

# Bases de Datos

German Jesus Pereira Muñoz PhD.



9/18/2024

German Jesus Pereira Muñoz PhD.

1

1

---

---

---

---

---

---

---

## ¿Qué es Data?

- De acuerdo con la Oxford **"Datos" es una información distinta, generalmente formateada de una manera especial.** Los datos se pueden medir, recolectar, reportar y analizar, con lo cual a menudo se visualizan usando gráficos, imágenes u otras herramientas de análisis.
  - Los datos brutos (datos no procesados) pueden ser una recopilación de números o caracteres antes de que se haya limpiado y corregido por los investigadores.
  - Los datos de campo se recogen en un entorno in situ.
  - Los datos experimentales se generan en la observación de investigaciones científicas.
- El procesamiento de datos ocurre comúnmente en etapas, por lo que los datos procesados de una etapa también podrían considerarse los datos brutos de las etapas posteriores.

18/09/2024

German Jesus Pereira Muñoz PhD.

2

2

---

---

---

---

---

---

---

## ¿Qué es la información?

- Son datos que han sido tratados, organizados o estructurados de una manera que los hace significativos, valiosos y útiles.
- Son los datos que se les ha dado contexto, relevancia y propósito.
- Da conocimientos, comprensión y percepciones que pueden ser utilizados para la toma de decisiones, la resolución de problemas, la comunicación y varios otros propósitos.

### ¿Por qué los datos son importantes ?

- Los datos ayudan a tomar mejores decisiones.
- Los datos ayudan a resolver problemas al encontrar la razón del bajo rendimiento.
- Los datos ayudan a evaluar el rendimiento.
- Los datos ayudan a uno a mejorar los procesos.
- Los datos ayudan a entender a los consumidores y al mercado.

18/09/2024

German Jesus Pereira Muñoz PhD.

3

3

---

---

---

---

---

---

---

### Categorías de datos

Los datos pueden ser de dos categorías en dos partes principales.

**Datos estructurados:** se organizan en formato específico, lo que facilita la búsqueda, el análisis y el proceso, se encuentran en una base de datos relacionales que incluye información como números, datos y categorías.

**Datos desestructurados:** Los datos no se ajustan a una estructura o formato específicos. Puede incluir algunos documentos de texto, imágenes, videos y otros datos que no se organizan o analizan fácilmente sin procesamiento adicional.

18/09/2024

Germán Jesús Pereira Muñoz PhD.

4

4

---

---

---

---

---

---

---

### Tipos de datos

Generalmente los datos pueden clasificarse en dos partes:

**1.Datos de Categorías:** Datos que tienen una categoría definida, por ejemplo: Estado civil, Partido Político, Color de ojos

**2.Datos numéricos:** Pueden clasificarse en dos categorías:

- **Datos Discretos:** Contienen los datos que tienen valores numéricos discretos, por ejemplo, Número de niños, defectos por hora, etc.
- **Datos continuos:** Contienen los datos que tienen valores numéricos continuos, por ejemplo, peso, voltaje, etc.

**3.Escala nominal:** Clasifica los datos en varias categorías distintas en las que no se implica ningún criterio de clasificación. Por ejemplo, Género, Estado civil.

18/09/2024

Germán Jesús Pereira Muñoz PhD.

5

5

---

---

---

---

---

---

---

### Tipos de datos

**1. Escala ordinaria:** clasifica los datos en categorías distintas durante las cuales la clasificación está implícita.

- Licenciatura en la facultad: Profesora, Profesora Asociada, Profesora Asistente
- Estudiantes grado: A, B, C, D.E.F

**2. Escala intercalada:** intervalo puede ser una escala ordenada durante la cual la diferencia entre mediciones es una cantidad significativa, pero las mediciones no tienen un punto cero verdadero. Por ejemplo: Temperatura en Fahrenheit, Celsius y Años

**3. Escala de coeficiente:** relación puede ser una escala ordenada durante la cual la diferencia entre las mediciones es una cantidad significativa y por lo tanto las mediciones tienen un punto cero verdadero. Por lo tanto, se puede realizar operaciones aritméticas en datos a escala real. Por ejemplo: Pesado, Edad, Salario, etc.

18/09/2024

Germán Jesús Pereira Muñoz PhD.

