

Normalización de Tablas en Bases de Datos Relacionales

¡Bienvenidos al Mundo de las Bases de Datos Ordenadas!

Imagina que eres el encargado de organizar la información de tu escuela técnica. Tienes que guardar datos sobre los alumnos, los profesores, las materias, las notas, etc. Si lo haces de manera desordenada, pronto tendrás un caos: información repetida, errores, y será muy difícil encontrar lo que buscas. ¡Ahí es donde entra la normalización!

La normalización es como poner orden en el caos de la información. Es un proceso que nos ayuda a organizar los datos en una base de datos de manera eficiente, evitando repeticiones innecesarias y asegurando que la información sea correcta y consistente. En este resumen, vamos a explorar qué es la normalización, por qué es importante y cómo se hace.

¿Qué es una Base de Datos Relacional?

Antes de hablar de normalización, necesitamos entender qué es una base de datos relacional. Piensa en una base de datos como un conjunto de tablas, donde cada tabla guarda información sobre un tema específico. Por ejemplo, una tabla podría guardar información sobre los alumnos, otra sobre los profesores, y otra sobre las materias.

En una base de datos relacional, estas tablas están relacionadas entre sí. Por ejemplo, la tabla de alumnos puede estar relacionada con la tabla de materias, indicando qué materias cursa cada alumno. Estas relaciones son las que hacen que la base de datos sea "relacional".

Analogía: Piensa en una tienda de videojuegos. Tienes una tabla para los juegos, otra para los clientes, y otra para las ventas. La tabla de ventas relaciona a los clientes con los juegos que compraron.

¿Por Qué Necesitamos Normalizar?

Imagina que tienes una tabla donde guardas información sobre los alumnos y las materias que cursan. Podrías tener una tabla como esta:

Tabla: Alumnos y Materias (Sin Normalizar)

ID Alumno	Nombre Alumno	Apellido Alumno	Materia 1	Materia 2	Materia 3	Profesor Materia 1	Profesor Materia 2	Profesor Materia 3
1	Juan	Pérez	Matemáticas	Física	Química	Ana López	Carlos Gómez	Laura Martínez
2	María	García	Matemáticas	Biología	Inglés	Ana López	Pedro Sánchez	Sofía Ruiz
3	Pedro	López	Física	Química	Informática	Carlos Gómez	Laura Martínez	Javier Torres

En esta tabla, hay varios problemas:

- **Repetición de Datos (Redundancia):** El nombre del profesor se repite varias veces. Si Ana López cambia de apellido, ¡tendrías que cambiarlo en varias filas!
- **Dificultad para Actualizar:** Si quieres añadir una nueva materia, tendrías que añadir una nueva columna. ¡Imagina si tuvieras 20 materias!
- **Inconsistencia:** Si por error escribes "Matematicas" en una fila y "Matemática" en otra, la base de datos lo considerará como dos materias diferentes.
- **Espacio Desperdiciado:** Se guarda mucha información repetida, ocupando espacio innecesario.

Estos problemas se conocen como **anomalías**. La normalización nos ayuda a evitar estas anomalías.

Analogía: Imagina una lista de contactos en tu teléfono donde tienes el mismo contacto repetido varias veces con diferentes números o información incompleta. Es difícil de mantener y puede llevar a errores.

¿Qué es la Normalización?

La normalización es un proceso que consiste en organizar los datos en una base de datos para:

- **Minimizar la Redundancia:** Evitar que la misma información se repita en varios lugares.
- **Asegurar la Consistencia:** Garantizar que la información sea correcta y no haya contradicciones.
- **Facilitar la Actualización:** Hacer que sea fácil modificar la información sin generar errores.
- **Mejorar la Eficiencia:** Hacer que las consultas a la base de datos sean más rápidas y eficientes.

