

Introducción a la Teoría de Bases de Datos

Contenido

- Concepto
- ¿Por qué usar bases de datos?
- Evolución
- Modelo Relacional
- Formas Normales
- Lenguajes de Gestión de Bases de Datos
- Ms Access

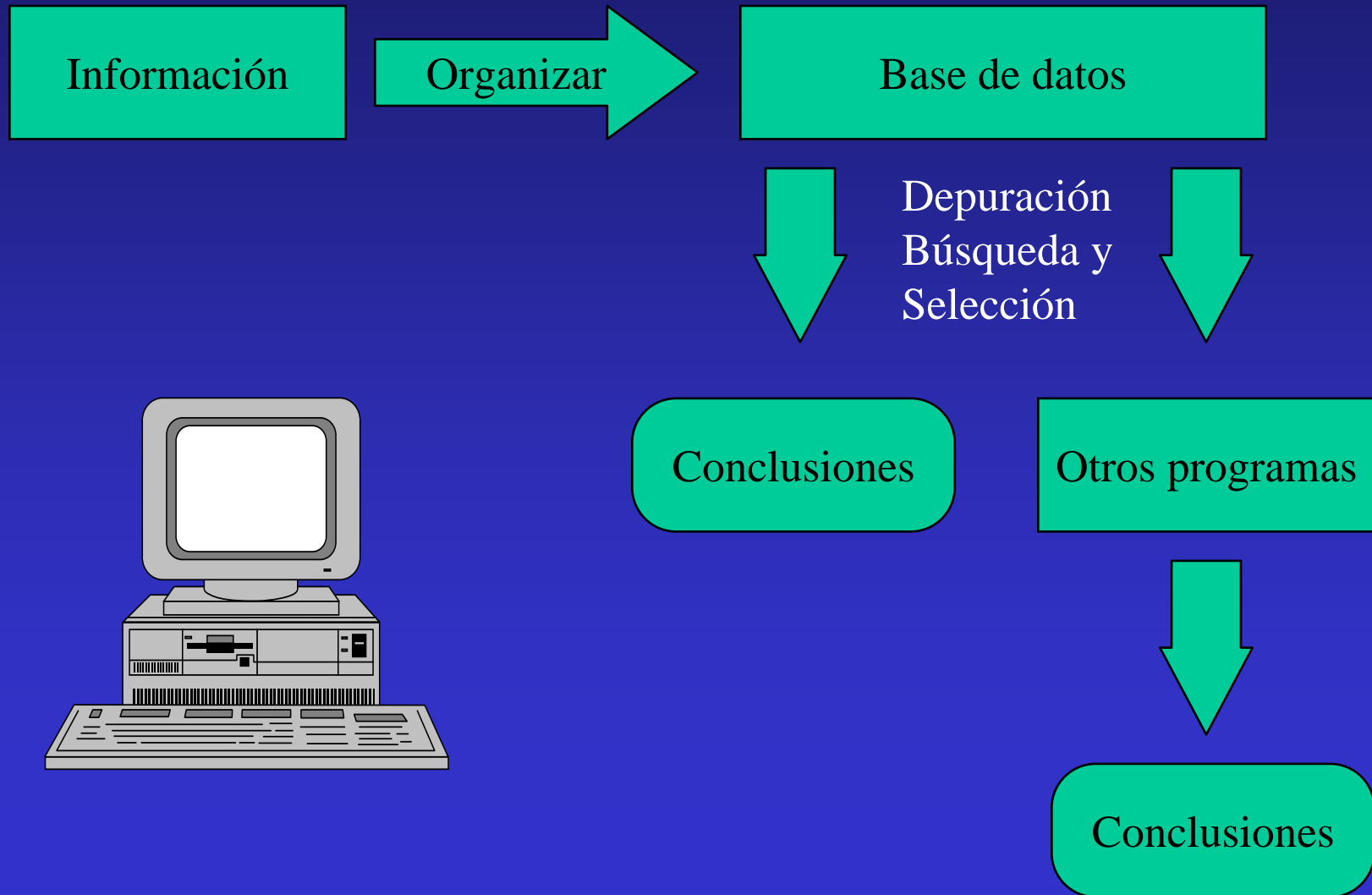
Concepto de Base de Datos

- Recopilación de datos (información) sobre un contexto específico:
 - Ficheros, registros y campos
- Características deseables:
 - Información correcta
 - Integridad
 - Ausencia de redundancias
- La teoría de base de datos intenta
 - Organizar de forma eficiente la información almacenada,
 - Facilitar mecanismos de búsquedas,
 - Evitar errores en la información

