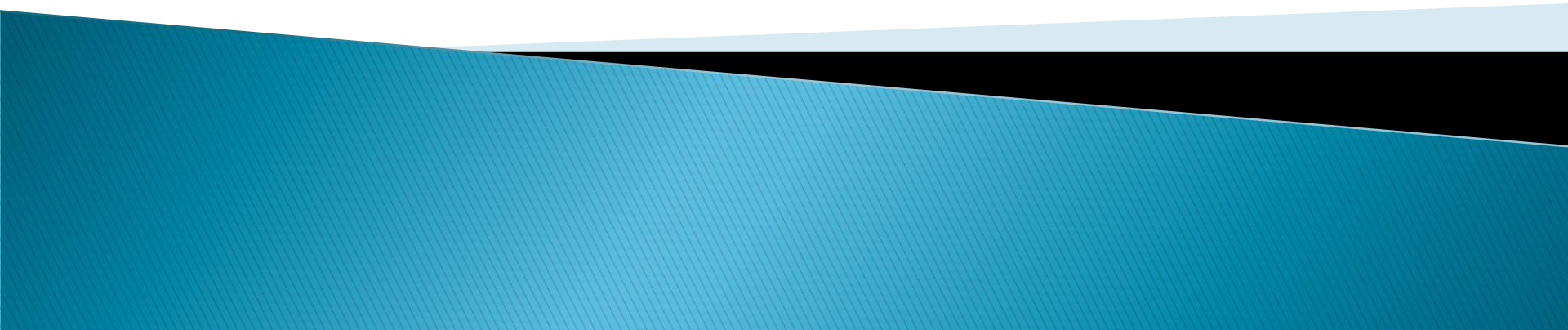
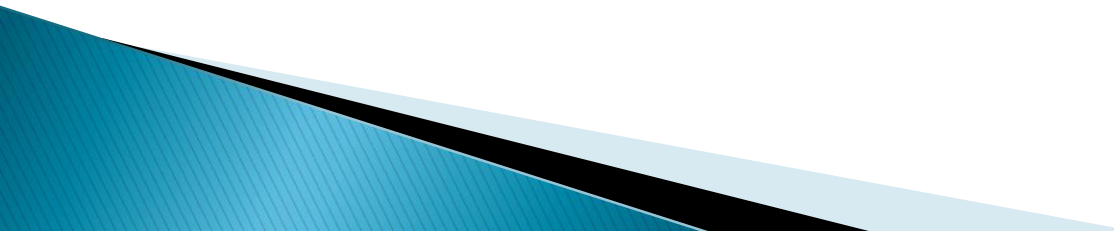


Normalización de base de datos

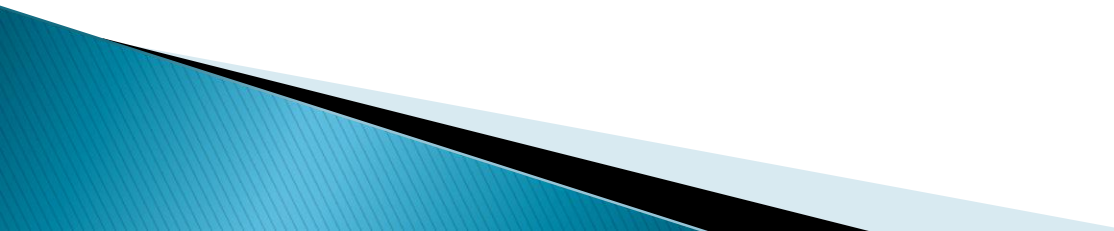
Sergio Sánchez



Agenda

- ▶ ¿Que es normalización?
 - ▶ Normalización de una base de datos
 - ▶ Grados de normalización: Primera Forma
 - ▶ Grados de normalización: Segunda Forma
 - ▶ Grados de normalización: Tercera Forma
 - ▶ Otras formas de normalización
 - ▶ Ejemplo
- 

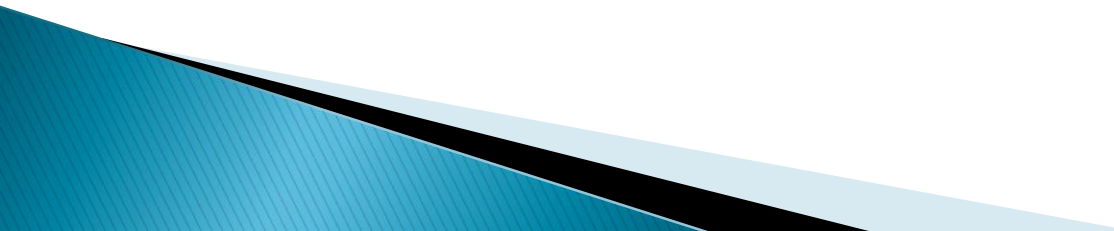
¿Que es normalización?

