

# **NORMALIZACION**

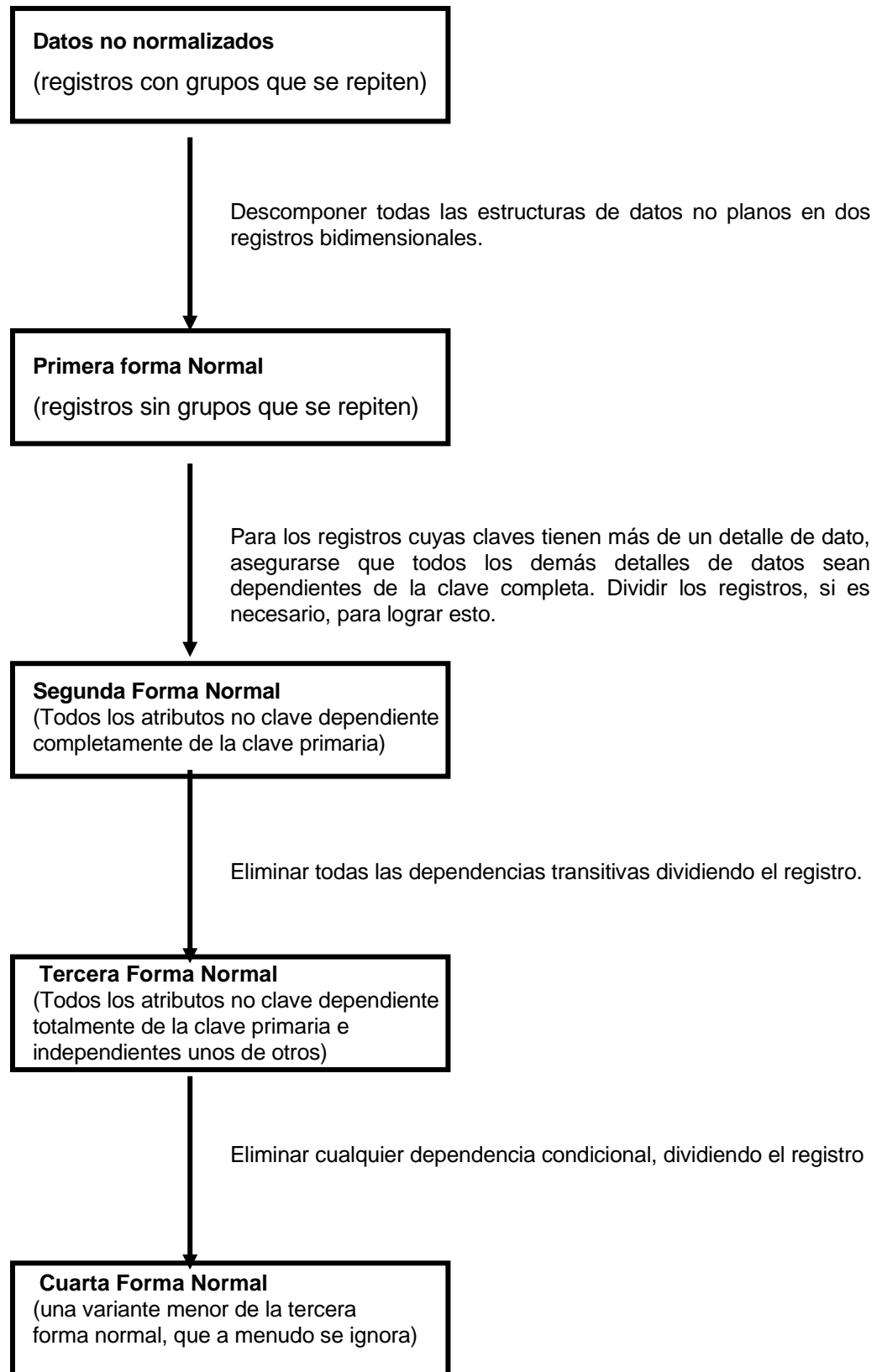
## **1. INTRODUCCION**

El término normalización de los datos se refiere a la forma en que los detalles de los datos son agrupados en estructuras de registro. La cuarta (o tercera) forma normal es un agrupamiento de datos diseñado para evitar anomalías y problemas que pueden ocurrir con los datos. El concepto se originó con las matemáticas de E.F.Codd.

