MBD

Máster en Big Data 2023-2024

Tecnologías Business Intelligence



Técnica para modelar bases de datos simples y comprensibles para el usuario final.

Ítems a definir

- Hechos: Aquello que queremos estudiar (Cliente, venta, producto) ¿ Cuánto?
- Dimensiones: Las medidas del hecho (geografía, tiempo, cliente...) ¿ Donde?
 ¿Cuándo? ¿ Que?
- Gránulo: Grado de atomicidad del hecho en función del nivel de la dimensión.

Los miembros de una dimensión se agrupan de forma jerárquica. (dimensión geográfica: ciudad, provincia, autonomía, país...)



Kimball, R., & Ross, M. (2013). *The data warehouse toolkit* : the definitive guide to dimensional modeling (3rd ed.). John Wiley & Sons, Inc.

Caso de la cadena de supermercados



Cuatro pasos del diseño dimensional:

- 1. Seleccionar el proceso de negocio a trabajar
- 2. Definir la granularidad del proceso de negocio: Cuanto
- 3. Escoger las dimensiones: **Donde**, **Cuando**, **Quien**...
- 4. ¿Qué medimos? ¿Qué llena la tabla de hechos?



Descripción del negocio:

Cadena de supermercados.

- 100 tiendas repartidas por 5 estados
- SKU Stock Keeping Units. 60.000 productos
- UPC Universal Process Codes 55.000
- Negocio en torno a la gestión pedidos, stock, venta
- POS Point of Sale
- Interesa saber cómo afectan las políticas de precios y promociones:
- Rebajas temporales, anuncios en los periódicos, displays en los puntos de venta, cupones descuento ...

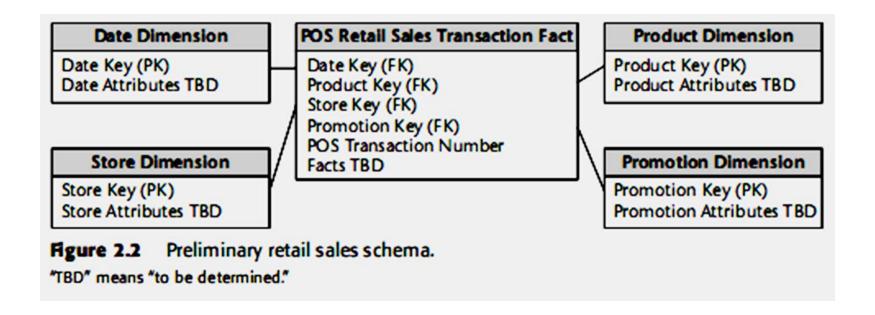


Cuatro pasos del diseño dimensional

- 1. Seleccionar el proceso de negocio a trabajar
 - Entender el proceso de compra entendido como el registro de los POS
- 2. Definir la granularidad del proceso de negocio: Cuanto
 - Grano: individual, ítem capturado en el POS
- 3. Escoger las dimensiones: Donde, Cuando...
 - Fecha (Temporal)
 - Producto (¿Qué?)
 - Tienda (Geográfica)
 - Promociones (¿Cómo?)
- 4. ¿Qué medimos? Que llena la tabla de hechos.