Tratamiento de la Información

- Aplicaciones varias
 - Hojas de Cálculo
 - Paquetes estadísticos
 - Procesadores de texto
 - Ficheros de texto
- Bases de datos
 - Organización más eficiente de la información
 - Permite el uso de múltiples tablas
 - Tienen motores de búsqueda y selección muy potentes

Procedimiento



Ejemplos

- Información sobre los empleados de una empresa: nombre y apellidos, DNI, nº de SS, fecha de nacimiento, departamento al que pertenece, categoría, nº de cuenta, domicilio, etc.
- Productos de una empresa: código, descripción, proveedor, precio, stock, etc.
- Proveedores de los productos: código, nombre, dirección, nº cuenta, CIF, etc.
- Pedidos de una empresa: código del producto, código del proveedor, cantidad solicitada, precio de compra, tipo de IVA, fecha pedido, etc.
- Cotizaciones en bolsa: título, fecha, precio inicial. Precio final, precio medio, volumen negociado, etc.

Ejemplos

- Calificaciones de una asignatura: nombre y apellidos del alumno, foto, trabajos obligatorios, trabajos voluntarios, parciales, finales, calificación junio, etc.
- Historial de un paciente: nombre, apellidos, edad, enfermedades, grupos de riesgo, etc.
- Bibliografías
- Recetas
- Base de datos de los DNI españoles
- Expedientes académicos
-

Base de datos de Títulos

Fecha	Titulo	Med	Spread	Max	Min	Ult	Cap_Neg
30/09/91	ACE	1266	0.3432	1275	1260	1265	137511110
30/09/91	ACX	7576	0.6020	7650	7540	7560	216716000
30/09/91	AGI	1225	1.5264	1250	1220	1220	9747165
30/09/91	AGN	4430	2.9468	4430	4430	4430	221500
30/09/91	AGR	1906	0.7458	1935	1890	1915	73571600
30/09/91	AGS	3769	0.5957	3780	3760	3780	33190350
30/09/91	AHV	170	0.8107	173	169	170	4794100
30/09/91	ALB	5287	1.0052	5360	5250	5330	14099850
30/09/91	ALI	5020	0.2081	5020	5020	5020	12550000
.....

- Seleccionar los títulos que hayan tenido una rentabilidad superior al 5% a lo largo del último trimestre de 1991
- Seleccionar las cotizaciones de los títulos por días de la semana

Cajas de un Supermercado

N° Caja	Fecha	Turno	Cod_Emp	Dinero	Tarjeta
01	23/12/96	Mañana	12345	2.547.465	5.678.900
01	23/12/96	Tarde	12346	1.253.852	3.896.900
..
17	10/01/97	Mañana	12345	562.960	675.800

- Resumen de ventas por empleado
- Resumen de ventas por un periodo de tiempo
- Ratio Dinero/Tarjeta a lo largo de los meses
- Resumen de ventas por turnos
- Tabla de ventas por empleado y turno
- Frecuencias de turnos por empleados

Evolución Histórica

- 1945: Primera unidad de cinta magnética
- 1957: Primer computador comercial instalado
- 1961: Primer SGBD
- 1965-1970:
 - Modelo jerárquico
 - Bases de datos multiusuario
- 1970-1980:
 - Desarrollo de varios sistemas comerciales de bases de datos
 - Desarrollo del modelo relacional.
 - Aparecen los primeros lenguajes de consulta.