6

6

---

---

---

---

---

---

---

## 2

7

## 3

8

## 2

9

## Archivos

- En el enfoque de archivos tradicionales (documentos) los datos se almacenan en archivos individuales, exclusivos para cada aplicación particular.
- Hoy en día los documentos tradicionales para el almacenamiento de datos es la planilla de Microsoft Excel o Google Sheet entre otros, la razón sigue siendo su simplicidad y rapidez para ingresar datos, para usuarios no necesariamente especialistas en TI.



9/18/2024

Germán Jesús Pereira Muñoz

10

10

## Archivos

- La información en estas planillas se ingresa en forma tabular con conceptos de fila y columnas que se han agregado también a su uso en las bases de datos.
- Es fácil ingresar datos en estas planillas de datos y luego cargar esta información para análisis en programas de análisis o en bases de datos especializadas. Generalmente debe exportarse estos datos en un formato más simple denominado CSV ("comma separated value").
- En estas planillas los datos pueden ser una alta fuente de error y la actualización de los archivos es *más lenta que en una base de datos*.



9/18/2024

Germán Jesús Pereira Muñoz PhD

11

11

## Archivos

Los problemas de utilización de estas planillas se resumen en (solo una muestra):

- **Redundancia** Al no existir algún tipo de control sobre el ingreso más que el del usuario, es normal que existan este tipo de errores de duplicidad en los registros.
- **Error de ingreso** Errores comunes en el ingreso manual de datos, errores de tipo ortográfico, números mal ingresados, etc.
- **Estandarización** se ejemplifica en el ingreso de fechas donde a pesar de poder regir el formato de entrada, no impide que se ingrese otros formatos que, si bien pueden ser correctos, interfieren en la forma de incluirse en una base de datos.

21-12-2021 o bien 21/02/2021, o 21/2/2021.

- **Seguridad** No hay un control de uso y acceso por usuarios a los datos, más que el control al archivo físico en el computador local o servidor.

9/18/2024


Germán Jesús Pereira Muñoz PhD

12

12

### Planillas de Datos

- El uso de planillas electrónicas como almacén de datos. Si bien presentan las desventajas antes descritas, también permiten en forma fácil realizar un ingreso de datos masivo y simple.
- Una planilla **no es una base de datos**.
- Contiene datos efectivamente, pero constituye un repositorio de datos.



9/18/2024

Germán Jesús Pereira Muñoz PhD.

13

13

---

---

---

---

---


---

---

---

### Enfoque de Base de Datos

- Para lograr un efectivo tratamiento del uso de Bases de Datos que permiten tener un conjunto de datos relacionados y almacenados en forma permanente y usados con variados propósitos por múltiples usuarios y que permiten:
  - **Integrar**: significa que los diferentes archivos de datos han sido lógicamente organizados para reducir la redundancia de datos y facilitar el acceso a ellos.
  - **Compartir**: significa que todos los usuarios calificados tienen acceso a los mismos datos, para usarlos en diferentes actividades.
- En una Base de Datos este conjunto de datos relacionados es almacenados en forma permanente y usados con variados propósitos por múltiples usuarios



9/18/2024

Germán Jesús Pereira Muñoz PhD.

14

14

---

---

---

---

---

---


---

---

### Información

Información son datos que han sido organizados o preparados en una forma adecuada para apoyar la toma de decisiones:

- Por ejemplo, una lista de productos y su stock sin ningún orden son datos, pero una lista de productos ordenados por stock (de menor a mayor) representa información para el encargado de compras de un supermercado.



9/18/2024

15

---

---

---

---

---

---

---

---

## Discusión

¿Puede en un momento dado un objeto considerarse como dato y en otro momento como información?

¿Por qué?

rut	ramo	nota1	nota2	nota3	nota4	final
9530897	CS78-21	6,5	4,8	6,0	6,3	5,9
DATO						INFORMACIÓN

rut	ramo	nota1	nota2	nota3	nota4	final
95308973	CS78-21	6,5	4,8	6,0	6,3	5,9
12349456K	CS78-21	4,5	5,6	5,5	6,5	5,5
158902356	CS78-21	6,2	6,1	5,2	5,8	5,8
DATO						INFORMACIÓN
Promedio de curso						5,8

9/18/2024

Germán Jesús Pereira Muñoz PhD.

