fecha	SKU	Volumen	Importe
05/04/2011	Agua 1.5 L	10	0
05/04/2011	Aqua 2.25 L	35	0
05/04/2011	Cola 1 L	15	0
06/04/2011	Cola 1.5 L	10	0
10/04/2011	Aqua 1.5 L	5	20

- ✓ El nuevo hecho respeta la granularidad ya definida.



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

- Supongamos que trabajo todas bebidas línea Coca Cola y me doy cuenta, que tener la *línea de productos* es importante:
 - ❑ ¿La línea de productos respeta la granularidad de la tabla?

Fecha	SKU	Volumen
05/04/2011	Agua 1.5 L	10
05/04/2011	Agua 2.25 L	35
05/04/2011	Cola 1 L	15
06/04/2011	Cola 1.5 L	10



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

- Si agrego la línea de productos a la tabla, los hechos no varían, entonces respeta la granularidad.

Fecha	Línea	SKU	Volumen
05/04/2011	Coca Cola	Agua 1.5 L	10
05/04/2011	Coca Cola	Agua 2.25 L	35
05/04/2011	Coca Cola	Cola 1 L	15
06/04/2011	Coca Cola	Cola 1.5 L	10

- ✓ Simplemente es agregar una columna y completarla con los valores correspondientes.



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

2. Agregar en una dimensión un atributo que no había sido planteado.
 - ❑ Supongamos que tengo la dimensión País y quiero agregar la densidad poblacional:

DIM PAÍS

Id_pais	Desc_pais
1	Argentina
2	Chile
3	Perú



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

- Es un atributo directamente relacionado con la dimensión.
- No tiene ningún impacto en la tabla de hechos.
- Se agrega un nuevo campo a la dimensión con el nuevo valor.

DIM PAÍS

Id_pais	Desc_pais	Desc_densi
1	Argentina	1000
2	Chile	700
3	Perú	500



VENTAJAS DEL MODELO DIMENSIONAL

Facilidad para adaptarse a cambios

3. Romper una dimensión existente llevándola a un nivel de granularidad menor, a partir de un momento dado.
 - ❑ Quiero quedarme con menos nivel de detalle.
 - ❑ Se puede ver que en la tabla de hechos la dimensión Tiempo está a nivel de fecha.
 - ❑ Dejar sólo el nivel superior de la dimensión (ej. mes)

Id_mes	Desc_mes
04	Abril
Fecha	Id_mes
05/04/2011	04
05/04/2011	04

↓

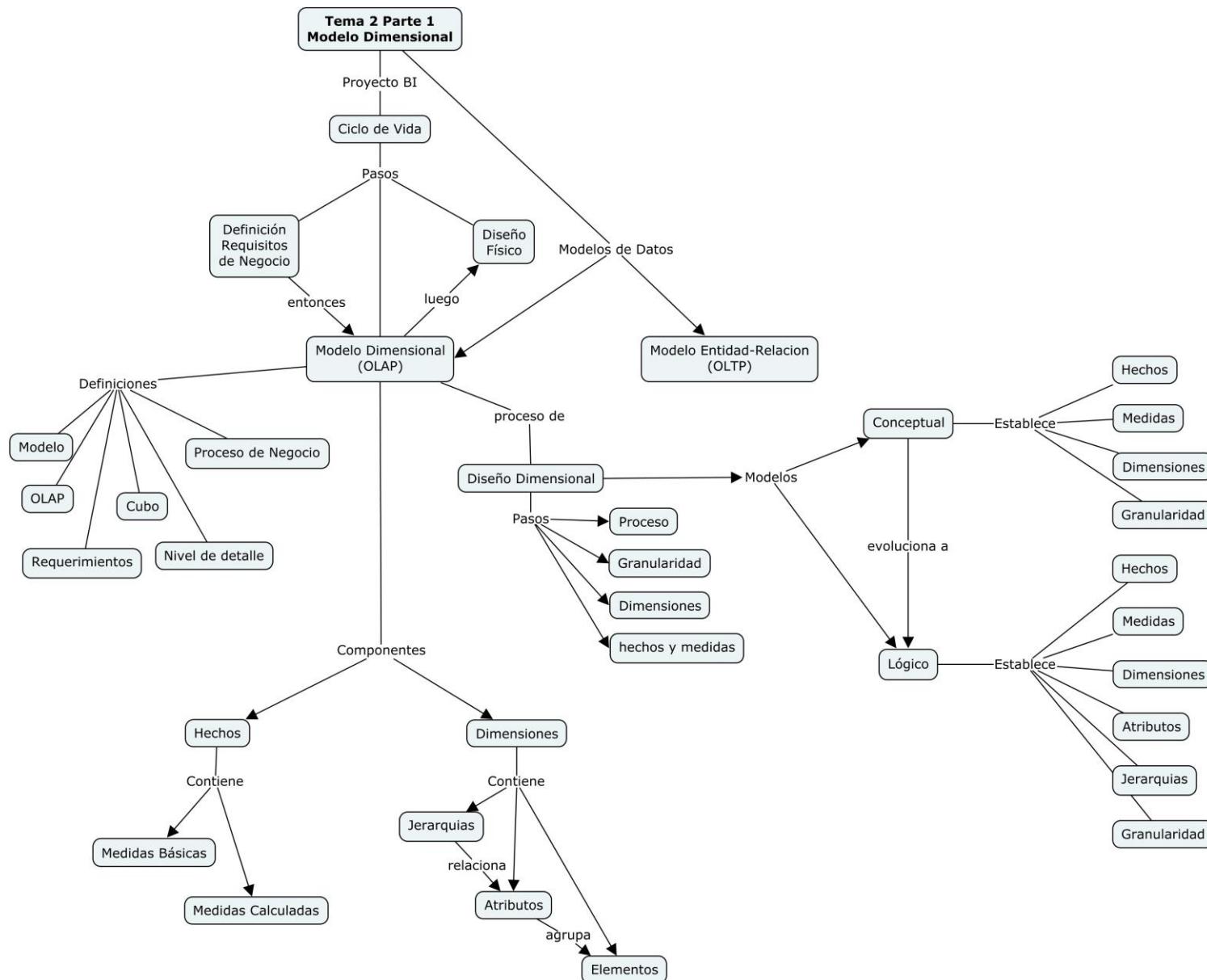
Dimensión Tiempo



Group by MES		SUM
Fecha	SKU	Volumen
05/04/2011	Agua 1.5 L	10
05/04/2011	Agua 2.25 L	35
05/04/2011	Cola 1 L	15
06/05/2011	Cola 1.5 L	10



RESUMEN TEMA 2 PARTE 1





TP CASO ESTUDIO: DORSA

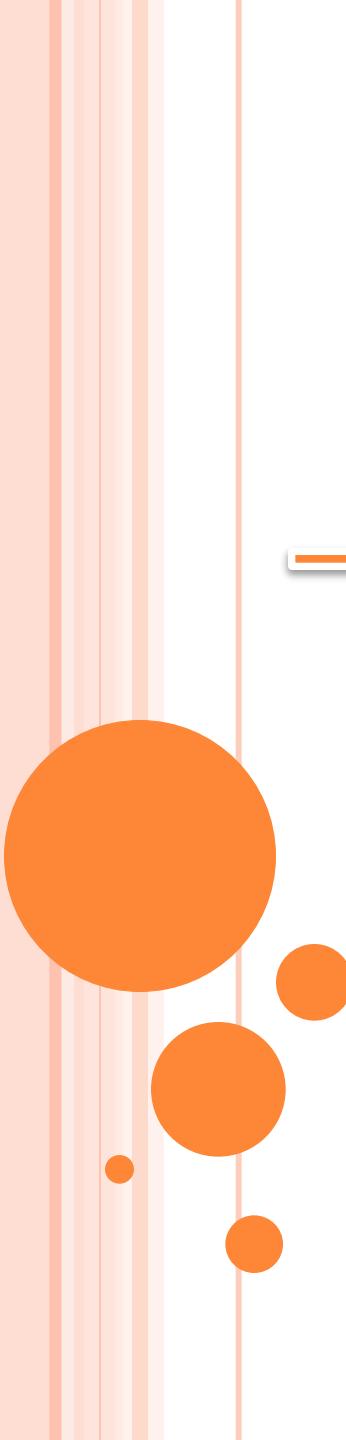
ENTREGA OPCIONAL (*)

- Listar las posibles dimensiones a partir de las necesidades de información enunciadas.
- Identificar los hechos con las medidas básicas y calculadas.
- Identificar el proceso de negocio a modelar.
- Construir el modelo dimensional conceptual.
- Identificar atributos y jerarquías para cada dimensión.
- Construir el modelo dimensional lógico.

(*)

- Los TPs Opcionales sirven para afianzar conceptos necesarios para realizar los TPs de Aplicación, serán corregidos en clase y/o mediante Autoevaluación.
- Los TPs de Aplicación tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción, serán corregidos en detalle por los docentes.
- Ver documentos: "Condiciones de Cursada en MIeL - sección: Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio" + "Círculo Entrega TPs Teams/MIeL"





UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

Modelo Dimensional – Parte 2

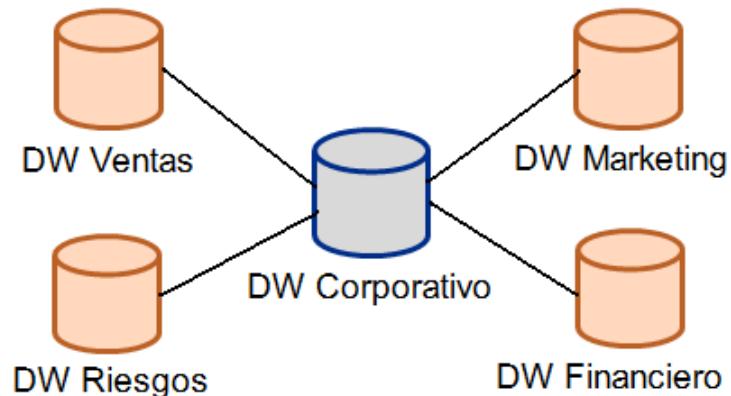
Docentes: ING. LORENA R. MATTEO

Autores ppt orig.: Lic. HUGO M. CASTRO / MG. DIEGO BASSO



DATA MART

- Subconjunto de los datos del DW con el objetivo de responder a un determinado análisis, función o necesidad, con un conjunto específico de usuarios.
- Cubre las necesidades de un grupo de trabajo o de una determinada área en una organización.
- El alcance es lo que diferencia un Data Mart de un Data Warehouse.
 - Independiente: Se genera a través de un proceso de ETL
 - Dependiente : Se toma del Data Warehouse



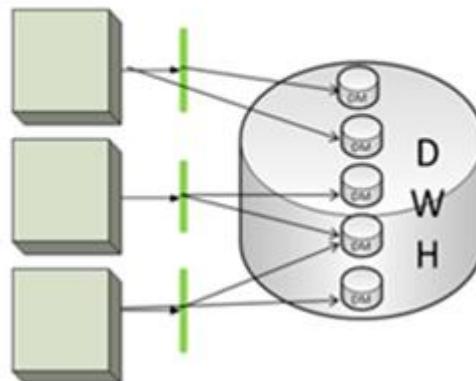


PARADIGMAS DE CONSTRUCCIÓN

○ Ralph Kimball

- Construir los Data Mart de acuerdo a las prioridades del negocio. El DW es la unión de todos los Data Mart de la organización.
- Ventajas
 - Soluciones más pequeñas y más fáciles de construir
 - Soluciones más fáciles de entender
- Desventajas
 - El mantenimiento es más complejo
 - Pueden existir problemas de sincronización

Bottom Up Ralph Kimball

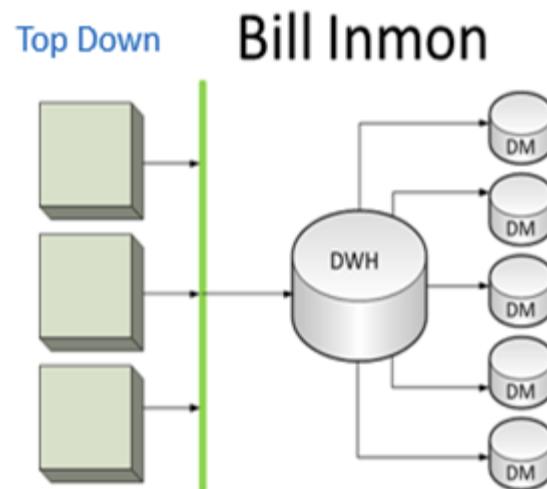




PARADIGMAS DE CONSTRUCCIÓN

○ Bill Inmon

- Construir el DW centralizado y luego extraer la información de negocio para cada Data Mart.
- Ventaja
 - Estructura de los datos del DW en 3FN, normalizada y limpia.
- Desventajas
 - Mayor tiempo de construcción
 - Costo muy alto





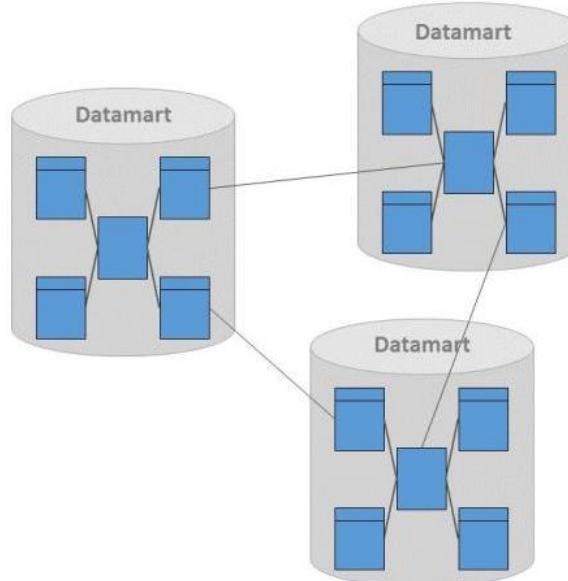
DIMENSIÓN CONFORMADA

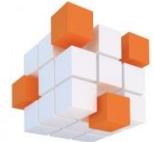
- Es una dimensión que tiene el mismo significado para todos los Data Mart que se armen.
 - Ejemplo: La **dimensión cliente** podría ser compartida por varios Data Mart (ventas, compras, pedidos, cobranzas, etc.)
- Ventajas
 - Evita duplicar dimensiones para distintos Data Mart.
 - Menor costo de mantenimiento. Los datos de la dimensión están organizados de la misma manera en todos los cubos.
- Desventajas
 - Deben emplearse del mismo modo en todos los cubos que usan estas dimensiones.
 - Un cambio implica modificar la dimensión en todos los cubos.



DIMENSIÓN CONFORMADA

- Es responsabilidad del equipo que diseña el DW establecer, publicar y mantener las dimensiones conformadas.
- Cualquier modificación o cambio en una dimensión conformada debe ser válida para todos los cubos que la empleen.
 - El DW debe funcionar como un todo integrado.





MÉTODO DE LA MATRIZ (Bus Matrix)

- Identificar todos los Data Mart y dimensiones.
- Armar una matriz e interceptar los **hechos** (fact) con las **dimensiones** que correspondan.
- Las dimensiones que se relacionen con varios hechos probablemente sean conformadas.

	Tiempo	Llamada	Documento	Geografía	Centro de Distribución
Reclamos	X	X		X	
Facturación	X		X	X	X
....					
Stock				X	X

- El nivel mínimo de granularidad de cada dimensión conformada debe ser el mismo.



SISTEMAS OLAP

- On Line Analytical Process
- Es una **tecnología** para el acceso y análisis de datos en línea, en el que se usan **herramientas analíticas**, que facilitan el **análisis de la información** del negocio.
- Es el proceso de almacenar y administrar datos sobre la base de las variables del negocio (dimensiones), para permitir a los profesionales de negocios visualizarlos y analizarlos para entender cuál es su significado.



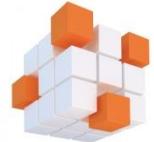
HERRAMIENTAS DE EXPLOTACIÓN DW

- El usuario formula **consultas** a la herramienta OLAP seleccionando **atributos** del modelo multidimensional, sin conocer la estructura interna del DW.
- La herramienta OLAP genera la correspondiente consulta y la envía al gestor de consultas del sistema.
 - Mediante una sentencia **SELECT**



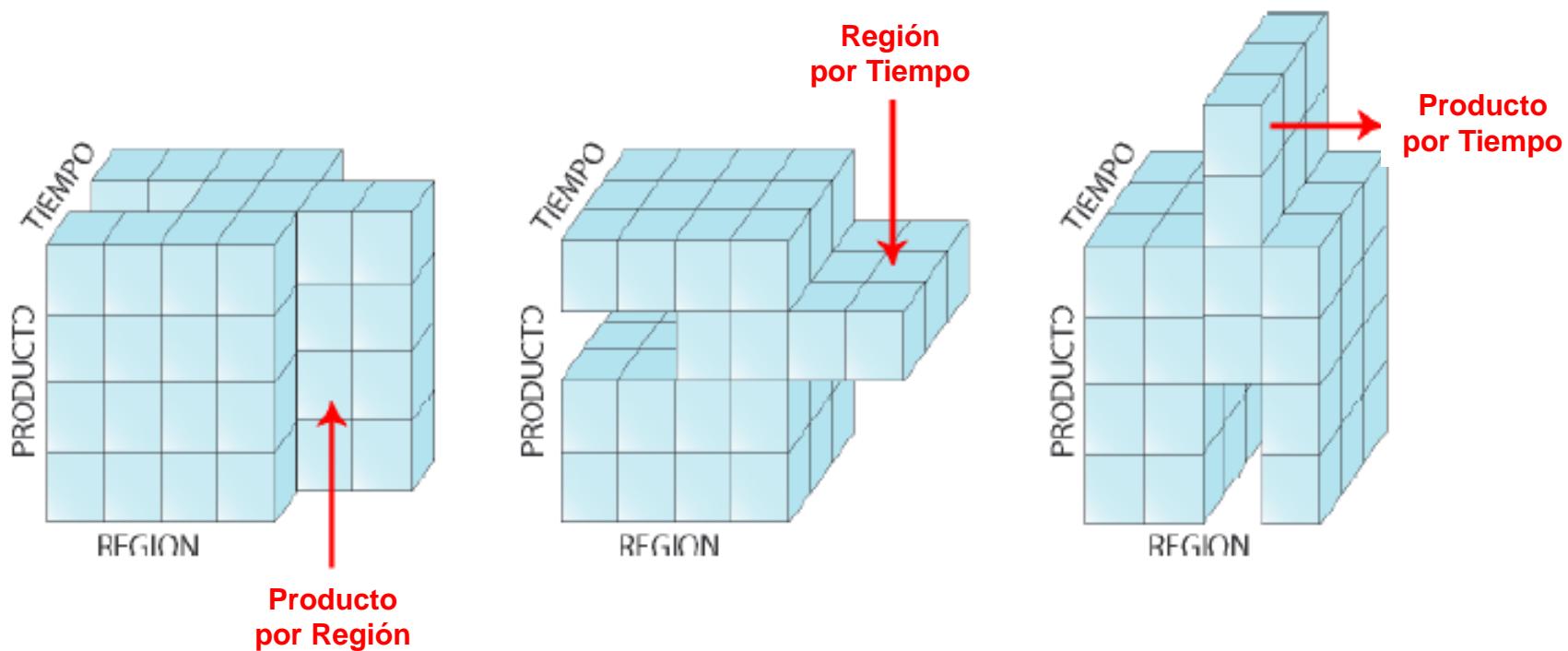
HERRAMIENTAS DE EXPLOTACIÓN DW

- Una consulta a un DW consiste en obtener **medidas** sobre los hechos, parametrizados por **atributos** de las **dimensiones** y restringidas por **condiciones** aplicadas sobre las dimensiones.
 - **Importe total de ventas** durante el año 2020 de los **productos** del área de limpieza, por **trimestre** y **categoría**.
 - *Medida:* **Importe Total Venta**
 - *Restricciones:* Productos del área de limpieza, ventas durante el Año 2020.
 - *Parámetros de la consulta:* Por **Categoría de Producto** y por **Trimestre**.



HERRAMIENTAS DE EXPLOTACIÓN DW

- Se pueden tomar rebanadas del cubo para responder a una diversidad de preguntas.

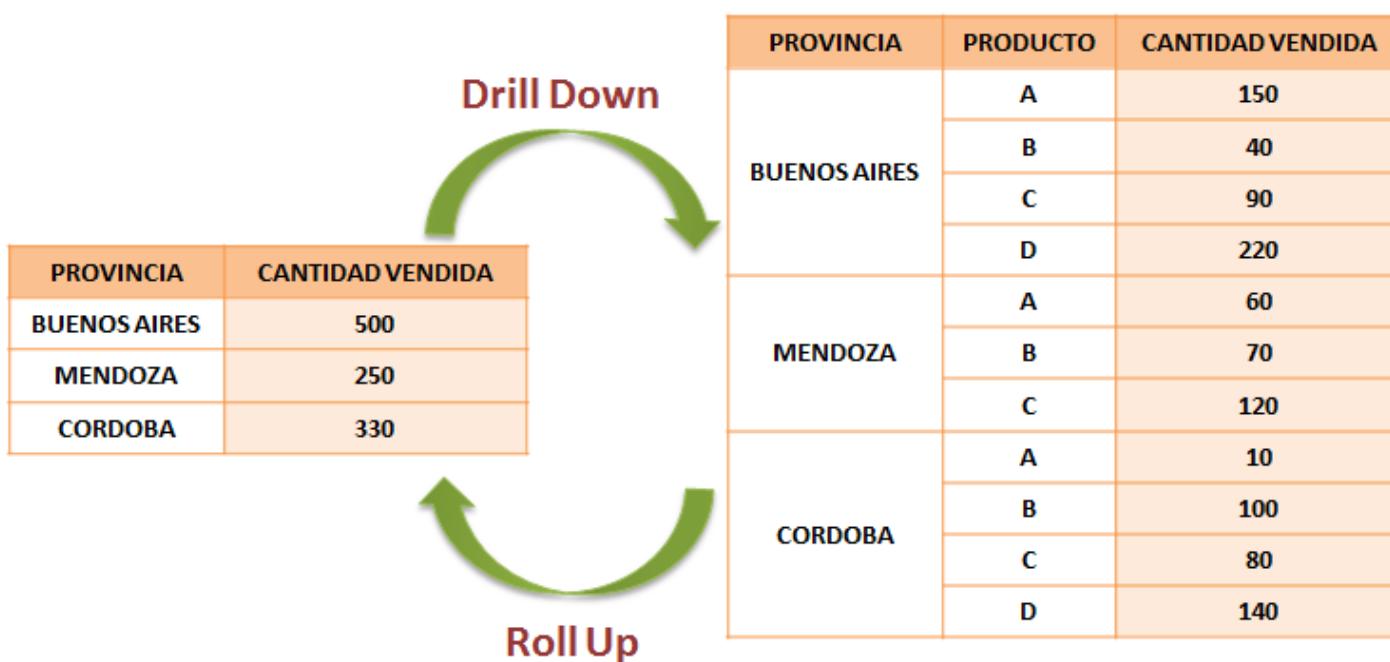




OPERACIONES OLAP

Relacionadas con la granularidad

- Drill Down (Roll Down): Mayor nivel de detalle
 - De datos sumados (o agregados) a datos desagregados
- Roll Up (Drill Up): Menor nivel de detalle
 - Mayor consolidación de datos



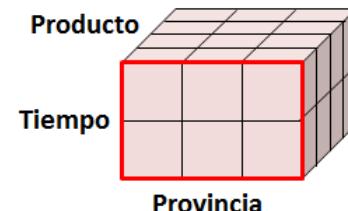
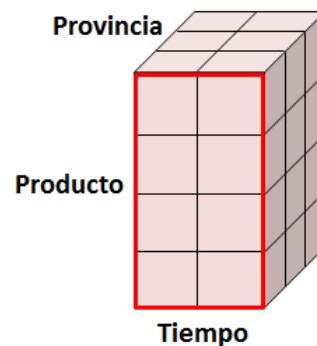
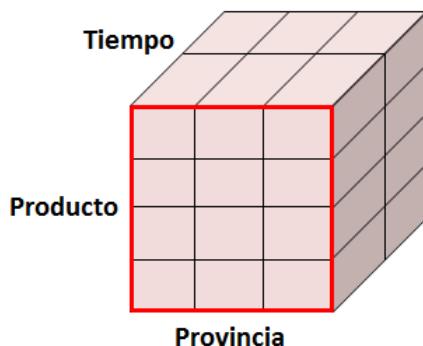


OPERACIONES OLAP

Navegación por las dimensiones

○ Slice and Dice

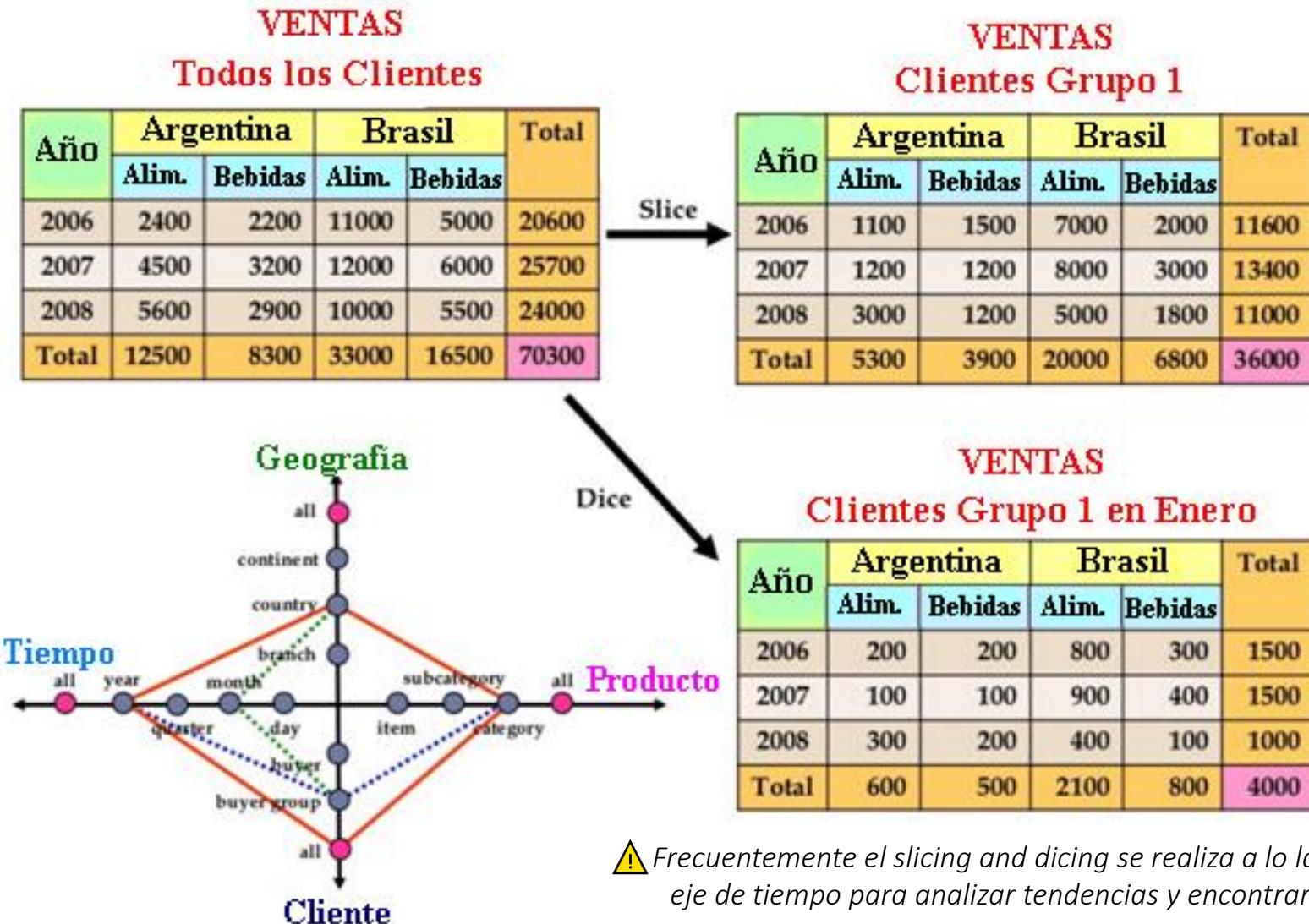
- “Segmentar y Desglosar” en terminología BI .
- Selección de valores específicos para una o varias dimensiones, cortes o subconjuntos de las vistas.
- Cambio en la orientación y forma de visualización del cubo.
- Se puede agrupar la información de una manera, analizarla, agruparla de otra y realizar otro análisis.
 - Provincia = Mendoza (filtro estático)
 - Los 2 productos más vendidos del año 2020 (filtro dinámico)
- El usuario posee diferentes perspectivas para el análisis.

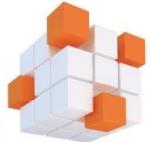




OPERACIONES OLAP

○ Slice and Dice – Ej.1



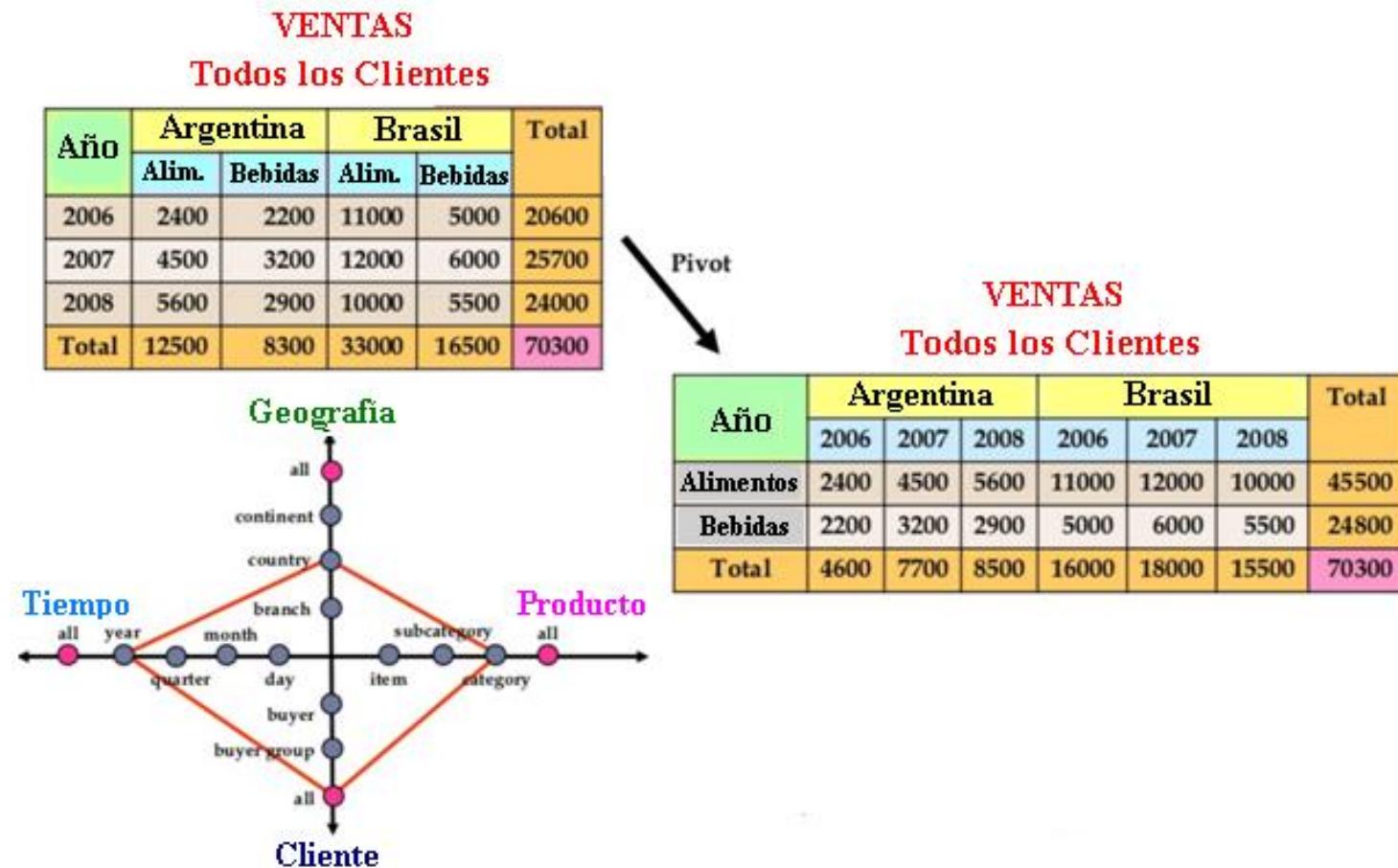


OPERACIONES OLAP

Navegación por las dimensiones

○ Pivot

- Reorientación de dimensiones.

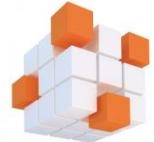




OTRAS OPERACIONES OLAP

○ Drill Across: Conexión con otras tablas de hechos

- A través de al menos una dimensión en común del cubo.
- Ejemplo: Información de Ventas: **fecha/sucursal/producto**
 - Al analizar el cubo se observan ventas sospechosas en la sucursal XX el día DD/MM/YYYY.
 - Se quiere ver quiénes eran las personas que trabajaron en esa sucursal y en esa fecha.
 - Puede existir también información de Asistencia de empleados: **fecha/sucursal/empleado**



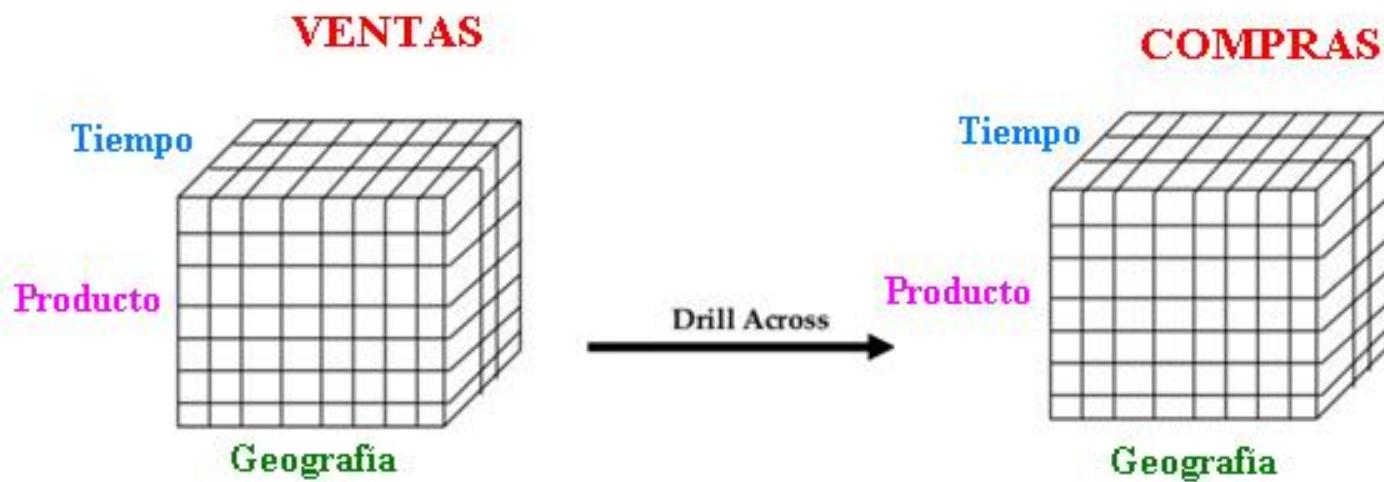
OTRAS OPERACIONES OLAP

○ Drill Across

VENTAS Todos los Clientes						
Año	Argentina		Brasil		Total	
	Alim.	Bebidas	Alim.	Bebidas		
2006	2400	2200	11000	5000	20600	
2007	4500	3200	12000	6000	25700	
2008	5600	2900	10000	5500	24000	
Total	12500	8300	33000	16500	70300	

Drill Across →

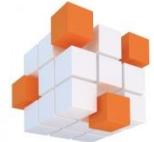
COMPRAS Todos los Clientes						
Time	Argentina		Brasil		Total	
	Alim.	Bebidas	Alim.	Bebidas		
2006	1500	1500	6000	3000	12000	
2007	3500	2500	8000	4000	18000	
2008	4000	1500	5000	3000	13500	
Total	9000	5500	19000	10000	43500	



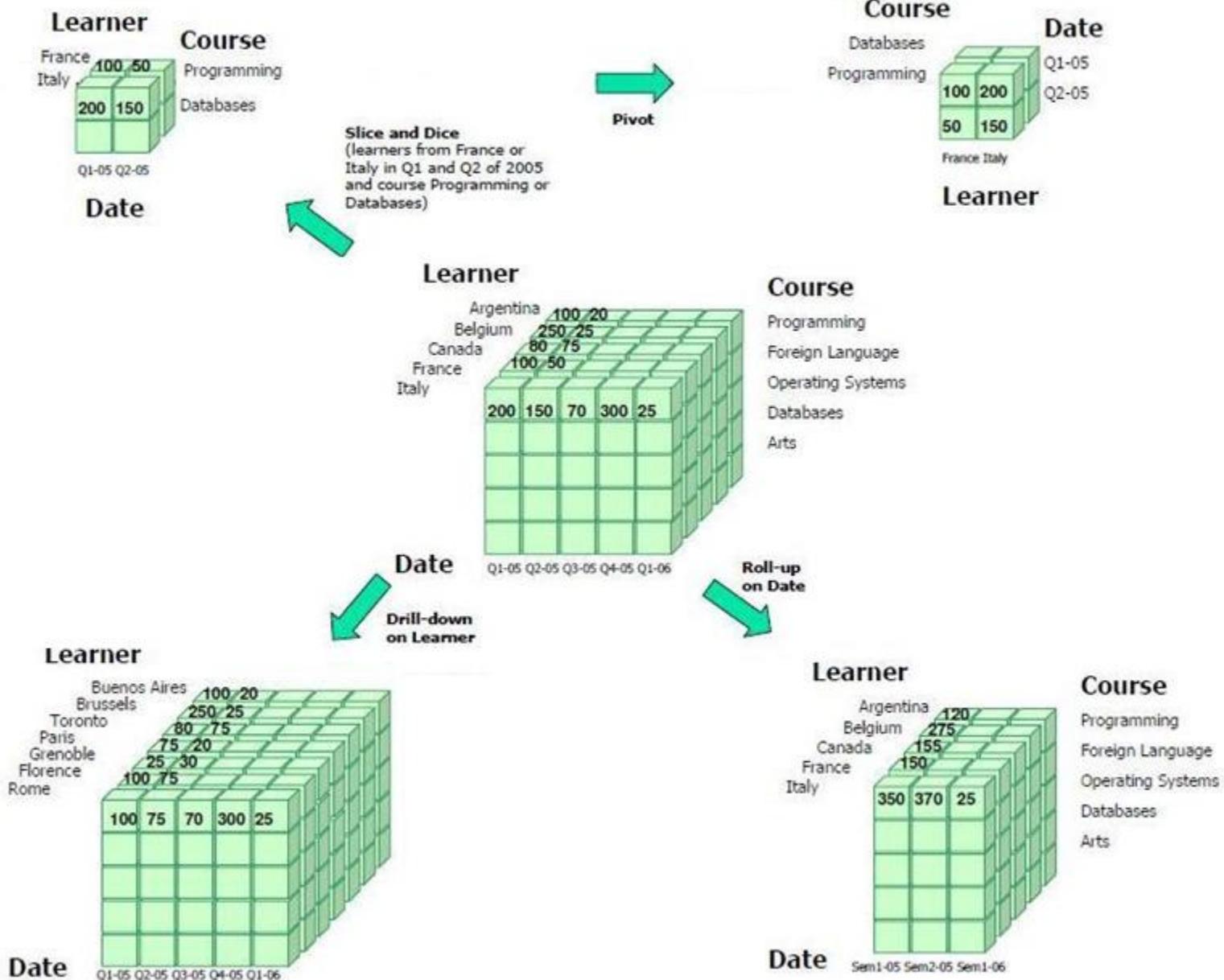


OTRAS OPERACIONES OLAP

- Drill Through/Drill Out: Conexión con tablas externas.
 - Desde el nivel más bajo de detalle del cubo se genera una consulta SQL a las tablas relacionales de donde se sacaron los datos, para llegar al nivel atómico.
 - Operación costosa. No debería ser lo habitual sino la excepción.
 - Si es habitual, está mal definido el cubo y hay que agregarle dimensiones para llegar al nivel atómico.



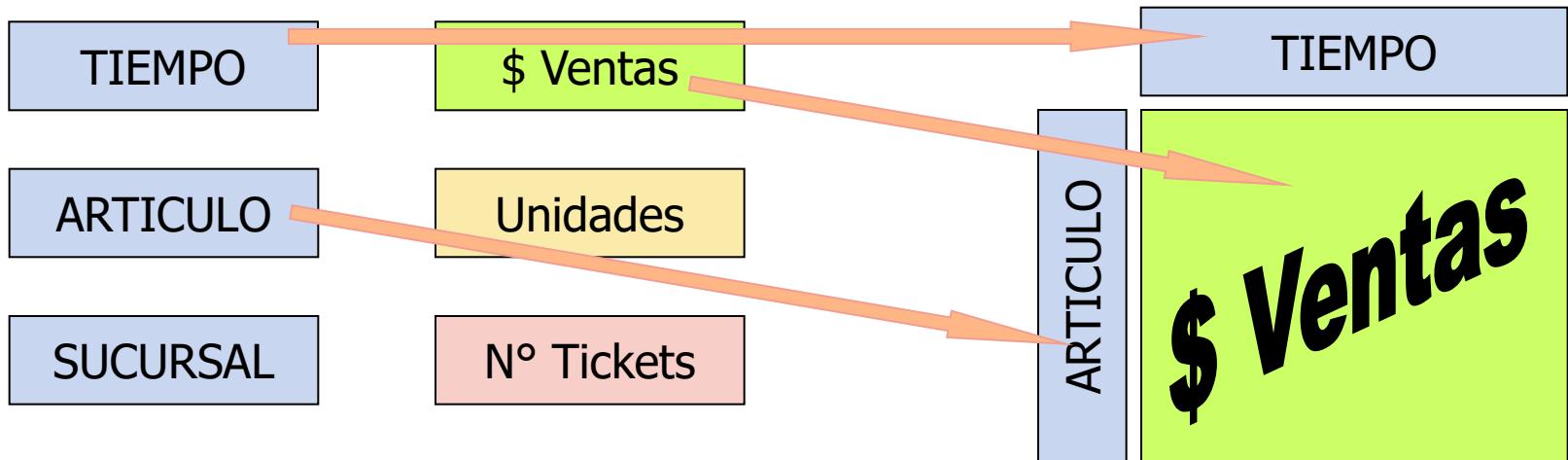
RESUMEN OPERACIONES OLAP



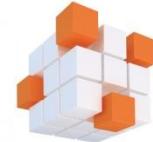


ANÁLISIS OLAP

DIMENSIONES MEDIDAS

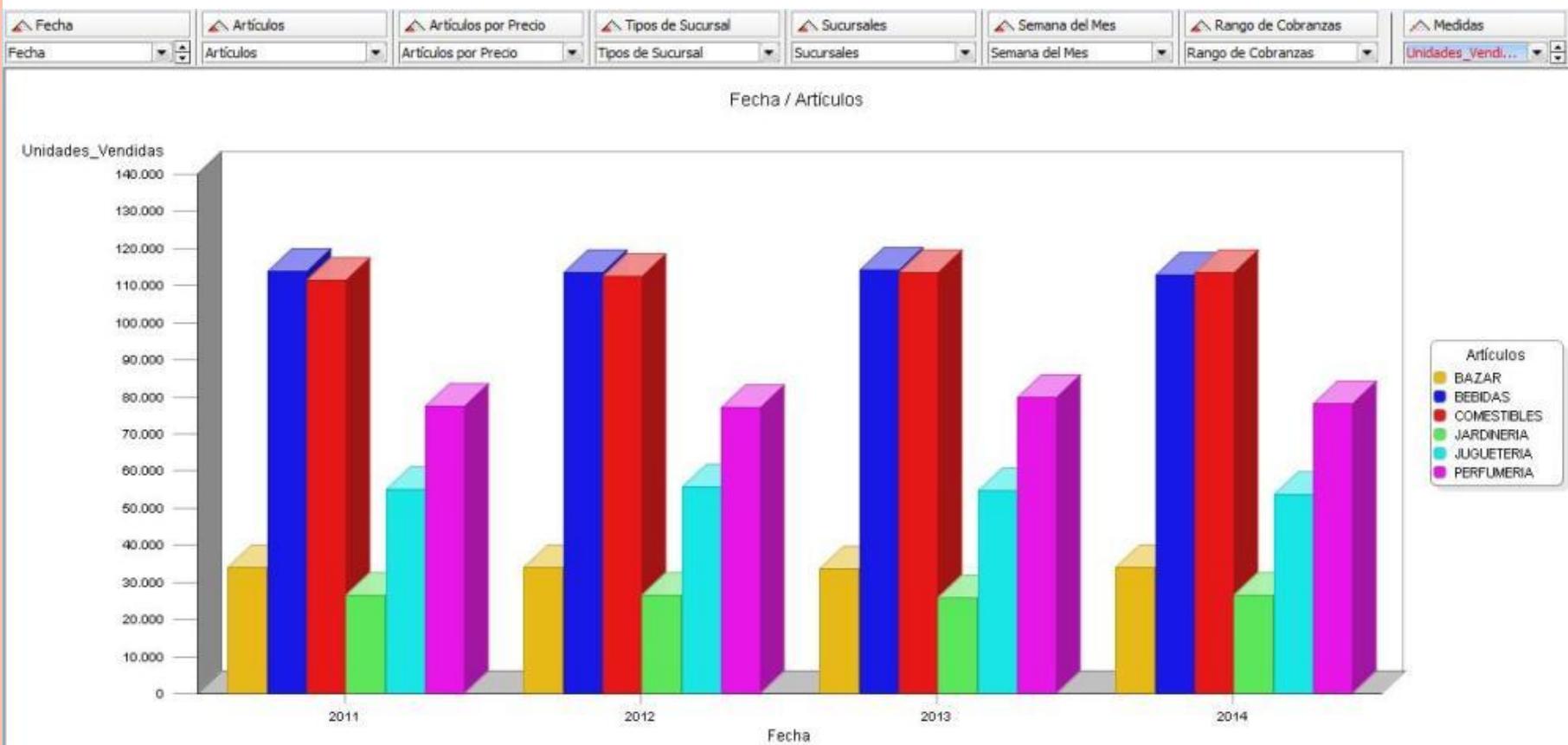


- Se elige la dimensión para las filas.
- Se elige la dimensión para las columnas.
- Se elige la medida a representar.