Evolución Histórica

- 1980-1990:
 - Desarrollo de SGBD para ordenadores personales (dbase, Paradox).
 - SQL
 - ORACLE, SYBASE, INFORMIX, etc.
- 1990- :
 - Extensión de las capacidades: multimedia, GIS, bases de conocimiento.
 - Bases de datos orientadas al objeto
 - Uso de multiproceso
 - Mejora de los SGBD para ordenadores personales

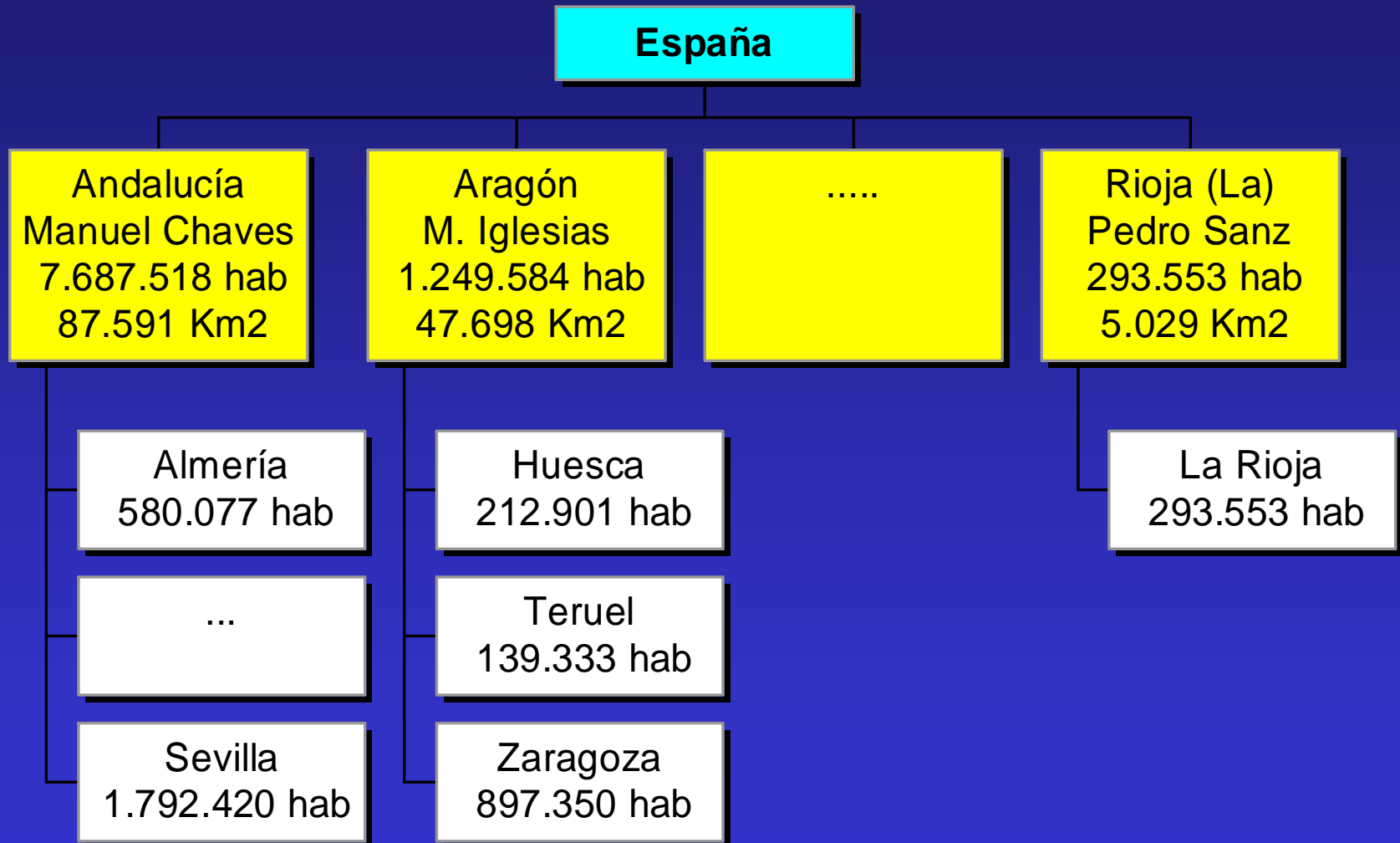
Evolución Histórica

- **Epoca de los FICHEROS**
 - El programador o responsable se debe ocupar de gestionar toda la información: índices, algoritmos de búsquedas, integridad de los datos, etc.
- **SISTEMAS GESTORES de BASES de DATOS**
 - El sistema se encarga de realizar todas las operaciones de mantenimiento físico.
 - El encargado del sistema solo debe diseñar la estructura de la base de datos y suministrar información correcta




Modelos de Bases de Datos

- Qué modelo se usa para representar la información:
- Los más usados a lo largo de la historia son
 - Jerárquico
 - En red
 - Relacional
- El modelo relacional es el que se usa en la actualidad en la mayoría de sistemas gestores de bases de datos

Geografía de España – Modelo Jerárquico






Geografía de España – Modelo Relacional (1)

Comunidad	Presidente	Población	Extensión	Bandera
Andalucía	Manuel Chaves	7.687.518	87.591	
Aragón	M. Iglesias	1.249.584	47.698	
...				
La Rioja	Pedro Sanz	293.553	5.029	

Geografía de España – Modelo Relacional (2)

Provincia	Población	Capital
Almería	580.077	Almería
Huesca	212.901	Huesca
La Rioja	293.553	Logroño
Sevilla	1.792.420	Sevilla
Teruel	139.333	Teruel
Zaragoza	897.350	Zaragoza

Geografía de España – Modelo Relacional (3)

