



UNIVERSIDAD NACIONAL DE JUJUY
FACULTAD DE INGENIERÍA

ESCUELA DE MINAS
CÁTEDRA DE BASES DE DATOS I

UNIDAD N° 6
DEPENDENCIAS FUNCIONALES Y
NORMALIZACIÓN

Mg. Ing. Héctor Pedro Liberatori
Año 2023

El problema en el diseño de Bases de Datos, consiste en elegir una estructura lógica conveniente para el mismo.

Sin embargo es preciso una medida formal, de porqué una agrupación de atributos para formar un esquema de relación, puede ser mejor que otra.

Figura 1: Dependencia Funcionales y Formas Normales

Existen 2 niveles en los que se puede evaluar la bondad de los esquemas de relación:

- **Nivel lógico (o conceptual):** es la forma en la que los usuarios interpretan los esquemas de relación y el significado de sus atributos.
- **Nivel de implementación (o de almacenamiento):** se refiere a como se almacenan y actualizan las tuplas de una relación base.

Figura 2: Evaluación de los esquemas de relación y atributos

Existen 4 medidas informales de calidad para diseñar esquemas de relación:

- **Semántica de los atributos.**
- **Reducción de los valores redundantes en las tuplas.**
- **Reducción de los valores nulos en la tuplas.**
- **Eliminación de la posibilidad de tuplas espurias.**

Figura 3: Pautas informales de diseño para los Esquemas de Relación

Una Dependencia Funcional (DF) es un vínculo muchos a uno que va de un conjunto de atributos a otro, dentro de un determinado esquema de relación.

Dado un esquema de relación R ,

la DF: $A \longrightarrow B$

es válida para R , si y sólo si, siempre que 2 o más tuplas de R coinciden en su valor A , tienen también el mismo valor de B .

Figura 4: Definición de Dependencia Funcional

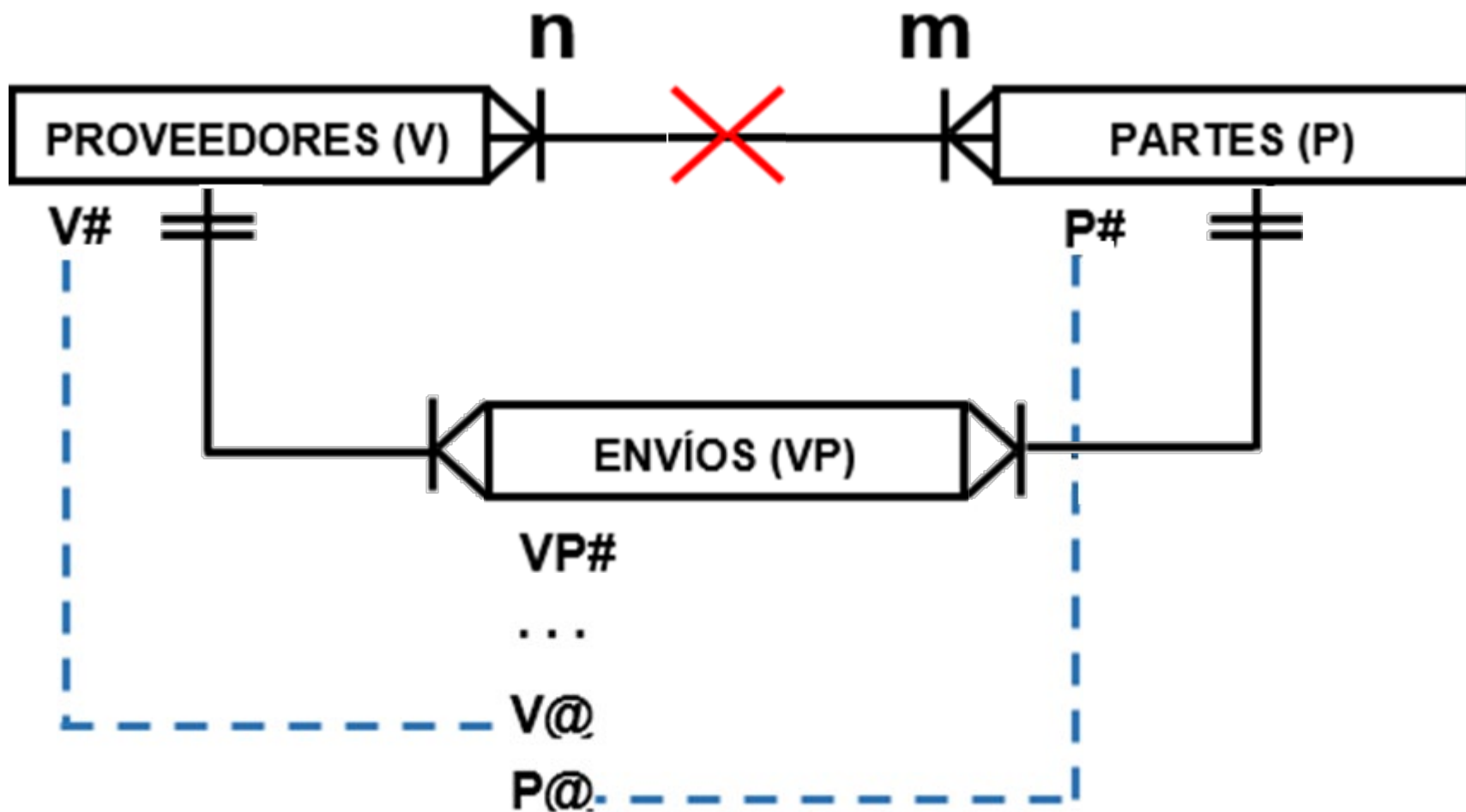


Figura 5: Base de Datos de Proveedores y Partes

V (Proveedores)			
V#	Proveedor	Status	Ciudad
V1	Smith	20	Londres
V2	Jones	10	París
V3	Blake	30	París
V4	Clark	20	Londres
V5	Adams	30	Atenas

P (Partes)				
P#	Parte	Color	Peso	Ciudad
P1	Tuerca	Rojo	12	Londres
P2	Perno	Verde	17	París
P3	Tornillo	Azul	17	Roma
P4	Tornillo	Rojo	14	Londres
P5	Leva	Azul	12	París
P6	Engrane	Rojo	19	Londres