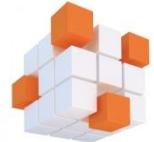


ANÁLISIS OLAP

- Unidades vendidas por rubro artículos entre 2011 y 2014



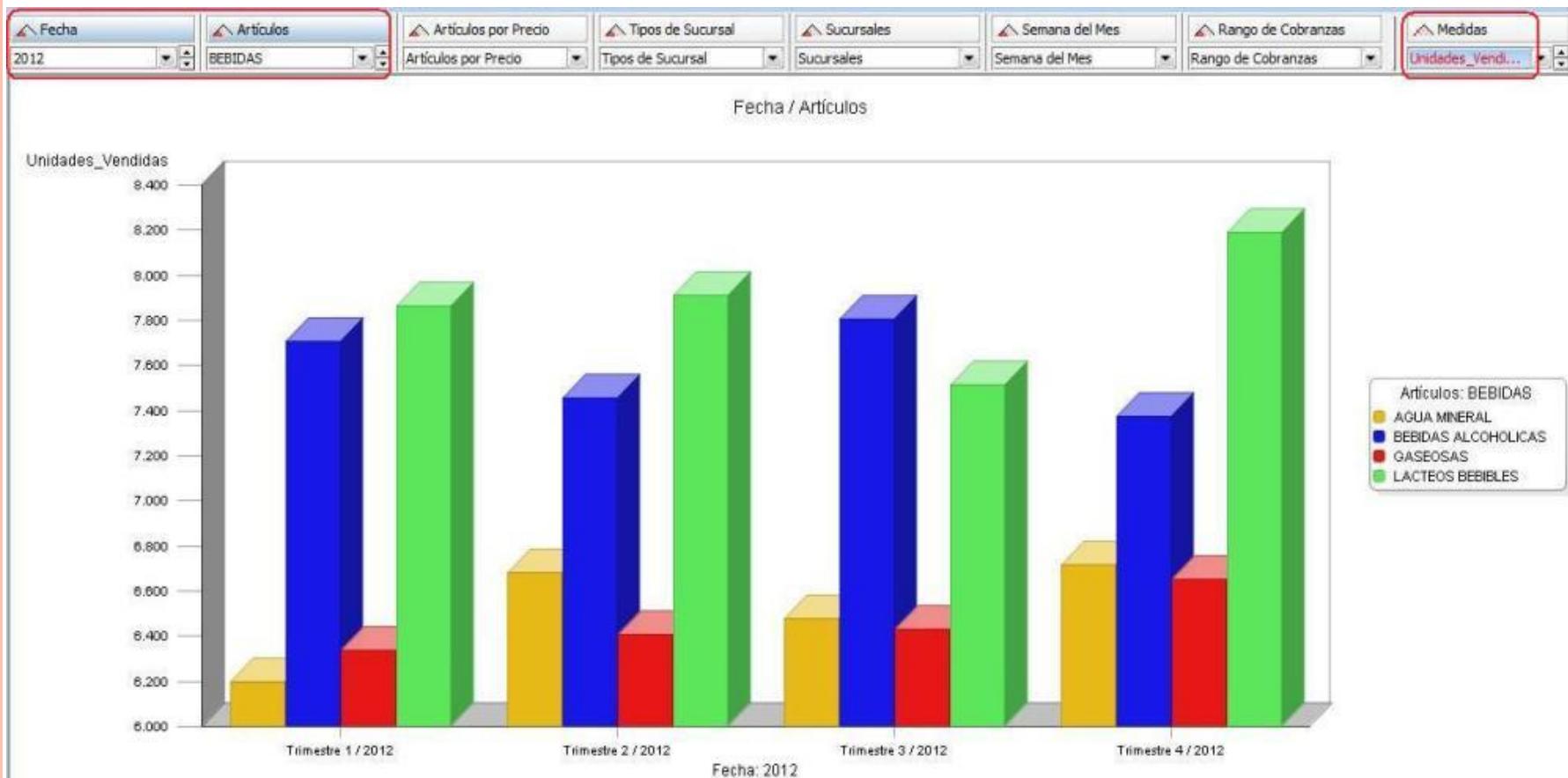
22/4/2025 IN2025

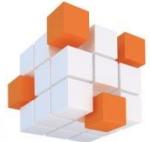


ANÁLISIS OLAP

- Unidades vendidas por artículos entre 2011 y 2014
 - Drill Down: Año 2012 y artículos de Bebidas

22/4/2025 IN2025



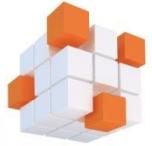


ANÁLISIS OLAP

- Ventas del 1º trimestre de 2014 por rubro de artículo

Fecha -> 2014 -> Trimestre 1 / 2014 / Articulos

Artículos	Fecha		
	Mes	ene / 2014	feb / 2014
BAZAR	43.826,56	39.569,17	44.305,97
BEBIDAS	34.415,11	29.392,49	36.281,93
COMESTIBLES	48.717,89	43.336,45	44.770,04
JARDINERIA	4.709,52	3.828,19	4.324,66
JUGUETERIA	64.339,10	58.069,97	68.276,26
PERFUMERIA	41.744,60	36.401,13	40.601,94



ANÁLISIS OLAP

- Ventas del 1º trimestre de 2014 por rubro de artículo
 - Agregamos Sucursales

Fecha -> 2014 -> Trimestre 1 / 2014 / Artículos, Sucursales

		Fecha			
Artículos	Sucursales	Mes	ene / 2014	feb / 2014	mar / 2014
BAZAR	Avellaneda		7.415,04	7.888,42	8.392,17
	Capital		14.510,94	12.763,32	13.740,90
	Córdoba		3.132,55	3.346,92	3.519,12
	Hurlingham		2.964,38	2.882,54	3.815,02
	Mar del Plata		8.054,26	7.210,76	6.806,98
	Rosario		4.411,15	2.419,55	4.071,08
	San Isidro		3.338,24	3.057,69	3.960,72
BEBIDAS	Avellaneda		5.422,68	5.279,61	6.360,58
	Capital		10.423,60	10.192,80	11.529,99
	Córdoba		3.255,62	2.615,63	2.626,01
	Hurlingham		2.486,82	2.400,79	2.819,45
	Mar del Plata		7.710,47	4.922,80	7.293,84
	Rosario		2.740,09	2.293,69	2.499,22
	San Isidro		2.375,84	1.687,17	3.152,86
COMESTIB...	Avellaneda		8.610,71	6.506,59	6.844,95
	Capital		17.413,63	15.611,91	14.760,34
	Córdoba		3.859,65	2.936,63	4.023,49
	Hurlingham		3.378,07	3.547,02	2.930,89
	Mar del Plata		8.481,91	8.426,26	9.584,42
	Rosario		3.441,46	3.198,11	3.684,04
	San Isidro		3.532,45	3.109,92	2.941,90



ANÁLISIS OLAP

- Ventas de art. de perfumería por nivel de precios y año.
 - Operación Slice and Dice

The screenshot shows a business intelligence interface with a toolbar at the top containing six filter dropdowns: Fecha, Artículos, Artículos por Precio, Tipos de Sucursal, Sucursales, and Semana del Mes. The 'Artículos' and 'Artículos por Precio' dropdowns are highlighted with a red border, and 'PERFUMERIA' is selected in both. Below the toolbar, a message reads 'Artículos por Precio, Fecha / Artículos -> PERFUMERIA'. The main area displays a pivot table with the following data:

Artículos	Artículos por Precio		Fecha														
	Familia	Nivel de Precio	Año	Alto				Medio				Bajo					
		2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014
CHAMPU		111.628,20	135.624,72	159.727,95	175.594,85	49.200,90	60.669,84	70.363,76	78.889,27	0,00	0,00	0,00	0,00				
DESODORANTES PERSONALES		0,00	0,00	0,00	0,00	47.683,30	56.486,52	65.602,04	76.536,06	0,00	0,00	0,00	0,00				
JABONES		0,00	0,00	0,00	0,00	10.384,40	11.871,60	13.562,64	15.930,59	23.546,30	27.570,60	32.465,07	35.130,91				
PERFUMES		43.101,90	52.090,44	60.121,85	68.696,93	23.260,50	27.064,80	29.767,50	36.095,62	0,00	0,00	0,00	0,00				

- Operación Pivot

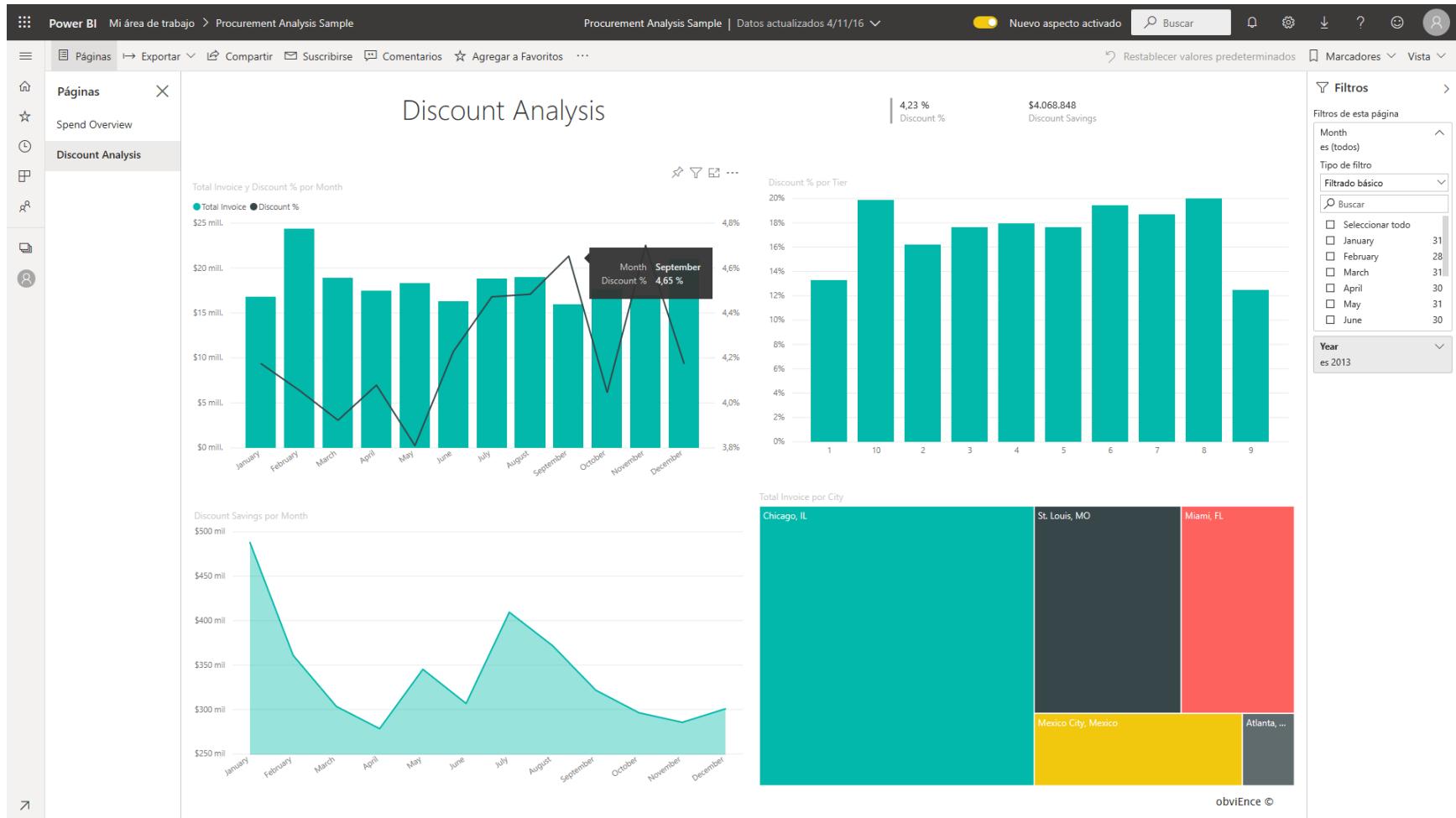
The screenshot shows a business intelligence interface with a toolbar at the top containing five filter dropdowns: Fecha, Artículos, Artículos por Precio, Tipos de Sucursal, and Tipos de Sucursal. The 'Artículos' and 'Artículos por Precio' dropdowns are highlighted with a red border, and 'PERFUMERIA' is selected in both. Below the toolbar, a message reads 'Fecha / Artículos -> PERFUMERIA, Artículos por Precio'. The main area displays a pivot table with the following data:

Artículos	Artículos por Precio		Fecha				
	Familia	Nivel de Precio	Año	2011	2012	2013	2014
		Alto	Medio	Medio	Medio	Medio	Medio
CHAMPU		111.628,20	135.624,72	159.727,95	175.594,85		
DESODORANTES PERSONALES		47.683,30	56.486,52	65.602,03	76.536,05		
JABONES		10.384,40	11.871,60	13.562,64	15.930,59		
PERFUMES		43.101,90	52.090,44	60.121,85	68.696,93		
		23.260,50	27.064,80	29.767,50	36.095,62		



ANÁLISIS OLAP

○ Análisis de Descuentos (Filtros y Vistas)



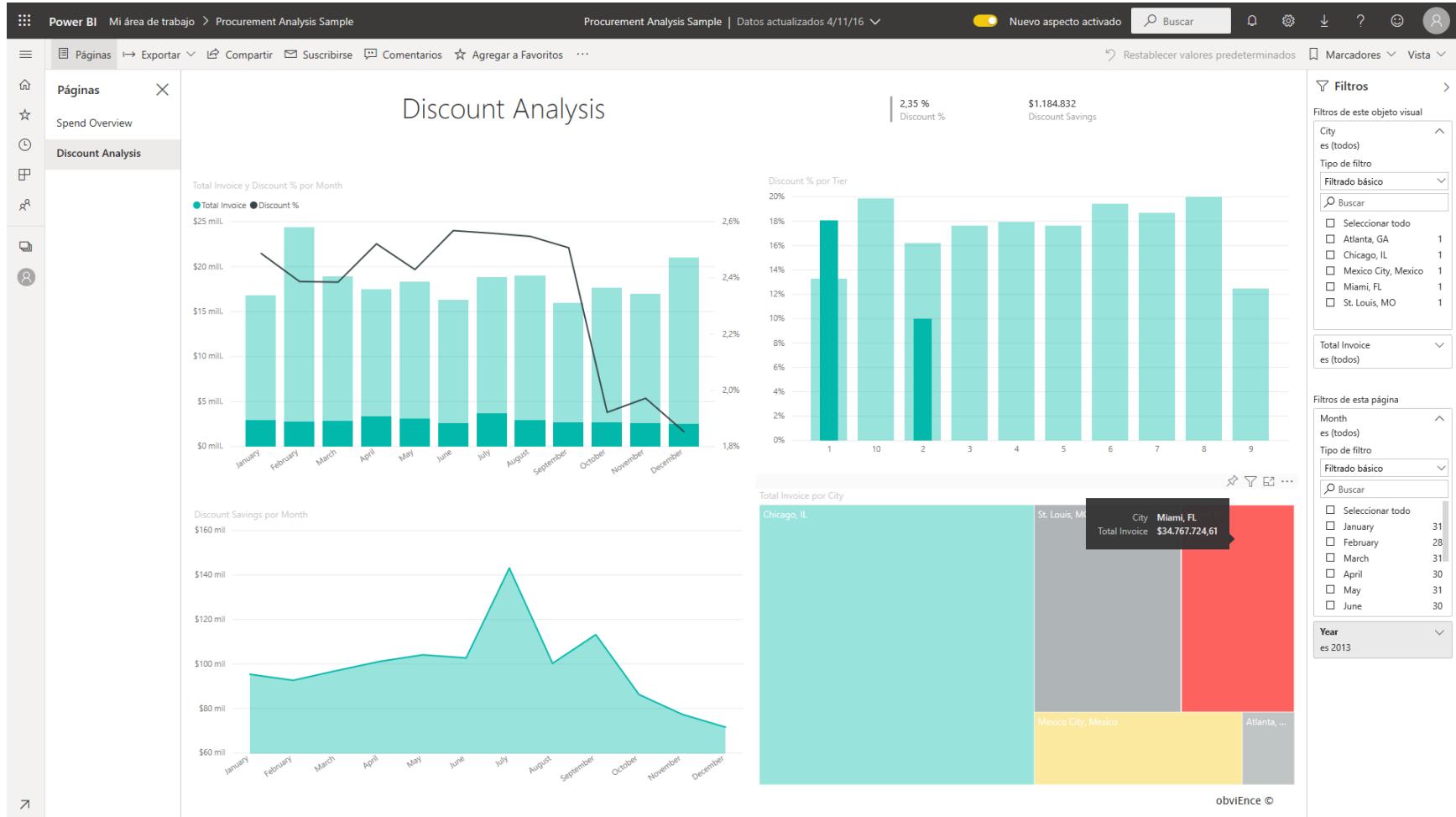
22/4/2025 IN2025

29



ANÁLISIS OLAP

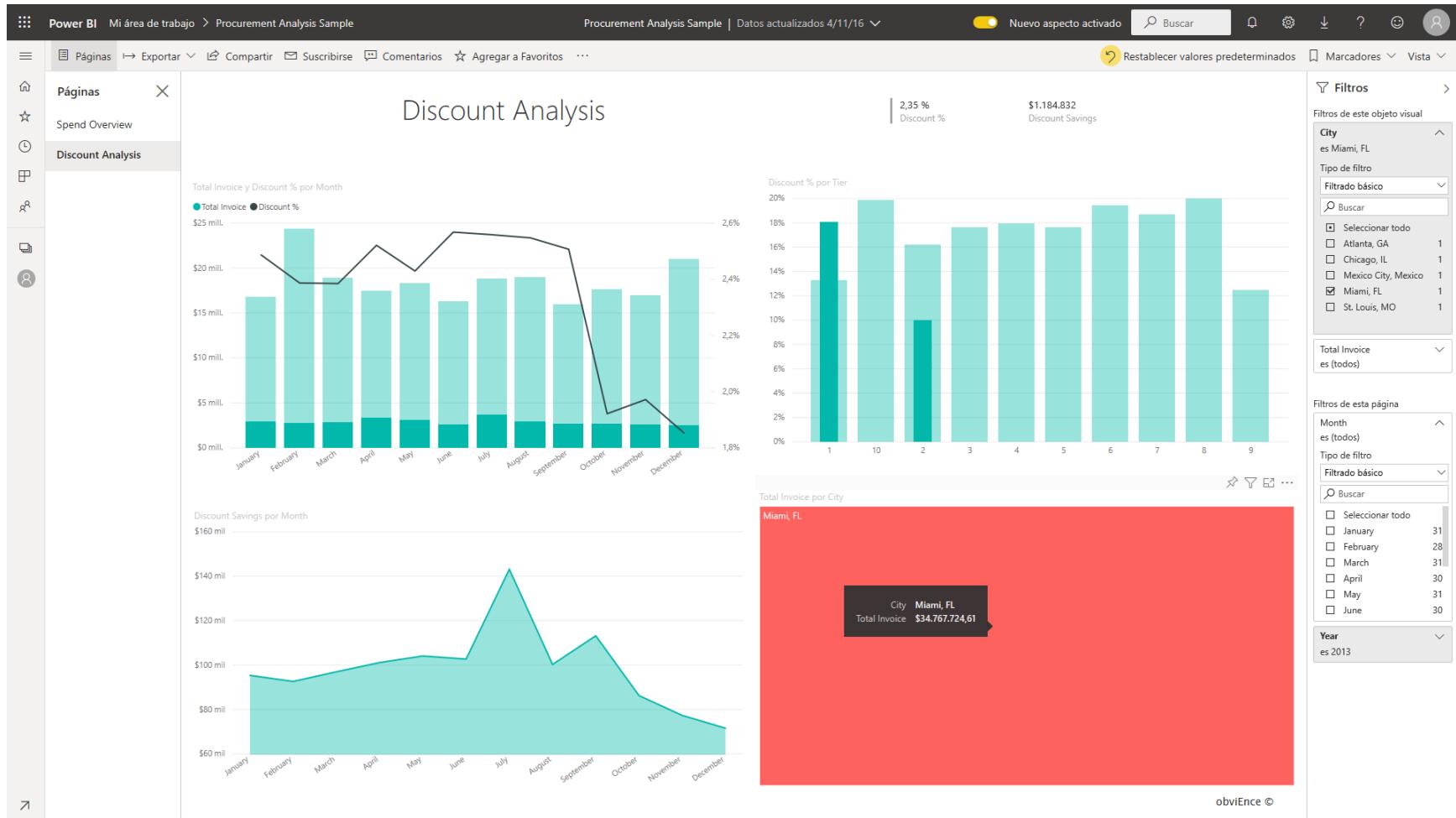
○ Análisis de Descuentos (Vista relativa a Miami, FL)





ANÁLISIS OLAP

○ Análisis de Descuentos (Filtro Miami, FL)





ANÁLISIS OLAP

○ Resumen de Gastos (Filtros Avanzados)

Power BI Mi área de trabajo > Procurement Analysis Sample | Datos actualizados 4/11/16 ▾ Nuevo aspecto activado Buscar Restablecer valores predeterminados Marcadores Vista

Páginas X

Spend Overview

Discount Analysis

Total Invoice por Month y Category

Category: Direct (azul), Indirect (negro), Logistics (rojo), Other (amarillo)

Total Invoice por Country/Region

ESTADOS UNIDOS

MÉXICO

Golfo de México

Nassau

La Habana CUBA

George Town HAITI

Cockburn Town JAMAICA

Ciudad de

Bing

Filtros

Year es 2013

es

no es

esta en blanco

no está en blanco

Aplicar filtro

Total Invoice por Sub Category

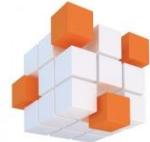
Categoría	Valor
Contracting & Services	\$2 mill.
Hardware	\$70 mill.
Software	\$62 mill.
Consulting	\$43 mill.
Travel	\$15 mill.
Other	\$21 mill.

Total Invoice por Tier

Tier	Valor
1	\$188 mill.
2	\$20 mill.
3	\$7 mill.
4	\$4 mill.
5	\$2 mill.
6	\$1 mill.
7	\$1 mill.
8	\$0 mill.
9	\$0 mill.
10	(\$1 mill.)

Presione F11 para salir del modo de pantalla completa.

obviEnce ©



○ Resumen de Gastos

Drill Down: Total Factura por Commodity y Ciudad

Power BI Mi área de trabajo > Procurement Analysis Sample

Procurement Analysis Sample | Datos actualizados 4/11/16

Nuevo aspecto activado

Buscar

Páginas Exportar Compartir Suscribirse Comentarios Agregar a Favoritos ... Restablecer valores predeterminados Marcadores Vista

Spend Overview

Total Invoice por Month y Category

Category: Direct (Blue), Indirect (Black), Logistics (Red), Other (Yellow)

Total Invoice por City

Total Invoice por Commodity

Total Invoice por Tier

Filtros

Filtros de este objeto visual

Commodity es (todos)

Tipo de filtro

Filtrado básico

Buscar

□ Seleccionar todo

□ Chemicals 5

□ Contingent Labor 24

□ Contracting & Servi... 41

□ EHS (Env, Health, ... 20

□ Electrical 230

□ Engineering Services 37

Commodity Detail es (todos)

Tipo de filtro

Filtrado básico

Buscar

□ CL - Security 1

□ Contracting 40

□ EHS - Consult 1

□ EHS - Electrical 1

□ EHS - Other 3

□ EHS - Recycling 1

□ EHS - Remediation 4

Sub Category es (todos)

Tipo de filtro

Filtrado básico

Buscar

□ Seleccionar todo

□ Contracting & Servi... 41

□ Hardware 3339

□ Indirect Goods & ... 850

□ Logistics 11

□ Other 2534

□ Outsourced 2077

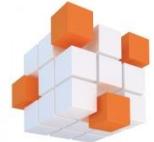
Total Invoice es (todos)



TIPOS DE ALMACENAMIENTO OLAP

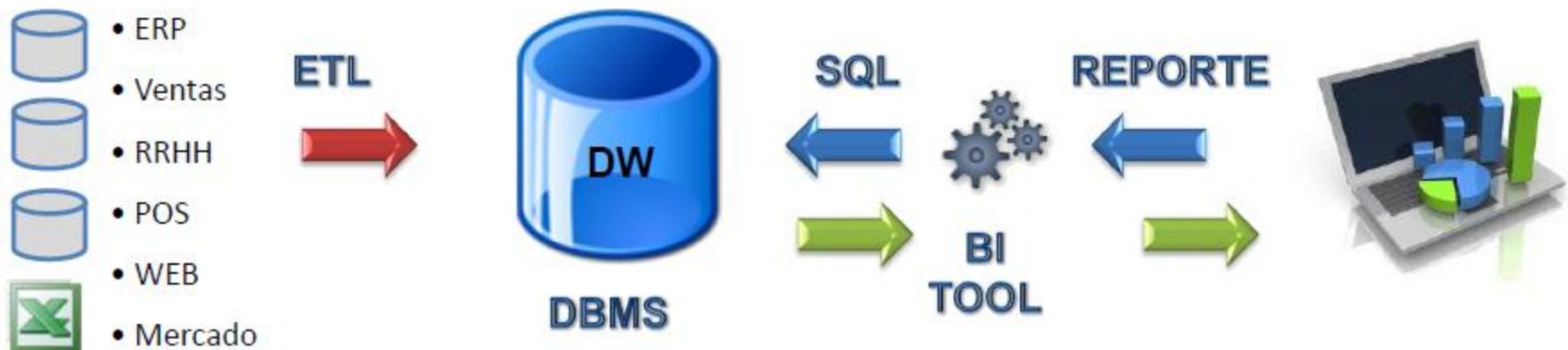
¿Dónde se guardan los datos?

- La base de datos OLAP puede tener distintas implementaciones de acuerdo a la tecnología.
 - ROLAP
 - MOLAP
 - HOLAP



ROLAP (OLAP RELACIONAL)

- Los datos del DW se guardan en BBDD relacionales.
- Se explota toda la información desde el mismo DW.
 - Las consultas multidimensionales generan SQL's al DW
 - Se pueden tener tablas con datos precalculados que aceleran las respuestas a las consultas.
- Escalable
 - Se pueden agregar nuevas tablas o nuevas columnas a tablas existentes.





ROLAP (OLAP RELACIONAL)

○ Ventajas

- Seguridad e integridad de la base de datos.
- Requiere menos espacio ya que no necesita replicar los datos.
- Los datos se almacenan en la base de datos, con lo cual no hay límite de crecimiento.

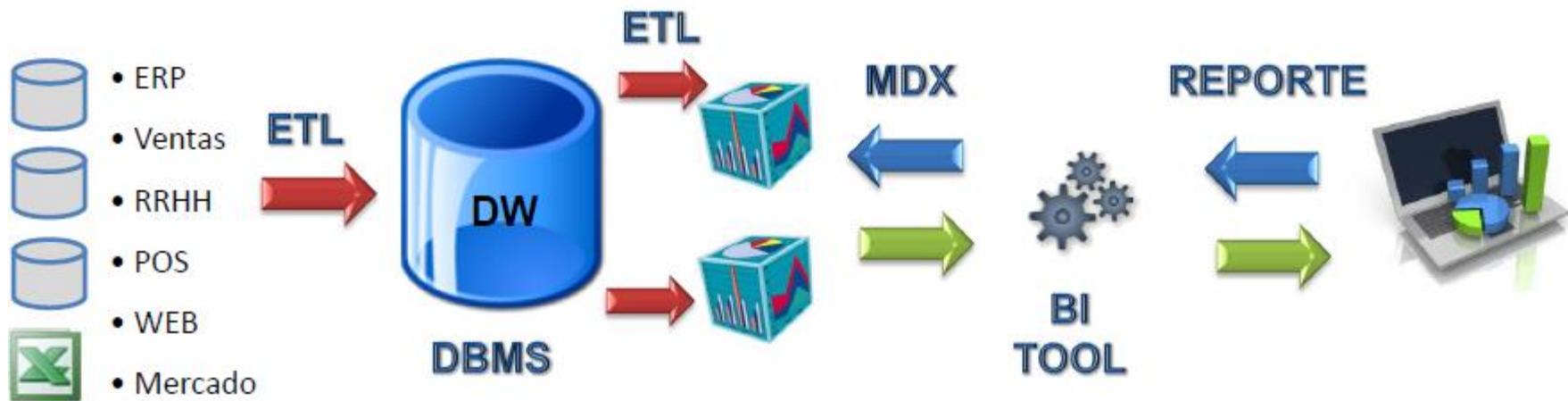
○ Desventajas

- La respuesta de las consultas es más lenta debido a que siempre tiene que ir a la base relacional a buscar los datos.
- Se debe mantener siempre conexión a la base de datos.
- Los cálculos están limitados a las funciones de la base de datos.



MOLAP (OLAP MULTIDIMENSIONAL)

- Los datos se guardan en una BBDD propietaria con estructura multidimensional.
 - Incluye datos precalculados y metadatos de las jerarquías.
- Todos los datos se cargan directamente a los cubos sin tener un DW y desde allí se consultan los datos.
- Buen tiempo de respuesta para las consultas.





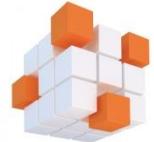
VENTAJAS MOLAP

- Mayor velocidad de acceso a los datos agregados y pre-calculados.
 - Se guardan en el servidor OLAP en un formato multidimensional sin necesidad de acceder al DW.
- Usa técnicas de compresión para guardar datos en el servidor OLAP, ocupando menos espacio que en una base relacional.
 - Actualmente, las BBDD usan técnicas de compresión.
- El modelo de almacenamiento en vectores/matrices proporciona indexación de tablas hash.
- Estructura muy optimizada para maximizar el rendimiento de las consultas.



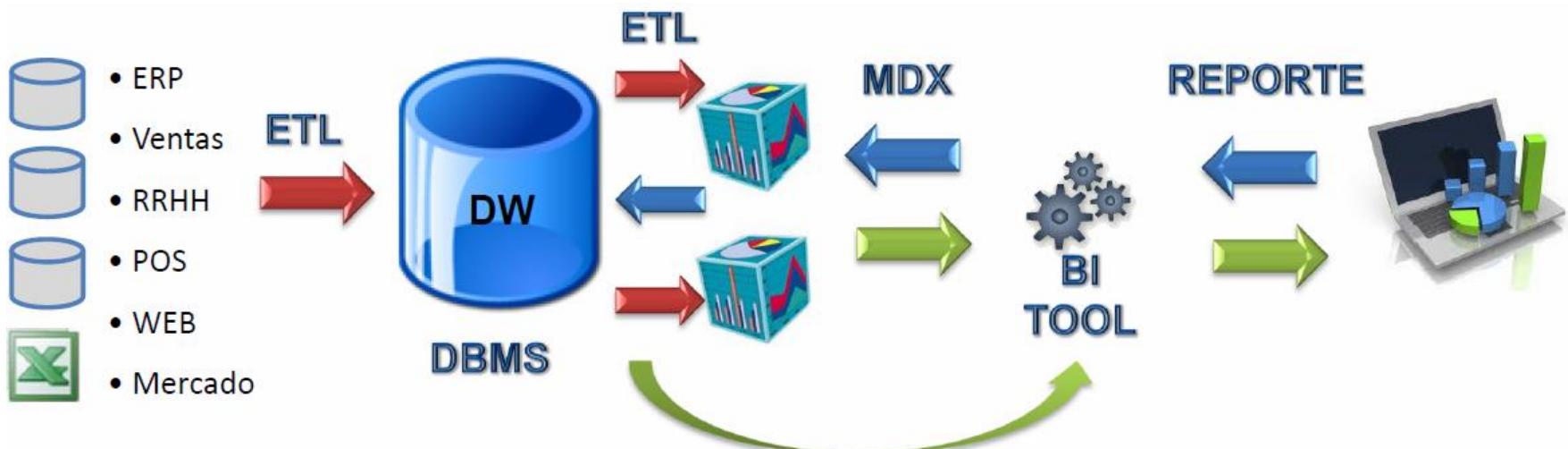
DESVENTAJAS MOLAP

- La carga de datos en estructuras multidimensionales puede durar mucho, sobre todo en grandes volúmenes de datos.
 - Esto se puede evitar realizando procesamiento incremental.
- Dificultad y complejidad para analizar modelos con más de 10 dimensiones.
- MOLAP guarda una copia de los datos relacionales en el servidor OLAP, por lo que se necesita invertir en espacio adicional.
- No puede acceder a datos que no están en el cubo.



HOLAP (OLAP HÍBRIDO)

- Utiliza una combinación de MOLAP y ROLAP.
- Los datos con menor nivel de detalle se almacenan en estructuras ROLAP.
- Los datos agregados y pre-calculados por diferentes dimensiones se almacenan en los cubos MOLAP.
- En una consulta, el motor OLAP primero busca si tiene los datos en el cubo. Si los datos no están en el cubo, dispara la consulta sobre la BBDD relacional.





TP CASO ESTUDIO: SUPERMERCADO PRECIOS JUSTOS

ENTREGA OPCIONAL (*)

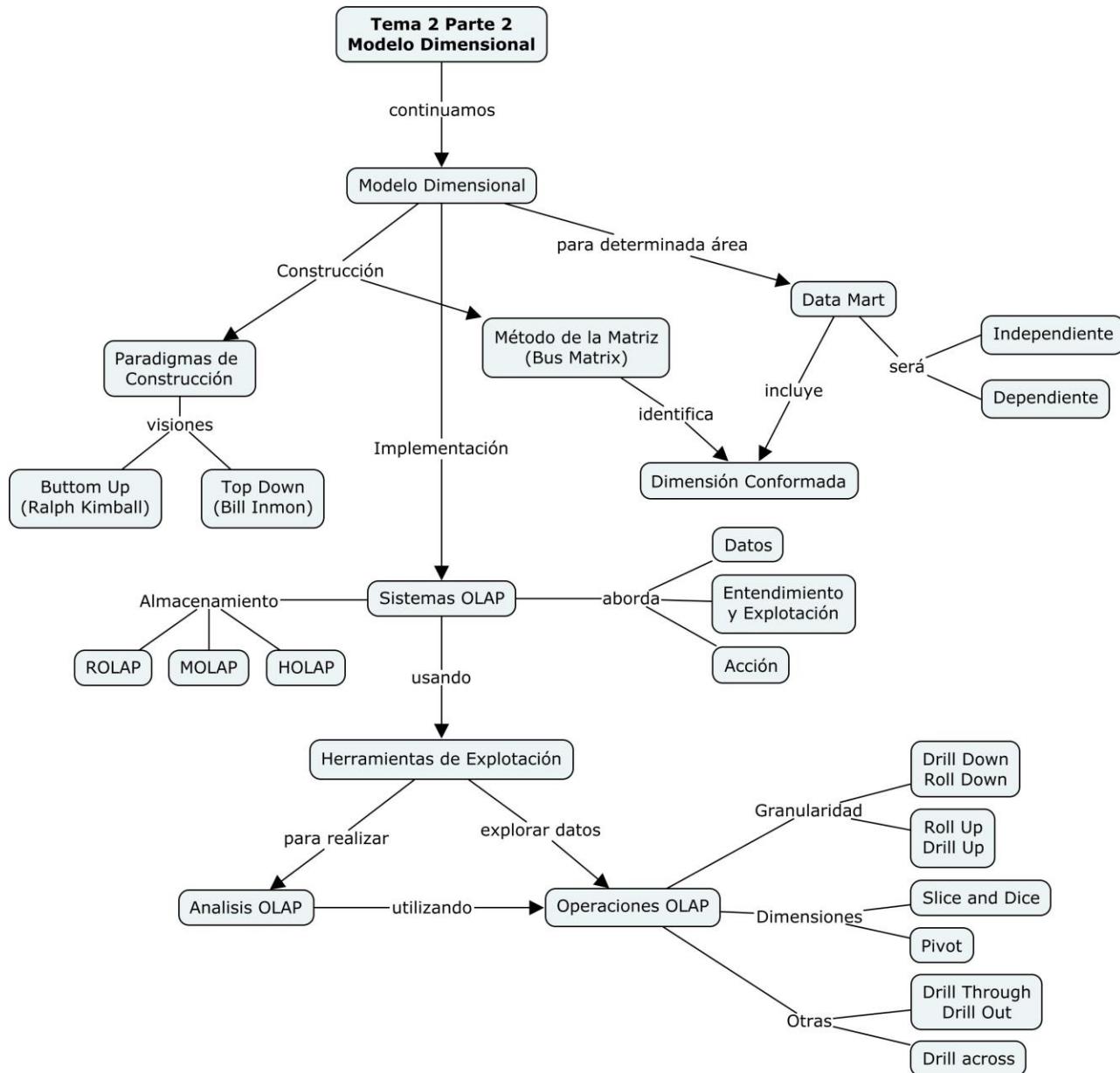
- Listar las posibles dimensiones a partir de las necesidades de información enunciadas.
- Identificar los hechos con las medidas básicas y calculadas.
- Identificar el proceso de negocio a modelar.
- Construir el modelo dimensional conceptual.
- Identificar atributos y jerarquías para cada dimensión.
- Construir el modelo dimensional lógico.

(*)

- Los TPs Opcionales sirven para afianzar conceptos necesarios para realizar los TPs de Aplicación, serán corregidos en clase y/o mediante Autoevaluación.
- Los TPs de Aplicación tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción, serán corregidos en detalle por los docentes.
- Ver documentos: "Condiciones de Cursada en MIeL - sección: Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio" + "Círculo Entrega TPs Teams/MIeL"



RESUMEN TEMA 2 PARTE 2







UNIVERSIDAD NACIONAL DE LA MATANZA

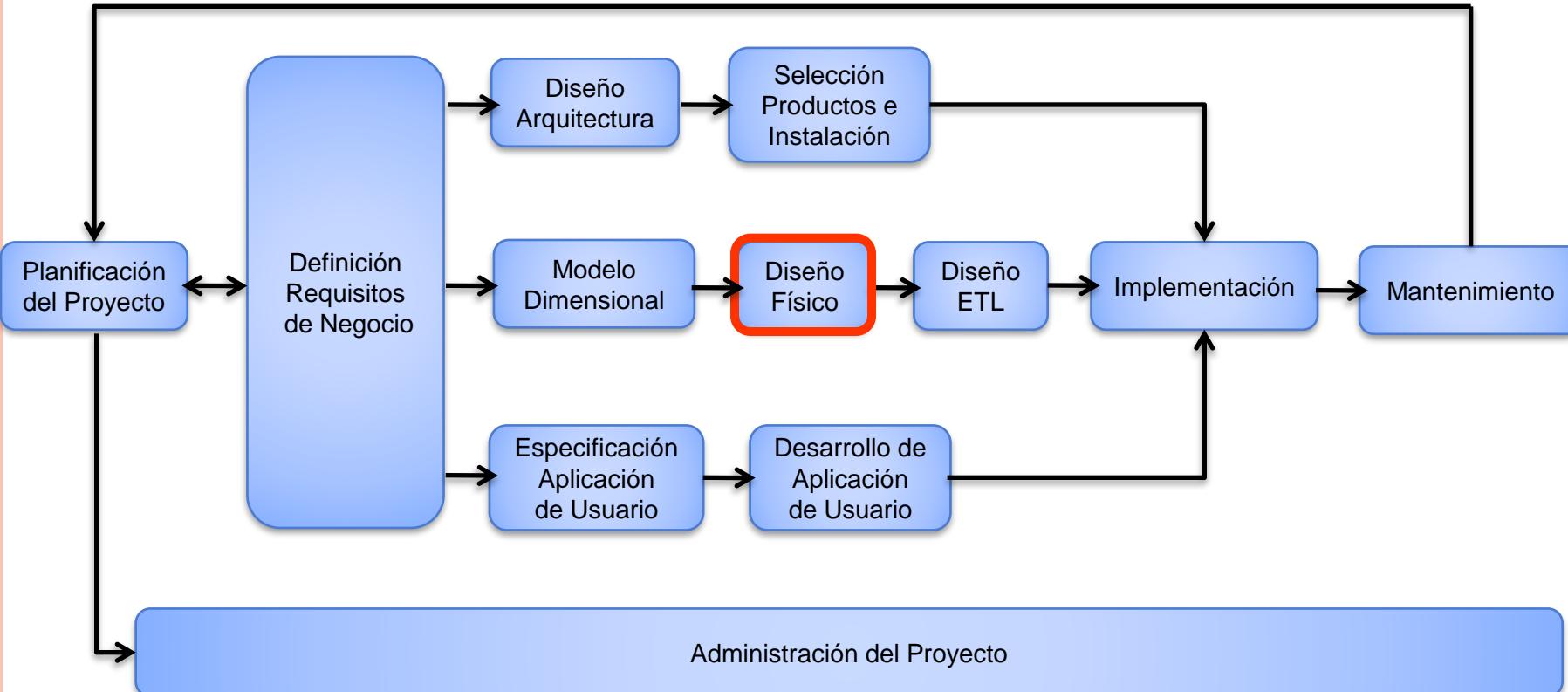
INTELIGENCIA DE NEGOCIOS

Modelo Físico
Técnicas de Modelado - Parte 1

Docentes: ING. LORENA R. MATTEO
Autores ppt orig.: Lic. HUGO M. CASTRO / Mg. DIEGO BASSO



CICLO DE VIDA DE UN PROYECTO DE BI



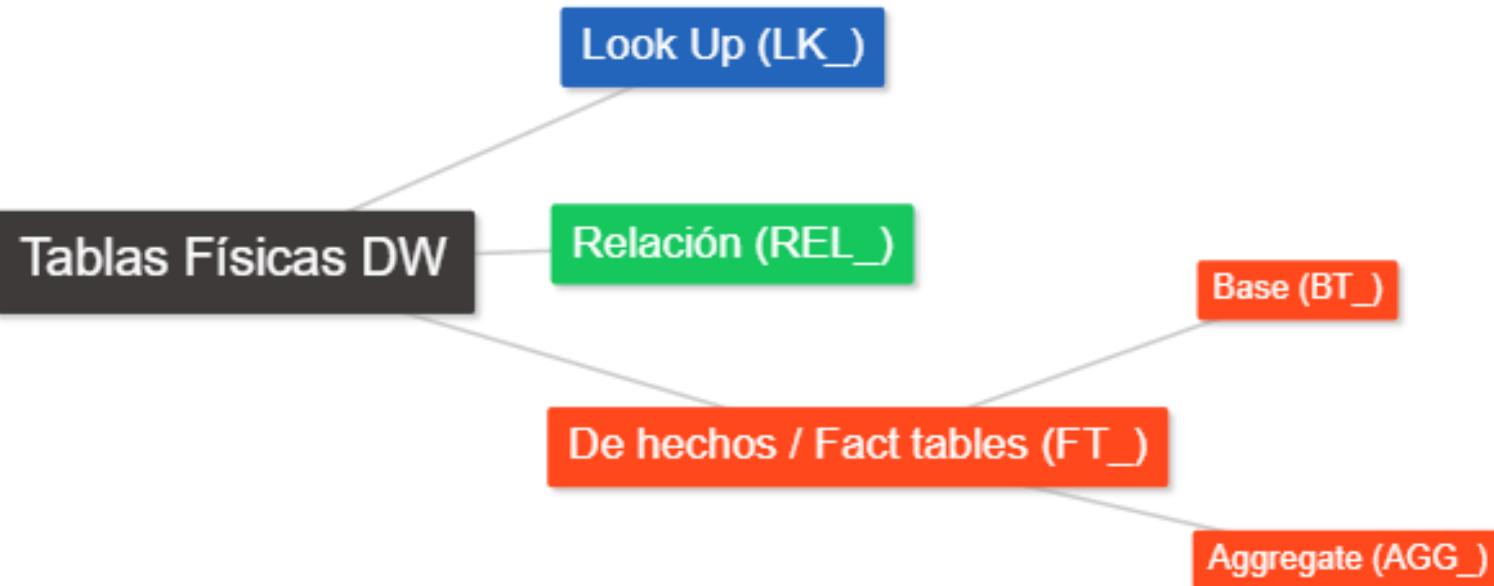


MODELO FÍSICO

- A partir del modelo dimensional lógico debemos armar el modelo físico.
- Son las tablas en las que se van a guardar los datos en el Data Warehouse.
 - Soporte a las dimensiones y grupo de hechos del modelo dimensional.
 - Se compone de tablas, registros y columnas.
- Diferentes técnicas de modelado, según la realidad del negocio.



ESTRUCTURA DE LAS TABLAS





ESTRUCTURA DE TABLAS

Tablas Look Up

- Son las tablas maestras donde se almacenan los elementos de un atributo descriptivo.
 - De tipo textual y discreto
 - Para seleccionar
 - Para agrupar
 - Para mostrar
- Se puede crear una tabla Look Up por cada **atributo** o una tabla por **dimensión**.
- No contiene valores que intervengan en cálculos.
- Valores numéricos categorizados.
- Tienen una columna por *attribute form* más una columna por cada *padre* del atributo.



EJEMPLO TABLA LOOK UP

○ Provincia

- ID_Provincia
- Desc_Provincia

} *Attribute form*

1, Buenos Aires
2, Córdoba
3, La Pampa

} *Elementos*

LK_PROVINCIA

ID_Provincia	Desc_Provincia
1	Buenos Aires
2	Córdoba
3	La Pampa
...	...

Una fila por **elemento**
del atributo

Una columna por **attribute form**



EJEMPLO TABLA LOOK UP

- Ciudad
 - ID_Ciudad
 - Desc_Ciudad



LK_CIUDAD

ID_Ciudad	Desc_Ciudad	ID_Provincia
11	Junín	1
12	Lobos	1
23	La Falda	2
35	Cosquín	2
38	Santa Rosa	3

Una columna por
attribute form

Una columna por
cada padre



ESTRUCTURA DE TABLAS

Tablas Relación

- Se utilizan cuando hay una relación *muchos a muchos* entre dos atributos.

Tabla de hechos (Fact table)

- Almacena los **valores** de las medidas de los hechos.
- Contiene una fila por cada acontecimiento que debe reflejar.
- Es la tabla de mayor cantidad de filas.
 - Ocupan más del 90% del Data Warehouse
- Tiene una columna por cada *medida* mas una columna por cada *referencia a las dimensiones* al cual se conocen los hechos.



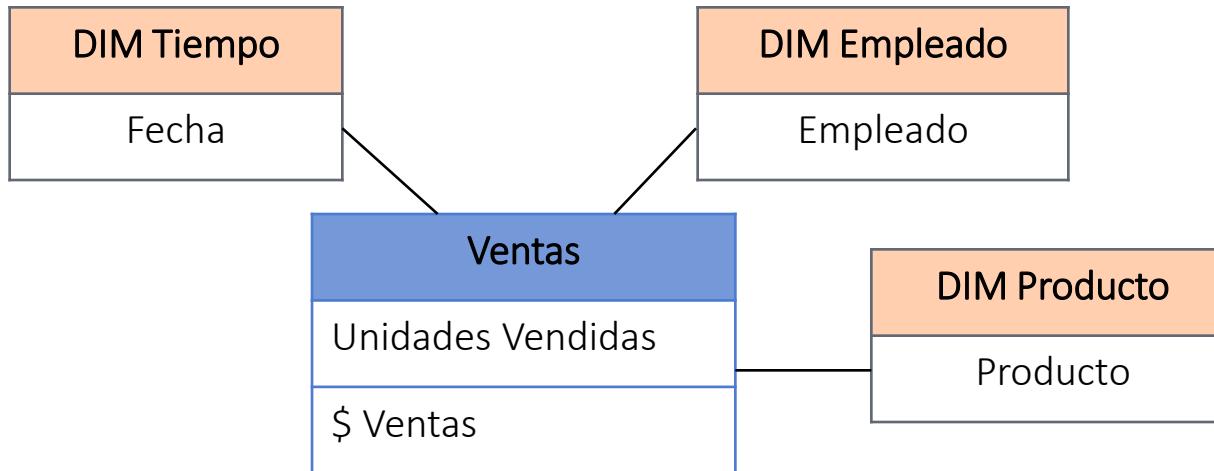
ESTRUCTURA DE TABLAS

Tabla de hechos **Base**

- Almacena los **valores** de las medidas de los hechos al *máximo nivel de detalle (Granularidad)*.
 - Esta tabla es imprescindible
- Sus filas se llenan a partir de los datos que provienen de los sistemas OLTP.



EJEMPLO DE TABLA DE HECHOS BASE



BT_VENTAS

ID_Fecha	ID_Emppleado	ID_Producto	Unid_Vend	\$ Ventas
01/04/2020	111	17	20	\$ 400
01/04/2020	112	17	45	\$ 900
01/04/2020	113	18	30	\$ 450
...

Dimensiones

Medidas



ESTRUCTURA DE TABLAS

Tabla de hechos **Agregadas**

- Se generan exclusivamente para reducir el tiempo de respuesta de las consultas.
- Su contenido se calcula a partir de la tabla de hechos base.
- Son prescindibles en términos de información.
 - No dicen nada que no sea dicho en la tabla de hechos base, pero lo dicen más rápido.
- En general, tienen pocas filas (respecto a la tabla de hechos base).



EJEMPLO DE TABLA DE HECHOS AGREGADA

BT_VENTAS

ID_Fecha	ID_Emppleado	ID_Producto	Unid_Vend	\$ Ventas
01/04/2020	111	17	20	\$ 400
08/04/2020	112	17	45	\$ 900
15/04/2020	112	17	30	\$ 450
19/04/2020	111	18	15	\$ 300

AGG_VENTAS_EMP_MES

ID_Mes	ID_Emppleado	Unid_Vend	\$ Ventas
202004	111	35	\$ 700
202004	112	75	\$ 1350
...

Contiene un atributo de alguna dimensión, por ejemplo TIEMPO





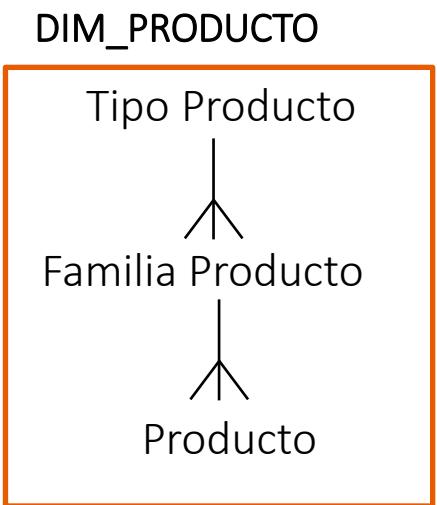
ESQUEMA ESTRELLA (STAR)

- Una tabla por cada dimensión.
- Menor cantidad de tablas involucradas.
- La tabla de hechos está en **3FN**.
 - No tiene filas repetidas.
- Las tablas de dimensión están en **2FN**.
 - Todos los productos de una misma familia llevan como atributo el nombre de la familia.
- Mayor espacio de almacenamiento.
- Las consultas son más fáciles de construir.
 - Joins entre tabla de hechos con cada tabla de dimensiones.
- Mejor rendimiento en consultas.
 - Se necesita acceder a menos tablas.