Cod.	Comunidad	Presid.	Población	Ext.	Bandera
1	Andalucía	M. Chaves	7.687.518	87.591	
2	Aragón	M. Iglesias	1.249.584	47.698	
...
17	La Rioja	Pedro Sanz	293.553	5.029	

Provincia	Población	Capital	Comunidad
Almería	580.077	Almería	1
Huesca	212.901	Huesca	2
La Rioja	293.553	Logroño	17
Sevilla	1.792.420	Sevilla	1
Teruel	139.333	Teruel	2
Zaragoza	897.350	Zaragoza	2

Geografía de España – Modelo Relacional (4)

Comunidades
CodComunidad
Comunidad
Poblacion
Extension
Presidente
Bandera

Provincias
CodProvincia
Provincia
Poblacion
Capital
CodComunidad

Modelo Relacional

- Reducción de los datos a estructuras planas o TABLAS con filas y columnas
- TABLAS-> RELACIONES o FICHEROS
- FILAS-> TUPLAS o REGISTROS
- COLUMNAS-> ATRIBUTOS o CAMPOS
- CARDINALIDAD-> N° de REGISTROS
- GRADO-> N° de ATRIBUTOS o CAMPOS

Propiedades

- Las relaciones son tablas en las que:
 - No hay tuplas iguales
 - El orden de las tuplas no es significativo
 - El orden de los atributos no es significativo
 - En una tupla determinada un atributo puede tomar sólo un valor
- El conjunto de posibles valores que puede tomar un atributo se llama DOMINIO. Tienen asociado un FORMATO
- Existe un valor especial: VALOR NULO

Claves

- **CLAVES CANDIDATAS:** atributo o conjunto de atributos que identifican unívoca y mínimamente cada tupla
- Una de ella será la **CLAVE PRIMARIA**. El resto son **CLAVES ALTERNATIVAS**
- **CLAVE AJENA:** Atributo o conjunto de atributos que son **CLAVE PRIMARIA** de otra **RELACION**
- **INDICES:** Atributos o conjunto de atributos sobre los que se pueden realizar “búsquedas rápidas”. La clave primaria debe estar indexada.