En la vida real los datos se encuentran en grupos de detalles de datos. Existen en facturas, orden de compra, formato de impuesto, licencias de conducir, etc. Estos agrupamientos usualmente no están en una forma normalizada. No es de sorprender que los analistas de sistemas a menudo hayan implementado registros de computadoras que tampoco están normalizados. Sin embargo, los datos que no están normalizados pueden conducir a diversos problemas en el futuro.

La experiencia ha mostrado que cuando los datos del computador están organizados en cuarta forma normal, las estructuras de datos resultantes son más estables y capaces de acomodarse a los cambios de negocio.

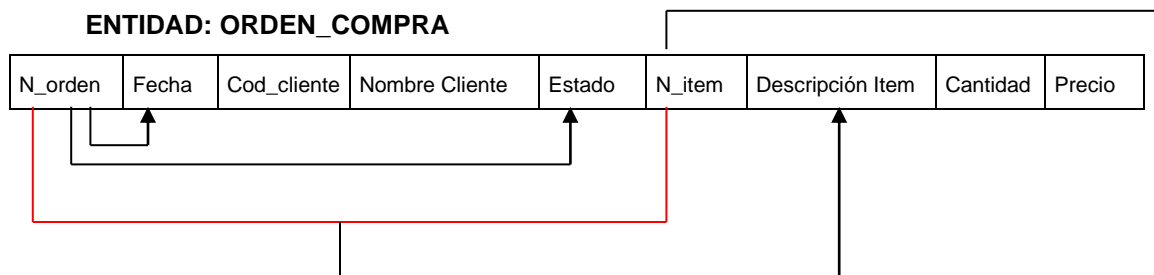
## 2. LA NORMALIZACIÓN DE LOS DATOS



### 3. DATOS SIN NORMALIZAR

Registros con grupos repetidos, es decir, no cumple con ninguna regla de normalización.

Ejemplo: se tiene los datos de la orden de compra.



Esta entidad, contiene un grupo repetido constituido por los siguientes atributos son: n\_item, descripción\_item, cantidad y precio.

**TABLA: ORDENES\_COMPRA**

N_orden	Fecha	Cod_cliente	Nombre Cliente	Estado	Cod_prod	Descripción Item	Cantidad	Precio
2340	6/6/01	001	Luis Soto	Aprob	01	Televisor	1	900.00
					02	DVD	1	850.00
					03	Equipo de sonido	1	650.00
2341	6/6/01	002	Liliana Vargas	Aprob	04	Mueble	1	300.00
2342	7/6/01	001	Luis Soto	Atend	05	Raqueta	1	65.00
					06	Funda	2	10.00

Como se puede observar esta tabla contiene un grupo repetido para n\_item, descripción\_item, cantidad y precio.

### PROBLEMAS

- El número de grupo repetidos por orden de compra tiene un límite, por lo que pasado este límite ya no se puede registrar más ocurrencias.
- Se reserva espacio para el máximo permitido de grupo repetido por orden de compra, pero las compra puede tener un solo grupo repetido desperdiándose de este modo espacio reservado en el disco con datos NULL.