ESQUEMA ESTRELLA (STAR)

- Producto
 - ID_Producto
 - Desc_Producto
- Familia_Producto
 - ID_FamiliaProducto
 - Desc_Familia
- Tipo_Producto
 - ID_TipoProducto
 - Desc_TipoProducto

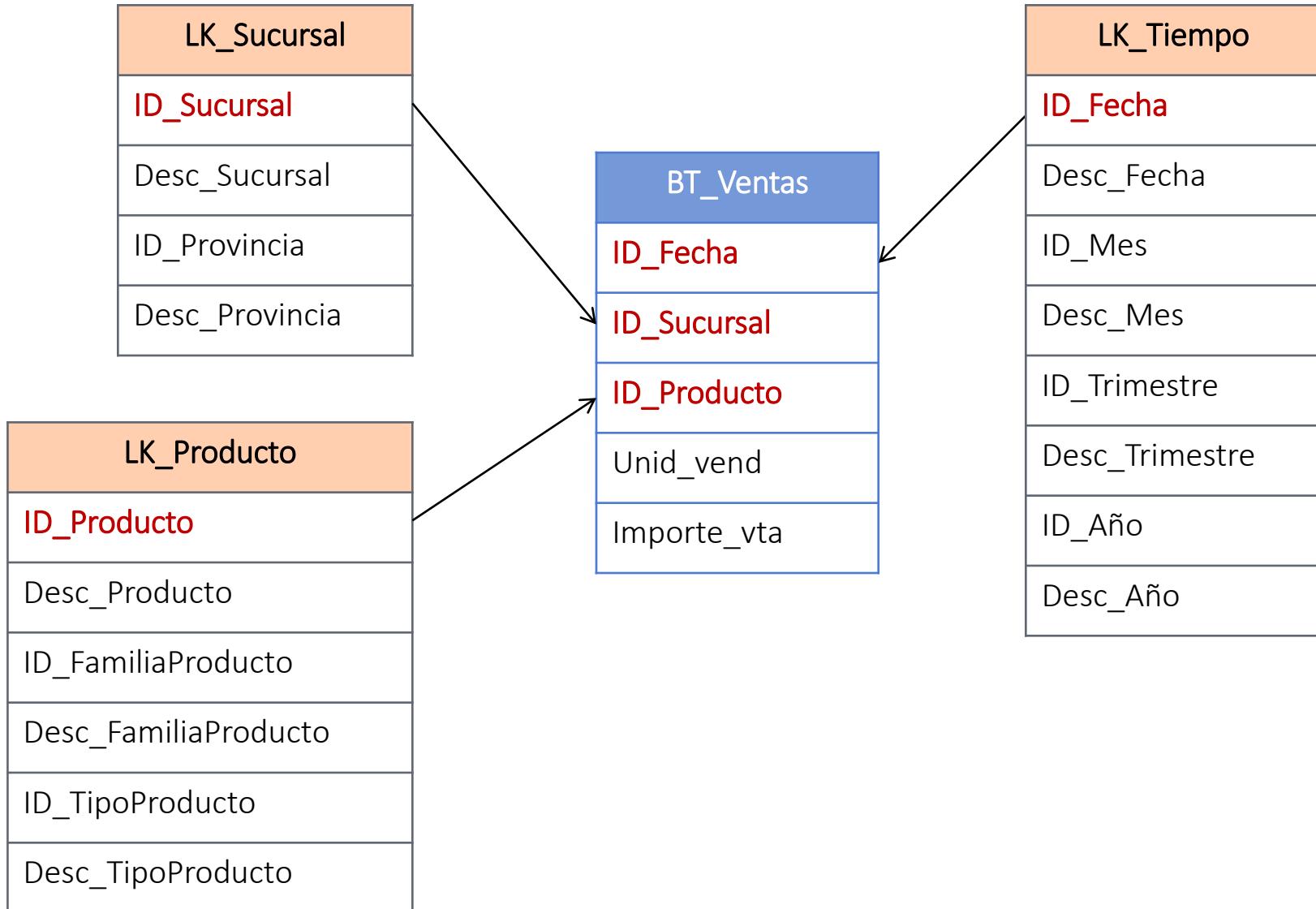


LK_Producto
ID_Producto
Desc_Producto
ID_FamiliaProducto
Desc_FamiliaProducto
ID_TipoProducto
Desc_TipoProducto

BT_Ventas
ID_Fecha
ID_Sucursal
ID_Empleado
ID_Producto
Unid_Vend
Importe_Vta



EJEMPLO ESQUEMA ESTRELLA





CONSULTA SQL A UN MODELO ESTRELLA

- Ejemplo: Para cada producto de tipo bazar calcular el importe vendido en las sucursales de CABA en el mes de febrero de 2020
- Dimensiones para:
 - Mostrar, filtrar, agrupar
- Medidas o Hechos para:
 - Calcular
- Join: Sólo entre tabla de hechos y tabla de dimensión



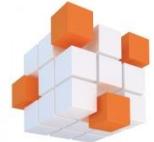
CONSULTA SQL A UN MODELO ESTRELLA

```
SELECT p.Desc_Producto, sum(v.Importe_vta) AS Importe  
FROM  
    BT_Ventas v,  
    LK_Sucursal s,  
    LK_Producto p,  
    LK_Tiempo t  
WHERE  
    v.ID_Fecha = t.ID_Fecha and  
    v.ID_Sucursal = s.ID_Sucursal and  
    v.ID_Producto = p.ID_Producto and  
    s.Desc_Provincia = 'CABA' and  
    t.ID_Mes = '022020' and  
    p.Desc_TipoProducto = 'Bazar'  
GROUP BY p.Desc_Producto
```



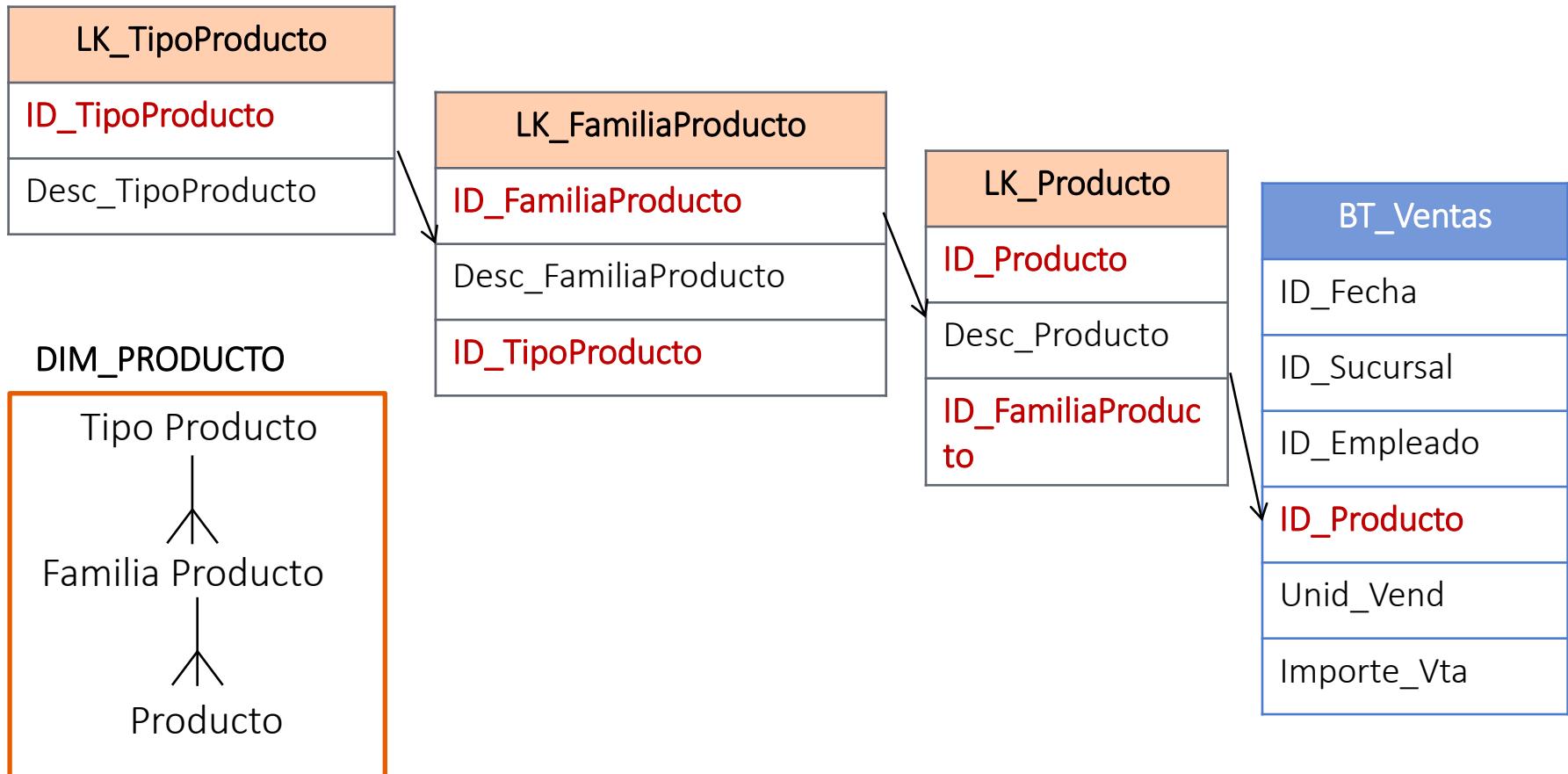
ESQUEMA COPO DE NIEVE (SNOWFLAKE)

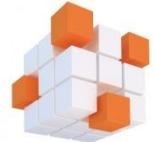
- Una tabla por atributo.
- Todas las tablas (dimensiones y hechos) están en **3FN**.
 - No tiene filas repetidas.
- Menor espacio de almacenamiento.
- Más fácil de entender para el usuario final.
- Más fáciles de llenar las tablas.
- Menor rendimiento de las consultas por los join.
- Aplicable para tablas de dimensiones con una gran cantidad de filas.
- Ampliamente usado en implementaciones ROLAP.



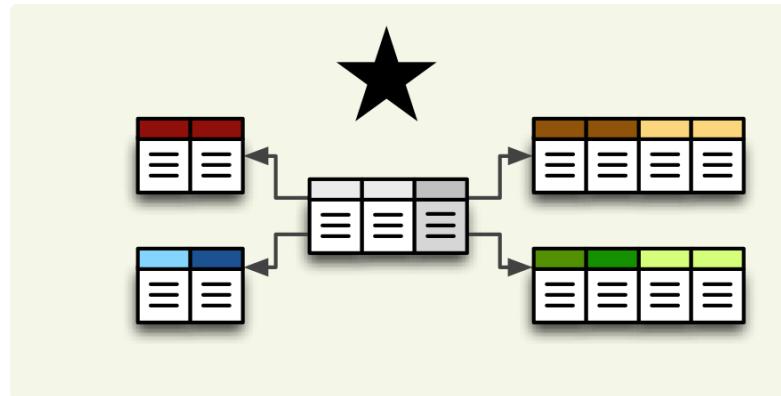
ESQUEMA COPO DE NIEVE (SNOWFLAKE)

- Producto
 - Producto_ID
 - Desc_producto
- Familia_Producto
 - ID_FamiliaProducto
 - Desc_Familia
- Tipo_Producto
 - ID_TipoProducto
 - Desc_TipoProducto

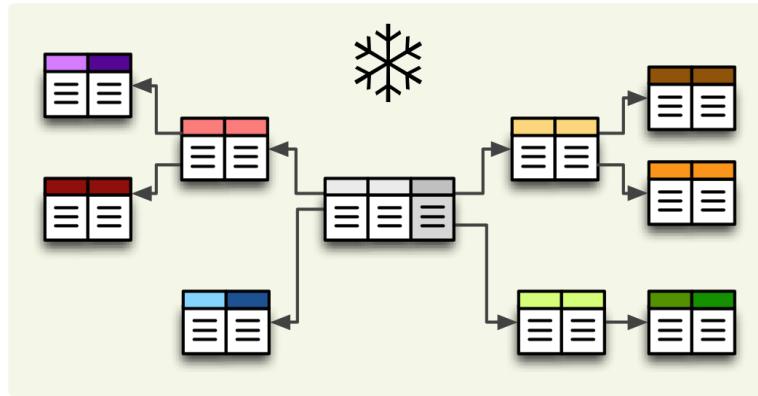




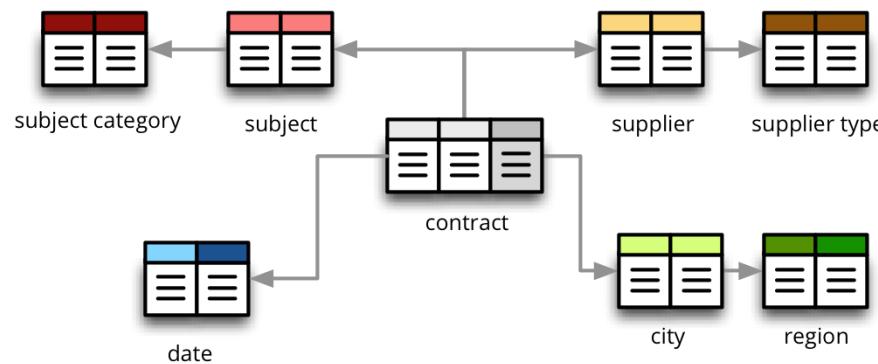
ESQUEMAS STAR VS SNOWFLAKE



Modelo Estrella



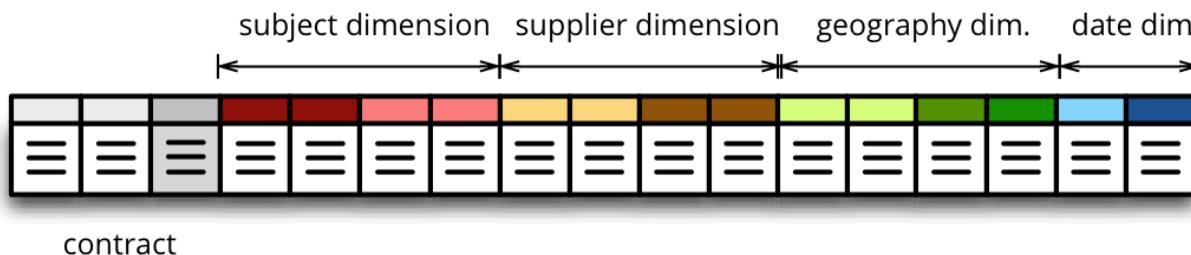
Modelo Copo de Nieve



Es posible pasar de este esquema **Copo de Nieve** ...



... a este en **Estrella** (desnormalizado) con todos los atributos del cubo entendiendo como relacionan las tablas a través de las **SKs**.

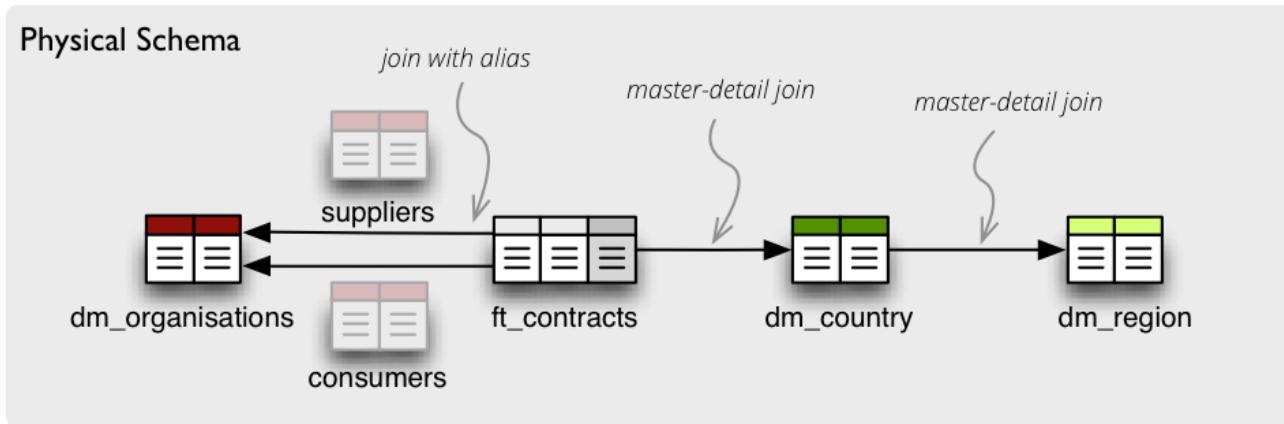


Fuente: <http://databrewery.org/cubes/doc/mapping.html#joins>



ESQUEMAS STAR VS SNOWFLAKE - ALIAS

- ¿Qué sucede si necesita unirse a la misma tabla dos o más veces?
 - Por ejemplo, tiene una lista de organizaciones y es necesario usarse como proveedor y consumidor de servicios.



Especificar Alias en los joins (pseudo):

```
"joins" =  
[  
  { "master" : "contratos.supplier_id",  
    "detalle" : "organizaciones.id",  
    "alias" : "proveedores"  
  },  
  { "master" : "contratos.consumer_id",  
    "detalle" : "organizaciones.id" ,  
    "alias" : "consumidores"  
  }  
]
```

Fuente: <http://databrewery.org/cubes/doc/mapping.html#joins>

Usar el alias especificado en los joins, y no el nombre real de la tabla.

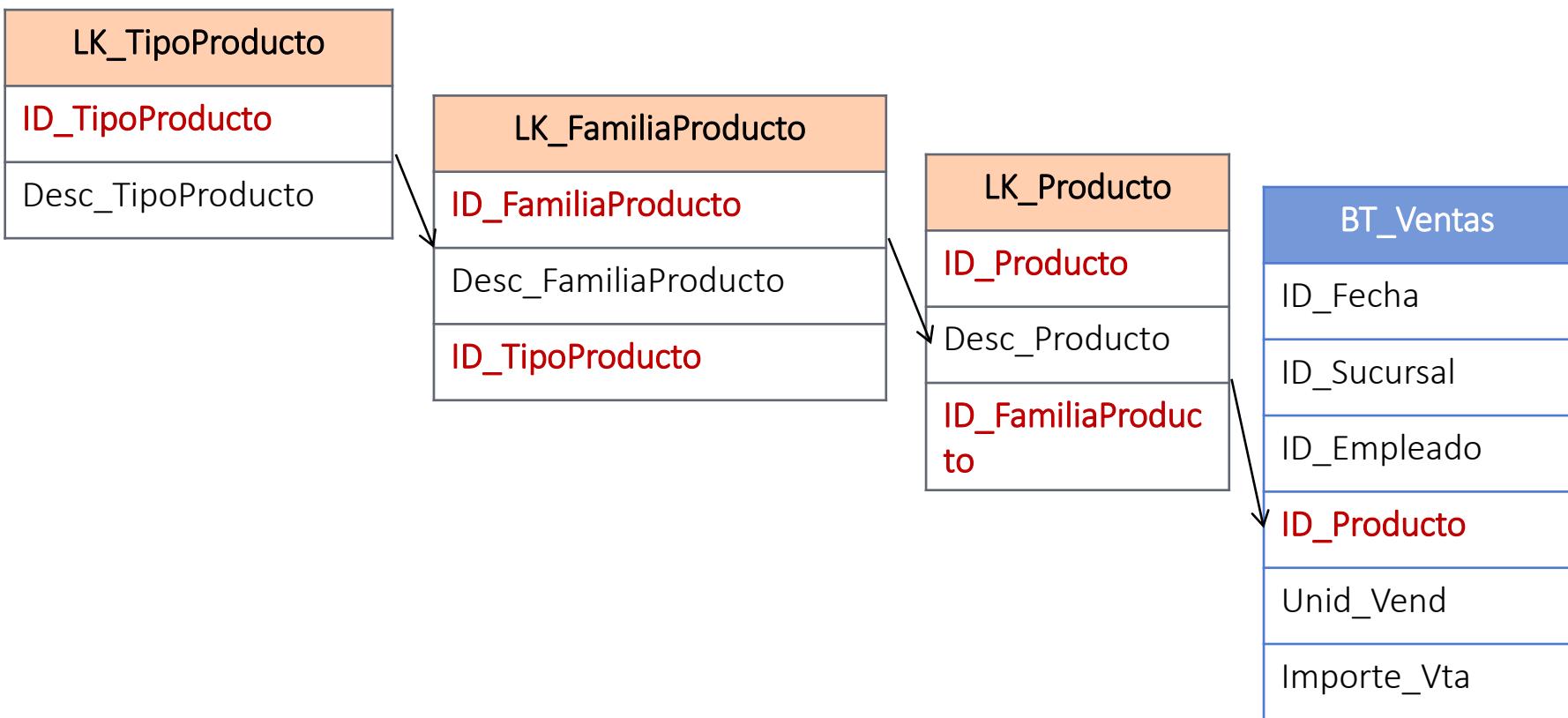
```
"asignaciones" :  
{  
  "proveedor.nombre" :  
  "proveedores.org_nombre",  
  "consumidor.nombre" :  
  "consumidores.org_nombre"  
}
```



NORMALIZACIÓN

○ Completamente Normalizado

- Las tablas Look Up tienen el ID propio, la descripción y el ID del padre.
- Minimiza la redundancia.
- Muchos joins para acceder a las tablas de más alto nivel.

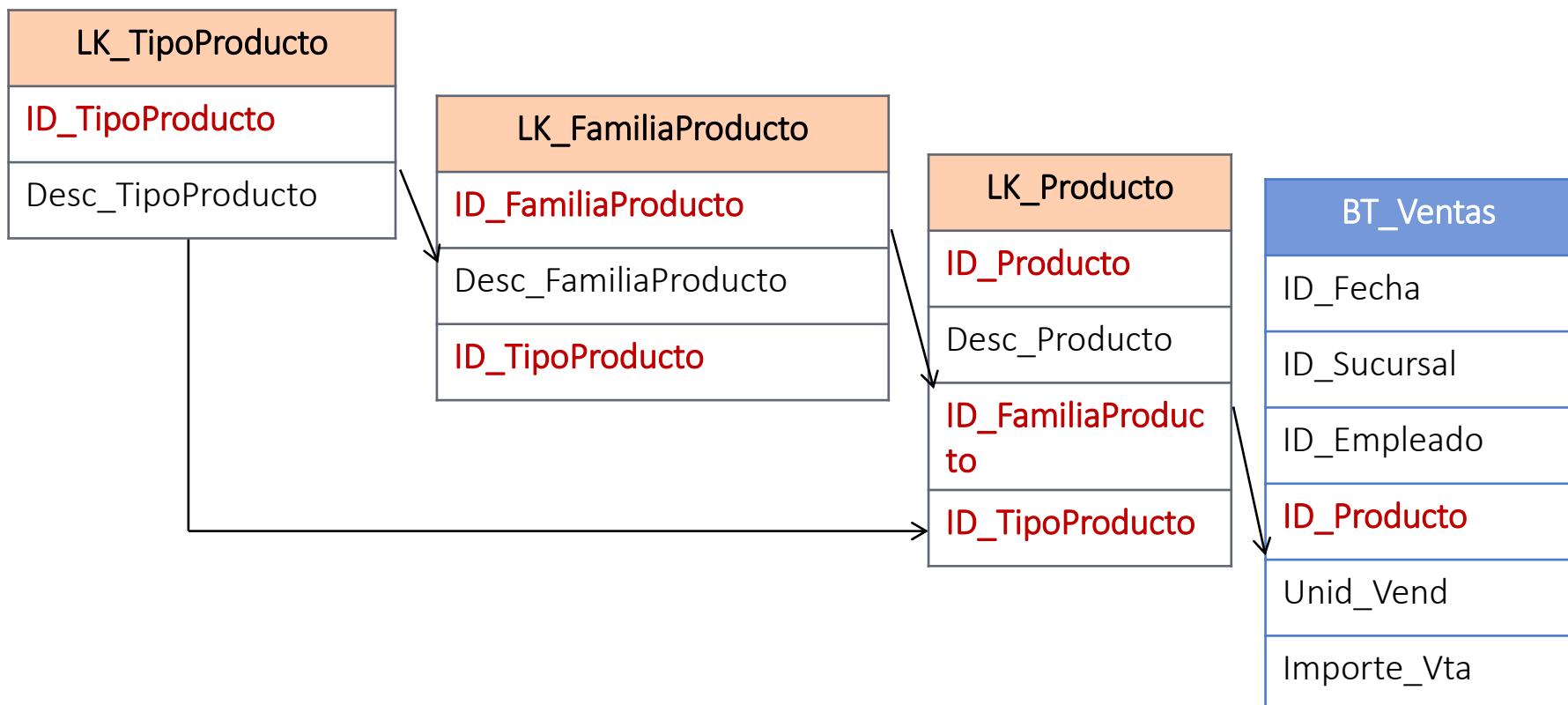




NORMALIZACIÓN

○ Moderadamente Normalizado

- Las tablas Look Up tienen el ID propio, la descripción y todos los ID de sus ancestros.
- Algo de redundancia.
- Reduce significativamente los joins para consultar datos dentro de la jerarquía.

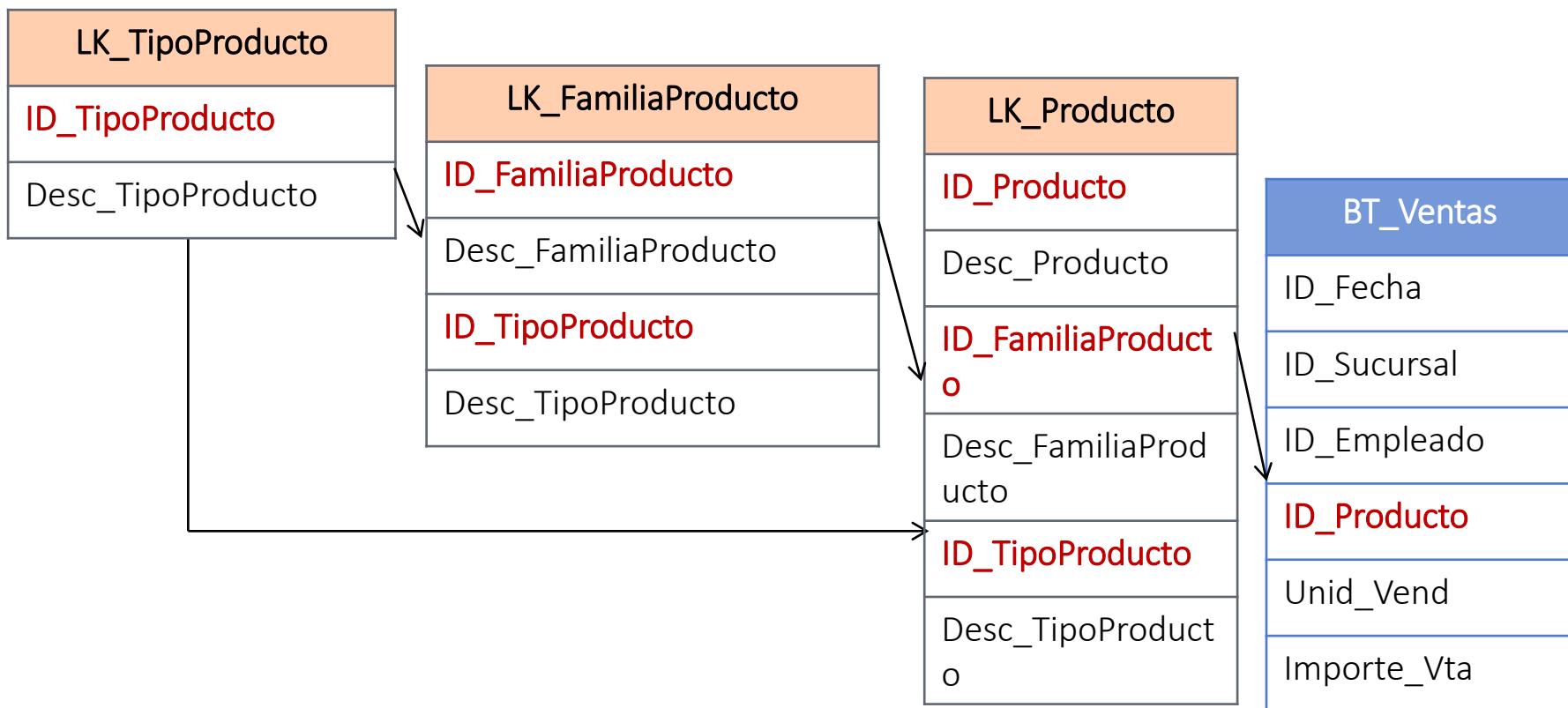




NORMALIZACIÓN

○ Completamente Desnormalizado

- Las tablas Look Up tienen el ID propio, la descripción y todos los ID y descripción de sus ancestros.
- Mucha redundancia y espacio de almacenamiento.
- Elimina los joins para consultar datos dentro de la jerarquía.





TIPOS DE MEDIDAS o HECHOS

- Aditivas
 - Se pueden sumar (SUM) por todas las dimensiones.
- Ejemplos
 - Importe de venta
 - Tiene sentido sumar el importe por producto, por sucursal, por fecha, por empleado.
 - Costo de venta
 - Ganancia o beneficio
 - Cantidad de ventas (en el día)
 - Cantidad de expedientes procesados (en una semana)
 - Cantidad de nacimientos (por mes)



TIPOS DE MEDIDAS o HECHOS

- Semiaditivas
 - Se pueden sumar en algunas dimensiones y en otras no.
 - Carece de sentido sumarla en otras dimensiones.
- Ejemplos
 - Unidades Vendidas
 - Tiene sentido sumar sólo si está la dimensión *Producto*.
 - Unidades en Stock (o inventarios de cualquier tipo)
 - Sólo se pueden sumar si son del mismo día, mes, año, etc.
 - No se puede sumar por *producto* ⇒ semánticamente incorrecto.
 - Saldo de Cuentas
 - Sólo se puede sumar si son del mismo día. Se obtiene el saldo x día.
 - Normalmente se obtiene a nivel del último día del mes.
 - Número de clientes
 - No se puede sumar por *producto* ⇒ semánticamente incorrecto.



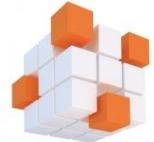
TIPOS DE MEDIDAS O HECHOS

- No aditivas
 - No tiene sentido sumarlas por ninguna dimensión.
 - Usan otras formas de consolidación o agregación.
- Ejemplos
 - Precio Unitario
 - No tiene sentido si lo sumo por día, producto o sucursal.
 - Margen o porcentaje de ganancia
 - No se puede sumarizar por ninguna dimensión.
 - Es una medida calculada obtenida con la herramienta OLAP.
 - Temperatura
 - Edad
 - Notas de exámenes
 - Tipo de cambio de moneda



FORMAS DE CONSOLIDACIÓN DE MEDIDAS

- Suma
- Promedio
- Máximo
- Mínimo
- Cantidad de casos
- Cantidad de casos distintos



¿COMO SE CONECTA LA TABLA DE HECHOS A LAS DE DIMENSIÓN?

- Verificar **integridad referencial** entre la tabla de hechos y las tablas de dimensión.
- En la tabla de dimensión:
 - La clave tiene que ser una clave primaria (PK)
- En la tabla de hechos:
 - Cada dimensión tiene su clave foránea (FK) que apunta a la fila que corresponde en la tabla de dimensión.
 - Todas las FK de las dimensiones forman la PK de la tabla de hechos.



CLAVE SUBROGADA (SURROGATE KEY - SK)

Problemas con las claves

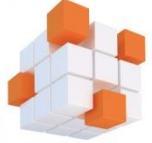
- Los sistemas fuentes utilizan sus propias claves (código de artículo, código de cliente, etc.), con un determinado formato y sentido para el negocio.
- Podrían cambiarse los formatos de las claves de los sistemas fuentes.
- Los datos de un DW vienen de fuentes heterogéneas.
- Las SK generan dentro del ámbito del DW una clave numérica única sin significado para el negocio.
 - Número entero asignado en forma secuencial.
- Aconsejable no usar como claves los códigos de los sistemas fuentes.



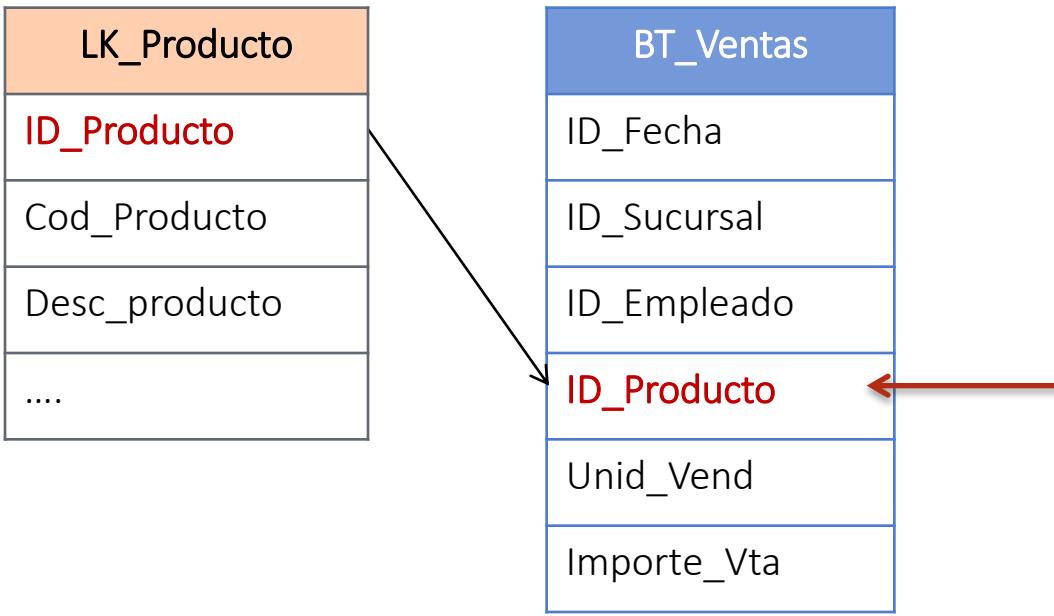
CLAVE SUBROGADA (SURROGATE KEY - SK)

Se necesita crear una SK en un DW

- Los sistemas fuentes pueden cambiar la descripción de un producto o cliente, sin cambiar el código del DW que lo representa.
- Si se cambian los formatos de las claves en los sistemas fuentes, la clave SK no se altera.
- Los sistemas fuentes pueden reutilizar códigos que se habían depurado y aún existen en el DW.
- Rendimiento
 - Las claves de las tablas ocuparán menos espacio.
- Más fáciles de mantener.



CLAVE SUBROGADA (SURROGATE KEY - SK)



Sólo se usa para conectar las tablas.

No tiene significado para el negocio



CLAVE SUBROGADA (SURROGATE KEY - SK)

Ventajas

- El DW se independiza de cambios en el manejo de claves de los sistemas fuente.
- Permite manejar dimensiones de cambio lento.

Desventajas

- Hay que manejar y administrar estas claves en el proceso de ETL.



LAS 10 REGLAS ESENCIALES DEL MODELADO

"10 Mandamientos de Kimball"

- Regla #1: Cargar los datos atómicos en estructuras dimensionales
- Regla #2: Estructura de los modelos dimensionales en función de los procesos de negocio
- Regla #3: Asegurarse de que cada tabla de hechos tiene una tabla de dimensión tiempo asociada
- Regla #4: Asegurarse de que todos los hechos de la tabla de hechos tienen el mismo nivel de detalle
- Regla #5: Resolver correspondencias muchos a muchos en tablas de hechos
- Rule #6: Resolver correspondencias muchos a muchos en tablas de dimensiones
- Regla #7: Almacenar las descripciones en las tablas de dimensión
- Regla #8: Asegurarse de que las tablas dimensionales usan claves subrogadas
- Regla #9: Crear dimensiones conformadas para integrar los datos de toda la empresa
- Regla #10: Valora constantemente los requerimientos y las realidades para proporcionar una solución DW/BI que sea aceptada por los usuarios de negocios y que apoye su proceso de decisiones



CONSEJOS NOMBRES DE OBJETOS DWH

... deben estar orientados a los usuarios de negocios, no al personal técnico.

- Regla #1. Sigue la convención de nomenclatura.
- Regla #2: Cada objeto tiene un nombre.
- Regla #3: Los nombres de los objetos son descriptivos.
- Regla #4: No se recomiendan abreviaciones y acrónimos.
- Regla #5: Los nombres de los objetos son atractivos.
- Rule #6: Los nombres de los objetos son únicos.
- Regla #7: Nombres de objetos no demasiado largos.
- Regla #8: Considera anteponer a las columnas de nombre, una tabla abreviada del nombre.
- Regla #9: Cambia los nombres en la capa de visualización si es necesario.
- Regla #10: ¡Se constante!

Fuente: <https://datawarehouse.es/2014/que-hay-en-un-nombre.html>



DIMENSIONAMIENTO DE UN MODELO

EJEMPLO

Dimensión	
Tiempo	2 Años x 365 días = 730 días
Sucursal	300 sucursales Informan ventas x día
Producto	30000 en cada sucursal, 4000 vendidos x día en todas las sucursales
Empleado	Solo 1 empleado x sucursal x día

Número Base de Registros en las Fact Tables

$730 \times 300 \times 4000 \times 1 = 876 \text{ Millones de registros}$

Número de los Campos Indices = **4**

Número de los campos de la Fact Table = **4**

Total de Campos = **8**

Tamaño Base de la Fact Table

$876 \text{ Millones} \times 8 \text{ Campos} \times 4 \text{ bytes } (*) = 28 \text{ Gb}$

(*) caso hipotético en que todos ocupen 4bytes.

TP CASO ESTUDIO: PLATAFORMA STREAMING SOUNDWAVE

ENTREGA OPCIONAL (*)



- Listar las posibles dimensiones a partir de las necesidades de información enunciadas.
- Identificar los hechos con las medidas básicas y calculadas.
- Identificar el proceso de negocio a modelar.
- Construir el modelo dimensional conceptual.
- Identificar atributos y jerarquías para cada dimensión.
- Construir el modelo dimensional lógico.
- Construir el modelo físico (en esquema Estrella y Copo de Nieve).

(*)

- Los TPs Opcionales sirven para afianzar conceptos necesarios para realizar los TPs de Aplicación, serán corregidos en clase y/o mediante Autoevaluación.
- Los TPs de Aplicación tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción, serán corregidos en detalle por los docentes.
- Ver documentos: "Condiciones de Cursada en MIeL - sección: Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio" + "Círculo Entrega TPs Teams/MIeL"



TRABAJO PRÁCTICO DE APLICACIÓN

ENTREGA PRÓXIMA CLASE (*)

Parte 1 (Modelo Dimensional)

- Definir el contexto de la empresa.
- Listar dimensiones, hechos con medidas (básicas y calculadas), atributos y jerarquías.
- Construir los modelos dimensional conceptual y lógico.
- Buscar nuevas preguntas de negocio que podría responder el modelo construido.

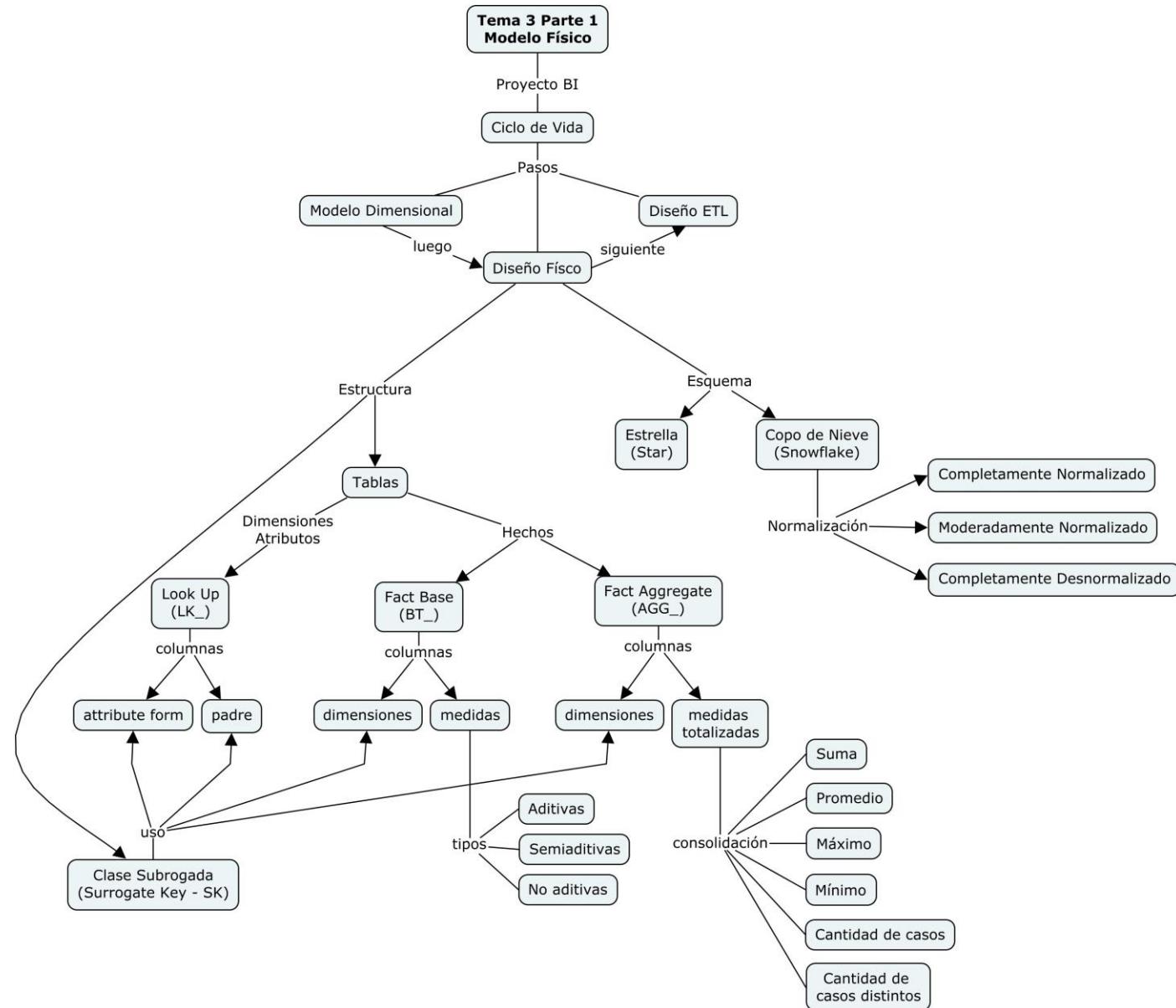
Parte 2 (Modelo Físico) (Próxima clase + 1)

- Construir el modelo físico (en esquema Estrella y Copo de Nieve) + identificar cada dimensión del modelo dimensional.

(*) Cada trabajo práctico tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción. (Ver Condiciones de Cursada en MIeL – Sección: "Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio")



RESUMEN TEMA 3 PARTE 1







UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

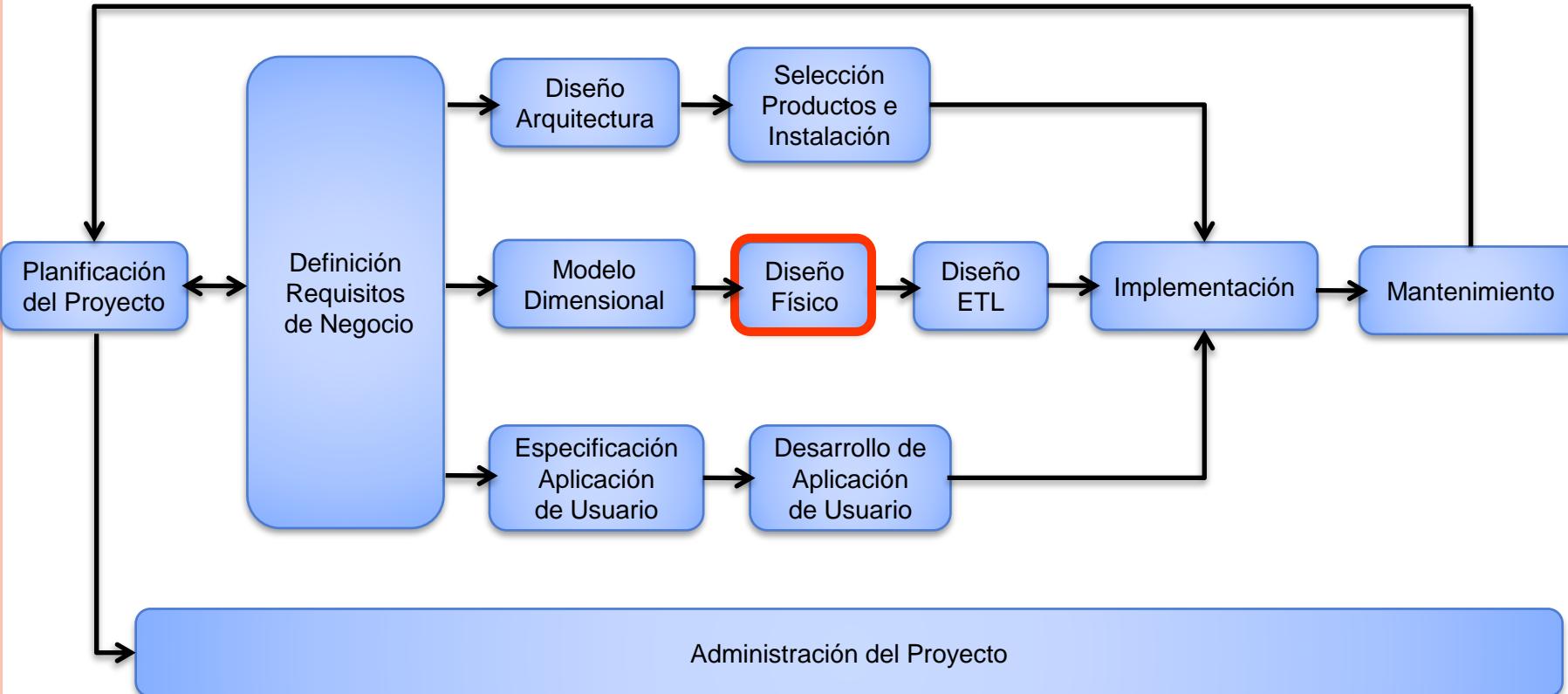
Modelo Físico
Técnicas de Modelado – Parte 2

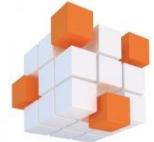
Docentes: ING. LORENA R. MATTEO
Autores ppt orig.: Lic. HUGO M. CASTRO / MG. DIEGO BASSO



CICLO DE VIDA DE UN PROYECTO DE BI

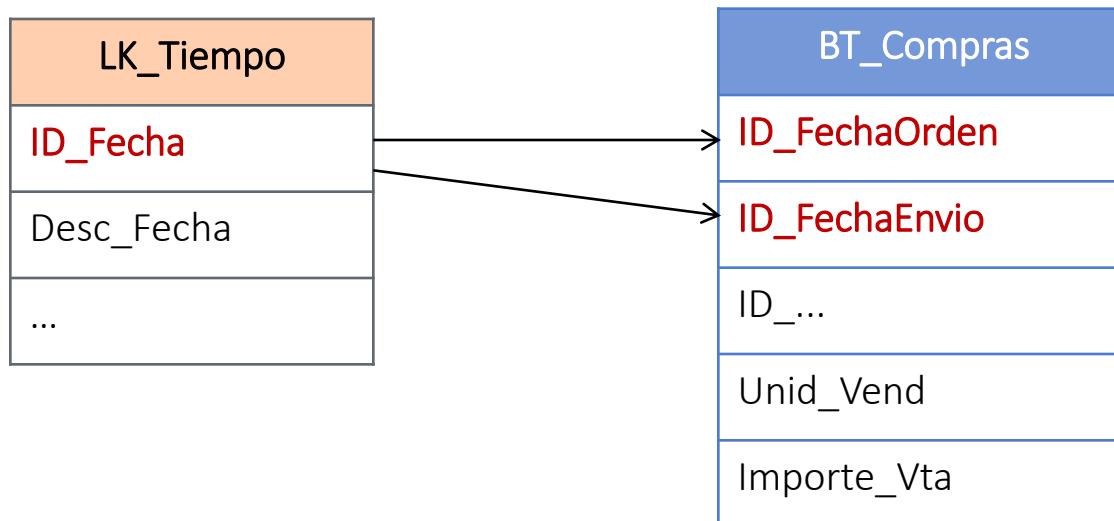
6/5/2025 IN2025





DIMENSIÓN MÚLTIPLE

- Existen situaciones que refieren a diferentes instancias de una misma dimensión en la tabla de hechos.
 - Ejemplo: Una orden es solicitada en una fecha determinada y enviada días después. La fecha cumple dos roles distintos.