VP (Envíos)		
V#	P#	Cant
V1	P1	300
V1	P2	200
V1	P3	400
V1	P4	200
V1	P5	100
V1	P6	100
V2	P1	300
V2	P2	400
V3	P2	200
V4	P2	200
V4	P4	300
V4	P5	400

Figura 6: Base de Datos de Proveedores y Partes

VPC			
V#	Ciudad	P#	Cant
V1	Londres	P1	300
V1	Londres	P2	200
V1	Londres	P3	400
V1	Londres	P4	200
V1	Londres	P5	100
V1	Londres	P6	100
V2	París	P1	300
V2	París	P2	400
V3	París	P2	200
V4	Londres	P2	200
V4	Londres	P4	300
V4	Londres	P5	400

DF: **V#** → **CIUDAD**

Figura 7: Varrel VP modificada, se agregó el atributo Ciudad (del proveedor)

Dada la siguiente DF: $A \longrightarrow B$, se dice que la misma puede ser:

DF Completa: $\{V\# \longrightarrow \{CIUDAD\}$

DF Trivial: $\{V\#,P\# \longrightarrow \{V\# \}$

DF Elemental: $\{V\#,P\# \longrightarrow \{CANT\}$

Figura 8: Tipos de Dependencias Funcionales

Antes de definir los tipos transitivos de DF, se debe tener en cuenta el concepto de descriptores equivalentes:

$$A \longrightarrow B \wedge B \longrightarrow A \equiv A \longleftrightarrow B$$

cuando el SGBD haga cumplir la DF de **B**, también hará cumplir la DF de **A**.

Figura 9: Descriptores equivalentes

Dados 3 descriptores **A**, **B**, y **C** del esquema de relación **R**.

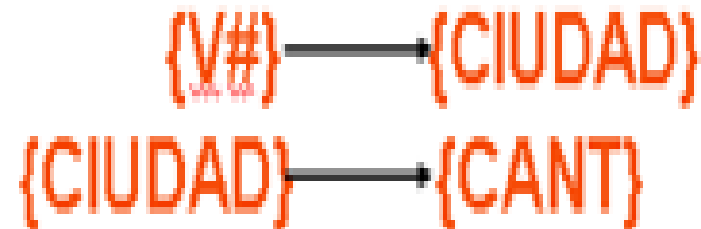
Existe dependencia transitiva de **C** respecto de **A**, a través de **B**, si se cumple:

$$A \longrightarrow B \wedge B \longrightarrow C$$

y además:

A y **B** no deben ser: $A \equiv B$

Figura 10: Dependencia Funcional Transitiva



$\{V\#\}$ y $\{CIUDAD\}$ no deben ser descriptores equivalentes, es decir: $\{V\#\} \longrightarrow \{CIUDAD\}$.

Todos los proveedores de una ciudad, proveen siempre la misma cantidad y ésta NO podrá ser la misma en otra ciudad.

Figura 11: Ejemplo de DF Transitiva sobre la Relación Envíos

Dados 3 descriptores **A**, **B**, y **C** del esquema de relación **R**.

Existe dependencia transitiva estricta de **C** respecto de **A**, a través de **B**, si se cumple:

$$A \longrightarrow B \wedge B \longrightarrow C$$

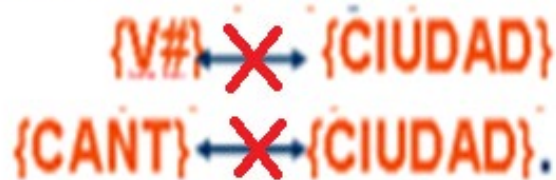
y además:

$$A \text{ y } B; B \text{ y } C \text{ no deben ser: } A \equiv B$$
$$C \equiv B$$

Figura 12: Dependencia Funcional Transitiva Estricta



No siendo equivalentes $\{V\# \}$ y $\{CIUDAD\}$ ni $\{CANT\}$ y $\{CIUDAD\}$, es decir:



Todos los proveedores de una ciudad, proveen siempre la misma cantidad y esta podrá ser la misma en otra ciudad.

Figura 13: Ejemplo de DF Transitiva Estricta sobre la Relación Envíos

Sea **S** un conjunto de DF, será irreducible si y sólo si se satisfacen las siguientes 3 propiedades:

- La parte derecha de toda DF en **S** involucra un solo atributo.
- La parte izquierda de toda DF en **S** es a su vez irreducible (DF irreducible a izquierda).
- No es posible descartar de **S** ninguna DF, sin pérdida de información.

DFs para el esquema de relación P (Partes):



Figura 14: Dependencia Funcional Irreducible

Al conjunto de todas las DFs implicadas en un conjunto S de DFs se lo denomina CIERRE de S , y se lo simboliza S^+ .

La cuestión de encontrar un equivalente irreducible para un conjunto determinado de DFs, es un problema práctico interesante.

Figura 15: Cierre de un Conjunto de Dependencias

La normalización es un proceso de análisis de los esquemas de relación, basados en sus DFs y claves 1º.

Para alcanzar las propiedades de:

- Minimizar redundancia.**
- Minimizar anomalías de actualización.**

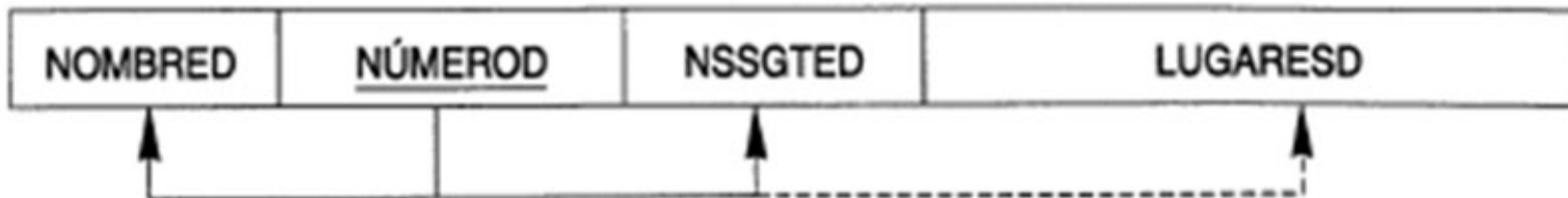
Figura 16: Formas Normales basadas en Claves 1º

La 1FN se definió para prohibir los atributos multivaluados, atributos compuestos (anidados), y sus combinaciones.

