

# Repaso de Base de Datos



# Repaso de Base de Datos

## I. Fundamento de Base de Datos - Repaso

### ¿Qué es un dato?

Un dato es información concreta sobre hechos, elementos, etc., que permite estudiarlos, analizarlos o conocerlos. Además, registrarlos y que tienen un significado implícito.

### ¿Qué es una base de datos?

Una base de datos es una colección coherente de datos relacionados sobre un tema, objetivo determinado, etc.

Se estructura de manera lógica que facilite su organización para evitar redundancias e inconsistencias, y que la misma vez permita consultarla de forma ágil y rápida. Cada dato pertenece a un tipo de datos (Int, float, char,String...)

### Características de las bases de datos

- a. Almacenamiento de gran cantidad de datos, que están relacionados entre si.
- b. Gestión segura de los datos.
- c. Acceso compartido y simultáneo por los usuarios.
- d. Lenguaje de definición de datos. (DDL siglas en inglés)

Este lenguaje permite a los programadores de un sistema gestor de base de datos definir las estructuras que almacenarán los datos así como los procedimientos o funciones que permitan consultarlos.



# Repaso de Base de Datos

## Características de las bases de datos

### d. Lenguaje de definición de datos. (DDL siglas en inglés) cont...

Para definir la estructura disponemos de tres sentencias:

- ✓ **CREATE**, se usa para crear una base de datos, tabla, vistas, etc.
- ✓ **ALTER**, se utiliza para modificar la estructura, por ejemplo añadir o borrar columnas de una tabla.
- ✓ **DROP**, con esta sentencia, podemos eliminar los objetos de la estructura, por ejemplo un índice o una secuencia.

### e. Lenguaje de manipulación de datos. (DML siglas en inglés)

Utilizando instrucciones de SQL, permite a los usuarios introducir datos para posteriormente realizar tareas de consultas o modificación de los datos que contienen las Bases de Datos.

Los elementos que se utilizan para manipular los datos, son los siguientes:

- ✓ **SELECT**, esta sentencia se utiliza para realizar consultas sobre los datos.
- ✓ **INSERT**, con esta instrucción podemos insertar los valores en una base de datos.
- ✓ **UPDATE**, sirve para modificar los valores de uno o varios registros.
- ✓ **DELETE**, se utiliza para eliminar las filas de una tabla.



# Repaso de Base de Datos

## Características de las bases de datos

### f. Lenguaje de Control de datos . (DCL siglas en inglés)

Es un lenguaje que incluye una serie de comandos SQL proporcionado por los sistemas gestores de bases de datos.

Estos comandos permiten al Administrador del sistema gestor de base de datos, controlar el acceso a los objetos, es decir, podemos otorgar o denegar permisos a uno o más roles para realizar determinadas tareas.

Los comandos para controlar los permisos son los siguientes:

- ✓ **GRANT**, permite otorgar permisos.
- ✓ **REVOKE**, elimina los permisos que previamente se han concedido.



# Repaso de Base de Datos

## II. Modelo Entidad-Relación

El modelo entidad-relación es una representación de la estructura de base de datos relacionales (o de representar los esquemas de estas).

Como herramienta, permite representar de manera simplificada los componentes que participan en un proceso de negocio y el modo en el que estos se relacionan entre sí.

Además de disponer de un diagrama que ayuda a entender los datos y como se relacionan entre ellos, debe de ser completado con un pequeño resumen con la lista de los atributos y las relaciones de cada elemento.

### Modelo entidad-relación y sus elementos

#### ✓ Entidad

Las entidades representan cosas u objetos (ya sean reales o abstractos), que se diferencian claramente entre sí. Por ejemplo **empleado** (objeto físico), **carro** (objeto físico), **taller** (objeto físico), etc.

