

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

- - Modelado lógico multidimensional (MD)
 - Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
 - □ El proceso de diseño de un DW

•INGP. 2019

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

□ Modelado lógico multidimensional (MD)

- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

3

Tema 4. Diseño lógico de almacenes de datos

Modelado multidimensional

- □ El modelado lógico multidimensional depende del tipo de servidor OLAP
 - MOLAP (Multidimensional OLAP)
 - Vectores o matrices multidimensionales
 - ROLAP (Relational OLAP)
 - Esquema estrella de R. Kimball
 - Tablas relacionales para representar hechos y dimensiones
 - Variantes esquema estrella
 - Constelaciones de hechos
 - Copos de nieve

•INGP. 2019

л

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

■ Modelado lógico multidimensional (MD)

- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

5

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

- □ Modelado lógico multidimensional (MD)
- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

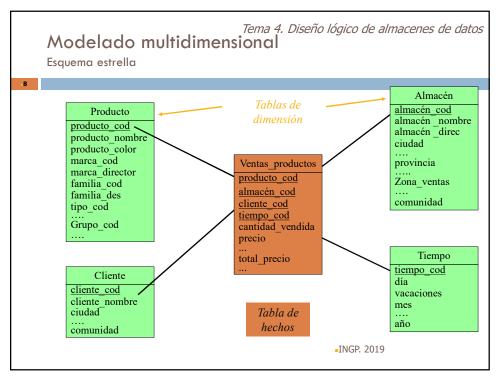
Esquema estrella

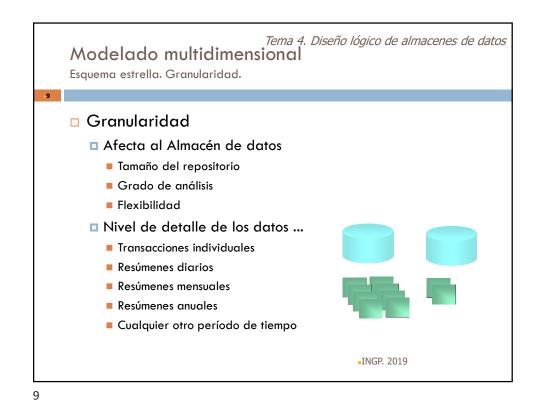
7

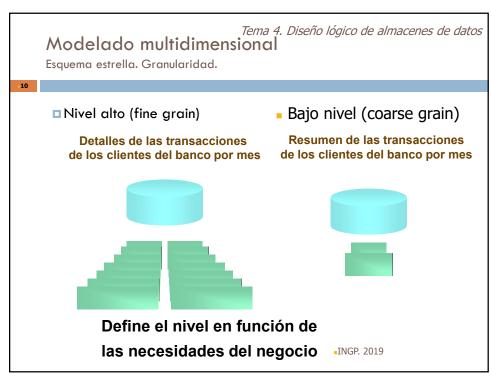
- □ Hechos → tablas de hechos
- □ Dimensiones → tablas de dimensiones
- □ Tabla de hechos es relación muchos a muchos con las tablas de dimensiones
 - Técnica utilizada → desnormalización
 - Clave primaria tabla de hechos compuesta por claves ajenas de tablas de dimensiones

