



# Criterios de evaluación

- Primer parcial 20%
- Segundo parcial 20%
- Tercer parcial 20%
- Tareas 15%
- Proyecto final 25% (obligatorio para aprobar la materia)

Proyecto final → en parejas  
- NO COPIAR EN EXÁMENES

## Introducción a Base de Datos

Una base de datos es un almacén que nos permite guardar grandes cantidades de información de forma organizada y relacionada sobre la cual podemos encontrar y utilizar fácilmente la información.

### Características:

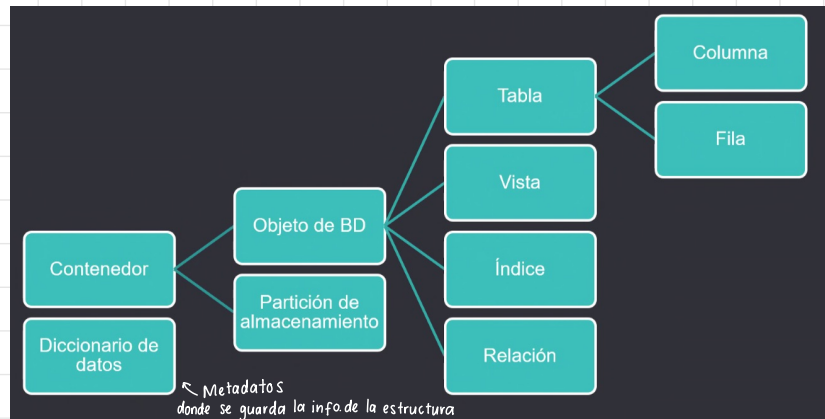
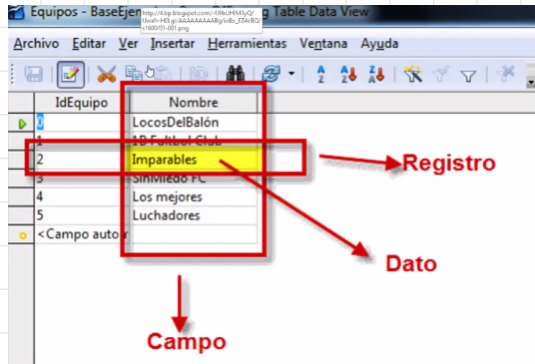
- Independencia lógica y física de los datos
- Acceso concurrente por parte de muchos usuarios
- Integridad de datos → si definimos un campo de núm. que solo acepte eso
- Redundancia mínima
- Consultas optimizadas
- Seguridad de acceso y auditoria  
el auditor puede ver quién entra, cuándo entra, qué checa
- Respaldo y recuperación
- Acceso a través de lenguajes de programación

### ¿Qué conforma una base de datos?

**DATOS:** representación simbólica, un atributo o característica de una entidad

**CAMPOS:** espacio de almacenamiento para un dato en particular

**REGISTROS:** conjunto de campos que contienen los datos que pertenecen a una misma repetición de entidad



### Ventajas:

**CONSISTENCIA:** un dato almacenado se actualiza una sola vez para todos los usuarios, y si esta duplicado, el sistema se quedara consistente

**COMPARTIR DATOS:** los usuarios autorizados pueden compartir la base de datos

**MANTENIMIENTO DE ESTÁNDARES:** estándares de documentación, procedimientos de actualización, reglas de acceso

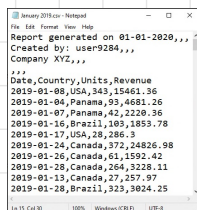
**BUENA INTEGRIDAD DE DATOS:** la validez y la consistencia de los datos almacenados se regulariza por reglas que no se pueden violar

**BUENA SEGURIDAD:** protección fuerte frente a usuarios no autorizados

**MEJORA A SERVICIOS DE COPIAS DE SEGURIDAD:** se minimiza la cantidad de trabajo perdido cuando se produce un fallo

**Base de datos plana (CSV):** se abren en cualquier editor de texto: hay dos dimensiones (estructura plana) de conjunto de datos. Hay una columna de información y dentro de esta columna, se supone que cada dato tendra que ver con la columna. **Columna con códigos postales**