Estas entidades se representan en un diagrama con un rectángulos, como los siguientes:

empleado

taller

carro

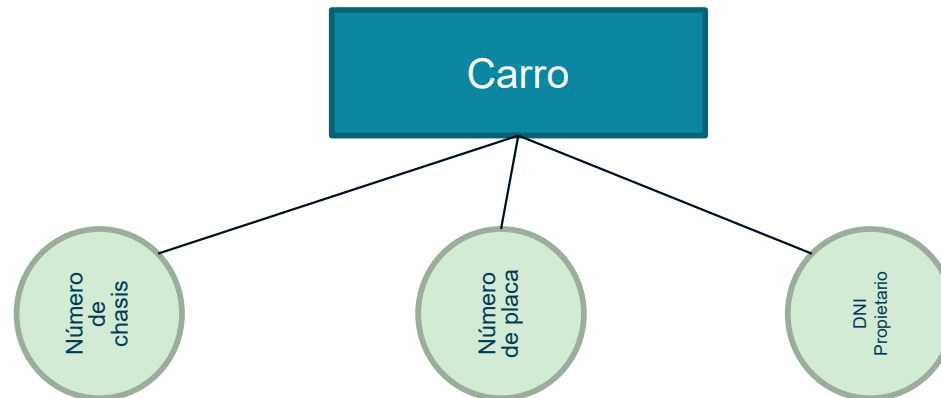
# Repaso de Base de Datos

## ✓ Atributos

Los atributos definen o identifican las características de entidad (es el contenido de esta entidad). Cada entidad contiene distintos atributos, que dan información sobre esta entidad. Estos atributos pueden ser de distintos tipos (numéricos, texto, fecha...).

Por ejemplo, para la entidad carro, podemos definir los siguientes atributos: *número de chasis*, *número de placa*, *identidad del propietario*, *marca*, *modelo* y muchos otros que complementen la información de cada coche.

Los atributos se representan como círculos que descienden de una entidad, y no es necesario representarlos todos, sino los más significativos, como a continuación.



# Repaso de Base de Datos

## ✓ Relación

Es un vínculo que nos permite definir una dependencia entre varias entidades, es decir, nos permite exigir que varias entidades compartan ciertos atributos de forma indispensable. Por ejemplo, al empleado se le asigna la reparación del carro en el taller. Las relaciones se muestran en los diagramas como rombos, que se unen a las entidades mediante líneas.



## ➤ Relaciones de cardinalidad

Podemos encontrar distintos tipos de relaciones según como participen en ellas las entidades. Es decir, en el caso anterior cada empleado puede tener al cargo una reparación, pero una misma reparación puede estar a cargo de varios empleados.

Esto complementa a las representaciones de las relaciones, mediante un intervalo en cada extremo de la relación que especifica cuantos objetos o cosas (de cada entidad) pueden intervenir en esa relación.

# Repaso de Base de Datos

## ➤ Relaciones de cardinalidad. Cont...

**Uno a uno:** Una entidad se relaciona únicamente con otra y viceversa. Por ejemplo, si tuviésemos una entidad con distintos chasis y otra con matrículas deberíamos de determinar que cada chasis solo puede tener una matrícula (y cada matrícula un chasis, ni más en ningún caso).



**Uno a varios o varios a uno:** determina que un registro de una entidad puede estar relacionado con varios de otra entidad, pero en esta entidad debe existir solo una vez.





# Repaso de Base de Datos

## ➤ Relaciones de cardinalidad. Cont...

**Varios a varios:** determina que una entidad puede relacionarse con otra con ninguno o varios registros y viceversa. Por ejemplo, en el taller un coche puede ser reparado por varios mecánicos distintos y esos mecánicos pueden reparar varios coches distintos.



Los indicadores numéricos indican el primero el número mínimo de registros en una relación y posteriormente el máximo (si no hay límite se representa con una "n").

# Repaso de Base de Datos

## ✓ Llave o Clave

Es el atributo de una entidad, al que le aplicamos una restricción que lo distingue de los demás registros (no permitiendo que el atributo específico se repita en la entidad) o le aplica un vínculo (exactamente como comentábamos en las relaciones). Estos son los distintos tipos:

- ❖ **Super llave:** aplica una llave o restricción a varios atributos (subconjunto) de la entidad, para así asegurarse que en su conjunto no se repitan varias veces y así no poder entrar en dudas al querer identificar un registro.
- ❖ **Llave primaria:** identifica inequívocamente un solo atributo no permitiendo que se repita en la misma entidad. Como sería la placa o el número de chasis de un carro (no puede existir dos veces el mismo)..
- ❖ **Llave externa o llave foránea:** este campo tiene que estar estrictamente relacionado con la llave primaria de otra entidad, para así exigir que exista previamente ese llave.

# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL

### ¿Qué es el Lenguaje SQL?

SQL (por sus siglas en inglés) o lenguaje de consulta estructurado, es un lenguaje de dominio específico, diseñado para administrar, y recuperar información de sistemas de gestión de base de datos relacionales.

Una de sus principales características es el manejo del álgebra y el cálculo relacional para efectuar consultas con el fin de recuperar, de forma sencilla, información de bases de datos, así como realizar cambios en ellas.

Este consiste en un lenguaje de definición de datos (DDL), un lenguaje de manipulación de datos (DML) y un lenguaje de control de datos (DCL). El alcance de SQL incluye la inserción de datos, consultas, actualizaciones y borrado, la creación y modificación de esquemas y el control de acceso a los datos. También el SQL a veces se describe como un lenguaje declarativo, también incluye elementos procesales.

SQL fue uno de los primeros lenguajes comerciales para el modelo relacional de Edgar Frank Codd como se describió en su artículo de investigación de 1970. El modelo relacional de datos para grandes bancos de datos compartidos. A pesar de no adherirse totalmente al modelo relacional descrito por Codd, pasó a ser el lenguaje de base de datos más usado.

SQL pasó a ser el estándar del Instituto Nacional Estadounidense de Estándares (ANSI) en 1986 y de la Organización Internacional de Normalización (ISO) en 1987. Desde entonces, el estándar ha sido revisado para incluir más características. A pesar de la existencia de ambos estándares, la mayoría de los códigos SQL no son completamente portables entre sistemas de bases de datos diferentes sin otros ajustes.

Los orígenes de SQL están ligados a las bases de datos relacionales, específicamente las que residían en máquinas IBM bajo el sistema de gestión System R, desarrollado por un grupo de la IBM en San José, California.

# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL cont...

### Características generales del lenguaje SQL

SQL es un lenguaje de acceso a bases de datos que explota la flexibilidad y potencia de los sistemas relacionales y permite así gran variedad de operaciones.

Es un lenguaje declarativo de "alto nivel" o "de no procedimental" que, gracias a su fuerte base teórica y su orientación al manejo de conjuntos de registros relacionales, permite una alta productividad en codificación y la orientación a objetos. De esta forma, una sola sentencia puede equivaler a uno o más programas que se utilizarían en un lenguaje de bajo nivel orientado a registros.

Tiene las siguientes características:

- **Lenguaje de definición de datos:** El DDL de SQL proporciona comandos para la definición de esquemas de relación, borrado de relaciones y modificaciones de los esquemas de relación.
- **Lenguaje interactivo de manipulación de datos:** El DML de SQL incluye lenguajes de consultas basado tanto en álgebra relacional como en cálculo relacional de tuplas.
- **Integridad:** El DDL de SQL incluye comandos para especificar las restricciones de integridad que deben cumplir los datos almacenados en la base de datos.
- **Definición de vistas:** El DDL incluye comandos para definir las vistas.
- **Control de transacciones:** SQL tiene comandos para especificar el comienzo y el final de una transacción.
- **SQL incorporado y dinámico:** Esto quiere decir que se pueden incorporar instrucciones de SQL en lenguajes de programación como: C++, C, Java, PHP, COBOL, Pascal y Fortran.



# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL cont...

### Características generales del lenguaje SQL cont..

- **Autorización:** El LDD incluye comandos para especificar los derechos de acceso a las relaciones y a las vistas.

