

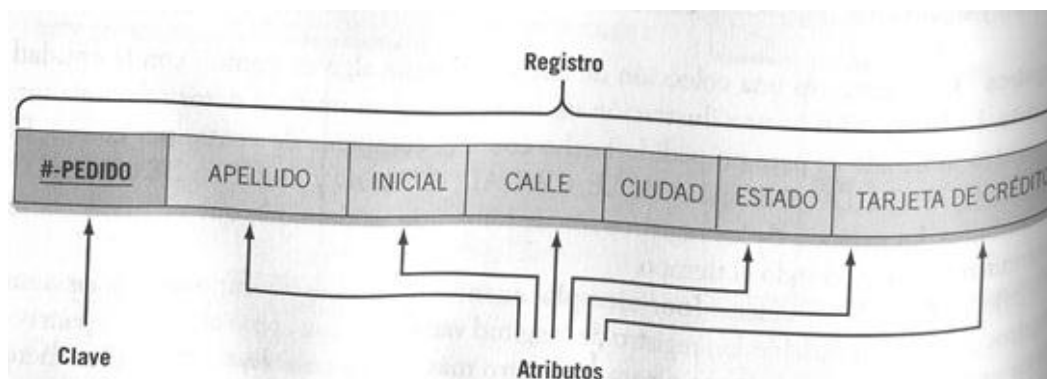
EFFECTIVIDAD DE LA BASE DE DATOS

- Asegurar que los datos se puedan compartir entre los usuarios para una diversidad de aplicaciones.
- Mantener datos que sean exactos y consistentes.
- Asegurar que todos los datos requeridos por las aplicaciones actuales y futuras se podrán acceder con facilidad.
- Permitir a la base de datos evolucionar conforme aumenten las necesidades de los usuarios.
- Permitir a los usuarios construir su vista personal de los datos sin preocuparse por la forma en que los datos se encuentren almacenados físicamente.

Al mundo real se le llamará realidad. En la realidad, los datos recopilados de personas, lugares o eventos se almacenarán eventualmente en un archivo o una base de datos. Para entender la forma y estructura de los datos, se necesita información sobre los datos mismos. A la información que describe los datos se le llama metadatos.

Entidades: Una entidad es cualquier objeto o evento sobre el cual alguien escoge recopilar datos. Una entidad podría ser una persona, lugar o cosa (por ejemplo, un vendedor, una ciudad o un producto). Cualquier entidad también puede ser un evento o unidad de tiempo tal como la avería de una máquina, una venta o un mes o año.

Tuplas o Registros: Un registro o tupla es una colección de datos que tiene algo en común con la entidad descrita. O se puede definir como una Tupla, un objeto que bien puede tener datos o diversos objetos.



Datos		Valor
Número del vendedor	N	5
Nombre del vendedor	A	20
Nombre de la compañía	A	26
Dirección	A	36
Ventas	N	9.2
Ancho	N	2
Alto	N	2
Longitud	N	2
Peso	N	2
Dirección de envío	A	36
Dirección de devolución	A	36
Producto(s)	A	4
Descripción(es)	A	30
Cantidad pedida	N	2
Apellido del cliente	A	24
Inicial	A	1
Calle	A	28
Ciudad	A	12
Estado	A	2
Código postal	N	9
Número de la tarjeta de crédito	N	10
Fecha en la que se hizo el pedido	D	8
Cantidad	\$	7.2
Estado	A	22

Campos

N Numérico
A Alfanumérico o texto
D Fecha MM/DD/AAAA
\$ Dinero
M Memorándum

7.2 significa que el campo ocupa 7 dígitos, dos de los cuales están a la derecha del decimal.

Se podrían especificar formatos especiales para los campos.

MM/DD/AAAA

Claves: Una clave es uno de los datos en un registro que se usa para identificar al registro. Cuando una clave identifica de forma única un registro, se llama clave primaria.

