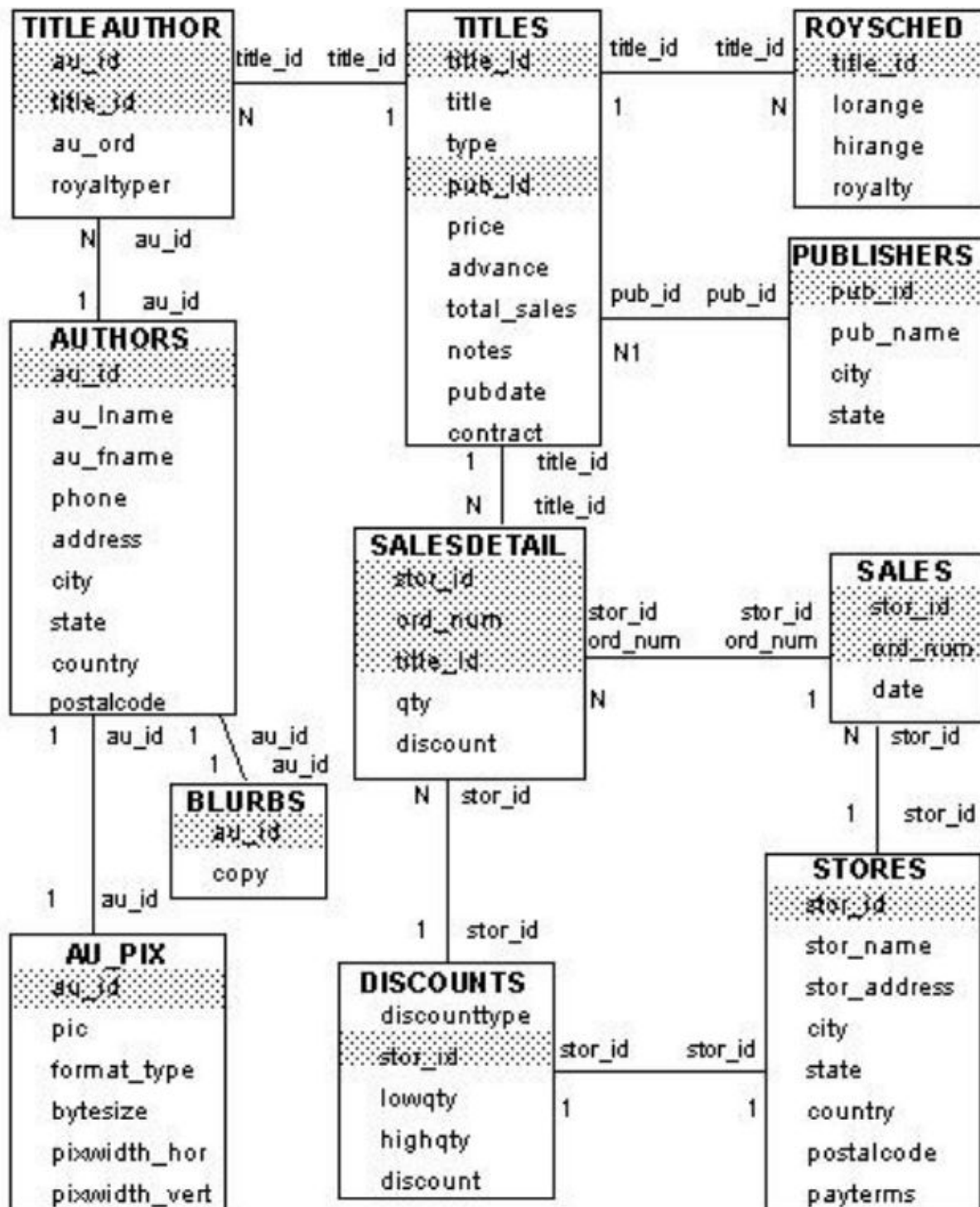


Una distribuidora de libros, al por mayor, tiene una base de datos como la de la figura. En ella se encuentra información de títulos, editoriales, autores, librerías, ventas, etc.



La descripción de la BD se encuentra en:

<http://infocenter.sybase.com/help/index.jsp?topic=/com.sybase.infocenter.dc32300.1570/html/sqlug/sqlug886.htm>

Siguiendo la metodología de Kimball diseña un almacén de datos. Recuerda que el diseño deben considerarse únicamente los datos con que se cuenta. En este caso, asume que sólo tienes los datos especificados en la BD que registra órdenes de compra de librerías a su proveedor, y aunque puede haber datos que no se requieran en el almacén. Si crees que a la BD le falta algún dato hazmelo saber, porque de otra manera no pueden modificar la BD.

En el diseño, especifica las llaves sustitutas(PK), las naturales (NK), las externas (FK), si tienen alguna dimensión degenerada (DD), en fin todo lo necesario para que sea muy claro tu diseño. Debes entregarme en la fecha especificada, un documento resultado de cada paso del diseño, en el que justifiques las decisiones de diseño que tomadas, así como el esquema resultante del diseño. Además especifica 5 consultas que se puedan resolver usando tu diseño y las operaciones (roll-up, drill-down, etc.) que deban usarse para resolverlas.

Diseño dimensional mediante la metodología de Kimball

1. Seleccionar el proceso a modelar.

La actividad desarrollada por la organización en la que nos interesa enfocarnos es la de **ventas**. Ya que nos proporcionan todos los datos necesarios para poder analizar este campo.

2. Definir la granularidad del proceso.

Total de ventas diarias por sucursal por día por libro(producto). Ya que nos permite diferenciar ventas entre semana o fines de semana, en diferentes editoriales, los libros en promoción que se vendieron, etc.