Definición: establece que el dominio de un atributo debe incluir sólo valores atómicos (simples e indivisibles).

Figura 17: Primera Forma Normal (1FN)

DEPARTAMENTO



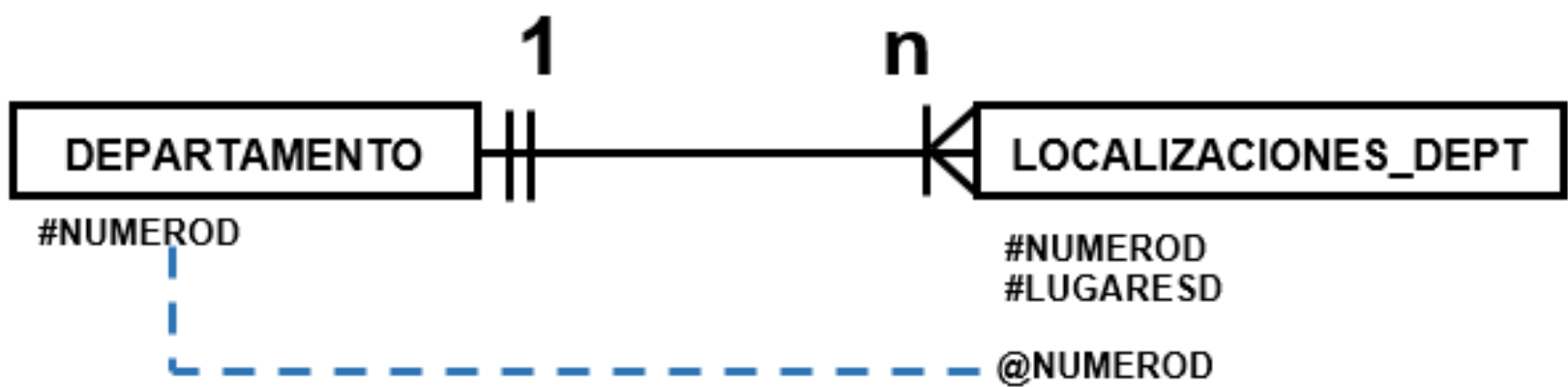
DEPARTAMENTO

NOMBRED	<u>NÚMEROD</u>	NSSGTED	LUGARESD
Investigación	5	333445555	(Belén, Sacramento, Higueras)
Administración	4	987654321	(Santiago)
Dirección	1	888665555	(Higueras)

DEPARTAMENTO

NOMBRED	<u>NÚMEROD</u>	NSSGTED	<u>LUGARESD</u>
Investigación	5	333445555	Belén
Investigación	5	333445555	Sacramento
Investigación	5	333445555	Higueras
Administración	4	987654321	Santiago
Dirección	1	888665555	Higueras

Figura 18: Normalización en Primera Forma Normal (1FN)



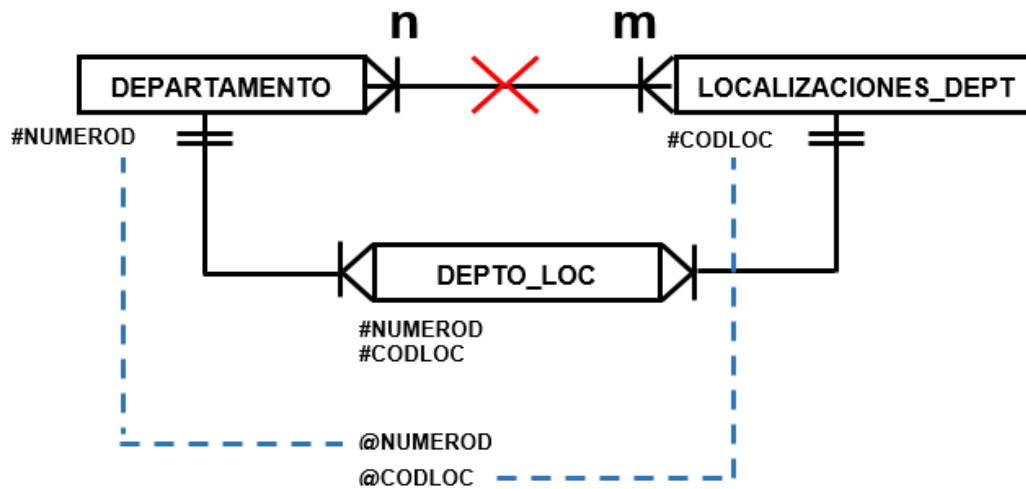
DEPARTAMENTO		
<u>NUMEROD</u>	NOMBRED	NSSGTED
5	Investigación	333446666
4	Administración	987654321
1	Dirección	888665555

LOCALIZACIONES_DEPT	
<u>NUMEROD</u>	<u>LUGARESD</u>
5	Belén
5	Sacramento
5	Higueras
4	Santiago
1	Higueras

Figura 19: Normalización 1FN, relación 1-n

DEPARTAMENTO					
<u>NUMEROD</u>	NOMBRED	NSSGTED	LUGARESD1	LUGARESD2	LUGARESD3
5	Investigación	333446666	Belén	Sacramento	Higueras
4	Administración	987654321	Santiago	NULL	NULL
1	Dirección	888665555	Higueras	NULL	NULL

Figura 20: Normalización 1FN, agregando atributos LUGARESD



DEPARTAMENTO		
<u>NUMEROD</u>	NOMBRED	NSSGTED
5	Investigación	333446666
4	Administración	987654321
1	Dirección	888665555

LOCALIZACIONES_DEPT	
<u>CODLOC</u>	LUGARESD
1	Belén
2	Sacramento
3	Higueras
4	Santiago

DEPTO_LOC	
<u>NUMEROD</u>	<u>CODLOC</u>
5	1
5	2
5	3
4	4
1	3

Figura 21: Normalización 1FN, relación n-m

DEPARTAMENTO				
<u>NUMEROD</u>	NOMBRED	NSSGTED	LOCALIZACIONES_DEP	
			CODLOC	LUGARESD
5	Investigación	333446666	1	Belén
			2	Sacramento
			3	Higueras
4	Administración	987654321	4	Santiago
1	Dirección	888665555	3	Higueras

DEPARTAMENTO (NUMEROD, NOMBRED, NSSGTED, {LOCALIZACIONES_DEPT(CODLOC, LUGARESD)})

Figura 22: Normalización 1FN, relación anidada

La 2FN se basa en el concepto de **DF Total**. Es decir que todos los atributos **NO** clave de una relación, deben depender de forma total de su clave 1º.