↑  
chechar archivo en Moodle



agosto  
12

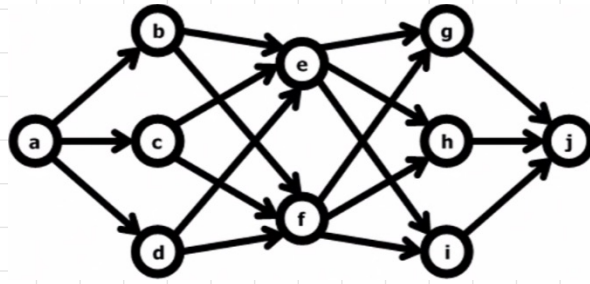
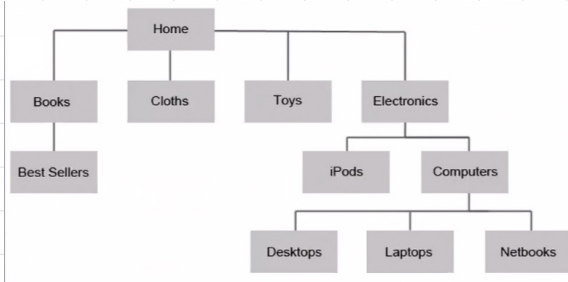
### Base de datos jerárquica:

- Su estructura se asemeja a un **árbol**. Carpetas, archivos

como buscador  
de windows

- Almacena un registro con un enlace a otros registros (como una red)

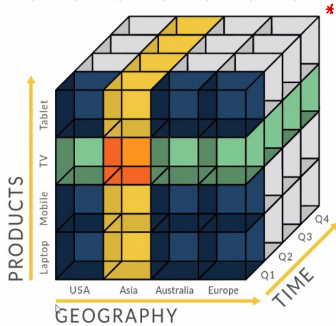
- Puede ser una variedad de diferentes tipos de información como números de nodo de un disco o la dirección



**Base de datos relacional (SQL):** son organizadas en forma de tablas. La belleza de estos cuadros es que la información se puede acceder o añadir sin reorganizar las tablas.

- Una tabla puede tener muchos registros y cada registro puede tener muchos campos.

**Base de datos multidimensionales:** son bases de datos ideadas para desarrollar aplicaciones muy concretas, como creación de Cubos OLAP\*. No se diferencian demasiado de las bases de datos relacionales, la diferencia está más bien a nivel conceptual; en las bases de datos multidimensionales los campos o atributos de una tabla pueden ser de dos tipos, o bien representan dimensiones de la tabla, o bien representan métricas que se desean estudiar



es como la plana, pero traspuesta

**Base de datos columnares:** es un modelo de datos plano, hay dos dimensiones (estructura plana) de conjunto de datos. Hay una columna de información y dentro de esta columna, se supone que cada dato tendrá que ver con la columna. **Columna con códigos postales**

usan JSON, JavaScript (no SQL), amazon las usa muy de moda

**Base de datos documentales:**

- Estos tipos de datos se basan en el almacenamiento de documentos en lugar de datos estructurados.

- Son buenas para datos no estructurados, como texto abierto de una carta o de un correo electrónico y para datos semi-estructurados como documentos académicos

- Tendrás que fijarte en ellas si estás pensando en análisis de texto de documentos demasiado grandes para las bases de datos convencionales.

parecidos a diccionarios de python

**Base de datos Key-Value**

- Usadas para desarrollo simple y fácil de aplicaciones

- Usadas para trabajar con aplicaciones que se puedan desarrollar rápidamente y donde todas las demás consideraciones son secundarias