3. Elegir las dimensiones.

- a. **Fecha:** La representación de la fecha puede ser de diferentes maneras ya que los que toman las decisiones no saben el lenguaje de SQL así como el tipo de dato DATE.

- i. **id_fecha(PK):** identificador de la fecha, este es la llave primaria.

- 1. Tipo de dato: Integer (numérico no aritmético).

- 2. Dominio: {1, 2, 3, ...}

- ii. **fecha:** fecha en formato de lenguaje sql.

- 1. Tipo de dato: Date.

- 2. Dominio: Fecha en formato sql.

- iii. **dia_semana.**

- 1. Tipo de dato: Cadena.

- 2. Dominio: {lunes, martes, ..., domingo}.

- iv. **numero_día_mes_calendario.**

- 1. Tipo de dato: Integer.

- 2. Dominio: {1, ..., 31}

- v. **numero_día_anio_calendario.**

Tarea 3: Diseño de un almacén de datos.

14 de Marzo de 2018

1. Tipo de dato: Integer.
2. Dominio: {1, 2, ..., 366}
- vi. nombre_mes_calendario.
 1. Tipo de dato: Cadena.
 2. Dominio: {Enero, Febrero, ..., Diciembre}.
- vii. numero_mes_anio.
 1. Tipo de dato: Integer.
 2. Dominio: {1, ..., 12}
- viii. fecha_anio-mes_(YYYY-MM).
 1. Tipo de dato: Cadena.
 2. Dominio: YYYY donde es el año a 4 dígitos y MM el mes a 2 dígitos.
- ix. trimestre_anio.
 1. Tipo de dato: Integer.
 2. Dominio: {1, 2, 3, 4}.

Jerarquía: fecha > numero_día_mes_calendario > día_semana > id_fecha

b. Libro. Identificar el objeto de la actividad.

- i. id_libro (PK): llave primaria que sustituirá a la llave natural por motivos de eficiencia.
 1. Tipo de dato: Integer.
 2. Dominio: {1, 2, 3, 4}.
- ii. id_titulo (NK)
 1. Tipo de dato: Integer.
 2. Dominio: {1, 2, 3, 4}.
- iii. autor_libro
 1. Tipo de dato: Cadena.
 2. Dominio: autor del libro.
- iv. editorial_libro
 1. Tipo de dato: Cadena.
 2. Dominio: editorial del libro.
- v. formato_libro
 1. Tipo de dato: Cadena.
 2. Dominio: tipo del formato del libro.

Jerarquía: autor_libro > formato_libro > id_titulo

c. Tienda. Donde se produce la actividad(dimensión geográfica).

- i. id_tienda (PK):
 1. Tipo de dato: Integer.
 2. Dominio: {1, 2, 3, ...}.
- ii. id_tienda_nk (NK):
 1. Tipo de dato: Char.
 2. Dominio: identificador de 4 caracteres.
- iii. nombre_tienda:

Tarea 3: Diseño de un almacén de datos.**14 de Marzo de 2018**

1. Tipo de dato: Cadena.
2. Dominio: nombre de la tienda.
- iv. direccion_calle_tienda:
 1. Tipo de dato: Cadena.
 2. Dominio: calle de la tienda.
- v. ciudad_tienda:
 1. Tipo de dato: Cadena.
 2. Dominio: ciudad de la tienda.
- vi. pais_tienda:
 1. Tipo de dato: Cadena.
 2. Dominio: pais de la tienda.
- vii. estado-ciudad_tienda:
 1. Tipo de dato: Cadena.
 2. Dominio: estado o ciudad de la tienda.
- viii. estado_tienda:
 1. Tipo de dato: Cadena.
 2. Dominio: estado de la tienda.
- ix. codigo_postal_tienda:
 1. Tipo de dato: Integer.
 2. Dominio: código postal de la tienda.

Jerarquia: pais_tienda > estado_tienda > direccion_calle_tienda > nombre_tienda > id_tienda

d. Promoción. Usada para ver si las promociones son adecuadas. Si anuncios, reducciones de precio, despliegues, cupones.

- i. stor_id(NK): llave natural.
 1. Tipo de dato: Char
 2. Dominio: identificador de 4 caracteres.
- ii. id_promocion(PK): llave sustituta para mejorar eficiencia.
 1. Tipo de dato: Integer
 2. Dominio: {1, 2, 3, ...}
- iii. costo_promocion:
 1. Tipo de dato: Double.
 2. Dominio: costo de la promoción que se descontará.

4. Identificar los hechos.**a. Ventas a menudeo.****Identificamos las medidas.**

1. Unidades vendidas
 - a. Tipo de dato: Integer.
 - b. Dominio: Ventas por unidad realizadas.
2. Unidades vendidas a precio regular.
 - a. Tipo de dato: Integer.
 - b. Dominio: número de unidades vendidas sin descuentos.
3. Unidades vendidas con descuento.
 - a. Tipo de dato: Integer.

- b. Dominio: número de unidades vendidas con descuentos.
- 4. Importe de ventas en pesos.
 - a. Tipo de dato: Double.
 - b. Dominio: cantidad recaudada de las ventas en pesos.

b. Ganancias.

Identificamos las medidas.

- 1. Ganancia bruta en pesos.
 - a. Tipo de dato: Double.
 - b. Dominio: número decimal que representa la ganancia bruta.
- 2. Ganancia neta en pesos.
 - a. Tipo de dato: Double.
 - b. Dominio: número decimal que representa la ganancia neta.
- 3. Precio neto.
 - a. Tipo de dato: Double.
 - b. Dominio: número decimal que representa el precio.
- 4. Importe de descuento en pesos.
 - a. Tipo de dato: Double.
 - b. Dominio: cantidad de descuento en pesos.