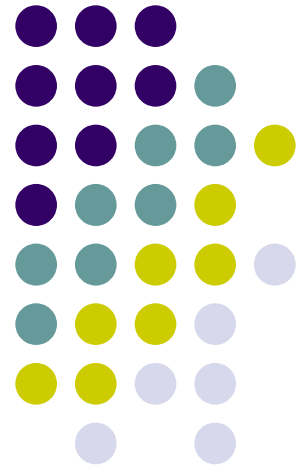


Data Warehousing Business Intelligence

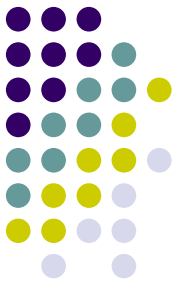
Parte 1:

Business Intelligence

Arquitectura Data Warehousing



Data Warehousing Business Intelligence



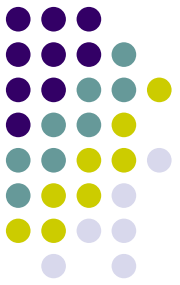
Business Intelligence

Consiste en una serie de conceptos, procesos, tecnologías y herramientas necesarias para transformar **DATO** en **INFORMACIÓN**, **INFORMACIÓN** en **CONOCIMIENTO** y **CONOCIMIENTO** en **SABIDURÍA** para generar **PLANES DE ACCIÓN** que sirvan de apoyo a la toma de decisiones del negocio.

Para ello **ALMACENA** y **PROCESA** grandes volúmenes de datos que analiza y explora con herramientas especializadas.

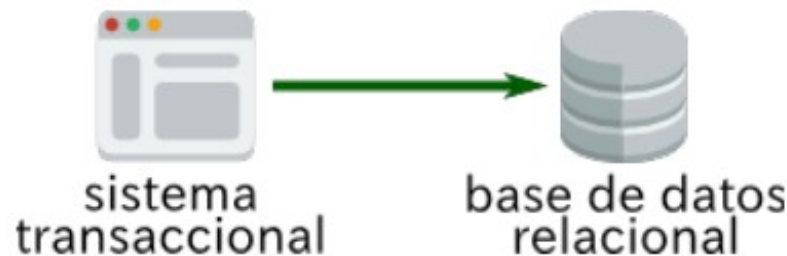


Data Warehousing Business Intelligence



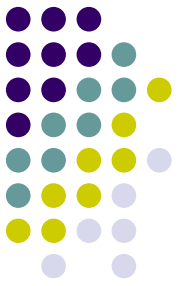
Business Intelligence

- Se generan y acumulan datos constantemente para operaciones y transacciones.
- Se administran con sistemas transaccionales y se almacenan en bases de datos relacionales (no excluyente).

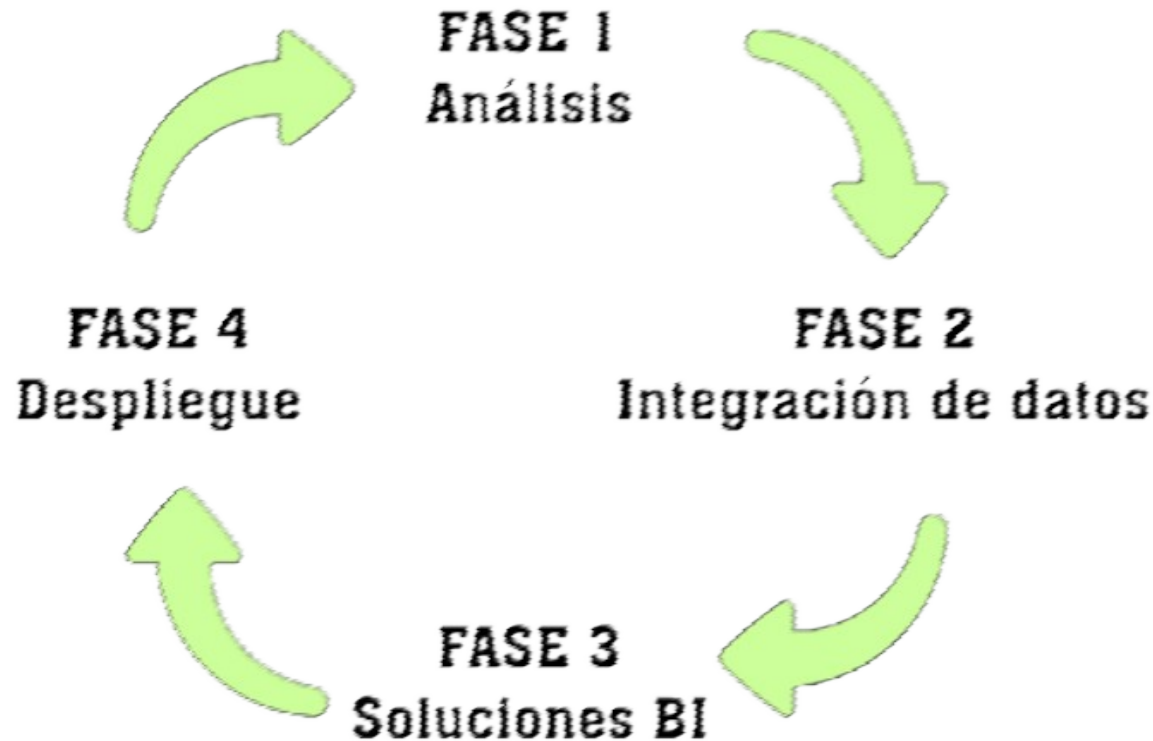


- Deben ser procesados y utilizados como fuente de información para la toma de decisiones mediante el uso de herramientas de presentación y estudio analítico.

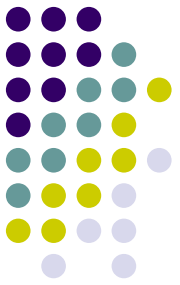
Data Warehousing Business Intelligence



Proceso BI

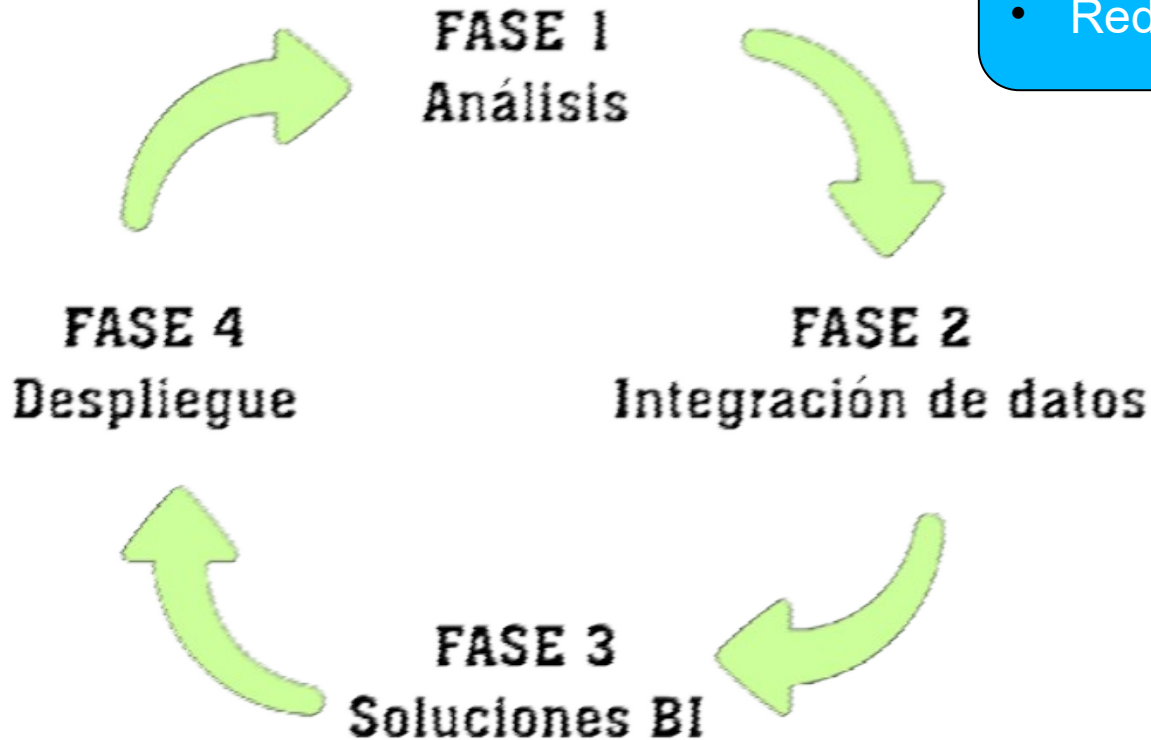


Data Warehousing Business Intelligence

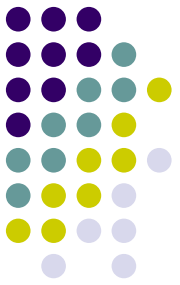


Proceso BI

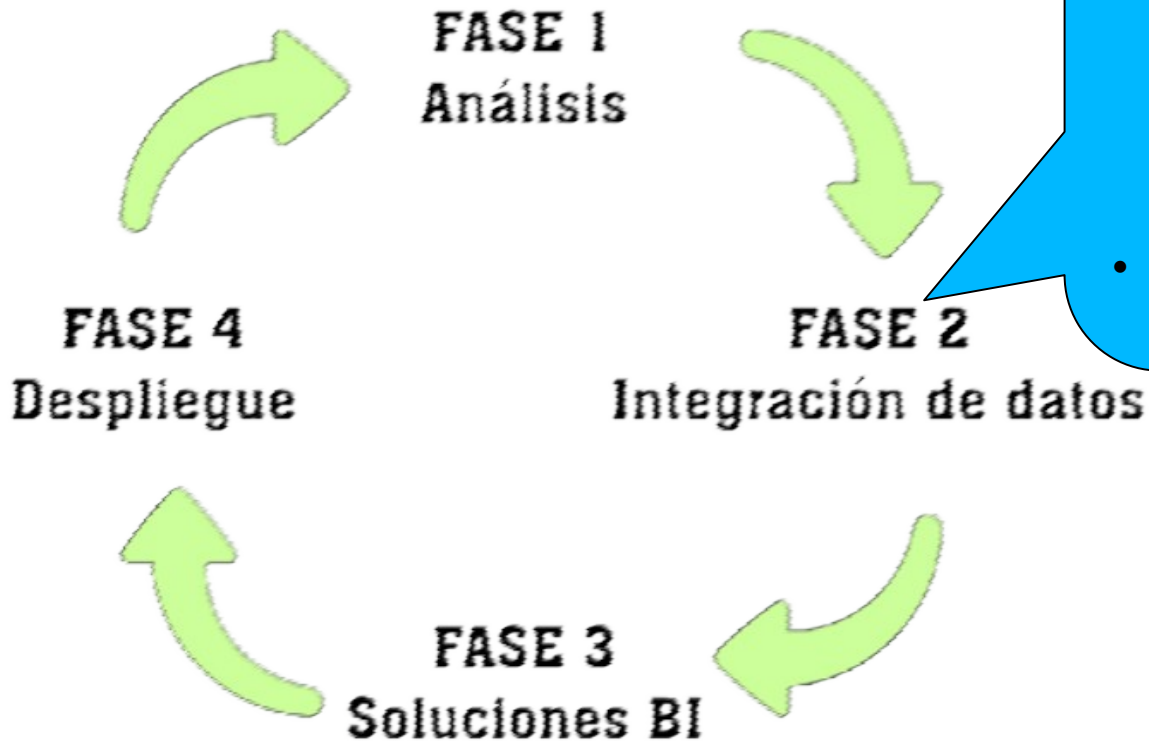
- Recolección de requerimientos
- Necesidades de información
- Redacción de preguntas



Data Warehousing Business Intelligence

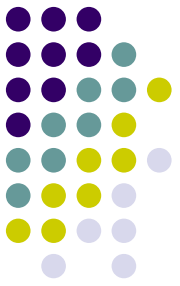


Proceso BI

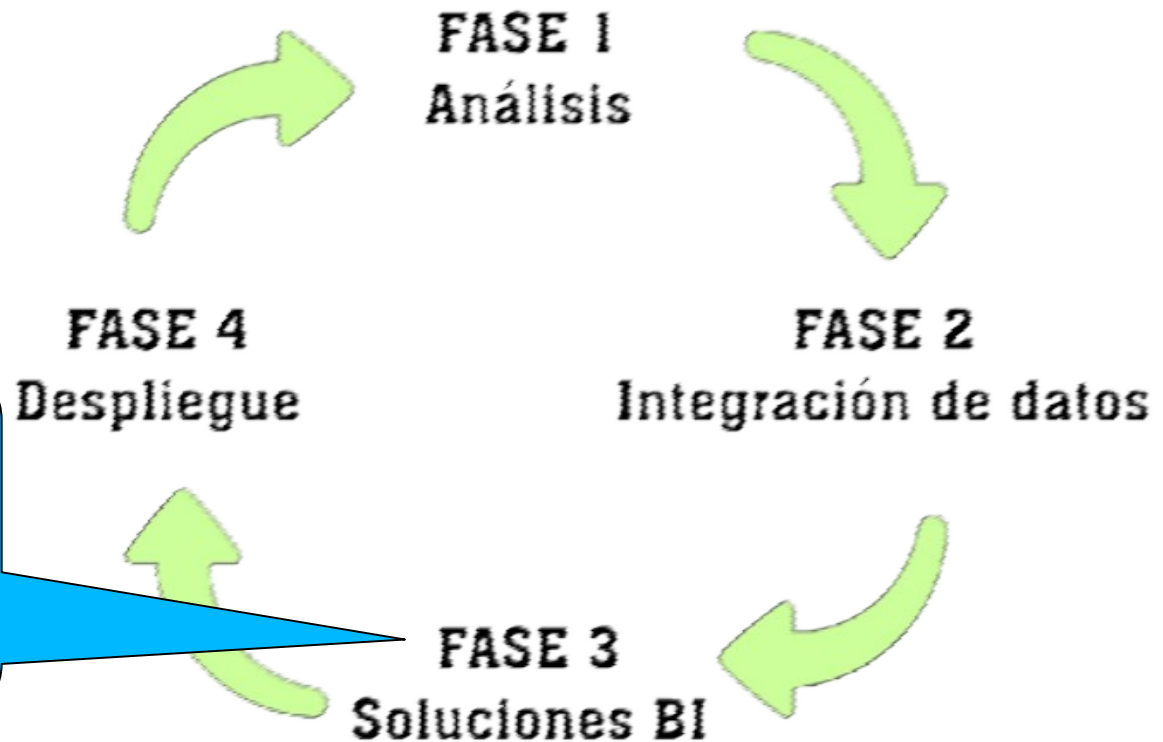


- Integración de datos **internos** y **externos**
- Transformación de esos datos que poseen la información para dar las respuestas a las preguntas de la Fase 1.
- Carga en la base de datos para servir al BI.

