



Seminario de Informática

Unidad V: Elementos conceptuales de Base de Datos.

INTRODUCCION

Aquí solo veremos nociones básicas de Base de Datos y de las Bases de Datos Relacionales puesto que se trata de una asignatura transversal en la formación de las carreras de Ciencias Económicas.

Una base de datos es un sistema informático a modo de almacén, que permite recopilar y organizar información. En este almacén se guardan grandes volúmenes de información sobre personas, productos, pedidos o cualquier otra cosa. Por ejemplo, imaginemos que somos una compañía telefónica y deseamos tener almacenados los datos personales y los números de teléfono de todos nuestros clientes, que posiblemente sean millones de personas.

Esta información es de gran volumen de tamaño: estamos hablando de veinte o treinta datos multiplicados por miles o millones de personas.

La antigua gestión de datos se basaba en archivos informáticos, pero para las necesidades de hoy en día, hacen falta sistemas más perfeccionados que son precisamente lo que se denomina sistema de base de datos.

Muchas bases de datos empiezan siendo una lista en un programa de procesamiento de texto o en una hoja de cálculo. A medida que crece la lista, empiezan a aparecer repeticiones e inconsistencias en los datos. Cada vez resulta más complicado comprender los datos presentados en la lista y existen pocos métodos para buscar o recuperar subconjuntos de datos para revisarlos. Cuando empiezan a observarse estos problemas, es aconsejable transferir la información a una base de datos creada mediante un sistema de administración de bases de datos (DBMS).

Llegamos así a la conclusión de que necesitaríamos una base de datos para automatizar el acceso a la información y poder acceder a ella de manera rápida y fácil, además de poder realizar cambios de una manera más eficiente.

Toda base de datos debe tener una serie de características tales como seguridad (sólo personas autorizadas podrán acceder a la información y realizar la operaciones según su perfil), integridad (la información se mantendrá sin pérdidas de datos), e independencia (esta característica es fundamental ya que una buena base de datos debería ser independiente del sistema operativo o programas que interactúen con ella). Hay más características que debe reunir una base de datos como ser consistente (es decir, que la información se guarde sin duplicidades y de manera correcta).

Finalmente, las bases de datos actuales permiten el manejo correcto de transacciones. Esto significa que se ha de permitir efectuar varias operaciones sobre la base de datos pero tratadas a modo de una sola. Es decir, si en el conjunto de las operaciones de una transacción se produce un error, entonces se deshacen todas las operaciones realizadas anteriormente y



Seminario de Informática

se cancela la transacción. Si pensamos en una transferencia bancaria que comprende varias operaciones como sacar dinero de una cuenta, anotarlo en el extracto de esta cuenta, añadirlo a otra cuenta y anotarlo en el extracto de esta otra cuenta. Supongamos ahora que después de sacar el dinero de la cuenta de origen se produce un error: la transacción (proceso completo) no se ha completado, y en este caso la base de datos revierte lo que se haya hecho y deja la situación tal y como estaba antes de comenzar la transacción.

Otro factor importante en las bases de datos es el tiempo de respuesta, que evidentemente debe ser lo más rápido posible en devolver o registrar las informaciones. Pensemos que una base de datos puede tener que estar enviando y registrando información correspondiente a decenas de personas conectadas a internet. Si la velocidad de respuesta no es buena, la página se quedaría “bloqueada”. Un ejemplo claro es cuando en el mes de febrero ustedes se reinscriben como alumnos, hay muchas computadoras conectadas a través de internet actualizando la información de los estudiantes.

HISTORIA

Los inicios de las bases de datos modernas se deben sobre todo al desarrollo realizado por el inglés Edgar Frank Codd, que propuso el primer modelo teórico relacional, es decir, definió cómo se debían relacionar los datos pertenecientes a una base de datos. Por otro lado IBM desarrolló la primera definición de lenguaje para base de datos (que definía cómo añadir y extraer información de una base de datos) llamado SEQUEL. Más tarde SEQUEL acabaría convirtiéndose en el lenguaje más utilizado hoy en día con bases de datos, el SQL.