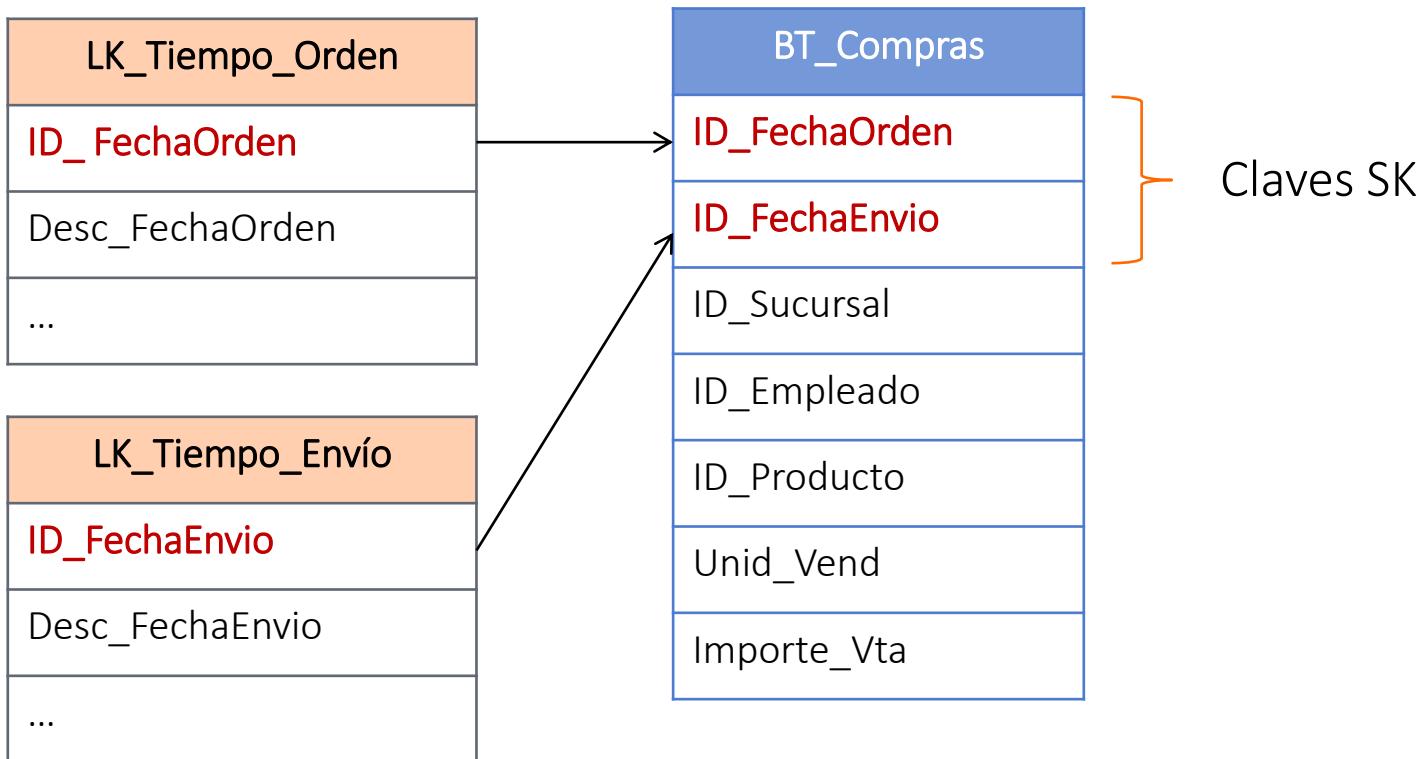


¿Es correcto este modelo?



DIMENSIÓN MÚLTIPLE

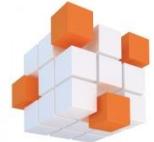
- Al momento de querer implementarlo tendríamos el problema de una dimensión de fecha que se relaciona con más de una SK de fecha en la tabla de hechos.
- Tantas tablas de dimensión (Alias) como SK en la fact table.





DIMENSIÓN DEGENERADA

- Se refiere a dimensiones que tienen un identificador transaccional (no necesariamente numérico) generado en el sistema fuente.
- Ejemplos
 - Número de órdenes de compra
 - Número de factura
 - Número de remito
 - Número de transacción
- Parece una dimensión, pero ¿cuáles son sus atributos?
 - Los atributos que podrían incluirse ya están en la propia dimensión (fecha, vendedor, sucursal, cliente).
 - Aparece cuando la granularidad es la misma que la de la tabla de hechos
 - Nivel de renglón de factura o de orden de compra



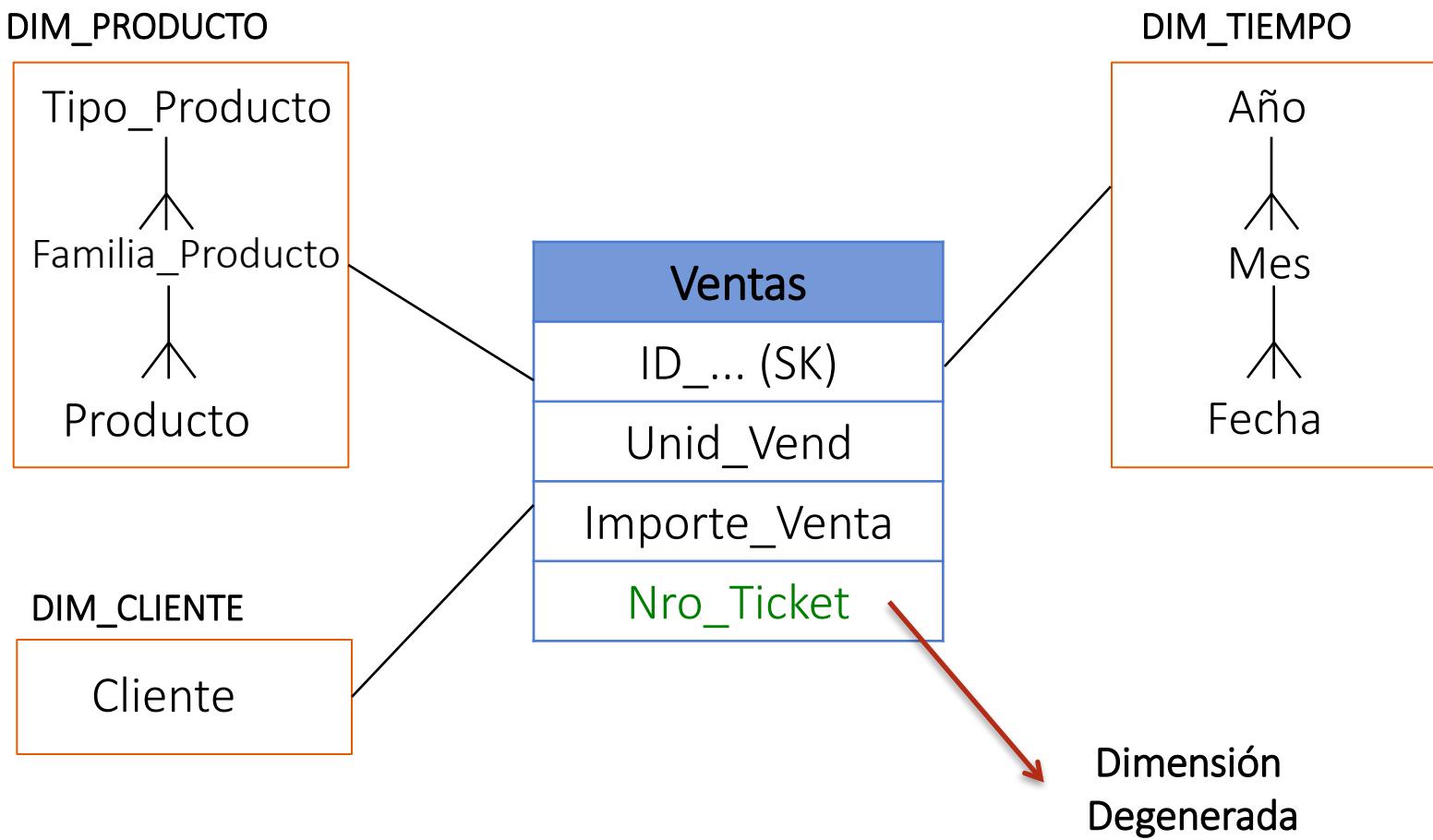
DIMENSIÓN DEGENERADA

- Esta dimensión no tiene una tabla asociada.
 - Atributo con sólo 2 columnas: Clave SK + N° transacción
- Es un atributo que será utilizado como criterio de análisis y que es almacenado en la *tabla de hechos*.
 - No permite agrupación ni summarización.
- ¿Para qué sirve?
 - Permite vincular todos los ítems pertenecientes a una misma transacción.
 - Sirve como fuente para procesos que requieran saber los productos que se compraron juntos en la misma transacción.
 - Da la posibilidad de localizar datos en los sistemas fuente.



DIMENSIÓN DEGENERADA

6/5/2025 IN2025





DIMENSIÓN DEGENERADA

6/5/2025 IN2025

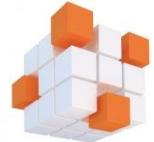
BT_VENTAS

ID_Fecha	ID_Cliente	ID_Producto	Unid_Vend	\$ Venta	Nro_Ticket
15/04/2017	1	10	3	\$ 90	1234
15/04/2017	1	11	2	\$ 100	1234
15/04/2017	1	12	5	\$ 150	1234
15/04/2017	3	10	2	\$ 60	2411
15/04/2017	3	13	10	\$ 200	2411

Dimensiones

Medidas

Dimensión
Degenerada



DIMENSIONES DE CAMBIO LENTO

- El horizonte temporal del DW es mayor que el de los sistemas transaccionales.
- El DW debe reflejar el paso del tiempo pero no perder la historia.
 - Un producto cambia de denominación.
 - Una sucursal cambia de distrito.
- Son dimensiones en las cuales sus datos tienden a modificarse a través del tiempo.
- Los atributos de las dimensiones son relativamente estáticos pero cambian a lo largo del tiempo.



DIMENSIONES DE CAMBIO LENTO

¿Qué hay que hacer?

- Para cada atributo hay que definir una estrategia de manejo de cambios.
- El modelo dimensional debe responder a los cambios.
- Interpretación del profesional de negocios.
 - ¿Cómo se quiere guardar la historia de cambios en los atributos?
 - ¿Con qué detalle?



DIMENSIONES DE CAMBIO LENTO

- Existen distintos tipos de técnicas para manejar cambios en los atributos.
- Se denominan **SCD (Slowly Changing Dimensions)**
 - Manejan en forma diferente la conservación de la historia.
 - No hay una que sea mejor que otra.
 - Se define para **cada atributo de cada dimensión**.



SCD TIPO 0

Se mantiene el valor original

- El atributo de la dimensión no permite cambios.
- Mantiene el valor original del atributo.
 - Los hechos siempre están asociados al valor original del atributo.
- Ejemplos
 - Atributos de la dimensión Tiempo
 - Datos del CUIT de la dimensión Clientes



SCD TIPO 1

Se reescribe el valor perdiendo el valor original

- No conserva la historia de cambios.
- Se modifica el dato en la tabla de dimensión (update).
- El atributo siempre contiene el último valor asignado.
- Es el más fácil de implementar.

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil
1	JC001	Jorge Cornejo	Soltero
2	CG015	Carlos González	Casado

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil
1	JC001	Jorge Cornejo	Casado
2	CG015	Carlos González	Casado



EJEMPLO SCD TIPO 1

ID_Producto	Cod_Producto	Desc_Producto	ID_TipoProducto	ID_RubroProducto
11	A101	Jeans	1	1

ID_Producto	Cod_Producto	Desc_Producto	ID_TipoProducto	ID_RubroProducto
11	A101	Jeans	1	2

- Evitar las inconsistencias en el DW.
- Si el atributo es un nivel por el que se suma (parte de una jerarquía), hay que recalcular agregaciones previas.



SCD TIPO 2

Se agrega una nueva fila o registro

- Cuando un atributo cambia de valor se agrega una nueva fila a la tabla de dimensión.
- Los nuevos hechos apuntan a la nueva fila (nueva SK).
- Los hechos anteriores continúan apuntando a la fila anterior.

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil
1	JC001	Jorge Cornejo	Soltero
2	CG015	Carlos González	Casado

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil
1	JC001	Jorge Cornejo	Soltero
2	CG015	Carlos González	Casado
3	<u>JC001</u>	<u>Jorge Cornejo</u>	<u>Casado</u>



SCD TIPO 2

- Es la técnica más usada.
- No requiere el recálculo de agregaciones.
- Permite guardar toda la historia de cambios.
- Los resultados previos al cambio siguen siendo los mismos.
- Se necesita colocar alguna *marca* o *fechas de vigencia* para saber cuándo se cambió o cuál se cambió primero.



EJEMPLOS SCD TIPO 2

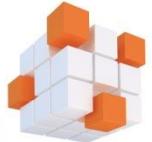
6/5/2025 IN2025

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil	FDesde	FHasta
1	JC001	Jorge Cornejo	Soltero	02/06/2005	10/09/2009
3	JC001	Jorge Cornejo	Casado	11/09/2009	05/02/2017
10	JC001	Jorge Cornejo	Divorciado	06/02/2017	

- Otra forma:

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil	Versión
1	JC001	Jorge Cornejo	Soltero	1
3	JC001	Jorge Cornejo	Casado	2
10	JC001	Jorge Cornejo	Divorciado	3

- Nota: si hay dudas en la elección del tipo SCD, es aconsejable elegir el 2.



SCD TIPO 3

Se agrega un nuevo atributo

- Guarda una cantidad limitada de valores históricos de atributos seleccionados.
- No incrementa el tamaño de la tabla.
- Sólo guarda el último valor
- Permite ver los datos recientes y los históricos por el *atributo actual* y el *anterior*.

ID_Producto	Cod_Producto	Desc_Producto	ID_TipoProducto	ID_RubroProducto	ID_RubroProductoAnt
11	A101	Jeans	1	1	-

ID_Producto	Cod_Producto	Desc_Producto	ID_TipoProducto	ID_RubroProducto	ID_RubroProductoAnt
11	A101	Jeans	1	2	1



EJEMPLO SCD TIPO 3

6/5/2025 IN2025

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil	FDesde	EstadoCivil_Ant
1	JC001	Jorge Cornejo	Soltero	02/06/2005	-

- Jorge se casa, entonces:

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil	FDesde	EstadoCivil_Ant
1	JC001	Jorge Cornejo	Casado	11/09/2009	Soltero

- Jorge se divorcia, entonces:

ID_Cliente	Cod_Cliente	NombreCliente	EstadoCivil	FDesde	EstadoCivil_Ant
1	JC001	Jorge Cornejo	Divorciado	06/02/2017	Casado

- **Nota:** Este tipo de SCD no es el más utilizado.



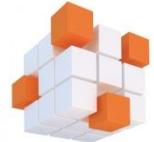
ATRIBUTOS CON DIFERENTE SCD

- Dado que la política SCD se define para cada atributo, una misma dimensión puede tener atributos con distinto tipo de SCD.
- Esto hay que tenerlo en cuenta al definir el ETL.



DIMENSIONES RÁPIDAMENTE CAMBIANTES

- Se denominan “Dimensiones Monstruo” (Rapidily Changing Monster Dimensions).
- Tablas de dimensiones con gran cantidad de filas.
- Los atributos descriptivos cambian frecuentemente.
- Causa un aumento desmedido de la cantidad de filas al usar SCD tipo 2.
- Son el tipo SCD 4.



SCD TIPO 4

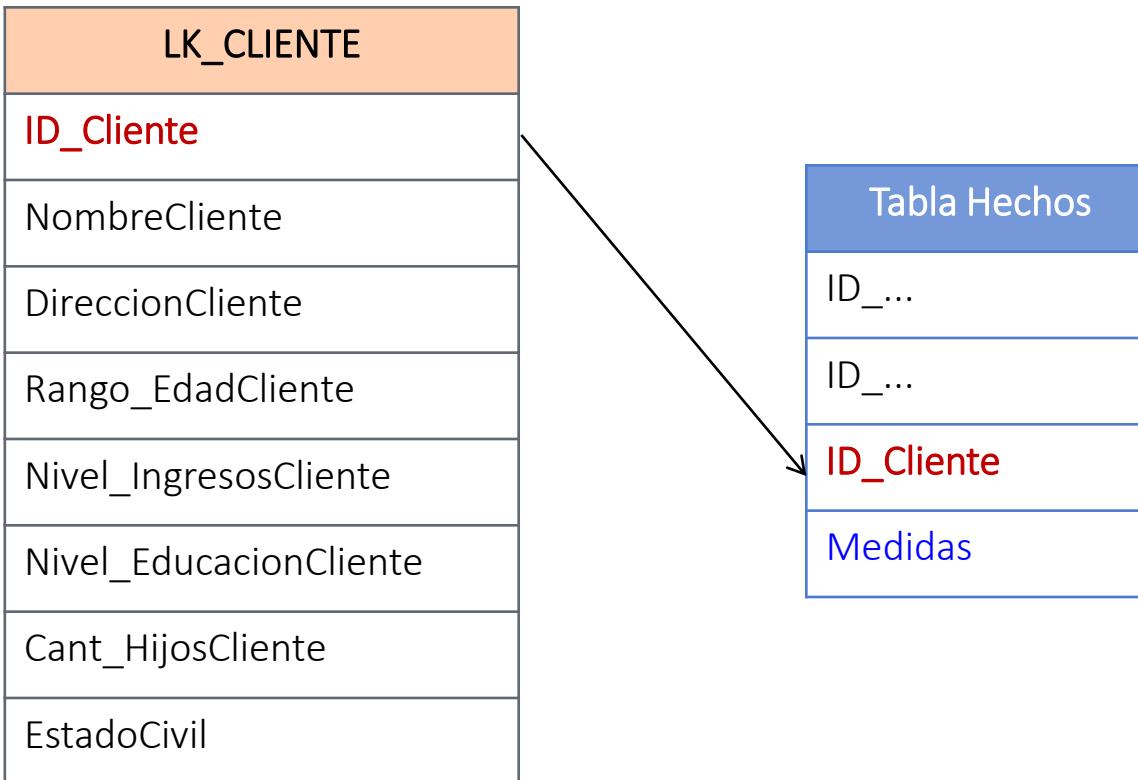
Se agrega una mini dimensión

- Abrir la dimensión original en 2 partes:
- Una dimensión con los atributos que cambian con frecuencia (mini dimensión).
- Otra dimensión con el resto de los atributos (dimensión primaria).



EJEMPLO SCD TIPO 4

- Dimensión Cliente con gran cantidad de filas.





EJEMPLO SCD TIPO 4

- Hay cambios de tipo sociodemográficos.
 - Rango de Edad
 - Nivel de Ingresos
 - Nivel de Educación
 - Cantidad de Hijos
 - Estado Civil
- Se crea una dimensión que agrupa a estos atributos.
 - Cada fila de esta tabla de dimensiones contiene un juego de valores posibles de cada uno de esos atributos.



EJEMPLO SCD TIPO 4

LK_CLIENTE
ID_Cliente
NombreCliente
DireccionCliente

Atributos
Estáticos

Tabla Hechos	
ID_...	
ID_...	
ID_Cliente	
ID_DemografiaCliente	
Medidas	

LK_DEMOGRAFIA
ID_DemografiaCliente
Rango_EdadCliente
Nivel_IngresosCliente
Nivel_EducacionCliente
Cant_Hijos
Estado_Civil

Atributos
Dinámicos



SCD TIPO 4

- La nueva dimensión contiene un conjunto de perfiles sociodemográficos.
- Tiene pocas filas.
- La tabla de hechos contiene:
 - La clave FK de la nueva dimensión.
 - La clave FK de la dimensión original.
 - Los hechos están asociados al valor que tenían los atributos de cambio rápido en el momento en que se generaron.
- **Recordar:** Transformar valores continuos en rangos
 - Cada atributo tendrá una cantidad pequeña de valores posibles.



DIMENSIÓN JUNK

- En las fuentes aparecen atributos con datos de baja cardinalidad (marcas por SI o NO).
 - Ejemplo: Encuestas con preguntas de Si / No
- Estos atributos no forman parte de ninguna dimensión.
- Los usuarios quieren conservarlos “**por las dudas**”.
- ¿Armamos una dimensión por cada uno de estos atributos?
 - Esto incide directamente en la tabla de hechos.
 - 10 dimensiones con indicadores Si / No $\Rightarrow 2^{10}$ registros.



DIMENSIÓN JUNK

- Una dimensión JUNK es una agrupación de atributos no relacionados que se llevan a una dimensión.
- Se utilizan para almacenar banderas o marcas.
- Solución
 - Definir una dimensión con las combinaciones de todos esos atributos y asignarle una clave subrogada a cada combinación.
 - Técnica similar a la de las mini dimensiones SCD tipo 4.
- Alternativas para la dimensión:
 - Guardar todas las combinaciones posibles.
 - Guardar sólo las que se presentan en los datos.



DIMENSIÓN JUNK

- Ejemplo alternativa para la dimensión:
 - Guardar todas las combinaciones posibles.

LK_Junk
ID_Junk
Promo1_IND
Descuento_IND

BT_Ventas
ID_Fecha
ID_Cliente
ID_Sucursal
ID_Producto
ID_Junk
Cant_UnidVendidas
Importe_Venta

- Combinación de Valores:

ID_Junk	Promo1_IND	Descuento_IND
1	S	S
2	S	N
3	N	S
4	N	N



DIMENSIÓN JUNK

- Otros ejemplos
 - Combinación de Valores:

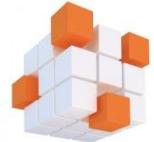
Estado del p...	1 2 3 ResellerSalesJunkKey	A ^B C Estado del pedido	A ^B C Estado de entrega
Presupuesto	1	Presupuesto	No entregado
Solicitado	2	Presupuesto	Entregado
Cancelado	3	Solicitado	No entregado
	4	Solicitado	Entregado
Estado de entr...	5	Cancelado	No entregado
Not Delivered	6	Cancelado	Entregado
Entregado			

DCA_KEY	DCA_COLOR	DCA_BODY_TYPE
1	BLACK	SEDAN
2	WHITE	SEDAN
3	RED	SEDAN
4	SILVER	SEDAN
...
101	BLACK	SUV
102	WHITE	SUV
103	RED	SUV
104	SILVER	SUV
...

Mystery_Dim_Key	Zortz	a3	uudl
1	TRUE	Confirmed	
2	TRUE	Confirmed	k
3	TRUE	Pending	
4	TRUE	Pending	k
5	FALSE	Confirmed	
6	FALSE	Confirmed	k
7	FALSE	Pending	
8	FALSE	Pending	k

DIM_JUNK

JUNK_ID	TXN_CODE	COUPON_IND	PREPAY_IND
1	1	Y	Y
2	2	Y	Y
3	3	Y	Y
4	1	Y	N
5	2	Y	N
6	3	Y	N
7	1	N	Y
8	2	N	Y
9	3	N	Y
10	1	N	N
11	2	N	N
12	3	N	N

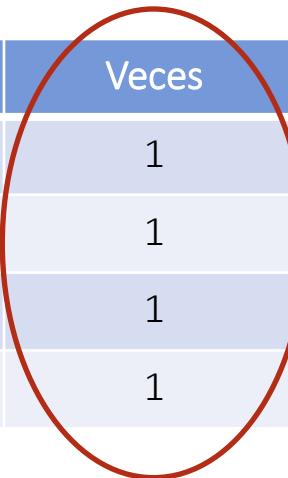


MEDIDAS O HECHOS - CASOS ESPECIALES

- La medida o hecho por sí solo no genera ninguna cantidad o valor para poder medir.
 - Ejemplo: Se necesita analizar la **cantidad de veces** que un usuario accede a un sitio web.
- En la tabla de hechos se podría agregar una medida “**veces**” con valor **1** para poder operar sobre ella.

BT_ACCESES

ID_Fecha	ID_Hora	ID_Usuario	Veces
15/04/2017	9:00	1	1
15/04/2017	9:30	2	1
15/04/2017	10:30	1	1
15/04/2017	14:00	3	1



Sumar

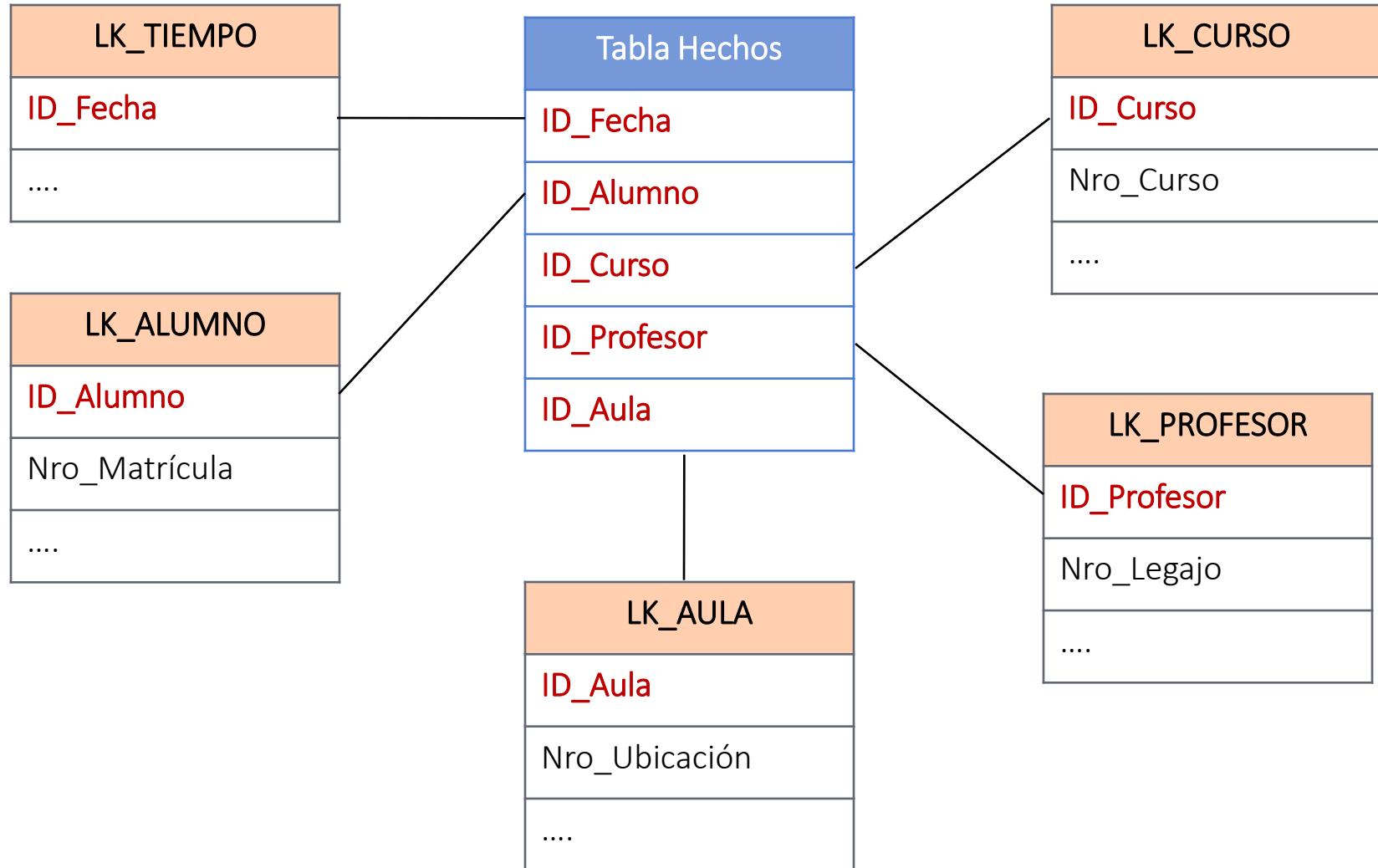


TABLAS DE HECHOS SIN MEDIDAS

- Factless Fact Table
- La tabla de hechos no tiene medidas, sino que registra un evento o acontecimiento que se da juntamente con todas las dimensiones en un determinado momento.
- Ejemplo
 - Asistencia de alumnos a distintos cursos
 - Censo
 - Atención médica
- No existen medidas numéricas para sumar.
- Se agregan contadores para facilitar las consultas.



EJEMPLO TABLAS DE HECHOS SIN MEDIDAS





TIPOS DE TABLAS DE HECHOS COMPLEMENTARIAS

Hay tres tipos fundamentales de tablas de hechos en el área de presentación del Data Warehouse:

- Tablas de hechos de transacciones
- Tablas de hechos de capturas periódicas
- Tabla de hechos de capturas acumulativa

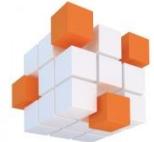




TIPOS DE TABLAS DE HECHOS COMPLEMENTARIAS

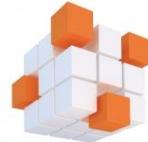
Hay tres tipos fundamentales de tablas de hechos en el área de presentación del Data Warehouse:

- **Tablas de hechos de transacciones**
- Son una visión fundamental y básica de las operaciones del negocio. Se utilizan para representar la ocurrencia de un evento en cualquier punto instantáneo del tiempo.
- Captura de detalles de eventos y actividades.
 - Transacciones de ventas | Inscripciones a cursos | Asistencia de estudiantes | Llamadas telefónicas
 - Fecha de la transacción, filtros FechaTrx (Desde/Hasta)
- **Tablas de hechos de capturas periódicas (Snapshot)**
- Ofrecen una visión del estado de los hechos en un momento determinado, es una «imagen del momento», una foto. (Snapshot). Miden el efecto de una serie de acontecimientos o actividades.
- La medición se realiza en un intervalo definido.
 - Saldos bancarios | Niveles de inventario | Matrículas/notas al final del trimestre
 - Fecha de la Foto, filtros FechaFoto (1 por vez)



TIPOS DE TABLAS DE HECHOS COMPLEMENTARIAS

- **Tabla de hechos acumulativa de instantáneas**
- Resume los eventos que ocurren en pasos predecibles entre el comienzo y el final de un proceso. Para realizar el seguimiento de una entidad a lo largo del tiempo.
- Permiten medir los tiempos de retraso entre cada paso del proceso.
- A veces hay una columna adicional que contiene la fecha que muestra cuándo se actualizó la fila por última vez.
 - Estudiantes a medida que cursan una carrera | Solicitudes de hipotecas | Compras desde el pedido hasta la entrega
 - Fechas de Estados X, filtros FechaEstadoX (Desde/Hasta)



TIPOS DE TABLAS DE HECHOS COMPLEMENTARIAS

	Transaction	Periodic Snapshot	Accumulating Snapshot
Periodicity	Discrete transaction point in time	Recurring snapshots at regular, predictable intervals	Indeterminate time span for evolving pipeline/workflow
Grain	1 row per transaction or transaction line	1 row per snapshot period plus other dimensions	1 row per pipeline occurrence
Date dimension(s)	Transaction date	Snapshot date	Multiple dates for pipeline's key milestones
Facts	Transaction performance	Cumulative performance for time interval	Performance for pipeline occurrence
Fact table sparsity	Sparse or dense, depending on activity	Predictably dense	Sparse or dense, depending on pipeline occurrence
Fact table updates	No updates, unless error correction	No updates, unless error correction	Updated whenever pipeline activity occurs



EJEMPLO TABLAS DE HECHOS COMPLEMENTARIAS

Fact_Sales	
P * Date_of_Sale_Key	
P * Customer_Key	
P* Product_Key	
P* Store_Key	
Quantity_Sale	
Amount_Sale	

1

Monthly_sales	
P * Month_Date_key	Number
P* Customer_Key	Number
P * Product_Key	Number
P * Store_Key	Number
Quantity_Sale	Number
Amount_Sale	Number

2

Fact_Claim	
P * Claim_Key	Number
P * Customer_Key	Number
P * Claim_Date	Date
P * Policy-Key	Number
Investigation_date	Date
Review_date	Date
Payment_Date	Date

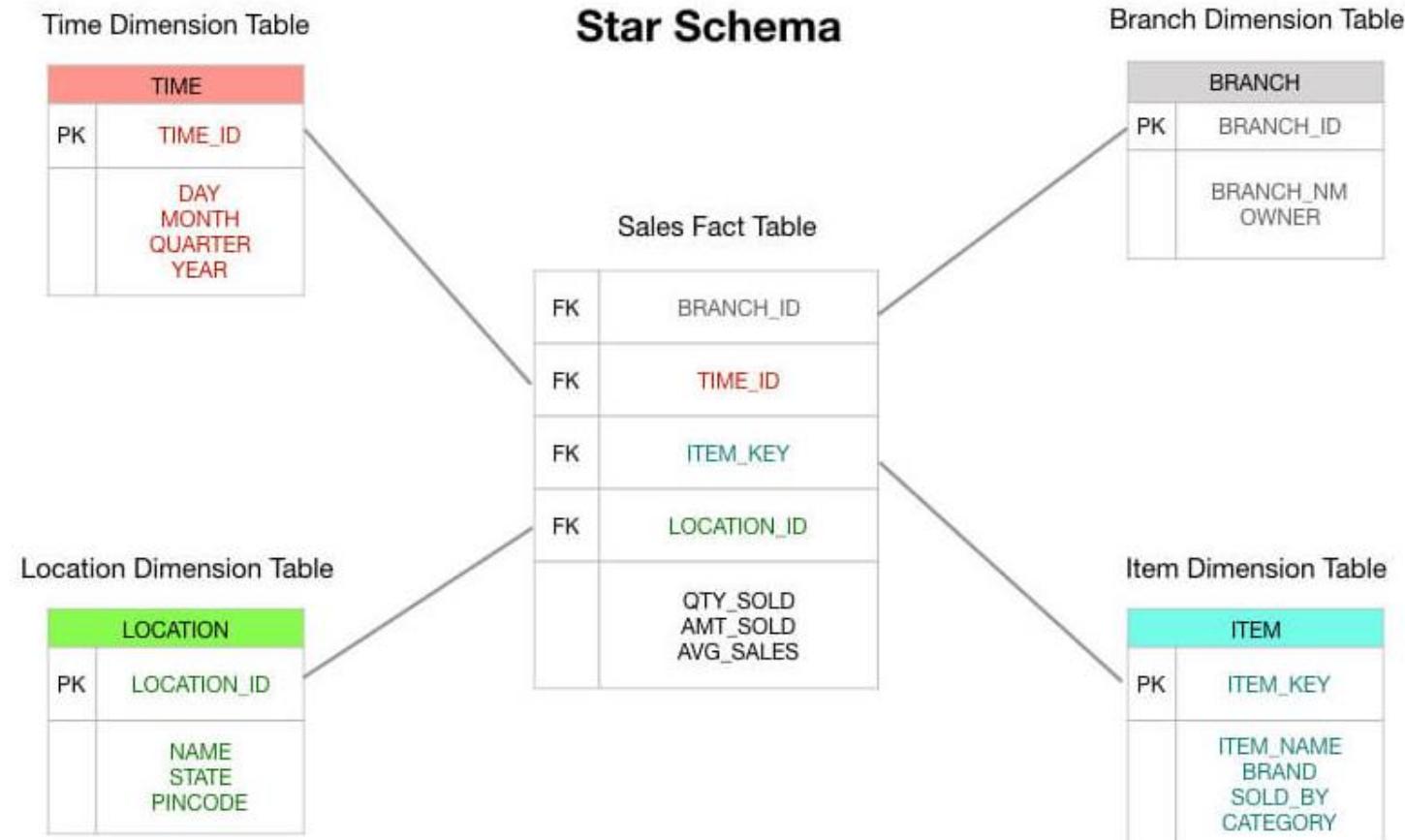
3





EJEMPLOS FT TRANSACCIONES

1



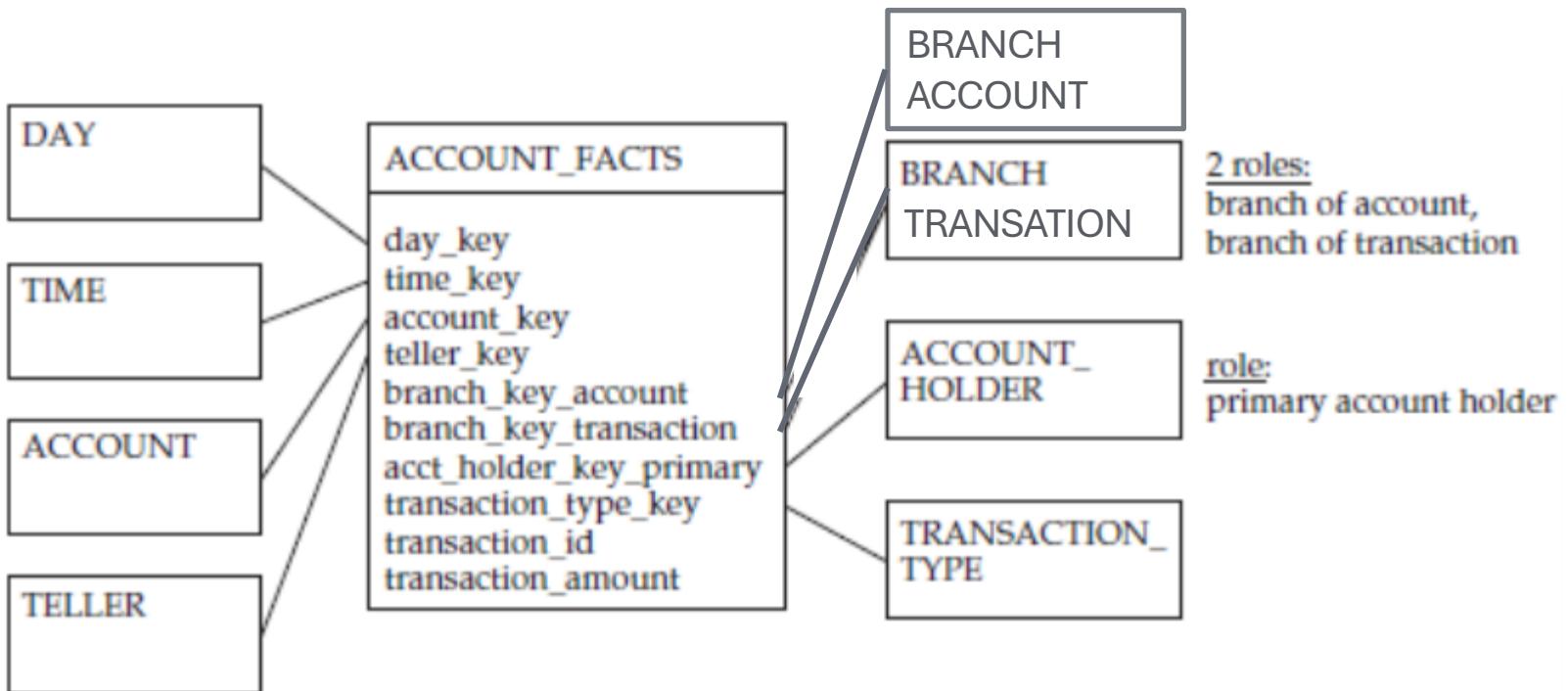


1

EJEMPLOS FT TRANSACCIONES

Bank account example

A schema with a transaction fact table:





EJEMPLOS FT TRANSACCIONES

1

Example of transactions in transaction fact table for account
7922-3002

From 2/1/2009 to 2/14/2009

Granular transaction data stored in star:

Day	Transaction Type	Transaction Amount
2/1/2009	Initial Deposit	2000.00
2/2/2009	Withdrawal	(20.00)
2/3/2009	Check	(35.50)
2/3/2009	Check	(17.02)
2/6/2009	Check	(75.00)
2/6/2009	Deposit	75.00
2/7/2009	Check	(800.00)
2/10/2009	Check	(68.29)
2/14/2009	Withdrawal	(100.00)



Same day

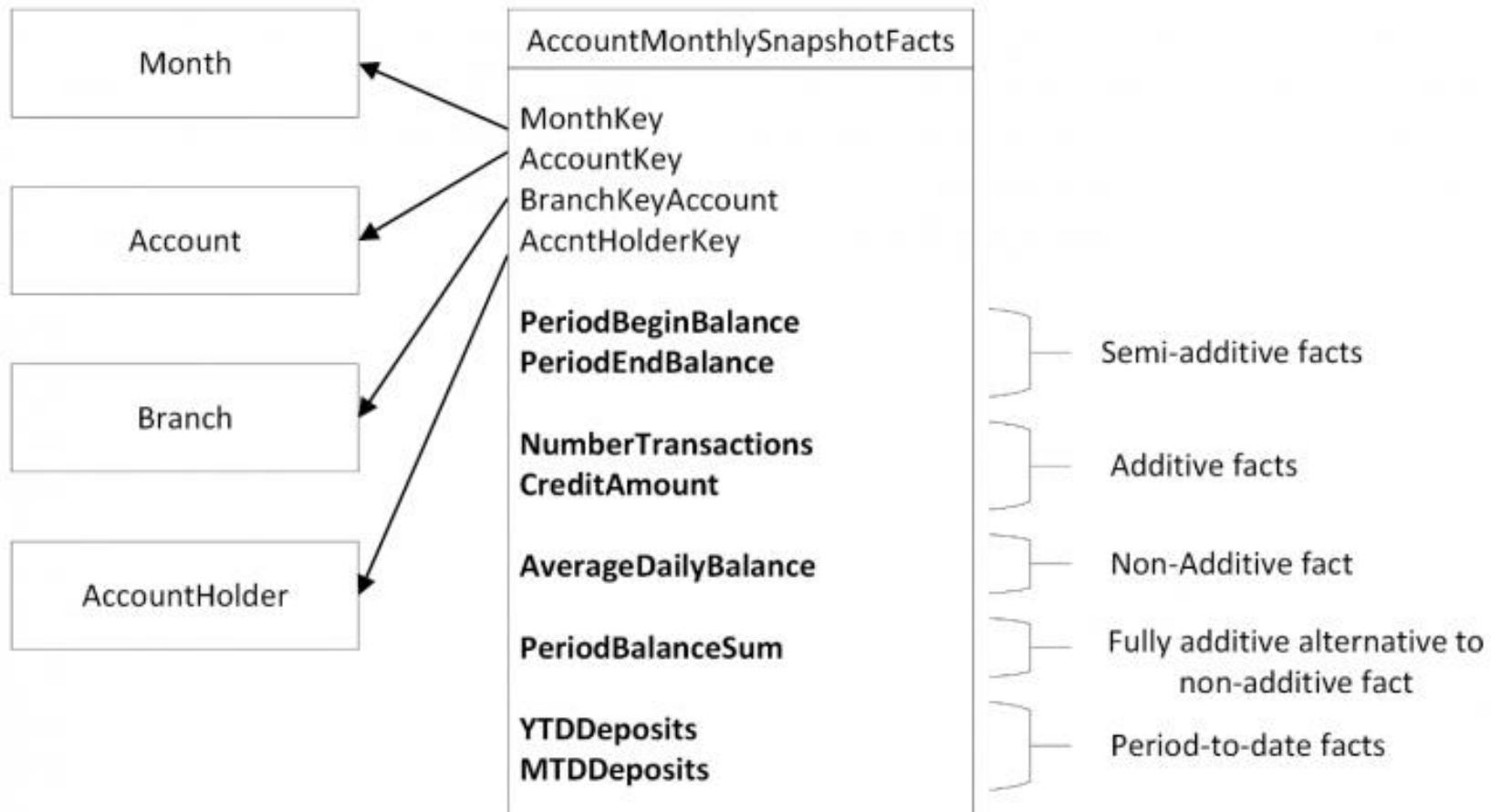
Same day

What is the balance
on Feb 9?



EJEMPLOS FT CAPTURA PERIÓDICA

2



EJEMPLOS FT CAPTURA PERIÓDICA

2



7/5/2025 IN2025

Consider a snapshot fact table

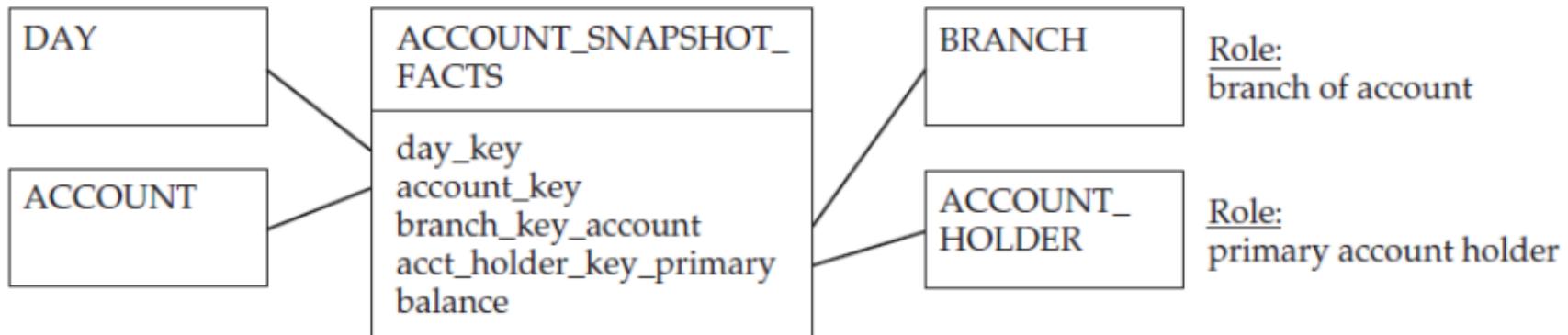


Figure 11-2 A snapshot fact table tracks account status

The balance of each account is recorded in the fact table at the end of each day. There is a measurement for each account on each day – more dense than the transaction fact table



3

EJEMPLO FT CAPTURA ACUMULATIVA

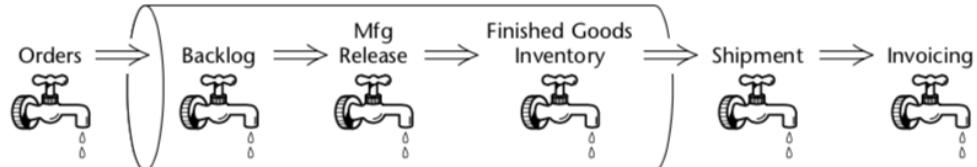


Figure 5.9 Order fulfillment pipeline diagram.

Para ese proceso, Kimball propuso la siguiente tabla de instantáneas acumulativas. Cada fila de la tabla representa un pedido o un lote de pedidos.

Se espera que cada una de estas filas se actualice varias veces a medida que avanzan a través de la canalización de cumplimiento de pedidos. Notar la gran cantidad de campos de fecha en la parte superior del esquema (roles de la Dimensión Tiempo – atributo Fecha).

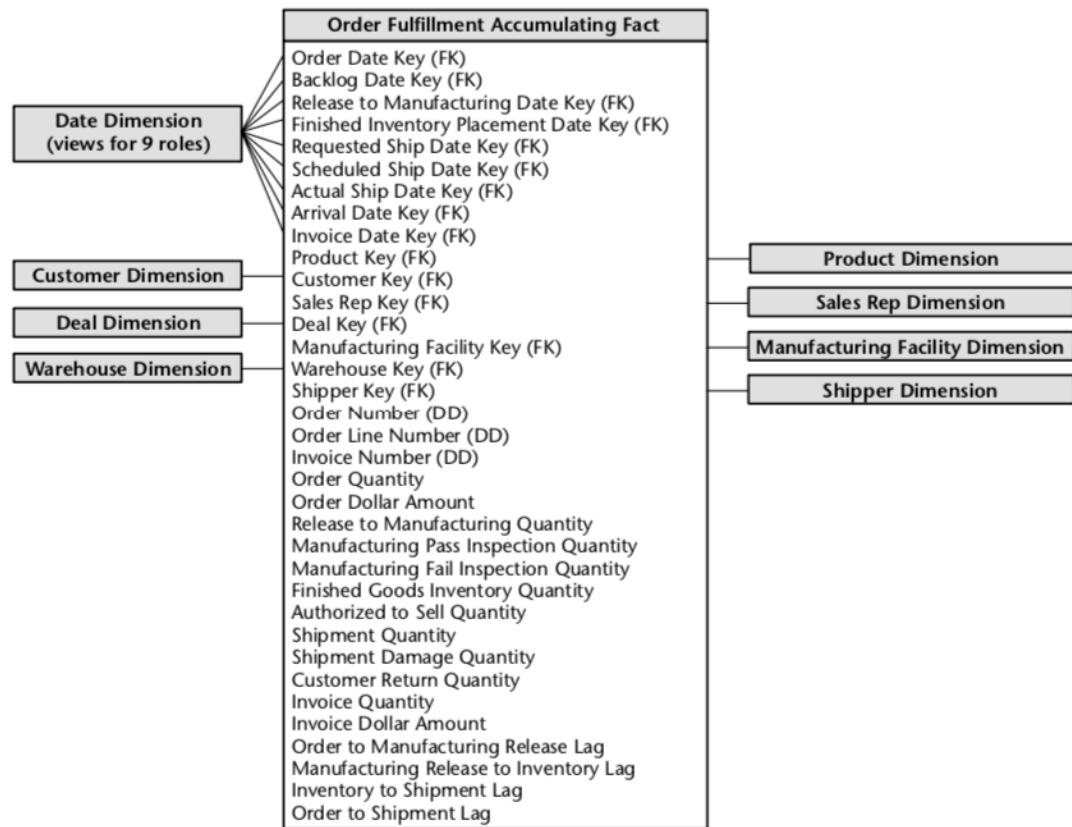
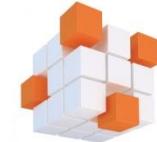


Figure 5.10 Order fulfillment accumulating snapshot fact table.