archivos de código fuente que abrimos en páginas web

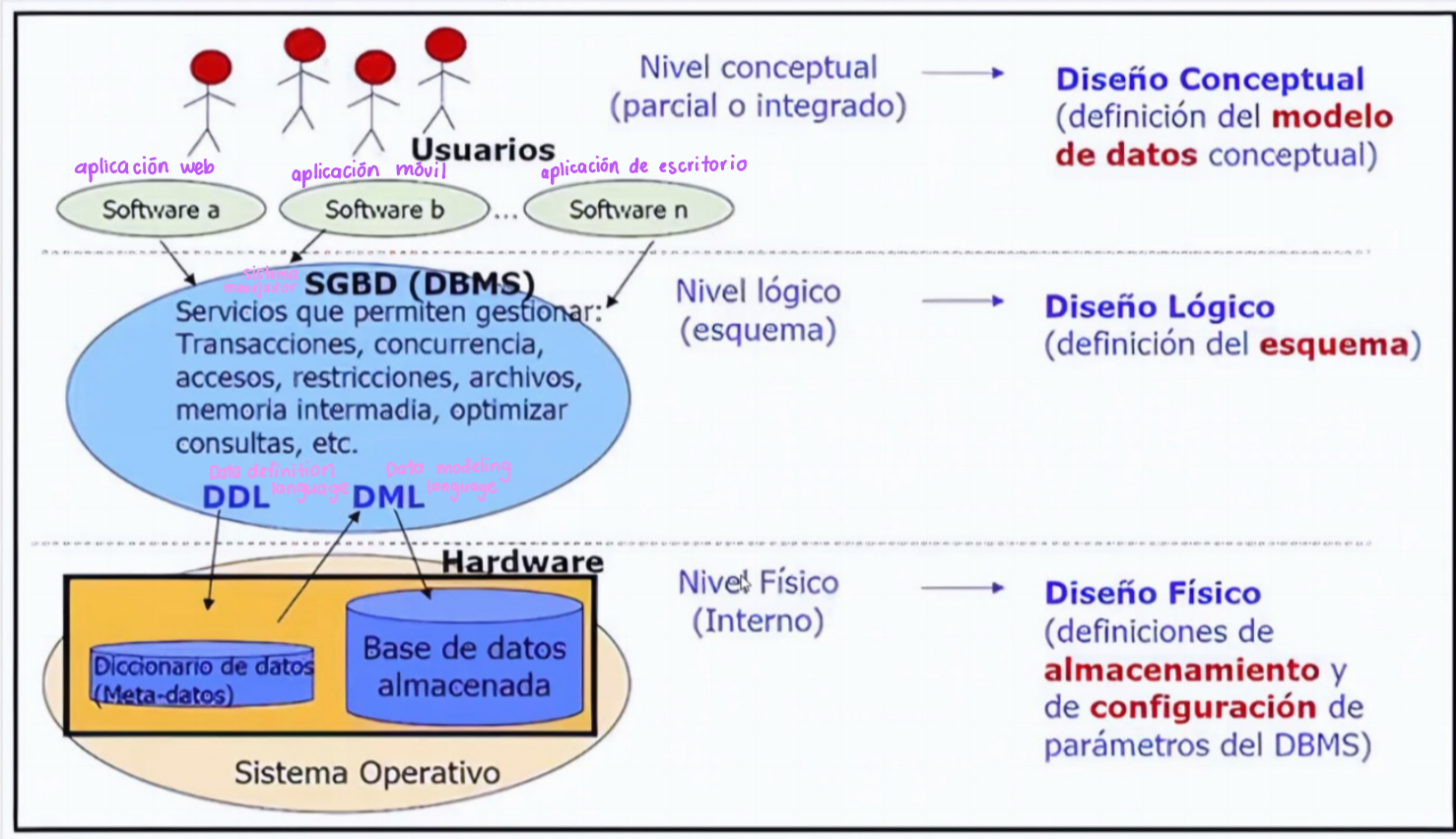
**Base de datos XML**

- Son buenas para la gestión de datos que no se puede obtener con cualquier otro tipo de base de datos, y un buen partido cuando se tiene una gran cantidad de datos en formatos no tradicionales, como video y audio.