En la prueba de la 2FN se verifican las DF de los atributos del miembro izquierdo que son parte de la clave 1º:



Figura 23: Segunda Forma Normal (2FN)

Si la clave 1º contiene un sólo atributo, entonces la relación ya se encuentra en 2FN, si los atributos no clave tienen una DF total de su clave 1º.

EMP_PROY está en 1FN (Fig. 26), pero no está en 2FN: los atributos no clave NOMBREE, NOMBREPR, y LUGARP; no tienen DF Total de la clave 1º.

Figura 24: Segunda Forma Normal (2FN)

Si un esquema de relación no está en 2FN, al normalizarlo da lugar a varias relaciones 2FN.

En estas relaciones 2FN los atributos no clave tienen una DF Total de la clave 1º (Fig. 26).

Figura 25: Segunda Forma Normal (2FN)

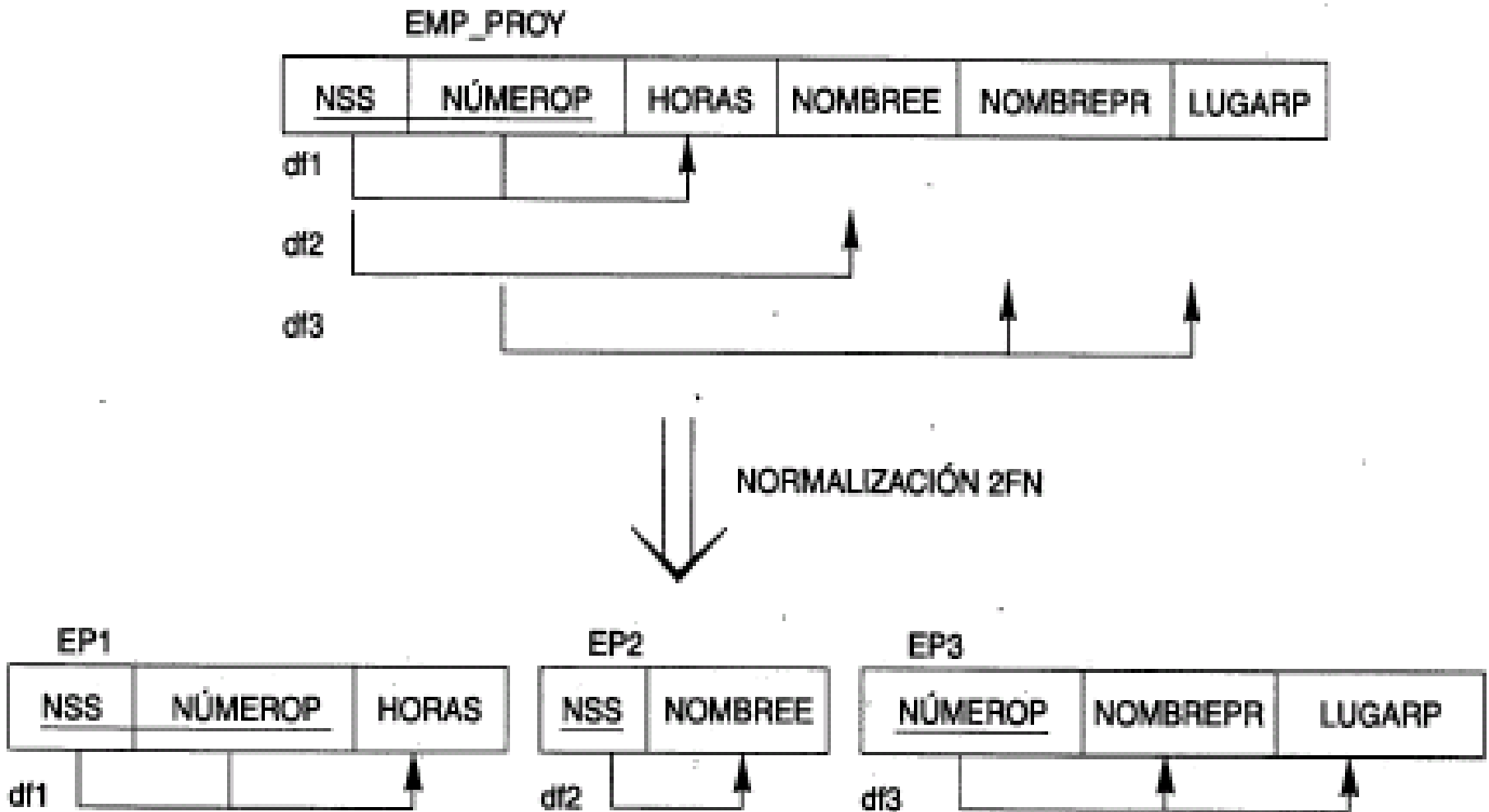


Figura 26: Normalización en 2FN

Definición: un esquema de relación **R** está en 2FN, si también está en 1FN, y todos los atributos no clave de **R** tienen una DF Total de su clave 1º.

Una operación de JUNTA sobre EP1, EP2, y EP3, debe obtener la relación EMP_PROY.

Figura 27: Normalización en 2FN

La 3FN se basa en el concepto de **DF Transitiva**. Una DF: **A** \longrightarrow **C** en un esquema de relación **R** es transitiva, si existe un conjunto de atributos **B** (no clave) de **R**, y se cumple:

$$\mathbf{A} \longrightarrow \mathbf{B} \wedge \mathbf{B} \longrightarrow \mathbf{C}$$

Figura 28: Tercera Forma Normal (3FN)

La DF: **NSS** → **NSSGTED** (Fig. 30) es transitiva a través de **NUMEROD**, porque se cumple:

NSS → **NUMEROD** ∧

NUMEROD → **NSSGTED**

Y **NUMEROD** no es un atributo clave.