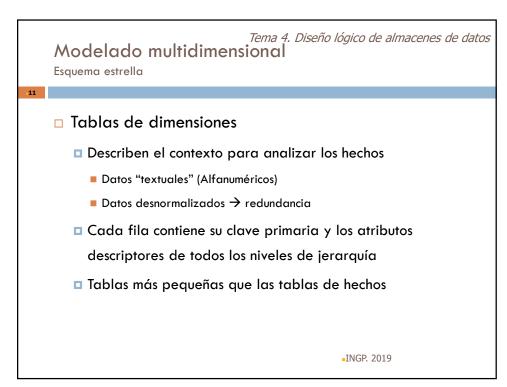
•INGP. 2019

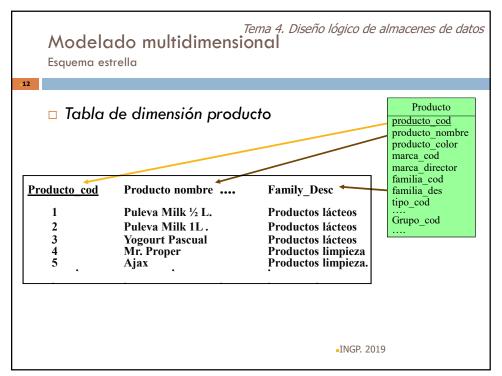
7



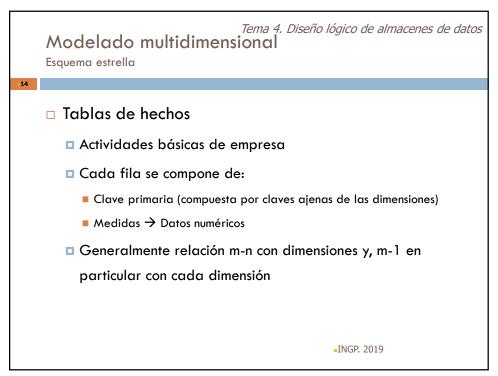


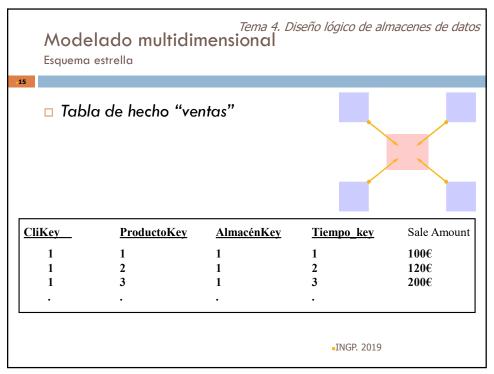


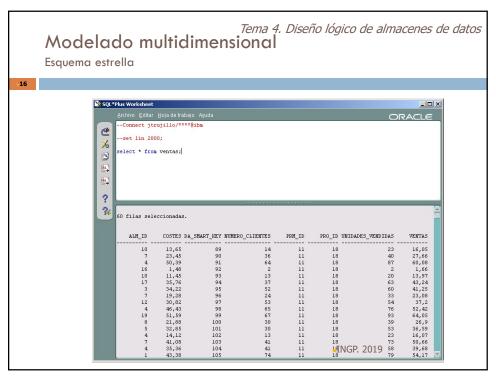


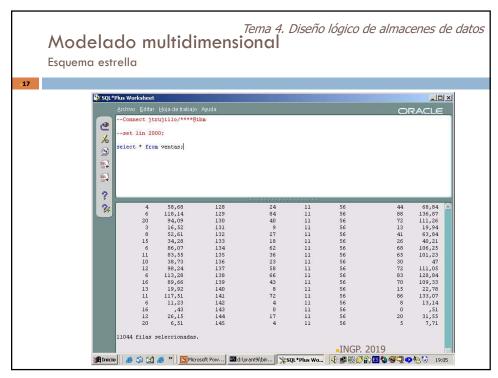


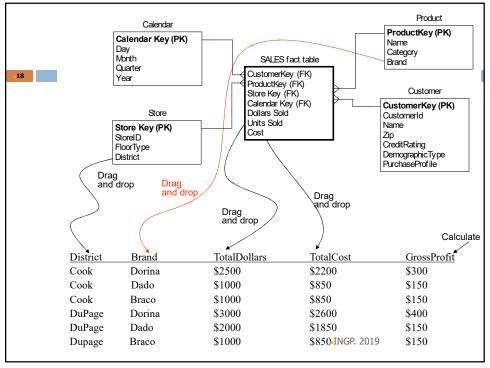


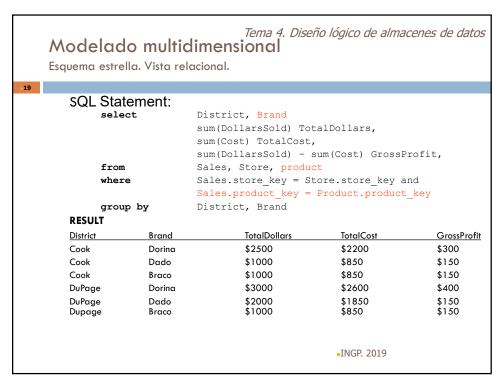


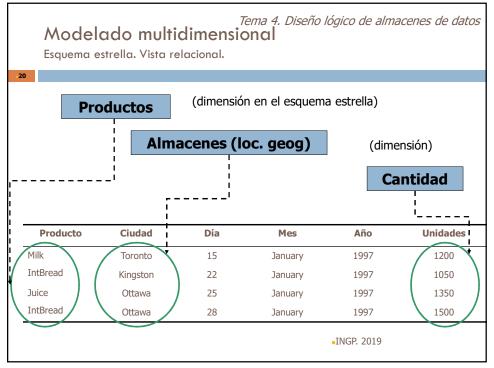


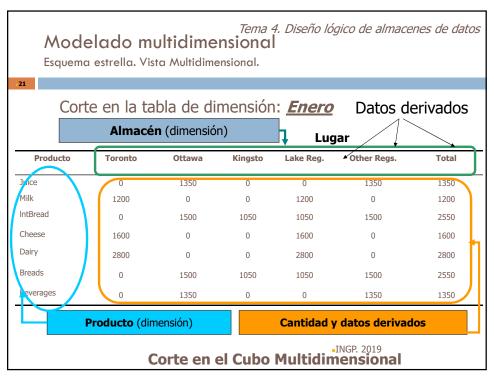


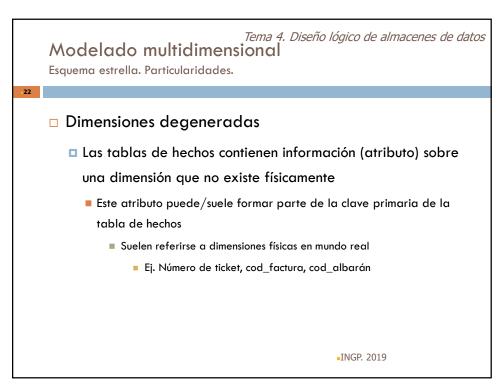


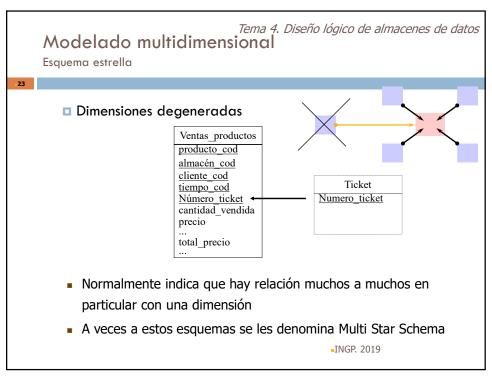


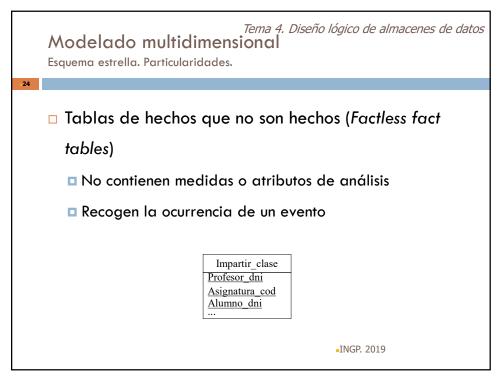


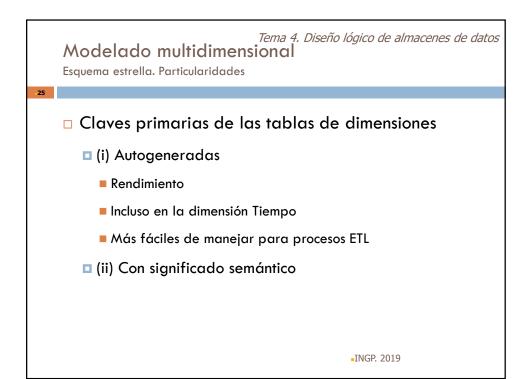


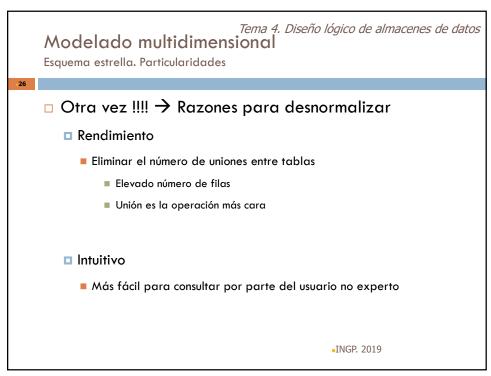












Esquema estrella. Particularidades

27

- □ Dimensiones que cambian lentamente
 - Las instancias de dimensiones se pueden considerar fijas a lo largo del tiempo
 - Sin embargo, a veces deseamos registrar cambios que puedan suceder en algunas dimensiones para seguir analizando datos históricos
 - Normalmente cambios en descripción de productos o de clientes
 - □ Kimball propone tres soluciones básicas y dos híbridas

-INGP. 2019

27

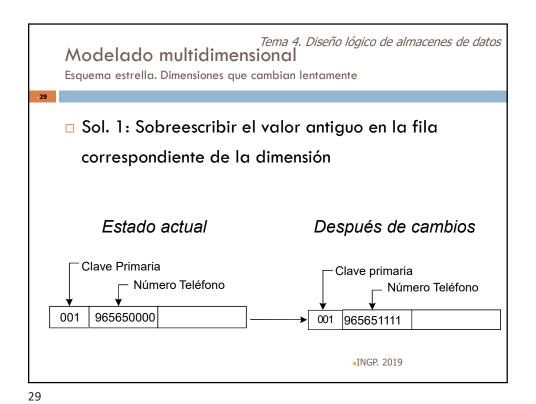
Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema estrella. Dimensiones que cambian lentamente

28

- □ Sol. 1: Sobreescribir el valor antiguo en la fila correspondiente de la dimensión
 - La fila registra el último valor válido
 - Perdemos la habilidad de analizar "la historia"
 - Se suele utilizar cuando el valor antiguo deja de tener significado o cuando se corrigen errores
 - Aproximación simple
 - □ Ejemplos: descripción productos, nombres nuevos, etc.

INGP. 2019



Tema 4. Diseño lógico de almacenes de datos
Modelado multidimensional
Esquema estrella. Dimensiones que cambian lentamente