Fuente: <https://www.holistics.io/blog/the-three-types-of-fact-tables/>



EJEMPLO FT CAPTURA ACUMULATIVA

3

6/5/2025 IN2025

Accumulating Snapshot Fact Tables

transactions

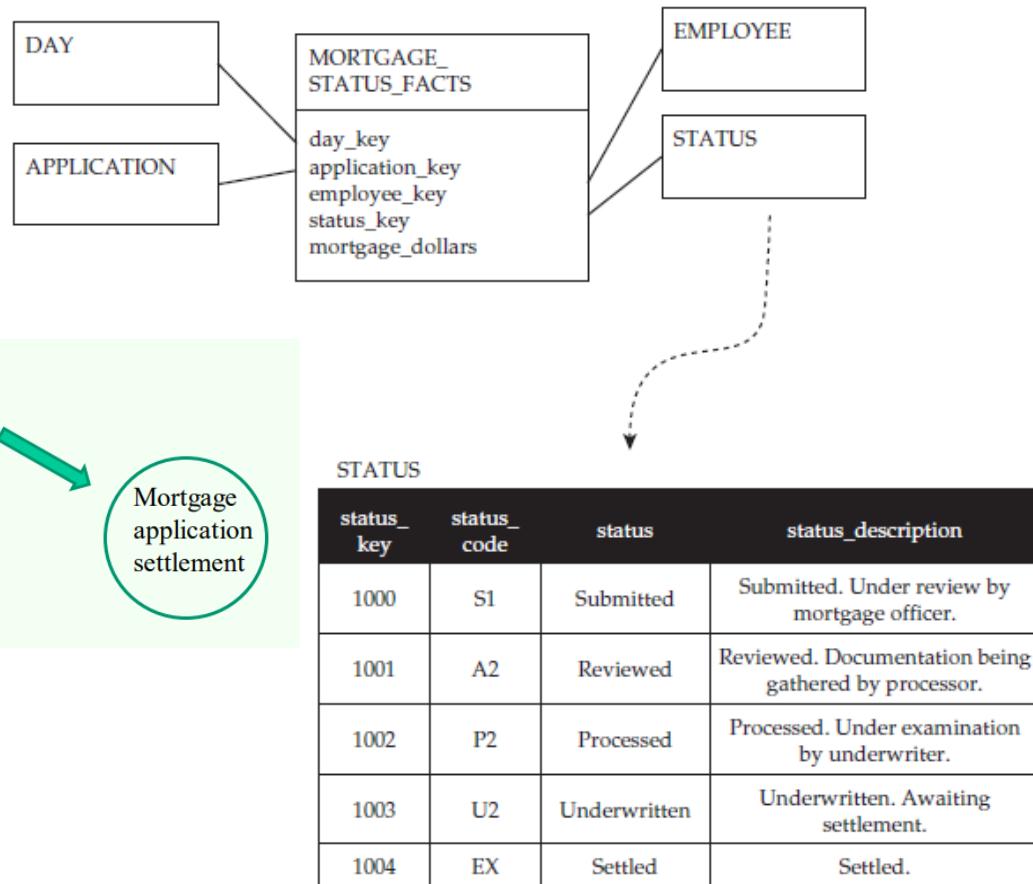


Figure 11-5 A transaction fact table captures status changes

Fuente: <https://courses.acs.uwinnipeg.ca/4904-001/ch11%20Fact%20table%20types.pdf>

46



EJEMPLO FT CAPTURA ACUMULATIVA

3

Tracking
progress via
accumulating
snapshot:
One row per
application

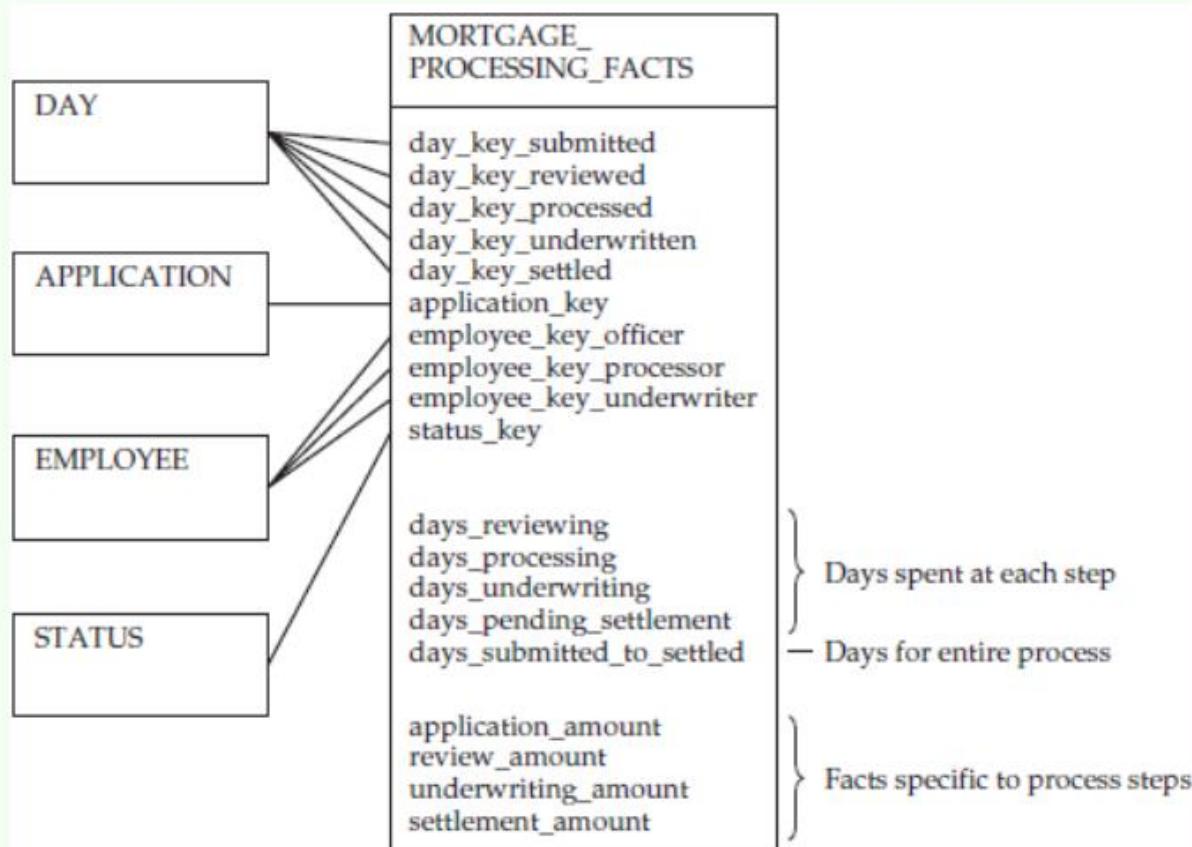


Figure 11-6 An accumulating snapshot with one row per application

Fuente: <https://courses.acs.uwinnipeg.ca/4904-001/ch11%20Fact%20table%20types.pdf>



3

EJEMPLO FT CAPTURA ACUMULATIVA

Each time an application reaches a milestone a row can be updated

Accumulating Snapshot Fact Tables

On Day 1 (Submitted; under review by officer):

application_key	day_key_submitted	day_key_reviewed	day_key_processed	day_key_underwritten	day_key_closing	application_amount	review_amount	underwriting_amount	days_reviewing	days_processing	--
1011	1021	0000	0000	0000	0000	100,000	0	0	0	0	--

Day 2 (No status change):

application_key	day_key_submitted	day_key_reviewed	day_key_processed	day_key_underwritten	day_key_closing	application_amount	review_amount	underwriting_amount	days_reviewing	days_processing	--
1011	1021	0000	0000	0000	0000	100,000	0	0	1	0	--

Days 3–9 (not shown)...

Day 10 (Reviewed; documents being gathered by processor):

application_key	day_key_submitted	day_key_reviewed	day_key_processed	day_key_underwritten	day_key_closing	application_amount	review_amount	underwriting_amount	days_reviewing	days_processing	--
1011	1021	1031	0000	0000	0000	100,000	90,000	0	9	0	--

Day 11 (No status change):

application_key	day_key_submitted	day_key_reviewed	day_key_processed	day_key_underwritten	day_key_closing	application_amount	review_amount	underwriting_amount	days_reviewing	days_processing	--
1011	1021	1031	0000	0000	0000	100,000	90,000	0	9	1	--

Remaining steps...

Figure 11-7 Evolution of a row in an accumulating snapshot



3

EJEMPLO FT CAPTURA ACUMULATIVA

Using the accumulating snapshot e.g., average processing time for applications in January 2023:

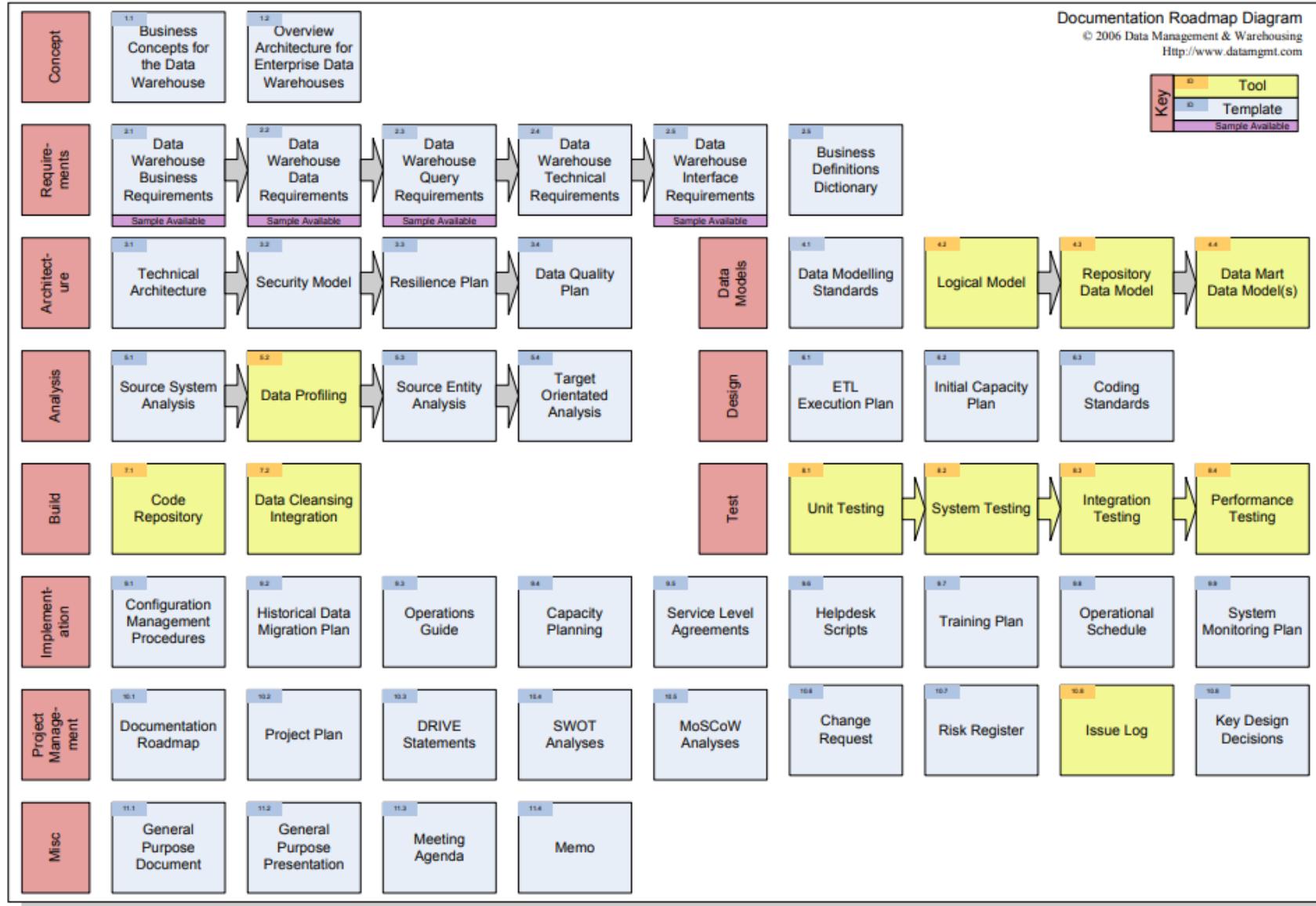
```
SELECT avg( days_processing )
FROM mortgage_processing_facts, day
WHERE mortgage_processing_facts.day_key_processed = day.day_key
      AND day.month = "January"
      AND day.year = 2023
```

Using the accumulating snapshot e.g., the average time spent reviewing and processing an application in January 2023:

```
SELECT avg( days_reviewing + days_processing )
FROM mortgage_processing_facts, day AS day_processed
WHERE      mortgage_processing_facts.day_key_processed = day_processed.day_key      AND
day_processed.month = "January"
      AND day_processed.year = 2023
```



DATA WAREHOUSE DOCUMENTATION ROADMAP



TP CASO ESTUDIO: PLATAFORMA STREAMING SOUNDWAVE

ENTREGA OPCIONAL (*)



- Listar las posibles dimensiones a partir de las necesidades de información enunciadas.
- Identificar los hechos con las medidas básicas y calculadas.
- Identificar el proceso de negocio a modelar.
- Construir el modelo dimensional conceptual.
- Identificar atributos y jerarquías para cada dimensión.
- Construir el modelo dimensional lógico.
- Construir el modelo físico (en esquema Estrella y Copo de Nieve).

(*)

- Los TPs Opcionales sirven para afianzar conceptos necesarios para realizar los TPs de Aplicación, serán corregidos en clase y/o mediante Autoevaluación.
- Los TPs de Aplicación tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción, serán corregidos en detalle por los docentes.
- Ver documentos: "Condiciones de Cursada en MIeL - sección: Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio" + "Círculo Entrega TPs Teams/MIeL"



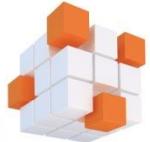
TRABAJO PRÁCTICO DE APLICACIÓN (PARTE2)

ENTREGA PRÓXIMA CLASE (*)

Parte 2 (Modelo Físico)

- Construir el modelo físico (en esquema Estrella y Copo de Nieve).
- Identificar en cada uno de estos modelos físicos cada una de las dimensiones del modelo dimensional.
- Identificar y/o aplicar las estrategias necesarias en el modelo físico del DWH para gestionar cambios históricos dimensionales (SCD).

(*) Cada trabajo práctico tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción. (Ver Condiciones de Cursada en MIeL – Sección: "Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio")



TRABAJO PRÁCTICO DE APLICACIÓN (PARTE3)

ENTREGA PRÓXIMA CLASE + 1 (*)

Parte 3 (Power BI Desktop)

- Construya el Modelo de Inteligencia de Negocios del caso propuesto. Se proveerán las fuentes de datos correspondientes.
- Construya diferentes Tableros, Reportes, Gráficos, Filtros que permitan dar respuesta a los interrogantes de información planteados por la empresa.
- Cree una página de Sugeridos con visualizaciones que mejor respondan a las nuevas preguntas detectadas en la Parte 1.

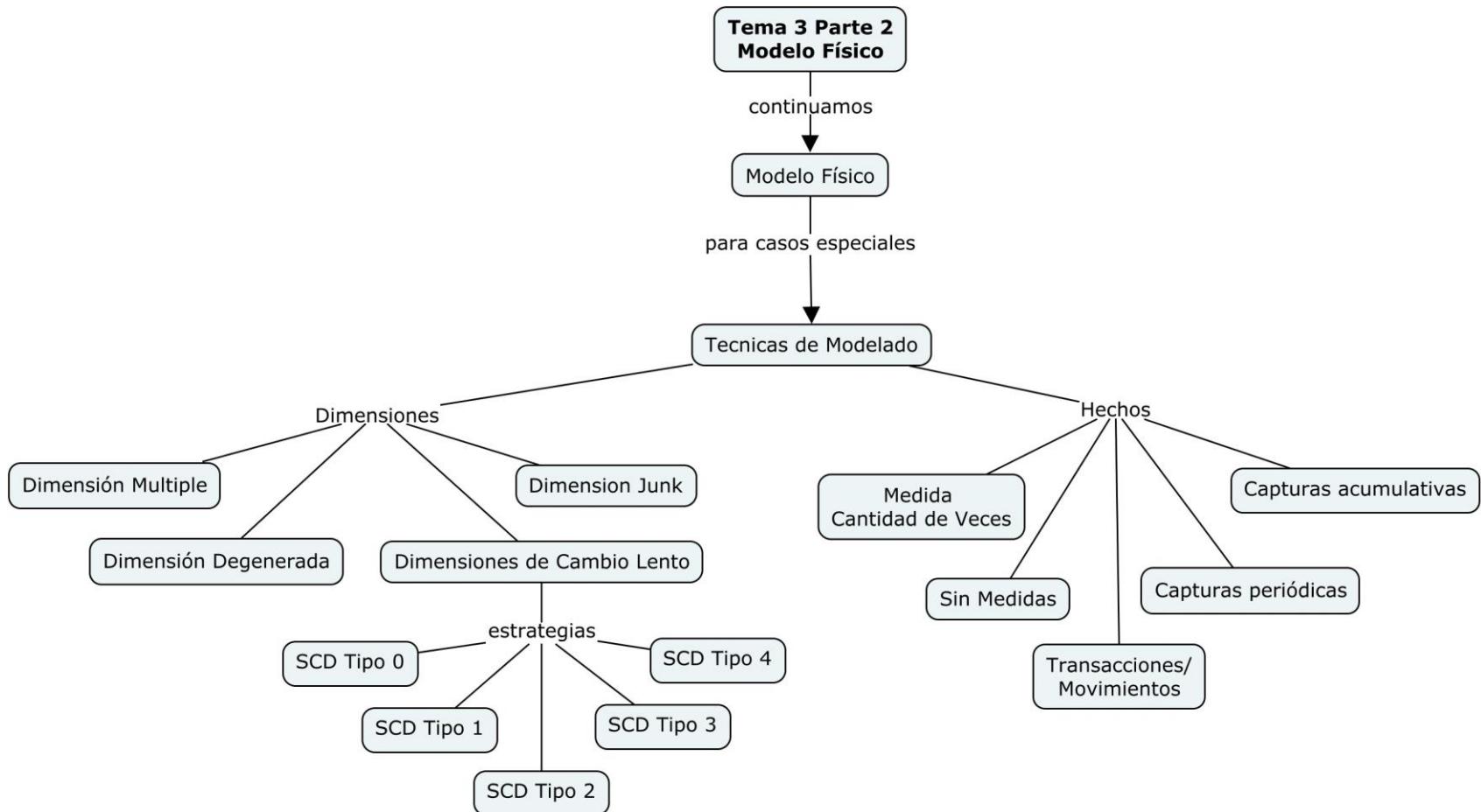
Links de Interés:

- Obtener Power BI Desktop - [Power BI | Microsoft Docs](#)
- Documentación de Power BI - [Power BI | Microsoft Docs](#)
- Power BI en Microsoft Learn | [Microsoft Docs](#)
- Documentación de introducción a Power BI - [Power BI | Microsoft Docs](#)
- Blog de Power BI: actualizaciones y novedades | [Microsoft Power BI](#)
- Galleries - [Microsoft Power BI Community](#)

(*) Cada trabajo práctico tiene una Fecha Límite de Entrega que deberá ser cumplida sin excepción. (Ver Condiciones de Cursada en MIeL – Sección: "Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio")



RESUMEN TEMA 3 PARTE 2







UNIVERSIDAD NACIONAL DE LA MATANZA
INTELIGENCIA DE NEGOCIOS

Material Power BI

Docentes: ING. LORENA R. MATTEO

Curso IN2025

Fecha última actualización.: 06/05/2025



SOFTWARE A UTILIZAR PARA LA ASIGNATURA

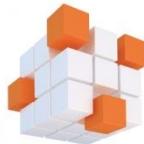
Microsoft Power BI

Herramienta para crear y explotar modelos de datos e informes de inteligencia empresarial

Enlaces:

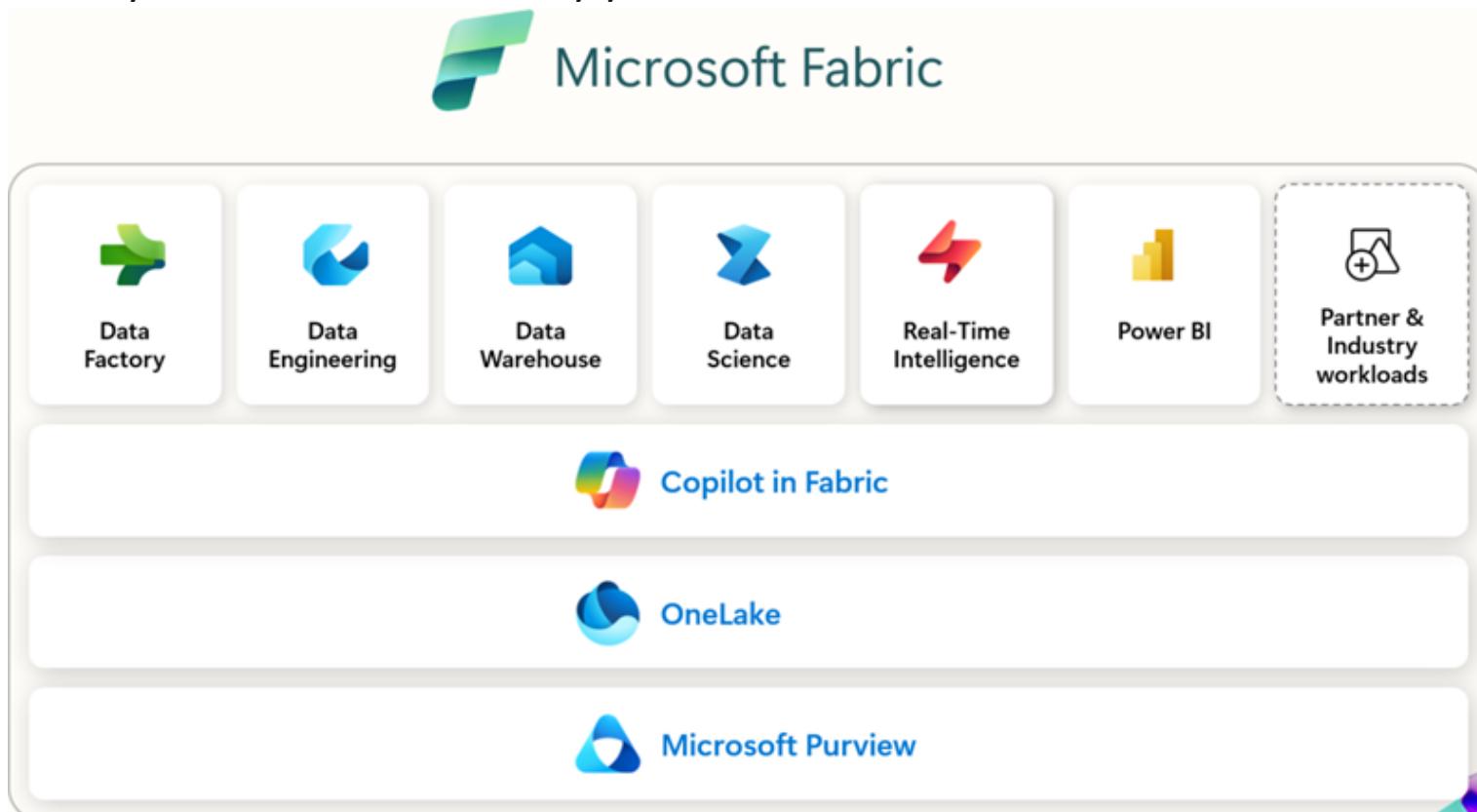
[Obtener Power BI Desktop - Power BI | Microsoft Docs](#)

[Documentación de Power BI - Power BI | Microsoft Docs](#)



POWER BI – PARTE DE LA FAMILIA MICROSOFT FABRIC

- Microsoft Fabric se basa en una plataforma de software como servicio (SaaS).
- Unifica los componentes nuevos y existentes de Power BI, Azure Synapse Analytics, Azure Data Factory y mucho más en un único entorno.





¿QUÉ ES POWER BI?

- Power BI es una colección de servicios de software, aplicaciones y conectores que funcionan conjuntamente para convertir orígenes de datos sin relación entre sí en información coherente, interactiva y atractiva visualmente.

Componentes de Power BI

- Power BI consta de varios elementos que funcionan de manera conjunta, empezando por estos tres conceptos básicos:
 - Aplicación de escritorio de Windows llamada **Power BI Desktop**.
 - Servicio de software como servicio (SaaS) en línea denominado **servicio Power BI**.
 - Aplicaciones para **Power BI Mobile** para dispositivos Windows, iOS y Android.



¿QUÉ ES POWER BI?

Además, Power BI también incluye otros dos entornos:

- **Power BI Report Builder**, para crear informes paginados y compartirlos en el servicio Power BI.
En el siguiente link pueden obtener más información sobre los informes paginados.
- **Power BI Report Server**, un servidor de informes local en el que puede publicar los informes de Power BI, después de crearlos en Power BI Desktop.
En el siguiente link pueden obtener más información sobre Power BI Report Server.



¿QUÉ ES POWER BI? - HERRAMIENTAS

The diagram illustrates the three components of Microsoft Power BI:

- Power BI Desktop**: The desktop application where you build your data models and reports. It shows a dashboard with various visualizations like a bar chart and a scatter plot.
- Power BI service**: The cloud-based service where your reports are published and shared. It shows a similar dashboard to the desktop version, with a bar chart and a scatter plot.
- Power BI Mobile**: The mobile application for iOS and Android, allowing you to access your reports on the go. It shows a mobile screen with a dashboard featuring a bar chart and a pie chart.

Yellow arrows indicate the flow from the desktop application to the service, and from the service to the mobile application.



¿QUÉ ES POWER BI? - HERRAMIENTAS

The image displays three separate Power BI dashboards arranged diagonally.

- Store Dashboard:** A dashboard for a bicycle store. It includes a navigation sidebar with filters for Location (Canada, United Kingdom, England, United States), Company (A Bike Store, A Cycle Shop, A Great Bicycle Company, A Typical Bike Shop, Acceptable Sales & Service, Action Bicycle Specialists), and Product (Accessories, Bib-Shorts, Bike Racks, Bike Stands, Bikes, Bottles and Cages). The main area shows a summary card with Sales (\$956.30K!), Orders (32), and Customers (32). Below are charts for Sales by Month, Orders by Month, and Customers by Month. A Q&A feature is present at the bottom.
- Corporate Spend Sample PBIX Dashboard:** A YTD IT Spend Trend Analysis dashboard. It features a 3D bar chart showing Actual and Plan by IT Area (Infrastructure, IT Support, Functional IT Area, Government, OEM). A map of Europe shows Var Plan % by Country/Region. A line chart tracks Actual by Period and Business Area (Distribution, Infrastructure, Marketing).
- COVID-19 Cases Dashboard:** An interactive map of the United States showing Confirmed Cases and Total Deaths by state. A color scale indicates case density from low (blue) to high (red). Below the map are two line charts: Daily Cases and Daily Deaths, showing trends from January 2021 to February 2023. A methodology section provides details about data sources and frequency.



MICROSOFT POWER BI TUTORIAL

[Tutorial: De modelo dimensional a impresionante informe en Power BI](#)

[Desktop - Power BI | Microsoft Learn](#)

The screenshot shows the Microsoft Learn website with a dark header. The header includes the Microsoft logo, a 'Learn' button, and navigation links for 'Descubra', 'Documentación del producto' (underlined), 'Lenguajes de desarrollo', and 'Temas'. Below the header, a navigation bar for 'Power BI' has links for 'Introducción', 'Trabajar con datos', 'Crear y compartir', 'Administración y empresas', 'Developer', and 'Recursos'. On the left, a sidebar lists articles under 'Información general', 'Visita al editor de informes', '¿Qué son los informes paginados?', and 'Tutoriales'. The 'Power BI Desktop: modelo dimensional a un informe' article is highlighted with a red border. The main content area displays the title 'Tutorial: De modelo dimensional a impresionante informe en Power BI Desktop', the date 'Artículo • 29/05/2024 • 8 colaboradores', and a section 'En este artículo' with links to 'Requisitos previos', 'Obtención de datos: Descarga del ejemplo', 'Preparar los datos', 'Modelar los datos', and 'Mostrar 7 más'. At the bottom left, there's a link 'Descargar PDF'.

El futuro es suyo

Regístrate ahora >

Microsoft Build - 19–22 may 2025

Learn | Microsoft

Power BI Introducción Trabajar con datos Crear y compartir Administración y empresas Developer Recursos

Filtrar por título

Learn / Power Platform / Power BI /

Tutorial: De modelo dimensional a impresionante informe en Power BI Desktop

Artículo • 29/05/2024 • 8 colaboradores

En este artículo

- Requisitos previos
- Obtención de datos: Descarga del ejemplo
- Preparar los datos
- Modelar los datos
- Mostrar 7 más

Power BI Desktop: de Excel a un informe

Power BI Desktop: modelo dimensional a un informe

Servicio Power BI: de Excel a un informe

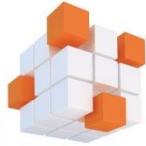
Exploración de un ejemplo de Power BI

Crea un árbol de descomposición

Curso en línea: Informes paginados

Creación y carga de un informe paginado

Descargar PDF



MICROSOFT POWER BI TUTORIAL

Obtención de ejemplos para Power BI - Power BI | Microsoft Learn

The screenshot shows a Microsoft Learn page for the 'Power BI' section. At the top, there's a dark banner with the text 'El futuro es suyo' and 'Regístrate ahora >'. Below the banner, the Microsoft Build date 'Microsoft Build - 19–22 may 2025' is visible. The main navigation bar includes 'Learn', 'Descubra', 'Documentación del producto', 'Lenguajes de desarrollo', and 'Temas'. Under 'Power BI', sub-navigation links include 'Introducción', 'Trabajar con datos', 'Crear y compartir', 'Administración y empresas', 'Developer', and 'Recursos'. A search bar with the placeholder 'Filtrar por título' is present. The breadcrumb navigation shows 'Learn / Power Platform / Power BI /'. The main article title is 'Obtención de ejemplos para Power BI', with a subtitle 'Artículo • 20/12/2024 • 11 colaboradores'. To the right, there are 'Comentarios' and other sharing icons. The article content starts with 'En este artículo' and lists several sections: 'Archivo .pbix de ejemplo de ventas y devoluciones', 'Muestras actualizadas', 'Ocho ejemplos originales', 'Instalación de los ejemplos integrados', and 'Mostrar 4 más'. Below this, under 'SE APLICA A:', 'Power BI Desktop' and 'Servicio Power BI' are checked. The main text explains that the article provides examples for Power BI users who want to learn its features. It highlights various types of examples like sales reports, AI models, and regional sales. A sidebar on the left lists all available examples, with 'Obtención de ejemplos para Power BI' being the selected one. At the bottom of the sidebar, there's a 'Descargar PDF' link.

7/5/2025

IN2025

10



VERSIONES POWER BI Y LICENCIAMIENTO

6/5/2025

IN2025

Características principales



- 1
- 2
- 3
- 4

Es una herramienta que encontramos en nuestro licenciamiento de Office 365 de manera **gratuita**

Tiene **3 versiones principales** el producto: Power BI Desktop, Power BI Service y Power BI Mobile. Cada una con un objetivo único

Tiene **3 licenciamientos principales**: Power BI Free, Power BI Pro y Power BI Premium por capacidad o por usuario

No necesitas ser ingeniero de datos, solo conocer tu información e **interactuar** con la herramienta



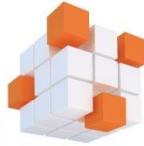
EXCEL VS. POWER BI



- Pensado para informes tabulares
- Cantidad de datos limitada
- Alto mantenimiento para la actualización
- Cantidad limitada de gráficos
- Conexión limitada a otras aplicaciones

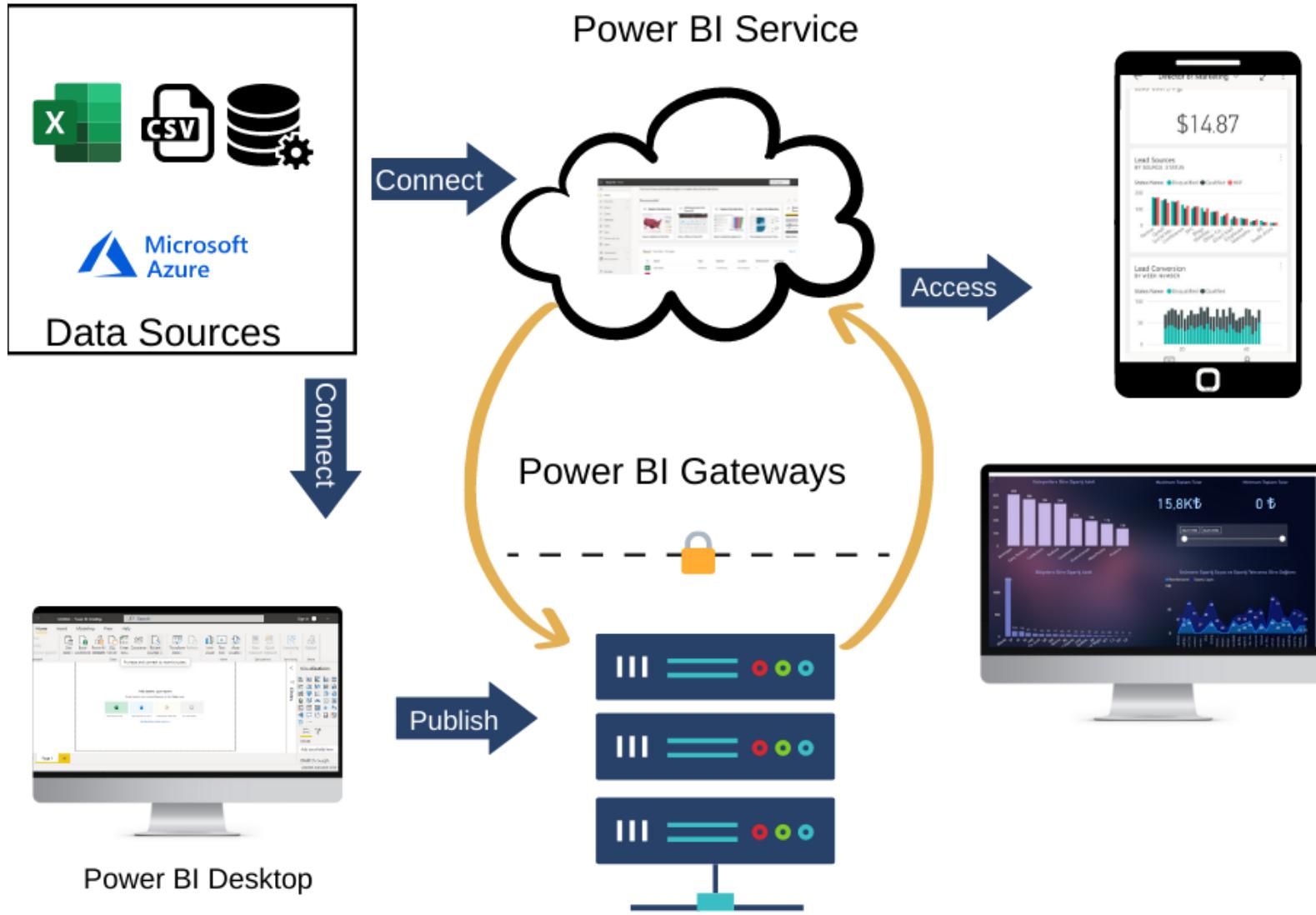


- No pensado para informes tabulares
- Cantidad de datos ilimitada
- Actualización automática
- Mayor cantidad de gráficos
- Conexión a casi cualquier aplicación, BD y herramientas.



ARQUITECTURA BÁSICA

6/5/2025 IN2025





COMPONENTES POWER BI

6/5/2025

IN2025

○ Data Sources

Power BI supports a wide range of data sources. Data can be imported from cloud-based online data sources or directly connected to live connections. There is a 1GB limit when importing data from on-premises or online services.

○ Power BI Desktop

This free platform allows you to connect to data sources, clean and transform data, model data, and report on it.

○ Power BI Gateway

Gateways aid in the extraction of data to Power BI platforms via secure channels for analysis and reporting in organizations where data is secure and monitored. It is typically used to connect to and access on-premise data.

○ Power BI Mobile

These apps are used to view views like reports and dashboards on mobile.

○ Embedded Power BI

Offers APIs for embedding images into custom applications.

○ Power BI Service

It is a web-based cloud platform for sharing Power BI Desktop reports, collaborating with other users, and creating dashboards.

○ Power BI Service Architecture

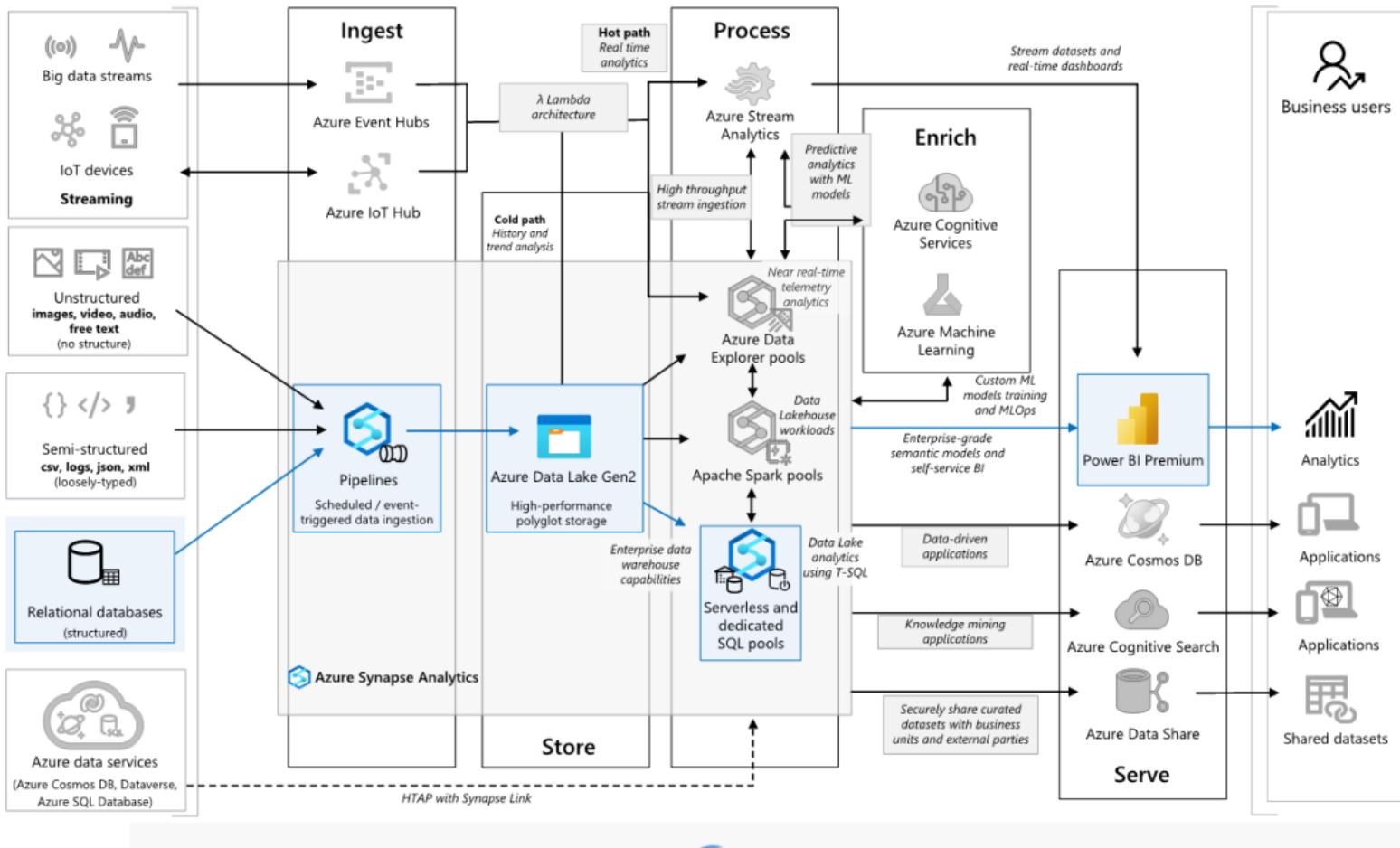
Microsoft's Azure cloud computing infrastructure and platform serve as the foundation for the Power BI service. Power BI's service architecture is divided into two groups: the Web Front-End cluster and the Backend cluster.

- The Web Front-End is in charge of the initial connection and authentication to the Power BI service.
- The Back-End is in charge of all subsequent user interactions.

14



ARQUITECTURA BI MICROSOFT



Platform



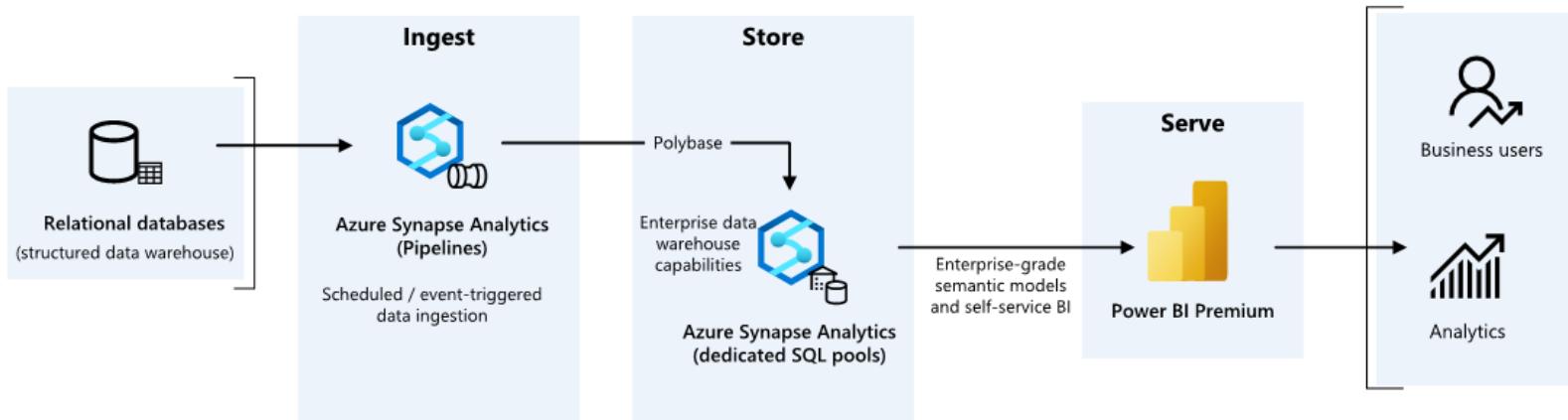
15



ARQUITECTURA BI MICROSOFT SIMPLIFICADA

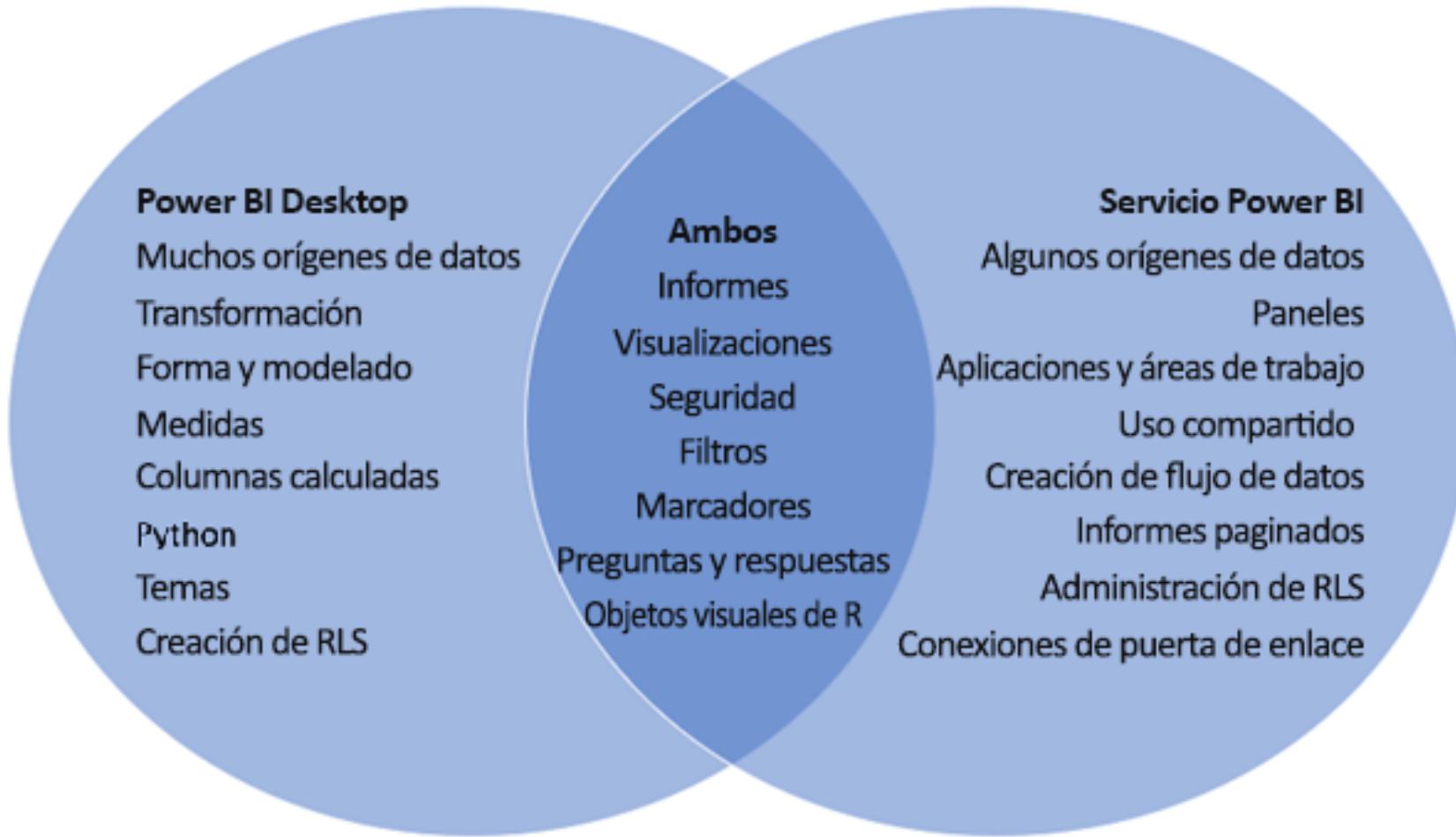
6/5/2025

IN2025





DIFERENCIAS POWER BI DESKTOP VS SERVICE





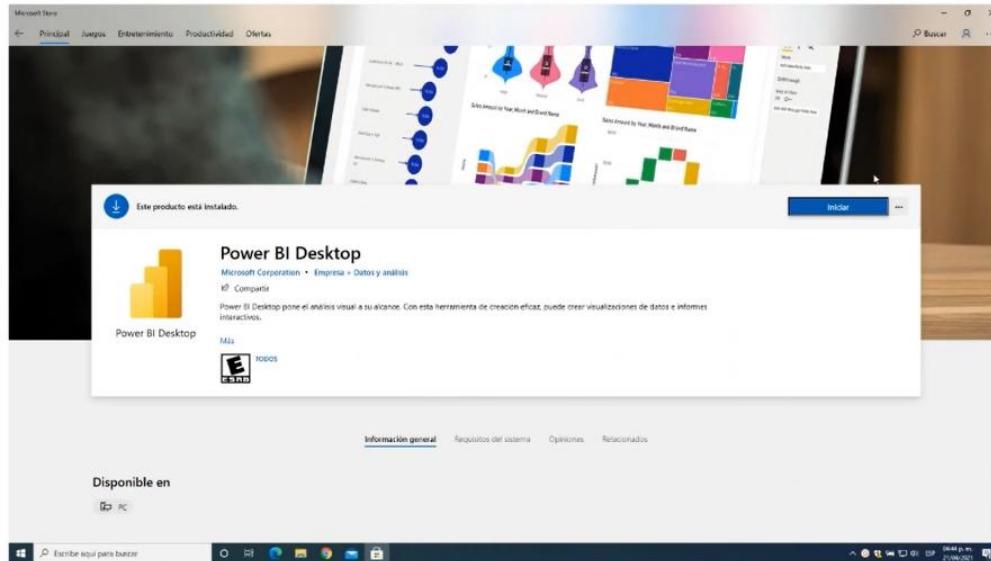
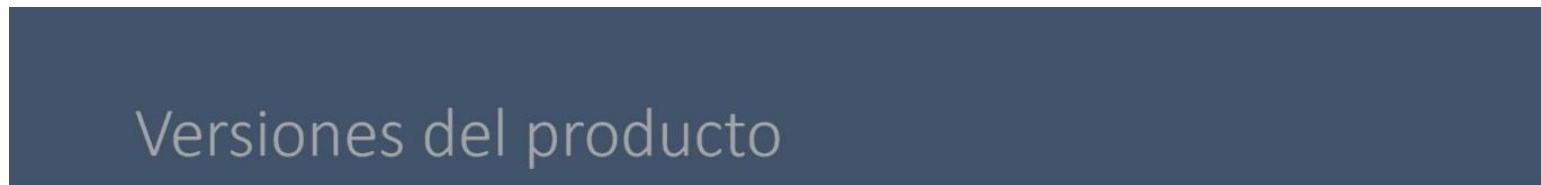
INSTALACIÓN POWER BI DESKTOP

6/5/2025

IN2025

Para obtener Power BI Desktop, se puede usar uno de los siguientes métodos:

- Instalarlo como una aplicación desde Microsoft Store.
- Descargarlo directamente como un archivo ejecutable e instalarlo en el equipo.