- ▶ Normalización es el proceso de organizar de manera eficiente los datos dentro de una base de datos. Esto incluye la creación de tablas y el establecimiento de relaciones entre ellas según reglas pre-diseñadas tanto para proteger los datos y la base de datos, como para hacer más flexible al eliminar la redundancia y dependencia incoherente.
- 

¿Que es normalización?

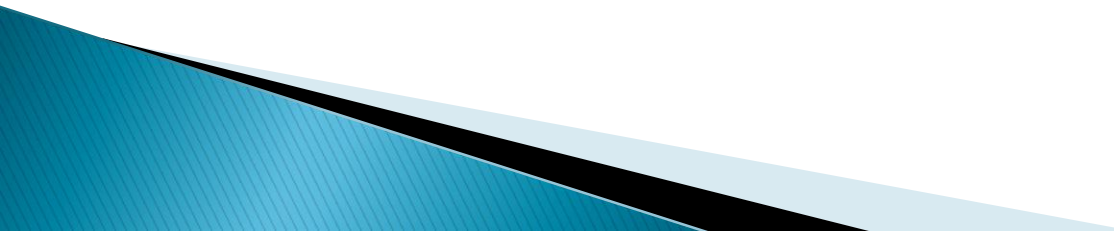
- ▶ Los principales objetivos de la normalización son:
 - La eliminación de datos redundantes, los cuales ocupan mas espacio en disco y crean problemas de mantenimiento; por ejemplo, cambio de la dirección del cliente es mucho más fácil de implementar si los datos se almacenan sólo en la tabla Clientes y en ninguna otra base de datos.
 - Evitar problemas de actualización de los datos en las tablas.
 - Garantizar que las dependencias que tienen los datos entre ellos, sean lógicas y presenten algún sentido.

¿Que es normalización?

- ▶ Estos puntos reducen la cantidad de espacio en la base de datos y aseguran que estos son almacenados de manera lógica (integridad).
 - ▶ La normalización también se puede entender como el proceso mediante el cual se transforman datos complejos a un conjunto de estructuras de datos más pequeñas, que además de ser más simples y más estables, son más fáciles de mantener.
- 

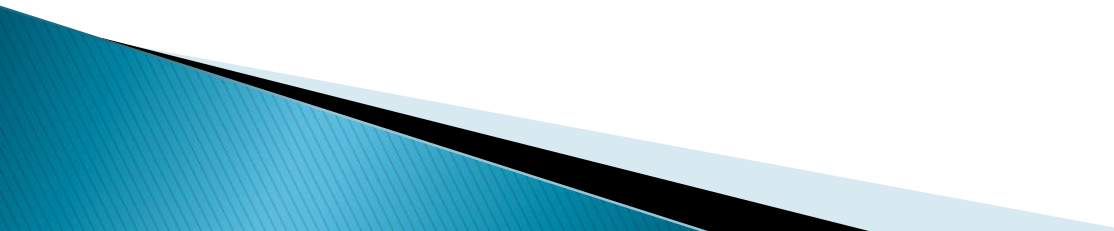
Normalización de una base de datos

Existen algunas reglas para la normalización de bases de datos. Cada regla se denomina "forma normal". Si dentro de la base de datos se observa la primera regla se dice que está en "primera forma normal". Si las tres primeras reglas se observan, la base de datos se considera en "tercera forma normal". Aunque es posible tener otros niveles de normalización, la tercera forma normal es considerado el más alto nivel necesario para la mayoría de aplicaciones.



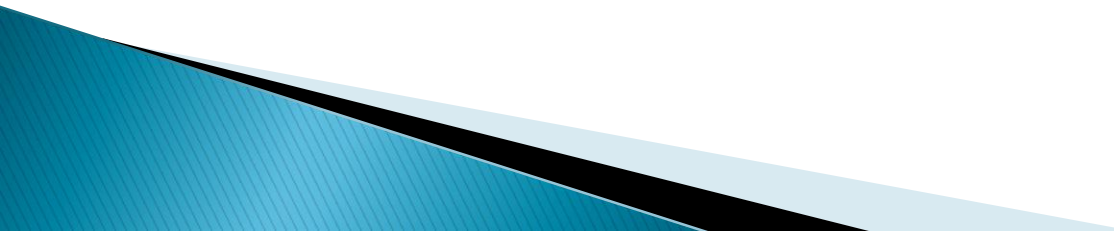
Normalización de una base de datos

Como ocurre con muchas reglas y especificaciones, en la vida real no siempre es factible el cumplimiento de estas. En general, la normalización requiere tablas adicionales y para algunos clientes esto es engorroso. Si se decide violar una de las tres primeras reglas de normalización, tenga por seguro que su aplicación presentara problemas, como los datos redundantes y dependencias incoherentes.