SQL se ha convertido en un lenguaje estándar para todas las bases de datos. Hoy en día es empleado por prácticamente todas las bases de datos existentes. Como en todo lenguaje ha habido mejoras y por tanto nuevas versiones a lo largo del tiempo, pero aunque cada sistema de base de datos tenga sus propias particularidades, todos comparten muchas características comunes.

DOS ÁREAS

Dentro de SQL hay dos grandes áreas llamadas DDL y DML. DDL (en inglés Data Definition Language) es la parte del lenguaje que permite la definición de datos, por tanto son funciones que definen cómo van a ser los datos. Por ejemplo nosotros podemos definir que el dato de “edad” de una persona va a ser un número entero mientras que el dato “nombre” va a ser una cadena de texto. Los datos en una base de datos se almacenan en tablas formadas por filas y columnas. Las columnas nos indican el nombre de los datos y las filas contendrán los valores de los datos propiamente dichos.

Un ejemplo de tabla muy simple podría ser la tabla de contactos telefónicos que tenemos en nuestro celular, donde vamos a almacenar la información de su nombre y teléfono. Tendremos por lo tanto dos columnas y tantas filas como contactos existan en nuestra base de datos

Nombre	Teléfono
Nahuel Díaz	154842365



Seminario de Informática

Jesús Córdoba	155958234
Juan Pérez	156364025

Figura 1: Agenda de contactos telefónicos

Por otra parte está el área de DML (en inglés Data Manipulation Language) o Lenguaje de Manipulación de Datos, que como su nombre indica nos permite manipular la información y que básicamente se compone de estas instrucciones o funciones:

- Select: función que permite solicitar que se nos devuelva un dato o serie de datos.
- Insert: función para insertar nuevas filas de información en una tabla.
- Update: función para modificar una o varias filas ya existentes previamente.
- Delete: función que borra una fila o un conjunto de filas de una tabla.

TIPOS DE BASE DE DATOS

Hay diversos tipos de base de datos dependiendo de los objetivos de su uso. Por ejemplo, son distintos objetivos mantener el historial médico de los pacientes de un hospital o el registro de operaciones financieras de un banco. Pero en general los sistemas de bases de datos más populares son las bases de datos relacionales, aunque también se usan otros tipos de bases de datos, entre los que cabe destacar las denominadas “bases de datos orientadas a objetos”.

Las bases de datos relacionales se basan en la idea fundamental del uso de relaciones para definir los tipos de datos o información. Así en nuestro ejemplo anterior podemos ver que un contacto queda definido por un nombre y un teléfono es decir la relación de nombre y teléfono nos da un contacto.

BASE DE DATOS RELACIONAL

Este modelo intenta representar la base de datos como un conjunto de tablas. Aunque las tablas son un concepto simple e intuitivo, existe una correspondencia directa entre el concepto informático de una tabla, y el concepto matemático de relación, lo cual es una gran ventaja, pues permite efectuar formalizaciones de una forma estricta mediante las herramientas matemáticas asociadas, como puede ser el álgebra relacional en el ámbito de las consultas.

Como ya se dijo, no será necesario enfrentarnos con todos estos formalismos propios de los matemáticos, sino que dispondremos de unas herramientas fáciles de manejar que nos permitirán interactuar con la base de datos.

Los conceptos básicos del modelo relacional son:

- Registro: Es algo así como cada ficha de un fichero convencional.
- Tabla: Es un conjunto de fichas de un mismo tipo.

Con estos dos conceptos es posible crear cualquier tipo de datos, y asociarlos entre sí. Por ejemplo, si necesitamos diseñar una base de datos para una agencia de alquiler de autos, necesitaremos una tabla en la que se guarde información sobre los autos, como puede verse en la figura siguiente.