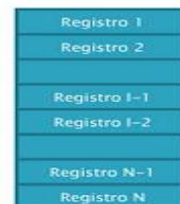
Clave foránea o secundaria: Es una limitación referencial entre dos tablas. La clave foránea identifica una columna o grupo de columnas en una tabla (tabla hija o referendo) que se refiere a una columna o grupo de columnas en otra tabla (tabla maestra o referenciada). Es decir una clave secundaria o foránea, es uno o más campos de una tabla que están relacionados con la clave principal de los campos de otra tabla.

Tipos de archivo: Los archivos se pueden usar para almacenar datos por un periodo indefinido, o se pueden usar para almacenar datos temporalmente para un propósito específico. Los archivos maestros y de tabla se usan para almacenar datos por un periodo largo. Los archivos temporales normalmente se llaman archivos de transacción, archivos de trabajo o archivos de reporte.

- **Archivos maestros:** Los archivos maestros contienen registros para un grupo de entidades. Con frecuencia los atributos se podrían actualizar, pero los registros en sí son relativamente permanentes. Estos archivos son propensos a tener registros grandes que contienen toda la información sobre una entidad de datos. Cada registro normalmente contiene una clave primaria y varias claves secundarias.

- **Archivos de tabla:** Un archivo de tabla contiene datos usados para calcular más datos o medidas de desempeño. Un ejemplo es una tabla de tasas de correos usada para determinar los gastos de envío de un paquete. Otro ejemplo es una tasa de impuestos. Normalmente los archivos de tabla se leen únicamente por un programa.
- **Archivos de transacción:** Un archivo de transacción se usa para hacer cambios que actualizan el archivo maestro y producen informes. Suponga que el archivo maestro de un suscriptor de periódico necesita ser actualizado; el archivo de transacción contendría el número del suscriptor y un código de transacción tal como E para extender la suscripción, C para cancelar la suscripción o A para cambiar la dirección. Después sólo se necesita introducir la información relevante para la actualización; es decir, la longitud de renovación si es E o la dirección si es A. Si la suscripción fue cancelada no se necesitaría ninguna información adicional. El resto de la información ya existe en el archivo maestro.
- **Archivos de reporte.** Cuando se necesita imprimir un informe y no hay ninguna impresora disponible (por ejemplo, cuando la impresora está ocupada imprimiendo otros trabajos), se usa un archivo de reporte. Enviar la salida a un archivo en lugar de a una impresora se denomina *spooling*. Después, cuando el dispositivo está listo, el documento se puede imprimir. Los archivos de reporte son muy útiles, debido a que los usuarios pueden tomar los archivos de otros sistemas de cómputo y enviarlos a dispositivos especializados tales como graficadores, impresoras láser, unidades de microficha e incluso máquinas de composición tipográfica computarizadas.

Organización secuencial: Cuando los registros están físicamente en orden en un archivo, se dice que éste es un archivo secuencial. En este tipo de archivo, un formato fijo es usado para los registros. Todos los registros tienen el mismo tamaño, constan del mismo número de campos de tamaño fijo en un orden particular.



Listas enlazadas: Cuando los archivos se almacenan en dispositivos de acceso directo tal como un disco, las opciones se extienden. Los registros se pueden ordenar lógicamente, en lugar de físicamente, usando listas enlazadas. Las listas enlazadas se logran usando un conjunto de indicadores para dirigirlo al próximo registro lógico ubicado en cualquier parte del archivo.