Puedes descargarla
desde la tienda de
aplicaciones de
Microsoft



Nota: No disponible
para Mac ni Linux

Fuente: Cursos "Formación MS Teams" MicroSoft LATAM 2023 - Curso 6 | "Toma las mejores decisiones para tu clase con PowerBI y Teams"



INICIANDO POWER BI DESKTOP

Podemos apreciar la **pantalla de inicio** con algunos atajos para **seleccionar orígenes de datos**.

The screenshot shows the Power BI Desktop application window titled "Sin título - Power BI Desktop". On the left, there's a sidebar with icons for "Inicio" (Home), "Abrir" (Open), "Cerrar sesión" (Sign Out), "Opciones y configuración" (Options and Configuration), and "Acerca de" (About). The main area has two sections: "Seleccione un origen de datos o empiece con un informe en blanco" (Select a data source or start with a blank report) and "Recomendado" (Recommended). The "Seleccione un origen de datos" section contains six buttons: "Informe en blanco" (Blank Report), "Catálogo de OneLake" (OneLake Catalog), "Libro de Excel" (Excel Workbook), "SQL Server" (SQL Server), "Empezar con datos de muestra" (Start with sample data), and "Obtener datos de otros orígenes" (Get data from other sources). The "Recomendado" section includes four cards: "Introducción" (Introduction) with a preview of data visualization components; "Has visto esto" (Items you've seen) showing a dashboard icon; "Modificado recientemente" (Recently modified) showing a file icon for "Grupo02_TPAppP3-Reentrega.pbix"; and "Has visto" (Items you've seen) showing a file icon for "Grupo01_1". At the bottom, there are filters for "Recientes" (Recent), "Compartido contigo" (Shared with me), "Nombre" (Name), "Ubicación" (Location), and "Abierto" (Open). There are also search and filter buttons.

7/5/2025

IN2025

19



INICIANDO POWER BI DESKTOP

Podemos apreciar la **pantalla de inicio** con algunos atajos para **Abrir modelos recientes o existentes**.

The screenshot shows the Power BI Desktop application window titled "Sin título - Power BI Desktop". On the left, there's a sidebar with "Inicio" and "Abrir" buttons. The main area has two sections: "Seleccione un origen de datos o empiece con un informe en blanco" and "Recomendado". Under "Recomendado", there are tabs for "Recientes" (selected) and "Compartido conmigo". A search bar and filter button are also present. Below these are tables for "Nombre", "Ubicación", and "Abierto". The "Abierto" column shows several recent reports, all of which have been open for less than a month. At the bottom of the sidebar are links for "Cerrar sesión", "Opciones y configuración", and "Acerca de".

Nombre	Ubicación	Abierto
...	...	hace 13 días
...	...	hace 17 días
...	...	hace 1 mes
...	...	hace 1 mes
...	...	hace 8 meses
...	...	hace 10 meses



INICIANDO POWER BI DESKTOP

En el lado izquierdo de Power BI Desktop, aparecen los iconos de las vistas de Power BI Desktop: **Informe, Datos, Modelo y DAX**, en orden descendente. La vista actual se indica mediante la barra amarilla situada a la izquierda y puede cambiar de vista si selecciona cualquiera de los iconos.

La vista **Informe** es la predeterminada.

The screenshot shows the Power BI Desktop interface. The ribbon at the top has 'Informe' selected. The left sidebar has four icons: 'Report' (highlighted in orange), 'Data', 'Model', and 'DAX'. The main workspace says 'Agregar datos al informe' and shows options to import from Excel, SQL Server, or clipboard. The right sidebar shows the 'Visualizaciones' and 'Filters' panes.



INTERFASE POWER BI DESKTOP

6/5/2025

IN2025

Vistas:
Reporte
Data
Modelo



Barra de herramientas

Vistas: Reporte Data Modelo

Panel Principal:

Análisis de Utilidad y Márgenes
Periodo 2017-2019

Utilidad: \$671 mill.

Margen: 19,95 %

Detalle Utilidad por País:

País	Utilidad	Margen
Estados Unidos	\$185.000.000	20,47%
China	\$183.40.559	16,73%
Alemania	\$157.41.341	20,75%
India	\$127.083.919	18,22%
Canada	\$11.144.943	20,57%
Brasil	\$17.902.043	15,66%
Nigeria	\$14.641.355	18,26%
Indonesia	\$11.207.869	10,30%
Japón	\$11.460.256	21,35%
Méjico	\$10.371.733	18,02%
Total	\$671.109.409	19,95 %

Visualizaciones:

- Utilidad y Margen por Continente
- Utilidad y Margen por Año y Continente
- Utilidad y Margen por País
- Utilidad por Trimestre y Continente

Campos:

- Table Datos
- Table Productos

Filtros:

Página(s) de Reporte:

Fuente: Guía de Power BI – DatData | PDF Power BI - Introducción, Tutorial, y Tips para aprender Power BI (En Miel) | <https://www.udemy.com/user/datdata/>



FASES DE TRABAJO EN POWER BI

6/5/2025

IN2025

GD

Get Data

Obtener Datos

Tablas de Excel, PDF, Bases de Datos, Fuentes Web, Servicios en Línea, etc.

DP

Data Preparation

Preparación de Datos

De los tablas de datos obtenidas - Limpia y organiza en Query Editor

DM

Data Modeling

Modelado de Datos

De las tablas de datos - Crea estructuras (modelos) que permitan relacionar datos.

DV

Data Visualization

Visualización de Datos

Representaciones de datos en forma de gráficos, matrices, y más visualizaciones.

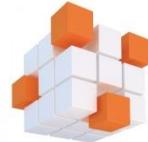
DR

Data Reporting

Reporteo de Datos

Estructura y Formato de visualizaciones y elementos que darán lugar a un Reporte.

 datdata



FASES DE TRABAJO EN POWER BI

6/5/2025 IN2025

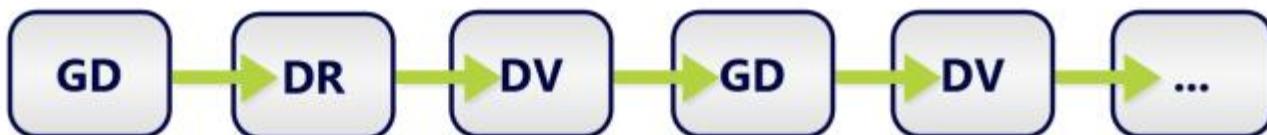
Fases de Power BI

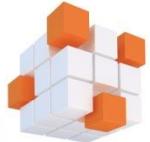
Consideraciones

- Puede haber proyectos donde una o más fases no sean necesarias. Ejemplo: Una tabla con la estructura adecuada, pueda ser que no requiera Preparación ni Modelado de Datos.



- Las fases no son secuenciales, son iterativas. En cada reporte que realicemos vamos a estar pasando de una fase a otra, sin importar su frecuencia de uso, ni orden.





INICIANDO POWER BI DESKTOP

Pasemos a la herramienta....



The screenshot displays a Power BI Desktop interface with a dashboard titled "Sales for Top 5 Categories". The dashboard features four visualizations:

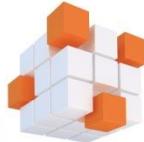
- A stacked bar chart showing Sales Amount for Germany, Canada, France, and Great Britain.
- A bar chart titled "Average Sale Amount by Class" comparing Deluxe, Economy, and Regular classes.
- A scatter plot titled "Units by Class & Brand" showing the relationship between Avg. RePurch and Avg. NSAT.
- A bubble chart showing Sales Amount tooltip, Brand & Class tooltip, and Satisfaction.

The left sidebar lists "VISUALIZATIONS" and "FIELDS". The right sidebar lists "FILTERS". The bottom navigation bar includes tabs for Overview, Germany, Canada, France, Great Britain, Sales Amount tooltip, Brand & Class tooltip, Satisfaction, and a plus sign. The status bar at the bottom indicates "PAGE 1 OF 10".

6/5/2025

IN2025

26



GD – CONECTAR Y ACTUALIZAR DATOS

GD - Conectar y Actualizar Datos

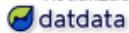
Es la función con la que empezamos cada proyecto debido a que con ella integramos los datos que serán necesarios en nuestro reporte. Esta función no solo se utiliza al inicio de cada proyecto, sino cada vez que sea requerido integrar más fuentes de datos.

Get Data = Obtener Datos; Realizar consultas (Queries); Importar Datos; Conexión a Datos; etc.

De las **fuentes de datos más comunes para Power BI** se encuentran:



Una vez obtenidos los datos (conectados) en Power BI, todas las fuentes o rutas a éstos quedarán almacenadas en el reporte, lo cual permitirá a Power BI actualizar los nuevos registros para ser visualizados en el reporte cuando sea requerido.



GD - Conectar y Actualizar Datos

Consideraciones

Fuente de datos: Cualquier ruta o destino que cuente con registros provenientes de documentos o archivos digitales del tipo Excel, PDF, Sitios Web, Bases de datos/Servidores, Servicios en línea, etc.

En ocasiones nuestras fuentes contaran con más **elementos que no son parte de la tabla de datos** que deseamos importar, o bien inconsistencias en los datos que requieren de modificación, y para **corregir** temas de este tipo será necesario de avanzar a la 2da. fase "**Preparación de datos**".

Actualizar Datos: Debido a que Power BI está conectado a fuentes de datos, al estar alimentando estas fuentes con nuevos registros, Power BI puede actualizarse y mostrar estos datos posteriormente en el reporte.

Nota: El cambiar la estructura de las fuentes de datos puede afectar en el proceso de obtención. Asimismo, cambiar el nombre de archivo o de carpeta o cualquier factor que afecte la ruta.

Tip: Organizar una estructura efectiva: Asignar nombres consistentes y que no vayan a ser cambiados con el paso del tiempo, tanto en las tablas, así como nombres de archivos o carpetas, modelo de base de datos, etc. La intención es prevenir problemas futuros por cambios de este tipo.



Fuente: Guía de Power BI – DatData | PDF Power BI - Introducción, Tutorial, y Tips para aprender Power BI (En Miel) | <https://www.udemy.com/user/datdata/>



DP – QUERY EDITOR

6/5/2025 IN2025

DP - Conociendo el Query Editor

El Query Editor o Editor de consultas, es donde se hacen los ajustes correspondientes a las tablas y sus registros para que Power BI los reconozca y posteriormente ser utilizados en el reporte. Es mayormente utilizado para corrección y limpieza de datos, no para realizar cálculos/operaciones.

Para abrir el Query Editor, click en Transform data desde Home. Un nuevo apartado aparecerá en otra ventana de Power BI:

The screenshot shows the Power BI desktop application. In the top ribbon, under the 'Home' tab, there is a 'Transform data' button highlighted with a green circle and arrow. Below the ribbon, there's a 'Tables' section in the main workspace. To the right, a large window titled 'Barra de herramientas del Query Editor' (Query Editor toolbar) is open, showing a preview of a table with data from 'Table_Acciones' and various transformation steps listed in the 'Applied Steps' pane. A watermark for 'datdata' is visible at the bottom left of the slide.

DP - Conociendo el Query Editor

Consideraciones

Tipos de datos: Importante revisar que cada campo cuente con el tipo de dato que le corresponde. Por lo general Power BI los asigna en automático, pero en ocasiones pueden no ser los correctos.

Pasos aplicados: Cada corrección o ajuste realizado queda "grabado" en un paso en el Query Editor, y cada uno de estos pasos son realizados en el mismo orden cuando actualicemos nuestro reporte. Es decir:



Nota: El Query Editor modifica la estructura a ser utilizada en Power BI, pero no afecta en lo absoluto a la estructura o registros de las fuentes de datos. Recordemos que Power BI sólo se conecta a las fuentes de datos y no provocará ningún cambio en éstas.





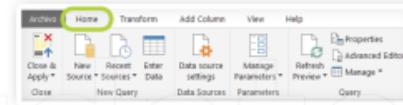
DP – QUERY EDITOR

6/5/2025 IN2025

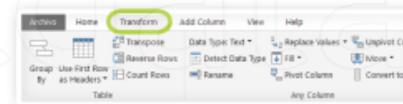
DP - Herramientas del Query Editor

En el Query Editor podemos hacer modificaciones a nuestras tablas de datos de multiples formas, y para esto haremos uso de las 3 pestañas principales de la barra de herramientas:

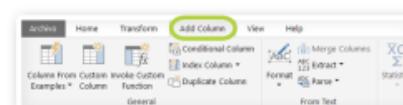
Home - Pestaña Principal donde se pueden realizar ajustes generales y transformaciones comunes (Modificar fuentes de datos, Remover columnas o filas, Combinar Tablas, etc.)



Transform - Los ajustes realizados desde esta pestaña se reflejan para las columnas seleccionadas (Reemplazar valores, Extraer Caracteres, Pivot/Unpivot, etc.)



Add Column - Se agregan nuevas columnas por lo general con referencia a los datos de otra columna (Indices, Extraer Mes/Año de columna de fecha, Columnas Condicionales, etc.)



DP - Herramientas del Query Editor

Consideraciones

Diferencia entre Transform y Add Column

- Transform: Por lo general es utilizado para modificar datos o eliminar datos innecesarios de determinada columna.

- Add column: Nos ayudará a extraer datos desde una columna, o bien para hacer referencia o ver desde otra perspectiva los datos de una columna.

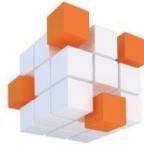
Pasos entre pasos aplicados

Podemos aplicar pasos entre los distintos pasos que ya realizamos, siempre y cuando estando consciente de que no afectarán los pasos futuros. El orden de los pasos es importante debido a que en este orden es como se volverán a aplicar cada vez que actualicemos los datos de Power BI.

Tip: Buscar eficiencia - Entre menos pasos se aplique, mejor.



Fuente: Guía de Power BI – DatData | PDF Power BI - Introducción, Tutorial, y Tips para aprender Power BI (En Miel) | <https://www.udemy.com/user/datdata/>



DM – RELACIONES DE TABLAS

DM - Relaciones de Tablas

[udemy.com/u/datdata/](https://www.udemy.com/u/datdata/)

Una de las funciones principales en el Modelado de Datos es la generación de relaciones entre tablas, o más específico entre sus campos. Existen distintas formas de relacionar, y asimismo varios propósitos para hacerlo. Vamos a ver un ejemplo común:

Tabla 1

Elemento	Cantidad
Manzana	5
Naranja	10
Zanahoria	15
Platano	5
Ciruela	10
Apio	5
Lechuga	3
Piña	1
Cebolla	3
Total	57



datdata

Tabla 2

Elemento	Categoría
Manzana	Fruta
Naranja	Fruta
Zanahoria	Verdura
Platano	Fruta
Ciruela	Fruta
Apio	Verdura
Lechuga	Verdura
Piña	Fruta
Cebolla	Verdura

Relación de Tablas

Categoría	Cantidad
Fruta	31
Verdura	26
Total	57

Ambas tablas cuentan con un campo en común 'Elemento'. Al relacionarlos podemos tomar el campo de 'Categoría' de una tabla y mostrar valores resumidos de 'Cantidad' proveniente de otra tabla.



DM – INTRODUCCIÓN A DAX

DM - Introducción a DAX

¿Qué es DAX?

DAX: Data Analysis Expressions

Podemos entender a DAX en un instante como el equivalente a realizar formulas en Excel. Los DAX son expresiones que nos permitirán hacer desde operaciones sencillas como sumas, divisiones, etc. Hasta más complejas operaciones y cálculos para múltiples propósitos.

Nota: No siempre será necesario realizar DAX, ésto debido a que Power BI en automático nos genera operaciones o también llamadas **medidas**, entre las cuales tenemos las siguientes:

- **Medidas Implícitas:** Sólo con arrastrar campo en visualización nos generá la operación (básicas)
- **Medidas Explícitas:** Nosotros las generamos utilizando DAX.

Asimismo, para **crear DAX** es importante conocer la diferencia entre hacerlo con las funciones de:

- Nueva Medida
- Nueva Columna



17

DM - Introducción a DAX

DAX - Nueva Columna vs Nueva Medida

Elemento	Ventas	Inicial	Restante
Manzana	4	10	$10 - 4 = 6$
Manzana	3	10	$10 - 3 = 7$
Manzana	1	5	$5 - 1 = 4$
Total	8	25	??

Queremos conocer el número de elementos restantes entre lo vendido y lo que había inicialmente. El Total Restante ¿Sumaría cada una de las restas, o restaría los totales?

DAX Nueva Columna			
Total	8	25	$6 + 7 + 4 = 17$

Realiza la operación en cada fila y seguido los representa en una Sumatoria (implícita).

DAX Nueva Medida			
Total	8	25	$25 - 8 = 17$

Realiza la operación después de haber realizado la Sumatoria de todas las filas.



INTRODUCCIÓN A DAX

6/5/2025

IN2025

DM - Introducción a DAX

Consideraciones

Diferencia entre Nuevas Columnas y Nuevas Medidas



Nuevas columnas: En la mayoría de los casos, el crear una nueva columna para hacer una operación o cálculo no es recomendable. Las nuevas columnas son generalmente utilizadas para agregar un dato estático o una categoría de acuerdo a ciertas condiciones que pudieran darse en determinada(s) columna(s) de la misma fila donde se está creando la nueva columna.



Nuevas medidas: Recomendables mayormente para operaciones o cálculos que puedan ser analizados desde el área de valores en las visualizaciones. Asimismo, pueden existir otros propósitos complejos en donde es mejor considerar medidas, en lugar de nueva columna.

DESCARGAR



DIFERENCIAS COLUMNA Y MEDIDA

6/5/2025

IN2025

Data Modeling - Introducción a DAX

DAX - Nueva Columna vs Nueva Medida

Elemento	Ventas	Inicial	Restante
Manzana	4	10	10 - 4 = 6
Manzana	3	10	10 - 3 = 7
Manzana	1	5	5 - 1 = 4
Total	8	25	??

Queremos conocer el número de elementos restantes entre lo vendido y lo que había inicialmente. El Total Restante ¿Sumaría cada una de las restas, o restaría los totales?

DAX Nueva Columna			
Total	8	25	$6 + 7 + 4 = 17$

Realiza la operación en cada fila y seguido los representa en una Sumatoria (implícita). 

DAX Nueva Medida			
Total	8	25	$25 - 8 = 17$

Realiza la operación después de haber realizado la Sumatoria de todas las filas. 

Data Modeling - Introducción a DAX

DAX - Nueva Columna vs Nueva Medida

Elemento	Ventas	Inicial	% Ventas
Manzana	4	10	$4/10 = 40\%$
Manzana	3	10	$3/10 = 30\%$
Manzana	1	5	$1/5 = 20\%$
Total	8	25	??

Queremos conocer el porcentaje de ventas entre lo vendido y lo que había inicialmente. ¿Cómo sería calculado el porcentaje general (Total)?

DAX Nueva Columna			
Total	8	25	$40\%+30\%+20\% = 90\%$

En este caso, primero divide cada fila y después las suma. **90% de ventas?** 

DAX Nueva Medida			
Total	8	25	$8/25 = 32\%$

En este caso, primero suma todas las filas y después divide el total. 

33



CREANDO VISUALIZACIONES

6/5/2025

IN2025

DV - Creando Visualizaciones

Visualización: Es toda representación de datos en forma de gráfico, diagrama, matriz o tabla, un número o texto, y todos esos elementos que aplicados correctamente nos ayudaran a comprender una situación, y en conjunto nos permitirán una toma de decisiones informada y eficiente.

Crear visualizaciones en Power BI: Una visualización básica consistirá de un valor, y este valor lo podemos segmentar por una categoría o categorías de distintos campos. Entre más compleja la visualización tendrá más campos y distintos tipos de valores. Por lo general las visualizaciones tienen las siguientes áreas de campos, aunque con ciertas variaciones:

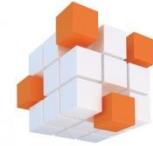
Segmentar por Campo de eje (Axis), Ubicación, Categoría, Detalles, etc.

Leyenda para diferenciar por colores las distintas categorías del campo colocado.

Campo de valores a representar con una operación/medida (Suma, Conteo, etc.)

Tooltips para mostrar más info al colocar mouse encima de segmento de dato.





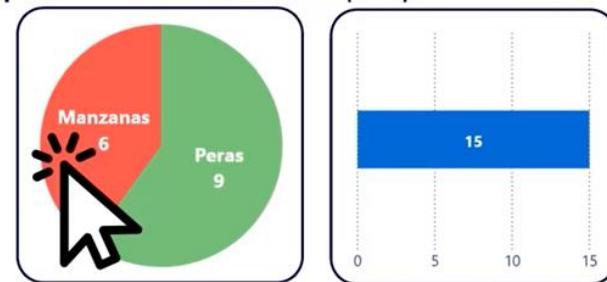
VISUALIZACIONES – INTERACCIONES Y FILTROS

6/5/2025 IN2025

Data Visualization - Interacciones y Filtros

Al seleccionar algún elemento de una visualización, existen **3 tipos de interacciones** que pueden impactar en las demás visualizaciones. Por ejemplo:

Tenemos dos visualizaciones: Una con un gráfico de pastel el cual nos muestra el total de ventas de distintos elementos.
Tenemos una segunda visualización de gráfico de barra el cual nos muestra el total de ventas de todos los elementos.



Al seleccionar 'Manzanas' se pueden seleccionar las siguientes interacciones en gráfico de barra:

Filter (Filtro)	Highlight (Resaltar)	None (Ninguna)
<p>Muestra datos únicamente para el elemento seleccionado (manzanas).</p>	<p>Muestra datos para el elemento seleccionado en proporción a todos los datos.</p>	<p>No tiene ningún efecto en la visualización.</p>



FORMATO DE VISUALIZACIONES

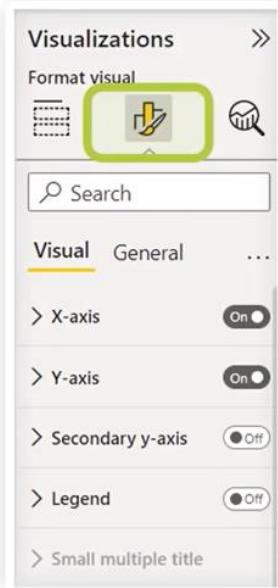
6/5/2025

IN2025

DV - Formato de Visualizaciones

Con formato de visualizaciones nos referimos a todos esos elementos que podemos modificar, remover o agregar de las visualizaciones. Ejemplo: Colores de elementos, títulos, fondos, etc.

Panel de formato



Tips de formato de visualizaciones:

Simpleza – Una visualización tiene que cumplir con el propósito de ser comprendida fácilmente. No debe demorarnos más de 5 segundos en comprender en qué consiste y su propósito.

Consistencia – Fuentes del mismo estilo y tamaño, colores/leyendas de acuerdo a categorías, contornos, margenes, etc. La intención es mantener una vista consistente para cualquier usuario.



FORMATO DE REPORTE

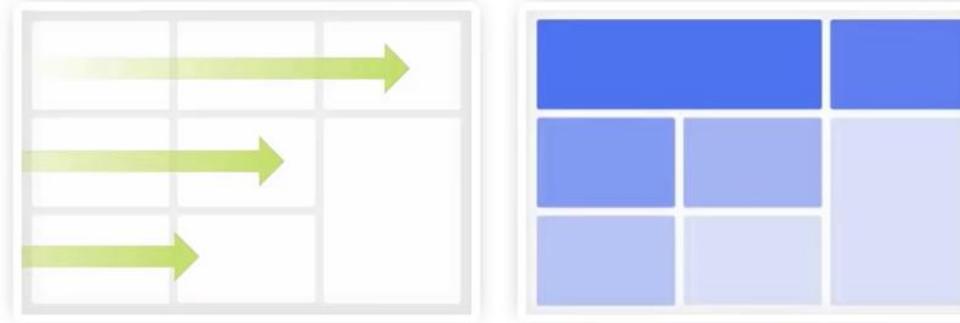
DR - Formato de Reporte

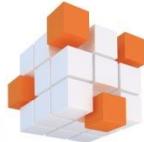
El formato de reporte esta ciertamente ligado al de las visualizaciones, pero con más enfoque en la estructura general: Acomodo de visualizaciones, botones, elementos, etc. que ayuden al usuario.

Tips de formato de reporte:

Interface de reporte – Limpia y sencilla. Colores relacionados a branding de la compañía, o temática de las categorías en cuestión. Recomendable agregar títulos o etiquetas que ayuden al reporte a ser más intuitivo para cualquier usuario nuevo. Margenes y distribución de elementos consistente.

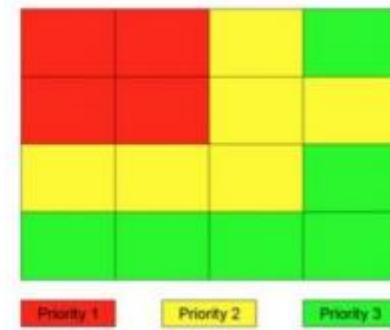
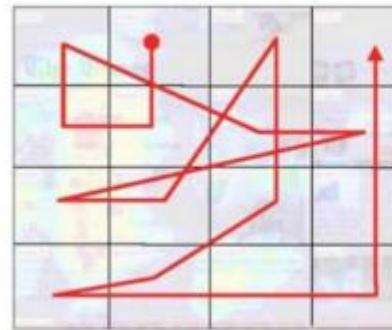
Lectura de reporte – General a específico, y de la forma en que leemos. Colocar indicadores clave desde el apartado superior izquierdo, y desde ahí partir a desglosar los detalles de los indicadores.

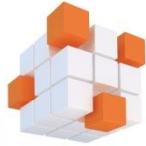




PRINCIPIO DE PRIORIDAD DE LA VISTA

- La vista humana, al encontrarse con un documento realiza instintivamente un recorrido preestablecido.
- Hay que conocer este recorrido para que nuestros gráficos y tableros respeten el orden de prioridad.





THE VISUAL REFERENCE – FOR MICROSOFT POWER BI

Categorías de Visualización

- **Comparación / Comparison** : para comparar la magnitud de las medidas.
- **Cambio a lo largo del tiempo / Change over time** : para mostrar la tendencia cambiante de las medidas
- **Parte a todo / Part-to-whole** : para identificar las partes de una medida total
- **Flujo / Flow** : para mostrar un flujo o relaciones dinámicas
- **Clasificación / Ranking** : para clasificar las medidas en un orden
- **Espacial / Spatial** : para mostrar medidas sobre mapas espaciales
- **Distribución / Distribution** : para mostrar la distribución de valores
- **Correlación / Correlation** : para mostrar correlaciones entre medidas.
- **Único / Single** : para presentar valores únicos
- **Narrativa / Narrative** : Contar una historia con datos
- **Filtro / Filter** : para controlar los filtros de informes
- **Misceláneas / Miscellaneous**

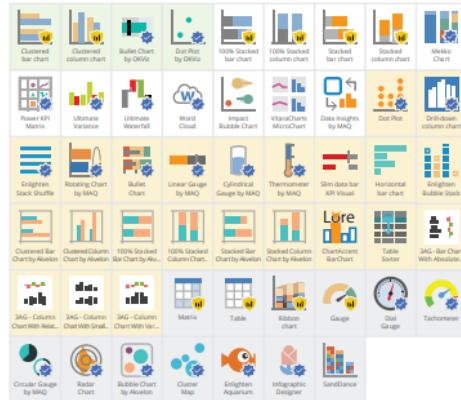


6/5/2025 IN2025

THE VISUAL REFERENCE – FOR MICROSOFT POWER BI

COMPARISON

To compare the magnitude of measures



CHANGE OVER TIME

To display the changing trend of measures



RANKING

To rank measures in an order



SPATIAL

To display measures over spatial maps



FLOW

To display a flow or dynamic relations



PART-TO-WHOLE

To identify the parts making up a measure total



DISTRIBUTION

To display the distribution of values



CORRELATION

To show correlations between measures



SINGLE

To present single values



FILTER

To control report filters



NARRATIVE

To tell a story with data



MISCELLANEOUS



Recommended

There is a better alternative

Don't use in the category



Built-in visual

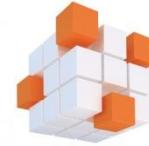


Certified visual



R required

2018 © SQLBI AND OKVIZ ARE TRADEMARK OF SQLBI CORP. / POWER BI IS A TRADEMARK OF MICROSOFT CORP.



TIPOS DE VISUALIZACIONES

6/5/2025

IN2025

HIGHLIGHT A PERSON'S COMMENT		
ALIGN WITH A GOAL OR OUTCOME		
SHOW A PROCESS		
PRESENT THEMES		
SHOW CHANGE OVER TIME		
DISPLAY PARTS OF A WHOLE		
FOR MORE SEE	STEPHANIEEVERGREEN.COM/BLOG • PRESENTING DATA EFFECTIVELY • EFFECTIVE DATA VISUALIZATION	

QUALITATIVE CHART CHOOSER 1.0

BY JENNIFER LYONS & STEPHANIE EVERGREEN

REMEMBER THIS IMPORTANT NUMBER

23%

COMPARE 2 OR MORE THINGS

COMPARE TO A TARGET

SHOW SURVEY RESPONSES

THESE ARE THE PARTS OF THIS WHOLE

VISUALIZE OPEN-ENDED COMMENTS

HEY, THINGS CHANGED OVER TIME

THIS THING CHANGES WHEN THAT THING DOES

Big Number

Icon Array

Pie/Donut

Bar/Column

Side by Side

Slopegraph

Back-to-Back

Dot Plot

Benchmark Line

Combo

Bullet Chart

Indicator Dots

Stacked Bar

Small Multiples

Diverging Bar

Bar/Column

Number & Icon

Nested

Pie/Donut

Stacked Bar

Histogram

Tree Map

Map

Quote & Pic

Word cloud

Stock photo Rep

Heat Map

Prezi

Line

Stacked Column

Deviation Bar

Slopegraph

Dot Plot

Scatterplot

Draw It

CHART CHOOSER 3.0
BY STEPHANIE EVERGREEN

WATCH FOR OVERLAPPING POINTS

45% their jobs

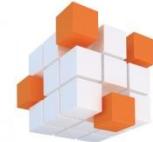
FOR BRANCHING QUESTIONS!

ALSO GOOD FOR SHOWING FLOW

FOR MORE SEE
STEPHANIEEVERGREEN.COM/TAG/STEP-BY-STEP
STEPHANIEEVERGREEN.COM/BLOG
PRESENTING DATA EFFECTIVELY

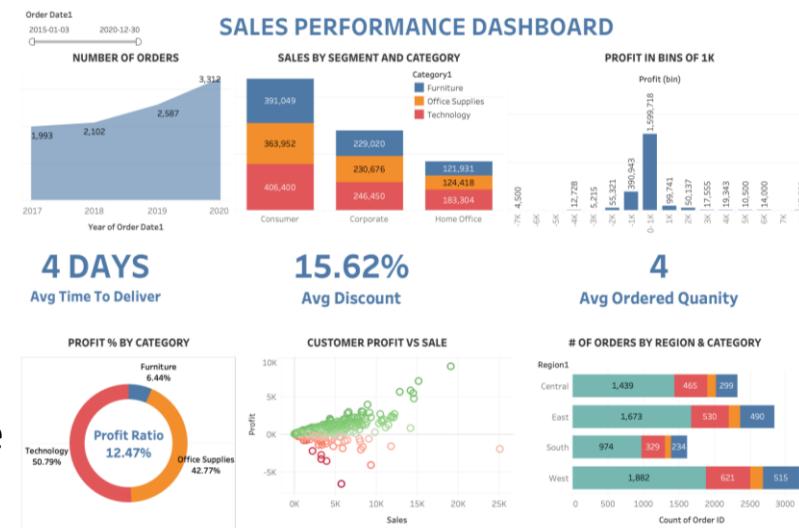
41

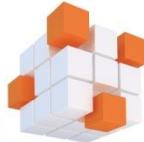
Fuente: <https://stephanieevergreen.com/>



TABLEROS DE CONTROL O DASHBOARDS

- Es la representación gráfica de los principales indicadores que intervienen en el desarrollo de los objetivos de negocio, y que está orientada a la toma de decisiones para optimizar la estrategia de la empresa.
- Los dashboards están integrados por gráficos, es por eso que es fundamental saber aplicar los gráficos adecuados a cada indicador.
- Un tablero en el ámbito de los negocios no es muy diferente al tablero de un auto. Tiene indicadores fijos, que son siempre los mismos, que nos permiten obtener información clave en un golpe de vista y poder tomar decisiones.

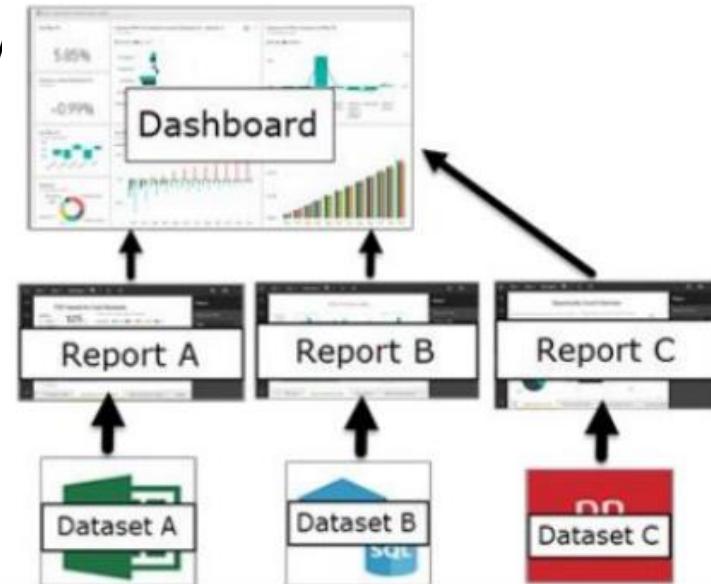


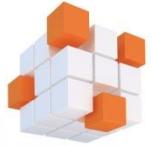


TABLEROS DE CONTROL O DASHBOARDS

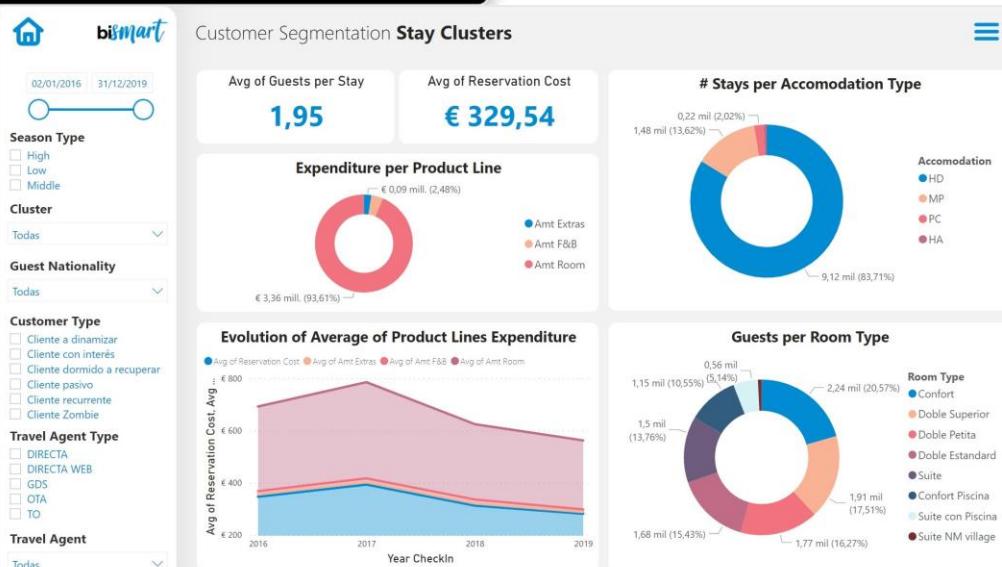
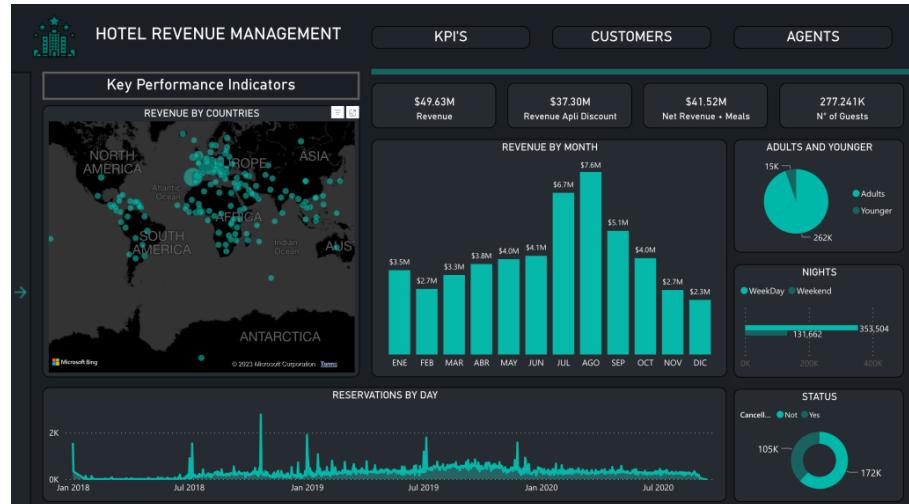
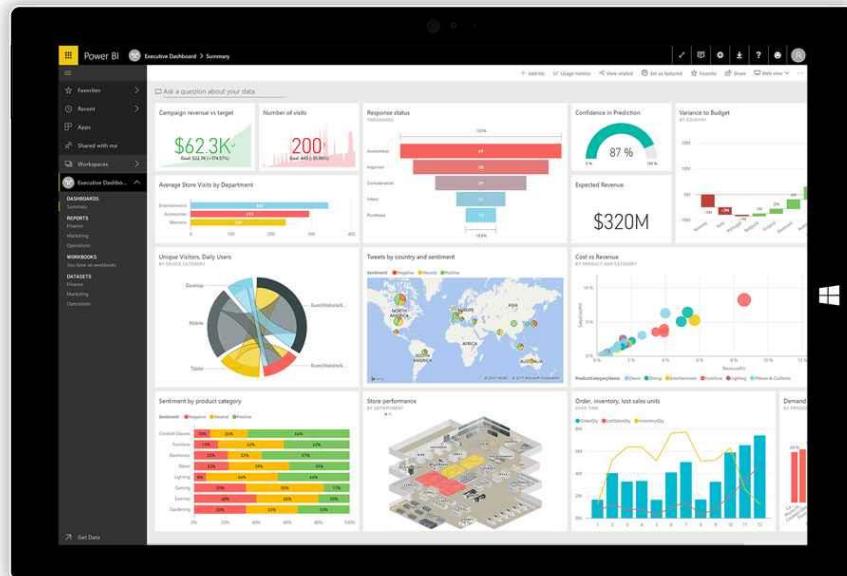
Un tablero debe:

- Estar bien organizado atendiendo la evolución del tema a monitorear.
- Ser resumido y visualmente económico, de nada sirve un tablero lleno de colores y elementos distractores.
- Resaltar excepciones.
- Estar diseñado específicamente para el usuario que lo va a utilizar.
- Mostrar conceptos concisos y claros.





TABLEROS DE CONTROL O DASHBOARDS



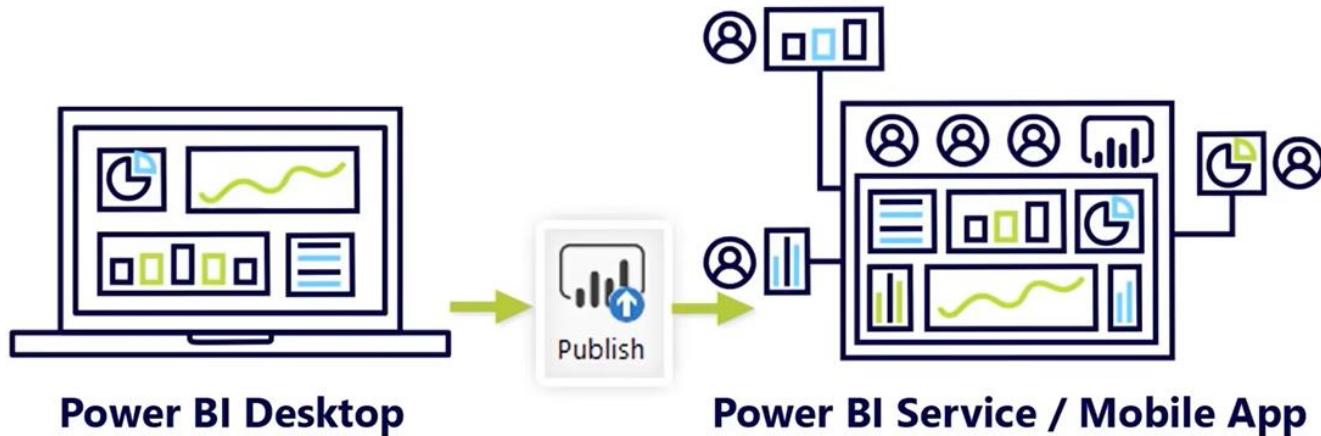
Fuente: [Microsoft Power BI - KPI Dashboard Software - Zilicus Blog | Project Management Software | Power BI KPI: Your Guide to Key Performance Indicators Visuals \(edureka.co\)](https://www.zilicus.com/power-bi-kpi-dashboards/) | <https://blog.bismart.com/10-mejores-power-bi-dashboards-2021> | <https://blog.enterpriseDNA.co/18-incredible-power-bi-dashboard-examples/>



REPORTES EN LÍNEA

DR - Viendo Reporte en Línea

Publicar en web: Esta función nos permitirá ver el reporte creado desde Power BI Desktop en línea desde cualquier dispositivo. El reporte es publicado en Power BI Service y nos genera un enlace web:



Nota: El generar un enlace web es gratis, y es público y cualquiera podría tener acceso al reporte. Power BI Pro es para aquellos que quieren mejorar su seguridad de información y compartir de forma privada sus reportes con accesos específicos a distintos usuarios.

QUESTION



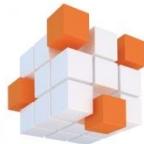
REPORTES EN LÍNEA



¿Qué es Power BI Service?

Es un software como servicio (SaaS) que nos permitirá almacenar nuestros informes, conjunto de datos, paneles, etc. Para realizar su respectiva gobernanza y posteriormente poder compartirlo



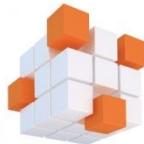


REPORTES EN LÍNEA

¿Como llevo mi informe a la nube de Power BI?

El proceso de publicación debe realizarse desde Power BI Desktop, también es posible la publicación de este directamente desde la nube de Power BI.





REPORTES EN LÍNEA

Ventajas de usar **Power BI Service**

- Los informes o paneles se pueden compartir rápidamente.
- Datos actualizados constantemente.
- Posibilidad de enviar alertas y correos electrónicos.
- Herramienta fácil de utilizar.
- Datos accesibles desde cualquier sistema o lugar.
- Datos seguros.
- Gobernanza de datos empresariales.



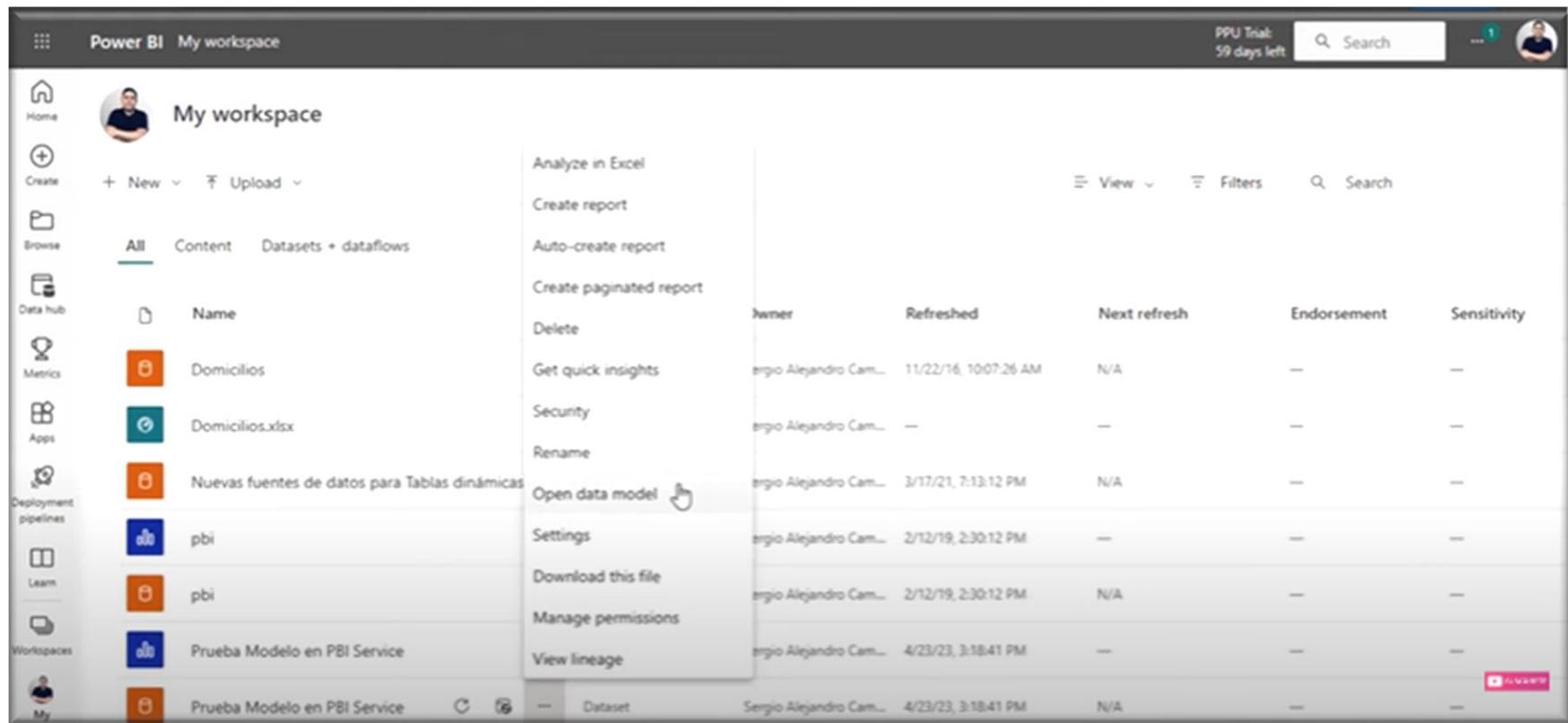
6/5/2025

IN2025



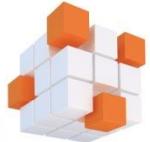
REPORTES EN LÍNEA

INDISTINTAMENTE PARA USUARIOS DE
WINDOWS – MAC - LINUX



The screenshot shows the Power BI desktop application's 'My workspace' view. On the left is a sidebar with icons for Home, Create, Browse, Data hub, Metrics, Apps, Deployment pipelines, Learn, Workspaces, and My. The main area displays a list of datasets under 'All'. A context menu is open over the dataset 'Domicilios'. The menu options are: Analyze in Excel, Create report, Auto-create report, Create paginated report, Delete, Get quick insights, Security, Rename, Open data model (with a cursor icon), Settings, Download this file, Manage permissions, and View lineage. The dataset 'Domicilios' has the following details: Owner: Sergio Alejandro Cam., Refreshed: 11/22/16, 10:07:26 AM, Next refresh: N/A, Endorsement: —, Sensitivity: —.

Name	Owner	Refreshed	Next refresh	Endorsement	Sensitivity
Domicilios	Sergio Alejandro Cam...	11/22/16, 10:07:26 AM	N/A	—	—
Domicilios.xlsx	Sergio Alejandro Cam...	—	—	—	—
Nuevas fuentes de datos para Tablas dinámicas	Sergio Alejandro Cam...	3/17/21, 7:13:12 PM	N/A	—	—
pbi	Sergio Alejandro Cam...	2/12/19, 2:30:12 PM	—	—	—
pbi	Sergio Alejandro Cam...	2/12/19, 2:30:12 PM	N/A	—	—
Prueba Modelo en PBI Service	Sergio Alejandro Cam...	4/23/23, 3:18:41 PM	—	—	—
Prueba Modelo en PBI Service	Sergio Alejandro Cam...	4/23/23, 3:18:41 PM	N/A	—	—



REPORTES EN LÍNEA

POWER BI SERVICES: EDITAR MODELO DE DATOS CREAR RELACIONES Y MEDIDAS EN DAX

6/5/2025 IN2025

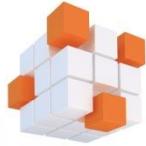
Crear relaciones entre Tablas

The screenshot shows the Power BI Services interface in the 'Data' view. It displays two tables: 'Tabla 1' and 'visitas'. 'Tabla 1' contains fields 'Alpha-2 code', 'Country', and 'Collapse ^'. 'visitas' contains fields 'FECHA', 'PAIS', and 'Σ VISTAS'. A 'Properties' panel on the right side shows settings for 'Cards' and 'Data'.