Seminario de Informática

Marca	Modelo	Color	Patente	Situación
Ford	Fiesta	Rojo	BPW-768	Alquilado
Peugeot	307	Gris plata	TBM-723	Disponible
Fiat	Siena	Amarillo	ZXD-438	En taller
Renault	Fluence	Verde	TUF-232	Alquilado
VW	Vento	Negro	HKL-839	Disponible
Ford	Ecosport	Blanco	WRG-395	En taller
Chevrolet	Corsa	Azul	QBS-374	Alquilado

Figura 2: Tabla de autos de una agencia de alquileres

De esta forma, vemos que cada tabla está compuesta por filas, también llamadas tuplas o registros, cada uno de los cuales posee una serie de campos o columnas (marca, modelo, color, patente, situación) en los que se almacenan los datos básicos. El esquema de una tabla nos indica los nombres de cada uno de los campos que contiene, así como el tipo de información que debe contener (números, texto, fecha, etc.).

Una tabla es para nosotros un conjunto de registros; por tanto, los registros no pueden repetirse.

Para poder acceder a un registro concreto, es necesario hacer una consulta a través de algún campo que identifique a dicho registro, como puede ser por ejemplo el número de la patente. A este campo especial que identifica cada registro se le llama clave del registro.

La figura siguiente ilustra una tabla de clientes.

Apellidos	Nombres	Documento	Edad
González Diez	Ana Belén	25836934	27
López Arias	Miguel Ángel	18836746	43
Díaz Aranda	María José	22665358	36
Ruíz Rojo	Juan Pablo	23667228	35

Figura 3: Tabla de clientes de una agencia de alquiler de autos

¿Cómo podemos hacer ahora para indicar qué cliente se hace responsable de cada auto alquilado?. Fácilmente, a través de una nueva tabla que relaciona los clientes con los autos. Para ello, dado que cada registro queda identificado por su clave, nos basta con incluir en esta nueva tabla a las claves de ambas tablas, en lugar de todos sus campos. Así, podemos obtener una nueva tabla de alquileres que contenga la patente del auto, y el documento del cliente, tal como se ve en la figura siguiente.

Patente	Documento
BPW-768	25836934
TUF-232	18836746
TUF-232	22665358
QBS-374	18836746

Figura 4: Tabla de alquileres que relaciona los autos con los clientes

En esta última tabla podemos observar varias cosas interesantes. Por un lado, un cliente se puede responsabilizar de más de un auto, o sea, puede alquilar más de un auto,



Seminario de Informática

vemos así que Miguel Ángel López Arias ha alquilado tanto el Renault Fluence como el Chevrolet Corsa. Pero, a su vez, más de una persona puede hacerse cargo de un auto: Miguel Ángel López Arias y María José Díaz Aranda comparten el alquiler del Renault Fluence.

De esta forma, el modelo relacional soluciona el problema de indicar los alquileres de una manera intuitiva a través de las tablas. Este método de expresar los datos facilita además las consultas, que se realizan ahora a través de estas tablas especiales que relacionan a otras tablas. Por ejemplo, si queremos saber los autos que ha alquilado López Arias Miguel Ángel, basta con buscar su clave en la tabla de clientes (18836746), y a continuación ver que patentes tiene asociadas en la tabla de alquileres (TUF-232 y QBS-374); a continuación, buscamos en la tabla de autos cuales son los autos que poseen esas claves, y obtenemos como resultado: Renault Fluence y Chevrolet Corsa.

INDICES

Un índice en informática es como el índice de un libro donde tenemos los capítulos del libro y la página donde empieza cada capítulo. No vamos a entrar ahora en cómo se implementan los índices internamente sino que vamos a ver unas breves nociones de cómo se definen, para qué sirven y cuándo hay que utilizarlos y cuando no.

Si una tabla tiene definido un índice sobre una columna se puede localizar mucho más rápidamente una fila que tenga un determinado valor en esa columna.

Recuperar las filas de una tabla de forma ordenada por la columna en cuestión también será mucho más rápido.

Al ser el índice una estructura de datos adicional a la tabla, ocupa un poco más de espacio en disco.

Cuando se añaden, modifican o se borran filas de la tabla, el sistema debe actualizar los índices afectados por esos cambios lo que supone un tiempo de proceso mayor.

Por estas razones no es aconsejable definir índices de forma indiscriminada.

Los inconvenientes comentados en este punto no son nada comparados con las ventajas si la columna sobre la cual se define el índice es una columna que se va a utilizar a menudo para buscar u ordenar las filas de la tabla.