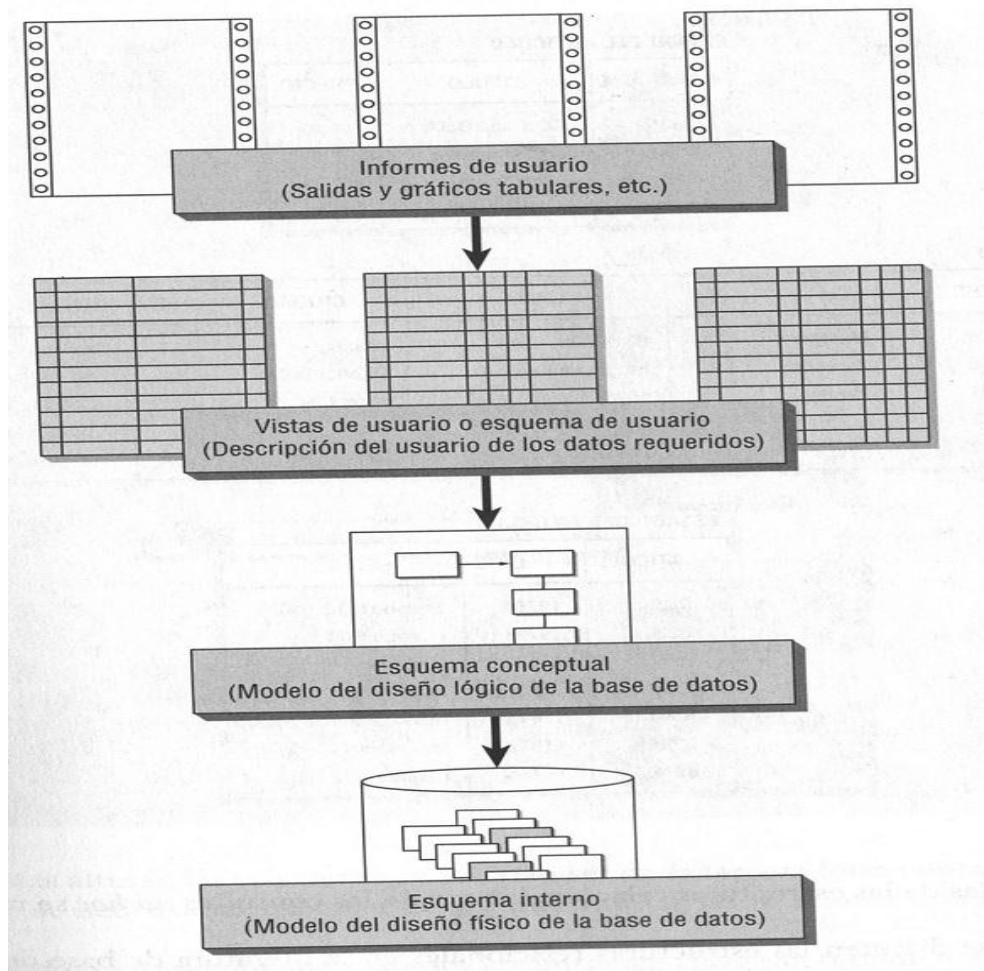
VISTAS LÓGICAS Y FÍSICAS DE DATOS

Una base de datos, a diferencia de un archivo, es diseñada para ser compartida por muchos usuarios. Está claro que todos los usuarios ven los datos de formas diferentes. Es decir a la forma en que un usuario visualiza y describe los datos como una vista de usuario. Sin embargo, el problema es que diferentes usuarios tienen vistas de usuario distintas. El analista de sistemas debe examinar estas

vistas y debe desarrollar un modelo lógico global de la base de datos. Finalmente, dicho modelo lógico se debe transformar en el diseño físico correspondiente de la base de datos. El diseño físico describe la forma como se almacenan y relacionan los datos, así como también la forma en que se acceden.