### Tipos de datos

Algunos de los tipos de datos básicos de SQL son:

- **Números enteros:**
  - **TINYINT(tamaño):** -128 a 127 normal. 0 a 255.
  - **SMALLINT(tamaño):** -32768 a 32767 normal. 0 a 65535.
  - **MEDIUMINT(tamaño):** -8388608 a 8388607 normal. 0 a 16777215.
  - **INT(tamaño):** -2147483648 a 2147483647 normal. 0 a 4294967295.
  - **BIGINT(tamaño):** -9223372036854775808 a 9223372036854775807 normal. 0 a 18446744073709551615.
- **Números en punto flotante:**
  - **FLOAT(tamaño, d):** Un pequeño número con un punto decimal flotante.
  - **DOUBLE (tamaño, d):** Un número grande con un punto decimal flotante.
  - **DECIMAL (tamaño, d):** Un DOUBLE almacenado como una cadena, lo que permite un punto decimal fijo.



# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL cont...

### Tipos de datos cont...

#### ➤ Fechas y marca de tiempo (timestamp):

- DATE (): Una fecha. Formato: AAAA-MM-DD Nota: el rango admitido es de '1000-01-01' a '9999-12-31'
- DATETIME (): \* Una combinación de fecha y hora. Formato: AAAA-MM-DD HH: MI: SS Nota: el rango admitido es de '1000-01-01 00:00:00' a '9999-12-31 23:59:59'
- TIMESTAMP (): \* Una marca de tiempo. Los valores de TIMESTAMP se almacenan como el número de segundos desde la época de Unix ('1970-01-01 00:00:00' UTC). Formato: AAAA-MM-DD HH: MI: SS Nota: el rango admitido es de '1970-01-01 00:00:01' UTC a '2038-01-09 03:14:07' UTC
- TIME (): Un tiempo. Formato: HH: MI: SS Nota: el rango admitido es de '-838: 59: 59' a '838: 59: 59'
- YEAR (): Un año en formato de dos o cuatro dígitos. Nota: Valores permitidos en formato de cuatro dígitos: de 1901 a 2155. Valores permitidos en formato de dos dígitos: 70 a 69, que representan los años de 1970 a 2069

# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL cont...

### Tipos de datos cont...

#### ➤ Cadena de caracteres:

- **CHAR (tamaño):** Tiene una cadena de longitud fija (puede contener letras, números y caracteres especiales). Puede almacenar hasta 255 caracteres
- **VARCHAR (tamaño):** Tiene una cadena de longitud variable (puede contener letras, números y caracteres especiales). Puede almacenar hasta 255 caracteres. Nota: si agrega un valor mayor que 255, se convertirá en un tipo de texto
- **TINYTEXT:** Tiene una cadena con una longitud máxima de 255 caracteres
- **TEXT:** Tiene una cadena con una longitud máxima de 65.535 caracteres
- **BLOB:** Para BLOB (Objetos grandes binarios). Almacena hasta 65.535 bytes de datos
- **MEDIUMTEXT:** Tiene una cadena con una longitud máxima de 16,777,215 caracteres
- **MEDIUMBLOB:** Para BLOB (Objetos grandes binarios). Tiene capacidad para 16.777.215 bytes de datos.
- **LONGTEXT:** Tiene una cadena con una longitud máxima de 4.294.967.295 caracteres
- **LOB:** Para BLOB (Objetos grandes binarios). Tiene capacidad para 4.294.967.295 bytes de datos..



# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL cont...

### Tipos de datos cont...

#### ➤ Enum y set:

- **Enum (x, y, z, etc.):** Permite ingresar una lista de valores posibles.
- **Set:** Similar a ENUM, excepto que SET puede contener hasta 64 elementos de lista y puede almacenar más de una opción

#### ➤ Binarios:

- bit: Entero que puede ser 0, 1 o NULL

### Optimización

SQL es un lenguaje declarativo, lo que significa que especifica qué es lo que se quiere y no cómo conseguirlo, por lo que una sentencia no establece explícitamente un orden de ejecución.