**Base de datos de objetos**

- Este modelo, bastante reciente, y propio de los modelos informáticos orientados a objetos, trata de almacenar en la base de datos los objetos completos, facilitando el desarrollo e integración dentro de las aplicaciones

# Vision de datos



**Nivel físico:** es la representación del nivel más bajo de abstracción, en este se describe en detalle la forma en como se almacenan los datos en los dispositivos de almacenamiento

**Nivel conceptual:** describe que datos son almacenados en la base de datos y las relaciones que existen entre los mismos, describe la base de datos completa en términos de su estructura de diseño

**Definición de los datos:** se describen el tipo de datos y la longitud de campo todos los elementos direccionales en la base

**Relaciones entre datos:** se definen las relaciones entre datos para enlazar tipos de registros relacionados para el procesamiento de archivos múltiples

## Modelado de datos

Representación, clasificación, modelado

El modelo de datos representa:

- Cosas (o entidades u objetos) importantes para un modelo de negocio.
- Relación existente entre las cosas
  - Se crea y se modifica hasta que representa bien al modelo de negocio y puede crear un sistema
  - También se utiliza flujo de datos, modelo de estado, etc., en el **lenguaje de modelado unificado (UML)**

**Las estructuras de datos de la base:** el tipo de los datos que hay en la base y la forma en que se relacionan

**Las restricciones de integridad:** un conjunto de condiciones que deben cumplir los datos para reflejar la realidad deseada

**Operaciones de manipulación de los datos:** típicamente, operaciones de agregado, borrado, modificado y recuperación de los datos de la base

### Clasificación de los modelos de datos

**Modelado de datos lógicos:** son orientados a las operaciones más que a la descripción de una realidad. Usualmente están implementados en algún manejado de Base de Datos (modelo relacional)

**Modelado de datos físicos:** son estructuras de datos a bajo nivel implementadas dentro del propio manejado (árboles binarios)

**Definición:** método de descubrimiento de datos, relaciones y reglas de una empresa, llamados en su conjunto reglas de negocio

**Propósito:** los modelados de negocio debe crear modelos lógicos de datos para que usuarios y desarrolladores puedan entender las reglas de negocio

# Conceptos del modelado de datos

## Entidades

**Definición:** son cosas, siempre físicas, que tienen hechos asociados. Las descripciones de los tipos de entidades deben de ser lo más extensas en el desarrollo del modelo

**Mala descripción:** Proveedor: a quien compramos los productos

**Buena descripción:** Proveedor: empresa a la que acudimos para la obtención de productos. Los proveedores se evalúan según el precio, la calidad, el rendimiento de entregas y la estabilidad financiera

## Atributos

**Definición:** son datos o propiedades asociadas a un tipo de entidad

- Nombre
- Apellido
- ID
- CURP
- Dirección
- Teléfono

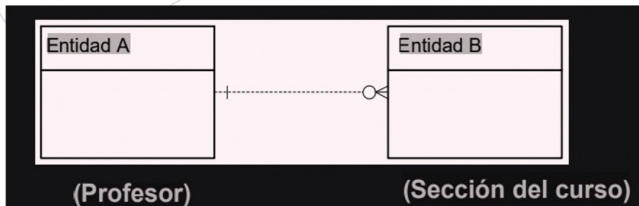
## Relaciones

**Definición:** son las líneas entre un conjunto de entidades.

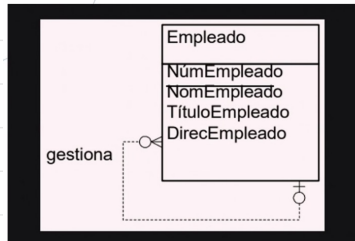
**Cardinalidad:** es el número esperado de instancias relacionadas entre las dos entidades de la relación

**Relaciones recurrentes:** la relación recurrente se da entre dos tablas: la real y la copia virtual. En este caso, una tabla de directores y una de empleados

### Relaciones



### Relaciones recurrentes



## Claves

**Primarias:** uno o más atributos que identifican unívocamente un registro

- Problemas al escogerla
- Estabilidad
- Control
- CURP
- Usar un sistema de claves generadas, si es posible