Sol. 2: Añadir una fila nueva utilizando un nuevo valor
de clave autogenerada

Permite registrar el cambio del valor de un atributo en una
dimensión
Representa un cambio físico real en la dimensión
Imprescindible utilizar claves autogeneradas
Añadir un campo que identifique al valor actual
Añadir un campo que sea un "flag" al valor actual

Esquema estrella. Dimensiones que cambian lentamente

31

- Sol. 2: Añadir una fila nueva utilizando un nuevo valor de clave autogenerada...
 - Se puede utilizar dos atributos añadidos para registrar la fecha de comienzo y fin de validez de cada instancia
 - Inconvenientes
 - Necesita almacenamiento extra para los atributos extras
 - Realizar test para excluir las versiones antiguas
 - □ Ej. Dirección, Región...

•INGP. 2019

31

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema estrella. Dimensiones que cambian lentamente

32

- Sol. 2: Añadir una fila nueva utilizando un nuevo valor de clave autogenerada...
 - Segmenta la historia
 - Particiona la tabla de hechos
 - Dos "productos" distintos con claves generadas distintas estarán registrados en las ventas
 - Queries...
 - Definir restricciones sobre atributo que ha cambiado el valor
 - Si queremos mismo producto, actuaremos sobre clave primaria original

■INGP. 2019

Esquema estrella. Dimensiones que cambian lentamente

33

- Sol. 2: Añadir una fila nueva utilizando un nuevo valor de clave autogenerada...
 - Suele ser la más utilizada para las verdaderamente Dimensiones que cambian lentamente

•INGP. 2019

33

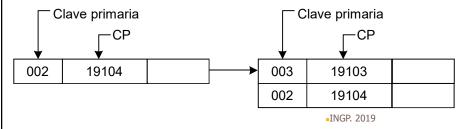
Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema estrella. Dimensiones que cambian lentamente

34

□ Sol. 2: Añadir una fila nueva utilizando un nuevo valor de clave autogenerada...