Arrastrar el campo por ej:
“país” llevamos a Tabla 1

The screenshot shows the Power BI Services interface with the 'Create Relationship' dialog open. It lists 'Tabla 1' and 'visitas' with their respective columns 'PAIS' and 'Alpha-2 code'. The 'Create Relationship' dialog includes fields for 'Cardinality' (set to 'Many to one (*:1)'), 'Cross filter direction' (set to 'Single'), and checkboxes for 'Make this relationship active' (checked) and 'Apply security filter in both directions' (unchecked). A 'Confirm' button is at the bottom right.

Confirmar



REPORTES EN LÍNEA

6/5/2025

IN2025

POWER BI SERVICES: EDITAR MODELO DE DATOS CREAR RELACIONES Y

The screenshot shows the Power BI Service interface for editing a data model. On the left, the navigation bar includes Home, Help, Create, Browse, Data hub, Metrics, Apps, Deployment pipelines, Learn, and Workspaces. The main area displays two tables: 'Tabla 1' and 'visitas'. A relationship is established between them, indicated by a line connecting the 'Country' field in 'Tabla 1' to the 'PAÍS' field in 'visitas'. The 'Properties' pane on the right shows settings for the 'Cards' section, including options to show the database in the header when applicable (set to No) and to show related fields when a card is collapsed (set to Yes). The 'Data' pane lists the tables 'Tabla 1' and 'visitas'.



REPORTES EN LÍNEA

6/5/2025 IN2025

The screenshot shows the Power BI Service interface. On the left, there's a sidebar with options like 'Inicio', '+ Crear' (which is highlighted with a red box), 'Examinar', 'Centro de datos', 'Metrics', 'Aplicaciones', 'Canalizaciones de implementación', 'Más información', 'Áreas de trabajo', and 'Mi área de trabajo'. The main area has a yellow callout box titled 'Creación del primer informe' with three steps: 'Agregar y preparar los datos', 'Crear automáticamente un informe pregrabado', and 'Personalizar para adaptarlo a sus necesidades'. Below this, there's another section titled 'Agregar datos para empezar' with two options: 'Pegue e especifique manualmente los datos.' and 'Diga un conjunto de datos publicado.'. At the bottom, it says '¿No ve el origen que busca? Descubra la aplicación de escritorio o pruebe estas opciones.'

Se pueden crear informes en Power BI Service pero con algunas restricciones por lo que al momento aún se recomienda crear el modelo desde Power BI Desktop, generar el proceso de ETL con las correspondientes transformaciones y luego publicarlo en Power BI Service.
No permite crear áreas de trabajo a las versiones Free.

Para más detalles:

<https://powerbi.microsoft.com/es-es/blog/edit-your-data-model-in-the-power-bi-service-public-preview-opt-in/>

<https://learn.microsoft.com/es-es/power-bi/transform-model/service-edit-data-models>

30 CONSEJOS Y BUENAS PRÁCTICAS PARA CREAR UN PROYECTO DE POWER BI EXITOSO



7/5/2025

IN2025

- 1) Modela correctamente tus datos antes de comenzar a desarrollar cualquier informe. Con Power BI se recomienda trabajar con modelos en estrella, reducen el tamaño de almacenamiento y optimizan los tiempos de consulta.
- 2) Valora la posibilidad de crear modelos agregados y el uso de índices para optimizar el rendimiento de las consultas. Si no usas herramientas ETL, puedes hacerlo desde Power Query o bien desde Power BI Desktop.
- 3) Evitar las relaciones uno a uno (1:1) y muchos a muchos (n:n). En este sentido, evitar también el filtro cruzado bidireccional siempre que sea posible. Si lo consigues, mejoraras el rendimiento del informe y la integridad de tus datos.
- 4) A pesar de que Power Query permite hacer operaciones de transformación del dato, se recomienda que el dato venga procesado desde el origen en la medida que sea posible. Las operaciones necesarias para realizar las operaciones de transformación en Power Query pueden repercutir en el rendimiento del informe.
- 5) Usa el método de conexión Direct Query para informes en tiempo real o con una alta volumetría de datos. Recuerda que Direct Query no permite conectar a todos los orígenes de datos permitidos en Power BI. Consulta las fuentes disponibles [en este enlace](#).
- 6) Si usas Direct Query, revisa el plegado de consultas (query folding) en Power Query.
- 7) Si tu modelo de datos cuenta con una variable tipo DATETIME, se recomienda separar las fechas de las horas en Power Query ya que mejora la capacidad de compresión.
- 8) A la hora de trabajar con Power Query, el primer paso es filtrar las filas y eliminar las columnas que no vayas a necesitar para reducir tu conjunto de datos. Luego, continúa aplicando las operaciones de transformación que necesites. Elimina lo que no vayas a utilizar.

30 CONSEJOS Y BUENAS PRÁCTICAS PARA CREAR UN PROYECTO DE POWER BI EXITOSO



6/5/2025

IN2025

- 9) Puedes combinar las ventajas del uso de Direct Query y las ventajas del uso de método Import con el uso de Modelos Diales.
- 10) Tipifica correctamente los datos. Evitar tipificar los datos de manera distinta al origen pues afecta al rendimiento.
- 11) Para crear columnas en el modelo, mejor desde Power Query (lenguaje M) que con columnas calculadas (lenguaje DAX). Lo mismo ocurre con las tablas.
- 12) Apóyate en los dataflows para definir el origen de datos en el servicio y para reutilizarlos en todos los informes de la organización. Los dataflows se configuran en Power BI Service y funcionan como una capa entre el origen de datos y el informe que permite hacer Power Query en el servicio (en cloud) de tal manera que se puede homogeneizar el origen de datos en el servicio, reutilizarlo y compartirlo entre varios informes. Se requiere cuenta PRO.
- 13) Usa variables en tus fórmulas DAX complejas. Evita crear una medida que luego únicamente vayas a utilizar en otra medida.
- 14)(14) Si deseas aplicar inteligencia de tiempo a diversas medidas, no crees una fórmula de inteligencia de tiempo para cada medida, usa grupos calculados desde Tabular Editor.
- 15) Analiza el rendimiento de una fórmula DAX con el analizador de rendimiento de Power Query o desde DAX Studio.
- 16) Usa DIVIDE para divisiones numéricas con DAX.
- 17) Para contar filas usa COUNTROWS en vez de COUNT.
- 18) Evitar convertir valores BLANK a valor. Controla los nulos y valores en blanco con lenguaje DAX.
- 19) A la hora de filtrar por una columna, es utiliza KEEPFILTERS en vez de FILTER ya que FILTER evalúa para cada elemento de una tabla mientras que KEEPFILTERS lo hace mediante una expresión booleana.
- 20) Haz un buen uso de las funciones de error ya que muchas fórmulas DAX traen excepciones ante errores.

55

30 CONSEJOS Y BUENAS PRÁCTICAS PARA CREAR UN PROYECTO DE POWER BI EXITOSO

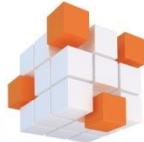


7/5/2025

IN2025

- 21) Crea una tabla de medidas en Power BI Desktop donde ir creando las diferentes medidas del modelo con DAX. Si tus medidas se importan del origen de datos, marca la opción “No resumir”, ocúltala y crear una medida DAX con esa operación de agregación
- 22) En Power BI Desktop, oculta los campos de relación de tablas (campos id) en las tablas del modelo y oculta la tabla de hechos. Facilita el trabajo a los usuarios de negocio que vayan a trabajar en un futuro con un informe.
- 23) Si creas páginas de Tooltips o de navegación en detalle (drill through), ocúltalas.
- 24) Incluye comentarios a las medidas
- 25) Crea un template corporativo e impórtalo en todos tus informes
- 26) Revisar y deshabilitar la opción de AUTO DATE/TIME para reducir el tamaño del pbix.
- 27) No sobrecargar una página de elementos visuales y no sobrecargar un informe con muchas páginas. Combina tus necesidades con el uso de aplicaciones y paneles en el servicio.
- 28) Evitar el uso de elementos visuales no certificados
- 29) El elemento visual matriz suele tener un tiempo de cómputo mayor que otros. Evitarlo si hay otra alternativa.
- 30) Si se crean tablas calendario, utilizar como rango de fechas, el valor mínimo y máximo de la tabla de hechos.

56



MEJORAS MENSUALES DE LA FUNCIONALIDAD

Mensualmente Microsoft anuncia las nuevas funcionalidades incorporadas a la solución. Las mismas se publican en el Blog de Microsoft Power BI:

Abril 2025:

[Power BI April 2025 Feature Summary](#) | [Blog de Microsoft Power BI](#) | [Microsoft Power BI](#)
[Novedades de la actualización más reciente de Power BI - Power BI](#) | [Microsoft Learn](#)

Abril 2024:

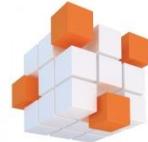
[Power BI April 2024 Feature Summary](#) | [Blog de Microsoft Power BI](#) | [Microsoft Power BI](#)

Marzo 2024:

[Power BI March 2024 Feature Summary](#) | [Blog de Microsoft Power BI](#) | [Microsoft Power BI](#)

Abril 2023:

[Edit your data model in the Power BI Service \(Preview\)](#) | [Blog de Microsoft Power BI](#) | [Microsoft Power BI](#)



LINKS DE INTERÉS

Material Oficial Microsoft Power BI

Guía de Instalación

- [Obtener Power BI Desktop - Power BI | Microsoft Docs](#)

Documentación

- [Home Power BI \(microsoft.com\) | Documentación de Power BI - Power BI | Microsoft Docs](#)
- [Documentación de introducción a Power BI - Power BI | Microsoft Docs](#)

Power BI Desktop

- [¿Qué es Power BI Desktop? - Power BI | Microsoft Docs](#)
- [Comparación de Power BI Desktop y el servicio Power BI - Power BI | Microsoft Docs](#)

Aprendizaje Guiado

- [Power BI en Microsoft Learn | Microsoft Docs](#)
- [Obtención de ejemplos para Power BI - Power BI | Microsoft Docs](#)

Blog de la Herramienta

- [Blog de Power BI: actualizaciones y novedades | Microsoft Power BI](#)

Galería de Temas, Webinars, Datos, Modelos, Scripts, etc

- [Galleries - Microsoft Power BI Community](#)

Nuevas funcionalidades Versiones:

- Abril 2025: [Power BI April 2025 Feature Summary | Blog de Microsoft Power BI | Microsoft Power BI](#)
- Abril 2024: [Power BI April 2024 Feature Summary | Blog de Microsoft Power BI | Microsoft Power BI](#)



LINKS DE INTERÉS

Canales Power BI en YouTube y + (Inglés)

Canal YouTube: Power BI

- [Microsoft Power BI – YouTube | Analyze & Visualize Data with Power BI - YouTube](https://www.youtube.com/@PowerBI/streams)

Canal YouTube: Guy in a Cube

- [https://www.youtube.com/@GuyInACube | Getting Started with Power BI - YouTube](https://www.youtube.com/@GuyInACube/streams)

Canal YouTube: SQL BI

- [SQLBI – YouTube](https://www.youtube.com/@SQLBI/streams)

Curso Introductorio Power BI – Microsoft Learn

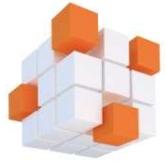
- [Introducción a la compilación con Power BI - Training | Microsoft Learn](https://learn.microsoft.com/en-us/power-bi/developer/introduction-to-power-bi)

Ejemplos PBI – Github Microsoft Power BI

- [microsoft/powerbi-desktop-samples: Power BI Desktop sample files for the monthly release. Here you can find the PBIX files used in the monthly release videos. \(github.com\)](https://github.com/microsoft/powerbi-desktop-samples)

Buscar Playlists más recientes de Microsoft Power BI en YouTube

- https://www.youtube.com/results?search_query=playlist+power+BI

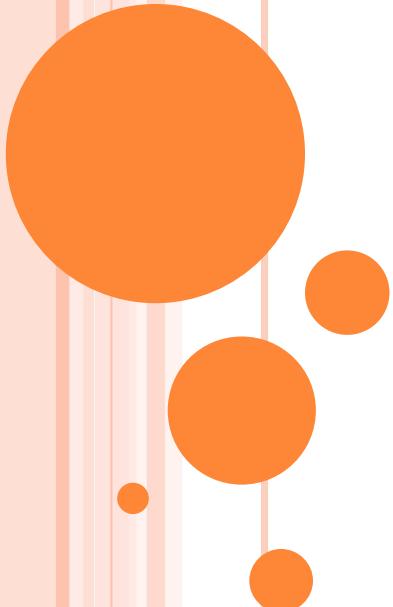


UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

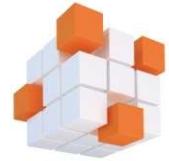
Resolución TPs

Docentes: ING. LORENA R. MATTEO



CASO DORSA

PROCESO DE DISEÑO DIMENSIONAL



Pasos

1. Elegir el **proceso de negocios** a modelar (ventas diarias, manejo de stock, etc.).
2. Elegir las **dimensiones** que van a intervenir (tener en claro los **atributos**, **jerarquías** y elementos de atributos).
3. Elegir la **granularidad** (nivel de detalle) del proceso de negocios con que se van a guardar los datos en el DW.
4. Elegir los **hechos** y **medidas** que se van a utilizar en la tabla de hechos (básicas y calculadas).

*Metodología de Ralph Kimball para
el armado incremental del DW*

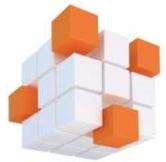
CASO DORSA

PROCESO DE DISEÑO DIMENSIONAL



Pasos

1. Elegir el **proceso** de negocios a modelar (ventas diarias, manejo de stock, etc.).
2. Elegir la **granularidad** (nivel de detalle) del proceso de negocios con que se van a guardar los datos en el DW.
3. Elegir las **dimensiones** que van a intervenir (tener en claro los **atributos** y elementos de atributos).
4. Elegir los **hechos** y **medidas** que se van a utilizar en la tabla de hechos (básicas y calculadas).



CASO DORSA

PROCESO DE DISEÑO DIMENSIONAL

Modelado Dimensional

Ejercicio 1 – Caso DORSA

DORSA (Dos Ruedas S. A.) es una empresa que se dedica a **vender** accesorios para bicicletas. Tiene dos **zonas de venta: Sur y Oeste**, y en cada una de ellas tiene un plantel de **vendedores**. Dichos vendedores tienen con DORSA **relaciones de dos tipos: Los vendedores exclusivos venden solamente productos de DORSA**, mientras que los no exclusivos dedican sólo parte de su tiempo a la venta de productos de DORSA.

DORSA vende dos **clases de productos**: los que son para la **bicicleta (como el espejo)** y los que son para el **ciclista (como el casco)**.

En la actualidad la empresa vende principalmente siete **productos**: **bocinas, espejos, cubre-manubrios y caja de cambios para la bicicleta y cascos, muñequeras y remeras para los ciclistas**.

A partir del **año 2010** (cuando empiezan a registrar información de ventas) se atienden dos **tipos de mercados**: el **mercado local y la exportación**.

El análisis de datos tiene como objetivo usar la información disponible para responder a preguntas tales como:

- ¿Cuántas **unidades** de cada producto se vendieron en el mercado local y en el de exportación?
- ¿Cómo **evolucionaron** los **precios promedio** desde 2010 a 2021?
- ¿Cómo fueron las **ventas** de cada tipo de producto en cada zona?
- ¿Vendieron más los vendedores exclusivos que los no exclusivos?
- ¿Cómo fue la evolución de las ventas en los últimos doce **meses**?
- ¿Cómo vendió cada vendedor los productos de cada **nivel de precios**?

La información de ventas se tiene acumulada por mes y comprende la **cantidad vendida y el precio por unidad**. La información para el sistema de Soporte a la Decisión proviene de la base de datos de Ventas, que consta fundamentalmente de tres grandes tablas:

- Productos
- Vendedores
- Ventas

Las tablas contienen la siguiente información:

Productos

- Código de producto
- Descripción del producto
- Código de Tipo de producto

Otra tabla permite conocer la denominación de cada tipo de producto.

Vendedores

- Identificador del vendedor
- Nombre del vendedor
- Código de zona del vendedor
- Código de **relación del vendedor con DORSA**

- Identificación del **primer día de cada mes** (Ej. 01/04/2021)
- Código de producto vendido
- Identificador del vendedor
- Código de mercado
- **Cantidad del producto vendida**
- **Precio unitario**

El código de producto y el identificador del vendedor remiten a las tablas de productos y de vendedores respectivamente, mientras que hay una tabla que permite conocer la denominación de cada mercado.

A partir del enunciado anterior, se pide:

- Liste las posibles dimensiones.
- Liste las medidas o hechos, indicando cuáles corresponden a medidas básicas y calculadas. ¿En qué proceso de negocio englobaría los hechos detectados?
- Construya el modelo dimensional conceptual indicando la granularidad del modelo.
- Identifique atributos para cada dimensión con sus jerarquías.
- Construya el modelo dimensional lógico.

Fecha Límite de Entrega: Opcional (*)¹

¹(*)

- Los TPs Opcionales sirven para afianzar conceptos necesarios para realizar los TPs de Aplicación y serán corregidos en clase y/o mediante Autoevaluación.
- Los TPs de Aplicación tienen una Fecha Límite de Entrega que deberá ser cumplida sin excepción y deben entregarse siguiendo lo indicado en el documento: "1325 Inteligencia de Negocios - Circuito Entrega TPs (Mapeo)". Serán corregidos en detalle por los docentes.
- Para las REENTREGAS: conservar el mismo documento durante las sucesivas correcciones manteniendo los comentarios efectuados por los docentes, agregando y resaltando los cambios solicitados para su posterior validación.

Ver Condiciones de Cursada en Miel – Sección: "Plazos y condiciones de Entrega Trabajos Prácticos y Casos de Estudio"

Unificar conceptos equivalentes:

- **Tipo Relación y Tipo Vendedor = Tipo Vendedor**



CASO DORSA

1. **Proceso de negocios: Ventas Mensuales**
2. **Dimensiones, atributos, jerarquías y elementos de atributos**

- **DIM VENDEDOR**

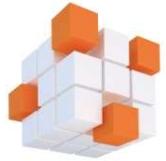
- **Zona de venta** (sur, oeste)
- **Tipo de vendedor** (exclusivo, no exclusivo)
- **Vendedor**

- **DIM PRODUCTO**

- **Clase de producto** (bicicleta, ciclista)
- **Producto** (bocinas, espejos, cubre-manubrios, caja de cambios, cascos, muñequeras, remeras)

- **DIM TIEMPO**

- **Año**
- **Mes**



CASO DORSA

2. Dimensiones, atributos, jerarquías y elementos de atributos

- DIM_MERCADO

- Mercado (local, exportación)

3. Granularidad:

- Vendedor x Producto x Mes x Mercado

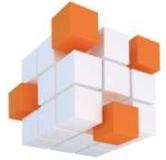
4. Hechos y Medidas (básicas y calculadas)

- Medidas Básicas:

- Cantidad vendida (unidades)
 - Precio Unitario (\$) (No fue requerido pero se usa para cálculo)

- Medidas Calculadas:

- Importe Venta (\$) = Cantidad vendida * Precio Unitario



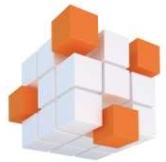
CASO DORSA

- No se tiene información sobre el **Nivel de precios** en el sistema fuente. No es posible responder esta pregunta. Si bien en un caso real es fácil de relevar y con Power BI podría definirse el atributo como un Grupo según la necesidad del negocio (por ej. Elementos: 0-99,100-999, etc. o bien Caro/Económico, etc.).
- **Precios Promedio** sería una medida calculada por la herramienta OLAP según el contexto del reporte, no es necesario agregar la medida al modelo, salvo aclaración/definición para el usuario.
- Las ventas se identifican con el primer día del mes. Podemos considerar por **Mes** y no por fecha. Aunque también se podría generar el atributo Fecha siempre y cuando sea factible completar con la info diaria.



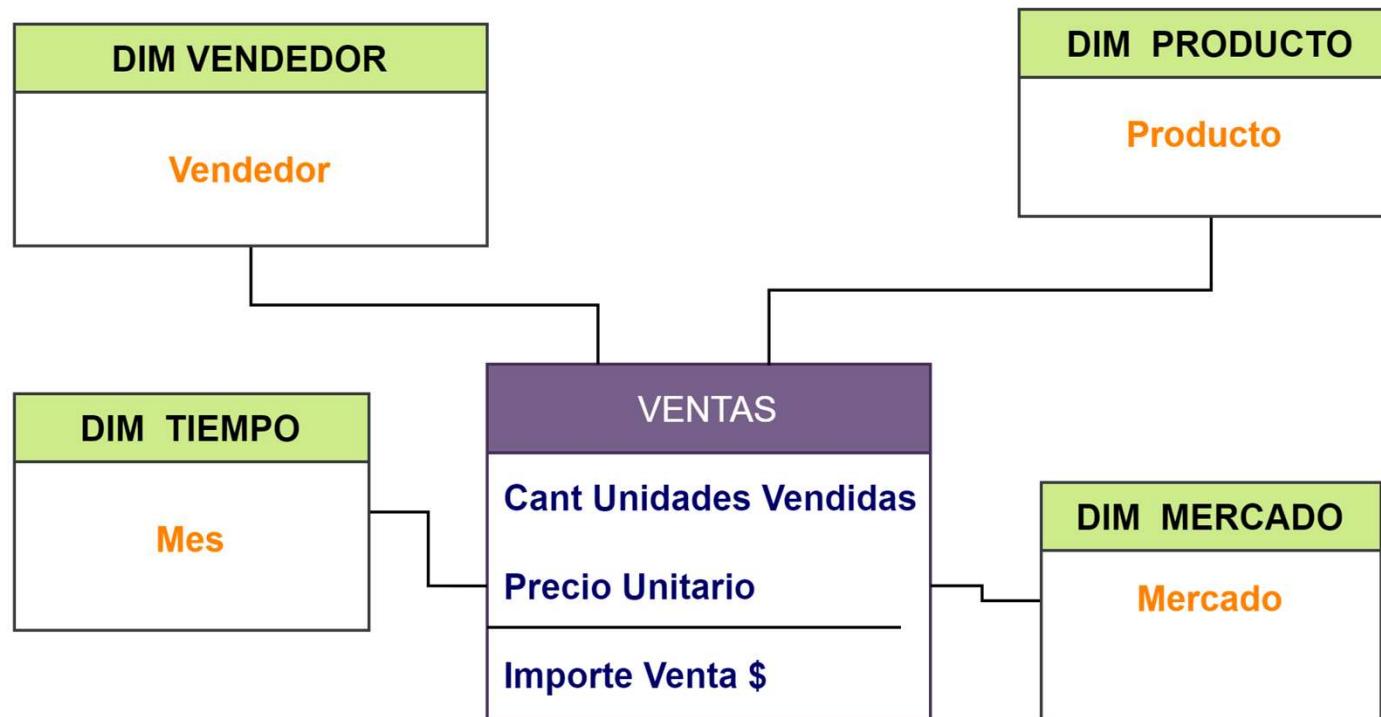
CASO DORSA

Hecho a medir: Venta de Productos				
VENTAS MENSUALES	Dimensiones			
Medidas	Tiempo	Mercado	Sucursal	Producto
Cantidad Unidades Vendidas	X	X	X	X
Precio Unitario	X	X	X	X
Importe Ventas	X	X	X	X

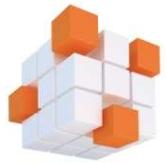


CASO DORSA

Modelo Dimensional CONCEPTUAL

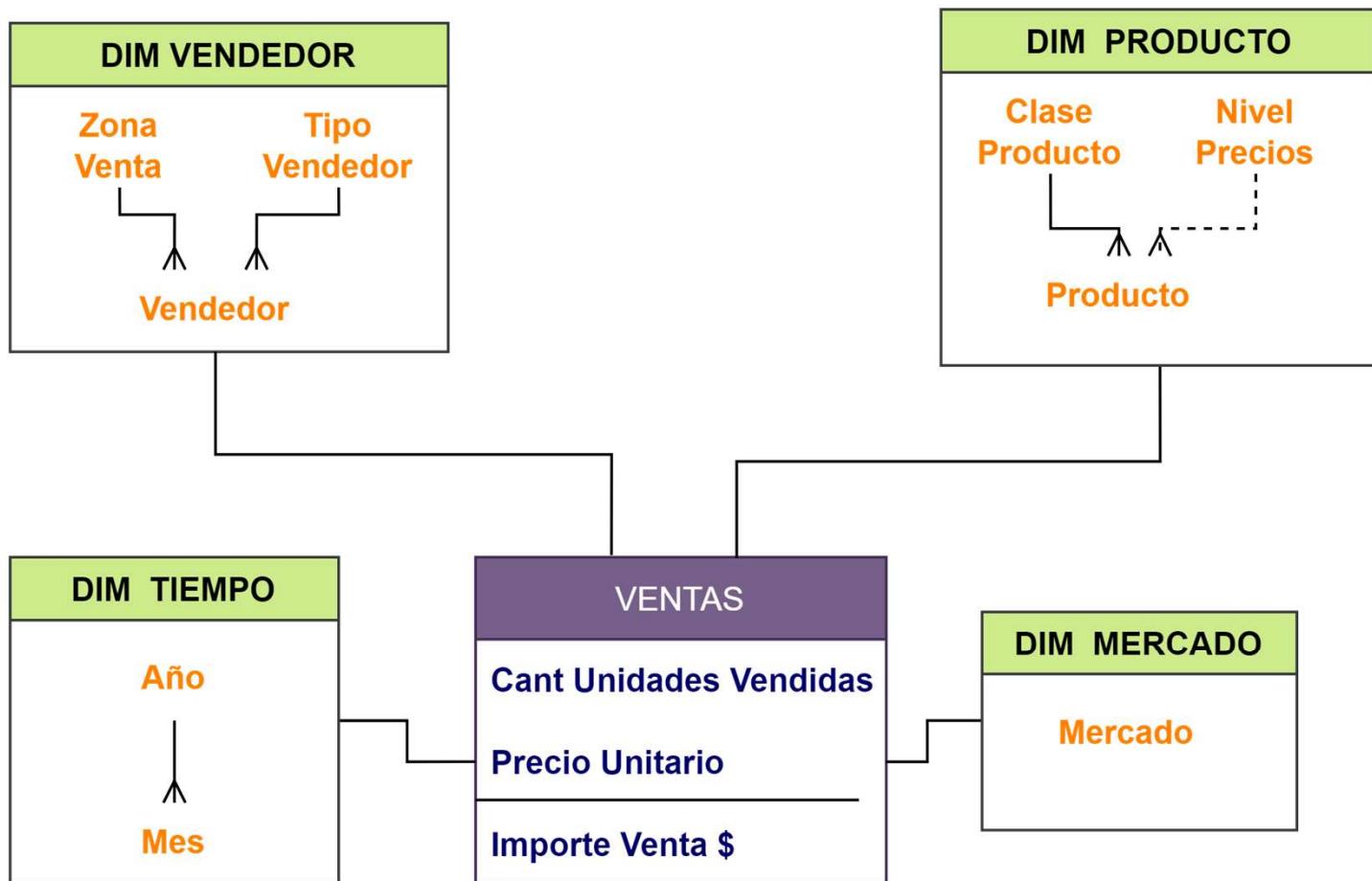


Granularidad: Vendedor x Producto x Mes x Mercado

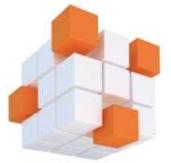


CASO DORSA

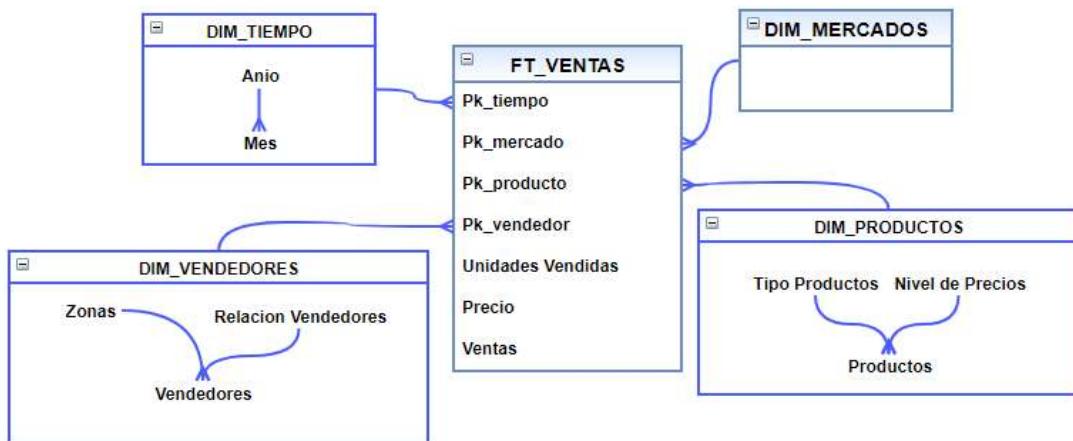
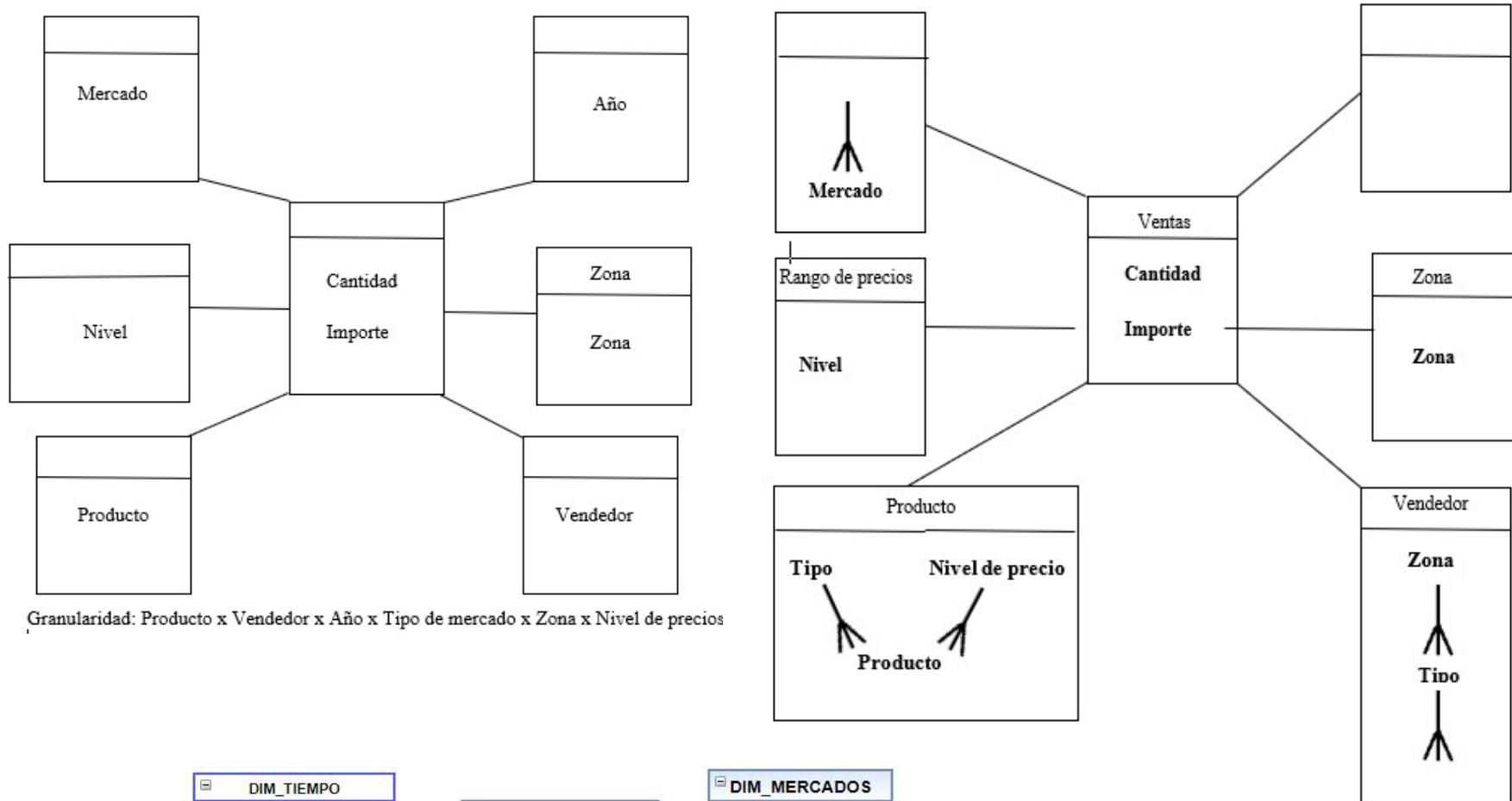
Modelo Dimensional LÓGICO

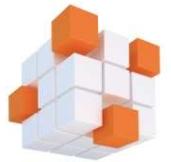


Granularidad: Vendedor x Producto x Mes x Mercado

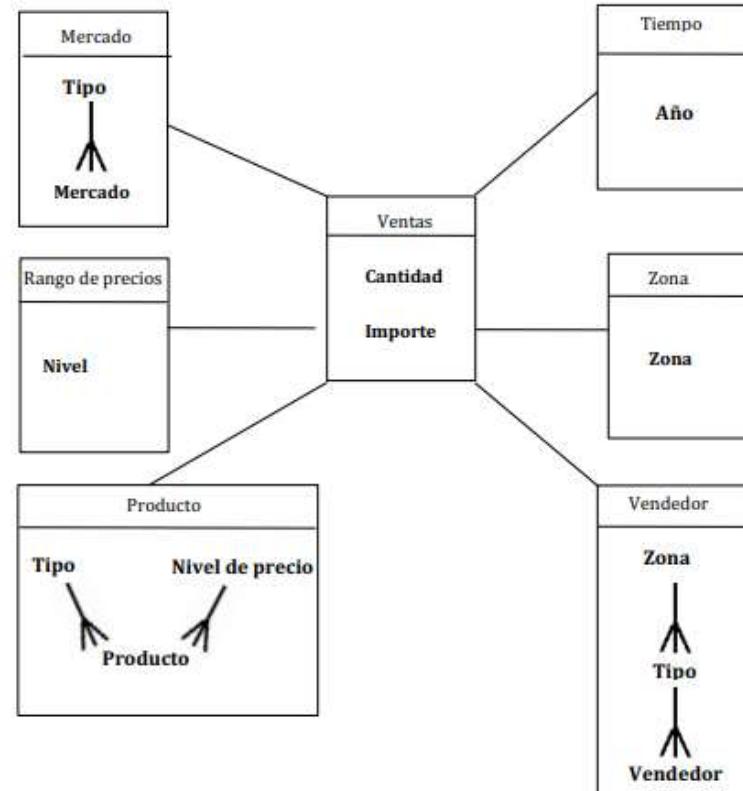
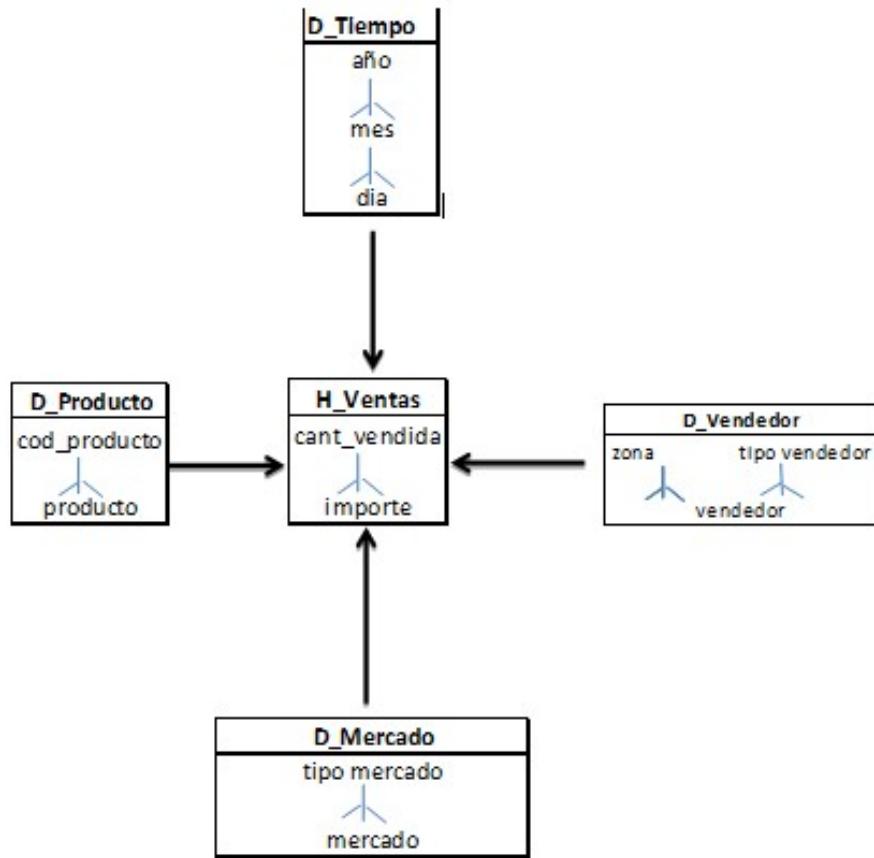


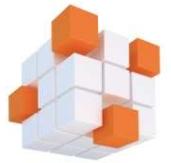
CASO DORSA – ERRORES MÁS COMUNES



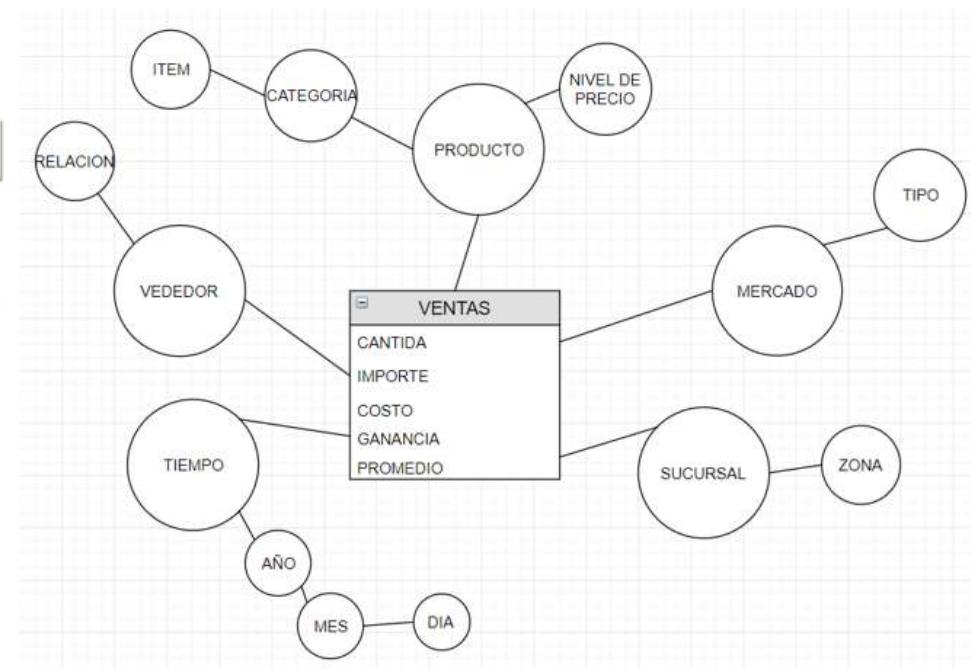
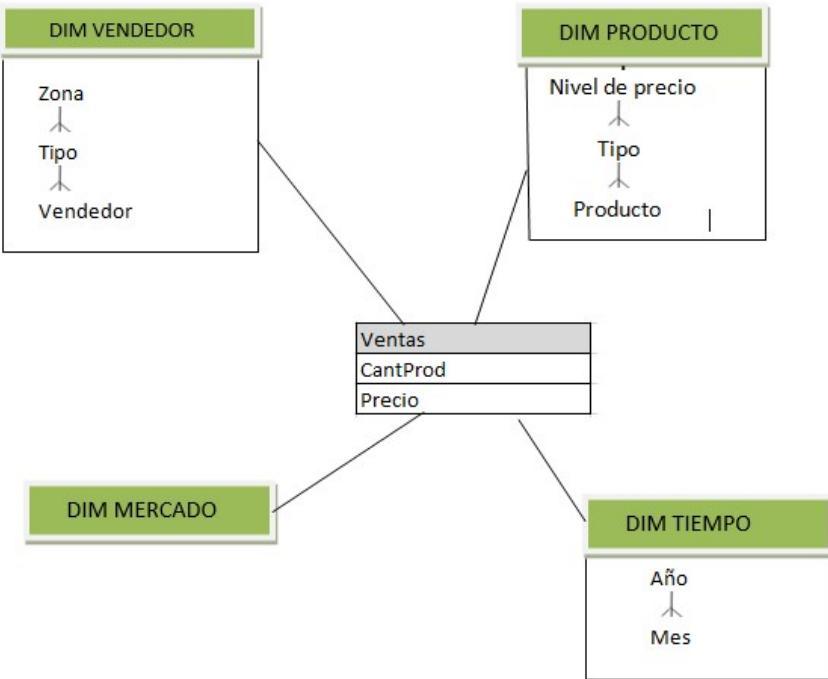