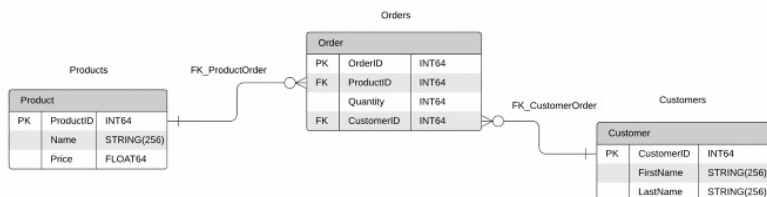
**Foráneas:** es la clave principal de una entidad padre y que se conserve como atributo ajeno a la clave principal en la entidad hijo

	A	B	C	D	E	F	G	H	I	J
1		IdAlumno	Curp	Nombre	correo	Sexo			Sexo	
2		1		Juan	juan@gmail.com	1			ID	Descripción
3		2		Jose	jose@gmail.com	1			1	Hombre
4		3		Sergio	sergio@gmail.com	1			2	Mujer
5		4		Juan	juan2@gmail.com	1				
6		5		Sergio	sergio@gmail.com	2				
7		6		Caro		2				
8		7		Sandra		2				

Llave primaria de Entidad 1

Llave foránea

## Claves foráneas de relaciones múltiples



**Integridad referencial:** mantiene la validez de las claves externas cuando la clave principal de la tabla padre cambia

- Cada clave foranea coincide con una principal o es nula

### Tipos de integridad referencial

**Débil:** Si en una tulla de R todos los valores de los atributos de K tienen un valor que no es el nulo, entonces debe existir una tulla en S que tome esos mismos valores en los atributos de J

**Parcial:** si en una tupla de R algun atributo de K toma el valor nulo, entonces debe existir una tupla en S que tome en los atributos de J los mismos valores que los atributos de K con valor no nulo

**Restriccion en cascada:** los filas de claves principales no se pueden eliminar, a no ser que las filas correspondientes de las tablas de claves externas se hayan eliminado. *Al eliminar un departamento no se eliminan todos los empleados.*

**Eliminacion en cascada:** si las filas de la tabla de claves principales se eliminan, las filas asociadas de las tablas de claves externas también se eliminaran. *Al eliminar un pedido, se eliminan todos los articulos*

**Actualizacion en cascada:** si las filas (claves) de la tabla de claves principales se actualizan, las filas asociadas de los tablas de claves externas también se actualizan. *Al cambiar un numero un número de departamento, se cambian también los numeros del empleado del departamento*

### DMBS (Sistema Manejador de Base de Datos)

- El DBMS es un conjunto de programas que se encargan de la creacion y todos los accesos a las bases de datos
- Se compone de un lenguaje de definicion de datos Y de un lenguaje de manipulacion de datos

### Lenguajes de BD (DDL/DML)

Es un conjunto de programas que se encargan de manejar la creacion y todos los acceso a las bases de datos

Existen 2 tipos de lenguajes, que son:

#### - El lenguaje de definicion de datos (DDL)

- Es utilizado para describir todas las estructuras de informacion y los programas que se usan para construir, actualizar e introducir la informacion que contiene una base de datos
- El DDL permite al administrador de la base especificar los elementos de datos que la integran, su estructura y las relaciones que existen entre ellos, las reglas de integridad, los controles a efectuar antes de autorizar el acceso a la base

**CREATE:** este comando crea un objeto dentro de la base de datos. Puede ser una tabla, vista, indice, trigger, funcion, procedimiento o cualquier otro objeto que el motor soporte

**ALTER:** este comando permite modificar la estructura de un objeto. Se pueden agregar quitar campos a una tabla, modificar el tipo de un campo, agregar, quitar indices de una tabla, etc

**DROP:** este comando permite eliminar un objeto de base de datos

#### - El lenguaje de manipulacion de datos (DML)

Es utilizado para escribir programas que crean, actualizan y extraen informacion de las bases de datos. Siempre de acuerdo con las especificaciones y normas de seguridad dictadas por el administrador

Un lenguaje de manipulacion de datos es un lenguaje que permite a los usuarios acceder o manipular los datos organizados mediante un modelo de datos apropiado.