16

16

---

---

---

---

---

---

---

---

---

---

## Ciclo de Vida



- **Análisis** plantea la definición de objetivos y genera una especificación de requerimientos en base a las necesidades de la organización.
- **Diseño y Modelamiento de Datos** Como resultado de los procesos de análisis se definen en forma creciente (en términos de especificidad) los elementos que constituirán el sistema de información a implementar. esta etapa define los objetos (entidades) que deben ser utilizados.

Germán Jesús Pereira Muñoz PhD.

17

17

---

---

---

---

---

---

---

---

---

---

## Ciclo de Vida



**Implementación** Según los objetivos a realizar y los objetos definidos en la etapa previa darán paso a la selección de la base de datos y metodología a implementar para la captura de datos y la generación de información. Esta etapa debe realizarse en base a varias premisas para determinar:

- implementación on-premise o en cloud.
- uso de servicios cloud (GCP, AWS u otros).
- modelo de bases de datos: relacional, no-relacional.
- niveles de privacidad.
- interoperabilidad.

**Operación y captura de datos** Operación o Utilización de la Base de Datos: la cual involucra a los usuarios finales acceder constantemente, y a los desarrolladores de sistemas realizando programas que permitan mantenerla actualizada y responder a nuevos requerimientos de los usuarios.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

18

18

---

---

---

---

---

---

---

---

---

---

### Modelamiento de Datos

- Los modelos facilitan la comunicación entre el diseñador de base de datos y los usuarios finales.
- Los modelos son fáciles de utilizar y cambiar, ya que son sólo una imagen muy simplificada del sistema de información que se desea desarrollar.
- Actualmente en la participación de la construcción del modelo de datos se involucran muchos actores.

```
graph TD; A[Especificación de requisitos] --> B[DISEÑO CONCEPTUAL]; B --> C[Esquema conceptual]; C --> D[DISEÑO LÓGICO]; E[Normalización] --> D; D --> F[Esquema lógico]; F --> G[DISEÑO FÍSICO]; G --> H[Esquema físico];
```

9/18/2024

Germán Jesús Pereira Muñoz PhD.

19

19

---

---

---

---

---

---

---

---

### Modelamiento de Datos

- **Conceptual:** incluye la identificación de las entidades del sistema, claves de nivel superior y sus relaciones, que definen el ámbito del problema que tratará el sistema. Elementos de modelado del perfil UML.
- **Lógica:** incluye el perfeccionamiento de las entidades del sistema de alto nivel de la fase conceptual en entidades lógicas más detalladas. Se pueden definir, en un modelo lógico de datos mediante la utilización de los elementos de modelado del perfil UML para el diseño de bases de datos.
- **Física:** incluye la transformación de los diseños de la clase lógica en diseños de tablas de bases de datos físicas detalladas y optimizadas.

```
graph TD; A[Especificación de requisitos] --> B[DISEÑO CONCEPTUAL]; B --> C[Esquema conceptual]; C --> D[DISEÑO LÓGICO]; E[Normalización] --> D; D --> F[Esquema lógico]; F --> G[DISEÑO FÍSICO]; G --> H[Esquema físico];
```

9/18/2024

Germán Jesús Pereira Muñoz PhD.

20

20

---

---

---

---

---

---

---

---

### Calidad de Datos

No existen estandarizaciones en lo que se refiere a Calidad de datos (**Data Quality**).

Mantener la exactitud y la integridad de todos los tipos de datos en toda la organización es trabajar por su aptitud para cumplir con su propósito en un contexto dado.

Es la cualidad de un conjunto de información recogida en una base de datos, un sistema de información que reúne entre sus atributos los siguientes:

- Exactitud, Complejidad, Integridad, Actualización
- Coherencia, Relevancia, Accesibilidad, Confiabilidad

9/18/2024

Germán Jesús Pereira Muñoz PhD.

21

21

---

---

---

---

---

---

---

---

**Aseguramiento de la Calidad de los Datos** proceso de verificación de la fiabilidad y efectividad de los datos, que debe realizarse periódicamente.