Estado actual Después de cambios



Esquema estrella. Dimensiones que cambian lentamente

35

- Sol. 3: Añadir atributos para registrar el valor nuevo y el inmediatamente anterior
 - □ Suele utilizarse en cambios mínimos como por ejemplo redefinición de descripción de producto
 - Utilizado cuando deseamos seguir la historia de tanto los valores antiguos como nuevos
 - Se puede / suele añadir un atributo de fecha_efectiva
 - □ Poco recomendada / utilizada

•INGP. 2019

35

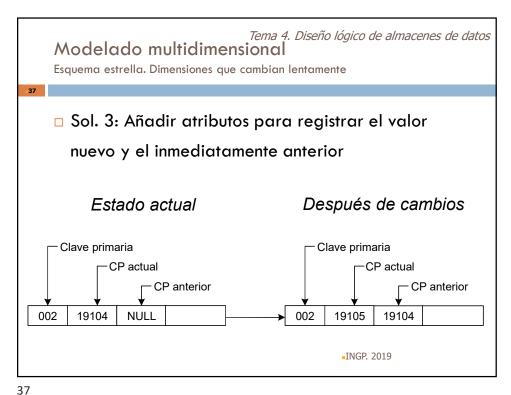
Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema estrella. Dimensiones que cambian lentamente

36

- Sol. 3: Añadir atributos para registrar el valor nuevo y el inmediatamente anterior
 - Inconveniente:
 - Sólo permite analizar los datos de la tabla de hechos o bien por el valor antiguo o bien por el nuevo
 - Ya que se utiliza la misma clave autogenerada

•INGP. 2019



Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema estrella. Dimensiones que cambian lentamente

38

- Sol híbridas son más flexibles aunque proporcionan mayor complejidad de manejo
 - Muy seguro de ser adecuado para el negocio y los posibles requerimientos que se pueden lanzar
- Sol. Híbrida 1: cambios predecibles con versiones múltiples que se superponen
 - Existen cambios predecibles y regulares
 - Usar la aproximación de tipo 3 generalizada

•INGP. 2019

Esquema estrella. Dimensiones que cambian lentamente

39

 Sol. Híbrida 1: cambios predecibles con versiones múltiples que se superponen

Sales Rep Dimension

Sales rep key
Sales rep name
Sales Rep address
Current district
District 2001
District 2000
District 1999

- Conocen los años de análisis de ventas
- Se puede consultar el estado del distrito en cada año

•INGP. 2019

39

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema estrella. Dimensiones que cambian lentamente

40

- □ Sol. Híbrida 2: cambios no predecibles con versiones simples que se superponen
 - Existen cambios irregulares e impredecibles
 - Necesidad de preservar los valores históricos
 - Ninguna de las soluciones estándares anteriores contempla estas dos características

•INGP. 2019

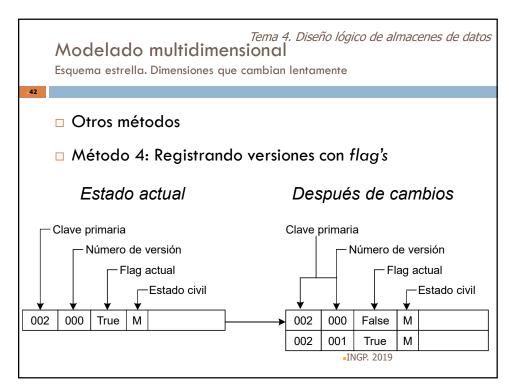
Esquema estrella. Dimensiones que cambian lentamente

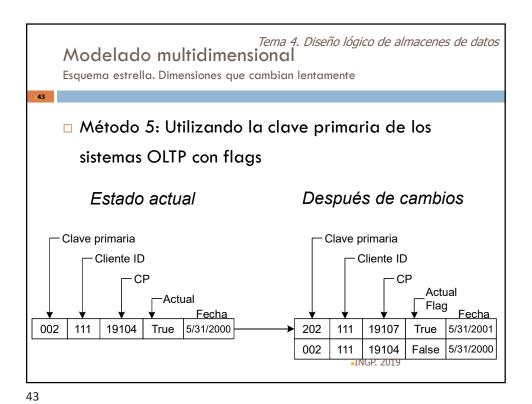
41

- Sol. Híbrida 2: cambios no predecibles con versiones simples que se superponen
 - □ Utilizar los tipo 1 + tipo 2 + tipo 3
 - Añadir una nueva fila para capturar los cambios
 - Añadir una nueva columna para controlar los valores actuales
 - Utilizar la aproximación de tipo 1 para cambios de valores muy recientes donde la historia no interesa
 - Más información: cap. 4 (R. Kimball), pag. 104

•INGP. 2019

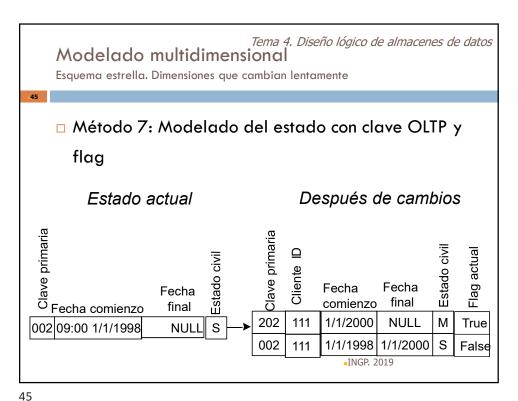
41





Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional Esquema estrella. Dimensiones que cambian lentamente □ Método 6: Modelar eventos con flag's □ Utiliza claves de OLTP. Prácticamente tipo 2 Después de cambios Estado actual Clave primaria Estado civil Clave primaria Estado civil Flag actual Momento del evento Momento del evento 202 15:00 5/31/2001 True 002 13:00 1/1/2000 002 13:00 1/1/2000 False

-INGP. 2019



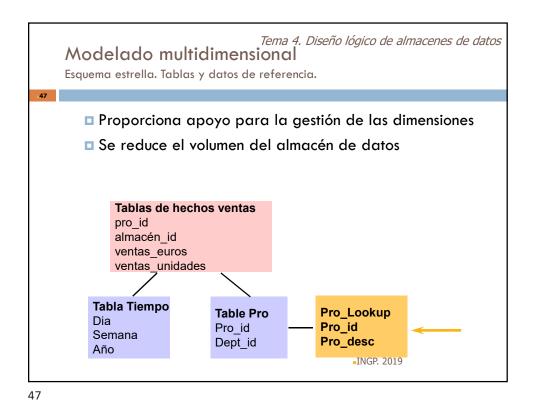
. .