Grados de normalización: Primera Forma Normal

Los principales objetivos son:

- ▶ Eliminar grupos de datos repetidos en tablas individuales.
 - ▶ Crear una tabla separada para cada conjunto de datos relacionados.
 - ▶ Identifique cada conjunto de datos relacionados con una clave principal. Ejemplo ID, Primary Key, FK.
- 

Grados de normalización: Primera Forma Normal

No utilizar varios campos en una sola tabla para almacenar datos similares. Por ejemplo, para el seguimiento de un artículo del inventario que proviene de dos fuentes diferentes, el registro puede contener campos para el código de proveedor 1 y un código de proveedor 2.

¿Qué sucede cuando se agrega un tercer proveedor? Agregar un campo no es la respuesta, ya que requiere de programación y modificación de tablas y la necesidad de repetirlo cada vez que se agregué a un nuevo proveedor. En su lugar, se deberá poner toda la información del proveedor en una tabla independiente denominada Proveedores, y vincular el inventario con los proveedores por medio de una clave o de sus claves.

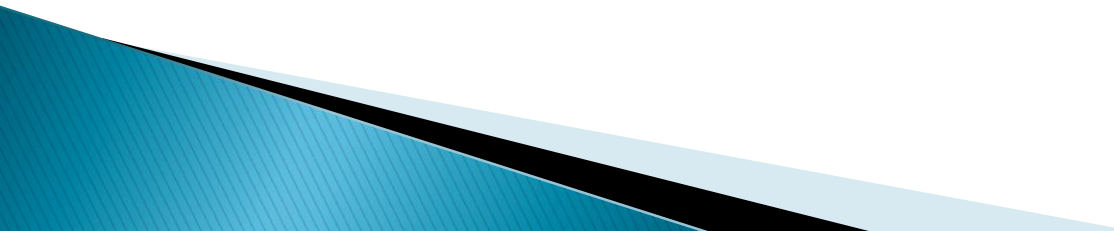
Grados de normalización: Segunda Forma Normal

Los principales objetivos son:

- ▶ Crear tablas separadas para aquellos conjuntos de valores que se aplican a varios registros. Ejemplo ciudades, profesión.
- ▶ Relacionar estas tablas por medio de una clave externa, ejemplo ID, Primary Key, FK.

Grados de normalización: Segunda Forma Normal

Los registros no deben depender de nada que no sea la clave primaria de una tabla (una clave compuesta, si es necesario). Por ejemplo, consideremos la dirección de un cliente en un sistema contable. La dirección no solo se necesita en la tabla de clientes, sino también para los pedidos, envío, facturas, cuentas por cobrar, e inclusive en las ordenes. En lugar de almacenar la dirección del cliente como una entrada independiente en cada una de estas tablas, guárdela en un lugar, ya sea en la tabla Clientes o en una tabla de direcciones separada.



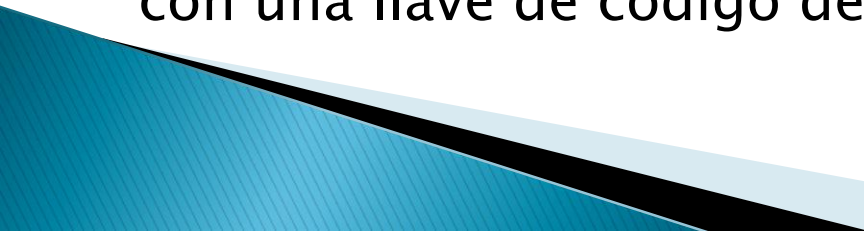
Grados de normalización: Tercera Forma Normal

Los principales objetivos son:

- ▶ Eliminar los campos que no dependan de las claves.

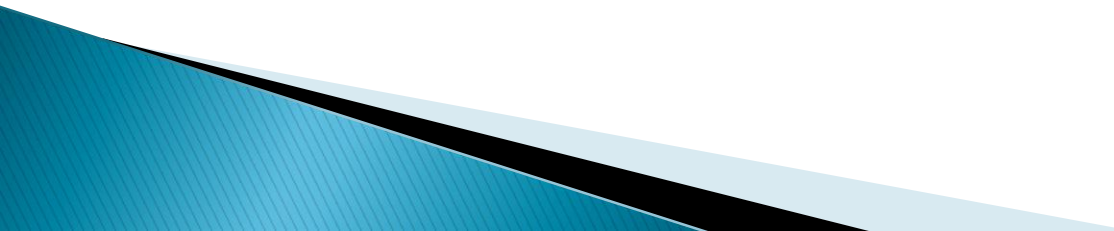
Los valores de un registro que no forman parte de la clave de registro no tienen cabida en la tabla.

