

Bases de Datos

Introducción a las Bases de Datos

Lic. Víctor M. Cardentey Fundora
Dra. Lucina García Hernández

Departamento de Computación
Facultad de Matemática y Computación
Universidad de La Habana

12 de septiembre de 2023

¿Qué tan grande es YouTube?

- ▶ 500 horas de video subidas cada minuto

¿Qué tan grande es YouTube?

- ▶ 500 horas de video subidas cada minuto
- ▶ 720000 horas de video subidas cada día

¿Qué tan grande es YouTube?

- ▶ 500 horas de video subidas cada minuto
- ▶ 720000 horas de video subidas cada día
- ▶ Se necesitan **82 años** para ver el contenido de un solo día





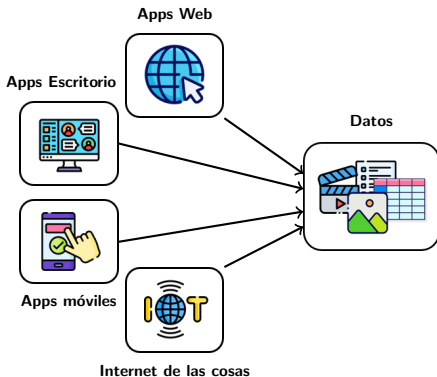
Motivación

Problemas a resolver

Motivación

Problemas a resolver

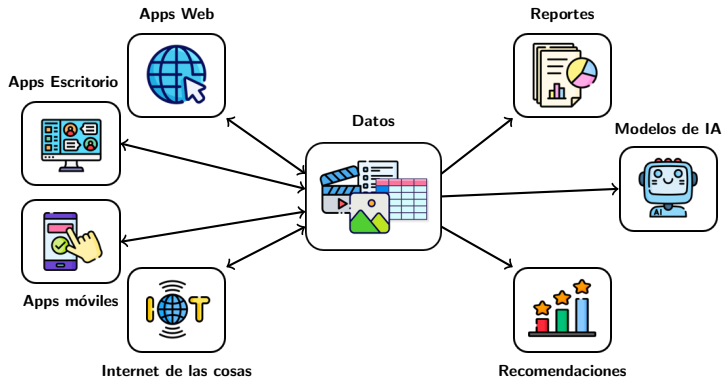
- Garantizar la persistencia de los datos generados por aplicaciones y dispositivos.



Motivación

Problemas a resolver

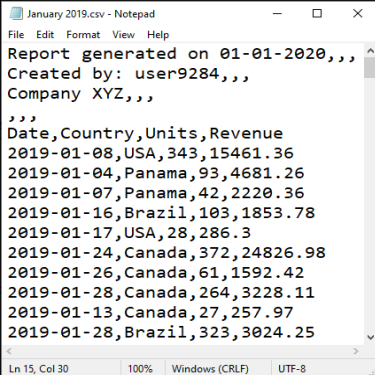
- ▶ Garantizar la persistencia de los datos generados por aplicaciones y dispositivos.
- ▶ Utilizar grandes cantidades de datos de forma eficiente.



¿Cómo?

¿Cómo?

La forma más sencilla de almacenar datos es escribirlos en un fichero



```
January 2019.csv - Notepad
File Edit Format View Help
Report generated on 01-01-2020,,,
Created by: user9284,,,
Company XYZ,,,
,,,
Date,Country,Units,Revenue
2019-01-08,USA,343,15461.36
2019-01-04,Panama,93,4681.26
2019-01-07,Panama,42,2220.36
2019-01-16,Brazil,103,1853.78
2019-01-17,USA,28,286.3
2019-01-24,Canada,372,24826.98
2019-01-26,Canada,61,1592.42
2019-01-28,Canada,264,3228.11
2019-01-13,Canada,27,257.97
2019-01-28,Brazil,323,3024.25
Ln 15, Col 30 100% Windows (CRLF) UTF-8
```

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se terminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se terminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se terminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia
- ▶ Gran redundancia de los datos

