



Creación y Gestión eficiente de bases de datos relacionales.

Del Camino del diseño a la implementación.



Lic. Darío Oscar Villegas
Prof. Bases de Datos I | FaCENA - UNNE



Contenido

- Bases de datos relacionales
- **Diseño de bases de datos**
- SQL. Búsqueda y recuperación de datos
- SQL. Presentación de datos (vistas) y búsquedas avanzadas.

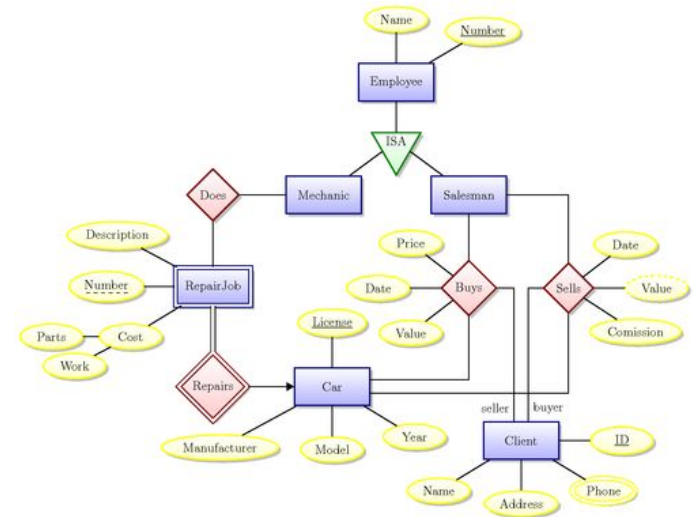
Introducción al diseño de BD

Modelos de Datos

Un **modelo de datos** es una definición **lógica, independiente y abstracta** de los objetos, operadores y demás elementos que en conjunto constituyen la máquina abstracta con la que interactúan los usuarios.

Los objetos nos permiten modelar la estructura de los datos. Los operadores nos permiten modelar su comportamiento.

Un **modelo de datos** es un conjunto de estructuras descriptivas de datos de **alto nivel** que oculta muchos detalles de almacenamiento de bajo nivel.





Modelo de Datos

Una base de datos es una **descripción de una colección particular de datos**, utilizando un modelo de datos determinado.

La mayor parte de los sistemas actuales de gestión de bases de datos se basan en el **modelo relacional de datos**.

Los **modelos semánticos de datos** son modelos de datos de alto nivel más abstractos que facilitan que los usuarios obtengan una buena descripción inicial de los datos de las empresas.

Un modelo semántico de datos muy utilizado, denominado **modelo entidad-relación (ER)** nos permitirá denotar de manera gráfica las entidades y las relaciones existentes entre ellas.

Modelo E-R

El **modelo de datos entidad-relación (ER)** permite describir los datos implicados en empresas/organizaciones reales en términos de **objetos y de sus relaciones**, y se emplea para desarrollar el diseño preliminar de las bases de datos.

El modelo de datos de entidad-relación (ER) se basa en una **percepción de un mundo real** que consiste en un conjunto de objetos básicos llamados **entidades** y de relaciones entre estos objetos.





Etapas del diseño

1- Análisis de requisitos

El primer paso del diseño de aplicaciones de bases de datos es **comprender los datos** que se deben guardar en la base de datos, las **aplicaciones** que se deben construir sobre ellos y las **operaciones** que son más frecuentes e imponen requisitos de rendimiento.

2- Diseño conceptual de la BD

La información reunida en el análisis de requisitos se emplea para desarrollar una descripción de alto nivel de los datos que se van a guardar en la base de datos, junto con las restricciones que se sabe que se impondrán sobre esos datos. El **modelo ER** es uno de los modelos de datos de alto nivel, o semánticos, empleados en el diseño de bases de datos.

3- Diseño lógico de la base de datos

Hay que escoger un SGBD que implemente nuestro diseño de la base de datos y transformar el diseño conceptual de la base de datos en un esquema de base de datos del modelo de datos del SGBD elegido. Sólo se considerarán SGBD relacionales.



Etapas del diseño

4- Refinamiento de los esquemas

Es el análisis del conjunto de relaciones del esquema relacional de la base de datos para identificar posibles problemas y refinarlo. El refinamiento del esquema se puede guiar por la teoría de **normalización** de relaciones.

5- Diseño físico de la BD

Se toman en consideración las cargas de trabajo típicas esperadas que deberá soportar la base de datos y se refinará aún más el diseño de la base de datos para garantizar que cumple los criterios de rendimiento deseados.