El orden de ejecución interno de una sentencia puede afectar seriamente a la eficiencia del sistema de gestión de base de datos, por lo que se hace necesario que este lleve a cabo una optimización antes de su ejecución. Muchas veces, el uso de índices acelera una instrucción de consulta, pero ralentiza la actualización de los datos. Dependiendo del uso de la aplicación, se priorizará el acceso indexado o una rápida actualización de la información. La optimización difiere sensiblemente en cada motor de base de datos y depende de muchos factores.

Los sistemas de bases de datos modernos poseen un componente llamado optimizador de consultas. Este realiza un detallado análisis de los posibles planes de ejecución de una consulta SQL y elige aquel que sea más eficiente para llevar adelante la misma.





# Repaso de Base de Datos

## III. Introducción al Lenguaje SQL cont...

### Lenguaje de definición de datos (DDL):

es el que se encarga de la modificación de la estructura de los objetos de la base de datos. Incluye órdenes para modificar, borrar o definir las tablas en las que se almacenan los datos de la base de datos. Existen cuatro operaciones básicas: CREATE, ALTER, DROP y TRUNCATE

### Lenguaje de manipulación de datos (DML):

es un lenguaje proporcionado por el sistema de gestión de base de datos que permite a los usuarios llevar a cabo las tareas de consulta o manipulación de los datos, organizados por el modelo de datos adecuado.

El lenguaje de manipulación de datos más popular hoy día es SQL, usado para recuperar y manipular datos en una base de datos relacional. Algunas operaciones principales son SELECT, INSERT, UPDATE, DELETE/TRUNCATE

### Lenguaje de control de datos (DCL):

es un lenguaje proporcionado por el sistema de gestión de base de datos que permite al Administrador del sistema gestor de base de datos, controlar el acceso a los objetos, es decir, podemos otorgar (GRANT) o denegar (REVOKE) permisos a uno o más roles para realizar determinadas tareas.



# Repaso de Base de Datos

## IV. Creación de modelos relacionales (lógicos)

El diseño de una base de datos consiste en definir la estructura de los datos que debe tener un sistema de información determinado. Para ello se suelen seguir por regla general unas fases en el proceso de diseño, definiendo para ello el modelo conceptual, el lógico y el físico.

- a. **Diseño conceptual:** se hace una descripción de alto nivel de la estructura de la base de datos, independientemente Sistema Gestor de Bases de Datos (SGBD) que se vaya a utilizar para manipularla. Su objetivo es describir el contenido de información de la base de datos y no las estructuras de almacenamiento que se necesitarán para manejar dicha información.
- b. **Diseño lógico:** este es parte del resultado del diseño conceptual y da como resultado una descripción de la estructura de la base de datos en términos de las estructuras de datos que puede procesar un tipo de SGBD. El diseño lógico depende del tipo de SGBD que se vaya a utilizar, se adapta a la tecnología que se debe emplear, pero no depende del producto concreto. Por lo general, el diseño lógico consiste en definir las tablas que existirán, las relaciones entre ellas, normalizarlas, etc...
- c. **Diseño físico:** a partir del diseño lógico se obtiene una descripción de la implementación de una base de datos en las estructuras de almacenamiento y los métodos utilizados para tener un acceso eficiente a los datos. Aquí el objetivo es conseguir una mayor eficiencia, y se tienen en cuenta aspectos concretos del SGBD sobre el que se vaya a implementar. Por regla general esto es transparente para el usuario, aunque conocer cómo se implementa ayuda a optimizar el rendimiento y la escalabilidad del sistema.

# Repaso de Base de Datos

## IV. Creación de modelos relacionales (lógicos) cont...

### El modelo relacional

En el modelo relacional las dos capas de diseño conceptual y lógico, se parecen mucho. Generalmente se implementan mediante diagramas de Entidad/Relación (modelo conceptual) y tablas y relaciones entre éstas (modelo lógico). Este es el modelo utilizado por los sistemas gestores de datos más habituales (SQL Server, Oracle, PostgreSQL, MySQL...).

