

Módulo 1: Práctica

TP 1 - Introducción a las bases de datos

- **Alumno:** Sussini Patricio
- **Tutor:** Constanza Uño
- **Materia:** Bases de Datos I

Ejercicio 1:

1. Diferencie entre "Dato" e "Información". Explique por qué un dato por sí solo carece de significado contextual y cómo se transforma en información útil. Proporcione un ejemplo original de cada uno.
 - Cuando nos referimos a Dato, lo hacemos en forma individual, por el contrario la Información en todo momento es plural, hace ilusión a un contenido. Dicho contenido ordenado y organizado son los datos individuales, cuyo orden y análisis otorgan a su conjunto característica de información. Un dato por su cuenta sin contexto no es capaz de ser dinámico, por ende no cumple la función de informar ni apoyar a los actores en la toma de decisiones o a su desarrollo. Un conjunto de datos desordenados tampoco, pero si a ese conjunto de datos se le aplica contexto y orden, se les otorga dinamismo y la imagen se convierte en video. Ahora los datos son información.

Ejemplos:

- *Dato*: Una palabra que dice **false** y el nombre propio **Francisco**
 - *Información*: Se pregunta si el Alumno Francisco, llegó a clase temprano hoy, para llenar el cuaderno de asistencias. Una vez generado el contexto llega la información: FRANCISCO ¿Temprano? = false.
-
2. Defina qué es un Sistema de Información (SI). Enumere y describa brevemente sus cuatro componentes principales.
 - “Un **sistema de información** es el conjunto de agentes, códigos y procesos que interactúan coordinadamente entre sí con un fin común o propósito.”
 - El *Hardware* es la parte física, los componentes electrónicos que se encargan de almacenar y procesar los datos.
 - El *Software* es el sistema que permite transformar y procesar la información.
 - Los *Datos* representan cada parte individual de la información.
 - La *Red* es el medio por el cual se comunican las diferentes partes.
 - Las *Personas* son las encargadas de tomar las decisiones, administrar y supervisar el funcionamiento de dicho sistema.

3. ¿Qué es una Base de Datos (BD)? Mencione y describa al menos cinco características esenciales de una Base de Datos.
- Una Base de Datos es una colección organizada de datos estructurados, almacenada electrónicamente en un sistema informático.
 - **Almacenamiento:** el almacenamiento de una base de datos es electrónico.
 - **Organización:** Los datos están de una forma organizada, en filas y columnas o árboles y redes.
 - **Software de Gestión:** Se usa un *Data Base Management System*
 - **Acceso de Múltiples Usuarios:** Permite que múltiples usuarios accedan y modifiquen información en simultáneo.
 - **Consultas:** Posibilidad de buscar y obtener información específica de la DB de forma eficiente.

Ejercicio 2:

1. En la década de los 50 y 60, se utilizaban los "Sistemas de archivos planos". Mencione los tres problemas principales que estos sistemas presentaban y que impulsaron la búsqueda de nuevas soluciones.
- **Dificultad para manejar grandes volúmenes de datos**
 - Cuando la cantidad de registros crecía, la búsqueda y actualización de información se volvía muy lenta, porque había que recorrer secuencialmente todo el archivo.
 - **Falta de organización jerárquica o estructurada**
 - Todos los datos estaban en el mismo "nivel", lo que dificultaba agrupar, clasificar y acceder a la información de manera lógica o temática.
 - **Redundancia y duplicación de datos**
 - Al no haber mecanismos para normalizar ni relacionar registros, la misma información se repetía en múltiples lugares, aumentando el uso de espacio y el riesgo de inconsistencias.

2. Antes del Modelo Relacional, ¿qué dos modelos de bases de datos surgieron en la década del 60? Describa brevemente cómo manejaba cada uno las relaciones entre los datos y cuál era su principal limitación.
- Surgieron los modelos:
 - Jerárquicos: Estructura de árbol con relaciones padre-hijo. Implementado en IMS de IBM.
 - **Rigidez estructural:** si querías cambiar la jerarquía (añadir un nuevo tipo de relación o mover un nodo), había que rediseñar todo el esquema y los programas.
 - **Acceso limitado:** solo se podía recorrer en el orden padre → hijo; acceder a un dato que estaba “fuera de la rama” requería recorridos costosos.
 - **Poca flexibilidad en consultas:** era difícil formular consultas que no siguieran la jerarquía natural; no había algo como un SQL flexible.
 - De Red: Relaciones complejas entre registros. Estandarizado por CODASYL para mejorar flexibilidad.
 - **Alta complejidad de programación:** los desarrolladores debían conocer exactamente las rutas y relaciones para poder acceder a los datos; era más “procedimental” que declarativo.
 - **Dependencia fuerte del esquema físico:** cualquier cambio en la forma de almacenar datos afectaba directamente a todas las aplicaciones.
 - **Mantenimiento costoso:** las relaciones complejas podían volverse muy difíciles de gestionar a medida que el sistema crecía.
3. ¿Quién fue el pionero en proponer el Modelo Relacional y en qué año? ¿Qué innovación fundamental introdujo que sentó las bases para su predominio?
- El pionero del modelo relacional fue Codd y lo publicó a finales de 1969. Lo que era innovador y fundamental que sentó las bases para su predominio fue el sistema de tablas con claves primarias y foráneas.

