



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

FUNDAMENTOS DE BASES DE DATOS - 7094

T A R E A 1

EQUIPO:

DEL MONTE ORTEGA MARYAM MICHELLE - 320083527

SOSA ROMO JUAN MARIO - 320051926

CASTILLO HERNÁNDEZ ANTONIO - 320017438

ERIK EDUARDO GÓMEZ LÓPEZ - 320258211

JULIO CÉSAR ISLAS ESPINO - 320340594

FECHA DE ENTREGA:

23 DE AGOSTO DE 2024

PROFESOR:

M. EN I. GERARDO AVILÉS ROSAS

AYUDANTES:

LUIS ENRIQUE GARCÍA GÓMEZ

KEVIN JAIR TORRES VALENCIA

RICARDO BADILLO MACÍAS

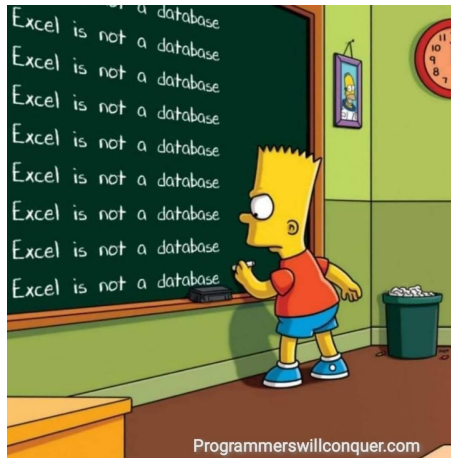
ROCÍO AYLIN HUERTA GONZÁLEZ



Tarea 1

- a. Describe las principales características del enfoque de bases de datos y contrástalo con el enfoque basado en hojas de cálculo. ¿En qué casos no tendría sentido utilizar una hoja de cálculo? ¿En qué condiciones resultaría mejor opción utilizar las hojas de cálculo?

Primero vamos a repasar un poco las diferencias entre estas 2 tecnologías para posteriormente ver en que situaciones nos conviene usar cada una.



[4]

Hay que aclarar antes que nada que en realidad estos 2 sistemas manejadores de la información comparten mas cosas de las que las diferencian; esto pues ambas tienen un propósito similar, el de guardar datos de manera estructura y de manera superficial ambas guardan "tablas". Sin embargo es cuando nos metemos a ver los detalles de como funcionan y que nos dejan hacer que empezamos a ver el porque de usar cada.

Característica	Hojas de calculo	Bases de datos
Integridad de los datos	Es difícil de mantener	Es mas fácil mantener
Redundancia de datos	Es fácil guardar varias veces lo mismo por falta de atomicidad	Cuando esta bien hecha tiene pocos datos redundantes
Validez	Es difícil asegurar que nuestros datos tengan formato valido	Es algo mas fácil pasar nuestros datos a una forma normal
Visualización	Es fácil perder datos al tener que recorrer la pantalla	Con consultas SQL la maquina recorre por nosotros
Rendimiento y capacidad	Como vimos en clase excel no puede tener mas de 1,048,576 filas por 16,384 columnas y la velocidad a la que procesa incluso estos datos es lenta	Oracle tiene un limite teórico de 8 exabytes aunque depende del hardware y la implementación, la velocidad esta ultra optimizada.
Seguridad	Suelen ser difíciles de proteger	Tienen estándares de seguridad y información distribuida

Tabla 1: Diferencias entre Hojas de Calculo y BDD [2]

Entre otras cosas, como multiuso o su compatibilidad con la programación, las BDD parece como si le sacaran una gran ventaja a las hojas de calculo, entonces ¿porque usar hojas de calculo?

Bueno la verdad es que es una respuesta algo simple, y es que las bases de datos son muy caras y/o complejas para cosas simples, y es que para personas que no sean muy ávidas a la tecnología, utilizar una BDD y un SMBD es algo bastante complicado; además de esto, hay que aclarar que la mayoría bases de información no son realmente tan grandes, digamos que como persona es raro exceder las capacidades del poderosísimo excel y no es hasta que escalamos el problema a cientos de miles de registros o seamos una empresa que necesita ciertas de estas características que vale la pena empezar a usar una BDD.

Entonces de manera concisa cuando usar cada uno:

Hojas de calculo

Mayormente para usos pequeños, de bajo coste y que no necesiten escalar, como puede ser tus finanzas personales, las de empresas o si se quiere una solución mas simple. Por ejemplo no tendría sentido intentar por ejemplo entrenar una IA con datos de hojas de calculo entre varias razones por el volumen de información, el rendimiento y la validez de la información.