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional Esquema estrella

46

- Qué información podemos saber a partir del esquema estrella
 - □ ¿Cuál es la jerarquía de clasificación en las dimensiones?
 - □ ¿Es la jerarquía entre almacén y zona_Ventas estricta?
 - Hay demasiada información explícita → necesidad de documentar el diseño

•INGP. 2019



Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

Modelado lógico multidimensional (MD)

Modelado multidimensional relacional

Esquema estrella

Esquema constelaciones de hecho

Esquema copos de nieve

Resumen/comparativa final

La dimensión tiempo

Otras consideraciones de diseño

El proceso de diseño de un DW

INGP. 2019

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

49

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

49

Tema 4. Diseño lógico de almacenes de datos

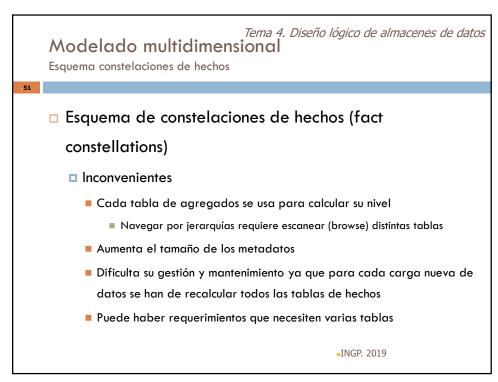
Modelado multidimensional

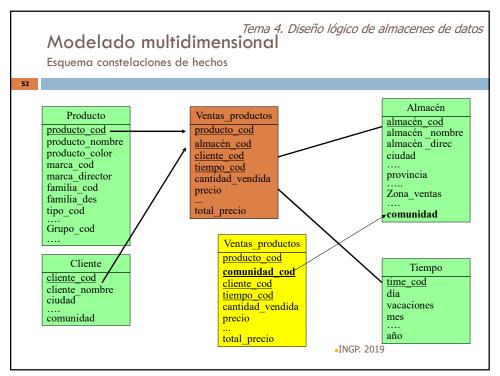
Esquema constelaciones de hechos

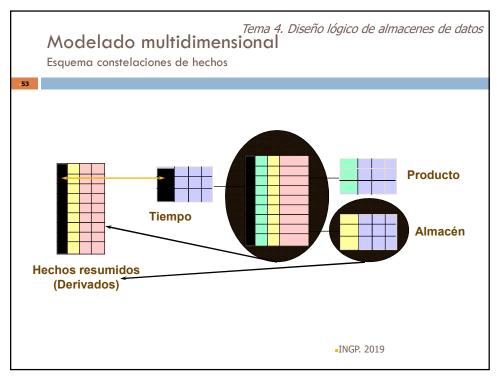
50

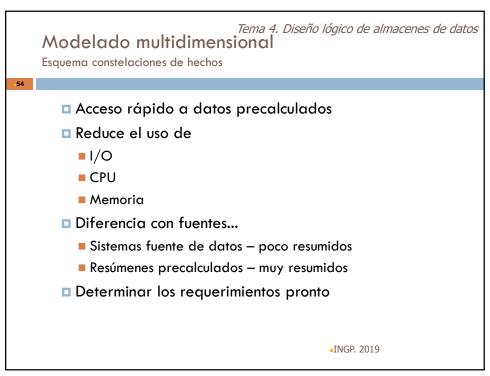
- Esquema de constelaciones de hechos (fact constellations)
 - Además de la tabla de hechos base, se define una tabla de hechos para cada nivel de agregado
 - A veces se les denomina Multi Fact Tables por contener más de una tabla de hechos
 - Acelera la consulta de datos agregados

•INGP. 2019









Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

55

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

55

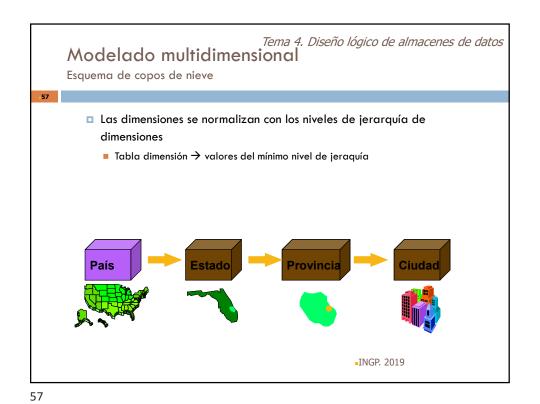
Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

56

- □ Modelado lógico multidimensional (MD)
- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019



Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional Esquema de copos de nieve Almacén Ciudad Comunidad País Cod_almacén Cod ciudad Cod comunidad Cod país descrip_ciudad nombre descrip_comunidad descrip_país superficie habitantes_ciudad habitantes_ciudad habitamtes_ dirección cod_comunidad · cod_país < teléfono cod_ciudad •INGP. 2019

Esquema de copos de nieve

59

Ventajas

- Podría salvar espacio en disco, pero no demasiado
- □ Consultas (browsing) de atributos simples es rápido
- Mejora el rendimiento cuando mayoría de requisitos solicitan niveles de agregación superiores
 - Disminuye el tamaño de tablas a escanear

•INGP. 2019

59

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema de copos de nieve

60

Inconvenientes

- □ Aumenta el número de tablas → aumenta el número de uniones (join)
 - Algunos requisitos pueden demorarse en exceso
- Consultar (browsing) atributos de más de una tabla es más lento
- Demasiado complejo para que los usuarios finales definan sus propias consultas ad-hoc a partir de ellos

•INGP. 2019

Esquema de copos de nieve

61

Inconvenientes

- Requiere una clave primaria más por cada nivel de jerarquía normalizado
- Aumenta la complejidad de diseño y mantenimiento
- No soportado por todas las herramientas del mercado