6- Diseño de aplicaciones y de la seguridad

Se toman en consideración los aspectos de la aplicación. Determinar claramente las entidades y procesos relacionados. Los roles de usuarios, y el acceso o restricciones a los datos.

Diagrama conceptual del modelo E-R

Una **entidad** es un objeto del mundo real que puede distinguirse de otros objetos

Cada entidad se describe empleando un conjunto de **atributos**. La selección de los atributos refleja el nivel de detalle con el que se desea representar la información relativa a las entidades. Para cada atributo asociado con un conjunto de entidades hay que identificar el **dominio** de valores posibles.

Una **clave** es un conjunto mínimo de atributos cuyos valores identifican de manera unívoca a cada entidad del conjunto. Puede haber más de una **clave candidata**; en ese caso, se escogerá una de ellas como **clave principal**.

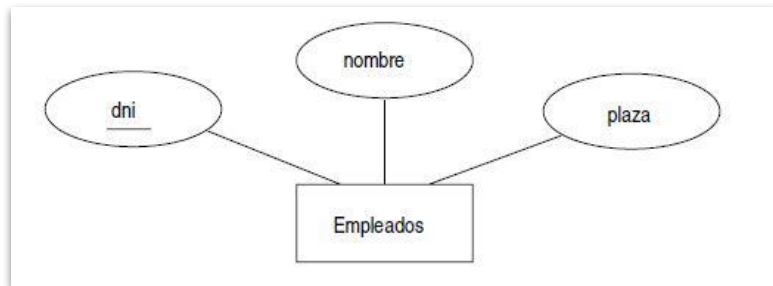
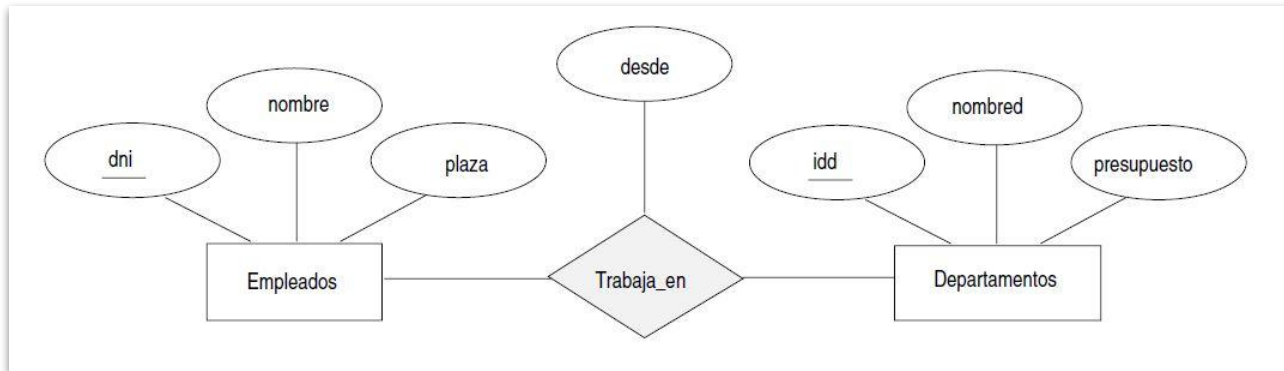



Diagrama conceptual del modelo E-R

Una **relación** es una asociación entre dos o más entidades.

Las relaciones también pueden tener **atributos descriptivos**. Los atributos descriptivos se emplean para registrar la información sobre la relación, más que sobre las entidades participantes.





**Elementos
básicos**
(modelo original)

Entidades




Relaciones entre entidades (interrelaciones)



Atributos





**Otros
elementos**
(añadidos para
mejorar la
expresividad)

Jerarquías de generalización



Atributos compuestos



Identificadores





Otras características del modelo E-R

Cardinalidad de mapeo

- **Uno a Uno:** Una ocurrencia de la entidad A está asociada con a lo sumo una ocurrencia de la entidad B y una ocurrencia de la entidad B está asociada con a lo sumo una ocurrencia de la entidad A. Ej.: el DNI con el Legajo de un empleado Bancario.-
- **Uno a Muchos:** Una ocurrencia de la entidad A está asociada con cualquier número de ocurrencias de la entidad B, pero toda ocurrencia de una entidad B solo puede estar asociada con una ocurrencia de la entidad A. Un País tiene muchas provincias, pero la provincia pertenece a un solo País.-
- **Muchos a Muchos:** Cualquier ocurrencia de la entidad A, está asociada con varias ocurrencias de la entidad B, y cualquier ocurrencia de la entidad B está asociada con varias ocurrencias de la entidad A. Ej.: Un alumno cursa varias materias y una materia tiene varios alumnos cursantes.-

El modelo relacional

Introducción al modelo relacional

La entidad es una **tabla**. El esquema de la entidad describe las cabeceras de las columnas de esa tabla.

