

# Bases de datos

## Introducción

Elizabeth León Guzmán, Ph.D.

[eleonguz@unal.edu.co](mailto:eleonguz@unal.edu.co)

Jonatan Gómez Perdomo, Ph. D.

[jgomezpe@unal.edu.co](mailto:jgomezpe@unal.edu.co)

Arles Rodríguez, Ph.D.

[aerodriguezp@unal.edu.co](mailto:aerodriguezp@unal.edu.co)

Camilo Cubides, Ph.D. (c)

[eccubidesg@unal.edu.co](mailto:eccubidesg@unal.edu.co)

Carlos Andres Sierra, M.Sc.

[casierrav@unal.edu.co](mailto:casierrav@unal.edu.co)

Research Group on Data Mining – Grupo de Investigación en Minería de Datos – (Midas)

Research Group on Artificial Life – Grupo de Investigación en Vida Artificial – (Alife)

Computer and System Department

Engineering School

Universidad Nacional de Colombia

# Agenda

## 1 Datos e información

- Datos
- Información

## 2 DBMS(DataBase Management System) - Sistema Administrador de Bases de Datos

## 3 Modelo relacional

## 4 MySQL



# Agenda

- 1 Datos e información
  - Datos
  - Información
- 2 DBMS(DataBase Management System) - Sistema Administrador de Bases de Datos
- 3 Modelo relacional
- 4 MySQL



# Datos I

## ¿Qué es un dato?

### Definición (Población)

- Conjunto con elementos que tienen propiedades en común.
- Conjunto bien definido (es posible identificar cuáles elementos pertenecen al conjunto y cuáles no).

### Características de la población

Son propiedades de la población que se pueden medir. En estadística, las propiedades medibles se llaman Variables).



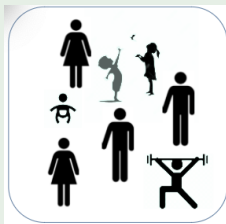
# Datos II

¿Qué es un dato?

## Ejemplo

**Población:** Personas que viven en mi barrio.

Las personas tienen características o propiedades en común que pueden ser *medibles*



**Característica de la población (variable):**

$$\text{edad} = \{25, 15, 6, 75, 15, \dots\}$$



# Datos III

## ¿Qué es un dato?

### Variable

$$edad(x) = ?$$

- La variable tiene rango.
- La variable aleatoria es una función.
- La posibilidad de que se tome un  $x$  y  $x$  tome un valor  $y$ , se le llama *evento*.



# Datos IV

## Datos

Los datos son:

- Hechos individuales acerca de algo de interés para alguien.
- Representación simbólica de una variable numérica o categórica.

## Ejemplos

Temperatura : 17, 28, 15

Ciudad : Bogotá, Cartagena

Fecha : Julio 20 2020, Julio 20 2021



# Datos V

## Tipos de Datos

- Las variables cuantitativas son las que se expresan mediante un número. Se puede realizar operaciones aritméticas con ellas.
- Las variables cualitativas son las que expresan características o cualidades, y no pueden ser medidas con número

### Ejemplos

Cuantitativas: temperatura, edad, altura

Cualitativa: nombre, título, profesión

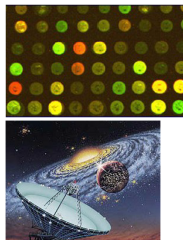




# Generación de Datos



Tomado de [2]



## Comerciales

- Web (*e-commerce*, *e-learning*).
- Supermercados (compras).
- Bancos (transacciones con tarjetas, *web*).

## Científicos

- Satélites
- Telescopios
- *Microarrays* (información genética)
- Simulaciones



# Agenda

## 1 Datos e información

- Datos
- Información

## 2 DBMS(DataBase Management System) - Sistema Administrador de Bases de Datos

## 3 Modelo relacional

## 4 MySQL



# Información I

## ¿Qué es información?

- Datos relacionados y estructurados.
- Almacenados generalmente en Bases de Datos.
- Consultas para obtener información.