Data Warehousing Business Intelligence

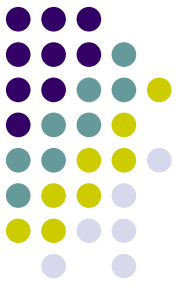


Proceso BI



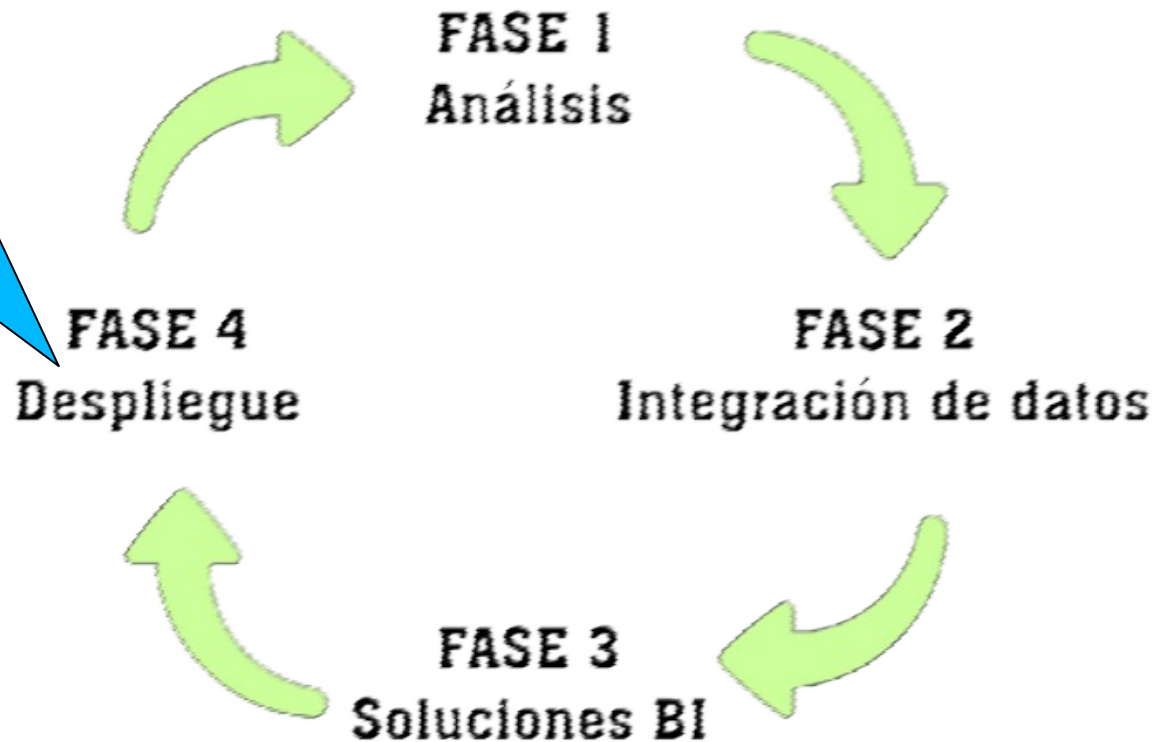
- Uso de herramientas y técnicas de explotación.
- Reportes, tableros, indicadores, gráficos estadísticos, etc..

Data Warehousing Business Intelligence

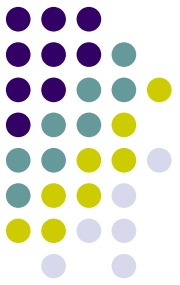


Proceso BI

- Entrega al usuario.
- Capacitación.
- Obtención de respuestas
- Nuevas preguntas ...



Data Warehousing Business Intelligence



Data Warehousing & Data Warehouse

Data Warehousing (DWH):

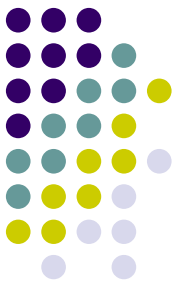
- **Proceso** → extracción, transformación, consolidación, integración y centralización de los datos **internos** y **externos relacionados** en una base de datos para permitir el análisis y exploración y dar soporte al proceso de toma de decisiones.

Data Warehouse (DW):

- **Base de datos** → almacenamiento de los datos para el análisis. Es una base de datos que posee una estructura multidimensional.



Data Warehousing Business Intelligence

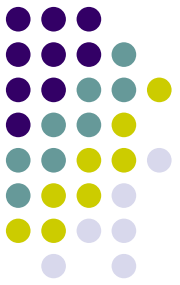


Data Warehousing & Data Warehouse

Data Warehouse (DW):



Data Warehousing Business Intelligence



Data Warehousing & Data Warehouse

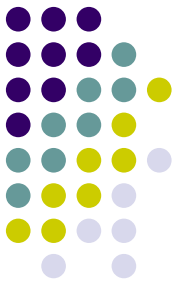
Data Warehouse (DW):



Sólo ingresarán **datos relevantes** , análisis y toma de decisiones. Muchos datos **NO son tenidos en cuenta para el DW** por carecer de valor analítico.

Se manejan conceptos de alto nivel en una estructura multidimensional que administre los datos de **entidades tipo PERMANENTE y de tipo MOVIMIENTO** en tablas denominadas **DIMENSIONES** y **HECHOS** respectivamente.

Data Warehousing Business Intelligence



Data Warehousing & Data Warehouse

Data Warehouse (DW):

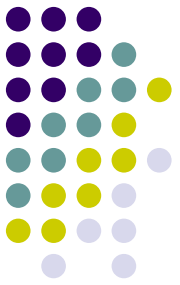


Todos los datos de orígenes heterogeneos se analizan para asegurar su calidad y pureza para consolidarse luego en el DW. Se materializa con las tareas de ETL.

Orígenes de los datos:

- Usuarios: Operacionales, Medios, Gerenciales.
- Áreas de la organización.
- Diferentes Data Sources: fuentes internas, fuentes externas.

Data Warehousing Business Intelligence



Data Warehousing & Data Warehouse

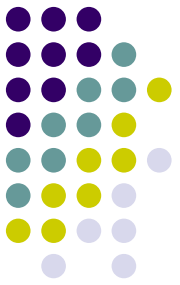
Data Warehouse (DW):



variante en
el tiempo

- Los datos actuales son almacenados junto a los datos históricos cada uno con su **marca de tiempo**.
- Permite acceder a distintas versiones de la misma situación, de esta manera, se podrá avanzar y retroceder en el tiempo manteniendo el foco de atención sobre la **situación analizada**.

Data Warehousing Business Intelligence



Data Warehousing & Data Warehouse

Data Warehouse (DW):



no volátil

La información **ES ESTABLE** → Los datos que ingresan al DW **NO CAMBIAN**.



DATA SOURCES



Consultar



Insertar



Eliminar



Modificar



DATA WAREHOUSE

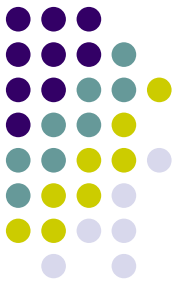


Consultar



Insertar

Data Warehousing Business Intelligence



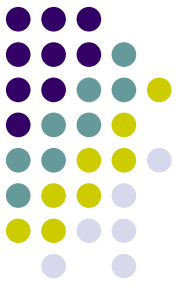
Data Warehousing & Data Warehouse

Granularidad:

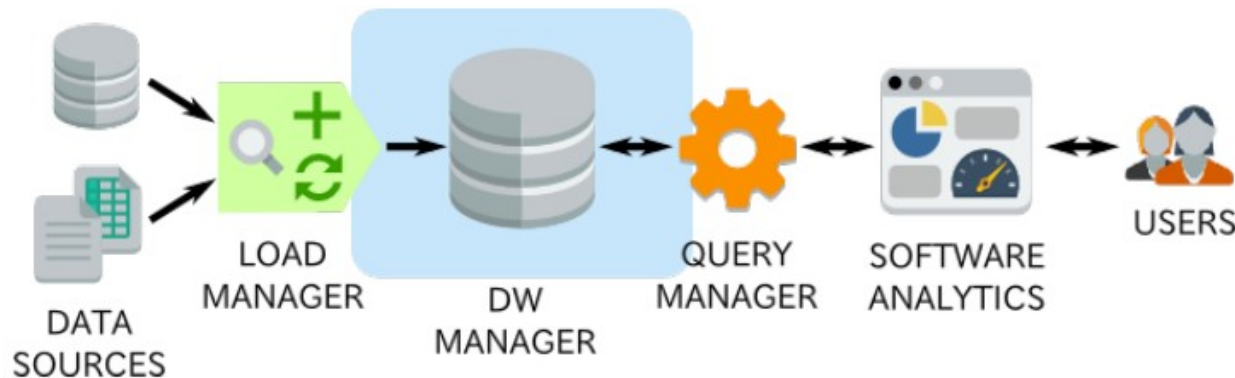
Es el nivel de detalle con el que se desea almacenar la información

- A mayor nivel de detalle de los datos, más posibilidades analíticas,
- Los datos que posean granularidad fina (nivel de detalle) podrán ser resumidos para obtener una granularidad media o gruesa.
- Trabaja juntamente con la agregación (cálculo que de varias filas produce un único resultado) que es una situación **NO REVERSIBLE**.

Data Warehousing Business Intelligence

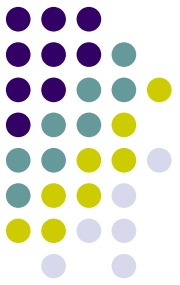


Arquitectura del Data Warehousing

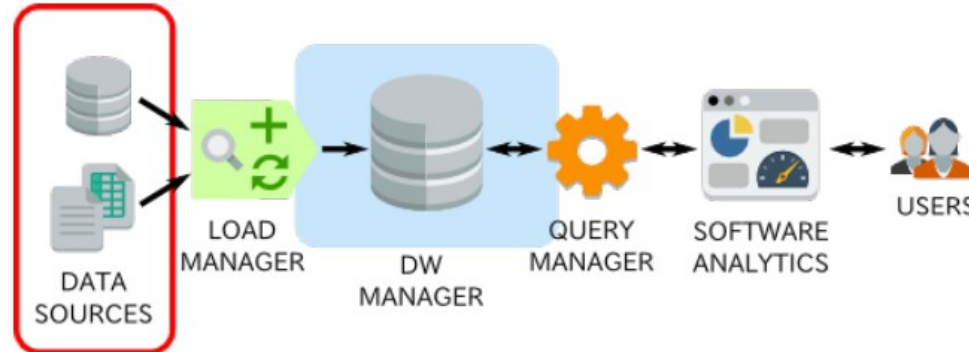


- Los datos son extraídos desde distintas fuentes.
- Los datos son integrados, limpiados, depurados y transformados y se cargan en el DW donde se construirán cubos multidimensionales y/o business models.
- Los usuarios accederán a los cubos multidimensionales o business models del DW, utilizando diversas herramientas de consulta, exploración, análisis, reportes, etc.

Data Warehousing Business Intelligence

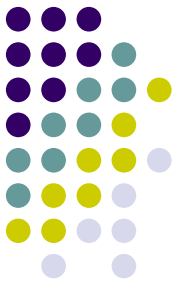


Arquitectura – Data Sources

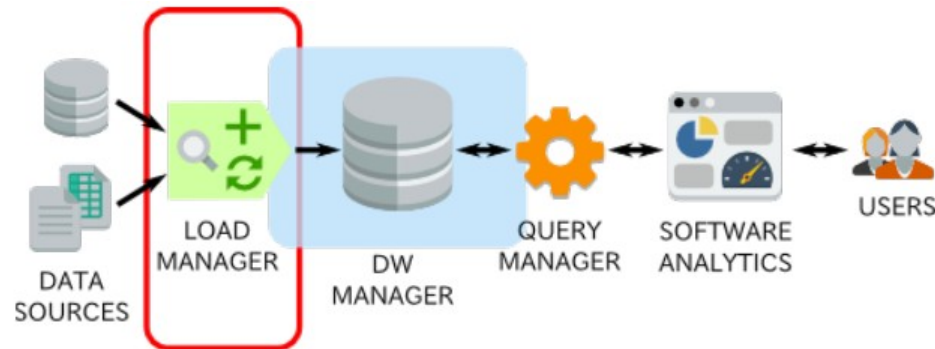


- Archivos de textos.
- Hojas de cálculos.
- Informes semanales, mensuales, anuales, etc.
- Bases de datos transaccionales, SQL y *NoSQL* .
- Información no estructurada (páginas web, mails) Redes sociales.
- Web Services.
- Etc...

Data Warehousing Business Intelligence

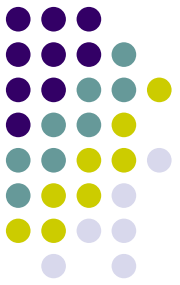


Arquitectura – Load Manager

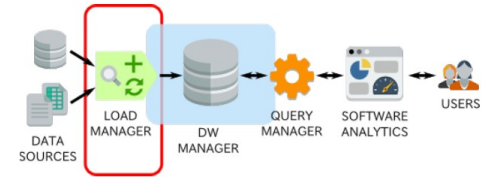


- **INTEGRACIÓN:** procesos de extracción y transformación (manipulación, control y depuración). Luego los resultados actualizan el DW (carga).
- La **Extracción** incluirá técnicas enfocadas a obtener desde diversas fuentes **solamente los datos relevantes** y mantenerlos en una **Staging Area**.
- La **Transformación** incluirá técnicas encargadas de compatibilizar formatos, filtrar y clasificar datos, relacionar diversas fuentes, depurar, etc.
- La **Carga** incluirá técnicas propias de la carga y actualización del DW.