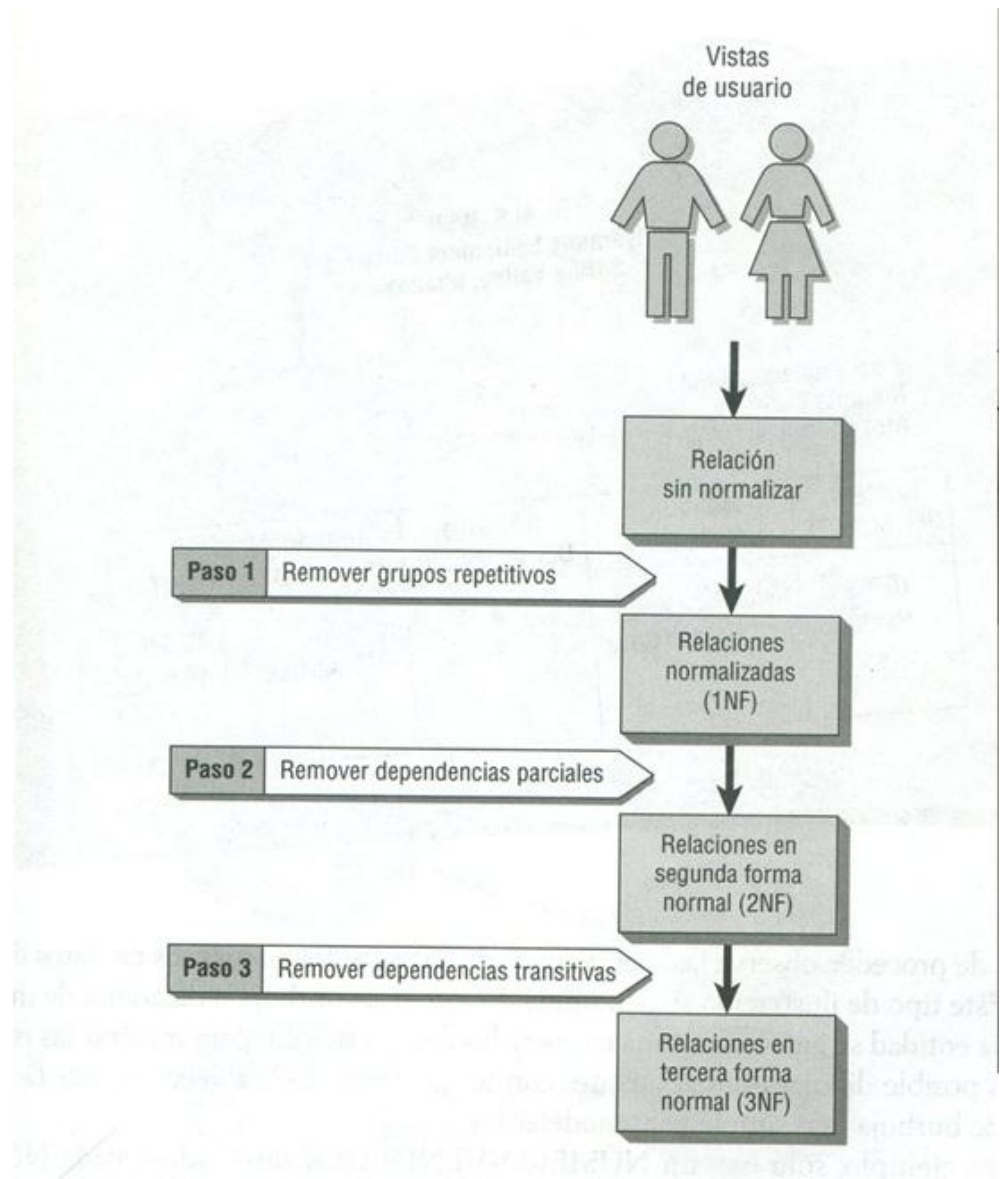
Existen los siguientes tipos principales de bases de datos:

- **Estructuradas de forma lógica jerárquica y red:** se pueden encontrar en sistemas heredados (antiguos). Hoy en día, un analista típicamente diseñaría una base de datos relacional.
- **Estructuras relacionales de datos:** Una estructura relacional de datos consiste en una o más tablas bidimensionales, las cuales se denominan relaciones. Las filas de la tabla representan registros y las columnas contienen atributos. En esta es importante tener claridad que las vistas se denominan esquema tal como muestra la siguiente figura con su respectivo informe de usuario y la vista de usuario (esquema de usuario) se relaciona al modelo lógico (esquema conceptual) y al diseño físico (esquema interno).