Geografía de España - Claves

Comunidades
CodComunidad
Comunidad
Poblacion
Extension
Presidente
Bandera

Provincias
CodProvincia
Provincia
Poblacion
Capital
CodComunidad

Cajas de un Supermercado

Nº Caja	Fecha	Turno	Cod_Emp	Dinero	Tarjeta
01	23/12/96	Mañana	12345	2.547.465	5.678.900
01	23/12/96	Tarde	12346	1.253.852	3.896.900
..
17	10/01/97	Mañana	12345	562.960	675.800

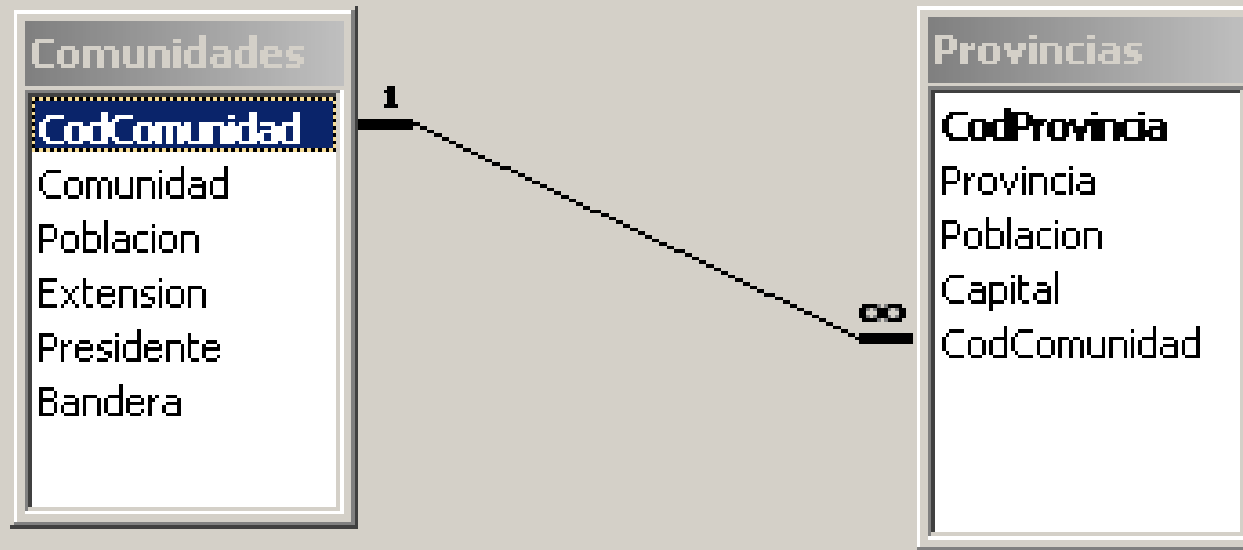
Cajas de un Supermercado

Nº Caja	Fecha	Turno	Cod_Emp	Dinero	Tarjeta
01	23/12/96	Mañana	12345	2.547.465	5.678.900
01	23/12/96	Tarde	12346	1.253.852	3.896.900
..
17	10/01/97	Mañana	12345	562.960	675.800

Integridad

- **INTEGRIDAD DE ENTIDAD:** Ningún atributo que forme parte de una clave primaria puede tomar valor nulo
- **INTEGRIDAD REFERENCIAL:** se establece entre dos relaciones y se usa para mantener la consistencia entre las tuplas de las dos relaciones.
- Una restricción de integridad referencial se introduce al establecer claves ajenas.

Geografía de España – Integridad Referencial



Otros Tipos de Integridad

- Integridad semántica: la información sea real
- Integridad de los datos individuales: que los valores de un atributo pertenezcan al dominio
- Integridad de los datos de un registro: coherencia entre los diferentes campos
- Integridad relativa a transiciones de estados: que una fecha siempre aumente
- Hay restricciones que se pueden dejar de cumplir transitoriamente

Objetivos del Modelo Relacional

- Disponer de una **gran calidad de la información**
- Facilitar mecanismos de recuperación de la información
- Para ello es fundamental definir correctamente las relaciones

Operaciones con los Registros

- **Insertar:** añadir nuevos registros. Puede fallar por:
 - Duplicar claves
 - Dar valor nulo a algún atributo de la clave principal
 - Hacer referencia a una clave que no exista en otra relación
- **Borrar:** Eliminar un registro. Puede producir problemas de integridad referencial. Para eliminarlos:
 - Cancelar, eliminar en cascada o modificar los valores afectados.
- **Modificar:** Cambiar el valor de uno o mas atributos de una tupla de una relación. Puede dar los mismos problemas que insertar y borrar.