•INGP. 2019

61

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Esquema de copos de nieve

62

Recomendaciones

- Normalmente no se recomienda
 - Realmente cuando el espacio en disco es un problema
- Normalmente se recomienda normalizar una o dos de las dimensiones más grandes
 - Existen un gran número de filas
- Suele aplicarse cuando muchos atributos caracterizan a los niveles más altos de las jerarquías

•INGP. 2019

Esquema de copos de nieve

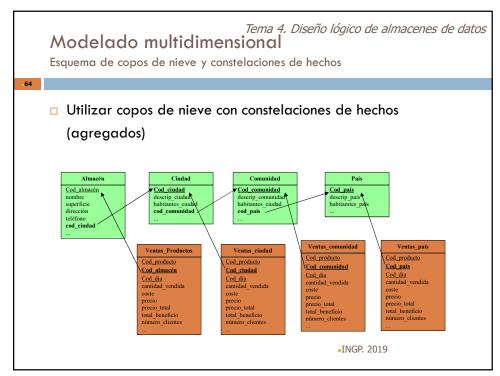
63

Recomendaciones

- Cuando se pueden utilizar índices bitmap para los atributos de granularidad mínima
- Utilizar sólo cuando las ventajas son muy explícitas:
 - Ahorro en disco significativo
 - Muchos atributos en los niveles más altos de jerarquías
- Estadísticamente, el espacio en disco ahorrado utilizando snowflake schemas es del 1% del espacio total en disco

•INGP. 2019

63



Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

65

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

65

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

66

- □ Modelado lógico multidimensional (MD)
- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

Resumen/comparativa final del esquema estrella y sus variantes

67

- Esquema estrella
 - Ventajas
 - Fácil de entender por los usuarios
 - Reduce número de uniones físicas
 - Respuestas rápidas para la mayoría de las consultas
 - Fácil de entender por usuarios finales
 - Metadatos sencillos
 - Soportado por la inmensa mayoría de aplicaciones

•INGP. 2019

67

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Resumen/comparativa final del esquema estrella y sus variantes

68

- Esquema estrella
 - Inconvenientes
 - El aumento del tamaño de la tabla de hechos con datos agregados puede empeorar el rendimiento general
 - Por ello se recomienda tablas de hechos agregados al margen
 - Las dimensiones tienen un tamaño enorme
 - Alrededor de 50 atributos (Kimball)
 - La necesidad de manejar el atributo de nivel para datos agregados en algunos sistemas
 - Es poco robusto o susceptible a cambios
 - Más lento de construir

•INGP. 2019

Resumen/comparativa final del esquema estrella y sus variantes

69

- □ Esquema constelaciones de hechos
 - Ventajas
 - No se necesita el atributo de nivel en las tablas de dimensiones que lo utilizan explícitamente
 - Los datos agregados no se almacenan con los del nivel de detalle más bajo en las tablas de hechos
 - Aumenta el rendimiento al hacer browse de datos almacenados precalculados (sumados)

•INGP. 2019

69

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Resumen/comparativa final del esquema estrella y sus variantes

70

- □ Esquema constelaciones de hechos
 - Inconvenientes
 - Un gran número de tablas de agregados
 - Complica el mantenimiento de los metadatos
 - Necesidad del usuario de conocer la existencia de dichas tablas
 - Algunos requisitos pueden necesitar consultar datos de varias tablas y mermar el rendimiento general del sistema

•INGP. 2019

Resumen/comparativa final del esquema estrella y sus variantes

71

- □ Esquema copos de nieve
 - Ventajas
 - No se necesita el atributo de nivel en los sistemas que lo gestionan explícitamente
 - Más flexible y adecuado para requerimientos
 - Carga de datos (ETL) más rápida y sencilla

•INGP. 2019

71

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Resumen/comparativa final del esquema estrella y sus variantes

72

- □ Esquema copos de nieve
 - Ventajas
 - Directamente implementados por algunas jerarquías
 - No Todas
 - Mejora considerablemente el rendimiento cuando un gran número de requisitos solicita datos agregados o de niveles superiores de jerarquías
 - Los requerimientos escanean un reducido número de filas

•INGP. 2019

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Resumen/comparativa final del esquema estrella y sus variantes

73

- □ Esquema copos de nieve
 - Inconvenientes
 - Aumenta la complejidad de mantener los metadatos
 - Debido al aumento del número de tablas
 - Si no se dispone de la suficiente cantidad de tablas de agregados, el rendimiento general podría disminuir

•INGP. 2019

73

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

74

- □ Modelado lógico multidimensional (MD)
- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

Diseño lógico de DW

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

75

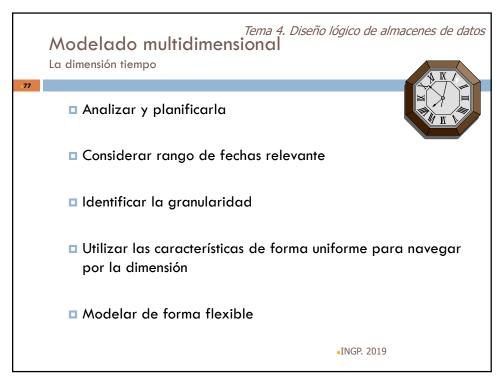
Tema 4. Diseño lógico de almacenes de datos