- Este proceso recupera los datos en bruto y comprueba su calidad, elimina los duplicados y, cuando es posible, corrige los valores erróneos y completa los valores vacíos, es decir se transforman los datos **-siempre que sea posible-** para reducir los errores de carga.
- Se obtienen datos limpios y de alta calidad.
- Entre los tipos de datos que mayores efectos adversos pueden provocar en términos de calidad se encuentran los tres siguientes:

**Datos oscuros, Datos sucios, Datos No estructurados**

9/18/2024

Germán Jesús Pereira Muñoz PhD.

22

22

---

---

---

---

---

---

---

---

- **Datos oscuros** se recopilan, procesan y almacenan como parte de las actividades comerciales cotidianas, pero que la organización no utiliza con ningún otro fin.

*Su existencia revela que el sistema de calidad de datos de la empresa no es el óptimo ni está lo suficientemente avanzado, puesto que permite que a la recopilación y gestión de datos les falte eficiencia y eficacia.*

- **Datos sucios** en modo local este tipo de datos supone un grave problema de calidad, en un entorno como la nube aún más, en especial en lo que respecta al Internet de las Cosas (IoT).

*En el ámbito de sistemas automatizados, los datos sucios pueden causar a la organización un daño real, al obligarla a incurrir en un costo económico real causado por las acciones automáticas que dan inicio con datos que no son válidos.*

9/18/2024

Germán Jesús Pereira Muñoz PhD.

23

23

---

---

---

---

---

---

---

---

**Datos No estructurados** son de relativa importancia sobre todo dentro de sistemas relacionales donde todos los datos deben ser estructurados.

- Reciben menos importancia en sistemas **NoSQL** que tienen capacidad de recibir datos más heterogéneos.
- En ocasiones, los datos están disponibles, pero no están preparados para su uso.
- Deben ser enriquecidos de alguna manera para poder considerarse compatibles con el sistema que los va a consumir. Si no se hace, los problemas de calidad empezarán a aparecer.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

24

24

---

---

---

---

---

---

---

---



### Limpieza de datos

- Limpieza de datos siempre está asociado a los objetivos determinados en el proceso de captura de datos.
- Un buen diseño inicial de captura de datos que esta alineado con los requerimientos establecidos, hará prácticamente innecesario un proceso que es bastante costoso en recursos humanos.
- La limpieza de los datos incluye 5 actividades principales:
  - Depurar
  - Corregir
  - Estandarizar
  - Relacionar
  - Consolidar

9/18/2024

Germán Jesús Pereira Muñoz PhD.

25

25

---

---

---

---

---

---

---

---

### Depurar

- Según las necesidades es requisito depurar los datos obtenido sobre todo si estos difieren de la forma a ser tratados posteriormente.
- Siempre será necesaria cuando no haya congruencia entre el método de captura y el dato definido a capturar.

nombre		direccion				
CARLOS IVAN ARANEDA SOTO		COSTA SAN FRANCISCO 890 DEPARTAMENTO SS-D VILLA SAN PEDRO CONCEPCION				

nombres	apellidos	calle	numero	depto	sector	comuna
CARLOS IVAN	ARANEDA SOTO	COSTA SAN FRANCISCO	890	SS-D	VILLA SAN PEDRO	CONCEPCION

paterno	materno
ARANEDA	SOTO

9/18/2024

Germán Jesús Pereira Muñoz PhD.

26

26

---

---

---

---

---

---

---

---

### Corregir

- Este proceso corrige los valores individuales de los atributos usando algoritmos de corrección y fuentes de datos externas.
- Por ejemplo: comprueba una dirección y el código postal correspondiente.

nombres	apellidos	genero	calle	numero	sector	comuna
CARLOS IVAN	ERNADEZ	F	COSTA SAN FRANCISCO	890	VILLA SAN PEDRO	CONCEPCION

nombres	apellidos	genero	calle	numero	sector	comuna
CARLOS IVAN	HERNANDEZ	M	COSTA SAN FRANCISCO	890	VILLA SAN PEDRO	SAN PEDRO DE LA PAZ

9/18/2024

Germán Jesús Pereira Muñoz PhD.