El esquema especifica el nombre de la relación, el de cada **campo**, y el dominio de cada campo.

En la entidad se hace referencia al **dominio** por su nombre de dominio y tiene un **conjunto de valores** asociados.

Alumnos (ide: string, nombre: string, usuario: string, edad: integer, nota: real)

Nombres de campos

CAMPOS (ATRIBUTOS, COLUMNAS)

ide	nombre	usuario	edad	nota
50000	David	david@inf	19	6,6
53666	Jiménez	jimenez@inf	18	6,8
53688	Sánchez	sanchez@ii	18	6,4
53650	Sánchez	sanchez@mat	19	7,6
53831	Martínez	martinez@musica	11	3,6
53832	García	garcia@musica	12	4,0

TUPLAS
(REGISTROS,
FILAS)

El **grado**, también denominado aridad, de una relación es su número de campos. La **cardinalidad** de un ejemplar de la relación es el número de tuplas que contiene.



entonces ...

una base de datos relacional es un conjunto de tablas..... (relacionadas entre si)



RESTRICCIONES DE INTEGRIDAD

Una base de datos sólo es tan buena como la información almacenada en ella y, por tanto, el SGBD debe ayudar a evitar la introducción de información incorrecta.

Una **restricción de integridad** (IC) es una condición especificada en el esquema de la base de datos que restringe los datos que pueden almacenarse en los ejemplares de la base de datos.

Las restricciones de **dominio**, de **clave principal** y de **clave externa** se consideran una parte fundamental del modelo relacional de datos y se les presta especial atención en la mayor parte de los sistemas comerciales.

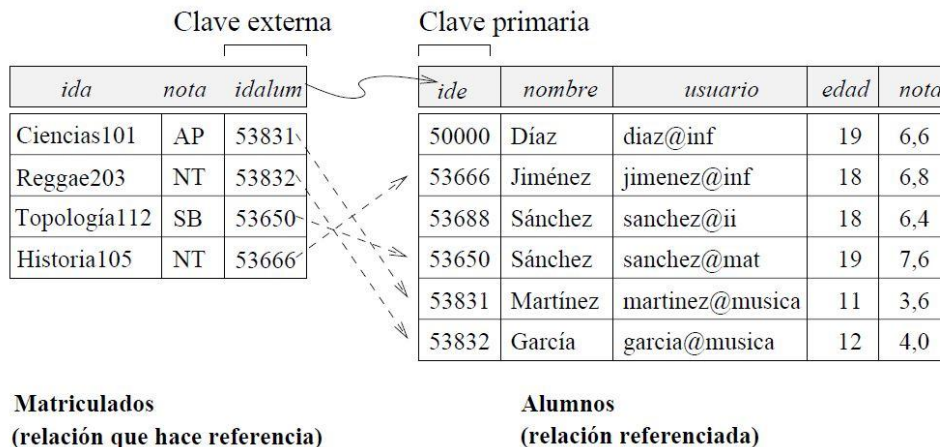
RESTRICCIONES DE INTEGRIDAD

Restricciones de clave

Una restricción de clave es una declaración de que un cierto subconjunto mínimo de los campos de una relación constituye un identificador único de cada fila.

Restricciones de clave externa

La RI que implica a dos relaciones más frecuente es la restricción de clave externa. Si se modifica una de las relaciones hay que comprobar la otra y, quizás, modificarla para hacer que los datos sigan siendo consistentes.



Se pueden aplicar otras **restricciones generales** como ser de unicidad, y control.



Valor “null”

- Se define como **null** el valor que puede tomar un atributo en una tupla cuando se desconoce dicho valor.
- **null** significa que dicho atributo no tiene asociado ningún valor o que “**su valor es desconocido**”

Integridad de la Entidad

Es el mecanismo que gobierna el comportamiento de la clave primaria de una relación, de modo que la representación de entidades en el modelo relacional garantice la identificación o distinción de las mismas en el universo de modelado.