Formas Normales

- Una base de datos está en primera forma normal si ninguna tabla contiene registros con grupos repetitivos
- Una base de datos está en segunda forma normal si está en primera forma normal y en todas la tablas los atributos dependen de la clave completa y no solo de una parte de la misma
- Una base de datos está en tercera forma normal si está en segunda forma normal y en todas la tablas los atributos dependen solo de la clave y no de ningún otro atributo

Fichero PEDIDOS	
Campo	Descripción
NUM_PED	Número de pedido
FECHA	Fecha del pedido
PROVEED	Nombre del proveedor
DIR_PROV	Dirección del proveedor
DES_PROD	Descripción del producto
PRECIO	Precio unitario del producto
CANTIDAD	Cantidad solicitada
TOT_PROD	Total producto
TOT_PED	Total pedido

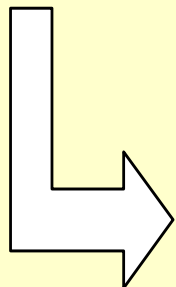
REPETIDOS

NUM_PED	FECHA	PROVEED	DIR_PROV	DES_PROD	PRECIO	CANTIDAD	TOT_PROD	TOT_PED
1	11-02-99	Pedro J.	Fresno s/n	Tornillo	12	200	2400	
				Tuerca	8	200	1600	4000
2	12-11-99	Martín	La Paz	Clavo	6	250	1500	
				Escarpia	9	100	900	
				Tirafondo	10	200	2000	4400
3	23-12-99	Pedro J.	Fresno s/n	Tornillo	12	100	1200	1200

Fichero PEDIDOS	
Campo	Descripción
NUM_PED	Número de pedido
FECHA	Fecha del pedido
PROVEED	Nombre del proveedor
DIR_PROV	Dirección del proveedor
TOT_PED	Total pedido

Fichero LINEAS	
Campo	Descripción
NUM_PED	Número de pedido
DES_PROD	Descripción del producto
PRECIO	Precio unitario del producto
CANTIDAD	Cantidad solicitada
TOT_PROD	Total producto

NUM_PED	FECHA	PROVEED	DIR_PROV	TOT_PED
1	11-02-99	Pedro J.	Fresno s/n	4000
2	12-11-99	Martín	La Paz	4400
3	23-12-99	Pedro J.	Fresno s/n	1200



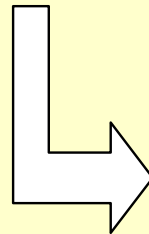
NUM_PED	DES_PROD	PRECIO	CANTIDAD	TOT_PROD
1	Tornillo	12	200	2400
1	Tuerca	8	200	1600
2	Clavo	6	250	1500
2	Escarpia	9	100	900
2	Tirafondo	10	200	2000
3	Tornillo	12	100	1200

Fichero PEDIDOS	
Campo	Descripción
NUM_PED	Número de pedido
FECHA	Fecha del pedido
PROVEED	Nombre del proveedor
DIR_PROV	Dirección del proveedor
TOT_PED	Total pedido

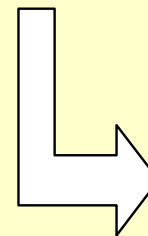
Fichero LINEAS	
Campo	Descripción
NUM_PED	Número de pedido
COD_PROD	Código del producto
CANTIDAD	Cantidad solicitada
TOT_PROD	Total producto

Fichero PRODUCTOS	
Campo	Descripción
COD_PROD	Código del producto
DES_PROD	Descripción del producto
PRECIO	Precio unitario del producto