27

27

---

---

---

---

---

---

---

---

### Estandarizar

- Este proceso aplica rutinas de conversión para transformar valores en formatos definidos (y consistentes) aplicando procedimientos de estandarización y definidos por las reglas del negocio.
- Por ejemplo: trato de Sr., Sra., etc. o sustituyendo los diminutivos de nombres por los nombres correspondientes.

nombres	apellidos	genero	calle	numero	sector	comuna
CARLOS IVAN	ERNAÑDEZ	H	O'HIGGINS	No. 890	CENTRO	SANTIAGO
JUAN	VERA	M	BERNARDO O'HIGGINS	700	MAIPU	STGO
MARCELA	VARAS SOLIS	F	ALAMEDA BERNARDO O'HIGGINS	#660	CENTRO	SANTIAGO

nombres	apellidos	genero	calle	numero	sector	comuna
			ALAMEDA BERNARDO O'HIGGINS	890		SANTIAGO



9/18/2024

28

28

---

---

---

---

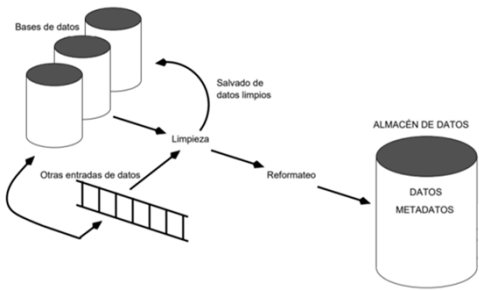
---

---

---

---


### Funcionamiento de un almacén de datos



9/18/2024

Germán Jesús Pereira Muñoz PhD.

29



29

---

---

---

---






---

---

---

---

### El Modelo Relacional




- Para el modelado y la gestión de bases de datos, es un modelo de datos basado en la lógica de predicados y en la teoría de conjuntos.
- Desarrolla un esquema de base de datos (data base schema) a partir del cual se podrá realizar el modelo físico o de implementación en el DBMS.
- Esta basado en que todos los datos están almacenados en tablas (entidades/relaciones) y cada una de estas es un conjunto de datos, por tanto una base de datos es un conjunto de relaciones.
- La agrupación se origina en la tabla: tabla -> fila (tupla) -> campo (atributo)

9/18/2024

Germán Jesús Pereira Muñoz PhD.

30



30

---

---

---

---

---

---

---

---

## El Modelo Relacional

El Modelo Relacional se ocupa de:

- La estructura de datos
- La manipulación de datos
- La integridad de los datos



Donde las relaciones están formadas por :

- Atributos (columnas)
- Tuplas (Conjunto de filas)

Existen dos formas para la construcción de modelos relacionales:

- Creando un conjunto de tablas iniciales y aplicando operaciones de normalización hasta conseguir el esquema óptimo,
- Convertir el modelo entidad relación (ER) en tablas, con una depuración lógica y la aplicación de restricciones de integridad

9/18/2024

Germán Jesús Pereira Muñoz PhD.

31

31

## Objetivos del modelo relacional

- **Independencia Física:** La forma de almacenar los datos no debe influir en su manipulación. Si el almacenamiento físico cambia, los usuarios que acceden a esos datos no tienen que modificar sus aplicaciones.
- **Independencia Lógica:** Las aplicaciones que utilizan la base de datos no deben ser modificadas por que se inserten, actualicen y eliminen datos.
- **Flexibilidad:** En el sentido de poder presentar a cada usuario los datos de la forma en que éste prefiera
- **Uniformidad:** Las estructuras lógicas de los datos siempre tienen una única forma conceptual (las tablas), lo que facilita la creación y manipulación de la base de datos por parte de los usuarios.
- **Sencillos:** Las características anteriores hacen que este Modelo sea fácil de comprender y de utilizar por parte del usuario final.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

32

32

## Definiciones

- **Relación:** Tabla bidimensional para la representación de datos. Ejemplo: Estudiantes.
- **Tuplas:** Filas de una relación que contiene valores para cada uno de los atributos (equivale a los registros). Ejemplo: 34563, José, Martínez, 19, Masculino. Representa un objeto único de datos implícitamente estructurados en una tabla. Un registro es un conjunto de campos que contienen los datos que pertenecen a una misma entidad.
- **Atributos:** Columnas de una relación y describe las características particulares de cada campo. Ejemplo: id estudiante
- **Esquemas:** Forma de representar una relación y su conjunto de atributos. Ejemplo: Estudiantes (id estudiante, nombre(s), apellido(s), edad, género)

9/18/2024

Germán Jesús Pereira Muñoz PhD.