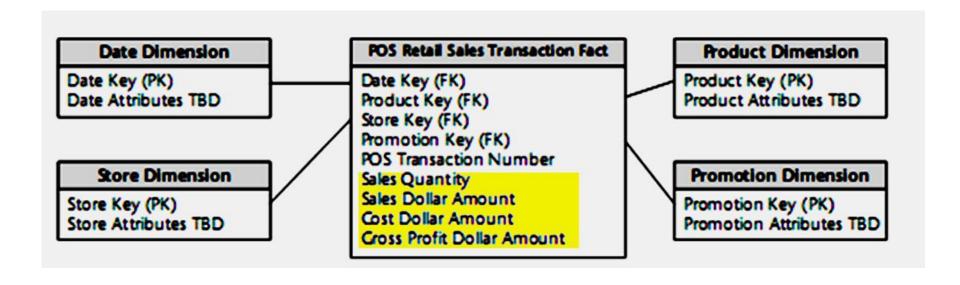


Primera aproximación al modelo de datos



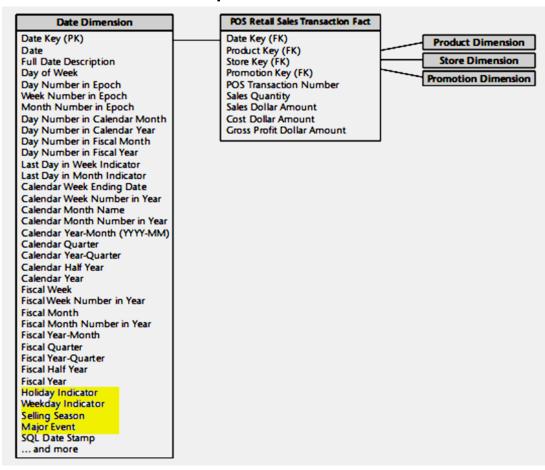


Ampliación tabla de hechos





Dimensión temporal



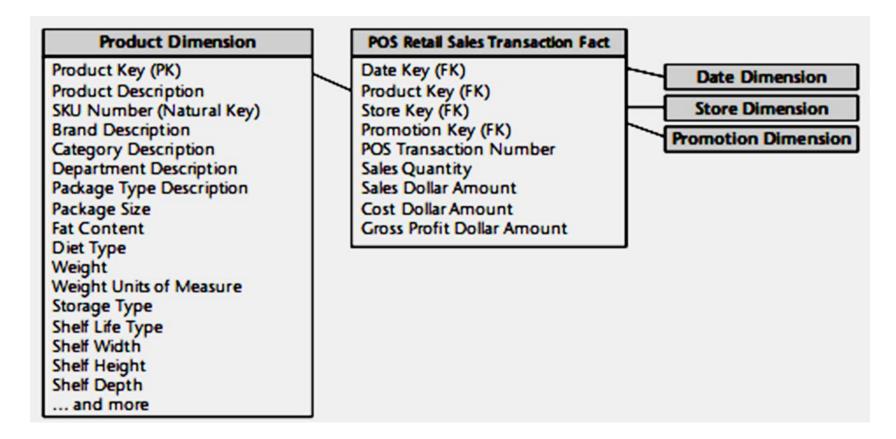


Detalle dimensión temporal

Date Key	Date	Full Date Description	Day of Week	Calendar Month	Calendar Year	Rscal Year- Month	Holiday Indicator	Weekday Indicator
1	01/01/2002	January 1, 2002	Tuesday	January	2002	F2002-01	Holiday	Weekday
2	01/02/2002	January 2, 2002	Wednesday	January	2002	F2002-01	Non-Holiday	Weekday
3	01/03/2002	January 3, 2002	Thursday	January	2002	F2002-01	Non-Holiday	Weekday
4	01/04/2002	January 4, 2002	Friday	January	2002	F2002-01	Non-Holiday	Weekday
5	01/05/2002	January 5, 2002	Saturday	January	2002	F2002-01	Non-Holiday	Weekend
6	01/06/2002	January 6, 2002	Sunday	January	2002	F2002-01	Non-Holiday	Weekend
7	01/07/2002	January 7, 2002	Monday	January	2002	F2002-01	Non-Holiday	Weekday
8	01/08/2002	January 8, 2002	Tuesday	January	2002	F2002-01	Non-Holiday	Weekday