“Ningún atributo que forme parte de la clave primaria de una relación puede tomar un valor nulo”

Diseño lógico de la BD: del modelo E-R al modelo relacional



Objetivo del diseño relacional:

- atomicidad
- coherencia/integridad
- redundancia

Diseño de un Diagrama E-R



Ejemplo

proveedor

articulo

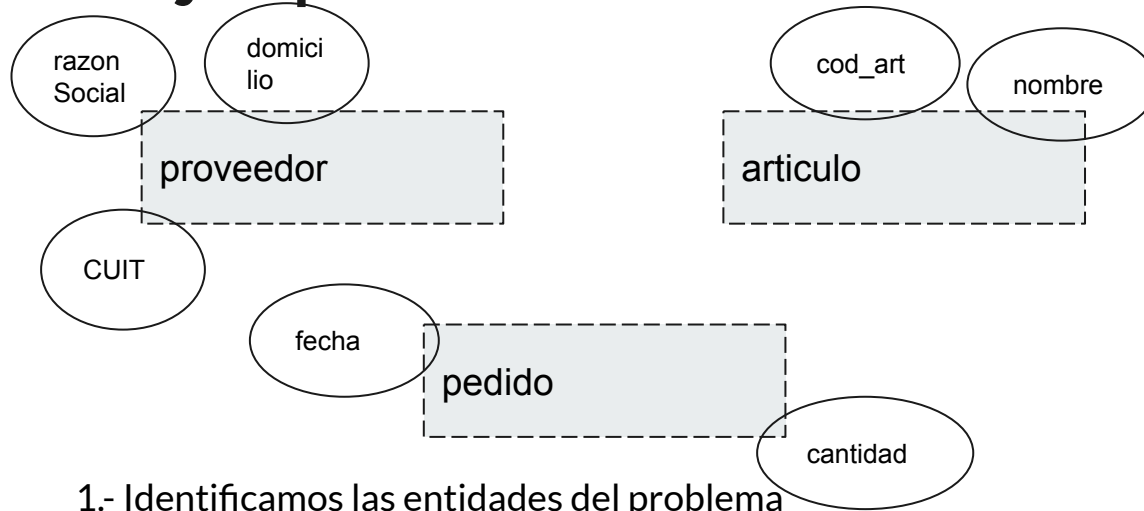
pedido

Caso:

El registro de pedidos que se realizan sobre los artículos que vende un comercio. Cada artículo es provisto por un proveedor.

1.- Identificamos las entidades del problema

Ejemplo

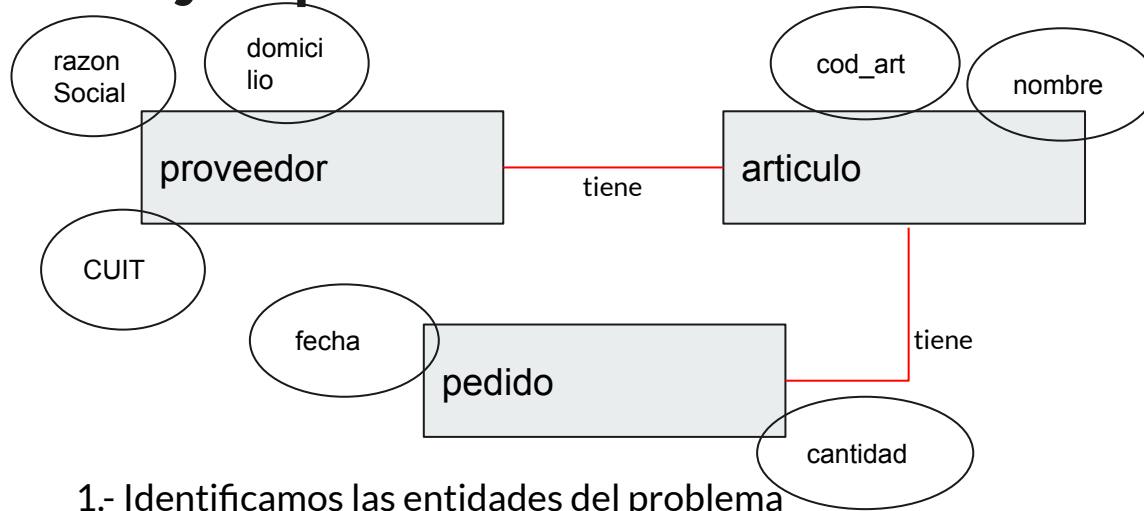


- 1.- Identificamos las entidades del problema
- 2.- Identificamos atributos posibles

Caso:

El registro de pedidos que se realizan sobre los artículos que vende un comercio. Cada artículo es provisto por un proveedor.