33

33

Definiciones

- **Claves:** Campo cuyo valor es único para cada registro. Principal, identifica una tabla, y Foránea, clave principal de otra tabla relacionada. Ejemplo: id estudiante.
- **Clave Primaria:** identificador único de una tupla.
- **Cardinalidad:** número de tuplas(m).
- **Grado:** número de atributos(n).
- **Dominio:** colección de valores de los cuales el atributo obtiene su atributo.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

34

34

---

---

---

---

---

---

---

---

Terminología Relacional	Terminología de Tablas	Terminología de Archivo
Relación	■ Tabla	■ Archivo
Tupla	■ Fila	■ Registro
Atributo	■ Columna	■ Campo
Grado	■ Número de columnas	■ Número de campos
Cardinalidad	■ Número de filas	■ Número de registros

EMPLEADO ← Nombre de la Relación

Clave Primaria

pasaporte	paombre	apaterno	apmaterno	fono	fecimiento
12095444	Alberto	Ólmez	Martínez	2345676	20/11/1969
9509590	Luisa	Jordán	Soto	3344567	12/09/2000
19456873	Cristian	Muñoz	Pereira	4567912	12/10/2010
20345765	Josefina	Carvajal	Durán	3456835	05/06/2011
15687490	Marcos	Ramírez	Ponce		28/02/1978

Cardinalidad

Atributos

Tuplas

Grado

9/18/2024

Germán Jesús Pereira Muñoz PhD.

35

35

---

---

---

---

---

---

---

---

Reglas de Integridad

Reglas o restricciones de validación que controlan que los datos a registrar sean correctos.

- **Integridad de Dominio:** Conjunto de valores válidos de un campo (propiedades del campo)
- **Integridad de Transiciones:** Define los estados por lo que un registro puede pasar válidamente (operación previa)
- **Integridad de Entidades:** Asegura la integridad de las tablas (claves, identificación)
- **Integridad Referencial:** Mantienen y protegen vínculos entre tablas (propiedades de las relaciones)
- **Integridad de Bases de Datos:** Referencian más de una tabla, gobiernan la DB como un todo.
- **Integridad de Transacciones:** Controlan la forma como se manipulan los datos entre una o varias BD.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

36

36

---

---

---

---

---

---

---

---

## Atributo

- Un Atributo en el Modelo Relacional representa una propiedad que posee esa Relación y equivale al atributo del Modelo E-R.
- Se corresponde con la idea de campo o columna.
- En el caso de que sean varios los atributos de una misma tabla, definidos sobre el mismo dominio, habrá que darles nombres distintos, ya que una tabla no puede tener dos atributos con el mismo nombre.

oficina	calle	area	telefono	fax
100	Lyon 2345	Las Condes	964201240	964201340
110	Alameda 234	Santiago Centro	964215760	964215670
120	Luis Thayer Ojeda	Providencia	964520250	964520255
130	Baldomero Lillo 2345	Puente Alto	964284440	
140	Calle Crucero 3456	La Dehesa	965678904	964252811

9/18/2024

Luis Thayer Ojeda PhD.

37

37

---

---

---

---

---

---

---

---

---

---

## Dominio

- El dominio dentro de la estructura del Modelo Relacional es el **conjunto de valores que puede tomar un atributo**.
- Un dominio contiene todos los posibles valores que puede tomar un determinado atributo. Dos atributos distintos pueden tener el mismo dominio.
- Un dominio es un **conjunto finito de valores del mismo tipo**.
- Los dominios poseen un nombre para poder referirnos a él y así poder ser reutilizable en más de un atributo.

Atributo	Nombre del Dominio	Descripción	Definición
oficina	NUM_OFICINA	Posibles valores de número de oficina	3 caracteres, rango 100 - 990
calle	NOM_CALLE	Nombres de calles y número de Santiago donde se ubica la oficina	25 caracteres
area	NOM_AREA	Área de Santiago en la que se encuentra ubicada la oficina	20 caracteres
telefono	NUM_TEL_FAX	Números de teléfono de Santiago	9 caracteres
fax	NUM_TEL_FAX	Números de teléfono de Santiago	9 caracteres

9/18/2024

Luis Thayer Ojeda PhD.