Data Warehousing Business Intelligence



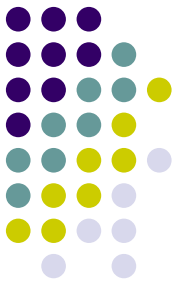
Arquitectura – Load Manager



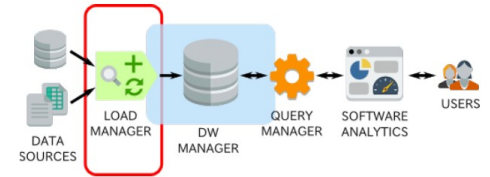
Extracción

- Seleccionar adecuadamente los datos relevantes.
- Si los Data Sources utilizan bases de datos relacionales, la extracción puede hacerse con consultas SQL o procedimientos almacenados.
- Los casos NO convencionales o NO estructurados la obtención se hará con elementos variados específicos.
- Una vez extraídos, los datos se persisten en una base de datos relacional **Staging** desconectándose de los Data Sources

Data Warehousing Business Intelligence



Arquitectura – Load Manager

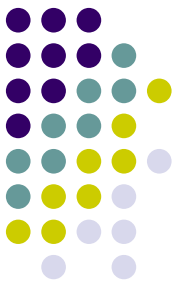


Extracción

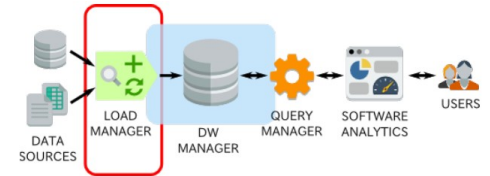
BASE DE DATOS DEL STAGING AREA

- Pueden ser varias bases de datos, esquemas o particiones dependiendo del DBMS a utilizar.
- Permiten manipular los datos sin interrumpir ni sobrecargar los Data Sources y el DW.
- Crean una capa de abstracción entre la lectura y la carga.
- Cuando se terminan las transformaciones en éste área continúa la población del DW.

Data Warehousing Business Intelligence

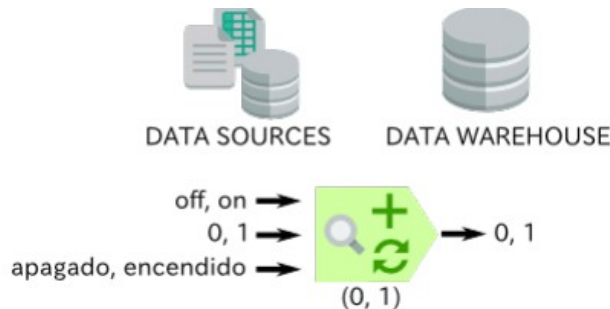


Arquitectura – Load Manager



Transformación →

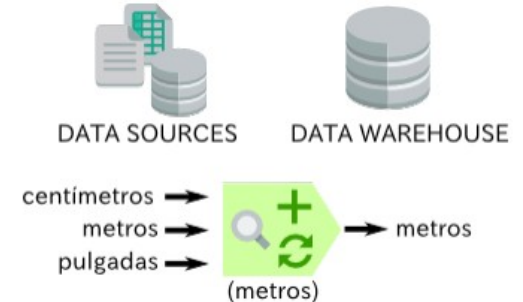
Datos Consistentes,
compatibles y congruentes



Codificación

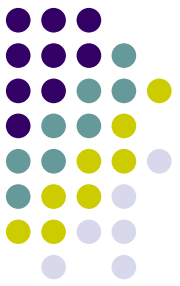


Fuentes múltiples

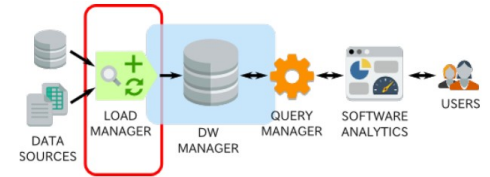


Medida de atributos

Data Warehousing Business Intelligence



Arquitectura – Load Manager



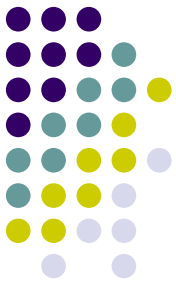
Transformación

- Procesos de **Limpieza** y **Calidad** de los datos.
- Identificar la razón de la anomalía y evitar repeticiones.

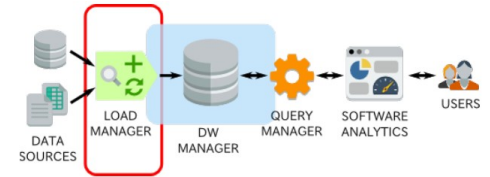
Acciones ante datos **anómalos** o **faltantes**:

- Ignorarlos.
- Eliminar la columna (se eliminan el 100% de los datos).
- Filtrar la columna (se eliminan algunos de los datos).
- Filtrar la fila errónea,
- Reemplazar el valor.
- Si falta, esperar a que esté disponible.
- Discretizar los valores de las columnas. Si una columna numérica presenta un valor de 1 a 2, se utiliza el texto **Bajo**; de 3 a 7, **Medio**; de 8 a 10, **Alto**. Al ocurrir un outlier se puede reemplazar por Bajo o Alto.

Data Warehousing Business Intelligence



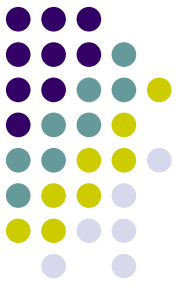
Arquitectura – Load Manager



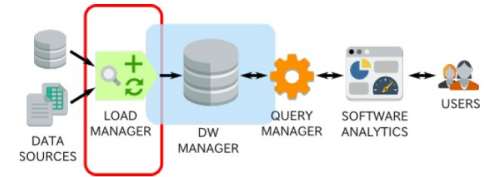
Carga

- **Carga inicial.**
- **Actualización periódica:**
 - Cotejar instancias de los Data Sources involucrados.
 - Uso de *triggers*
 - Utilizar *marcas de tiempo* en los Data Sources.
 - Comparar los datos entre Data Source y DW.
 - Técnicas mixtas.
 - **FULL LOAD** (vaciado previo)
- **Mantener elementos estructurales del DW:**
 - Claves subrogadas.
 - SCD.
 - Etc.

Data Warehousing Business Intelligence



Arquitectura – Load Manager

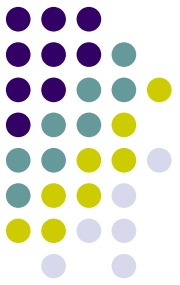


Resumen → ETL

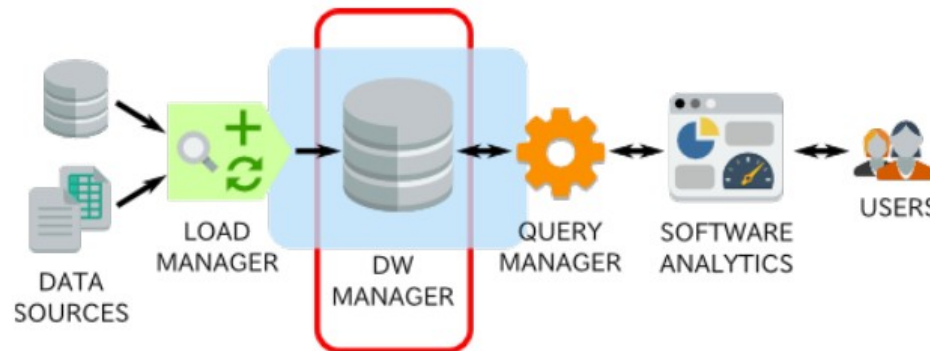
- Extracción de los datos desde Data Sources que se depositan en la **Staging Area**.
- Integración y transformación de los datos – tratamiento de las inconsistencias.
- Carga de los datos desde **Staging Area** al DW



Data Warehousing Business Intelligence



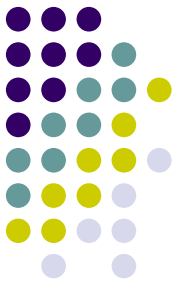
Arquitectura – DW Manager



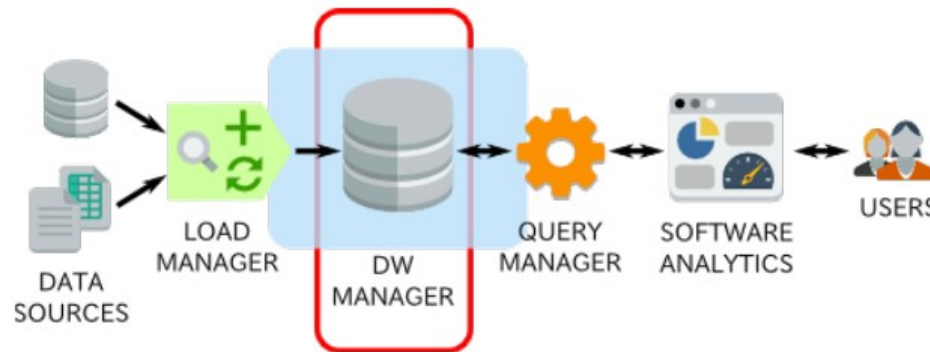
Compuesto por:

- el DW (DBMS),
- las conexiones a base de datos y otros Data Sources,
- las estructuras de datos (cubos multidimensionales, business models),
- información de autenticación y autorización (credenciales de acceso, users, roles, permisos, etc), y otros metadatos.

Data Warehousing Business Intelligence



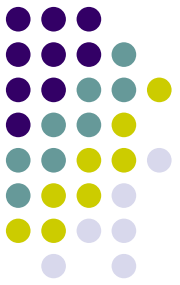
Arquitectura – DW Manager



Es una base de datos con estructura multidimensional que es una forma de almacenamiento con dos elementos principales:

- ✓ **Tablas de Hechos**
- ✓ **Tablas de Dimensiones**

Data Warehousing Business Intelligence



Modelos básicos



Star scheme (esquema en estrella)



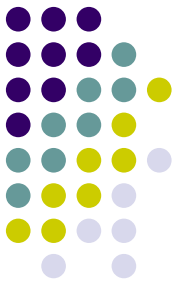
Snowflake scheme (esquema copo de nieve)



Starflake scheme (esquema constelación)

Estos modelos permiten facilitar el acceso a consultas complejas y con gran cantidad de agregaciones, es por ello que se encuentran **desnormalizadas** o **semi desnormalizadas**, reduciendo de esta manera al mínimo la cantidad de JOINS que deben emplearse para acceder a los datos requeridos.

Data Warehousing Business Intelligence



Tablas de Dimensiones

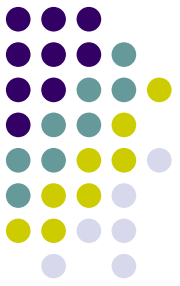
- Permiten analizar los datos en el contexto del negocio.
- Datos cualitativos
- Aspectos de interés para filtrar
- Perspectivas de visión

Columnas:

- Clave principal (primary key – por ejemplo **id**)
- Claves ajenas (solo para esquemas copo de nieve y constelación).
- Datos de referencia **primarios**: datos que identifican la Dimensión de manera básica (por ejemplo **nombre_infraccion**).
- Datos de referencia **secundarios**: datos que complementan la descripción de la Dimensión (por ejemplo **es_grave**) Estos datos no son significativos para tomar decisiones, pero son potencialmente valiosos para implementarla.

d_infracciones	
123	id
ABC	nombre_infraccion
ABC	es_grave

Data Warehousing Business Intelligence



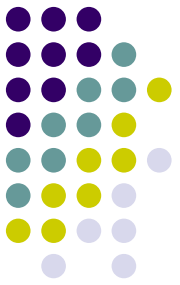
Tablas de Dimensiones

Recomendación !!!!!

Las claves principales de las tablas de Dimensiones deben ser **independientes** de las claves de los Data Sources ya que, si estos últimos son recodificados, el DW quedaría inconsistente.

Estas claves serán Claves Subrogadas

Data Warehousing Business Intelligence



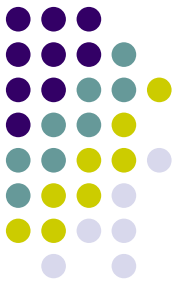
Dimensiones especiales

La dimensión **TIEMPO**

No es solamente una secuencia cronológica representada de forma numérica, sino que mantiene niveles jerárquicos especiales que permiten por ejemplo, analizar infracciones de tránsito ocurridas en determinado día de la semana, quincena, mes, trimestre, semestre, año, estación, etc.... **Y LA HORA** (o rangos).

d_fecha	
123	id
🕒	fecha
123	epoch
123	día_del_mes
ABC	día_nombre
ABC	día_nombre_corto
123	día_semana
123	día_trimestre
123	día_año
123	semana_mes
123	semana_año
123	mes_numero
ABC	mes_nombre
ABC	mes_nombre_corto
123	trimestre_numero
ABC	trimestre_nombre
ABC	trimestre_nombre_corto
ABC	bimestre_nombre
ABC	bimestre_nombre_corto
ABC	semestre_nombre
123	año_actual
🕒	fecha_primerdia_semana
🕒	fecha_ultimodia_semana
🕒	fecha_comienzo_trimestre
🕒	fecha_fin_trimestre
ABC	mes2_año4
ABC	mes2_día2_año4
<input checked="" type="checkbox"/>	es_finde
<input checked="" type="checkbox"/>	es_feriado
ABC	nombre_feriado
ABC	fechar

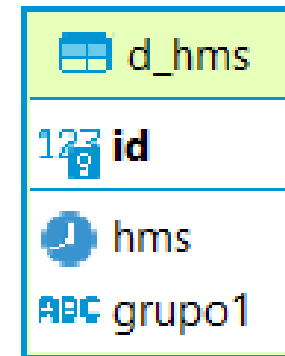
Data Warehousing Business Intelligence



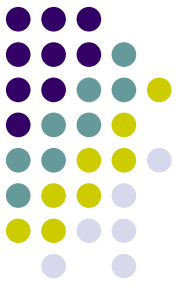
Dimensiones especiales

La dimensión **HORA**

Cuando se requiere la hora, se debe crear una dimensión independiente de la de fecha y conjuntamente con ésta última, determinan el momento de cualquier ocurrencia.