Ejemplo

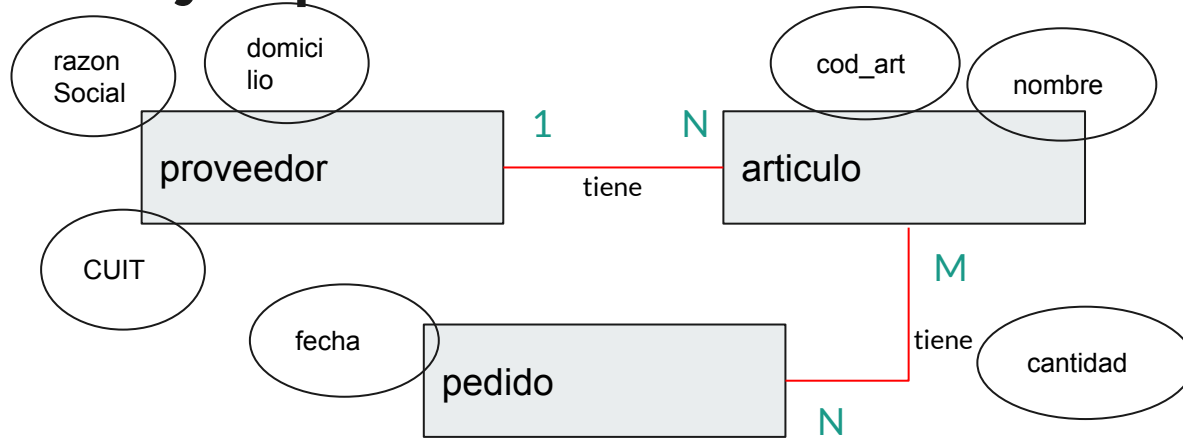


- 1.- Identificamos las entidades del problema
- 2.- Identificamos atributos posibles
- 3.- Identificamos las relaciones entre las entidades

Caso:

El registro de pedidos que se realizan sobre los artículos que vende un comercio. Cada artículo es provisto por un proveedor.

Ejemplo



Caso:

El registro de pedidos que se realizan sobre los artículos que vende un comercio. Cada artículo es provisto por un proveedor.

Cada pedido puede contener varios artículos.

- 1.- Identificamos las entidades del problema
- 2.- Identificamos atributos posibles
- 3.- Identificamos las relaciones entre las entidades
- 4.- Determinamos la cardinalidad de la relación (uno a uno, uno a muchos, mucho a muchos)