Los Niveles de Normalización: Las Formas Normales

La normalización se realiza en varios pasos, llamados **formas normales**. Cada forma normal elimina un tipo de problema. Las tres formas normales más comunes son:

Primera Forma Normal (1FN):

- **Regla:** Cada celda de la tabla debe contener un solo valor, no una lista de valores.
- **Ejemplo:** En la tabla anterior, cada alumno tiene varias materias en la misma fila. En 1FN, cada alumno tendría una fila por cada materia que cursa.

Tabla: Alumnos y Materias (1FN)

ID Alumno	Nombre Alumno	Apellido Alumno	Materia	Profesor
1	Juan	Pérez	Matemáticas	Ana López
1	Juan	Pérez	Física	Carlos Gómez
1	Juan	Pérez	Química	Laura Martínez
2	María	García	Matemáticas	Ana López
2	María	García	Biología	Pedro Sánchez
2	María	García	Inglés	Sofía Ruiz
3	Pedro	López	Física	Carlos Gómez
3	Pedro	López	Química	Laura Martínez
3	Pedro	López	Informática	Javier Torres

- **Explicación:** Ahora cada celda tiene un solo valor. Pero todavía hay repetición de datos (el nombre del profesor se repite).

Segunda Forma Normal (2FN):

- **Regla:** La tabla debe estar en 1FN y cada columna que no sea parte de la clave primaria debe depender completamente de la clave primaria.
- **Explicación:** En la tabla anterior, la clave primaria es una combinación de ID Alumno y Materia. El nombre del profesor depende solo de la materia, no del alumno. Para estar en 2FN, debemos separar la información de las materias y los profesores en una tabla aparte.

Tabla: Alumnos y Materias (2FN)

ID Alumno	Nombre Alumno	Apellido Alumno	Materia
1	Juan	Pérez	Matemáticas
1	Juan	Pérez	Física
1	Juan	Pérez	Química
2	María	García	Matemáticas
2	María	García	Biología
2	María	García	Inglés
3	Pedro	López	Física
3	Pedro	López	Química
3	Pedro	López	Informática

Tabla: Materias y Profesores (2FN)

Materia	Profesor
Matemáticas	Ana López
Física	Carlos Gómez
Química	Laura Martínez
Biología	Pedro Sánchez
Inglés	Sofía Ruiz
Informática	Javier Torres

- **Explicación:** Ahora tenemos dos tablas. En la tabla "Alumnos y Materias", solo guardamos la información del alumno y la materia que cursa. En la tabla "Materias y Profesores", guardamos la información de cada materia y su profesor.

Tercera Forma Normal (3FN):

- **Regla:** La tabla debe estar en 2FN y no debe haber dependencias transitivas.

- **Explicación:** Una dependencia transitiva ocurre cuando una columna depende de otra columna que no es la clave primaria.
- **Ejemplo de Dependencia Transitiva:** Imagina que en la tabla "Materias y Profesores" tuviéramos una columna "Departamento del Profesor". Si el departamento del profesor dependiera del profesor, pero no de la materia, tendríamos una dependencia transitiva (Materia -> Profesor -> Departamento del Profesor). Para estar en 3FN, deberíamos crear una tabla "Profesores" con el departamento y relacionarla con la tabla "Materias y Profesores".

Tabla: Materias y Profesores (Antes de 3FN)

Materia	Profesor	Departamento del Profesor
Matemáticas	Ana López	Ciencias Exactas
Física	Carlos Gómez	Ciencias Exactas
Química	Laura Martínez	Ciencias Exactas
Biología	Pedro Sánchez	Ciencias Naturales
Inglés	Sofía Ruiz	Humanidades
Informática	Javier Torres	Tecnología

Tabla: Profesores (3FN)

Profesor	Departamento del Profesor
Ana López	Ciencias Exactas
Carlos Gómez	Ciencias Exactas
Laura Martínez	Ciencias Exactas
Pedro Sánchez	Ciencias Naturales
Sofía Ruiz	Humanidades
Javier Torres	Tecnología

Tabla: Materias y Profesores (3FN)

Materia	Profesor
Matemáticas	Ana López
Física	Carlos Gómez
Química	Laura Martínez
Biología	Pedro Sánchez
Inglés	Sofía Ruiz
Informática	Javier Torres

- **Explicación:** En el ejemplo anterior, al aplicar 3FN, se separo la tabla "Materias y Profesores" en dos tablas, "Profesores" y "Materias y Profesores".