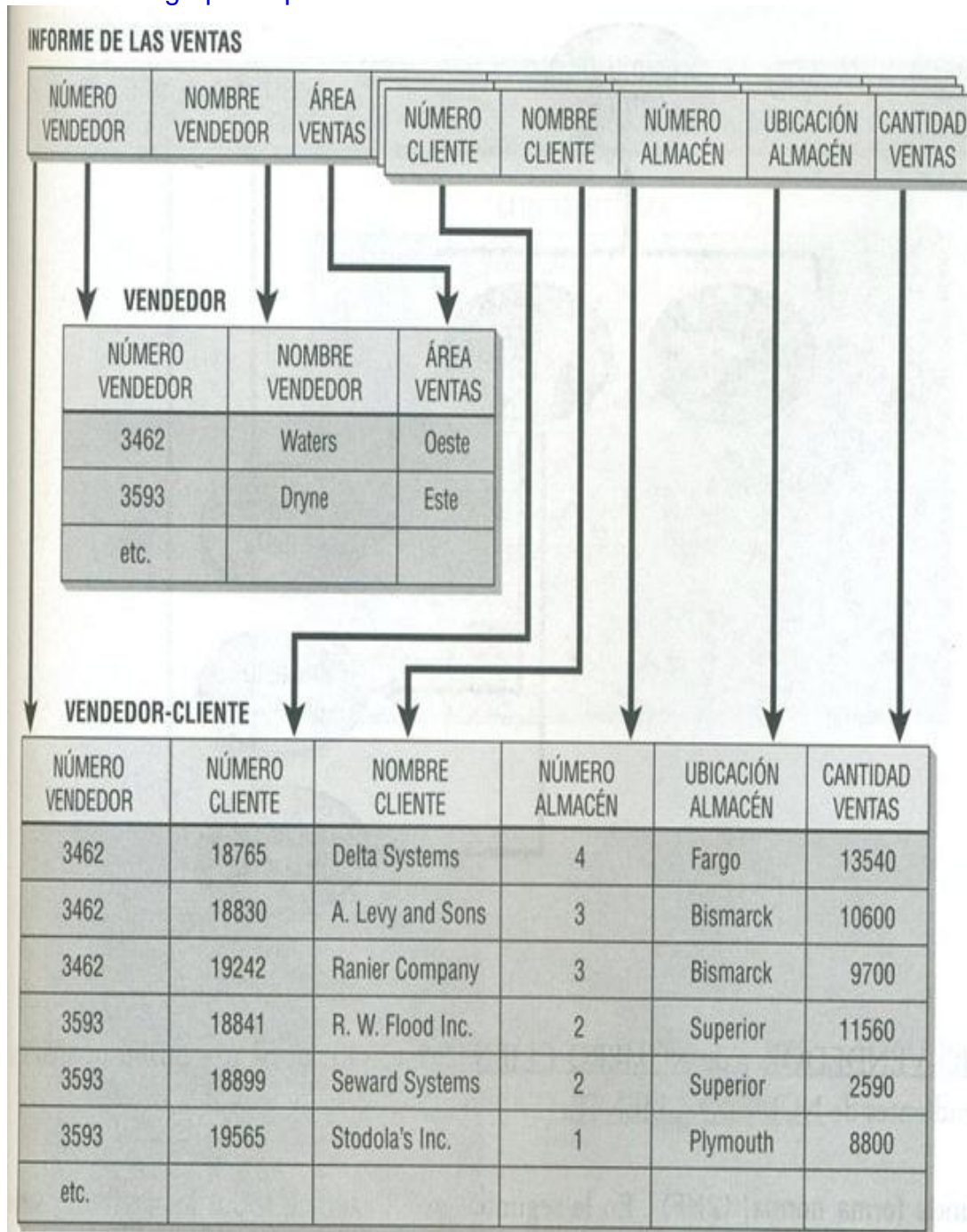


NORMALIZACIÓN

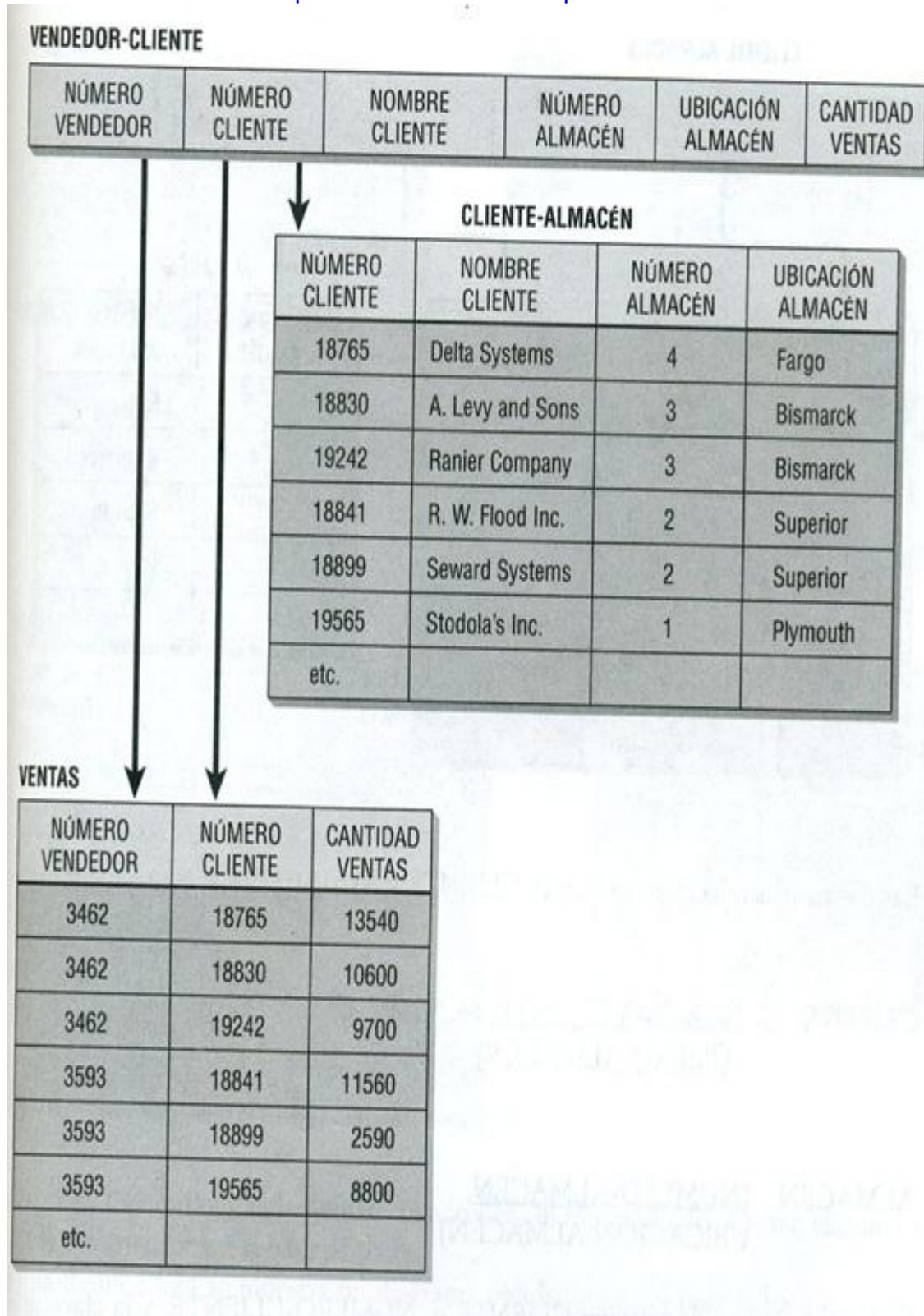
La normalización es la transformación de las vistas de usuario complejas y del almacén de datos a un juego de estructuras de datos más pequeñas y estables. Además de ser más simples y estables, las estructuras de datos normalizadas son más fáciles de mantener que otras estructuras de datos. Es importante tomar como referencia los siguientes pasos:



Primera forma normal (1NF): en el proceso incluye quitar todos los grupos repetitivos e identificar la clave primaria. Para ello, la relación se debe dividir en dos o más relaciones. Las relaciones ya podrían ser de la tercera forma normal, pero probablemente se necesitarán más pasos para transformar las relaciones a la tercera forma normal. Es decir el primer paso para normalizar una relación es remover los grupos repetitivos.



Segunda forma normal (2NF): En la segunda forma normal, todos los atributos serán funcionalmente dependientes de la clave primaria.



Tercera forma normal (3NF): Una relación normalizada está en tercera forma normal si todos los atributos sin clave son funcionalmente dependientes por completo de la clave primaria y si no hay dependencias transitivas (sin claves) es decir se aplica a tres o más variables.