Dimensión producto



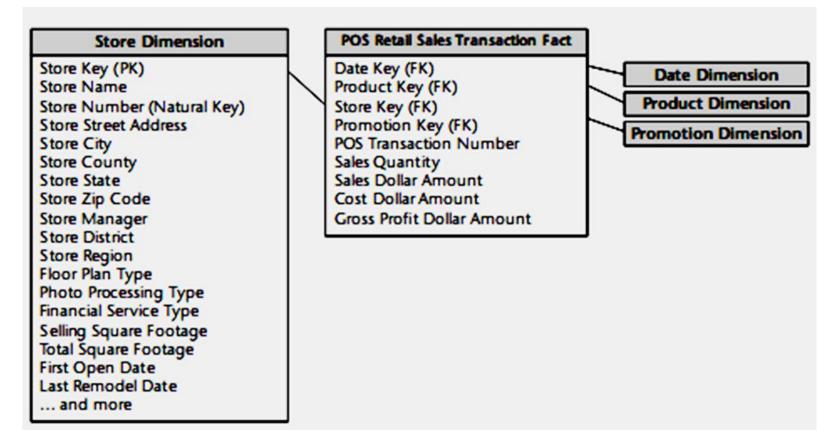


Ejemplo dimensión producto

Product Key	Product Description	Brand Description	Category Description	Department Description	Fat Content
1	Baked Well Light Sourdough Fresh Bread	Baked Well	Bread	Bakery	Reduced Fat
2	Fluffy Sliced Whole Wheat	Fluffy	Bread	Bakery	Regular Fat
3	Fluffy Light Sliced Whole Wheat	Fluffy	Bread	Bakery	Reduced Fat
4	Fat Free Mini Cinnamon Rolls	Light	Sweeten Bread	Bakery	Non-Fat
5	Diet Lovers Vanilla 2 Gallon	Coldpack	Frozen Desserts	Frozen Foods	Non-Fat
6	Light and Creamy Butter Pecan 1 Pint	Freshlike	Frozen Desserts	Frozen Foods	Reduced Fat
7	Chocolate Lovers 1/2 Gallon	Frigid	Frozen Desserts	Frozen Foods	Regular Fat
8	Strawberry Ice Creamy 1 Pint	lcy	Frozen Desserts	Frozen Foods	Regular Fat
9	Icy Ice Cream Sandwiches	lcy	Frozen Desserts	Frozen Foods	Regular Fat

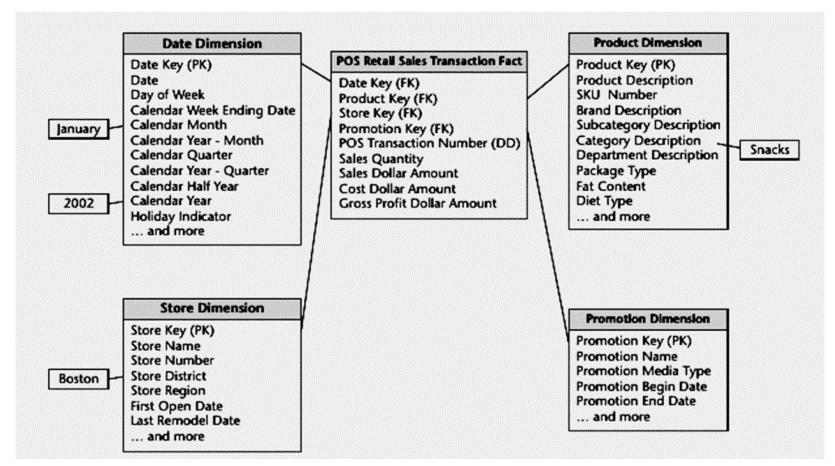


Dimensión geográfica



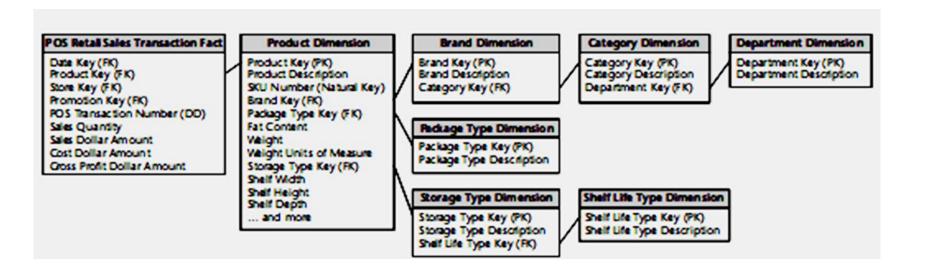


Ejemplo de busqueda



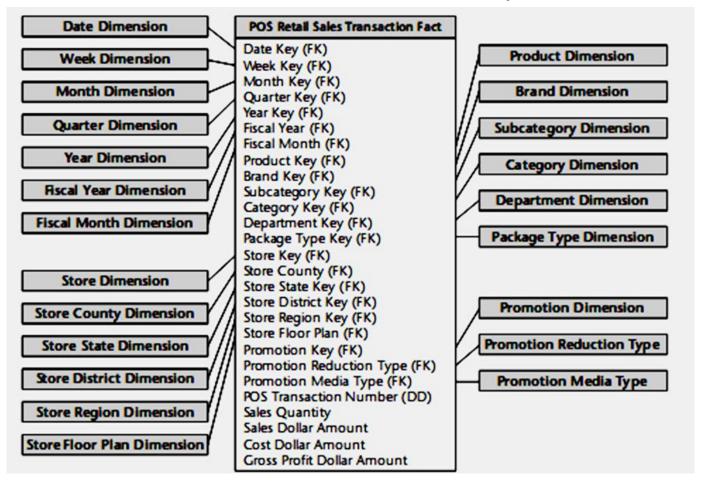


Del modelo de estrella al modelo de copo de nieve





Del modelo de estrella al modelo ciempies





MBD

Máster en Big Data

Tecnologías Business Intelligence