Data Warehousing Business Intelligence

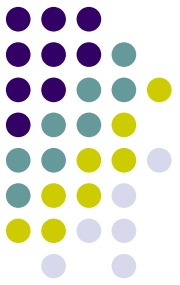


Tablas de Hechos

- Las tablas de Hechos contienen los **acontecimientos** o **eventos** que serán utilizados por los usuarios del DW para analizar y responder preguntas de negocio.
- Los Hechos contienen datos que son filtrados, agrupados y explorados a través de condiciones definidas en las tablas de Dimensiones.
- La tabla de Hechos posee un identificativo compuesto por la **concatenación de las claves de las tablas de las Dimensiones relacionadas**. No requiere de la definición de una primary key.

h_actas
ABC id_acta
ABC id_dominio
123 id_fecha
123 id_hms
123 id_sexo
123 id_edad
123 id_infraccion
123 id_tipo_vehiculo
123 id_estado
123 id_actor
123 id_origen
123 id_marca
123 id_modelo
ABC tiene_titulo_ejecutivo
ABC es_liquidacion_web
ABC pagada
ABC otra_jurisdccion
ABC fue_retenido
123 latitud
123 longitud
123 importe_teorico
123 importe_abonado
ABC tiene_georreferencia

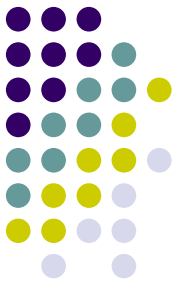
Data Warehousing Business Intelligence



Tablas de Hechos

- Se denominan también **hecho** a los atributos que cuantifican o califican al evento u acontecimiento que representa el Hecho.
- Los **hechos básicos** son los que están como columna en la tabla.
- Los **hechos derivados** son los que se calculan a partir de una expresión combinando hechos básicos.

Data Warehousing Business Intelligence



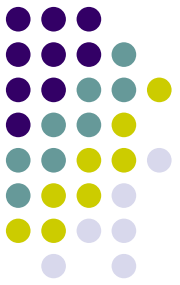
Tablas de Hechos

Indicadores/medidas

Hacen uso de los datos de los hechos para obtener valores analizables y se definen con los siguientes metadatos:

- nombre representativo, descripción;
- tipo de agregación al momento de crear una estructura de datos (SUM, MAX, MIN, COUNT, AVG, porcentajes, fórmulas, etc.)
- agregaciones alternativas;
- tipo de datos (siempre numéricos).
- Se calculan dinámicamente sin necesidad de persistir en la estructura de datos.

Data Warehousing Business Intelligence



Conceptos

Medidas:

son datos operativos que pueden estar agregados o detallados. Representan un hecho de la realidad; por ejemplo, una venta o descuento.

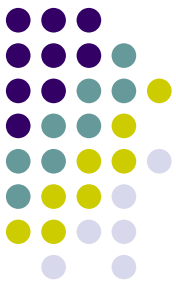
Métricas:

son valores cuantitativos que miden un desempeño. Son medidas dentro de un contexto definido. Continuando con el ejemplo de venta, podemos tener “venta año anterior” o “venta promedio”.

KPI (indicador):

son métricas o medidas ponderadas y evaluadas. Se necesita definir qué valor es “bueno”, cuál es “malo” y cuál es “aceptable”. Sobre la base de estas premisas podemos analizar el estado, su tendencia y su variación respecto a lo planificado o proyectado.

Data Warehousing Business Intelligence



estado
123 cod_estado
ABC descrip_estado
🕒 fecha_caducidad

t_vehiculos
123 cod_vehiculo
ABC descrip_vehiculo

marcas
123 cod_marca
ABC descrip_marca
123 cod_marca_altec
123 en_gaem

modelos
123 id_modelo
123 id_marca
ABC descripcion

actas_por_liquidacion
123 id_liquidacion
ABC id_acta
123 estado_acta
123 monto_acta
123 gasto_notificacion
123 descuento_monto

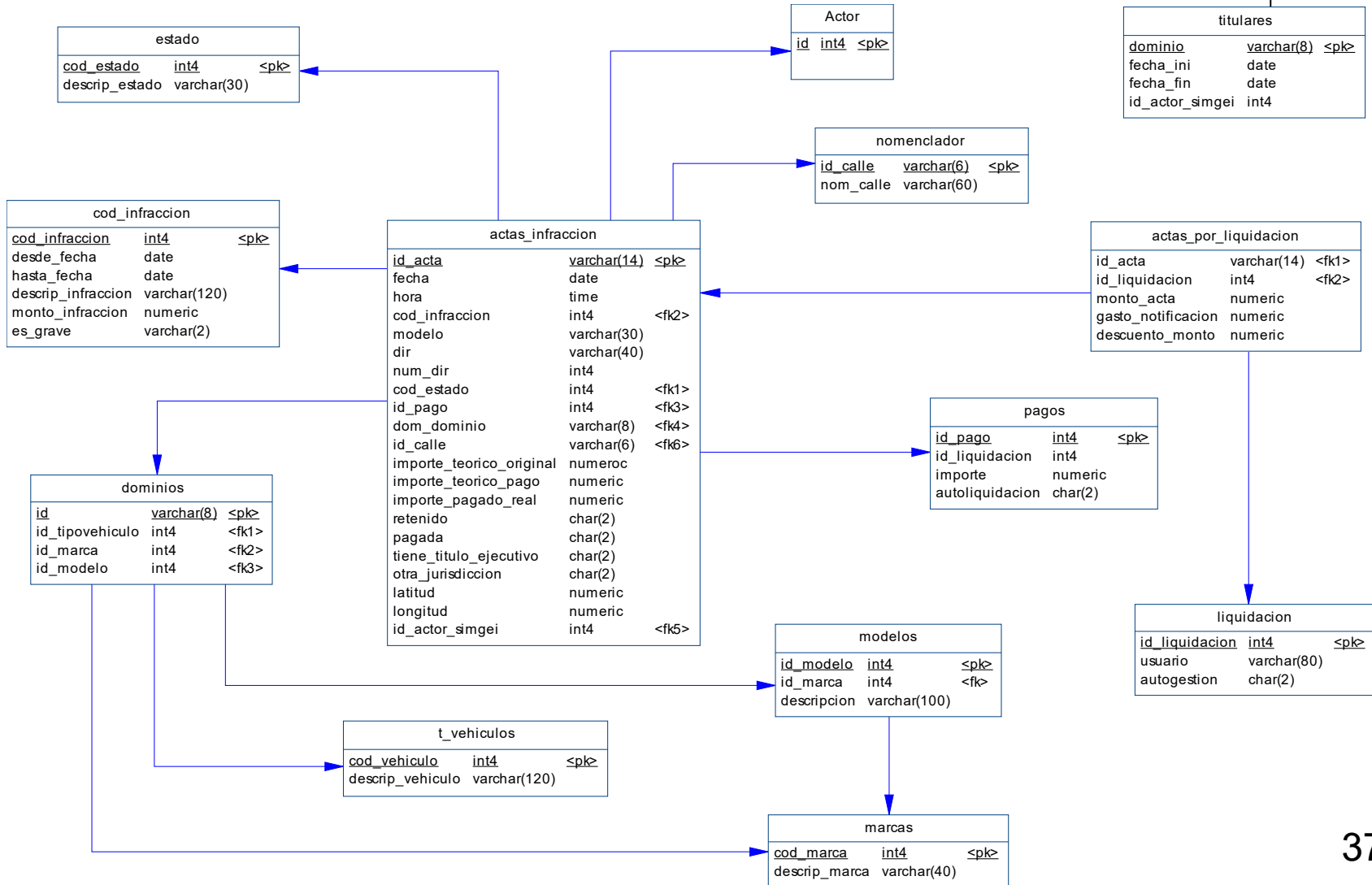
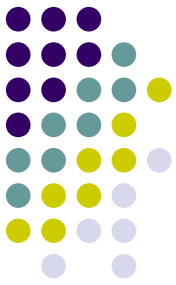
liquidacion
123 id_liquidacion
🕒 fecha_liquidacion
🕒 fecha_vencimiento
123 estado
ABC dominio
123 multa
123 sellados
123 notificaciones
123 estadia
123 grua
123 tipo_doc
123 cuit_cuil_tit
ABC nombre
ABC domicilio
123 juzgado
123 tipo_pago
ABC articulos
ABC acceso
ABC observaciones
123 descuento_monto
ABC usuario
ABC barra
123 id_sentencia

actas
ABC id_acta
ABC dominio
🕒 fecha
🕒 hora
123 cod_inspector
123 cod_infraccion1
123 cod_infraccion2
123 cod_infraccion3
123 cod_infraccion4
123 cod_infraccion5
🕒 fecha_entrada
🕒 fecha_carga
123 cod_vehiculo
123 cod_marca
ABC modelo
ABC dir
123 num_dir
ABC retenido
123 num_juzgado
ABC observacion
123 cod_estado
123 cod_noti
123 cod_cedula
🕒 fecha_pago
123 id_pago
123 redo_pago
123 cod_lugar
123 cod_pago
123 protocolo
123 notificar
123 coincide_api
123 informada
ABC usuario
123 ticket_prorroga
123 nro_parte
123 lote_fotomulta

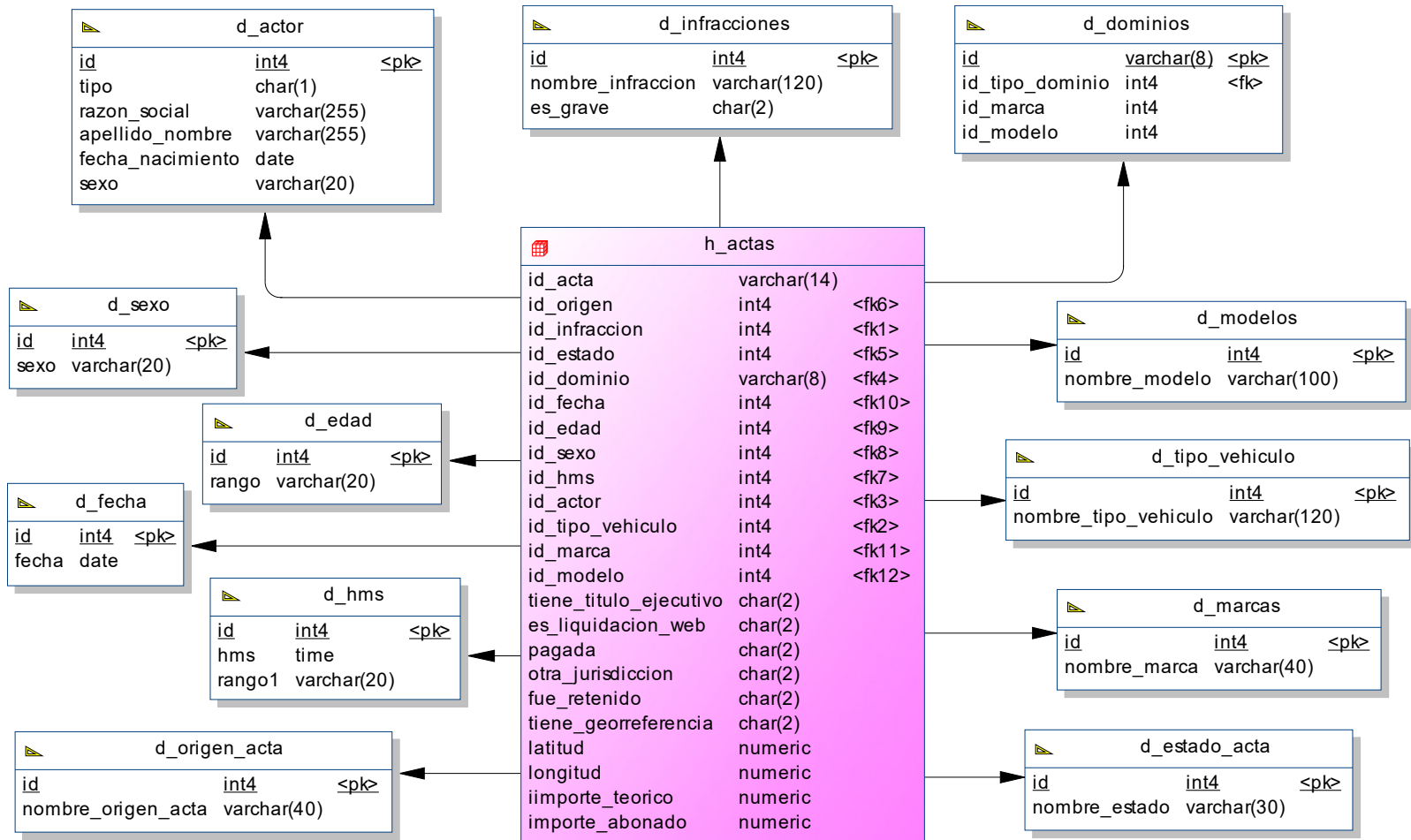
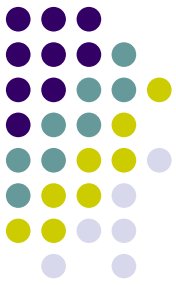
cod_infraccion
123 cod_infraccion
🕒 desde_fecha
🕒 hasta_fecha
ABC articulo
123 juez
123 ordenanza
123 decreto
ABC descrip_infraccion
123 monto_max
123 monto_min
123 monto_ext
123 monto_infraccion
ABC inf_grave
ABC no_mostrar_importe
123 no_imprimible

titulares
ABC dominio
🕒 fecha_ini
🕒 fecha_fin
ABC tipo_titular
ABC tipo_persona
ABC genero
123 tipo_documento
123 cuit_cuil_tit
123 cuit
123 id_actor_simgei
ABC nombre
ABC domicilio_c
123 domicilio_a
123 postal
ABC localidad
ABC observacion
ABC titular_infractor
ABC econotificable

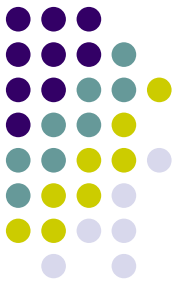
Data Warehousing Business Intelligence



Data Warehousing Business Intelligence



Data Warehousing Business Intelligence

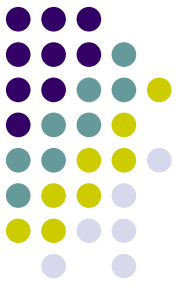


Tipos de tablas de Hechos

Transaction Fact Tables: Representan eventos que suceden en un determinado espacio-tiempo. Se caracterizan por permitir analizar los datos con el máximo detalle.