**INSERT:** una sentencia INSERT de SQL agrega uno o mas registros a una (y solo una) tabla en una base de datos relacional

**UPDATE:** una sentencia UPDATE de SQL es utilizada para modificar los valores de un conjunto de registros existentes en una tabla

**DELETE:** una sentencia DELETE de SQL borra uno o mas registros existentes en una tabla

#### - El lenguaje de Consulta Estructurado (SQL)

Es un lenguaje declarativo de acceso a bases de datos relacionales que permite especificar diversos tipos de operaciones sobre las mismas. Una de sus características es el manejo del algebra y el calculo relacional permitiendo lanzar consultas con el fin de recuperar informacion de interés de una base de datos, de una forma sencilla.

**Elementos del lenguaje de manipulacion de datos**

**Clausulas:** son condiciones de modificacion utilizadas para definir los datos que desea seleccionar o manipular

Cláusula	Descripción
FROM	Utilizada para especificar la tabla de la cual se van a seleccionar los registros
WHERE	Utilizada para especificar las condiciones que deben reunir los registros que se van a seleccionar
GROUP BY	Utilizada para separar los registros seleccionados en grupos específicos
HAVING	Utilizada para expresar la condición que debe satisfacer cada grupo
ORDER BY	Utilizada para ordenar los registros seleccionados de acuerdo con un orden específico

Agosto

26

se usa mucho para detectar duplicados



Operador	Uso
AND	Es el "y" lógico. Evalúa dos condiciones y devuelve un valor de verdad sólo si ambas son ciertas.
OR	Es el "o" lógico. Evalúa dos condiciones y devuelve un valor de verdad si alguna de las dos es cierta.
NOT	Negación lógica. Devuelve el valor contrario de la expresión.

Cláusula	Descripción
FROM	Utilizada para especificar la tabla de la cual se van a seleccionar los registros
WHERE	Utilizada para especificar las condiciones que deben reunir los registros que se van a seleccionar
GROUP BY	Utilizada para separar los registros seleccionados en grupos específicos
HAVING	Utilizada para expresar la condición que debe satisfacer cada grupo
ORDER BY	Utilizada para ordenar los registros seleccionados de acuerdo con un orden específico

Operador	Uso
<	Menor que
>	Mayor que
<>	Distinto de
<=	Menor ó Igual que
>=	Mayor ó Igual que
=	Igual que
BETWEEN	Utilizado para especificar un intervalo de valores.
LIKE	Utilizado en la comparación de un modelo
In	Utilizado para especificar los valores de los registros a seleccionar

## Usuarios y administradores

Los usuarios de una base de datos pueden clasificarse en diferentes roles:

- Usuarios con diferentes niveles de acceso
- Administradores de Base de Datos

### Tipos de usuarios:

**Usuarios normales:** usuarios no sofisticados, que interactúan con el sistema mediante la ejecución de programas específicos escritos por otras personas. Normalmente la interfaz consiste en formularios e informes generados.

**Programadores de aplicaciones:** profesionales informáticos que escriben los programas de aplicación que utilizan los usuarios. Para ello se suelen usar lenguajes convencionales, entornos de herramientas de desarrollo rápido de aplicaciones

**Usuarios sofisticados:** interactúan con el sistema sin usar aplicaciones específicas, usando directamente el lenguaje de consultas. Los analistas de datos que utilizan consultas para explotar los datos en la base de datos entran en esta categoría

**Administrador de BD:** es aquel profesional que administra las tecnologías de la información y la comunicación, siendo el responsable de los aspectos técnicos, tecnológicos, científicos, inteligencia de negocios y legales de bases de datos, así como encargado de la calidad de datos