Ciudad	Temperatura	Fecha
Bogotá	15	Julio 20 2020
Cartagena	28	Julio 20 2020
Bogotá	17	Julio 20 2021

Ej: los datos se relacionan en una fila de una tabla



# Información II



# Información III

Los datos se pueden procesar y modelar para guardarlos para luego recuperar información.

Los datos se pueden guardar en:

- Archivos txt
- Excel, pdf
- Estructuras
- Bases de datos

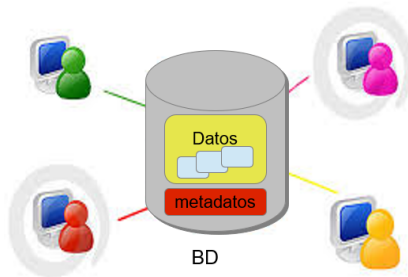


# Base de Datos

## Definición

Una colección de **datos relacionados**, y una descripción de estos datos, diseñados para cumplir con las necesidades de **información** de una organización.

*(Connolly & Begg)*



# Operaciones sobre los datos

Para cumplir con las necesidades de información las Organizaciones pueden realizar operaciones sobre los **datos** provenientes de su actividad que se encuentran almacenados en una base de datos:

- Almacenar
- Procesar
- Recuperar
- Actualizar
- Eliminar
- Intercambiar



# Actividades con requerimientos de datos

## Ejemplos

Algunos ejemplos de industrias y/o actividades que requieren tener almacenados los datos para satisfacer necesidades de información:

### Ejemplos

- Ventas
- Bancos
- Hospitales
- Educación
- Bibliotecas
- Finanzas
- Transporte





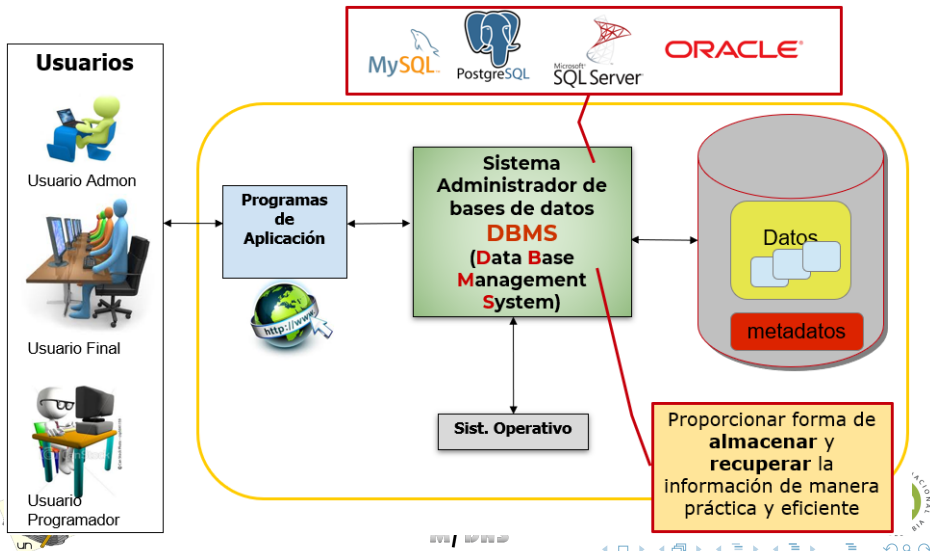
# Agenda

- 1 Datos e información
  - Datos
  - Información
- 2 DBMS(DataBase Management System) - Sistema Administrador de Bases de Datos
- 3 Modelo relacional
- 4 MySQL