PRESUPUESTO del mes de mayo 2021

ARTÍCULOS	INGRESOS	GASTOS	CAPACIDAD DE PAGO
Sueldo fijo Antonio	15000		
Sueldo fijo Mariana	15000		
Venta de gelatinas	3000		
Comida		4000	
Clases niños		6000	
Servicios		1500	
Mensualidad pago de auto		2000	
Despensa		2000	
Transporte		1500	
Renta		10000	
Ahorro		6000	
	33000	33000	

[1]

Bases de datos

Son mejores para todo lo demás; no es cierto, básicamente cualquier aplicación de la vida real, todas las empresas medianas en adelante, cosas de gobierno, todo lo relacionado a data, mucho de lo que usa la IA, videojuegos entre muchos otros.

- b. ¿Qué ventajas y desventajas encuentras al trabajar con un Sistema de Bases de Datos considerando que se planea implantar este sistema en una empresa de telemarketing?
- c. ¿Qué es la independencia de datos y por qué se carece de ella en los sistemas de archivos? ¿Cuál tipo de independencia de datos es más difícil de lograr? Justifica tu respuesta.

Sabiendo que una base de datos tiene niveles, entendemos por independencia de datos a la capacidad de modificar el esquema de una base de datos en un nivel sin afectar los esquemas en otros niveles, es decir, la independencia de los datos consiste en la capacidad de modificar la presentación de una base de datos sin tener que afectar o hacer cambios en las aplicaciones que se sirven de ella. Dicho esto, existen dos tipos de independencia de datos.

- a) Independencia física: Es la capacidad de modificar el esquema físico de la base, es decir, cómo los datos están almacenados físicamente en el hardware (índices, estructuras de almacenamiento, etc) sin tener que modificar el esquema lógico o las aplicaciones que acceden a los datos. Esto significa que podemos optimizar el almacenamiento o cambiar la forma en que los datos se almacenan y recuperan para mejorar el rendimiento, sin afectar a los usuarios finales o a las aplicaciones.
- b) Independencia lógica: Es la capacidad de modificar el esquema lógico, es decir, la estructura de tablas, relaciones, vistas, etc., sin necesidad de cambiar las aplicaciones o los programas que acceden a los datos. Por ejemplo, si se añade una nueva columna a una tabla o se cambian algunas relaciones, las aplicaciones no deberían verse afectadas si no utilizan esas partes del esquema que se han modificado.

Ahora bien, los sistemas de archivos no ofrecen la independencia de datos que un sistema de bases de datos. De hecho, es todo lo contrario, ya que las aplicaciones dependen del tipo de datos que manejan, por lo que editar o modificar la definición de los datos afecta directamente a las aplicaciones que trabajan con los datos modificados.

Por otro lado, podemos considerar la independencia lógica más difícil ya que trabajamos con la abstracción de los datos, además de que representa la estructura y organización de los datos tal como son percibidos por los usuarios y las aplicaciones, lo que significa que está mucho más ligado a cómo los datos son utilizados y comprendidos en la aplicación. Por lo que mantiene una estrecha relación entre el esquema lógico de la base de datos y las aplicaciones que dependen de ese esquema, lo que hace que cualquier cambio en el esquema lógico sea potencialmente disruptivo para esas aplicaciones.[5]

- d. Describe el papel que tienen los Sistemas Manejadores de Bases de Datos (SMBD) en el enfoque de bases de datos. ¿Por qué consideras que es importante (o no) que un administrador de bases de datos (DBA) conozca las características de un SMBD?
- e. Indica las responsabilidades que tiene un Sistema Manejador de Bases de Datos y para cada responsabilidad, explica los problemas que surgirían si dicha responsabilidad no se cumpliera.
- f. Investiga qué es la redundancia de datos. ¿Cuál sería la diferencia entre redundancia de datos controlada y no controlada?
- g. Investiga cuáles son las responsabilidades de un DBA. Si asumimos que el DBA nunca está interesado en ejecutar sus propias consultas, ¿necesita entender y/o conocer el modelo de datos lógico de la base de datos? Justifica tu respuesta.
- h. Investiga por qué surgieron los sistemas NoSQL en la década de 2000 y compara a través de una tabla sus características vs. los sistemas de bases de datos relacionales.