Diseño lógico de BD: del modelo E-R al relacional

Ejemplo

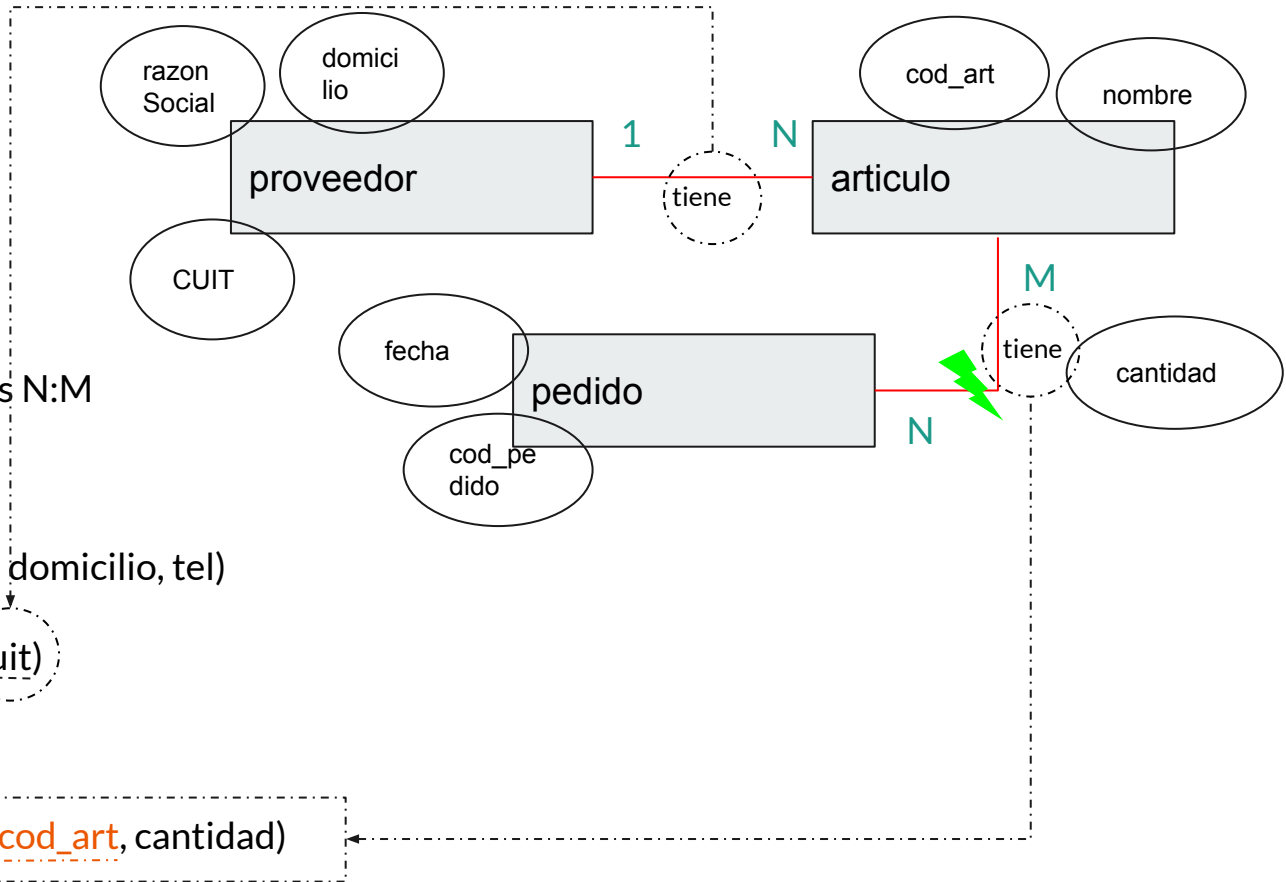
- 1.- Identificar las PK
- 2.- Establecer las FK
- 3.- Resolver las relaciones N:M

proveedor(**cuit**, razonSocial, domicilio, tel)

articulo (**cod_art**, nombre, **cuit**)

pedido (**cod_pedido**, fecha)

pedido_detalle(**cod_pedido**, **cod_art**, cantidad)



Ejemplo

proveedor(**cuit**, razonSocial, domicilio, tel)

Objetivo del diseño relacional:

- atomicidad
- coherencia/integridad
- redundancia

cuit	razonSocial	domicilio	tel
P001	Los Alamos	3 de abril 645, Corrientes	099-99782
P002	Albamonte	Gral San Martin 5435, Sto Tomé	09-951753
P003	La Trinchera	Las Heras 2450, Corrientes	099-12456

analizar →

Ejemplo

tabla → pedido

<u>cod_pedido (PK)</u>	fecha	cliente
100	01/05/2021	0233
101	12/03/2021	0301
102	feb-2021	0302
103	2021	0301

proveedor

campo

<u>cuit (PK)</u>	razonSocial	domicilio	tel
001	Los Alamos	3 de abril 645, Corrientes	099-99782
002	Albamonte	Gral San Martin 5435, Sto Tomé	09-951753
003	La Trinchera	Las Heras 2450, Corrientes	099-12456

dominio

FK

pedido_detalle

N

N FK

<u>cod_pedido (PK,FK)</u>	<u>cod_art (PK,FK)</u>	cantidad
100	A001	1000
100	A002	500
102	A002	2000
103	A004	350

valor

articulo

columna

<u>cod_art (PK)</u>	nombre	cuit (FK)
A001	Arroz	001
A002	Maíz	001
A003	Sorgo	002
A004	Soja	004

registro



Del diseño lógico al diseño físico



Diseño físico

Al igual que todos los demás aspectos del diseño de bases de datos, el diseño físico se debe guiar por la **naturaleza de los datos** y por el **fin al que se destinan**.

Es importante comprender la **carga de trabajo** habitual que debe soportar la base de datos, la carga de trabajo consiste en una mezcla de consultas y de actualizaciones.