Por eso una regla bastante acertada es definir índices sobre columnas que se vayan a utilizar a menudo para recuperar u ordenar las filas de una tabla.

El Sistema Gestor de Base de Datos crea automáticamente índices sobre las columnas claves principales y sobre los campos que intervienen en las relaciones ya que se supone que se utilizan a menudo para recuperar filas concretas.

SISTEMAS GESTORES DE BASES DE DATOS



Seminario de Informática

Con el uso y el incremento del tipo de base de datos se desarrollaron sistemas informáticos que gestionan toda la funcionalidad de la base de datos propiamente dicha intentando que sea de la manera más clara, directa y sencilla.

Normalmente cada compañía de base de datos trae su propio SGBD (Sistema Gestor de Base de Datos).

Algunos ejemplos de Sistemas de Base de Datos son:

- MS Access
- MySQL
- PostgreSQL
- Microsoft SQL Server
- Sybase
- Informix
- Oracle

MICROSOFT ACCESS

Una base de datos en Access es un contenedor de objetos. Una base de datos puede contener más de una tabla. Por ejemplo, un sistema de seguimiento de inventario que utiliza tres tablas no es un conjunto de tres bases de datos, sino una sola base de datos que contiene tres tablas. Excepto si se ha diseñado específicamente para utilizar datos o código de otro origen, una base de datos de Access almacena sus tablas en un solo archivo, junto con otros objetos, como formularios, informes, macros y módulos. Las bases de datos creadas con formato de Access 2007 tienen la extensión de nombre de archivo .accdb y las bases de datos creadas con formatos de versiones anteriores de Access tienen la extensión de nombre de archivo .mdb. Access 2007 se puede utilizar para crear archivos con formatos de versiones anteriores.

Con Access se puede:

- Agregar más datos a una base de datos, por ejemplo, un elemento nuevo en un inventario.
- Modificar datos existentes en la base de datos, por ejemplo, cambiar la ubicación de un elemento.
- Eliminar información, por ejemplo, si se ha vendido o retirado un artículo.
- Organizar y ver los datos de distintas formas.
- Compartir los datos con otros usuarios mediante informes, mensajes de correo electrónico, una intranet o Internet.

Partes de una base de datos de Access

A continuación se describen brevemente los componentes de una base de datos de Access típica:

- Tablas
- Formularios
- Informes
- Consultas



Seminario de Informática

- Macros
- Módulos

Tablas

Una tabla de una base de datos es similar en apariencia a una hoja de cálculo, en cuanto a que los datos se almacenan en filas y columnas. Como consecuencia, normalmente es bastante fácil importar una hoja de cálculo en una tabla de una base de datos. La principal diferencia entre almacenar los datos en una hoja de cálculo y hacerlo en una base de datos es la forma de organizar los datos.

Para lograr la máxima flexibilidad para una base de datos, la información tiene que estar organizada en tablas, para que no haya redundancias. Por ejemplo, si se almacena información sobre empleados, cada empleado se insertará una sola vez en una tabla que se configurará para contener únicamente datos de los empleados. Los datos sobre productos se almacenarán en su propia tabla, y los datos sobre sucursales también tendrán su tabla aparte.

Cada fila de una tabla se denomina registro. En los registros es donde se almacena cada información individual. Cada registro consta de campos (al menos uno). Los campos corresponden a las columnas de la tabla. Por ejemplo, puede trabajar con una tabla denominada "Empleados", en la que cada registro (fila) contiene información sobre un empleado distinto y cada campo (columna) contiene un tipo de información diferente, como el nombre, los apellidos, la dirección, o similares. Los campos se deben configurar con un determinado tipo de datos, ya sea texto, fecha, hora, numérico, o cualquier otro tipo.

Formularios

Los formularios se conocen a veces como "pantallas de entrada de datos". Son las interfaces que se utilizan para trabajar con los datos y, a menudo, contienen botones de comando que ejecutan diversos comandos. Se puede crear una base de datos sin usar formularios, editando los datos de las hojas de las tablas. No obstante, casi todos los usuarios de bases de datos prefieren usar formularios para ver, escribir y editar datos en las tablas.