Factless Fact Tables/Coverage Tables: Son tablas **que no tienen medidas** y tiene sentido dado que representan el hecho que el evento suceda. Suelen asociarse medidas que permita facilitar las consultas requeridas.

Data Warehousing Business Intelligence



Cubo dimensional

Representa los datos planos en una matriz de N-dimensional.

Se componen de:

Indicadores

- Valores analizables calculados con operaciones sobre los Hechos.

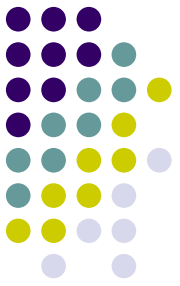
Atributos

- Criterios de análisis mediante los cuales se analizan los Indicadores. Se obtienen de los atributos de las dimensiones.

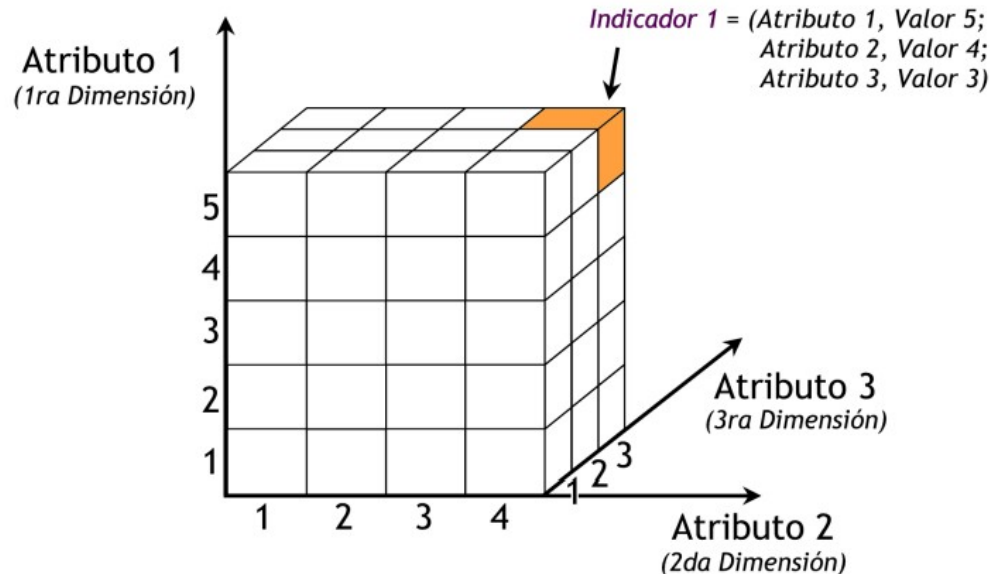
Jerarquías

- Relación lógica del tipo padre-hijo entre los Atributos que permiten analizar los datos desde el nivel más general al más detallado y viceversa (navegación en profundidad de los Hechos). Las Jerarquías manejan el nivel de agregación de los Hechos.

Data Warehousing Business Intelligence



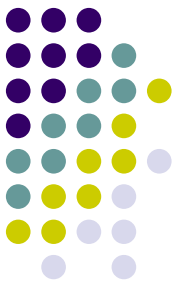
Cubo dimensional



El **Indicador 1**, representa el cruce del valor **5** de **Atributo 1**, el valor **4** de **Atributo 2** y el valor **3** de **Atributo 3**.

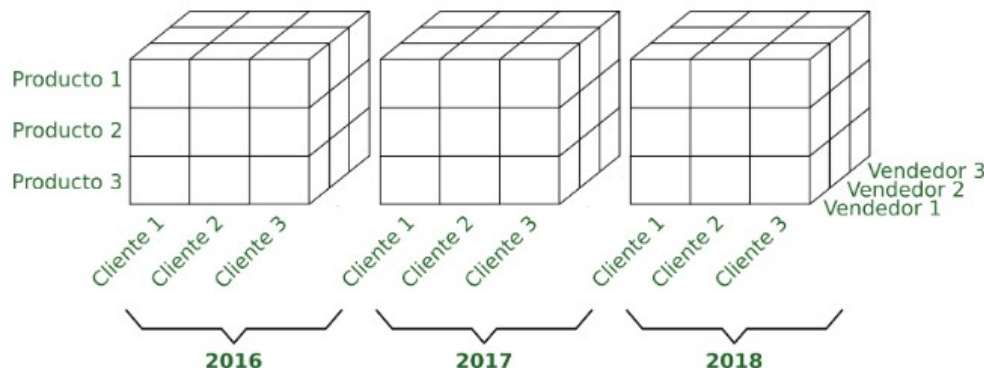
El resultado del análisis está dado por los cruces matriciales de acuerdo con los valores de las **Dimensiones** seleccionadas.

Data Warehousing Business Intelligence

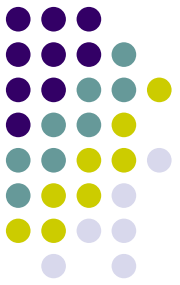


Cubo dimensional

1. El **CUBO ES EL HECHO** a analizar que corresponde a la tabla Ventas.
2. Las dimensiones se asocian a las tablas de cada una y se crean Jerarquías que utilizan alguna columna de la tabla.
3. Se crean los **Indicadores (Cantidad)** que puede estar asociado a un atributo de la tabla del CUBO.



Data Warehousing Business Intelligence



Business Models

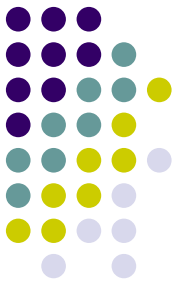
Son una representación de los datos desde una perspectiva empresarial, viendo la información y su interrelación como entidades de ALTO NIVEL.

Se componen de:

- Entidades
- Atributos
- Relaciones

Son similares en su comportamiento a los cubos dimensionales.

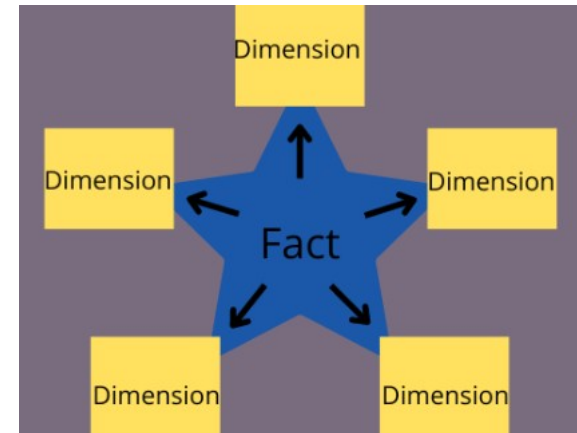
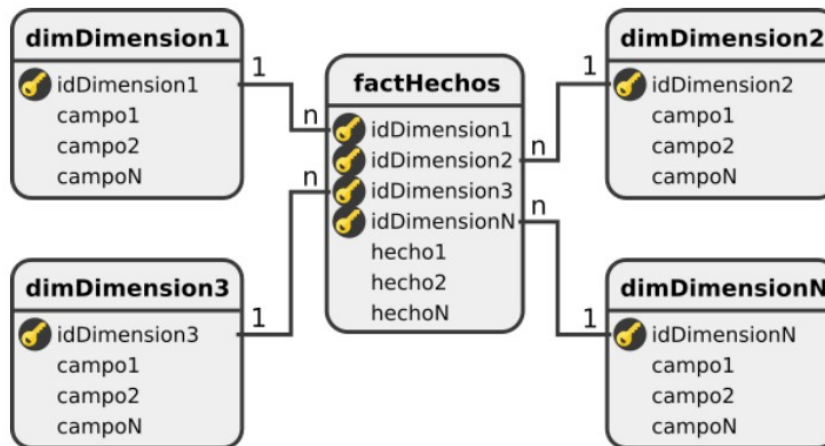
Data Warehousing Business Intelligence



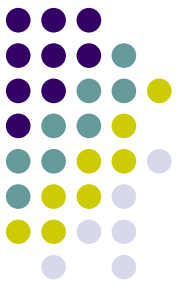
Modelos del DW

Esquema ESTRELLA

- **UNA** tabla de Hechos y
- **UNA o más** tablas de Dimensiones relacionadas a través de sus respectivas claves.