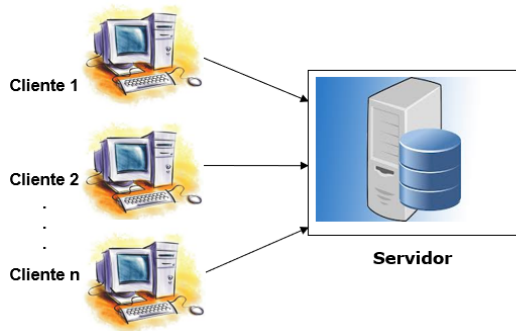


# DBMS (*DataBase Management System*)

Sistemas administrador de bases de datos



# Arquitectura Cliente Servidor



- **Servidor**: gestiona la base de datos (DBMS)
- **Cliente**: permite enviar comandos al servidor



# DBMS (*DataBase Management System*) I

## Definición (*DataBase Management System*)

UN DBMS es un conjunto de programas que **maneja** la **estructura** de la BD y **controla** el acceso a los datos guardados en ella.

- Permite crear y organizar los datos.
- Maneja los datos según las necesidades de los usuarios.
- Mantiene integridad y seguridad de los datos.
- Establece y mantiene rutas de acceso.
- Control de usuarios que acceden.



# DBMS (*DataBase Management System*) II

<b>Eficiencia</b>	Manejo de grandes cantidades de datos y de usuarios	<ul style="list-style-type: none"><li>• Multiusuario</li><li>• Manejo de concurrencia: para los usuarios es transparente que otros usuarios acceden a la base de datos simultáneamente.</li><li>• Miles de consultas por segundos</li></ul>
<b>Seguridad</b>	Proteger los datos de destrucciones maliciosas o cambios en los datos	<ul style="list-style-type: none"><li>• Robo, destrucción, intentos de falsificación</li></ul> <p><b>Ejemplo:</b> alguien que quiera incrementar el saldo de su cuenta bancaria</p> <ul style="list-style-type: none"><li>• <i>Datos deben ser consistentes a pesar de lo que pueda ocurrir</i></li></ul>
<b>Privacidad</b>	Garantizar acceso solo a los datos que se requieran	<ul style="list-style-type: none"><li>• Acceso a los datos pertinentes para cada perfil de usuario</li><li>• <b>Ejemplo:</b> los vendedores no deberían tener acceso a los datos de salario de los empleados de la compañía</li></ul>



# DBMS (*DataBase Management System*) III

## Recuperación de datos

Capacidad de recuperar los datos si estos son perdidos o dañados

- Fallas de hardware
- Desastre natural
- Persistencia: datos deben mantenerse. (Respaldo de los datos)

## Exactitud de Datos

Integridad y consistencia de los datos

- Evitar redundancia
- *Datos deben ser consistentes a pesar de lo que pueda ocurrir*

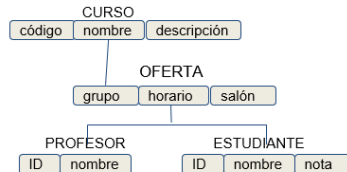


# Tipos de DBMS I

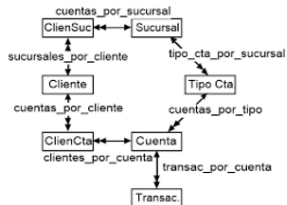
## Según modelo de datos

- DBMS jerárquico
- DBMS de red
- DBMS relacional
- DBMS orientados a objetos
- DBMS basada en clave-valor
  - Orientadas a Columnas
  - Documentales

- 1960
- IBM, UNIVAC, HoneyWell
- Computadores mainframe
- Legacy systems (datos históricos)



**Jerárquico:** Modelo de datos basado en árboles



**Red:** Modelo de datos basado en grafos

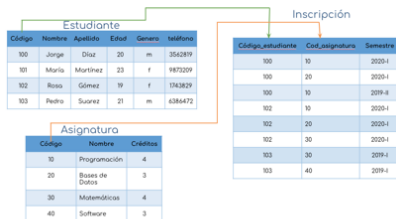


# Tipos de DBMS II

## Según modelo de datos

- DBMS jerárquico
- DBMS de red
- **DBMS relacional**
- DBMS orientados a objetos
- DBMS basada en clave-valor
  - Orientadas a Columnas
  - Documentales