Por ejemplo, en una tabla que contiene los datos de los candidatos a un puesto, el nombre del candidato, nombre de la universidad a la que asistió y la dirección pueden estar incluidos. Pero existen muchas universidades. Si la información de la universidad se almacena en la tabla de candidatos, no hay manera de listar las universidades que no tengan candidatos. La mejor opción es crear una tabla separada de Universidades y vincularlo a la tabla Candidatos con una llave de código de la universidad.



Otras formas de normalización

Cuarta forma normal, también llamada Boyce Codd Forma Normal (FNBC), y quinta forma normal, existen, pero rara vez se consideran en el diseño práctico. Haciendo caso omiso de estas reglas puede resultar en menos de diseño de base de datos perfecto, pero no debería afectar a la funcionalidad.



Ejemplo de normalización

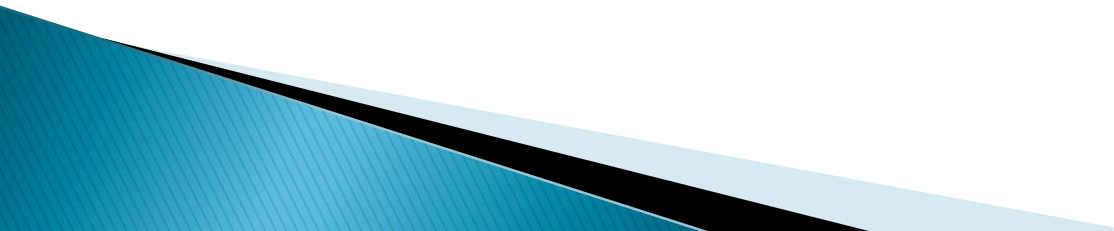
Aquí se muestra la normalización de una tabla de estudiantes.

Estudiante#	Nombre del titular	Salón	Clase1	Clase2	Clase3
102	Sr. Rodriguez	101	Matemáticas	Literatura	Química
412	Srita. Jimenez	201	Biología	Geografía	Cálculo

Ejemplo de normalización

Se aplica la primera forma.

Las tablas deben tener sólo dos dimensiones. Dado que los estudiantes tiene varias clases, estas clases deben ser listados en una tabla separada. Los campos Clase 1, Clase 2, y Clase 3 en los registros anteriores son indicios de problemas de diseño.



Ejemplo de normalización

Las hojas de cálculo suelen usar la tercera dimensión, pero las tablas no deben. Otra forma de ver este problema es con una relación uno-a-muchos. Cree otra tabla en la primera forma normal eliminando el grupo de repetición (clase), como se muestra a continuación:.

Estudiante#	Nombre del titular	Salón	Clase#
102	Sr. Rodriguez	101	Matemáticas
102	Sr. Rodriguez	101	Literatura
102	Sr. Rodriguez	101	Química
412	Srita. Jimenez	201	Biología
412	Srita. Jimenez	201	Geografía
412	Srita. Jimenez	201	Cálculo

Ejemplo de normalización

Segunda forma:

Tome en cuenta los múltiples valores para el campo Clase# por cada estudiante en la tabla anterior. El campo Clase# no es dependiente del campo Estudiante# (llave primaria) por lo que esta relación no esta en la segunda forma normal. Las siguientes tablas muestran como quedarían con la 2ª forma:

Estudiante#	Nombre del titular	Salón
102	Sr. Rodriguez	101
412	Srita. Jimenez	201

Estudiante#	Clase#
102	Matemáticas
102	Literatura
102	Química
412	Biología
412	Geografía
412	Cálculo

Ejemplo de normalización

Tercera forma normal: eliminar los datos que no dependen de la llave

En el último ejemplo, salón (salón/grupo asignado al asesor) es funcionalmente dependiente del atributo titular. La solución es mover dicho atributo de la tabla Alumnos a la tabla de Facultad, como se muestra a continuación:

Estudiante#	Nombre del titular
102	Sr. Rodriguez
412	Srita. Jimenez

Nombre del titular	Salón	Departamento
Sr. Rodriguez	101	A
Srita. Jimenez	201	B

Preguntas

ssanchez@informatica.com

<http://sesa78.wordpress.com/>