Restricciones del modelo de negocio

- Son todos aquellos requerimientos funcionales y no funcionales que se pueden trasladar al modelo de datos, como ser restricción de dominio, de clave foránea, claves únicas, etc. y que no fueron contemplados en la etapa del diseño lógico. (**refinamiento**)

Criterio en el diseño de la base de datos

- Son los aspectos que favorecen la funcionalidad en la gestión de la base de datos. Definición de nuevas claves primarias, de índices, construcción de vistas, políticas de acceso y seguridad, etc

proveedor

<u>id_prov (PK)</u>	cuit	razonSocial	domicilio	tel
---------------------	------	-------------	-----------	-----

UK

articulo

<u>id_articulo (PK)</u>	cod_art	nombre	id_prov (FK)
-------------------------	---------	--------	--------------

UK

INDEX

pedido

<u>cod_pedido (PK)</u>	fecha	cliente
------------------------	-------	---------

FK

pedido_detalle

<u>cod_pedido (PK,FK)</u>	<u>id_articulo (PK,FK)</u>	cantidad
---------------------------	----------------------------	----------

FK


CHECK > 0

Ejemplo

VIEW - Pedido

cod_art - nombre	cantidad	fecha	cliente
------------------	----------	-------	---------

Procesos de normalización



El proceso de normalización de bases de datos consiste en diseñar y aplicar una serie de reglas a las relaciones obtenidas tras el modelo entidad relación.

Las bases de datos se normalizan para:

- Evitar la redundancia de datos.
- Disminuir problemas de actualización de los datos en las tablas.
- Proteger la integridad de los datos.

Primera forma normal

- Una relación está en 1era FN si y sólo si cada atributo es atómico.

Alumnos							
ID	Nombre	Curso	Fecha Matrícula	Tutor	Localidad	Provincia Alumno	Teléfonos
11111111A	Eva	1ESO-A	01-Julio-2016	Isabel	Ecatepec	México	660111222
22222222B	Ana	1ESO-A	09-Julio-2016	Isabel	Ecatepec	México	660222333 660333444 660444555
33333333C	Susana	1ESO-B	11-Julio-2016	Roberto	Ecatepec	México	
44444444D	Juan	2ESO-A	05-Julio-2016	Federico	Aragón	CDMX	
55555555E	José	2ESO-A	02-Julio-2016	Federico	Aragón	CDMX	661000111 661000222

... la solución sería:



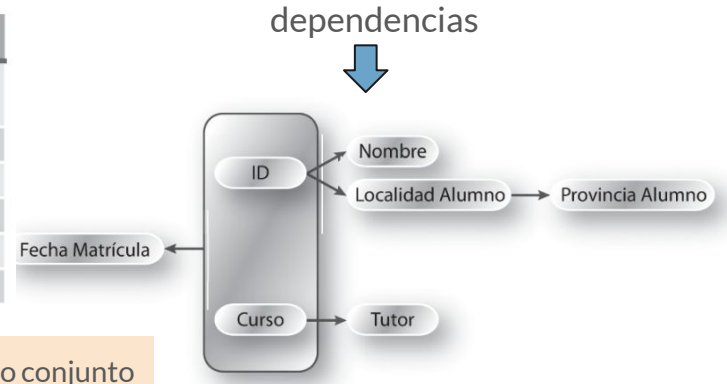
Alumnos						
ID	Nombre	Curso	Fecha Matrícula	Tutor	Localidad	Provincia Alumno
11111111A	Eva	1ESO-A	01-Julio-2016	Isabel	Ecatepec	México
22222222B	Ana	1ESO-A	09-Julio-2016	Isabel	Ecatepec	México
33333333C	Susana	1ESO-B	11-Julio-2016	Roberto	Ecatepec	México
44444444D	Juan	2ESO-A	05-Julio-2016	Federico	Aragón	CDMX
55555555E	José	2ESO-A	02-Julio-2016	Federico	Aragón	CDMX

Teléfonos	
DNI	Teléfono
11111111A	660111222
22222222B	660222333
22222222B	660333444
22222222B	660444555
55555555E	661000111
55555555E	661000222

Segunda forma normal

- Una relación está en 2FN si y sólo si está en 1era FN y todos los atributos que no forman parte de la clave principal tienen dependencia funcional completa de ella. Es decir, no existen referencias parciales.