- Viable en 1980
- Preferido desde entonces
- Lenguaje SQL



**Relacional:** Modelo de datos basado en tablas



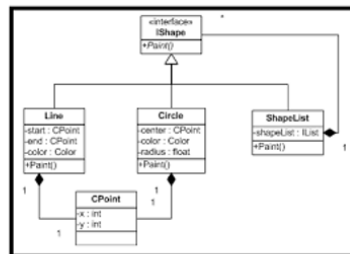


# Tipos de DBMS III

## Según modelo de datos

- DBMS jerárquico
- DBMS de red
- DBMS relacional
- **DBMS orientados a objetos**
- DBMS basada en clave-valor
  - Orientadas a Columnas
  - Documentales

- POO
- Se usa para aplicaciones que intuitivamente se manejan con objetos



**Orientado a Objetos:** Modelo de datos donde los datos se almacenan en objetos (clases)



# Tipos de DBMS IV

## Según modelo de datos

- DBMS jerárquico
- DBMS de red
- DBMS relacional
- DBMS orientados a objetos
- DBMS basada en clave-valor
  - Orientadas a Columnas
  - Documentales

- NoSQL
- Datos No estructurados o semiestructurados
- Big Data
- Preferibles para disponibilidad de datos
- Distribución de datos

```
{
  001:{
    personal: {
      nombre:{
        timestamp1 : "Juan"
      }
      telefono:{
        timestamp1 : 3129765437
        timestamp2 : 3114563523
      }
    }
    oficina: {
      dirección:
        timestamp1 : "Av 20 de Julio"
      }
      telefono: {
        timestamp1 : 3789029
      }
    }
  }
}
```



# Agenda

- 1 Datos e información
  - Datos
  - Información
- 2 DBMS(DataBase Management System) - Sistema Administrador de Bases de Datos
- 3 Modelo relacional
- 4 MySQL



# Modelo Relacional I

## Definición (Base de datos relacional)

Una base de datos relacional es una **colección de relaciones** que contienen los datos que describen un ambiente de negocios.

- Las relaciones se representan en **tablas** y cada una tiene un nombre exclusivo.
- Basado en **conceptos matemáticos**.



# Modelo Relacional II

Es el más usado desde el año 1970

Relación  $\longrightarrow$  Tabla

## Estudiante

Nombre	Apellido	Edad	Género	teléfono
Jorge	Díaz	20	m	3562819
María	Martínez	23	f	9873209
Rosa	Gómez	19	f	1743829
Pedro	Suarez	21	m	6386472



# Modelo Relacional III

Relación  $\longrightarrow$  Tabla

## Estudiante

Columna  
Campo  
Atributo



Nombre	Apellido	Edad	Género	teléfono
Jorge	Díaz	20	m	3562819
María	Martínez	23	f	9873209
Rosa	Gómez	19	f	1743829
Pedro	Suarez	21	m	6386472

fila  
registro  
tupla



# Modelo Relacional IV

## Varias tablas

### Estudiante

Código	Nombre	Apellido	Edad	Género	teléfono
100	Jorge	Díaz	20	m	3562819
101	María	Martínez	23	f	9873209
102	Rosa	Gómez	19	f	1743829
103	Pedro	Suarez	21	m	6386472

### Asignatura

Código	Nombre	Créditos
10	Programación	4
20	Bases de Datos	3
30	Matemáticas	4
40	Software	3

### Inscripción

Código_Estudiante	Cod_asignatura	Semestre
100	10	2020-I
100	20	2020-I
100	10	2019-II
102	10	2020-I
102	20	2020-I
102	30	2020-I
103	30	2019-I
103	40	2019-I

# Modelo Relacional V

## Varias tablas enlazadas por referencias

Estudiante

Código	Nombre	Apellido	Edad	Género	teléfono
100	Jorge	Díaz	20	m	3562819
101	María	Martínez	23	f	9873209
102	Rosa	Gómez	19	f	1743829
103	Pedro	Suarez	21	m	6386472