38

38

---

---

---

---

---

---

---

---

---

---

## Tupla, Grado y Cardinalidad

- **Tupla**: es cada una de las filas de la relación. Representa por tanto el conjunto de cada elemento individual (ejemplar o ocurrencia) de esa tabla. En la relación OFICINA, cada tupla tiene cinco valores, uno para cada atributo. Las tuplas de una relación no siguen ningún orden.
- **Grado**: número de columnas de la relación (número de atributos). La relación OFICINA es de grado seis porque tiene seis atributos. Esto quiere decir que cada fila de la tabla es una tupla con seis valores.
- **Cardinalidad**: número de tuplas de una relación (número de filas). Ya que en las relaciones se van insertando y borrando tuplas a menudo, la cardinalidad de las mismas varía constantemente.

pasaporte	pnombre	apaterno	apmaterno	fono	fnacimiento
12095444	Alberto	Gómez	Martínez	2345676	20/11/1969
9509590	Luisa	Jordán	Soto	3344567	12/09/2000
19456873	Cristian	Muñoz	Pereira	4567912	12/10/2010
20345765	Josefina	Carvajal	Durán	3456835	05/06/2011
15687490	Marcos	Ramírez	Ponce		28/02/1978

Cardinalidad

Tuplas

9/18/2024

Luis Thayer Ojeda PhD.

Grado

39

39

---

---

---

---

---

---


---

---

---

---

**SQL**



SQL es un lenguaje estructurado de consulta (Structured Query Language, en inglés), diseñado específicamente para administrar información en sistemas de gestión de bases de datos de tipo relacional.

Puede considerarse un lenguaje de programación como tal, ya que cuenta con uso de variables, tipos de datos, elementos condicionales y lógicos.

Es el estándar de facto para la gestión de datos y permite:

- Consultar, actualizar y reorganizar datos
- Crear y modificar la estructura de datos
- Controlar el acceso a los datos

El uso de este lenguaje es altamente imperativo para todo profesional que tenga por objetivo acceder a altos volúmenes y/o datos complejos

9/18/2024 Germain Jesús Pereira Muñoz PhD. 40

40

---

---

---

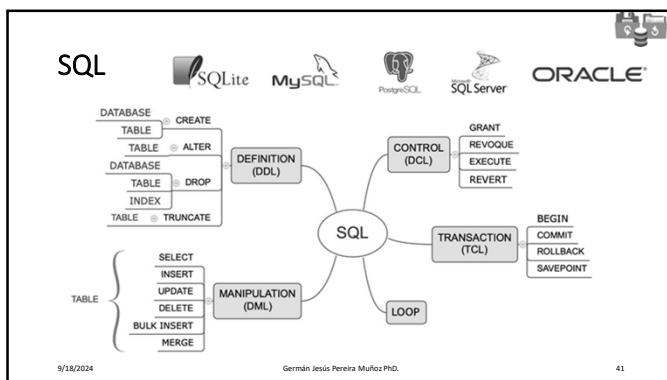
---

---

---

---

---



41

---

---

---

---


---

---

---

---

**NoSQL**



El modelo NoSQL, por “not only SQL” es un modelo de almacenamiento de datos que a diferencia del modelo relacional no es tabular, sino que permite diversos tipos según el modelo de datos.

Los principales son documentos, pares clave-valor (key-value), wide-column, grafos. Comparten en común que son bastante más flexibles a la hora de almacenar datos y son más fáciles de escalar con grandes volúmenes de datos.

Tipos de bases de datos NoSQL

1. Bases de datos clave-valor
2. Bases de datos documentales
3. Bases de datos en grafo
4. Bases de datos orientadas a objetos

9/18/2024 Germain Jesús Pereira Muñoz PhD. 42

42

---

---

---

---


---

---

---

---

## NoSQL





### Tipos de bases de datos NoSQL

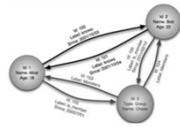
1. Bases de datos clave-valor
2. Bases de datos documentales
3. Bases de datos en grafo
4. Bases de datos orientadas a objetos

**Keys**

**Values**







9/18/2024      Germán Jesús Pereira Muñoz PhD.      43

43

---

---

---

---

---

---


---

---

---

---

## Similitudes entre NoSQL y SQL



Ambos tipos de bases de datos, tanto NoSQL como SQL, tienen el mismo objetivo principal: almacenar y administrar datos de manera eficiente y segura para ser utilizados en aplicaciones y sistemas.