*El modelo relacional de bases de datos se rige por algunas normas sencillas:*

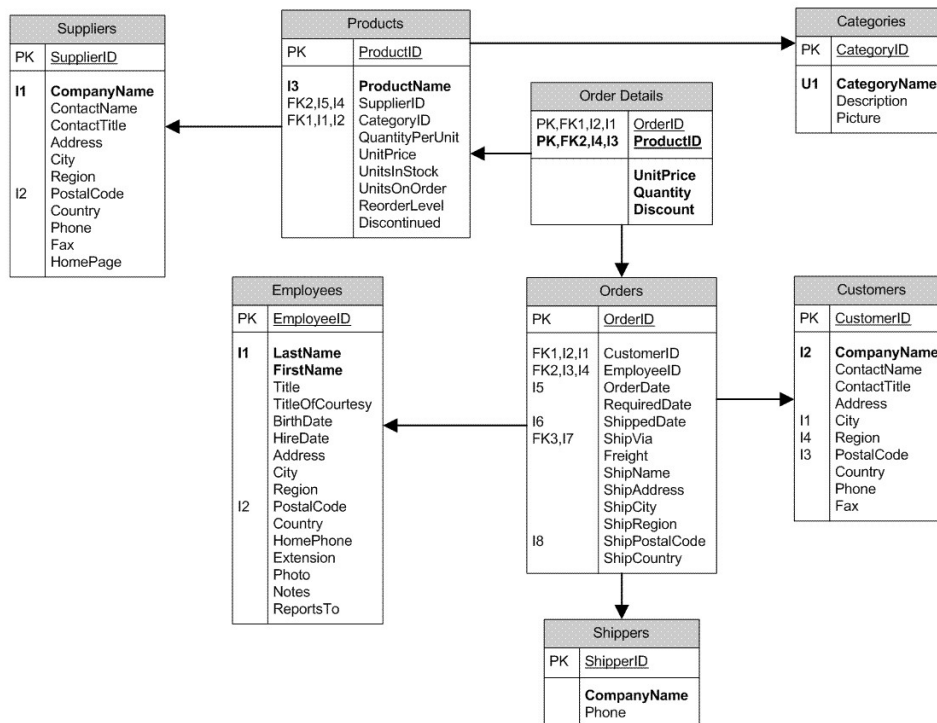
- Todos los datos se representan en forma de tablas (también llamadas “relaciones”). Incluso los resultados de consultar otras tablas. La tabla es además la unidad de almacenamiento principal.
- Las tablas están compuestas por filas (o registros) y columnas (o campos) que almacenan cada uno de los registros (la información sobre una entidad concreta, considerados una unidad).
- Las filas y las columnas, en principio, carecen de orden a la hora de ser almacenadas. Aunque en la implementación del diseño físico de cada SGBD esto no suele ser así. Por ejemplo, en SQL Server si añadimos una clave de tipo "Clustered" a una tabla haremos que los datos se ordenen físicamente por el campo correspondiente.
- El orden de las columnas lo determina cada consulta que se realizan usando SQL.
- Cada tabla debe poseer una llave primaria, esto es, un identificador único de cada registro compuesto por una o más columnas.
- Para establecer una relación entre dos tablas es necesario incluir, en forma de columna, en una de ellas la llave primaria de la otra. A esta columna se le llama llave externa. Ambos conceptos de llave son extremadamente importantes en el diseño de bases de datos.