Los formularios proporcionan un formato fácil de utilizar para trabajar con los datos. Además, se les puede agregar elementos funcionales, como botones de comando. Puede programar los botones para determinar qué datos aparecen en el formulario, abrir otros formularios o informes, o realizar otras tareas diversas. Por ejemplo, podría crear un formulario denominado "Formulario de cliente" para trabajar con datos de clientes. El formulario de cliente podría tener un botón para abrir un formulario de pedido en el que se pudiese escribir un pedido nuevo del cliente.

Los formularios también permiten controlar la manera en que otros usuarios interactúan con los datos de la base de datos. Por ejemplo, puede crear un formulario que muestre únicamente ciertos campos y que permita la ejecución solamente de determinadas operaciones. Así, se favorece la protección de los datos y se facilita la entrada correcta de los mismos.



Seminario de Informática

Informes

Los informes sirven para resumir y presentar los datos de las tablas. Normalmente, un informe responde a una pregunta específica, como "¿Cuánto dinero se ha facturado por cliente este año?" o "¿En qué ciudades están nuestros clientes?" Cada informe se puede diseñar para presentar la información de la mejor manera posible.

Un informe se puede ejecutar en cualquier momento y siempre reflejará los datos actualizados de la base de datos. Los informes suelen tener un formato que permita imprimirlos, pero también se pueden consultar en la pantalla, exportar a otro programa o enviar por correo electrónico.

Consultas

Las consultas son las que verdaderamente hacen el trabajo en una base de datos. Pueden realizar numerosas funciones. Su función más común es recuperar datos específicos de las tablas. Los datos que desea ver suelen estar distribuidos por varias tablas y, gracias a las consultas, puede verlos en una sola hoja de datos. Además, puesto que normalmente no desea ver todos los registros a la vez, las consultas le permiten agregar criterios para "filtrar" los datos hasta obtener solo los registros que desee. Las consultas a menudo sirven de origen de registros para formularios e informes.

Algunas consultas son "actualizables", lo que significa que es posible editar los datos de las tablas base mediante la hoja de datos de la consulta. Si trabaja con una consulta actualizable, recuerde que los cambios se producen también en las tablas, no solo en la hoja de datos de la consulta.

Hay dos tipos básicos de consultas: las de selección y las de acción. Una consulta de selección simplemente recupera los datos y hace que estén disponibles para su uso. Los resultados de la consulta pueden verse en la pantalla, imprimirse o copiarse al portapapeles. O se pueden utilizar como origen de registros para un formulario o un informe.

Una consulta de acción, como su nombre indica, realiza una tarea con los datos. Las consultas de acción pueden servir para crear tablas nuevas, agregar datos a tablas existentes, actualizar datos o eliminar datos.

Macros

Las macros en Access se pueden considerar como un lenguaje de programación simplificado, que se puede utilizar para aumentar la funcionalidad de la base de datos. Por ejemplo, puede adjuntar una macro a un botón de comando en un formulario, de modo que la macro se ejecute cuando se haga clic en el botón. Las macros contienen acciones que realizan tareas, como abrir un informe, ejecutar una consulta o cerrar la base de datos. Casi todas las operaciones de bases de datos que normalmente se realizan manualmente se pueden automatizar mediante macros, ahorrando así mucho tiempo.



Seminario de Informática

Módulos

Los módulos, como las macros, son objetos que sirven para aumentar la funcionalidad de la base de datos. Mientras que las macros en Access se crean seleccionando acciones de una lista, los módulos se escriben en el lenguaje de programación de Visual Basic para Aplicaciones (VBA). Un módulo es una colección de declaraciones, instrucciones y procedimientos que se almacenan conjuntamente como una unidad. Un módulo puede ser de clase o estándar. Los módulos de clase se adjuntan a formularios o informes, y normalmente contienen procedimientos específicos del formulario o el informe al que se adjuntan. Los módulos estándar contienen procedimientos generales que no están asociados a ningún otro objeto. Los módulos estándar se enumeran en Módulos en el panel de exploración, pero los módulos de clase no.