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se terminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia
- ▶ Gran redundancia de los datos
- ▶ Pobre control sobre los datos

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se terminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia
- ▶ Gran redundancia de los datos
- ▶ Pobre control sobre los datos
- ▶ Capacidades inadecuadas de manipulación de datos

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

98082200205	José	Pérez	22/08/98
-------------	------	-------	----------

*“... una base de datos es una colección **auto-descriptiva** de registros integrados.”*
—Allen Taylor

Auto-descriptiva: se almacenan metadatos (la descripción de su estructura) dentro del diccionario de datos de la propia base de datos.

Persona

CI : string	Nombre : string	Apellido : string	F. Nacimiento : date
98082200205	José	Pérez	22/08/98

*“... una base de datos es una colección auto-descriptiva de registros **integrados**.”*
—Allen Taylor

Integrados: no solo contiene los datos sino también las interrelaciones que se establecen entre estos.

Persona

CI : string	Nombre : string	Apellido : string	F. Nacimiento : date
98082200205	José	Pérez	22/08/98

Cuenta

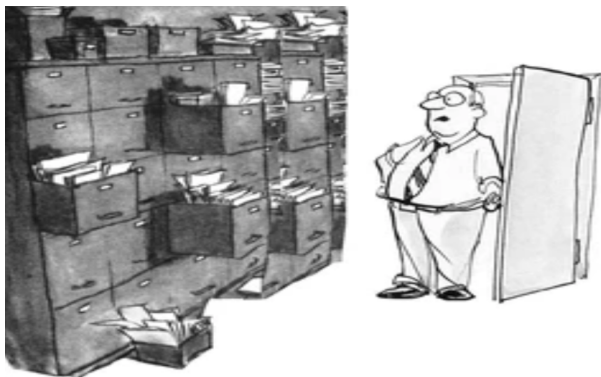
No. Cuenta : string	Balance : decimal	CI Dueño : string
8976	270.98	98082200205

Un cambio de enfoque

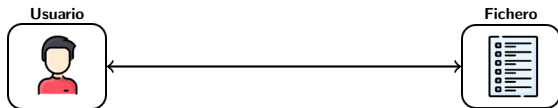
En los sistemas orientados a ficheros los humanos tienen el control sobre los ficheros

Un cambio de enfoque

En los sistemas orientados a ficheros los humanos tienen el control sobre los ficheros



Un cambio de enfoque



El usuario interactúa directamente con los ficheros



El usuario interactúa con la base de datos mediante *software*

¿Qué es este *software*?

Sistemas de gestión de bases de datos (SGBD)

*“Un **sistema de gestión de bases de datos** es un sistema computacional que proporciona **funcionalidades, medios o servicios para manipular** y, en particular, **manejar todos los accesos** a una base de datos o una colección de bases de datos.”*

—C. J. Date

Superando limitaciones

- ▶ **Persistencia:** los datos permanecen en memoria externa

Superando limitaciones

- ▶ **Persistencia:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos

Superando limitaciones

- ▶ **Persistencia:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos

Superando limitaciones

- ▶ **Persistencia:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos
- ▶ **Multi-usuarios:** protocolos para la gestión de accesos concurrentes

Superando limitaciones

- ▶ **Persistencia:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos
- ▶ **Multi-usuarios:** protocolos para la gestión de accesos concurrentes
- ▶ **Seguros:** consistentes ante accesos por usuarios no autorizados y fallos del sistema

Superando limitaciones

- ▶ **Persistencia:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos
- ▶ **Multi-usuarios:** protocolos para la gestión de accesos concurrentes
- ▶ **Seguros:** consistentes ante accesos por usuarios no autorizados y fallos del sistema
- ▶ **Confiables:** disponibles al 99.99999 %

Superando expectativas

Conveniencia de los SGBD

Conveniencia de los SGBD

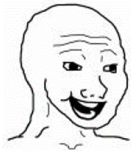
- ▶ **Independencia física de datos:** admiten el cambio de la forma de almacenamiento de los datos, pero la estructura de la base de datos y las operaciones definidas sobre ella no cambian.