Alumnos						
ID	Nombre	Curso	Fecha Matriculación	Tutor	Localidad	Provincia Alumno
11111111A	Eva	1ESO-A	01-Julio-2016	Isabel	Ecatepec	México
22222222B	Ana	1ESO-A	09-Julio-2016	Isabel	Ecatepec	México
33333333C	Susana	1ESO-B	11-Julio-2016	Roberto	Ecatepec	México
44444444D	Juan	2ESO-A	05-Julio-2016	Federico	Aragón	CDMX
55555555E	José	2ESO-A	02-Julio-2016	Federico	Aragón	CDMX



A la hora de establecer la clave primaria de una tabla debemos escoger un atributo, o conjunto de ellos, de los que dependan funcionalmente el resto de los atributos.

... la solución sería:



Alumnos			
ID	Nombre	Localidad	Provincia
11111111A	Eva	Ecatepec	México
22222222B	Ana	Ecatepec	México
33333333C	Susana	Ecatepec	México
44444444D	Juan	Aragón	CDMX
55555555E	José	Aragón	CDMX

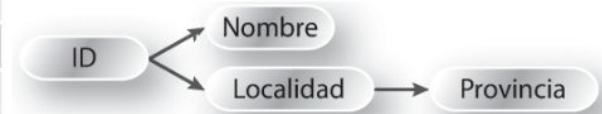
Matrículas		
ID	Curso	FechaMatrícula
11111111A	1ESO-A	01-Julio-2016
22222222B	1ESO-A	09-Julio-2016
33333333C	1ESO-B	11-Julio-2016
44444444D	2ESO-A	05-Julio-2016
55555555E	2ESO-A	02-Julio-2016

Cursos	
Curso	Tutor
1ESO-A	Isabel
1ESO-B	Roberto
2ESO-A	Federico

Tercera forma normal

- Una relación está en 3era FN si y sólo si está en 2da FN y no existen dependencias transitivas. Todas las dependencias funcionales deben darse respecto a la clave principal.

Alumnos			
ID	Nombre	Localidad	Provincia
11111111A	Eva	Ecatepec	México
22222222B	Ana	Ecatepec	México
33333333C	Susana	Ecatepec	México
44444444D	Juan	Aragón	CDMX
55555555E	José	Aragón	CDMX



Podemos observar, existe una dependencia funcional transitiva: ID-> Localidad -> Provincia.

... la solución sería:



Alumnos		
ID	Nombre	Localidad
11111111A	Eva	Ecatepec
22222222B	Ana	Ecatepec
33333333C	Susana	Ecatepec
44444444D	Juan	Aragón
55555555E	José	Aragón

Localidades	
Localidad	Provincia
Ecatepec	México
Aragón	CDMX

Resultado final

Alumnos		
ID	Nombre	Localidad
11111111A	Eva	Ecatepec
22222222B	Ana	Ecatepec
33333333C	Susana	Ecatepec
44444444D	Juan	Aragón
55555555E	José	Aragón

Matrículas		
ID	Curso	Fecha Matrícula
11111111A	1ESO-A	01-Julio-2016
22222222B	1ESO-A	09-Julio-2016
33333333C	1ESO-B	11-Julio-2016
44444444D	2ESO-A	05-Julio-2016
55555555E	2ESO-A	02-Julio-2016

Localidades	
Localidad	Provincia
Ecatepec	México
Aragón	CDMX

Teléfonos	
ID	Teléfono
11111111A	660111222
22222222B	660222333
22222222B	660333444
22222222B	660444555
55555555E	661000111
55555555E	661000222

Cursos	
Curso	Tutor
1ESO-A	Isabel
1ESO-B	Roberto
2ESO-A	Federico

i?

A la acción. Descripción del Caso

Una organización educativa ofrece diversos cursos destinados a personas con perfiles técnicos y estudiantes de carreras en sistemas o afines. Los estudiantes tienen la posibilidad de inscribirse en múltiples cursos y de expresar diferentes intereses de aprendizaje.

- Un estudiante puede tener múltiples Intereses.
- Un estudiante puede inscribirse en múltiples cursos.
- Un curso puede tener múltiples estudiantes inscritos.



Bibliografía

- Presentaciones de clases de Bases de Datos I (FaCENA - UNNE)
- Sistemas de gestión de bases de datos, Tercera Edición. Raghu Ramakrishnan; Johannes Gehrke. McGraw-Hill.
- Introducción a los Sistemas de Bases de Datos – 7ma Edición. Date, C. J.
- Pulido Romero, E. Escobar Domínguez, Ó. y Núñez Pérez, J. Á. (2019). Base de datos: (ed.). Ciudad de México, Grupo Editorial Patria.