NUM_PED	FECHA	PROVEED	DIR_PROV	TOT_PED
1	11-02-99	Pedro J.	Fresno s/n	4000
2	12-11-99	Martín	La Paz	4400
3	23-12-99	Pedro J.	Fresno s/n	1200



NUM_PED	COD_PROD	CANTIDAD	TOT_PROD
1	1	200	2400
1	2	200	1600
2	3	250	1500
2	4	100	900
2	5	200	2000
3	1	100	1200



COD_PROD	DES_PROD	PRECIO
1	Tornillo	12
2	Tuerca	8
3	Clavo	6
4	Escarpia	9
5	Tirafondo	10

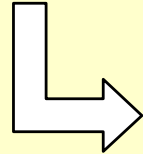
Fichero PEDIDOS	
Campo	Descripción
NUM_PED	Número de pedido
FECHA	Fecha del pedido
COD_PROV	Código del proveedor
TOT_PED	Total pedido

Fichero PRODUCTOS	
Campo	Descripción
COD_PROD	Código del producto
DES_PROD	Descripción del producto
PRECIO	Precio unitario del producto

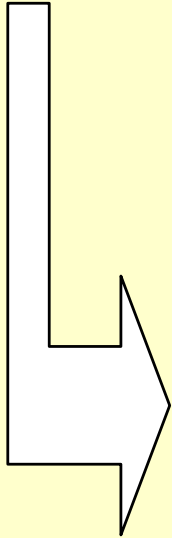
Fichero LINEAS	
Campo	Descripción
NUM_PED	Número de pedido
COD_PROD	Código del producto
CANTIDAD	Cantidad solicitada
TOT_PROD	Total producto

Fichero PROVEEDORES	
Campo	Descripción
COD_PROV	Código del proveedor
PROVEED	Nombre del proveedor
DIR_PROV	Dirección del proveedor

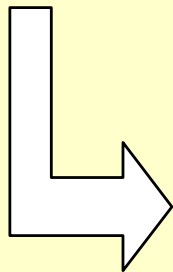
NUM_PED	FECHA	PROVEED	TOT_PED
1	11-02-99	1	4000
2	12-11-99	2	4400
3	23-12-99	1	1200



COD_PROV	PROVEED	DIR_PROV
1	Pedro J.	Fresno s/n
2	Martín	La Paz



NUM_PED	COD_PROD	CANTIDAD	TOT_PROD
1	1	200	2400
1	2	200	1600
2	3	250	1500
2	4	100	900
2	5	200	2000
3	1	100	1200



COD_PROD	DES_PROD	PRECIO
1	Tornillo	12
2	Tuerca	8
3	Clavo	6
4	Escarpia	9
5	Tirafondo	10

SQL

- Structured Query Language
- Estandar en Bases de Datos. Lo soportan casi todos los SGBD
- Ejemplos:
 - `SELECT * FROM EMPLEADOS WHERE NOMBRE="PEPE"`
 - `SELECT nombre, dni, ventas FROM empleados WHERE (ventas>1.000.000) ORDER BY nombre`
 - `SELECT COTIZ.Fecha, TITULOS.Nombre, COTIZ.Pre_Ult FROM COTIZ INNER JOIN TITULOS ON COTIZ.Titulo = TITULOS.Titulo WHERE (((COTIZ.Fecha) Between #12/1/91# And #12/31/91#)) ORDER BY COTIZ.Fecha;`

Multiusuario

- Los SGBD actuales permiten el acceso simultáneo por parte de varios usuarios.
- El problema puede surgir porque dos o mas personas intenten acceder al mismo tiempo a un mismo registro
- El SGBD se preocupa de resolver los conflictos
- En el contexto multiusuario podemos tener
 - Bases de datos centralizadas
 - Bases de datos distribuidas
 - Bases de datos replicadas

Microsoft Access

- Es un SGBD relacional
- Permite todas las funcionalidades que se han comentado en el modelo relacional
- Además de poder gestionar toda la base de datos permite
 - Elaboración de informes
 - Elaboración de formularios para la introducción de datos
 - Creación visual de consultas. SQL
 - Uso de macros
 - Módulos programables (BASIC)