Conveniencia de los SGBD

- ▶ **Independencia física de datos:** admiten el cambio de la forma de almacenamiento de los datos, pero la estructura de la base de datos y las operaciones definidas sobre ella no cambian.
- ▶ **Independencia lógica de datos:** proporcionan lenguajes de consulta declarativos, se define lo que se desea pero no cómo alcanzarlo.

Superando las expectativas

Usuario de SGBD



Lenguaje declarativo

Desarrollador de SGBD



Estructuras de datos

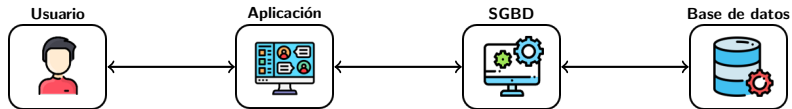
Algoritmos

Optimización

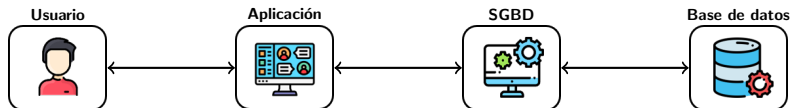
Compilación

Gestión de ficheros

Sistemas de bases de datos (SBD)



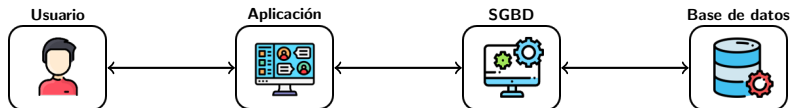
Sistemas de bases de datos (SBD)



*“Un **sistema de base de datos** es un **sistema computacional de mantenimiento de registros**, que se diseña para manejar grandes cantidades de información.”*

—C. J. Date

Sistemas de bases de datos (SBD)



*“Un **sistema de base de datos** es un **sistema computacional de mantenimiento de registros**, que se diseña para manejar grandes cantidades de información.”*
—C. J. Date