CASO DORSA – ERRORES MÁS COMUNES

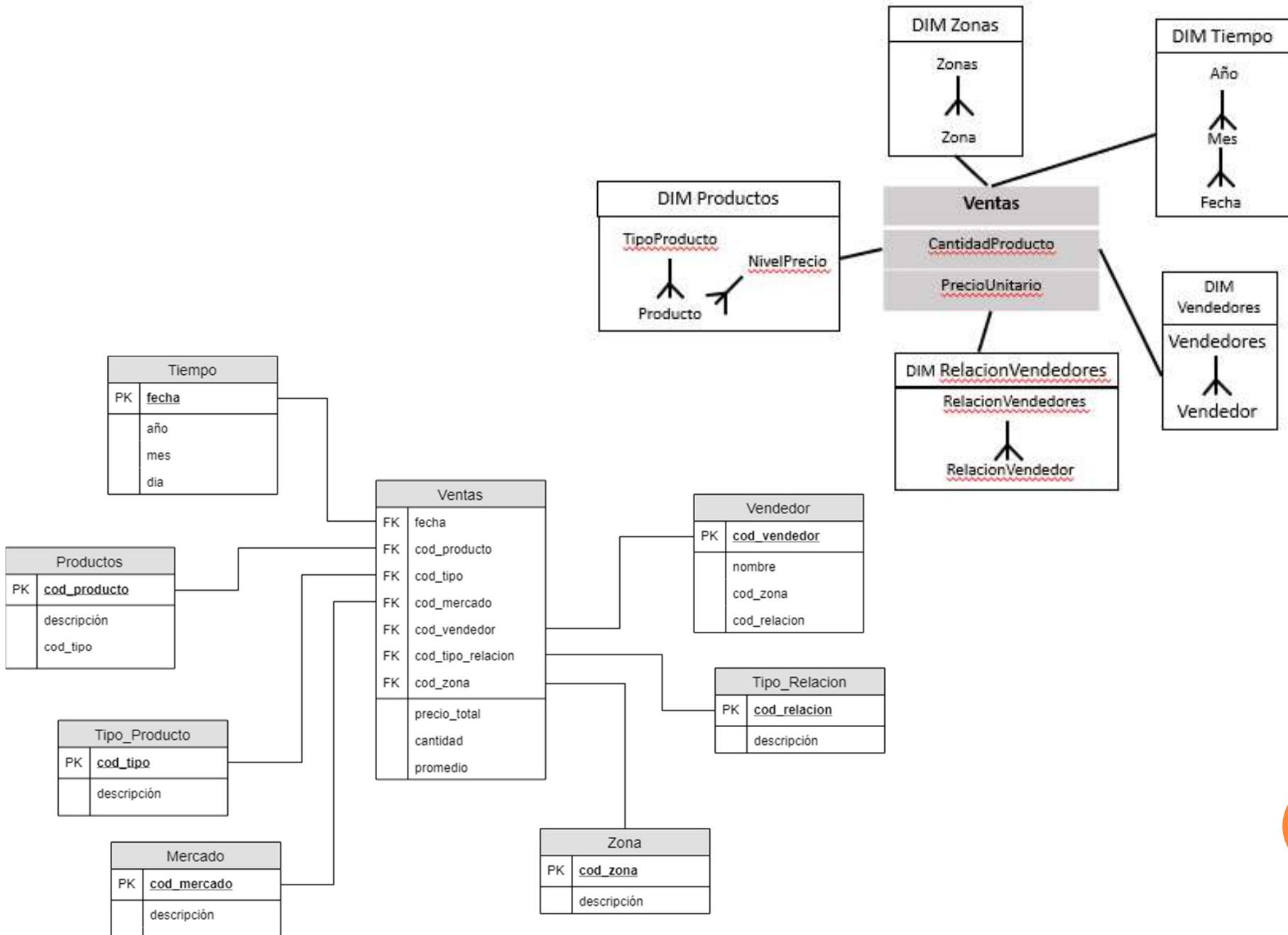


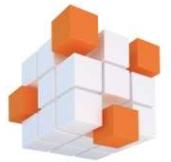


CASO DORSA – ERRORES MÁS COMUNES

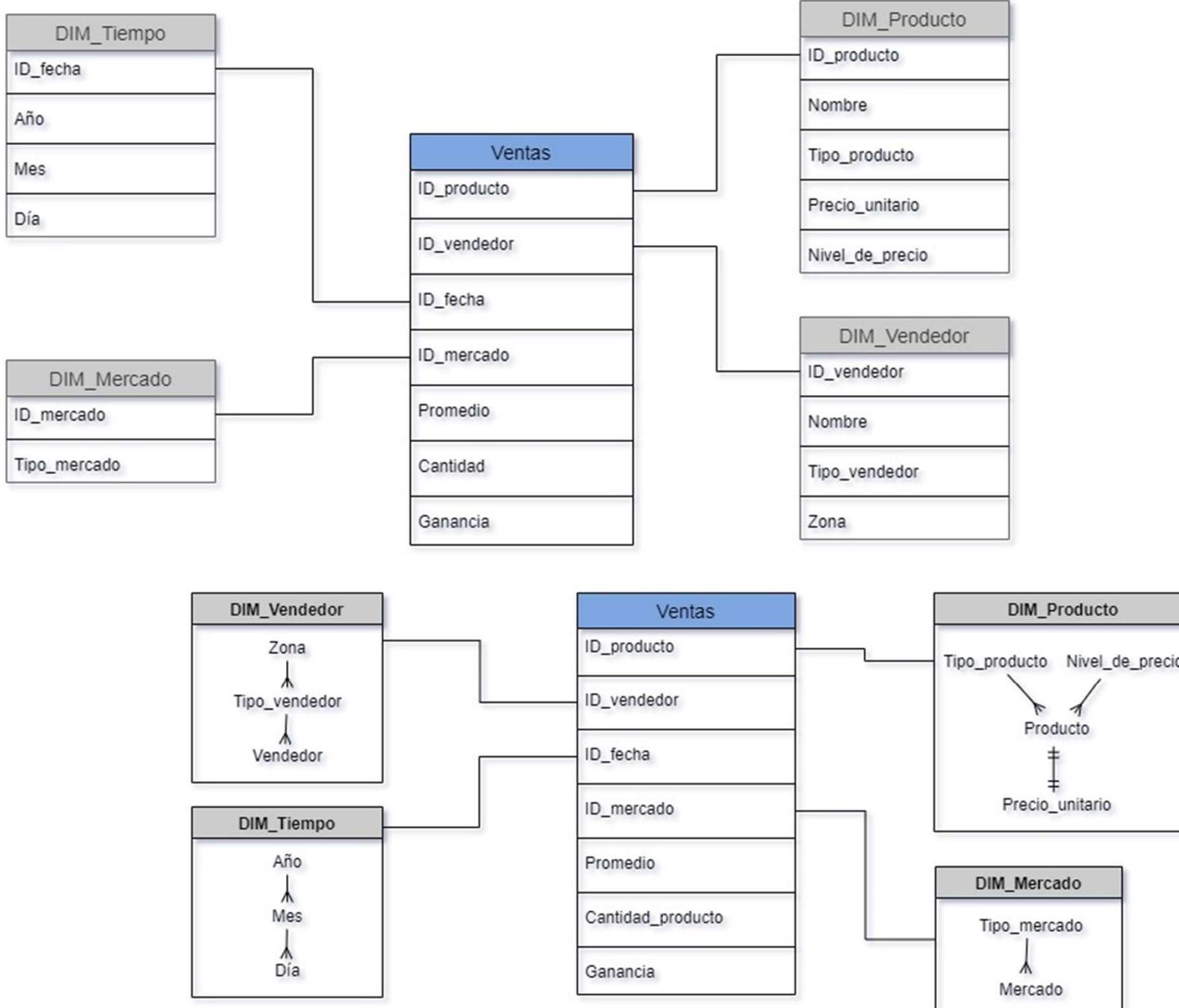


CASO DORSA – ERRORES MÁS COMUNES



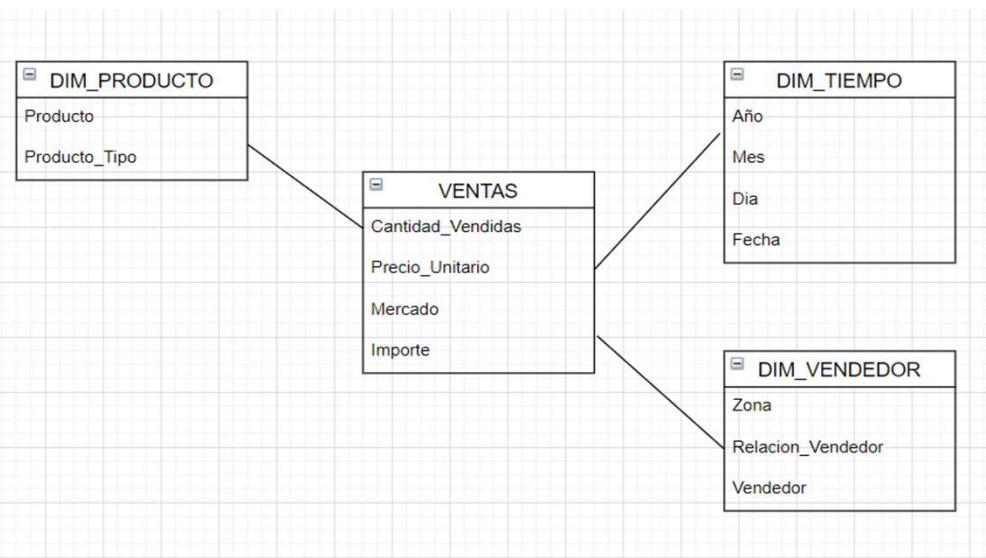
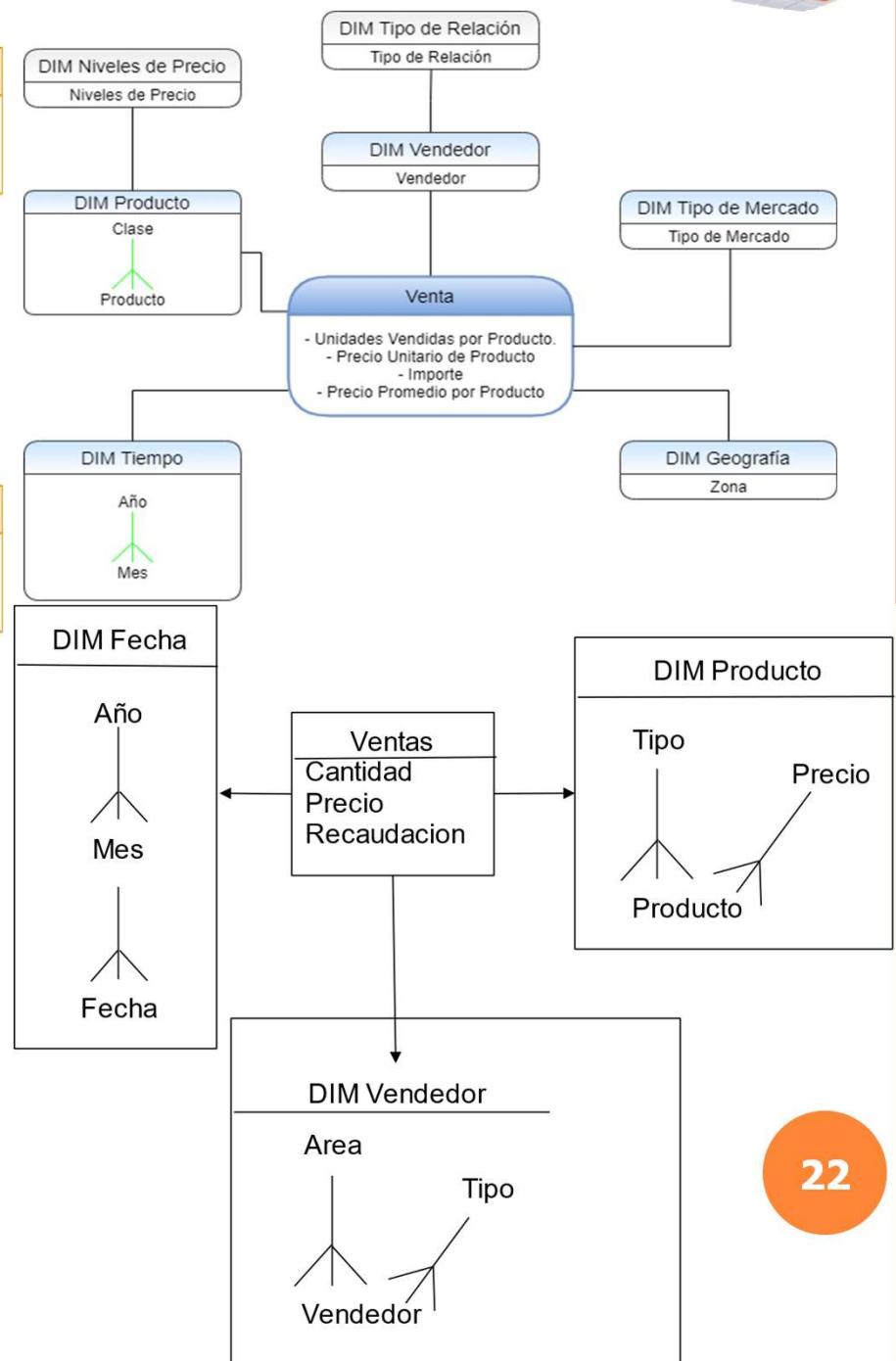
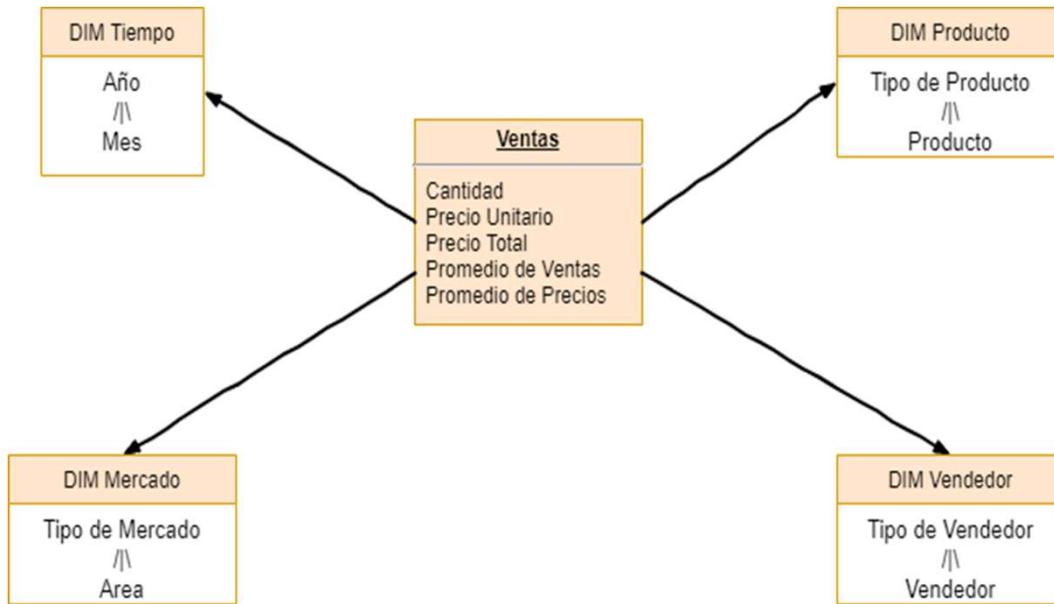


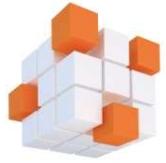
CASO DORSA – ERRORES MÁS COMUNES





CASO DORSA – ERRORES MÁS COMUNES



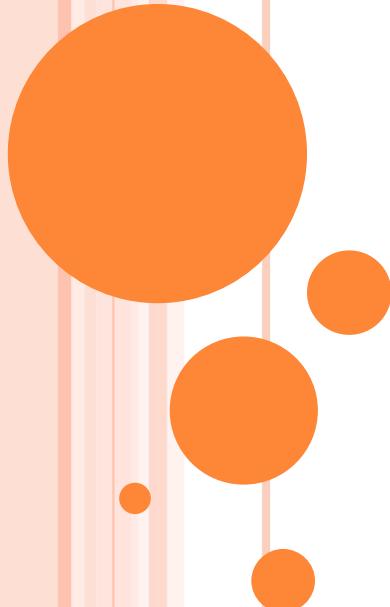


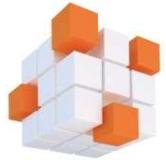
UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

Resolución TPs

Docentes: ING. LORENA R. MATTEO



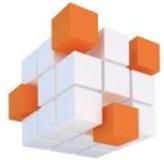


PLANILLA LITMUS

Data Warehouse Readiness Litmus Test Disposición de la empresa para implementar un datawarehouse

(The Datawarehouse Lyfecycle – Ralph Kimball)

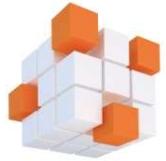
- Se debe evaluar el proyecto en una escala de 1 a 5, donde 5 representa una alta disposición, para cada uno de los ítems de los factores de éxito mencionados.
- Principales aspectos a evaluar:
 - Fuerte Apoyo de la Gerencia (Sponsor / Patrocinante)
 - Motivación Imperiosa del Negocio
 - Participación de gente del negocio y equipo TI
 - Cultura actual de análisis de información
 - Factibilidad



PLANILLA LITMUS

Data Warehouse Readiness Litmus Test Disposición de la empresa para implementar un datawarehouse (The Datawarehouse Lyfecycle – Ralph Kimball)

Puntaje	Evaluación
> 4	Alta predisposición de la empresa para implementar un DW.
3-4	Buena predisposición en general pero prestar atención a los ítems con puntaje <3
2-3	Límite. Evaluar el esfuerzo que implica mejorar las puntuaciones. La implementación del DW puede resultar muy trabajosa.
< 2	Comenzar nuevamente la planificación del proyecto. Bucar un nuevo patrocinador. Replantear la necesidad de negocio. Existe desinterés en el cambiar la cultura organizacional con la implementación del DW.

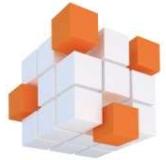


CASO PFIZER

Resultados y Conclusiones factibilidad de la empresa para implementar un DWH

(basado en la aplicación del Test de Litmus)

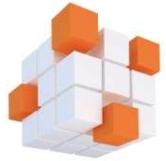
- De acuerdo a lo relevado y aplicando el Test de Litmus Pfizer presenta una alta predisposición para implementar un DWH. En general las ponderaciones de cada ítem se encuentran en promedio en 3,9
- *Puntuación 3-4 (prestar atención en los puntos < 3)*
- Factores determinantes en esa puntuación:
 - contar con un fuerte apoyo de la gerencia, fondos necesarios y recursos experimentados disponibles para el proyecto,
 - cultura organizacional receptiva a BI,
 - buena predisposición y trabajo en equipo de los distintos sectores,
 - creación de prototipos de utilidad para el negocio,
 - plan de implementación bien definido y convicción general respecto a la implementación de un DWH corporativo como solución necesaria



CASO PFIZER

Consideraciones importantes a la hora de determinar esta factibilidad

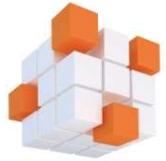
- Es de suma importancia comprender quiénes son los participantes involucrados en el proyecto para determinar el rol de cada uno dentro del mismo.
- Es clave entender cual es la necesidad del negocio y la oportunidad que se debe aprovechar con esta implementación.
- No solo se deben comprender cuales son los objetivos de negocio y del proyecto, también deben poder medirse a través de métricas concretas que demuestren que esos objetivos se están alcanzando o no.
 - Por ej. *mejorar la gestión y análisis de datos de ensayos clínicos, reducir % tiempos de desarrollo y aprobación de medicamentos, asegurar la calidad y consistencia de los datos regulatorios, agilizar la toma de decisiones estratégicas.*



CASO KODAK

Resultados y Conclusiones factibilidad de la empresa para implementar un DWH (basado en la aplicación del Test de Litmus)

- De acuerdo a lo relevado y aplicando el Test de Litmus Kodak presenta una baja predisposición para implementar un DWH. En algunos puntos deberían considerar ponderaciones que incluso no lleguen a 3.
- *Puntuación <2*
- Factores determinantes en esa puntuación:
 - bajo compromiso gerencial (la dirección no creyó en la fotografía digital ni en el valor de BI),
 - inflexibilidad en los tiempos y presupuesto,
 - cultura organizacional cerrada al cambio (dependencia excesiva del modelo de negocio tradicional),
 - falta de alineación estratégica (BI se implementó tarde y sin un propósito claro: falta de objetivos de negocio y del proyecto),
 - deficiencias en el proceso ETL (datos inconsistentes y desorganizados)



CASO KODAK

Consideraciones importantes a la hora de determinar esta factibilidad

Errores cometidos Antes y Durante el Proyecto:

- **Antes del Proyecto:**

- Kodak subestimó la importancia de la fotografía digital.
- Se basó en métricas de negocio obsoletas (ventas de películas en lugar de tendencias digitales).
- No diseñó una estrategia digital clara.

- **Durante el Proyecto:**

- Falta de visión estratégica en la implementación del DW.
- No se establecieron objetivos claros de negocio para la solución de BI.
- Mal diseño del proceso ETL.
- Falta de capacitación a los usuarios.

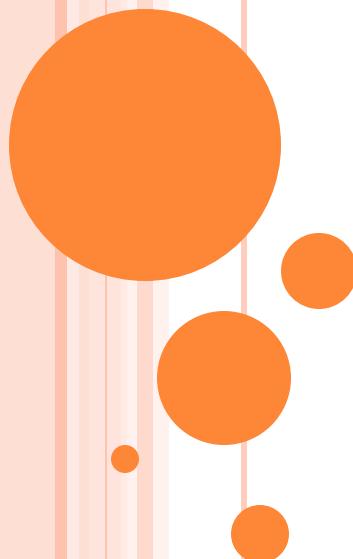


UNIVERSIDAD NACIONAL DE LA MATANZA

INTELIGENCIA DE NEGOCIOS

**Resolución TP
Caso SoundWave**

Docentes: ING. LORENA R. MATTEO/LIC. MARCELO N. SAIED



CASO PLATAFORMA STREAMING SOUNDWAVE



PROCESO DE DISEÑO DIMENSIONAL



CASO PLATAFORMA STREAMING SoundWAVE

PROCESO DE DISEÑO DIMENSIONAL



Pasos

- 1 Elegir el **proceso de negocios** a modelar (ventas diarias, manejo de stock, etc.).
- 2 Elegir las **dimensiones** que van a intervenir (tener en claro los **atributos, jerarquías y elementos de atributos**).
- 3 Elegir la **granularidad** (nivel de detalle) del proceso de negocios con que se van a guardar los datos en el DW.
- 4 Elegir los **hechos** y **medidas** que se van a utilizar en la tabla de hechos (básicas y calculadas).

*Metodología de Ralph Kimball para
el armado incremental del DW*

CASO PLATAFORMA STREAMING SOUNDWAVE

PROCESO DE DISEÑO DIMENSIONAL



Ejercicio Caso Práctico - Modelado Dimensional y Físico

Plataforma de Streaming SoundWave

La plataforma de streaming de música SoundWave ha experimentado un crecimiento exponencial desde su lanzamiento hace tres años. Con sede en Buenos Aires, **ofrece a sus usuarios acceso a más de 50 millones de canciones de diversos géneros musicales bajo diferentes modelos de suscripción:** Free (con publicidad), Basic (sin publicidad), Premium (sin publicidad y con descargas ilimitadas), y Family (hasta 6 usuarios). La plataforma está disponible en Argentina, Chile, Uruguay, Paraguay, Brasil, Colombia, México y España.

En SoundWave, los usuarios pueden crear playlists personalizadas, seguir a artistas, compartir música con amigos, y descubrir nuevas canciones a través del algoritmo de recomendación. Los artistas y sellos discográficos reciben regalías basadas en la cantidad de reproducciones que obtienen sus canciones.

El sistema de operaciones actual contiene información sobre cada reproducción de **canción**, con los siguientes datos: ID de reproducción, timestamp de inicio, duración de la escucha, ID de usuario, ID de canción, dispositivo utilizado (móvil Android, móvil iOS, web desktop, smart TV, smartwatch), si la reproducción fue completa o parcial, si se realizó online o como archivo descargado, y si formaba parte de una playlist.

También cuenta con información sobre las canciones: ID de canción, título, artista principal, artistas colaboradores (si los hay), álbum, género principal, subgénero, duración en segundos, fecha de lanzamiento, sello discográfico, y país de origen del artista principal.

Sobre los **usuarios**, el sistema almacena: ID de usuario, nombre, edad, género, país, ciudad, tipo de suscripción, fecha de registro en la plataforma, y dispositivos registrados.

Mariana Suárez, recién nombrada como Directora de Análisis de Datos, identificó que a pesar de contar con abundante información, la empresa tiene dificultades para obtener insights accionables que mejoren la experiencia de usuario y las decisiones de negocio. Actualmente, el equipo de análisis depende del departamento de TI para ejecutar consultas SQL complejas, lo que genera demoras y frustraciones.

CASO PLATAFORMA STREAMING SOUNDWAVE

PROCESO DE DISEÑO DIMENSIONAL



Tras varias reuniones donde Mariana presentó las limitaciones del enfoque actual y las posibilidades que ofrecería un sistema de BI, Santiago finalmente accedió a realizar una prueba. Mencionó varios análisis que le gustaría poder realizar fácilmente: patrones de escucha por hora del día, días de la semana y temporadas; tendencias de géneros musicales por región; comportamiento de usuarios según tipo de suscripción; artistas emergentes con crecimiento acelerado; canciones con mayor retención de escucha; impacto de las playlists en el descubrimiento de nuevos artistas; y proyección de ingresos por tipo de suscripción.

Santiago hizo énfasis en poder analizar tendencias por períodos de tiempo específicos (años, trimestres, meses, semanas, días), por ubicación geográfica a diferentes niveles (país, región/provincia, ciudad), por características del usuario (tipo de suscripción, rango etario, género), por características de la música (género, subgénero, artista, sello discográfico) y por dispositivo de acceso.

Mariana contactó a la consultora especializada en BI donde usted trabaja para llevar adelante este proyecto. Como consultor principal, usted debe diseñar e implementar la solución que transformará la manera en que SoundWave utiliza sus datos para tomar decisiones estratégicas.

Retención de escucha = Porcentaje de Reproducción

Temporadas = Agrupación de Semanas, meses , Estaciones



ANALISIS DESEABLES

Analisis deseables

- Patrones de escucha por hora del día, días de la semana y temporadas;
- Tendencias de géneros musicales por región
- Comportamiento de usuarios según tipo de suscripción;
- Artistas emergentes con crecimiento acelerado
- Canciones con mayor retención de escucha
- Impacto de las playlists en el descubrimiento de nuevos artistas
- Patrones de escucha por tipo de suscripción

- Por Periodos de tiempo específicos (años, temporadas ,trimestres, meses, semanas, días)
- Por ubicación geográfica a diferentes niveles (país, región/provincia, ciudad),
- Por características del usuario (tipo de suscripción, rango etario, género),
- Por características de la música (género, subgénero, artista, sello discográfico)
- Por dispositivo de acceso
- Por País del Artista
- Por tipo de suscripción

CASO PLATAFORMA STREAMING SOUNDWAVE

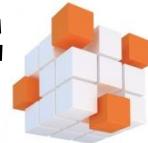
PROCESO DE DISEÑO DIMENSIONAL



A partir del enunciado anterior, se pide:

- a) ¿Cuál es el objetivo de negocios a cumplir?
- b) Liste las posibles dimensiones.
- c) Liste las medidas o hechos, indicando cuáles corresponden a medidas básicas y calculadas. ¿En qué proceso de negocio englobaría los hechos detectados?
- d) Construya el modelo dimensional conceptual indicando la granularidad del modelo.
- e) Identifique atributos para cada dimensión con sus jerarquías.
- f) Construya el modelo dimensional lógico.
- g) Arme el modelo físico del Data Warehouse en Esquema Estrella (Star)
- h) Arme el modelo físico del Data Warehouse en Esquema Copo de Nieve (Snowflake)

CASO PLATAFORMA STREAMING SOUNDWAVE



A) ¿CUÁL ES EL OBJETIVO DE NEGOCIOS A CUMPLIR?

El objetivo de negocios principal es transformar la toma de decisiones en SoundWave de un enfoque basado en reportes estáticos y dependientes de TI a un modelo ágil, autónomo y basado en datos en tiempo real, permitiendo a los diferentes departamentos obtener insights accionables para:

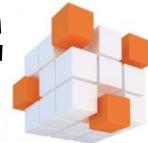
Mejorar la experiencia del usuario: Entender patrones de escucha, preferencias musicales, impacto de las recomendaciones y playlists para personalizar la oferta y aumentar la retención.

Optimizar las decisiones de negocio: Identificar tendencias de mercado, evaluar el rendimiento de artistas y sellos, optimizar estrategias de marketing y contenido, y proyectar ingresos.

Democratizar el acceso a la información: Empoderar a los diferentes departamentos para realizar sus propios análisis sin depender del equipo de TI, agilizando la identificación de oportunidades y la resolución de problemas.

En esencia, el objetivo es convertir los datos brutos de SoundWave en inteligencia de negocio para impulsar el crecimiento y la competitividad de la plataforma.

CASO PLATAFORMA STREAMING SOUNDWAVE



b) LISTE LAS POSIBLES DIMENSIONES.

- **Tiempo:** Para analizar tendencias a lo largo del tiempo.
- **Usuario:** Para segmentar y analizar el comportamiento de los usuarios.
- **Canción:** Para analizar las características de la música (genero, Tipo de reproducción, Estado) reproducida.
- **Artista:** Para analizar el rendimiento de los artistas.
- **Sello Discográfico:** Para analizar el rendimiento de los sellos.
- **Ubicación:** Para analizar el comportamiento por región geográfica.
- **Dispositivo:** Para analizar el uso de la plataforma según el dispositivo.
- **Suscripción:** Para analizar el comportamiento según el tipo de suscripción.
- **Playlist:** Para analizar si la reproducción provino de una playlist.
- **Estado de Reproducción:** Para analizar si el esatado de la reproducción es Completa / parcial
- **Tipó de Reproducción:** Para analizar si el Tipo de la reproducción es Online / Downloaded

CASO PLATAFORMA STREAMING SOUNDWAVE



c) LISTE LAS MEDIDAS O HECHOS, INDICANDO CUÁLES CORRESPONDEN A MEDIDAS BÁSICAS Y CALCULADAS. ¿EN QUÉ PROCESO DE NEGOCIO ENGLOBARÍA LOS HECHOS DETECTADOS?

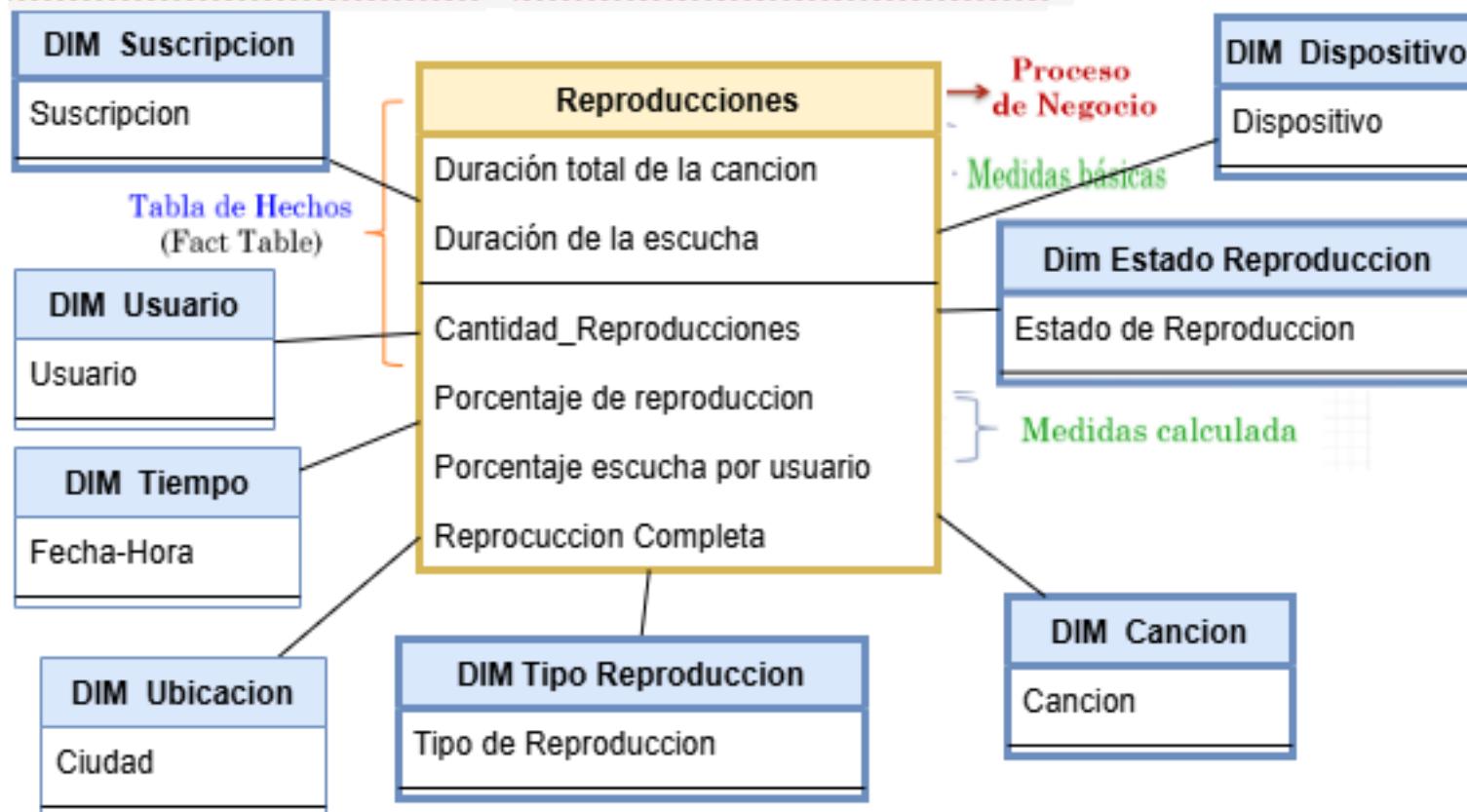
Medida	Descripción	Tipo	Fórmula	Forma de consolidación
Duracion Total de la Cancion	Tiempo Total de una canción	B		SUM
Duración total de escucha	Tiempo total de escucha en segundos	B	Suma de duraciones	SUM
Reproducciones completas	Número de reproducciones que se completaron	C	Contar reproducciones completas	SUM
Porcentaje de reproducción	Porcentaje de la canción que fue escuchada	C	(Duración de escucha / Duración total de la canción) * 100	AVG
Promedio escucha por usuario	Tiempo promedio de escucha por usuario	C	Suma de duraciones / Cantidad de usuarios	AVG
Cantidad de reproducciones	Número total de veces que se reprodujo una canción	C	Contar	SUM

CASO PLATAFORMA STREAMING SOUNDWAVE



D) CONSTRUYA EL MODELO DIMENSIONAL CONCEPTUAL INDICANDO LA GRANULARIDAD DEL MODELO.

STREAMING SOUNDWAVE (Conceptual)



Granularidad de la tabla de Hechos:

Usuario x Ciudad x Suscripción x Dispositivo x Canción x Playlist x
Estado_Reproducción x Tipo Reproducción x Fecha-hora

CASO PLATAFORMA STREAMING SOUNDWAVE



D) CONSTRUYA EL MODELO DIMENSIONAL CONCEPTUAL INDICANDO LA GRANULARIDAD DEL MODELO.

Granularidad de la tabla de Hechos:

**Usuario x Ciudad x Suscripcion x Dispositivo x Cancion x Playlist
x Estado_Reproduccion X Tipo_Reproduccion X Fecha-hora**

Granularidad del Modelo:

La granularidad de la tabla de hechos de "Reproducciones" será a nivel de cada evento individual de reproducción de una canción por un usuario en un momento específico. Esto significa que cada fila en la tabla de hechos representará una única reproducción. Esta granularidad permitirá realizar análisis detallados a nivel de reproducción y luego agregarlos a niveles superiores utilizando las dimensiones.

CASO PLATAFORMA STREAMING SoundWave



E) IDENTIFIQUE ATRIBUTOS PARA CADA DIMENSIÓN CON SUS JERARQUÍAS.

Dimensión Usuario:

Atributos:

- Rango etario
- (ej: 13-17,18-24,25-34,...)
- Género
- País
- Region/Provincia
- Ciudad
- Tipo de suscripción

Jerarquías:

- Rango etario -> Usuario
- Genero -> Usuario

Dimensión Estado

Reproducción:

Atributos:

- Estado de Reproducción (Completa/Parcial)

Jerarquías:

- Estado de Reproducción

Dimensión Tiempo:

Atributos:

- Año
- Temporada (Verano, Otoño, Invierno, etc..)
- Trimestre
- Mes
- Semana
- Día semana
- Fecha – Hora

Jerarquías:

- Año-> Trimestre -> Mes -> Fecha-Hora
- Año-> Temporada ->Semana-> Día samana-> Fecha-Hora

Dimensión Dispositivo:

Atributos:

- Tipo de dispositivo

Jerarquías:

- Tipo de dispositivo

Dimensión Tipo Reproducción:

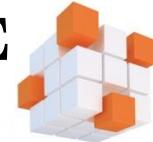
Atributos:

- Tipo de Reproducción (Online/Descarga)

Jerarquías:

- Tipo de Reproducción

CASO PLATAFORMA STREAMING SOUNDWAVE



E) IDENTIFIQUE ATRIBUTOS PARA CADA DIMENSIÓN CON SUS JERARQUÍAS.

Dimensión Canción:

Atributos:

- Título de la canción
- Nombre del artista
- País Artista
- Género Principal
- Subgénero
- Sello Discográfico

Jerarquías:

- Género Principal -> Subgénero -> Canción
- Sello Discográfico -> Artista -> Canción

Dimensión Ubicación:

Atributos:

- País
- Región/Provincia
- Ciudad

Jerarquías:

- País -> Región/Provincia -> Ciudad

Dimensión Suscripción:

Atributos:

- Tipo de suscripción
(Free/Premium/Family)

Jerarquías:

- Tipo de suscripción

Dimensión Playlist:

Atributos:

- Playlist

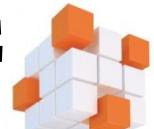
Jerarquías:

- Playlist

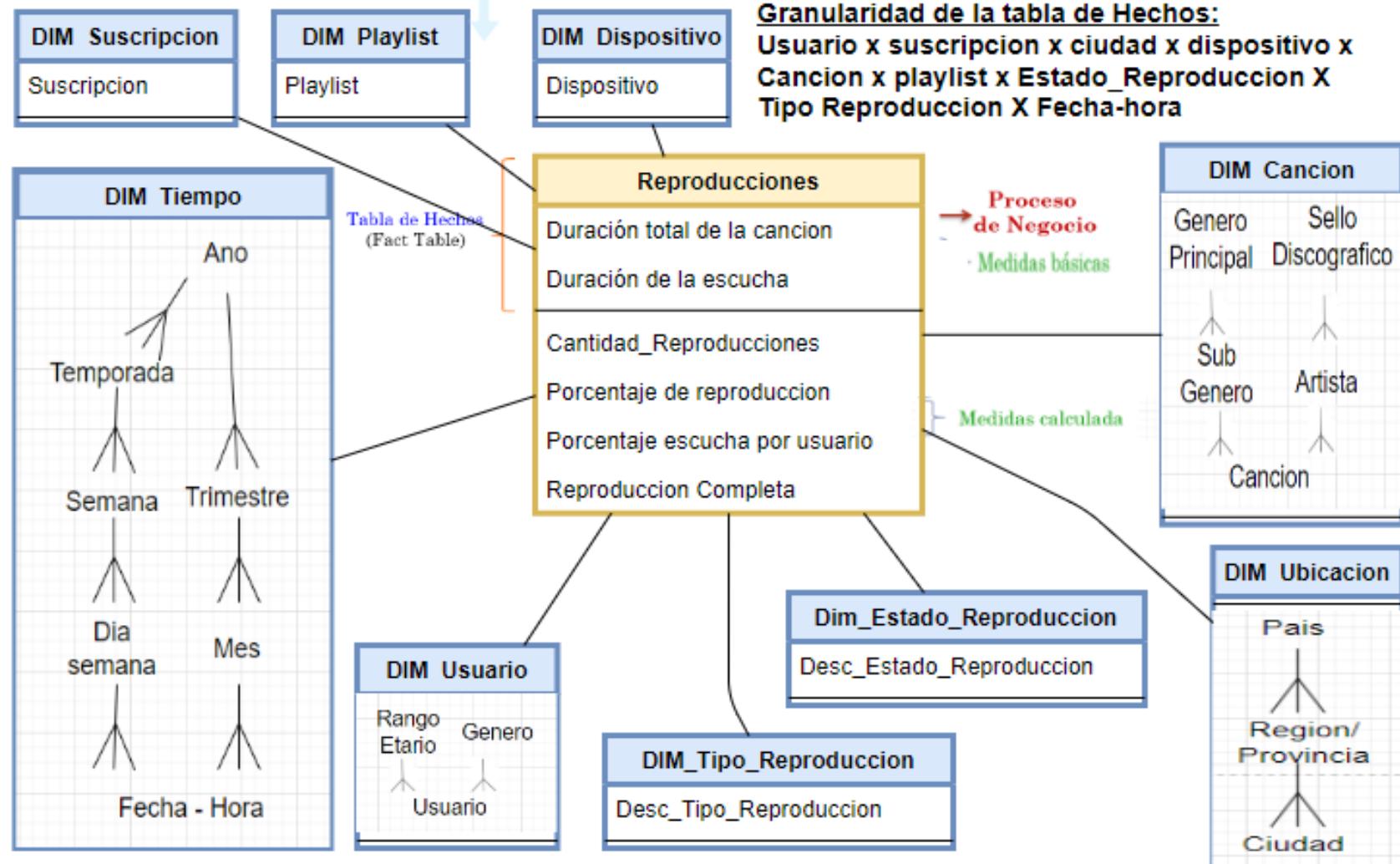
CASO PLATAFORMA STREAMING SOUNDWAVE

F)

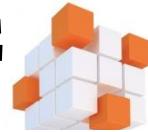
CONSTRUYA EL MODELO DIMENSIONAL LÓGICO.



STREAMING SOUNDWAVE (Logico) →



CASO PLATAFORMA STREAMING SOUNDWAVE



G) ARME EL MODELO FÍSICO DEL DATA WAREHOUSE EN ESQUEMA ESTRELLA (STAR)

STREAMING SOUNDWAVE (Fisico - Estrella)

LK_Cancion	
PK/SK	ID_Cancion
	Titulo_cancion
	ID_Artista
	Nombre_Artista
FK	ID_Pais_Artista
FK	Pais_Artista
	ID_Subgenero
	Subgenero
	ID_Genero_principal
	Genero_principal
	ID_Seloo_Discografico
	Sello_discografico
	Duracion

Granularidad de la tabla de Hechos:
Usuario x suscripcion x ciudad x dispositivo x Cancion x playlist x Estado_Reproduccion X Tipo Reproduccion X Artista x Fecha-hora
Alias:
Ciudad Artista = Ciudad Usuario

BT_Reproducciones	
PK/SK	ID_Reproduccion
FK	ID_Fecha
FK	ID_Usuario
FK	ID_Cancion
FK	ID_Dispositivo
FK	ID_Playlist
FK	ID_Suscripcion
FK	ID_Estado_Reproduccion
FK	ID_Tipo_Reproduccion
	Duracion_Escucha
	Diracion_Cancion

LK_Estado_Reproduccion	
PK/SK	ID_Estado_Reproduccion
	Desc_Estado_Reproduccion

LK_Tipo_Reproduccion	
PK/SK	ID_Tipo_Reproduccion
	Desc_Tipo_Reproduccion

LK_Usuario	
PK/SK	ID_Usuario
	Usuario_Nombre
	Edad
	Rango_Etario
FK	ID_Ciudad
	ID_Genero
	Genero

LK_Tiempo	
PK/SK	Fecha_ID
	Fecha-Hora
	ID_Mes
	Mes
	ID_Dia_Semana
	Dia_Semana
	ID_Semana
	Semana
	ID_Tempodada
	Tempodada
	ID_Trimestre
	Trimestre
	ID_Ano
	Año

LK_Ubicacion	
PK	ID_Ciudad
	Ciudad
	ID_Region_provincia
	Region_Provincia
PK	ID_Pais
	Pais

LK_Playlist	
PK/SK	ID_DPlaylist
	Playlist_Nombre

LK_Dispositivo	
PK/SK	ID_Dispositivo
	Dispositivo_Nombre

LK_Suscripcion	
PK/SK	ID_Suscripcion
	tipo_suscripcion

CASO PLATAFORMA STREAMING SoundWave



G) ARME EL MODELO FÍSICO DEL DATA WAREHOUSE EN ESQUEMA ESTRELLA (STAR)

Descripción del Modelo Físico en Esquema Copo de Nieve:

- **Tabla de Hechos (fact_reproducciones):** Permanece como la tabla central, conteniendo las medidas y las claves foráneas a las dimensiones de nivel superior.
- **Dimensiones Normalizadas:** Las dimensiones que en el esquema estrella eran tablas únicas y potencialmente grandes (como dim_usuario, dim_cancion, dim_sello_discografico, dim_ubicacion) ahora se han normalizado en múltiples tablas relacionadas.
- **dim_usuario y dim_ubicacion:** La información de ubicación del usuario (país, región, ciudad) se ha movido a una sub-dimensión dim_ubicacion, que a su vez se relaciona con tablas dim_pais, dim_region y dim_ciudad. Esto reduce la redundancia si muchos usuarios comparten la misma ubicación.
- **dim_artista y dim_ubicacion:** Similar a los usuarios, la información de ubicación del artista se normaliza a través de dim_ubicacion, dim_pais, dim_region y dim_ciudad.

CASO PLATAFORMA STREAMING SOUNDWAVE



G) ARME EL MODELO FÍSICO DEL DATA WAREHOUSE EN ESQUEMA ESTRELLA (STAR)

Descripción del Modelo Físico en Esquema Estrella:

- **Tabla de Hechos (fact_reproducciones):** Esta es la tabla central del esquema estrella. Contiene las medidas o hechos (como duracion_escucha) y las claves foráneas que referencian las claves primarias de cada tabla de dimensión. La clave primaria de esta tabla suele ser una clave artificial o una combinación de claves foráneas (aunque aquí usamos un id_reproduccion explícito).

Puntos Clave del Diseño Físico:

- **Claves Primarias (PK):** Cada tabla de dimensión tiene una clave primaria única para identificar cada registro. La tabla de hechos también tiene una clave primaria.
- **Claves Foráneas (FK):** La tabla de hechos contiene claves foráneas que enlazan cada registro de hecho con los registros correspondientes en las tablas de dimensiones. Esto asegura la integridad referencial y permite realizar joins para el análisis.
- **Tipos de Datos:** Se han utilizado tipos de datos comunes en bases de datos relacionales para Data Warehousing (INT, VARCHAR, BOOLEAN, TIMESTAMP, DATE, SMALLINT, TINYINT). La elección final del tipo de dato puede depender del sistema de base de datos específico.
- **Desnormalización:** Las tablas de dimensiones están típicamente desnormalizadas para optimizar las consultas de análisis. Esto significa que pueden contener atributos que podrían estar en tablas separadas en un modelo transaccional (como el nombre del artista dentro de la tabla dim_cancion).
- **Tabla de Ubicación:** Se ha creado una tabla dim_ubicacion separada y se ha referenciado desde dim_usuario, dim_artista y dim_sello_discografico. Esto permite una gestión más centralizada de la información geográfica y facilita el análisis por ubicación.
- **Tabla de Playlist:** Se ha creado una dimensión simple para indicar si la reproducción provino de una playlist. Este modelo físico en esquema estrella simplifica la estructura del Data Warehouse y optimiza las consultas de análisis, ya que las joins se realizan principalmente entre la tabla de hechos central y las tablas de dimensiones directamente conectadas. Esto facilita la comprensión del modelo y mejora el rendimiento de las consultas de Business Intelligence.

CASO PLATAFORMA STREAMING SOUNDWAVE



G) ARME EL MODELO FÍSICO DEL DATA WAREHOUSE EN ESQUEMA COPO DE NIEVE (SNOWFLAKE)

**STREAMING
SOUNDWAVE
(SNOWFLAKE)**

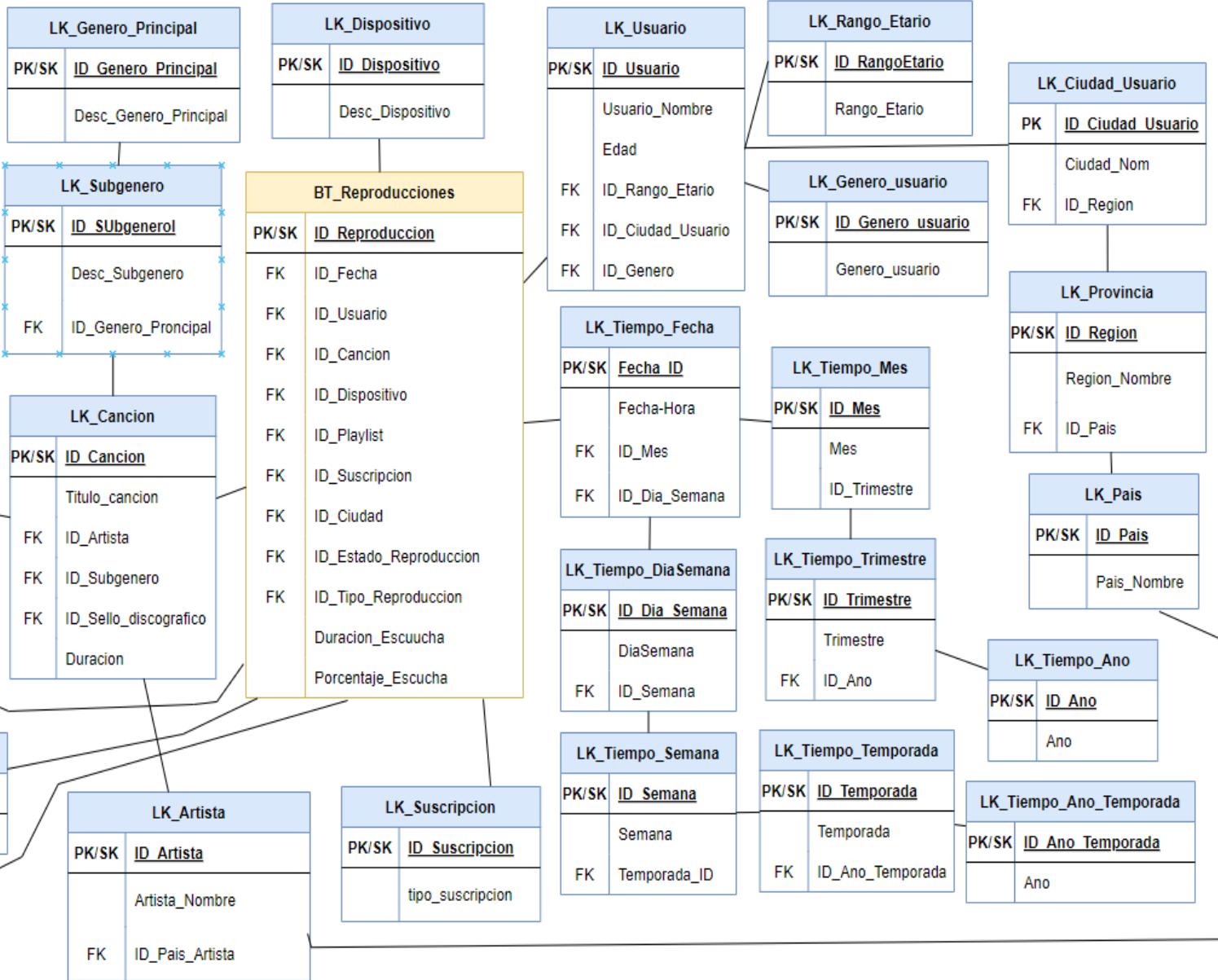
Granularidad de la tabla de Hechos:
Usuario x suscripcion x ciudad x dispositivo x
Cancion x playlist x
Estado_Reproduccion X
Tipo Reproduccion X
Fecha-hora

LK_Sello_discografico	
PK/SK	ID_Sello_discografico
	Desc_Sello_Discografico

LK_Playlist	
PK/SK	ID_DPlaylist
	Playlist_Nombre

LK_Estado_Reproduccion	
PK/SK	ID_Estado_Reproduccion
	Desc_Estado_Reproduccion

LK_Tipo_Reproduccion	
PK/SK	ID_Tipo_Reproduccion
	Desc_Tipo_Reproduccion



CASO PLATAFORMA STREAMING SoundWave



G) ARME EL MODELO FÍSICO DEL DATA WAREHOUSE EN ESQUEMA COPO DE NIEVE (SNOWFLAKE)

Ventajas del Esquema Copo de Nieve:

- **Reducción de la redundancia de datos:** Al normalizar las dimensiones, se evita la repetición de atributos, lo que puede ahorrar espacio de almacenamiento.
- **Mejora de la integridad de los datos:** Los cambios en los atributos de las dimensiones solo necesitan realizarse en una única tabla.
- **Mayor flexibilidad para cambios en las dimensiones:** Modificar la estructura de una dimensión normalizada puede ser más sencillo sin afectar directamente a la tabla de hechos.

Desventajas del Esquema Copo de Nieve:

- **Consultas más complejas:** Para obtener la información completa de una dimensión, a menudo se requieren más joins entre las tablas normalizadas, lo que puede aumentar la complejidad de las consultas y potencialmente afectar el rendimiento (aunque las bases de datos modernas suelen optimizar esto).
- **Mayor complejidad del modelo:** El esquema con más tablas puede ser más difícil de entender y mantener inicialmente.

CASO PLATAFORMA STREAMING SOUNDWAVE



G) ARME EL MODELO FÍSICO DEL DATA WAREHOUSE EN ESQUEMA COPO DE NIEVE (SNOWFLAKE)

Descripción del Modelo Físico en Esquema Copo de Nieve:

- **Tabla de Hechos (fact_reproducciones):** Permanece como la tabla central, conteniendo las medidas y las claves foráneas a las dimensiones de nivel superior.
- **Dimensiones Normalizadas:** Las dimensiones que en el esquema estrella eran tablas únicas y potencialmente grandes (como dim_usuario, dim_cancion, dim_artista, dim_sello_discografico, dim_ubicacion) ahora se han normalizado en múltiples tablas relacionadas.
- **dim_usuario y dim_ubicacion:** La información de ubicación del usuario (país, región, ciudad) se ha movido a una sub-dimensión dim_ubicacion, que a su vez se relaciona con tablas dim_pais, dim_region y dim_ciudad. Esto reduce la redundancia si muchos usuarios comparten la misma ubicación.
- **dim_artista y dim_ubicacion:** Similar a los usuarios, la información de ubicación del artista se normaliza a través de dim_ubicacion, dim_pais, dim_region y dim_ciudad.
- **dim_sello_discografico y dim_ubicacion:** La ubicación del sello discográfico también se normaliza de la misma manera.

- **Otras Dimensiones:** Dimensiones como dim_tiempo, dim_suscripcion, dim_dispositivo y dim_playlist pueden permanecer más desnormalizadas si la redundancia en sus atributos no es significativa o si la complejidad adicional de la normalización no se justifica por los beneficios.

Ventajas del Esquema Copo de Nieve:

- **Reducción de la redundancia de datos:** Al normalizar las dimensiones, se evita la repetición de atributos, lo que puede ahorrar espacio de almacenamiento.
- **Mejora de la integridad de los datos:** Los cambios en los atributos de las dimensiones solo necesitan realizarse en una única tabla.
- **Mayor flexibilidad para cambios en las dimensiones:** Modificar la estructura de una dimensión normalizada puede ser más sencillo sin afectar directamente a la tabla de hechos.

Desventajas del Esquema Copo de Nieve:

- **Consultas más complejas:** Para obtener la información completa de una dimensión, a menudo se requieren más joins entre las tablas normalizadas, lo que puede aumentar la complejidad de las consultas y potencialmente afectar el rendimiento (aunque las bases de datos modernas suelen optimizar esto).
- **Mayor complejidad del modelo:** El esquema con más tablas puede ser más difícil de entender y mantener inicialmente.