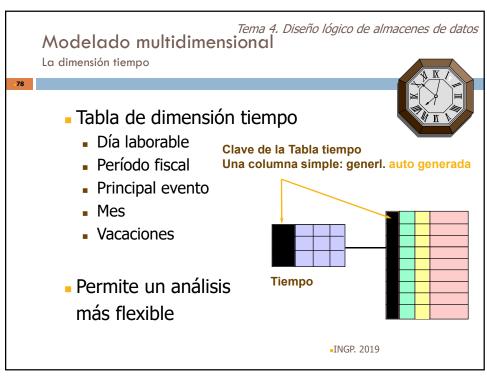
Modelado multidimensional

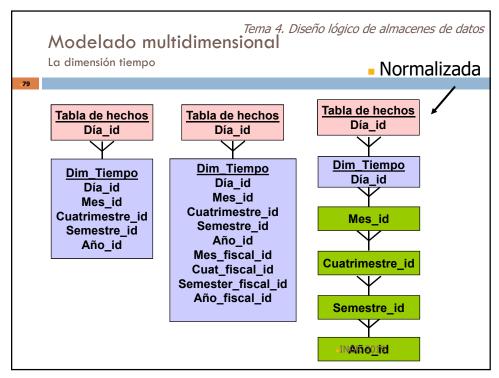
La dimensión tiempo

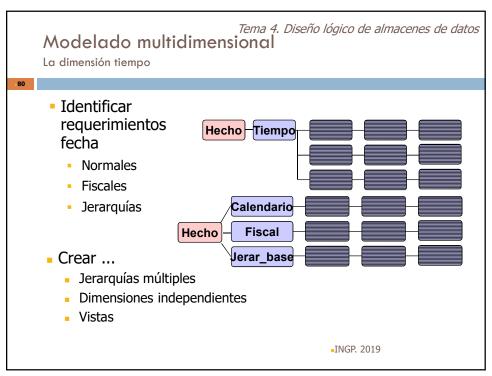
- La dimensión tiempo
 - Es obligatoria
 - Es única
 - Es potente
 - Necesita ser diseñada cuidadosamente
- Incluye períodos del tiempo de negocio
- Incluye fechas especiales

-INGP. 2019









Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional La dimensión tiempo □ Tipo de consultas sobre dimensión tiempo Absolutas Ej. Ventas desde julio 1999 hasta julio 2000

- Relativas Ej. Obtener ventas para el mes actual del año pasado



•INGP. 2019

81

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

La dimensión tiempo

- □ Algunos sistemas de forma interna
 - Añaden un atributo para indicar período actual y el previo
 - Suele hacerse en el mínimo nivel de granularidad
 - Aunque hay que definir muy bien los periodos de comparación
 - Proporcionar códigos de tiempo secuencial que identifiquen a los elementos de las jerarquías

-INGP. 2019

Modelado multidimensional

La dimensión tiempo

Cod		Cod		Cod		Cod		Periodo	100
Dia Dia	Día	Mes	Mes	Semestre	Semestre	Año	Año	actual	Agg level
1300	2	500	Sept	200	S3	50	1999		1
1301	3	500	Sept	200	S3	50	1999	P	1
1420	1	504	Enero	202	S1	51	2000		1
1421	2	504	Enero	202	S1	51	2000		1
1665	2	512	Sept	204	S3	51	2000		1
1666	3	512	Sept	204	S3	51	2000	s	1

□ Tabla de dimensión tiempo en Informik/Wetacube

83

Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

-INGP. 2019

Diseño lógico de DW

85

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

85

Tema 4. Diseño lógico de almacenes de datos

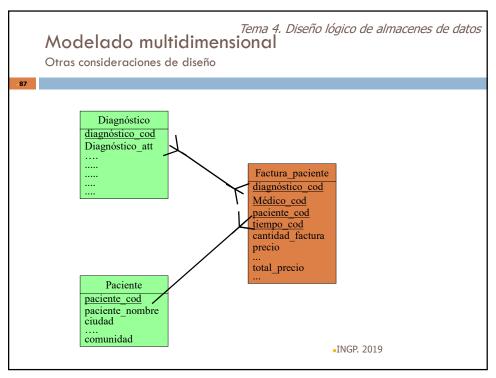
Modelado multidimensional

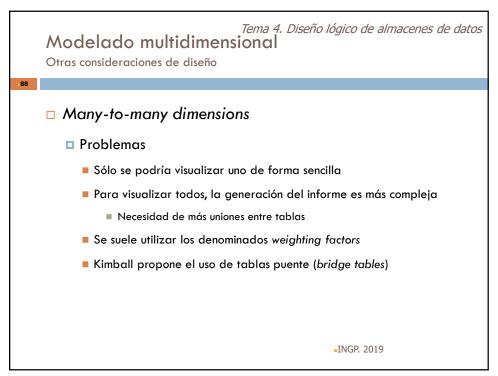
Otras consideraciones de diseño

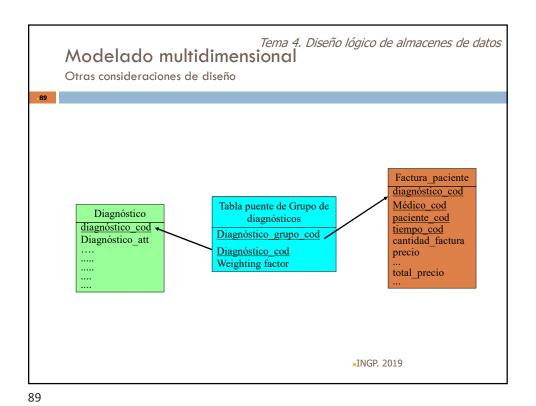
86

- □ Many-to-many dimensions
 - Cuando entre la tabla de hechos y una tabla de dimensión existe una relación de muchos a muchos (m-n)
 - □ Ejemplo: Health Care System
 - Un paciente puede tener varios diagnósticos
 - ¿Qué hacer cuando un paciente tiene varios diagnósticos?