Este tipo de sistemas se desarrollaron para abordar las limitaciones de los sistemas relacionales, que a menudo enfrentaban desafíos en términos de escalabilidad, flexibilidad y rendimiento al manejar grandes volúmenes de datos no estructurados. A medida que las aplicaciones web y móviles crecieron en popularidad, se necesitaban soluciones capaces de escalar horizontalmente, manejar datos diversos sin esquemas fijos y ofrecer alta disponibilidad con mejor tolerancia a fallos. Los sistemas NoSQL respondieron a estas necesidades al proporcionar modelos de datos más flexibles, mayor rendimiento y capacidades avanzadas para manejar datos en tiempo real y grandes volúmenes de información.

Veamos la comparación de algunas características [3]:

Característica	Sistemas NoSQL	Sistemas de Bases de Datos Relacionales
Modelo de Datos	Flexible (clave-valor, columnares, documentos, grafos)	Estructurado (tablas con esquema fijo)
Escalabilidad	Horizontal (agregar más servidores)	Vertical (mejorar hardware del servidor)
Rendimiento	Optimizado para alta disponibilidad y baja latencia	Puede ser más lento con grandes volúmenes de datos
Flexibilidad del Esquema	Esquema dinámico, datos semi-estructurados o no estructurados (más flexible)	Esquema fijo y predefinido (menos flexible)
Garantías ACID	A menudo sacrificadas por rendimiento y disponibilidad	Fuerte soporte para transacciones ACID
Disponibilidad	Diseñado para replicación y particionamiento distribuidos	Requiere configuraciones adicionales para alta disponibilidad
Usos	Aplicaciones web grandes, big data, datos en tiempo real	Aplicaciones empresariales tradicionales, sistemas transaccionales
Cuándo usar	Cuando el volumen de datos no crece tanto con el tiempo	Cuando el volumen de datos crece rápidamente

- i. Investiga tres noticias relacionadas con las hojas de cálculo, en donde se muestre que hay errores en el procesamiento/manejo/distribución de datos y realiza una discusión breve de cada una de ellas, contrastándolas con el enfoque de las bases de datos. ¿Por qué consideras que a pesar de los ejemplos mostrados, las hojas de cálculo siguen siendo la herramienta más prevaleciente en un ambiente de negocios típico?
- j. Supón que deseas crear una red social orientada al uso empresarial, negocios y empleo, similar a LinkedIn. Considera cada una de las desventajas indicadas en el documento “Purpose of Database Systems”, cuando se administran los datos en un sistema de archivos. Discute la relevancia de cada uno de los puntos indicados, con respecto al almacenamiento de datos de los perfiles profesionales: el usuario que lo subió, las fechas de los empleos que ha tenido, sus habilidades, sus postulaciones y los usuarios que vieron su perfil, cantidad de publicaciones realizadas, cantidad de reacciones, cantidad de compartidos, entre otros.

Bibliografía

Referencias

- [1] Crédito Maestro. Finanzas personales, ¿cómo hacer tu presupuesto en excel?, 2024. URL <https://www.creditomaestro.com/blog/finanzas-saludables/finanzas-personales-como-hacer-tu-presupuesto-en-excell>. Accedido: 19-08-2024.
- [2] KDE Documentation Team. Database and spreadsheet in kexi, 2024. URL <https://docs.kde.org/trunk5/es/kexi/kexi/database-and-spreadsheet.html>. Accedido: 19-08-2024.
- [3] Redón Yecid. Bases de datos relacionales vs. no relacionales, 2019. URL <https://www.pragma.com.co/academia/lecciones/bases-de-datos-relacionales-vs.-no-relacionales>. Accedido en agosto 8, 2024.
- [4] r/ProgramerHumor/splint3e. :(, 2020. URL https://www.reddit.com/r/ProgrammerHumor/comments/j7f8c7/_/#lightbox. Accedido en agosto 19, 2024.
- [5] Universidad Autónoma de México. Características de una base de datos, No especificado. URL [https://uapas2.bunam.unam.mx/matematicas/caracteristicas_base_datos/#:~:text=Se%20conoce%20como%20independencia%20de,\(Independencia%20%C3%B3gica%20de%20datos\)](https://uapas2.bunam.unam.mx/matematicas/caracteristicas_base_datos/#:~:text=Se%20conoce%20como%20independencia%20de,(Independencia%20%C3%B3gica%20de%20datos)). Accedido en agosto 8, 2024.