# Repaso de Base de Datos

## IV. Creación de modelos relacionales (lógicos) cont...

### El modelo relacional

Diagrama Entidad/Relación - Ejemplo



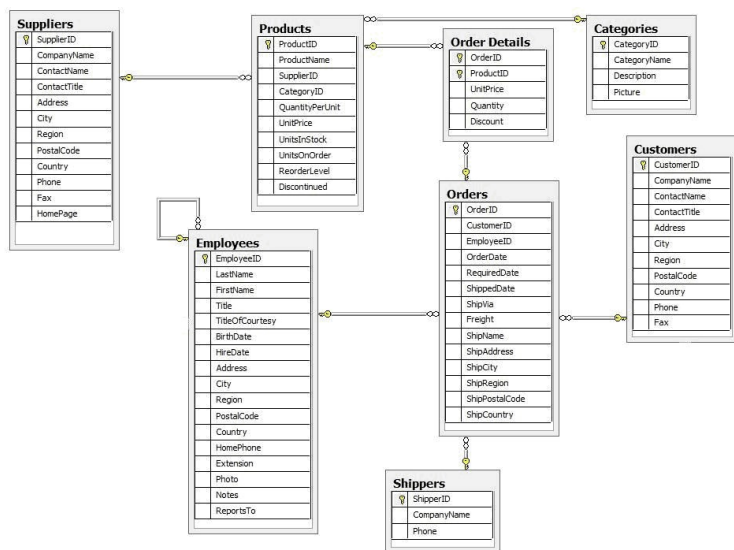
# Repaso de Base de Datos

## IV. Creación de modelos relacionales (lógicos) cont...

### Modelo lógico

Una vez tenemos claro el modelo E-R debemos traducirlo a un modelo lógico directamente en el propio sistema gestor de bases de datos (Oracle, MySQL, SQL Server...). Si hemos utilizado alguna herramienta profesional para crear el diagrama E-R, seguramente podremos generar automáticamente las instrucciones necesarias para crear la base de datos.

Entonces, todo este modelo conceptual se traduce en un modelo lógico que trasladaremos a la base de datos concreta que estemos utilizando y que generalmente será muy parecido.



En este caso, se ha creado cada tabla, una a una, siguiendo lo identificado en el diagrama E-R y estableciendo índices y demás elementos según las indicaciones de cada uno de los campos.

Su representación gráfica en la base de datos es muy similar, sin embargo el modelo físico (cómo se almacena esto físicamente), puede variar mucho de un SGBD a otro y según la configuración que le demos.

# Repaso de Base de Datos

## V. Migración de un modelo lógico a un modelo físico

### ¿Qué es un Lenguaje Unificado de Modelado?

El modelado sirve no solamente para los grandes sistemas, aun en aplicaciones de pequeño tamaño se obtienen beneficios de modelado, sin embargo es un hecho que entre más grande y más complejo es el sistema, más importante es el papel de que juega el modelado por una simple razón: "El hombre hace modelos de sistemas complejos porque no puede entenderlos en su totalidad".

UML es una técnica para la especificación sistemas en todas sus fases. Nació en 1994 cubriendo los aspectos principales de todos los métodos de diseño antecesores y, precisamente, los padres de UML son Grady Booch, autor del método Booch; James Rumbaugh, autor del método OMT e Ivar Jacobson, autor de los métodos OOSE y Objectory. La versión 1.0 de UML fue liberada en Enero de 1997 y ha sido utilizado con éxito en sistemas construidos para toda clase de industrias alrededor del mundo: hospitales, bancos, comunicaciones, aeronáutica, finanzas, etc.

El Lenguaje Unificado de Modelado (UML) desempeña un rol importante no solo en el desarrollo de software, sino también en los sistemas que no tienen software en muchas industrias, ya que **es una forma de mostrar visualmente el comportamiento y la estructura de un sistema o proceso.**

### Los principales beneficios de UML son:

- ✓ Mejores tiempos totales de desarrollo (de 50 % o más).
- ✓ Modelar sistemas (y no sólo de software) utilizando conceptos orientados a objetos.
- ✓ Establecer conceptos y artefactos ejecutables.



# Repaso de Base de Datos

## V. Migración de un modelo lógico a un modelo físico

### ¿Qué es un Lenguaje Unificado de Modelado? Cont ...

- ✓ Encaminar el desarrollo del escalamiento en sistemas complejos de misión crítica.
- ✓ Crear un lenguaje de modelado utilizado tanto por humanos como por máquinas.
- ✓ Mejor soporte a la planeación y al control de proyectos.
- ✓ Alta reutilización y minimización de costos

Links de referencial sobre UML:

<http://profesores.fi-b.unam.mx/carlos/aydoo/uml.html>

<https://www.microsoft.com/es-ww/microsoft-365/business-insights-ideas/resources/guide-to-uml-diagramming-and-database-modeling>

## V. Instalación del entorno de desarrollo