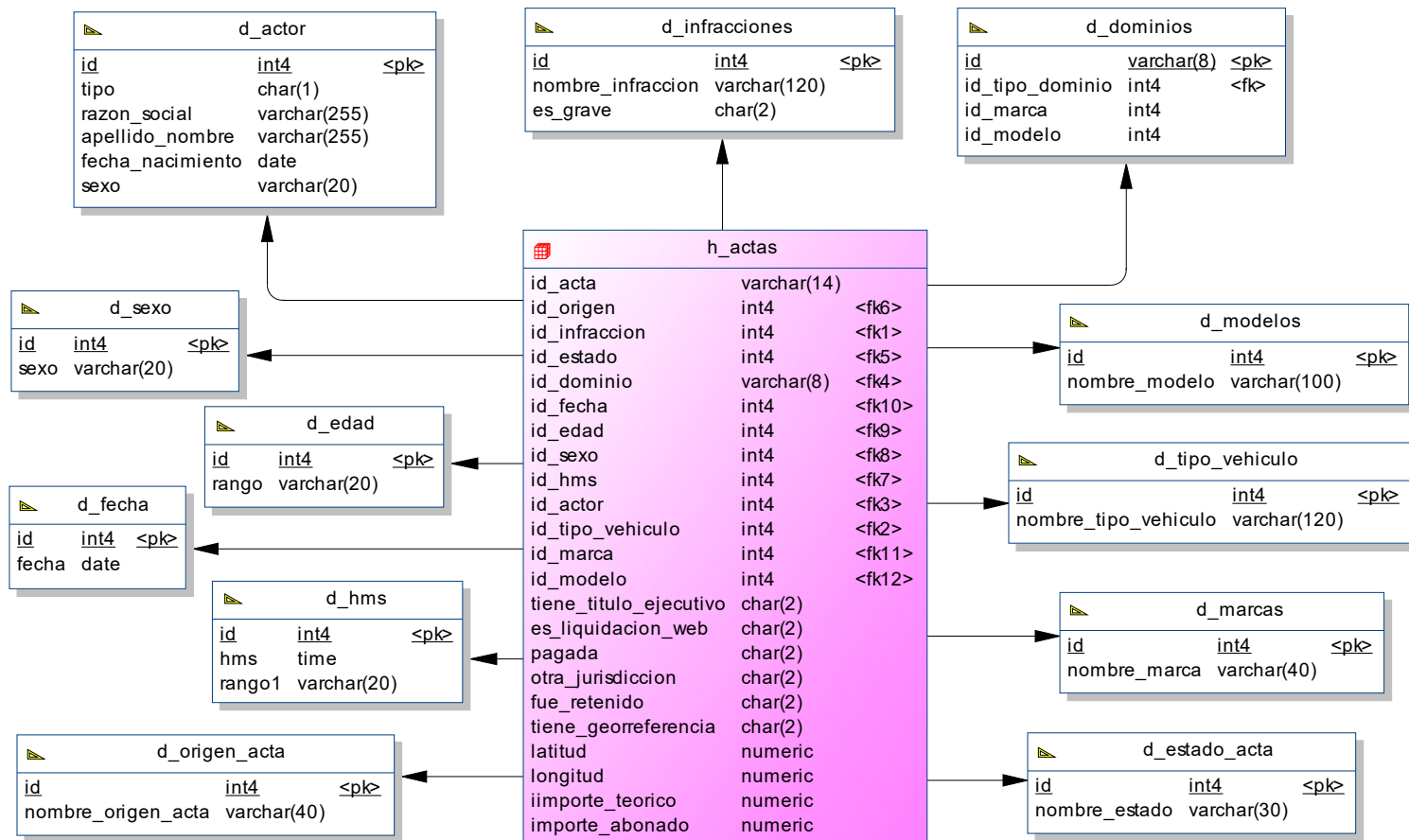


Data Warehousing Business Intelligence

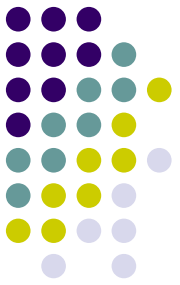


Modelos del DW

Esquema ESTRELLA



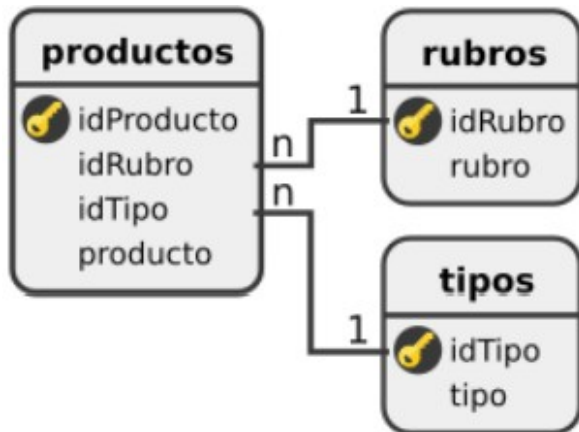
Data Warehousing Business Intelligence



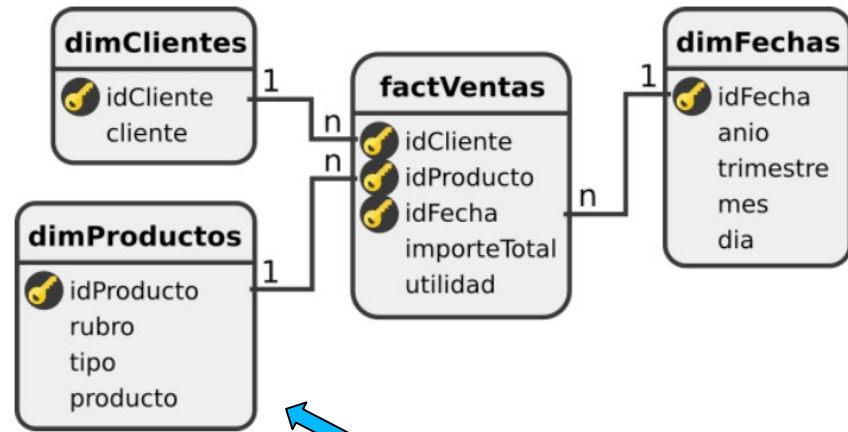
Modelos del DW

Esquema ESTRELLA

Data Source

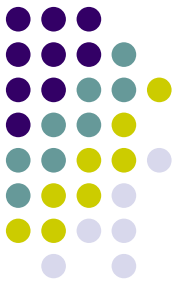


Modelo dimensional



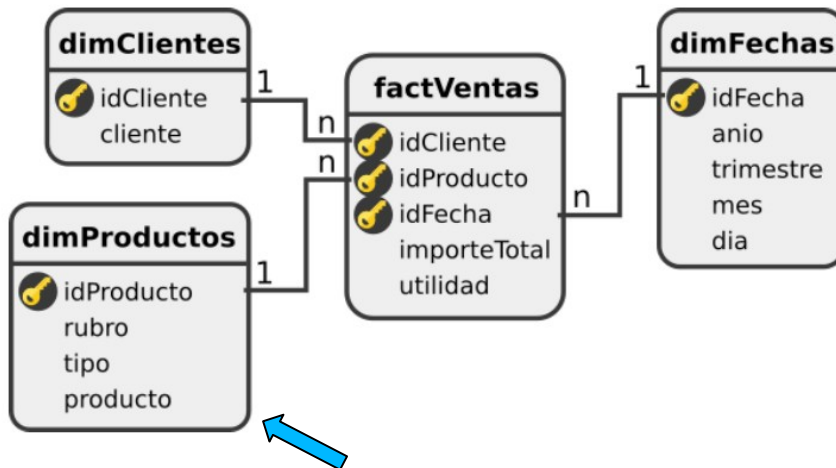
Desnormalizada

Data Warehousing Business Intelligence



Modelos del DW

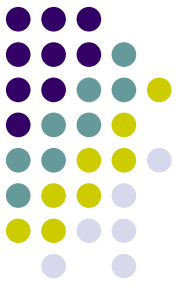
Esquema ESTRELLA



Desnormalizada

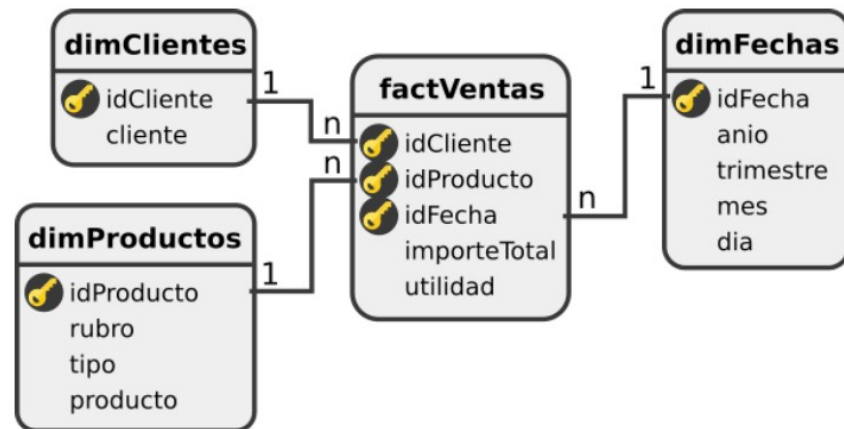
Los modelos multidimensionales requieren la desnormalización criteriosa para evitar uniones (JOIN) entre las tablas para las consultas.

Data Warehousing Business Intelligence



Modelos del DW

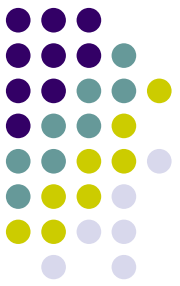
Esquema ESTRELLA



Características

- Es el más simple de interpretar.
- Posee los mejores tiempos de respuesta.
- Es soportado por todos los visores.
- Su diseño es sencillo de mantener y actualizar.
- Existe paralelismo entre su diseño y la forma que los usuarios visualizan y manipulan los datos.

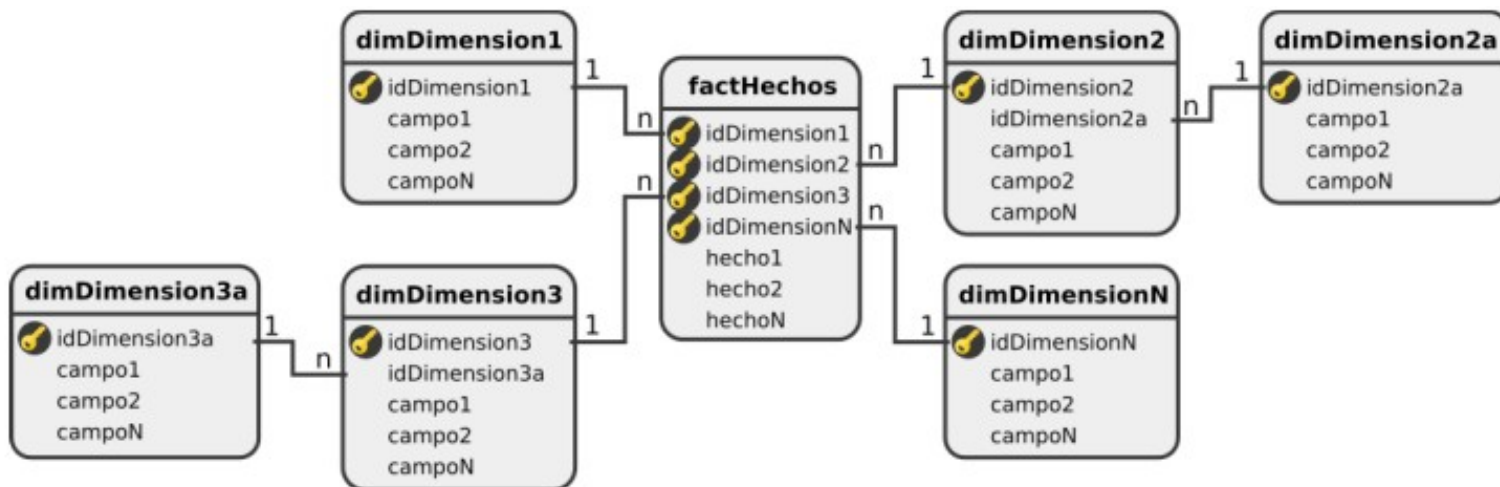
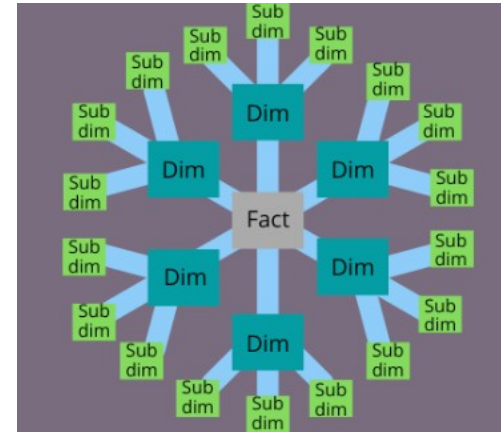
Data Warehousing Business Intelligence



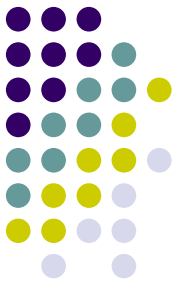
Modelos del DW

Esquema COPO DE NIEVE

- Extensión del esquema en ESTRELLA
- **UNA** tabla de Hechos
- **Más de una** tabla de Dimensión organizadas en jerarquías.

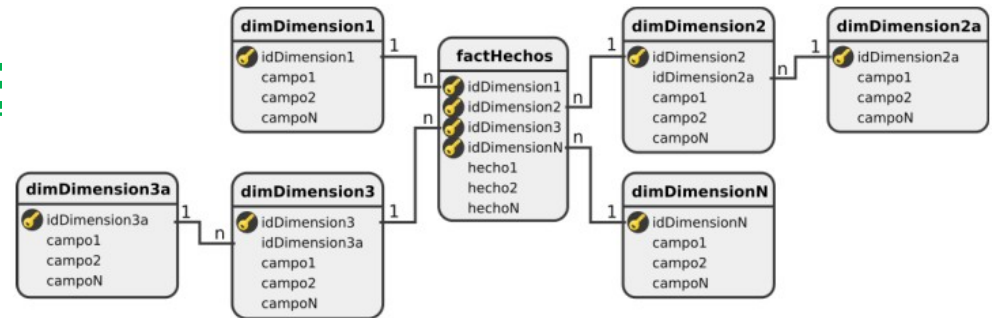


Data Warehousing Business Intelligence



Modelos del DW

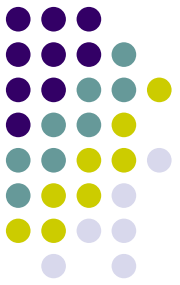
Esquema COPO DE NIEVE



Características

- Posibilita la segregación de los datos de las tablas de Dimensiones.
- Puede implementarse después de que se haya desarrollado un Esquema en Estrella.
- Posee mayor complejidad en su estructura.
- Utiliza menos espacio de almacenamiento.
- Es más eficiente en el caso de tablas de Dimensiones con gran cantidad de registros.
- Su semántica se ajusta a las representaciones de las diferentes Jerarquías de Dimensiones. Se deben planificar correctamente las uniones y el indexado, a fin de NO generar sobrecarga en la resolución de consultas.

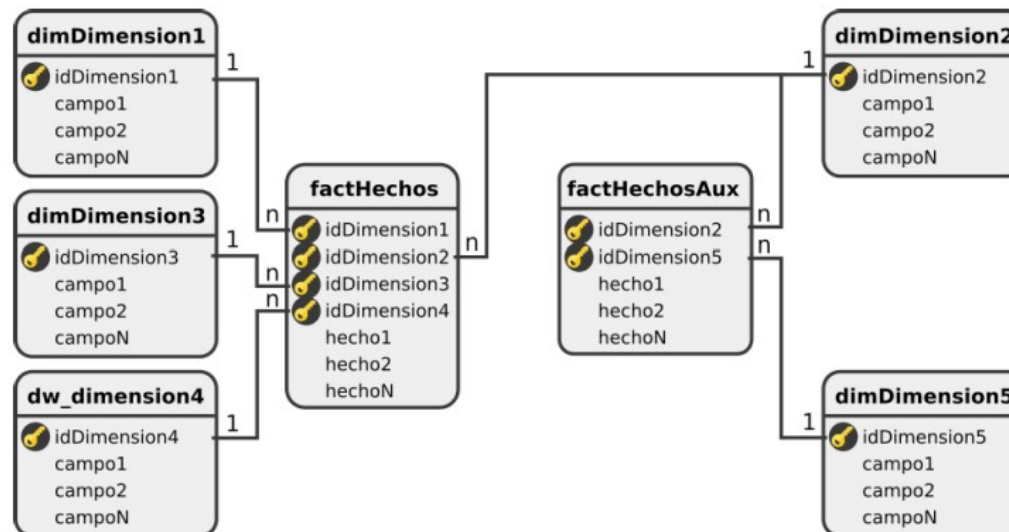
Data Warehousing Business Intelligence



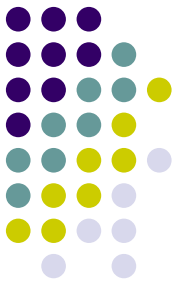
Modelos del DW

Esquema CONSTELACION

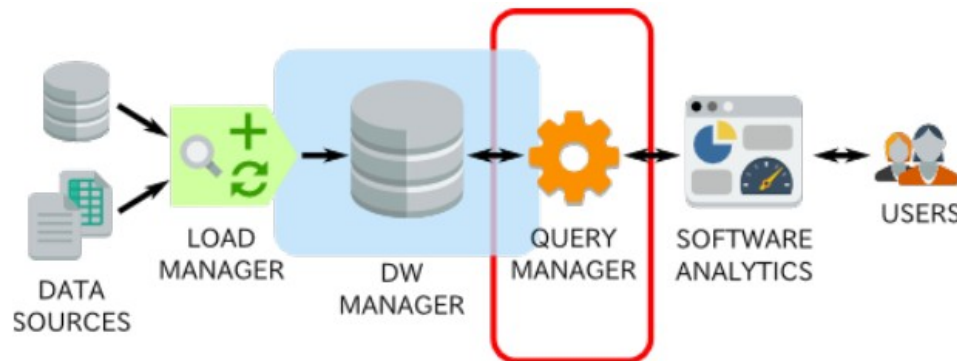
- Compuesto por varios esquemas en ESTRELLA
- **UNA** tabla de Hechos principal y **UNA o más** tablas de Hechos auxiliares (pueden ser agregaciones de la principal).
- **UNA o más** tablas de Dimensiones relacionadas a través de sus respectivas claves.



Data Warehousing Business Intelligence



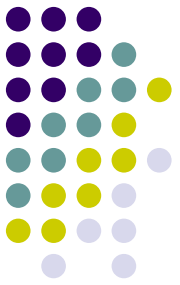
Arquitectura – Query Manager



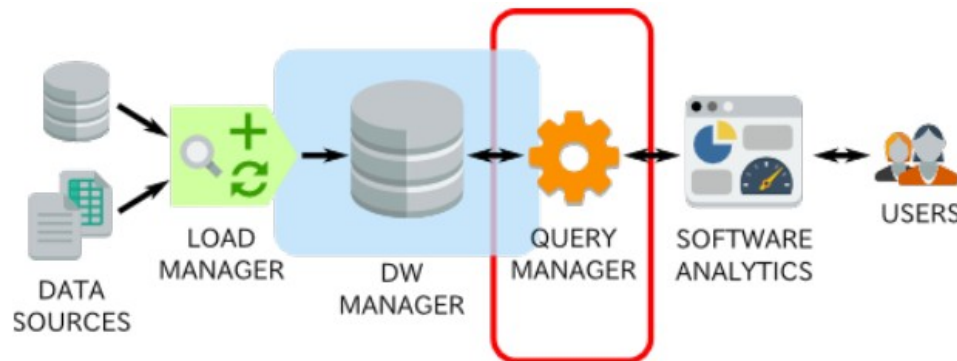
El Query Manager se encarga de resolver:

- consultas relacionales: como JOIN y agregaciones (SUM, COUNT, AVG, etc.)
- consultas propias del análisis de datos (DRILL-UP, DRILL-DOWN, DRILL-ACROSS, etc.)