- **Lenguajes de programación:** A pesar de que SQL utiliza un lenguaje de consulta estructurado, ambos tipos de bases de datos pueden interactuar con una amplia variedad de lenguajes de programación, como Python, Java, JavaScript, entre otros.
- **Soporte de la comunidad:** Tanto SQL como NoSQL cuentan con un amplio soporte de la comunidad de desarrolladores, lo que facilita la resolución de problemas y la obtención de recursos y documentación.
- **Herramientas y servicios:** Existen numerosas herramientas y servicios disponibles para trabajar con bases de datos SQL y NoSQL, lo que permite a los desarrolladores elegir la opción que mejor se adapte a sus necesidades y preferencias.

9/18/2024      Germán Jesús Pereira Muñoz PhD.      44

44

---

---

---

---

---

---


---

---

---

---

## Diferencias entre NoSQL y SQL



1. **Modelo de datos:** La principal diferencia es el modelo de datos que utilizan.
  - **SQL** emplea un modelo relacional, en el que los datos se almacenan en tablas y se relacionan entre sí.
  - **NoSQL**, en cambio, utiliza varios modelos de datos, como documentos, pares clave-valor, columnas y grafos, lo que permite una mayor flexibilidad y escalabilidad.
2. **Esquema:**
  - **SQL** requiere un esquema predefinido, lo que significa que debes especificar la estructura de las tablas y las relaciones entre ellas antes de agregar datos.
  - **NoSQL** es más flexible en este aspecto, ya que permite agregar y modificar datos sin necesidad de definir un esquema previamente.
3. **Consultas:**
  - **SQL** utiliza un lenguaje de consulta estructurado y uniforme (SQL) para interactuar con la base de datos.
  - **NoSQL**, no tiene un lenguaje de consulta universal, y las consultas pueden variar según el tipo de base de datos NoSQL que se utilice.

9/18/2024      Germán Jesús Pereira Muñoz PhD.      45

45

---

---

---

---

---

---

---

---

---

---

Diferencias entre NoSQL y SQL

3. Transacciones y consistencia:

- SQL se adhiere al modelo ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad), lo que garantiza que las transacciones sean confiables y seguras.
- NoSQL, en cambio, sigue el teorema CAP (Consistencia, Disponibilidad y Tolerancia a la partición), lo que significa que puede ser necesario sacrificar una de estas propiedades para obtener las otras dos.

4. Escalabilidad:

- SQL tiende a escalar verticalmente, lo que significa que debes agregar más recursos a un único servidor para manejar una mayor cantidad de datos.
- NoSQL, en cambio, escala horizontalmente, lo que permite distribuir los datos en múltiples servidores, facilitando la gestión de grandes volúmenes de información.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

46

46

---

---

---

---

---

---

---

---

Como elegir entre NoSQL y SQL

La elección entre NoSQL y SQL dependerá en gran medida de las necesidades y objetivos de tu proyecto. A continuación, presentamos algunas consideraciones para ayudarte a tomar una decisión informada:

- Si necesitas un esquema estricto y relaciones bien definidas entre los datos, SQL podría ser la mejor opción.
- Si buscas flexibilidad y escalabilidad, especialmente para manejar grandes volúmenes de datos, NoSQL podría ser la mejor elección.
- Si la consistencia y la integridad de los datos son cruciales para tu proyecto, SQL podría ser más adecuado.
- Si la disponibilidad y la tolerancia a fallos son factores importantes, NoSQL podría ser la solución ideal.

9/18/2024

Germán Jesús Pereira Muñoz PhD.

47

47

---

---

---


---

---

---

---

---



9/18/2024

Germán Jesús Pereira Muñoz PhD.

48

48

---

---

---

---

---

---

---

---

16