### SOLUCION

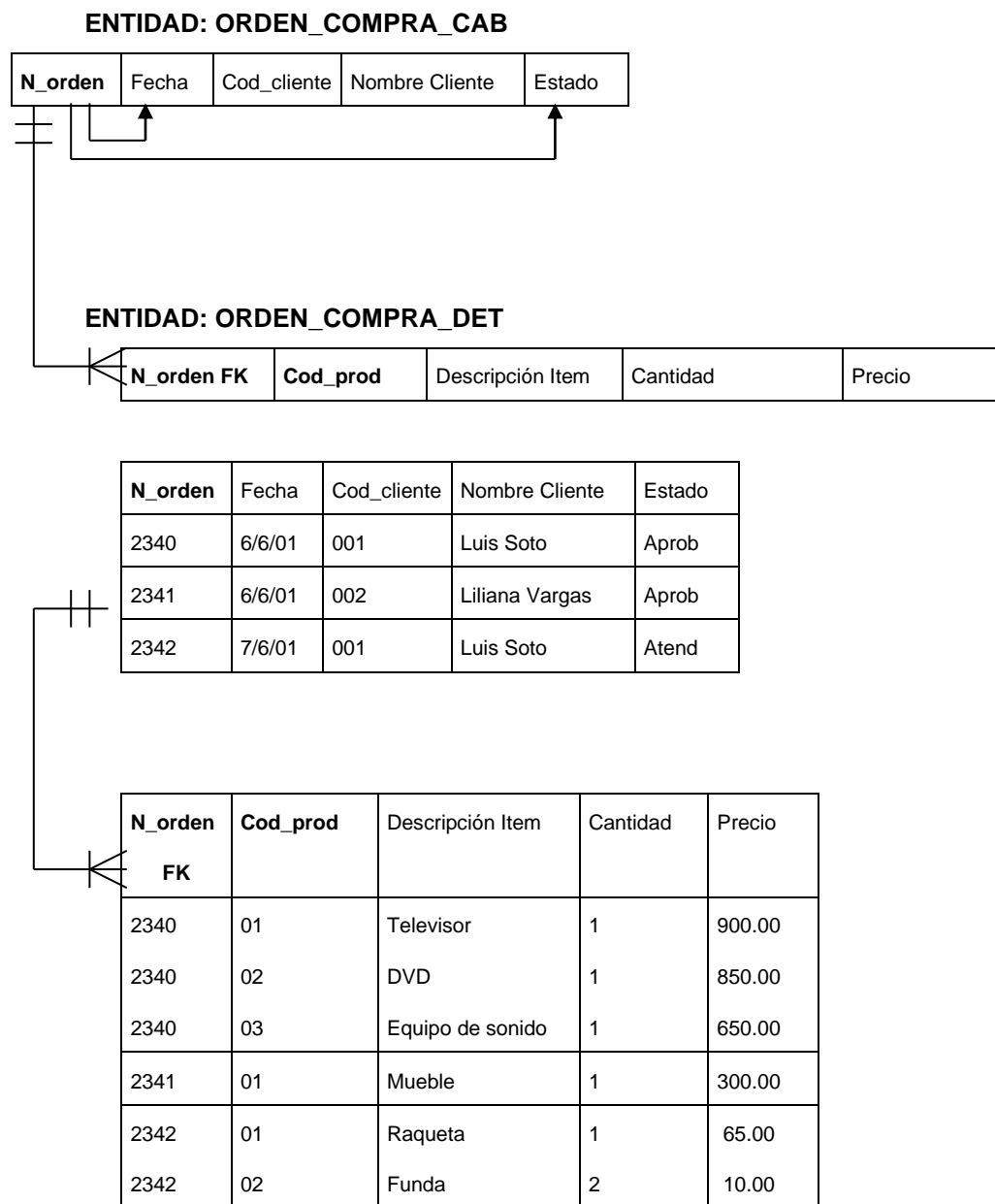
- Separar los grupos repetidos en una entidad y los grupos no repetidos en otra entidad.
- No olvidarse de la PK de cada entidad.

#### 4. PRIMERA FORMA NORMAL

La conversión a primera forma normal consiste en remover el grupo repetido.

##### PASOS

- Remover los grupos de repetición.
- Crear una entidad con la PK de la entidad base y el grupo de repetición.



## 5. SEGUNDA FORMA NORMAL

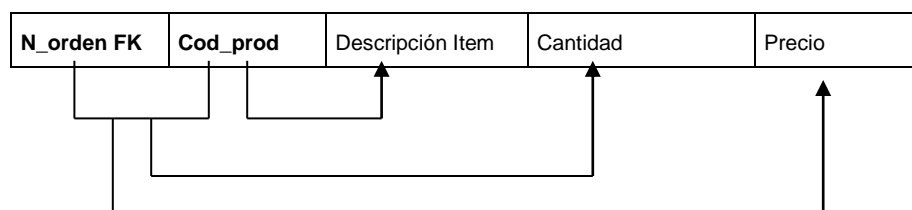
Remover cualquier atributo no clave que no dependa de la PK completa, se da en aquellas entidades que tienen su PK compuesta. En otras palabras los atributos no claves deben por completo funcionalmente depender de la PK.

### PASOS

- Determinar cuales atributos, que no son llaves, no dependen de la llave primaria completa de la entidad.
- Remover estas columnas de la entidad base.
- Crear una segunda entidad con esas columnas y la(s) columna(s) de la PK de la cual dependen.