Data Warehousing Business Intelligence

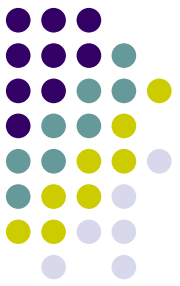


Arquitectura – Query Manager



1. Recibe consultas de los usuarios escritas en un lenguaje de alto nivel;
2. Lee los metadatos que describen los mapeos y reescribe las consultas para que sean ejecutadas en el sistema destino (por lo general SQL).
3. Obtiene los datos y utilizando nuevamente las estructuras de metadatos, los

Data Warehousing Business Intelligence

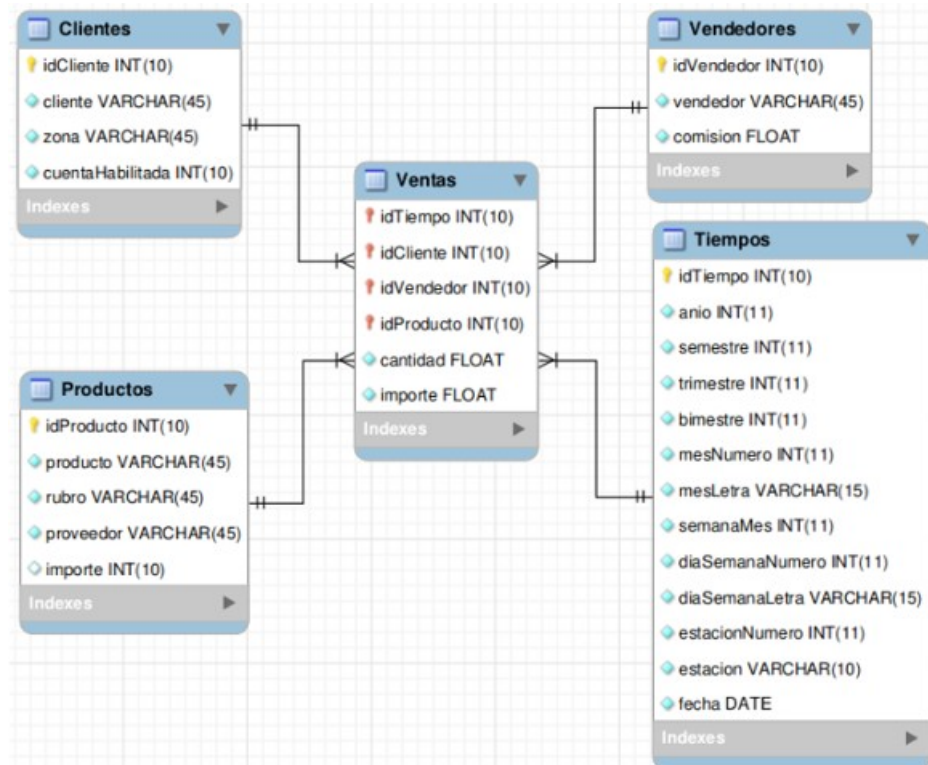


Arquitectura – Query Manager

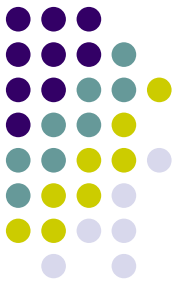


- Las principales operaciones que se realizan sobre modelos multidimensionales son:

- Drill-down, Drill-up
- Drill-across, Roll-across
- Pivot, Page
- Drill-through



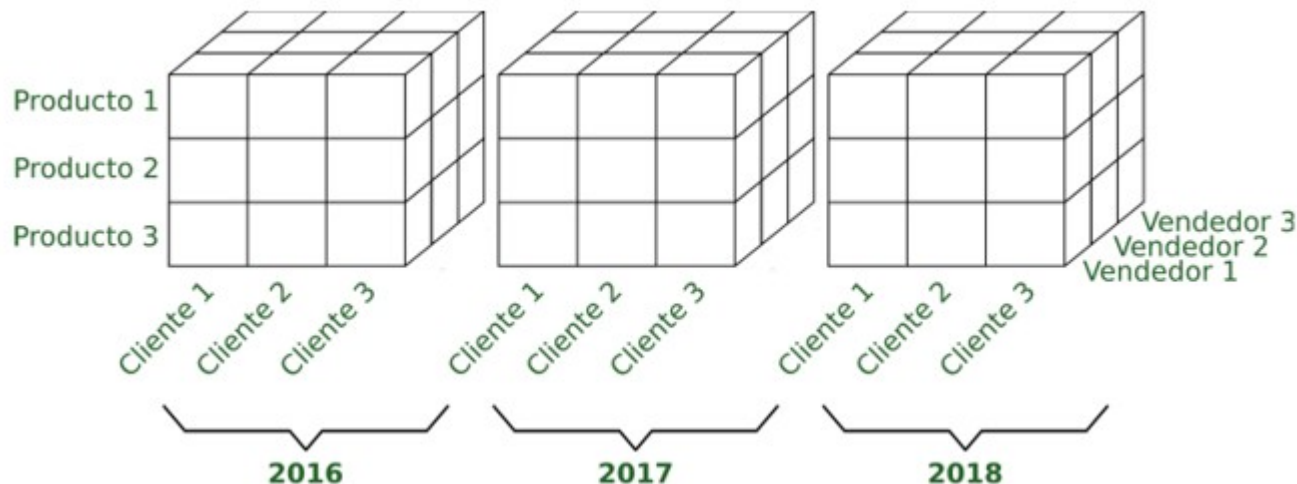
Data Warehousing Business Intelligence



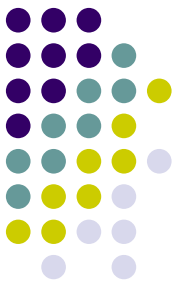
Arquitectura – Query Manager



El modelo de ejemplo consta de cuatro tablas de Dimensiones (**Cientes**, **Productos**, **Vendedores** y **Tiempos**) y una tabla de Hechos (**Ventas**). El cubo dimensional correspondiente será similar a:



Data Warehousing Business Intelligence



Arquitectura – Query Manager

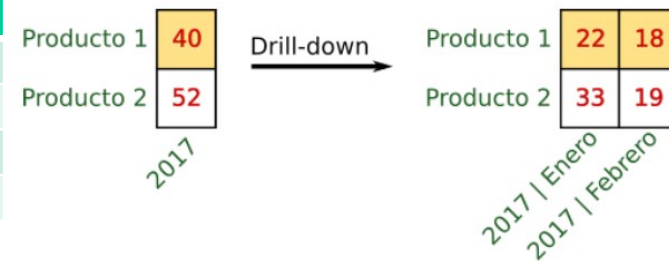


Drill-down: Va de lo general a lo específico para ver los datos con mayor nivel de detalle bajando por una **Jerarquía** definida en el cubo. Si existe la jerarquía de fechas en la dimensión **Tiempos** del tipo año→mes→día las operaciones son:

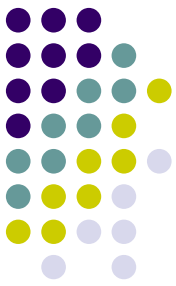
Dimensión Productos	Dimensión Fechas (año)	Hecho Ventas (cantidad)
Producto 1	2017	40
Producto 2	2017	52

Se aplica la operación Drill-down sobre la jerarquía añadiendo un nivel de detalle:

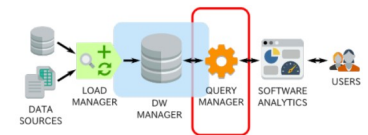
Dimensión Productos	Dimensión Fechas (año a mes)	Hecho Ventas (cantidad)
Producto 1	2017 – Enero	22
Producto 1	2017 – Febrero	18
Producto 2	2017 – Enero	33
Producto 2	2017 – Febrero	19



Data Warehousing Business Intelligence



Arquitectura – Query Manager

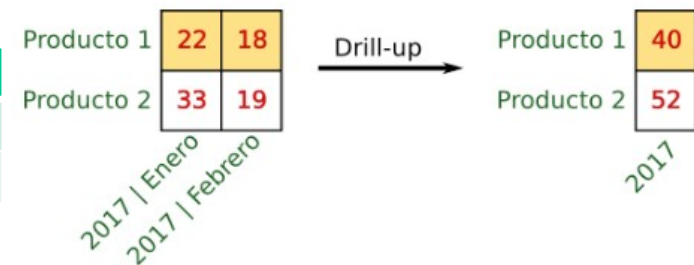


Drill-up: Va de lo específico a lo general para ver los datos con menor grado de detalle subiendo por una **Jerarquía** definida.

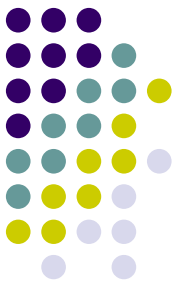
Dimensión Productos	Dimensión Fechas (año à mes)	Hecho Ventas (cantidad)
Producto 1	2017 – Enero	22
Producto 1	2017 – Febrero	18
Producto 2	2017 – Enero	33
Producto 2	2017 – Febrero	19

Se aplica la operación Drill-up sobre la jerarquía quitando un nivel de detalle, se obtiene la situación inicial:

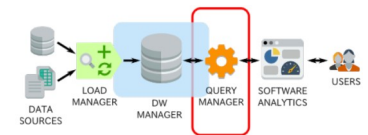
Dimensión Productos	Dimensión Fechas (año)	Hecho Ventas (cantidad)
Producto 1	2017	40
Producto 2	2017	52



Data Warehousing Business Intelligence



Arquitectura – Query Manager

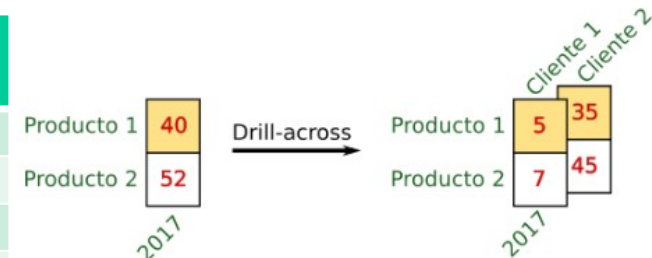


Drill-across: Va de lo general a lo específico para ver los datos con mayor nivel de detalle pero **no se aplica a una Jerarquía** sino que se agrega un atributo como nuevo criterio de análisis.

Dimensión Productos	Dimensión Fechas (año)	Hecho Ventas (cantidad)
Producto 1	2017	40
Producto 2	2017	52

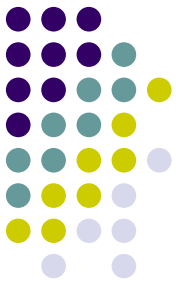
Si se aplica la operación Drill-across al agregar el atributo **Cliente** quedará:

Dimensión Productos	Dimensión Fechas (año)	Dimensión Cliente	Hecho Ventas (cantidad)
Producto 1	2017	Cliente 1	5
Producto 1	2017	Cliente 2	35
Producto 2	2017	Cliente 1	7
Producto 2	2017	Cliente 2	45



Roll-across: es la operación inversa a **Drill-across**.

Data Warehousing Business Intelligence



Arquitectura – Query Manager



Pivot: permite cambiar el orden de visualización de **Atributos** e **Indicadores** para analizar la información desde distintos puntos de vista. Se tiene inicialmente el orden **Producto, Año y Cliente**:

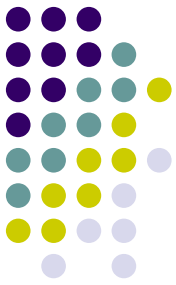
Dimensión Productos	Dimensión Fechas (año)	Dimensión Cliente	Hecho Ventas (cantidad)
Producto 1	2017	Cliente 1	5
Producto 1	2017	Cliente 2	35
Producto 2	2017	Cliente 1	7
Producto 2	2017	Cliente 2	45

Si se aplica la operación Pivot cambiando el orden:

Dimensión Cliente	Dimensión Fechas (año)	Dimensión Productos	Hecho Ventas (cantidad)
Cliente 1	2017	Producto 1	5
Cliente 1	2017	Producto 2	7
Cliente 2	2017	Producto 1	35
Cliente 2	2017	Producto 2	45

No tiene sentido una operación inversa ya que es otro **Pivot**.

Data Warehousing Business Intelligence



Arquitectura – Query Manager



Page: presenta la información dividida en secciones tomando como criterio los valores de un **Atributo**.