Figura 29: Tercera Forma Normal (3FN)

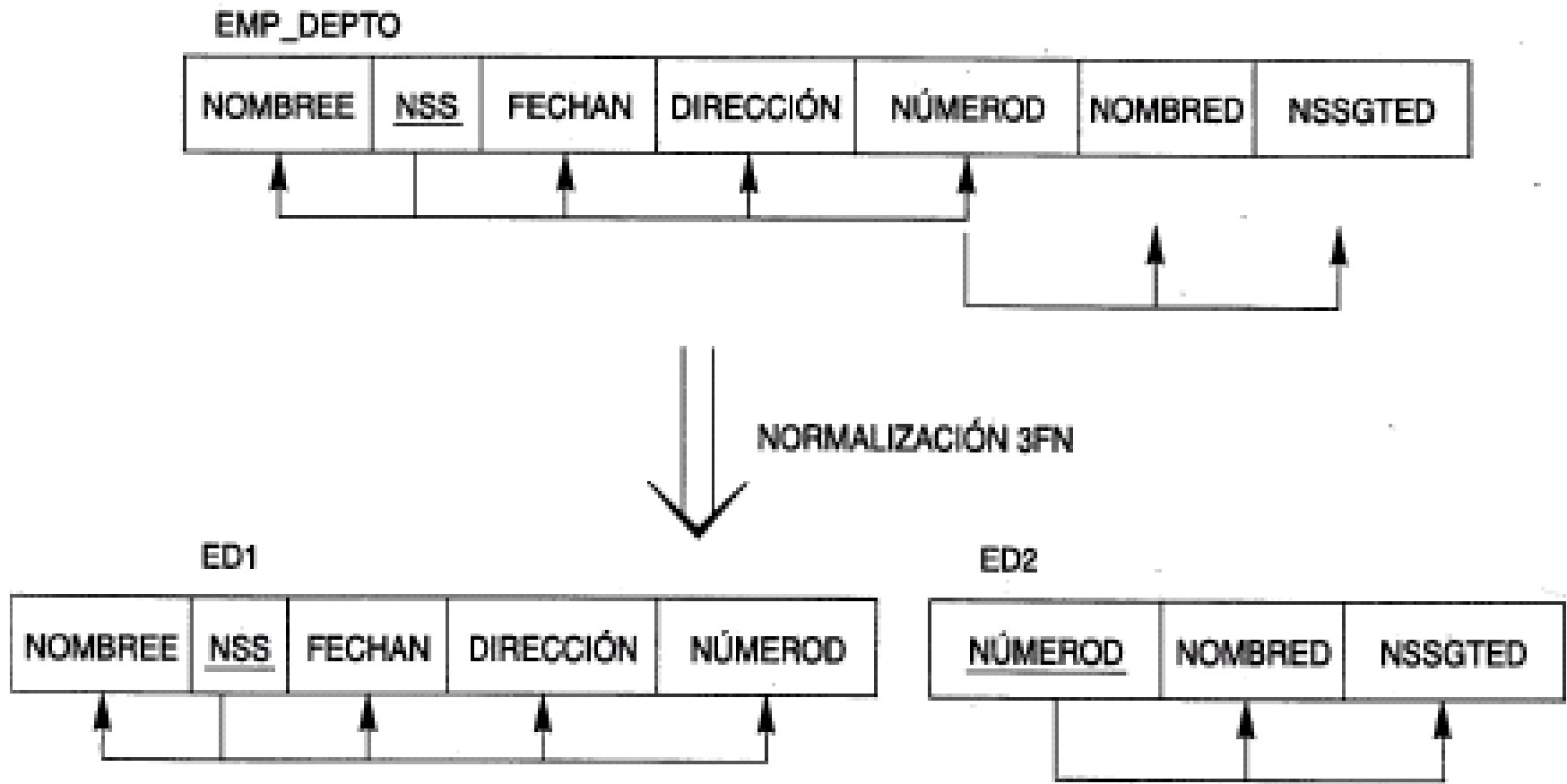


Figura 30: Normalización en 3FN

Definición: un esquema de relación **R** está en 3FN, si también está en 2FN, y ningún atributo no clave de **R** depende de forma transitiva de la clave 1°.

Una operación de JUNTA sobre ED1 y ED2, debe obtener la relación EMP_DEPTO.

Figura 31: Normalización en 3FN

Dada la siguiente relación con su esquema correspondiente:

PROVEEDOR {**V#**, **CIUDAD**, **STATUS**, **P#**, **DIMENSIÓN**, **CANT**}

Primary Key{**V#**, **P#**}

y además se declara la siguiente dependencia funcional:

DF: **CIUDAD** → **STATUS**

El diagrama de dependencias funcionales de PROVEEDOR es:

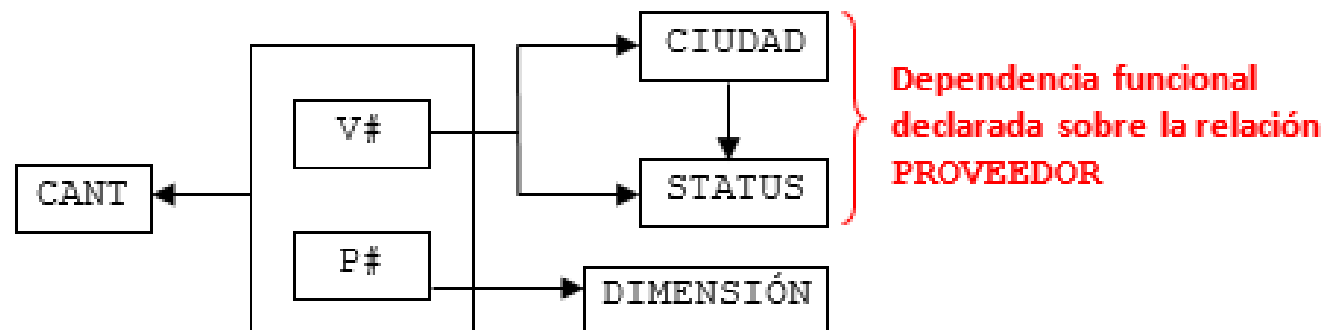


Figura 32: Normalización en Primera Forma Normal (1FN)

Solución del problema: Proyectar la relación PROVEEDOR:

PRIMERA{**V#**, **CIUDAD**, **STATUS**, **P#**, **CANT**}

Primary Key{**V#**, **P#**}

MEDIDA{**P#**, **DIMENSIÓN**}

Primary Key{**P#**, **DIMENSIÓN**}

Foreign Key{**P#**} References PRIMERA

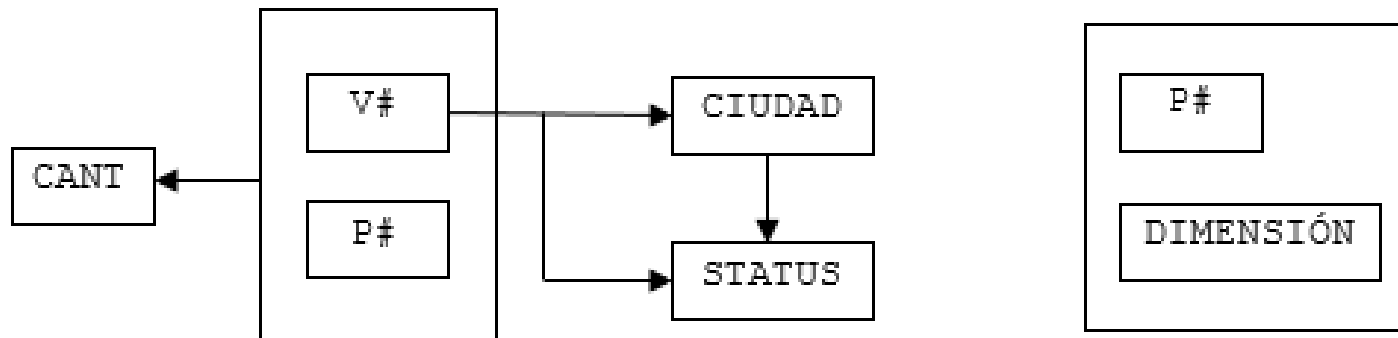


Figura 33: Normalización en Primera Forma Normal (1FN)

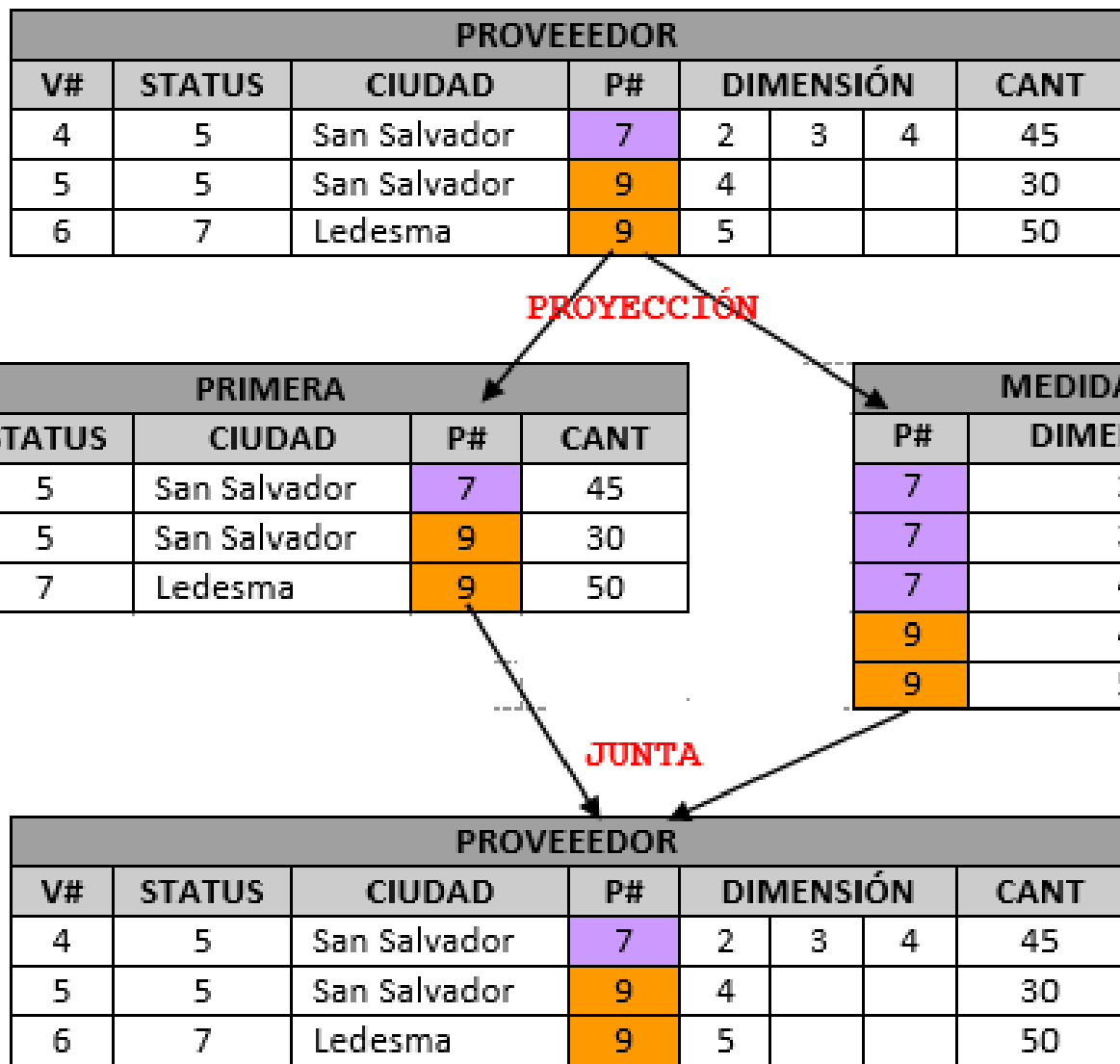


Figura 34: Normalización en Primera Forma Normal (1FN)

La relación MEDIDA no presenta problemas al estar en 2FN y en 3FN.

La relación PRIMERA los atributos STATUS y CIUDAD no presentan una dependencia funcional total de la clave primaria {V#, P#}.