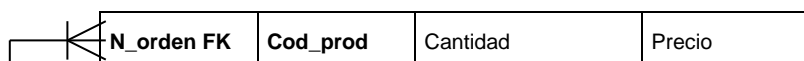
**ENTIDAD: ORDEN\_COMPRA\_CAB**, se encuentra en segunda forma normal, observe que la PK esta constituido por un atributo.

#### ENTIDAD: ORDEN\_COMPRA\_DET

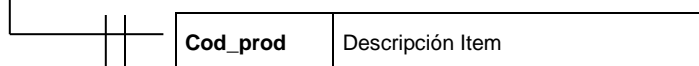


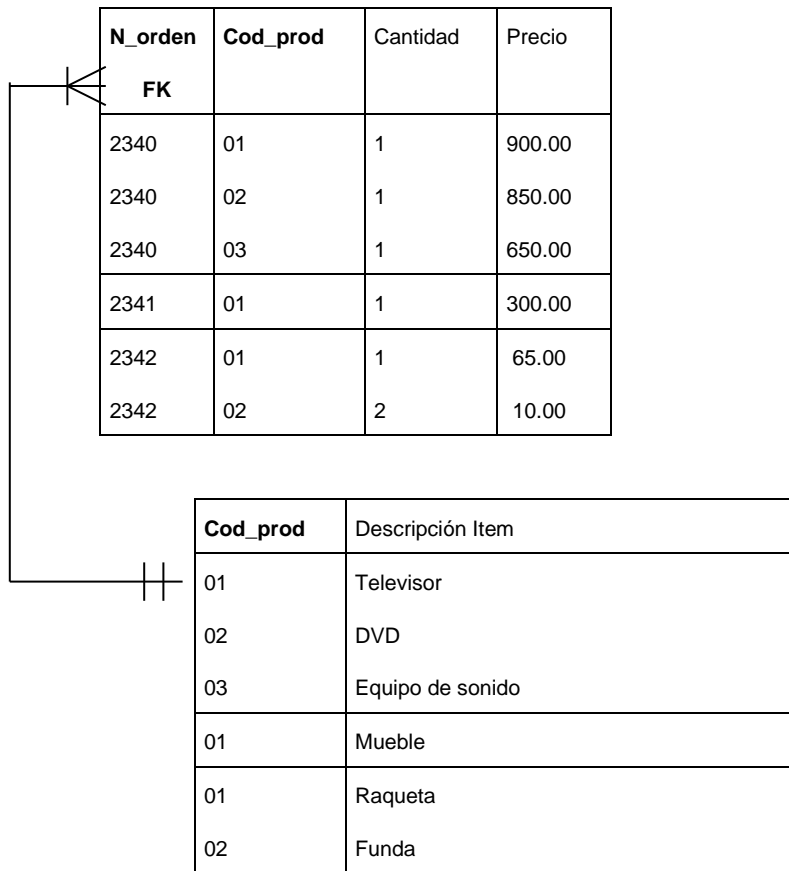
### SOLUCIÓN

#### ENTIDAD: ORDEN\_COMPRA\_DET



#### ENTIDAD: PRODUCTO



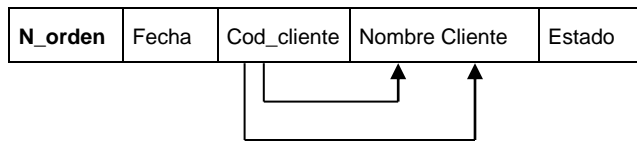
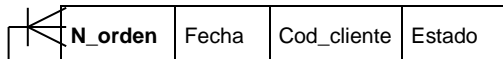
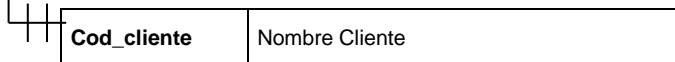


## 6. TERCERA FORMA NORMAL

Para aquellas entidades que se encuentren en 2FN, puede que no se encuentren en 3FN, esto se da cuando un atributo no clave depende de otro atributo no clave.

### PASOS

- Determinar los atributos no claves que dependen de otro atributo no clave.
- Remover estos atributos de la entidad base
- Crear una segunda entidad con estos atributos y con el atributo no clave de la cual son dependientes.