### Tareas de administrador BD

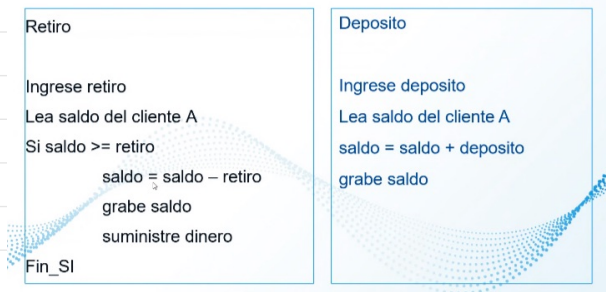
- Modelado de Datos y Diseño de Base de Datos
- Auditoría y seguridad
- Integración con aplicaciones
- Resguardo y recuperación de datos
- Inteligencia de negocios y almacenamiento de datos
- Administración de cambios
- Asegurar la alta disponibilidad
- Protección de datos
- Optimización del rendimiento
- Gestión de entornos de prueba
- Balanceo de carga
- Instalación y configuración del hardware
- Archivar datos

# Gestión de transacciones

**Transacción:** es un sistema de gestión de bases de datos es un conjunto de órdenes que se ejecutan formando una unidad de trabajo, es decir, en forma indivisible o atómica

- Retiro cajero electrónico
- Compra de producto en Amazon Mercado libre
- Registro en página Web aplicación móvil
- Solicitud de servicio en Uber Didi Bolt

## Ejemplo pseudocódigo



## Transacción sin control

Saldo Inicial 1000

Retiro 800	Retiro 700	Retiro 100	Deposito 300
Retiro = 800	Retiro = 700	Retiro = 100	Deposito = 300
Saldo = 1000	Saldo = 1000	Saldo = 1000	Saldo = 1000
Saldo = 200	Saldo = 300	Saldo = 900	Saldo = 1300

## Transacción con control

Saldo Inicial 1000

Retiro 800	Retiro 700	Retiro 100	Deposito 300
Retiro = 800	Retiro = 700	Retiro = 100	Deposito = 300
Saldo = 1000	Espera	Espera	Espera
Saldo = 200	Espera	Espera	Espera
Saldo = 200	no puede hacer el retiro de 700 porque no alcanza con el saldo	Saldo = 200	Espera
		Saldo = 100	Saldo = 100
			Saldo = 400

## Características de una transacción

- Una transacción es una o más sentencias que se toman como una unidad (todo termina bien o todo se aborta)
- Una transacción es una unidad lógica de trabajo
  - Definida para las reglas del negocio
  - Típicamente incluye al menos una modificación de datos
  - Pasa la base de datos de un estado consistente a otro
- Una transacción tiene dos posibles salidas:
  - **Committed:** todas las modificaciones quedan en firme, es decir todas las operaciones de actualización se aplican
  - **Rollback:** las modificaciones retornan a su estado inicial, es decir todas las operaciones de actualización se reversan

## Propiedades ACID

Una transacción se compone de 4 propiedades:

- Atomicidad
- Consistencia
- Aislamiento
- Durabilidad

**Consistencia:** esta propiedad garantiza que la base de datos quede en un estado consistente, independientemente si la transacción terminó de manera exitosa o fue abortada por algún error

**Aislamiento:** esta propiedad garantiza que todas las transacciones son independientes entre sí

**Durabilidad:** el sistema generador de base de datos asegura que perduran los cambios realizados por una transacción una vez que este termina con éxito

**Atomicidad:** se comportan como una sola unidad

## Rol de transacciones:

- Proteger los datos de las fallas del software, hardware
- Permitir el aislamiento de datos de tal forma que varios usuarios puedan acceder simultáneamente a los datos sin que existan interferencias entre ellos.

## Cuando usar transacciones?

Cuando un conjunto de sentencias se deben comportar como unidad



### Modo de transacción:

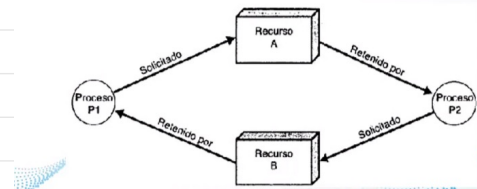
- En los manejadores de bases de datos existen dos modos de transacción:

- Commit explícito
- Auto commit

**Bloqueo:** mecanismo automático que tiene todos los motores de bases de datos cuya finalidad es aislar los datos para prevenir los conflictos cuando estos se están modificando y consultando al mismo tiempo