Ejemplo Práctico: El Taller de Robótica

Imagina que estás diseñando la base de datos para el taller de robótica de tu escuela técnica. Necesitas guardar información sobre los proyectos, los alumnos que participan y los materiales que se utilizan.

Tabla Inicial (Sin Normalizar):

ID Proyecto	Nombre Proyecto	ID Alumno	Nombre Alumno	Material 1	Material 2	Material 3
1	Robot Seguidor	1	Juan Pérez	Arduino	Motor DC	Cables
2	Brazo Robótico	2	María García	Arduino	Servomotor	Engranajes
3	Robot Seguidor	3	Pedro López	Arduino	Motor DC	Cables

- **Problemas:** Repetición del nombre del proyecto y los materiales.

Tablas Normalizadas:

Tabla: Proyectos

ID Proyecto	Nombre Proyecto
1	Robot Seguidor
2	Brazo Robótico

Tabla: Alumnos

ID Alumno	Nombre Alumno
1	Juan Pérez
2	María García
3	Pedro López

Tabla: Materiales

ID Material	Nombre Material
1	Arduino
2	Motor DC
3	Cables
4	Servomotor
5	Engranajes

Tabla: Proyectos_Materiales

ID Proyecto	ID Material
1	1
1	2
1	3
2	1
2	4
2	5
3	1
3	2
3	3

Tabla: Alumnos_Proyectos

ID Alumno	ID Proyecto
1	1
2	2
3	3

- **Beneficios:** No hay repetición de datos. Si se agrega un nuevo material, solo se agrega en la tabla "Materiales". Si un alumno se une a un proyecto, solo se agrega en la tabla "Alumnos_Proyectos".

Conclusión

La normalización es una herramienta fundamental para diseñar bases de datos eficientes y confiables. Al eliminar la redundancia y organizar los datos de manera lógica, podemos evitar errores, facilitar la actualización de la información y mejorar el rendimiento de las consultas. Como estudiantes de una escuela técnica, es importante que comprendan estos conceptos, ya que les serán muy útiles en su futuro profesional, especialmente si se dedican al desarrollo de software o a la gestión de información.

Recuerda que la normalización es como poner orden en tu taller: al principio puede parecer un poco complicado, pero una vez que lo haces, todo es más fácil de encontrar y mantener. ¡Así que no tengas miedo de normalizar tus bases de datos!

Tabla de Resumen de las Formas Normales

Forma Normal	Descripción	Beneficios
1FN	Cada celda contiene un solo valor.	Elimina la repetición de grupos de datos.
2FN	Está en 1FN y cada columna no clave depende completamente de la clave primaria.	Elimina la repetición de datos que dependen solo de una parte de la clave.

Forma Normal	Descripción	Beneficios
3FN	Está en 2FN y no hay dependencias transitivas.	Elimina la repetición de datos que dependen de columnas no clave.

Glosario

- **Base de Datos:** Un conjunto organizado de información.
- **Base de Datos Relacional:** Una base de datos donde la información se guarda en tablas relacionadas entre sí.
- **Tabla:** Una estructura que guarda información en filas y columnas.
- **Fila:** Un registro de información en una tabla.
- **Columna:** Un atributo o característica de la información en una tabla.
- **Clave Primaria:** Un identificador único para cada fila en una tabla.
- **Clave Foránea:** Una columna en una tabla que hace referencia a la clave primaria de otra tabla.
- **Redundancia:** Repetición innecesaria de información.
- **Anomalía:** Un error o inconsistencia en la información.
- **Dependencia Funcional:** Cuando el valor de una columna determina el valor de otra columna.
- **Dependencia Transitiva:** Cuando una columna depende de otra columna que no es la clave primaria.
- **Normalizacion:** Proceso de organizar los datos en una base de datos.