Ejercicio 3:

1. El modelo relacional se construye sobre la base de la teoría de conjuntos. Relacione los siguientes conceptos fundamentales del modelo relacional con sus equivalentes más conocidos en una tabla de base de datos: Relación, Tupla y Atributo
 - Claramente la teoría del modelo relacional se construye sobre la base de la teoría de conjuntos. Es así como una tupla puede considerarse un conjunto de elementos con posición propia, una relación está vinculada a la idea de producto cartesiano entre dos conjuntos, intersección, unión y demás operaciones que se pueden hacer. Por último un atributo es una característica que diferencia a cada elemento del conjunto o a cada conjunto de los demás.
2. Explique la función y la importancia de las "Claves" en el modelo relacional. Mencione y describa al menos cuatro garantías que las claves proporcionan a los datos.
 - Las claves en el modelo relacional funcionan como los ejecutores de las relaciones y operaciones entre los conjuntos matemáticamente. Son aquellos indicadores que nosotros agregamos a los elementos de las tablas para generar esas relaciones.
 - Las claves garantizan la integridad y unicidad de los datos en una base de datos. Proporcionan mecanismos para asegurar que cada fila sea identificable de forma única y que las relaciones entre tablas se mantengan consistentes.
3. ¿Qué es un "Índice" en el contexto de una base de datos relacional y cuál es su principal utilidad? Utilice un ejemplo para ilustrar cómo un índice mejora el rendimiento de las consultas.
 - Índice es una estructura de datos que mejora la velocidad de las consultas a una tabla al permitir un acceso más rápido a los registros. Actúa como un índice de libro, permitiendo encontrar la información deseada sin tener que revisar cada página. Su principal utilidad es acelerar las búsquedas.

- Imaginemos una tabla de clientes con 100,000 registros. Si queremos encontrar el cliente con ID 50,000, sin un índice, la base de datos tendría que revisar cada registro uno por uno hasta encontrarlo. Este proceso puede llevar tiempo.
4. Considerando el ejemplo de la "Tabla Alumnos" que se presenta en los materiales:
- a. Identifique un atributo que funcione como clave para esta tabla.
 - i. El ID alumno y materia.
 - b. Escriba una tupla completa de la tabla "Alumnos".
 - i. (id_materia, nombre, apellido, DNI)
 - c. ¿Cuál es el nombre de la relación (o tabla) en este ejemplo?
 - i. alumnos()

Ejercicio 4:

1. Utilizando la tabla comparativa de los documentos, contraste las Bases de Datos y los Archivos en términos de su estructura y la gestión de la seguridad.
 - Los archivos tienen organización lineal o secuencial mientras que las bases de datos tienen organización estructurada, jerárquica o según el modelo.
 - No existen relaciones entre los archivos, no hay jerarquía entre los datos, en las bases de datos existen relaciones lógicas entre datos.
 - Los archivos no poseen gestión de seguridad o es muy básica, todo el control es manual a nivel archivo completo. No hay control de secuencias. En las bases de datos existen mecanismos formales de control de acceso a nivel de usuario, tabla, registro o incluso campo. Validaciones automáticas y control de concurrencia para acceso simultáneo.

2. Describa en qué escenarios específicos se recomienda el uso de archivos y en cuáles el de bases de datos.
 - Es preferente usar archivos en los siguientes casos:
 - Datos simples y de bajo volumen.
 - Procesos secuenciales
 - Simple y sencilla portabilidad.
 - Costos y requisitos muy bajos, mínimos.
 - Es preferible usar Bases de Datos en los siguientes casos:
 - Grandes volúmenes de datos con estructura.
 - Múltiples usuarios que acceden simultáneamente a la información.
 - Necesidad de realizar consultas complejas y flexibles.
 - Integridad y consistencia de datos.
 - Escalabilidad a largo plazo.
 - Seguridad detallada.

Ejercicio 5:

1. Complete la siguiente tabla relacionando las operaciones del Modelo Relacional con sus equivalentes en el Teoría de Conjuntos:

Teoría de Conjuntos	Modelo Relacional
Selección	WHERE (filtrar tuplas)
Producto Cartesiano	JOIN (combinación de tablas)
Unión	UNION
Atributo	Columna (campo)
Elemento	Tupla (fila)
Intersección	INTERSECT
Proyección	SELECT columna1, columna2...
Diferencia de Conjuntos	EXCEPT o MINUS
Conjunto	Relación (tabla)

2. Mencione y explique brevemente al menos tres ventajas fundamentales que esta base matemática aporta al modelo relacional.
 - **Formalidad y Precisión:** como se basa en conceptos matemáticos bien definidos, el modelo relacional evita ambigüedades y asegura que las operaciones sobre los datos tengan un significado exacto y consistente.
 - **Operaciones predecibles y comprobables:** Cada operación tiene propiedades matemáticas conocidas, lo que permite predecir el resultado antes de ejecutarla y verificar su corrección mediante reglas lógicas.
 - **Independencia lógica de datos:** El modelo describe los datos en términos lógicos y no físicos, lo que permite cambiar la forma de almacenamiento sin alterar las consultas.

Ejercicio 6:

1. Basándose en las "Reflexión Final" y "Conclusión" de los documentos, explique por qué el Modelo Relacional, a pesar de las décadas, sigue siendo el más utilizado en la actualidad.
 - Considero personalmente que sigue siendo el modelo más utilizado por su estrecha y casi perfecta relación con conceptos matemáticos, lógicos que lo hacen una herramienta casi natural para su trabajo. Es difícil reinventar la rueda, no creo que nunca se deje de usar la rueda redonda.
2. Como futuros desarrolladores y programadores, y considerando la información en los materiales, ¿cómo la comprensión profunda de las claves, índices y estructuras de datos (como vectores, listas y árboles) influirá en su capacidad para diseñar sistemas de información eficientes y escalables?
 - Es sumamente importante conocer el funcionamiento de la lógica de los sistemas, para poder tomar decisiones acertadas sobre cómo empezar a desarrollar un sistema o poder mejorar los existentes. Es como saber las reglas del truco antes de empezar a jugar o como funciona un vehículo antes de aprender a manejar.