PRIMERA {V#, STATUS, CIUDAD, P#, CANT}

Primary Key{V#, P#}

DF: CIUDAD → STATUS

Figura 35: Normalización en Segunda Forma Normal (2FN)

Solución del problema: Proyectar la relación PRIMERA:

SEGUNDA {**V#**, **STATUS**, **CIUDAD**}

Primary Key {**V#**}

VP {**V#**, **P#**, **CANT**}

Primary Key {**V#**, **P#**}

Foreign Key {**V#**} References SEGUNDA

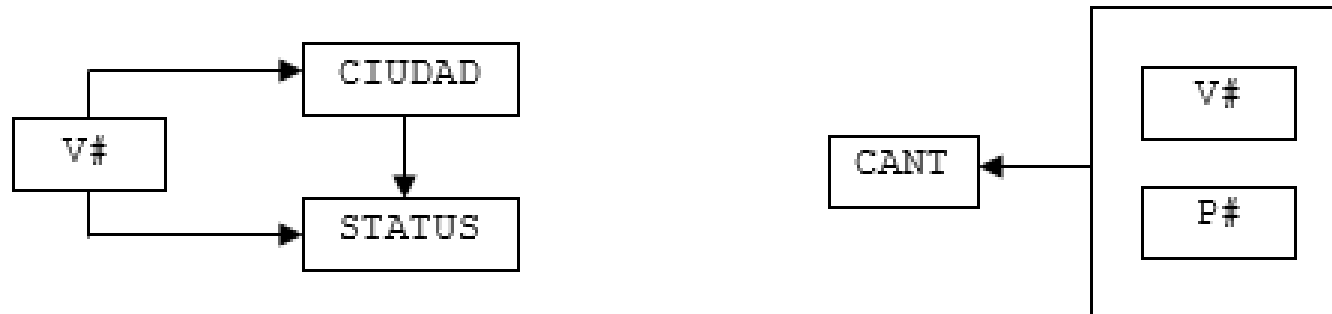


Figura 36: Normalización en Segunda Forma Normal (2FN)

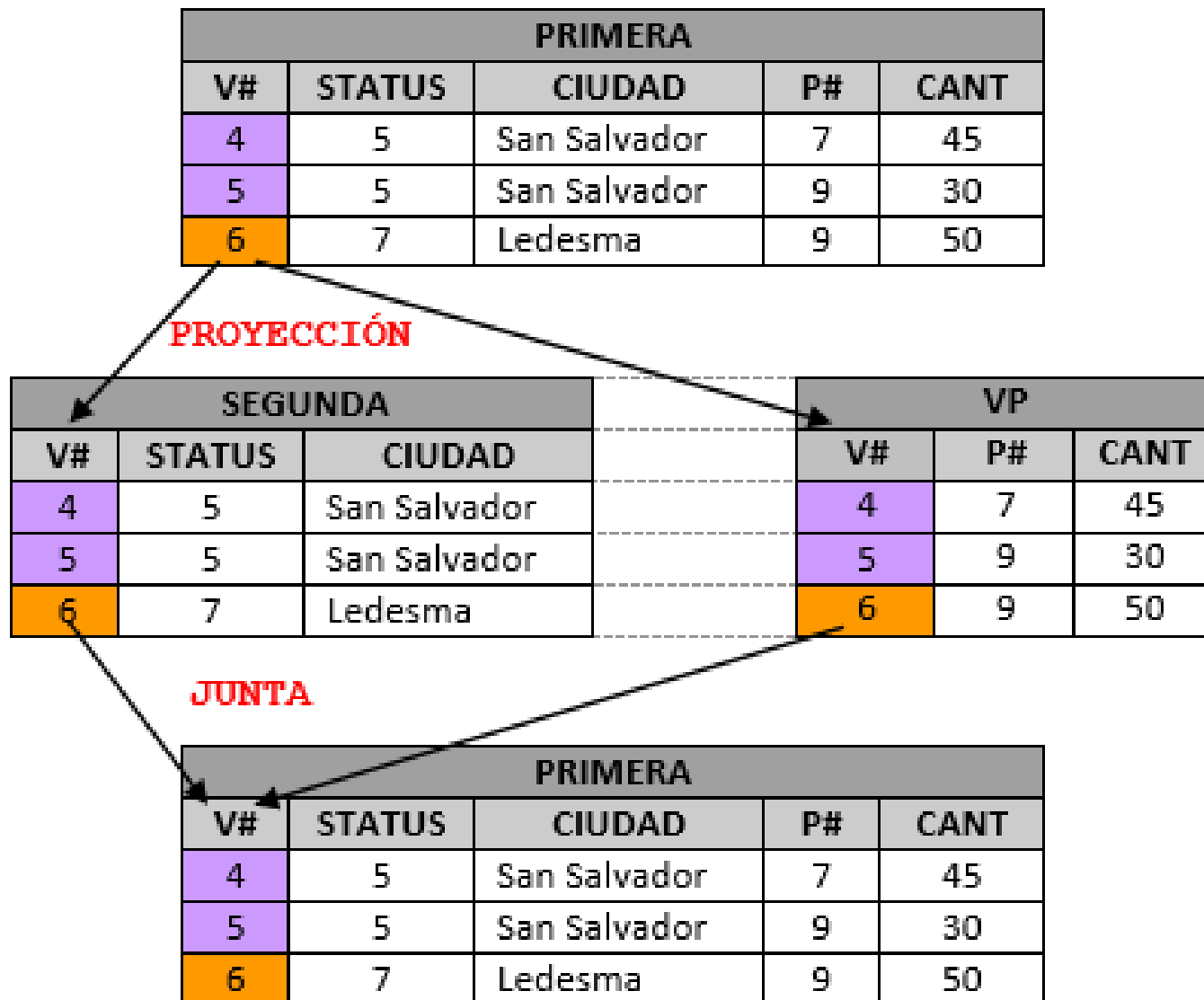


Figura 37: Normalización en Segunda Forma Normal (2FN)

La estructura de SEGUNDA todavía ocasiona problemas, por la falta de independencia mutua entre sus atributos no clave. La dependencia funcional:

V# → **STATUS**

es transitiva a través del atributo **CIUDAD**:

V# → **CIUDAD** ^ **CIUDAD** → **STATUS**

Figura 38: Normalización en Tercera Forma Normal (3FN)

Solución del problema: Proyectar la relación SEGUNDA:

CS {CIUDAD, STATUS}

Primary Key {CIUDAD}

VC {V#, CIUDAD}

Primary Key {V#}

Foreign Key {CIUDAD} References CS

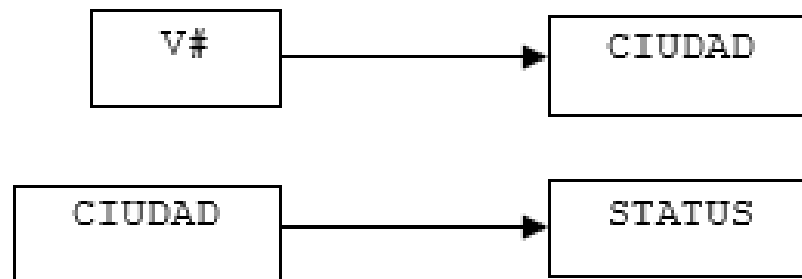


Figura 39: Normalización en Tercera Forma Normal (3FN)

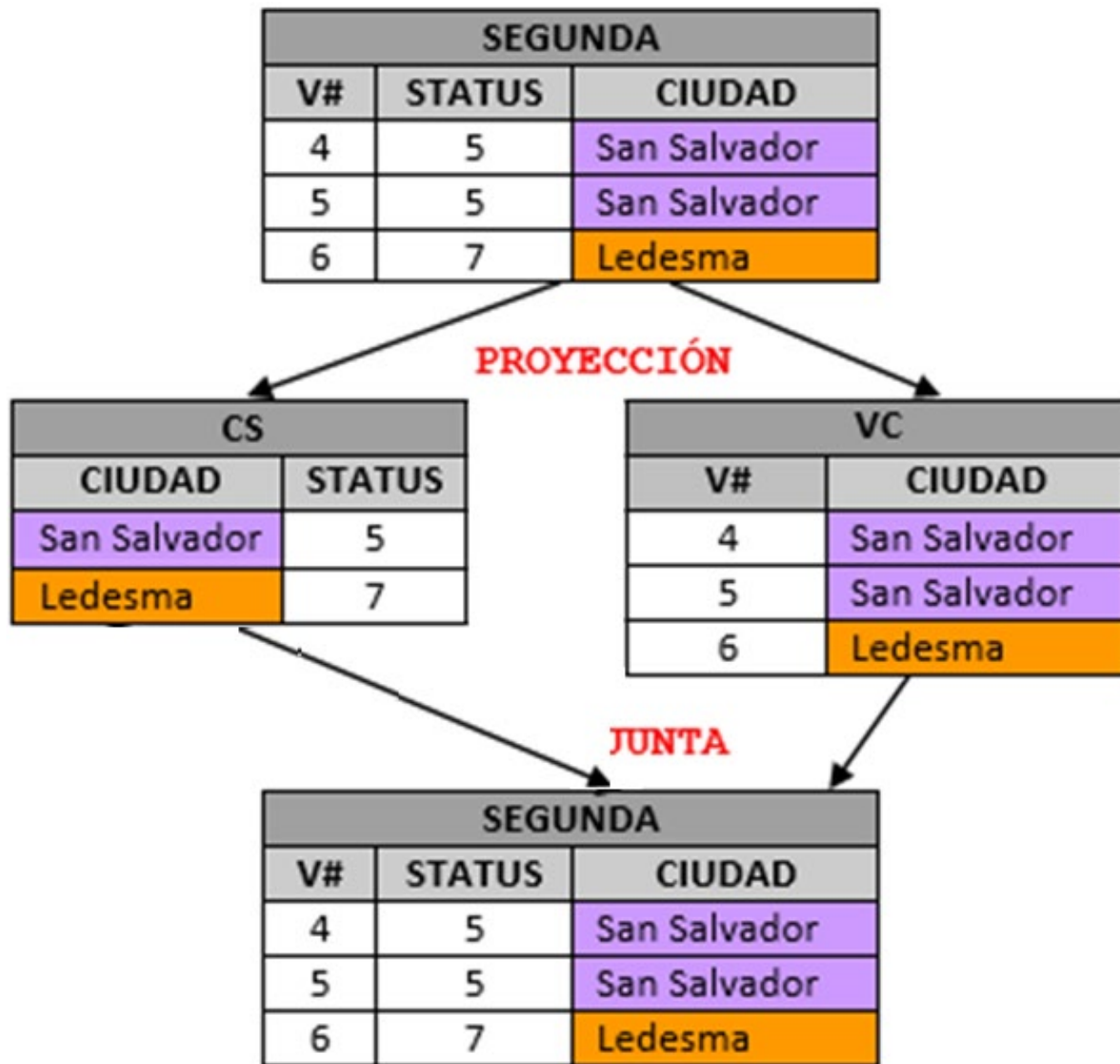


Figura 40: Normalización en Tercera Forma Normal (3FN)

Del esquema de relación original:

PROVEEDOR {V#, CIUDAD, STATUS, P#, DIMENSIÓN, CANT}

Primary Key{V#, P#}

DF: CIUDAD → STATUS

que presentaba el problema del atributo multivaluado

DIMENSIÓN, se fueron realizando sucesivas Proyecciones

hasta obtener como resultado los esquemas de relación:

Figura 41: Resumen

MEDIDA{P#, DIMENSIÓN}

Primary Key{P#, DIMENSIÓN}

Foreign Key{P#} references VP

VP{V#, P#, CANT}

Primary Key{V#, P#}

Foreign Key{V#} references VC

VC{V#, CIUDAD}

Primary Key{V#}

Foreign Key{CIUDAD} references CS

CS{CIUDAD, STATUS}

Primary Key{CIUDAD}

Figura 42: Resumen

Las relaciones entre las tablas se muestran en el siguiente diagrama:

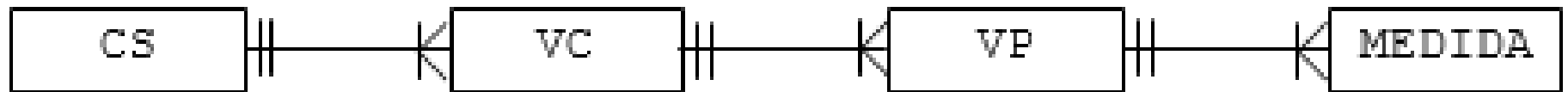


Figura 43: Resumen