**ENTIDAD: ORDEN\_COMPRA\_CAB****SOLUCION****ENTIDAD: ORDEN\_COMPRA\_CAB****ENTIDAD: CLIENTE**

N_orden	Fecha	Cod_cliente	Estado
2340	6/6/01	001	Aprob
2341	6/6/01	002	Aprob
2342	7/6/01	001	Atend

Cod_cliente	Nombre Cliente
001	Luis Soto
002	Liliana Vargas
001	Luis Soto

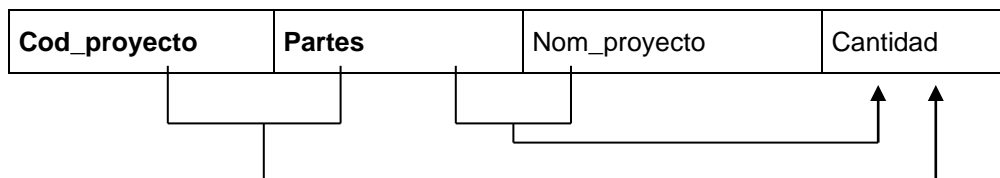
**7. FORMA NORMAL DE BOYCE CODD**

Para aquellas entidades que se encuentren en 3FN, pero que tienen varias claves candidatas compuestas y que tienen ciertos atributos comunes.

**PASOS**

- Identificar todas las claves candidatas y compuestas que tienen atributos comunes.
- Identificar una super clave para la entidad base y aquellos atributos que dependen funcionalmente de toda la super clave.

- Aquellos atributos que no dependen de la super clave remover en otra entidad.

**EJEMPLO****SOLUCION**

Cod_proyecto	Partes	Cantidad
--------------	--------	----------

Cod_proyecto	Nom_proyecto
--------------	--------------

**8. CUARTA FORMA NORMAL**

Una entidad no se encuentra en cuarta forma normal si tiene una dependencia multivaluada no trivial.

**EJEMPLO**

Persona	Proyecto	Parte	aptitud
Luis	P1	Tuerca	Sistemas
Luis	P1	Tornillo	Sistemas
Luis	P1	Tuerca	Industrial
Luis	P1	Tornillo	Industrial
Liliana	P1	Tuerca	Sistemas
Liliana	P1	Tornillo	Sistemas
Oriol	P2	Martillo	Contador
Oriol	P2	Serrucho	Contador
Emperatriz	P2	Martillo	Inglés
Emperatriz	P2	serrucho	Inglés

Pyto,parte: persona, aptitud: persona, proyecto



**9. QUINTA FORMA NORMAL**

Una entidad no se encuentra en cuarta forma normal si tiene una dependencia multivaluada no trivial.

# DESNORMALIZACIÓN

## 1. INTRODUCCION

Siempre es recomendable empezar con tablas en tercera forma normal o superior, se debe evaluar cuidadosamente cuando desnormalizar por ello:

- ♦ Ser extremadamente renuente a desnormalizar el diseño de default de la tabla.
- ♦ La desnormalización puede causar problemas de inconsistencia de datos.

La denormalización puede ser la solución para transacciones con requerimientos de performance como:

- ♦ High Throughput (Alto uso de recursos)
- ♦ Alta frecuencia
- ♦ Tiempo de respuesta rápido

Considerar todas las otras opciones antes de desnormalizar, especialmente el agregar o cambiar la estructura de índices.

## 2. COMBINAR TABLAS

Se puede juntar tablas como por ejemplo BANCO y CUENTA CORRIENTE en una sola tabla. Se puede establecer esta solución de existe un gran volumen de datos.

Tablas con pocos registros se pueden llevar a una tabla maestra.

## 3. CREAR VECTORES.

Un vector es un arreglo de una dimensión con un número compuesto de valores. Un vector de datos se puede representar como un conjunto de reglones o un conjunto de columnas.

### COLUMNAS:

AÑO	MES	PROVEEDOR	IMPORTE
2001	ENERO	001	2000.00
2001	ENERO	002	4000.00

### REGLONES:

PROVEEDOR	ENE	FEB	MAR
001	1000.00	1500.00	1800.00
002	4000.00	4500.00	2000.00