**Deadlock:** estos bloqueos permanentes suceden cuando dos procesos están compitiendo y bloqueando mutuamente recursos solicitados por ambos

## Deadlock



→ No nos va a pasar porque solo pasa en commit explícito y nosotros usaremos "autocommit"

**Lectura sucia** → Leyendo datos que NO están actualizados, por lo que están mal

- La transacción 1 modifica datos
- La transacción 2 lee los datos modificados antes de que la modificación haya terminado. Por lo que se puede decir que esta transacción lee datos uncommitted o dirty

- Ver un perfil (en Instagram) sin la foto que se acaba de subir

## Estructura de un sistema de Bases de Datos y arquitectura de aplicaciones

### 1. Estructura de un sistema de Bases de Datos

Un sistema de bases de datos se encuentra dividido en módulos, cada uno de los cuales controla una parte de la responsabilidad total del sistema

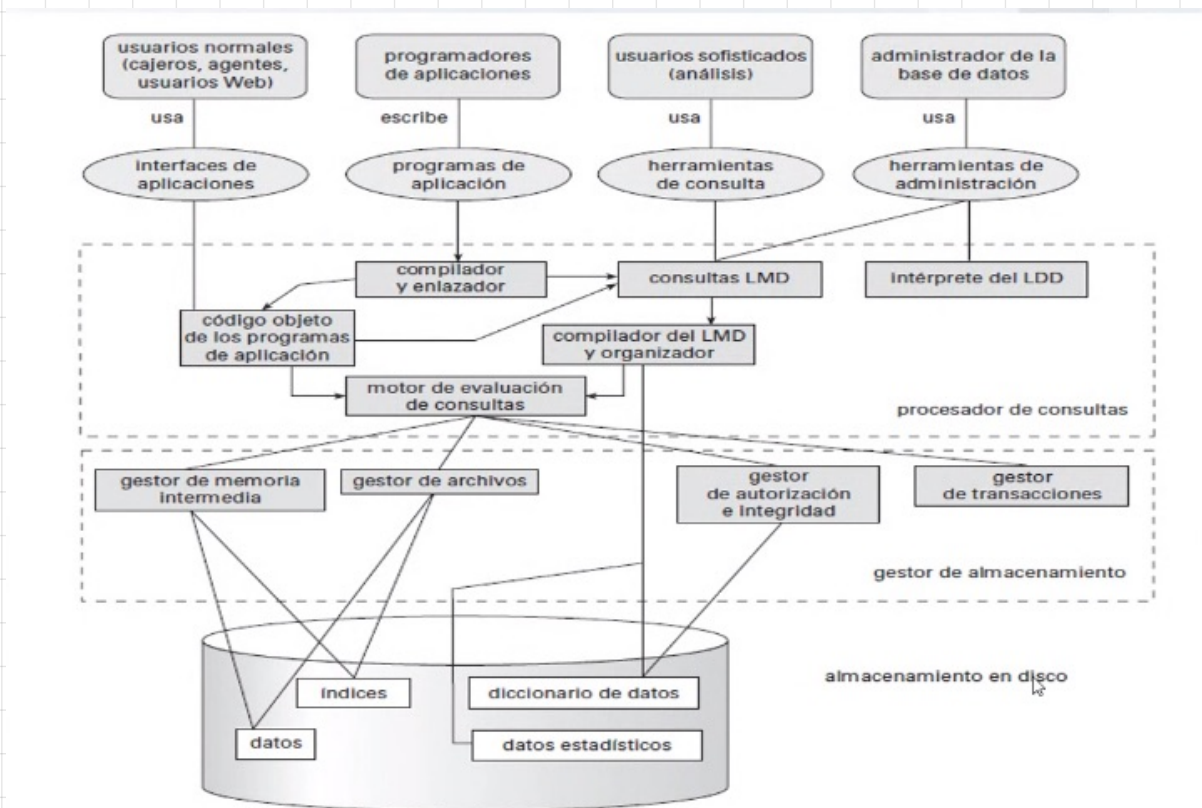


FIGURA 1.4. Estructura del sistema.

PENDIENTE