Funciones

- ▶ Insertar datos
- ▶ Editar datos
- ▶ Eliminar datos
- ▶ Consultar datos

Evolución de las tecnologías de la información

1980

1990

2000

2010



Primeros sistemas
gestores de bases de
datos relacionales
comerciales



Estándar de
lenguaje SQL



Primeros servidores
de bases de datos
relacionales



Procesamiento
analítico de datos

Evolución de las tecnologías de la información

1980

1990

2000

2010



Internet



Negocios online



Modelo orientado a
objetos
XML

Evolución de las tecnologías de la información

1980

1990

2000

2010



Data Warehousing



Redes sociales



Computación móvil



Inicios del Big Data

Evolución de las tecnologías de la información

1980

1990

2000

2010

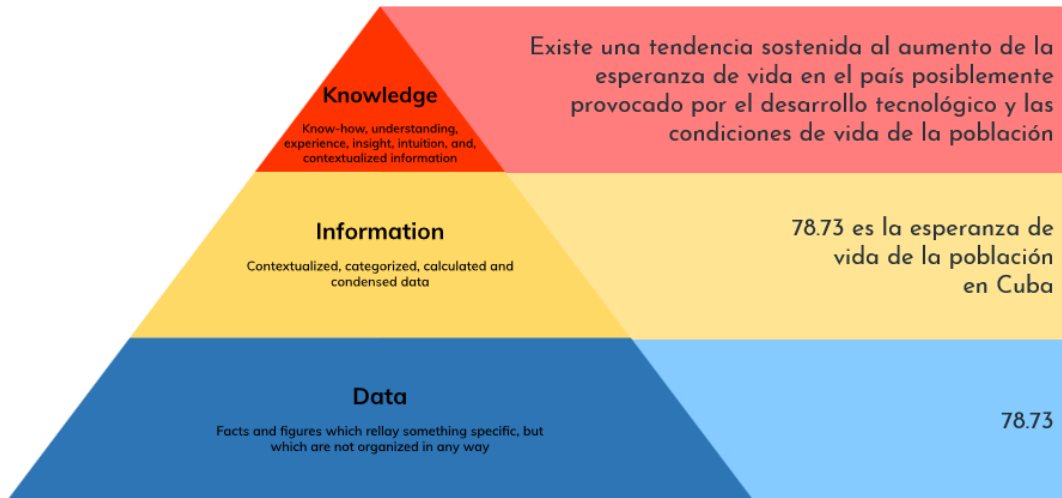


Sistemas gestores
de bases de datos
NoSQL



Computación en la
nube

Jerarquía del conocimiento



Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)
- ▶ Permiten el procesamiento transaccional de datos

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)
- ▶ Permiten el procesamiento transaccional de datos
- ▶ Base de los sistemas de información para la generación de conocimiento

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)
- ▶ Permiten el procesamiento transaccional de datos
- ▶ Base de los sistemas de información para la generación de conocimiento

Proporcionar un conjunto de métodos y herramientas para:

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)
- ▶ Permiten el procesamiento transaccional de datos
- ▶ Base de los sistemas de información para la generación de conocimiento

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)
- ▶ Permiten el procesamiento transaccional de datos
- ▶ Base de los sistemas de información para la generación de conocimiento

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas
- ▶ Evaluar la calidad de bases de datos específicas

Objetivos de la asignatura

Bases de datos relacionales

- ▶ Orientadas a almacenar datos estructurados (datos tabulares)
- ▶ Basadas en el modelo relacional (modelo matemático de datos)
- ▶ Permiten el procesamiento transaccional de datos
- ▶ Base de los sistemas de información para la generación de conocimiento

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas
- ▶ Evaluar la calidad de bases de datos específicas
- ▶ Identificar la vigencia del modelo relacional y sus limitaciones

Estructura de la asignatura

Análisis de Requerimientos

Estructura de la asignatura

Análisis de Requerimientos



Diseño Conceptual

Estructura de la asignatura

Análisis de Requerimientos

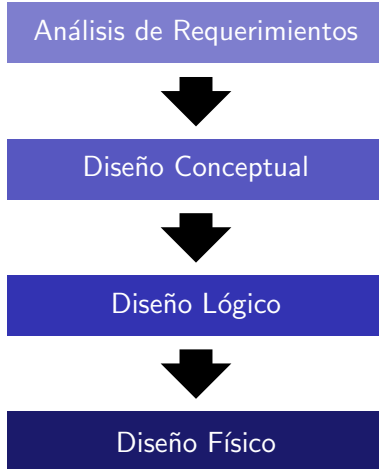


Diseño Conceptual



Diseño Lógico

Estructura de la asignatura



Evaluación de la asignatura

Evaluación Sistemática

- ▶ Trabajos de Control Parcial



Evaluación de la asignatura

Evaluación Sistemática

- ▶ Trabajos de Control Parcial
- ▶ Participación en clases
- ▶ Entrega de tareas orientadas
- ▶ Asistencia



Evaluación de la asignatura

Evaluación Sistemática

- ▶ Trabajos de Control Parcial
- ▶ Participación en clases
- ▶ Entrega de tareas orientadas
- ▶ Asistencia

Nota final

- ▶ Prueba Final
- ▶ Opinión del profesor



Bibliografía

- [1] Lucina García y Marta Montes de Oca. *Sistemas de Bases de Datos: Modelación y Diseño*. Universidad de La Habana, 2009.
- [2] Oracle Corporation. *MySQL 8.0 Reference Manual*.

Software

- ▶ MySQL 8.0
- ▶ Python
- ▶ Jupyter Notebook



@matcom_databases_cs

Bases de Datos

Introducción a las Bases de Datos

Lic. Víctor M. Cardentey Fundora
Dra. Lucina García Hernández

Departamento de Computación
Facultad de Matemática y Computación
Universidad de La Habana

20 de febrero de 2023