Dimensión Productos	Dimensión Fechas (año à mes)	Hecho Ventas (cantidad)
Producto 1	2017 – Enero	22
Producto 1	2017 – Febrero	18
Producto 2	2017 – Enero	33
Producto 2	2017 – Febrero	19

Si se aplica Page sobre el atributo **Producto** la salida será:

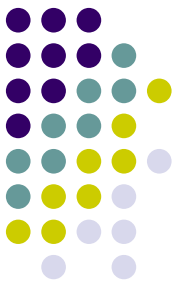
Página 1 – Producto
1

Dimensión Fechas (año à mes)	Hecho Ventas (cantidad)
2017 – Enero	22
2017 – Febrero	18

Página 2 – Producto
2

Dimensión Fechas (año à mes)	Hecho Ventas (cantidad)
2017 – Enero	33
2017 – Febrero	19

Data Warehousing Business Intelligence



Arquitectura – Query Manager



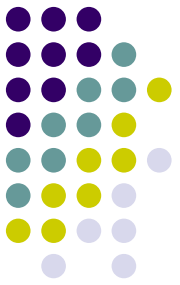
Drill-through: permite visualizar cuáles son los datos relacionados al valor de un Indicador. Los datos se mostrarán en su máximo nivel de detalle. Se tomará como referencia la siguiente tabla:

Dim Productos	Dim Fechas (año)	H Ventas - Indicador (cantidad)
Producto 1	2017	40
Producto 2	2017	52

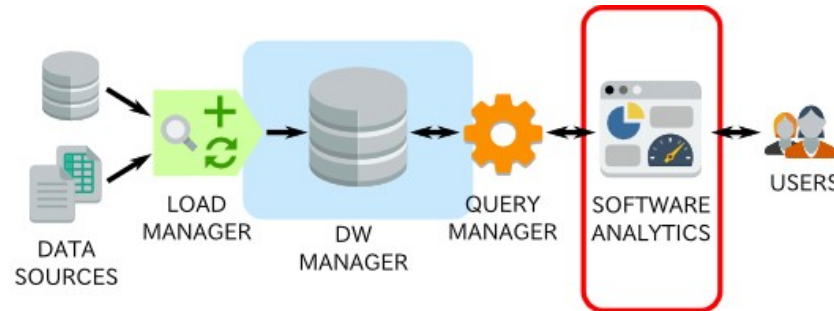
Si se aplica la operación Drill-through sobre el Indicador de la fila seleccionada, se puede obtener el detalle completo:

Dim Fechas (Jerarquía año → mes → día)	Dim Clientes	Dim Productos	Dim Vendedores	H Ventas – Indicador (cantidad)
2017→Enero→1	Cliente 1	Producto 1	Vendedor 1	3
2017→Enero→5	Cliente 1	Producto 1	Vendedor 1	2
2017→Enero→10	Cliente 2	Producto 1	Vendedor 1	10
2017→Enero→18	Cliente 2	Producto 1	Vendedor 1	6
2017→Enero→21	Cliente 2	Producto 1	Vendedor 1	1
2017→Febrero→4	Cliente 2	Producto 1	Vendedor 1	5
2017→Febrero→12	Cliente 2	Producto 1	Vendedor 1	4
2017→Febrero→16	Cliente 2	Producto 1	Vendedor 1	5
2017→Febrero→23	Cliente 2	Producto 1	Vendedor 1	3
2017→Febrero→24	Cliente 2	Producto 1	Vendedor 1	1

Data Warehousing Business Intelligence



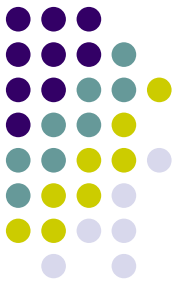
Arquitectura – Software Analytics



Estas herramientas constituyen el nexo entre el DW y los usuarios y son la parte más visible y/o tangible del proceso de Data Warehousing. Para la obtención de datos del DW, se utiliza principalmente:

- La metadata de las estructuras de datos que han sido creadas previamente (cubos multidimensionales, business models) y
- Conexiones a bases de datos

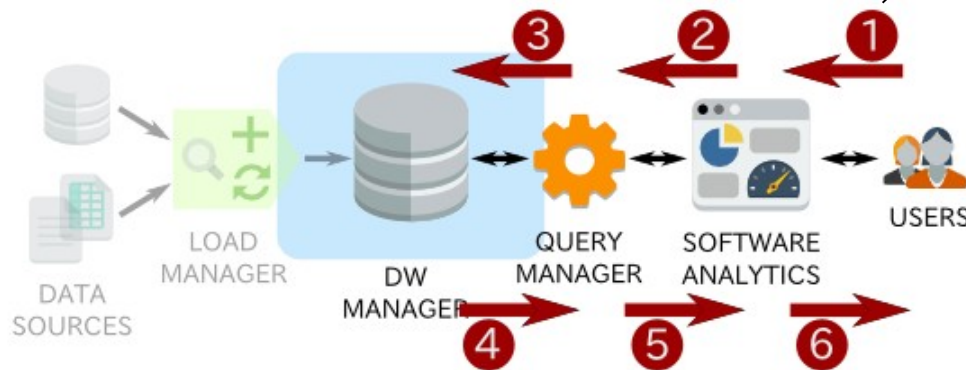
Data Warehousing Business Intelligence



Arquitectura – Software Analytics

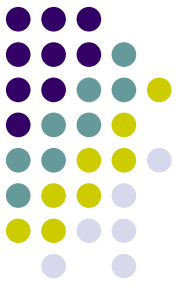


Cada vez que el usuario interactúa con el Software Analytics para explorar los datos del DW se llevan a cabo los siguientes pasos generales:



1. Los usuarios seleccionan o establecen qué datos desean obtener del DW, mediante la GUI (interfaz gráfica) del Software Analytics

Data Warehousing Business Intelligence

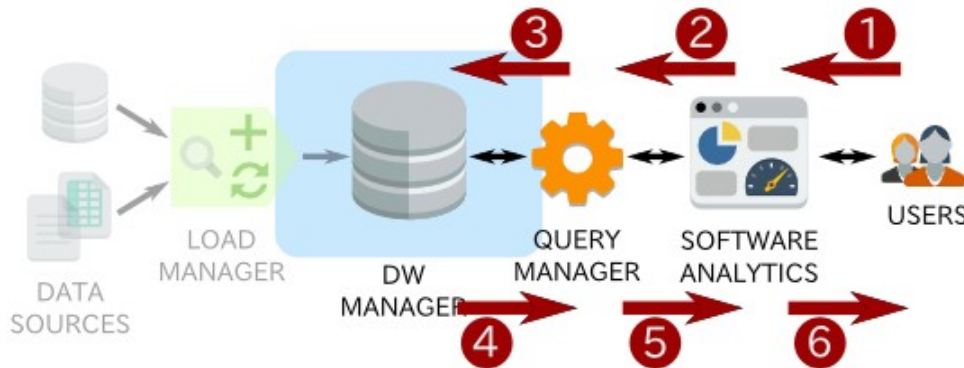


Arquitectura – Software Analytics

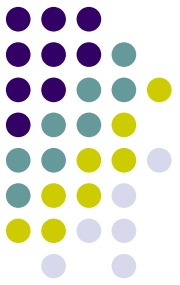


Cada vez que el usuario interactúa con el Software Analytics para explorar los datos del DW se llevan a cabo los siguientes pasos generales:

2. El Software Analytics procesa el pedido de los usuarios, construye las consultas (utilizando la metadata) y las envía al Query Manager.



Data Warehousing Business Intelligence

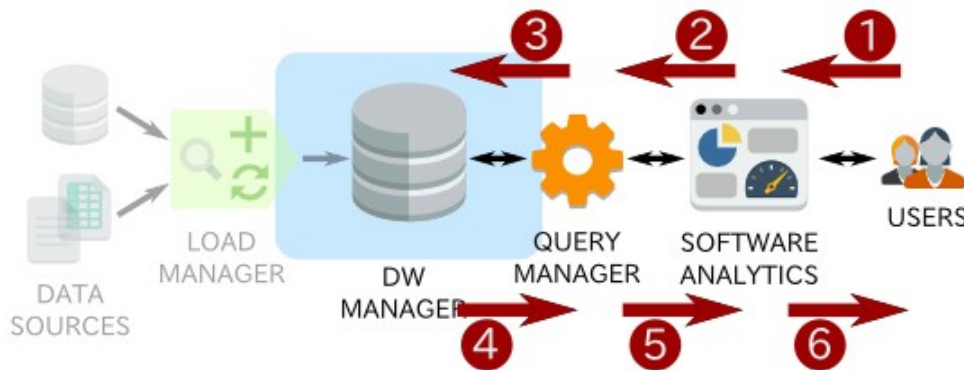


Arquitectura – Software Analytics

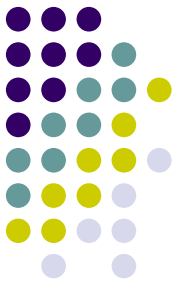


Cada vez que el usuario interactúa con el Software Analytics para explorar los datos del DW se llevan a cabo los siguientes pasos generales:

3. El Query Manager ejecuta las consultas sobre la estructura de datos con la que se esté trabajando (cubo multidimensional, business model, etc.).



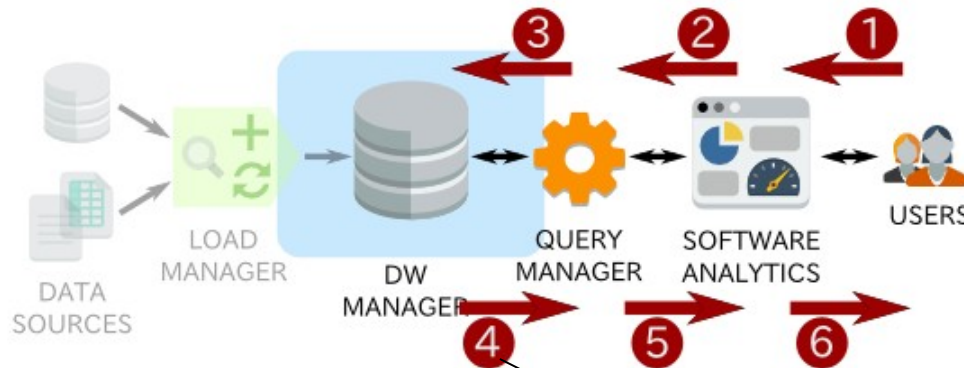
Data Warehousing Business Intelligence



Arquitectura – Software Analytics

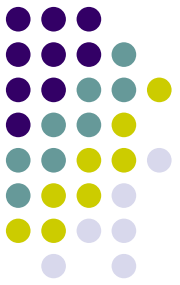


Cada vez que el usuario interactúa con el Software Analytics para explorar los datos del DW se llevan a cabo los siguientes pasos generales:



4. El Query Manager obtiene los resultados de las consultas.

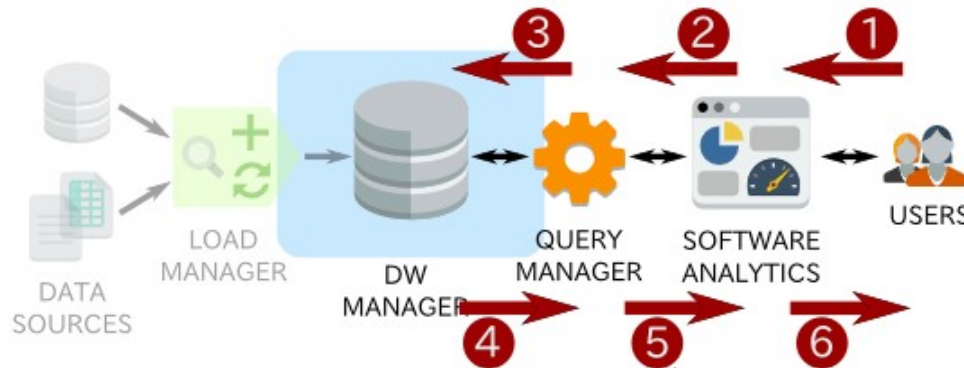
Data Warehousing Business Intelligence



Arquitectura – Software Analytics

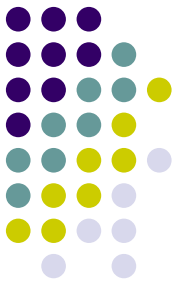


Cada vez que el usuario interactúa con el Software Analytics para explorar los datos del DW se llevan a cabo los siguientes pasos generales:



5. El Query Manager envía los datos al Software Analytics.

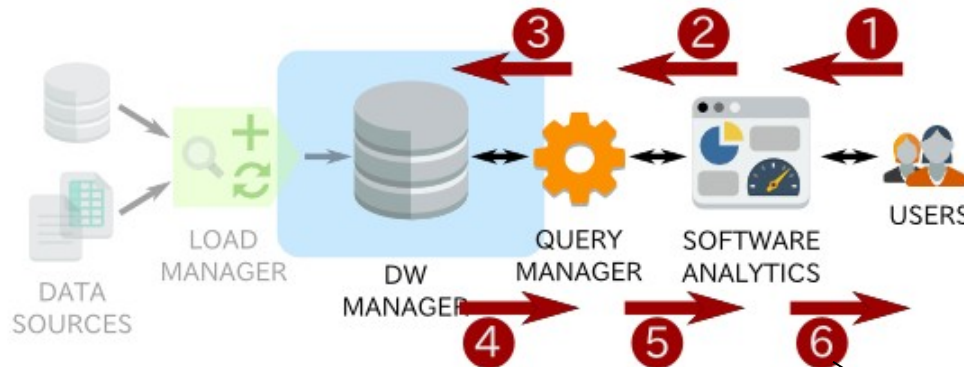
Data Warehousing Business Intelligence



Arquitectura – Software Analytics

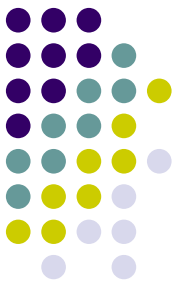


Cada vez que el usuario interactúa con el Software Analytics para explorar los datos del DW se llevan a cabo los siguientes pasos generales:



6. El Software Analytics presenta a los usuarios los datos requeridos.

Data Warehousing Business Intelligence



Arquitectura – Software Analytics



Herramientas de consulta y análisis

Dashboards (reporting en general: el análisis parte de la hipótesis del usuario)

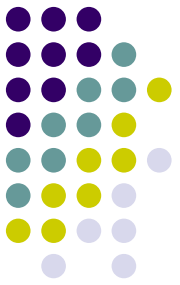
Colección de componentes gráficos de análisis como reportes, tablas, gráficos, consultas y análisis interactivos, etc. que hacen referencia a un tema en particular **y que están relacionados entre sí**. Presentan la información altamente resumida.

- Se componen de consultas, reportes, análisis interactivos, charts (barras, líneas, torta, áreas, etc., semáforos, Indicadores causa-efecto, etc.
- Permiten evaluar situaciones con un solo golpe de vista.
- Poseen un formato de diseño visual muy llamativo.

Data Mining (permite generar las hipótesis del usuario)

Se emplea para analizar factores de influencia, predecir comportamientos futuros, extraer conocimientos ocultos, agrupar o segmentar elementos similares etc. Permite inferir comportamientos y realizar predicciones. Es una técnica para descubrir patrones en grandes volúmenes de datos basados en IA y métodos matemáticos como **Redes neuronales, árboles de decisión** entre otros.

Data Warehousing Business Intelligence



Arquitectura – Users