Asignatura

Código	Nombre	Créditos
10	Programación	4
20	Bases de Datos	3
30	Matemáticas	4
40	Software	3

Inscripción

Código_estudiante	Cod_asignatura	Semestre
100	10	2020-I
100	20	2020-I
100	10	2019-II
102	10	2020-I
102	20	2020-I
102	30	2020-I
103	30	2019-I
103	40	2019-I



# Aplicaciones

**Ventas:** Clientes, productos y compras.

**Bancos:** Clientes, cuentas, préstamos y transacciones bancarias.

**Hospitales** Historias clínicas, citas, doctores, enfermeras y enfermedades.

**Educación (colegios, universidades)** Estudiantes, matrículas, cursos y profesores.

**Bibliotecas:** Libros, autores, editoriales, préstamos.

**Finanzas:** Acciones, bonos, ventas y compras.

**Transporte:** Reservas, horarios, inventario.



# SQL-Structured Query Language

## Definición (SQL)

Lenguaje propio de los **DBMS** para interactuar con los datos

- Crear
- Definir
- Consultar
- Modificar
- Actualizar



# Modelo Relacional Vs. NoSQL

RELACIONALES	NoSQL
Aplicaciones centralizadas (ERPs, CRMs, etc)	Aplicaciones descentralizadas (web, IoT, móvil, etc.)
Datos estructurados Volumen moderado	Datos semi-estructurados y No estructurados. "Big Data"
Alta disponibilidad	Disponibilidad Continua
Velocidad Moderada	Alta velocidad. Datos de sensores, gps, máquinas, etc.
Transacciones complejas	Transacciones simples
Escalabilidad Vertical	Escalabilidad Horizontal



# Agenda

- 1 Datos e información
  - Datos
  - Información
- 2 DBMS(DataBase Management System) - Sistema Administrador de Bases de Datos
- 3 Modelo relacional
- 4 MySQL



# MySQL I

- DBMS relacional
- Programada en C y C++
- Es el DBMS de código abierto más popular.
- Esta bajo licencia dual: Licencia pública general/Licencia comercial por Oracle Corporation

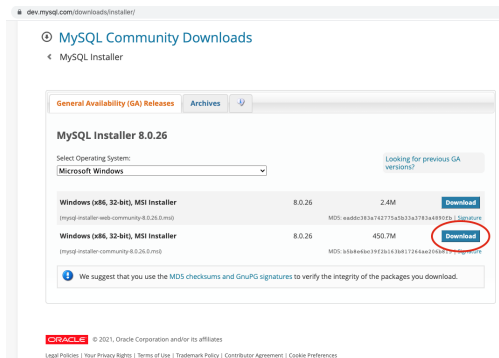


# MySQL II

## Instalación

Se instalará el servidor de MySQL y el cliente workbench

[Enlace](#) para bajar el instalador (incluye el servidor y el cliente) en windows.



dev.mysql.com/downloads/installer/

### MySQL Community Downloads

MySQL Installer

General Availability (GA) Releases Archives

#### MySQL Installer 8.0.26

Select Operating System:  
Microsoft Windows

Looking for previous GA versions?

Windows (x86, 32-bit), MSI Installer <small>(mysql-installer-web-community-8.0.26.0.msi)</small>	8.0.26	2.4M	<a href="#">Download</a>
Windows (x86, 32-bit), MSI Installer <small>(mysql-installer-community-8.0.26.0.msi)</small>	8.0.26	450.7M	<a href="#">Download</a>

We suggest that you use the MD5 checksums and GnuPG signatures to verify the integrity of the packages you download.

**ORACLE** © 2021, Oracle Corporation and/or its affiliates  
Legal Policies | Your Privacy Rights | Terms of Use | Trademark Policy | Contributor Agreement | Cookie Preferences



# Referencias I



Margaret H. Dunham, *Data mining: Introductory and advanced topics*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 2002.



Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, *Introduction to data mining, (first edition)*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.