•INGP. 2019







Tema 4. Diseño lógico de almacenes de datos
Modelado multidimensional
Otras consideraciones de diseño

The Data Warehouse Bus Arquitecture

Arquitectura detallada de cómo construir todo el DW corporativo

Cada Data Mart debería estar basado en la mínima granularidad del proceso de negocio que contempla

Dos aproximaciones:

Crear una arquitectura que defina todo el DW

Supervisar la construcción de cada Data Mart en particular

Utilizar las dimensiones y hechos comunes (conformed)

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Otras consideraciones de diseño

91

- □ Dimensiones comunes (Conformed dimensions)
 - Es una misma dimensión que existe en todos los DM
 - Ejplo. Tabla maestro de clientes o productos
 - Tablas que se mantienen de forma independiente a los DM
 - Cuando se definen dimensiones comunes
 - No obviarlas ni menospreciarlas
 - Definir el mínimo nivel de granularidad posible
 - Utilizar claves auto generadas (surrogate keys)
 - Este tipo de dimensiones puede ocupar el 80% del esfuerzo del desarrollo total

•INGP. 2019

91

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

Otras consideraciones de diseño

92

- □ Dimensiones comunes (Conformed dimensions)
 - Ventajas
 - Una misma dimensión se puede utilizar contra varios hechos
 - Interfaces de usuario y datos son consistentes
 - Permite navegar entre Data Marts (DM)
- □ Hechos comunes (Conformed facts)
 - Medidas utilizadas en más de un DM
 - Ejemplos: beneficio, coste, precio etc.

•INGP. 2019

Diseño lógico de DW

93

- □ Modelado lógico multidimensional (MD)
- □ Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019

93

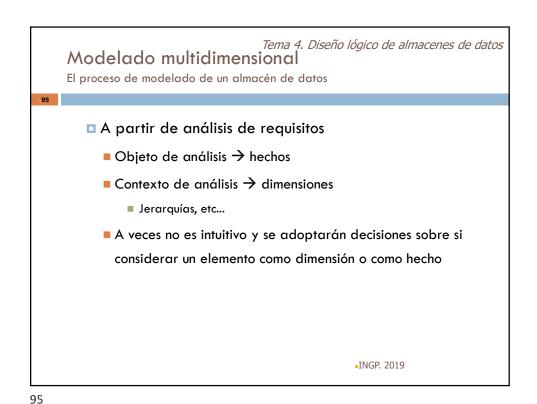
Tema 4. Diseño lógico de almacenes de datos

Diseño lógico de DW

94

- □ Modelado lógico multidimensional (MD)
- Modelado multidimensional relacional
 - Esquema estrella
 - Esquema constelaciones de hecho
 - Esquema copos de nieve
 - Resumen/comparativa final
 - La dimensión tiempo
 - Otras consideraciones de diseño
- □ El proceso de diseño de un DW

•INGP. 2019



Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional El proceso de modelado de un DW: ¿¿¿ Hechos o dimensiones ???? Unidades vendidas - Llamadas en un mes - Color Guardarlo en la tabla de Almacenar en la tabla dimensión si el atributo se de hechos si el percibe como un valor atributo varía estático o constante: continuamente Descripción Balance Localización Unidades vendidas Color Llamadas por mes Tamaño Coste producto -INGP. 2019

Tema 4. Diseño lógico de almacenes de datos Modelado multidimensional

El proceso de modelado de un DW

97

- □ żżż Diseño conceptual ???
 - Algunos diseñadores lo prefieren por la documentación
- Diseño lógico a partir del modelo corporativo/transac.
 - Extraer de la documentación o modelo → esquema estrella
 - Tablas de hechos
 - Tablas de dimensiones
 - Jerarquías de clasificación
 - Atributos
 - En función de los requisitos más comunes decidir
 - Normalizar algunas dimensiones → copos de nieve
 - Estudiar qué agregados definir → constelaciones de hechos

•INGP. 2019

97

Tema 4. Diseño lógico de almacenes de datos

Modelado multidimensional

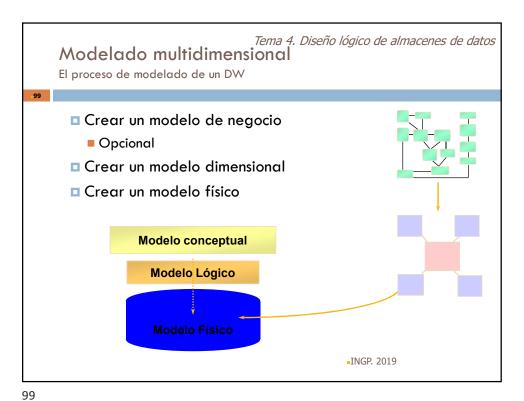
El proceso de modelado de un DW

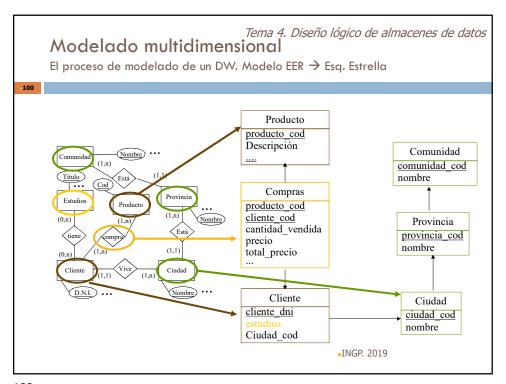
98

- Determinar los requerimientos de usuario
- Asistir a los usuarios en la comprensión de la tecnología
- Casi necesario...
 - Definir requerimientos analíticos
 - Objetivos de rendimiento
- Mantenerse informado con los cambios de requerimientos
- Proporcionar soporte para más requerimientos que los detectados inicialmente



•INGP. 2019





Bibliografía

101

- □ Giovinnazo (2000). Object-Oriented Data Warehouse Design: Building a star schema
- □ Inmon (2002). Building the Data Warehouse (3° ed.)
- □ Kimball (2002). The Data Warehouse Toolkit (3° ed.)
- ☐ Thomsen (2000). OLAP solutions: Building Multidimensional Information Systems

•INGP. 2019

101

