



Profesora: Erika Gutiérrez Beltrán

Tema 2: Bases de datos relacionales



Objetivo

Transformar los modelos conceptuales a lógicos. Construir bases de datos para datos estructurados

- Utilización de los modelos en el mundo real ✓
- Transformación del modelado a la base de datos ✓



Temario

- Clases de entidades
- Modelo relacional, restricciones de integridad y normalización ✓
- Arquitectura de los sistemas gestores de Bases de datos, OLTP, OLAP
- Lenguaje de consulta formal
- Lenguaje de consulta SQL



Logros

Capturar datos desde una pantalla de usuario real y almacenarlos en una base de datos. Manipulación de estos datos.



TOMÁTE UNA PAUSA

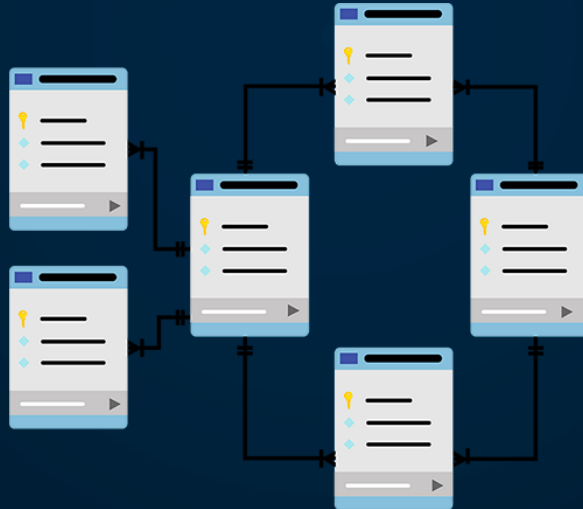
Vamos a analizar lo aprendido hasta el momento y que ha sido lo más complejo de entender. ¿Por qué crees tú que ese tema ha sido complejo de entender?, ¿falta mayor detalle?, ¿faltan más actividades en clase para fortalecer los conocimientos?, ¿crees que faltan otros conocimientos?

En este espacio de 10 minutos vamos a intentar resolver las cuestiones que tenemos para fortalecer nuestras debilidades y mejorar en el proceso de aprendizaje



Normalización

La normalización es el proceso de organizar y estructurar datos en la base de datos. Esto incluye las relaciones entre las tablas y las restricciones de integridad las cuales son declaración de variables, cardinalidad de la relación, participación mínima y máxima, integridad de dominio, integridad referencial, dependencias funcionales y multivaloradas.





Normalización

En este proceso de organización se almacenan solo datos relacionados entre sí y se ubican los datos no relacionados, eliminando los datos redundantes.

La normalización ofrece beneficios como:

- Reducir espacio de almacenamiento
- Mejora desempeño de las consultas con SQL
- Reducción de anomalías y errores en la manipulación de los datos



Normalización

La normalización tiene varias formas normales, para esta clase haremos el énfasis principalmente en 3 de las formas normales.

Algunos autores hacen referencia solo a las 3 primeras que veremos, otros a 5 y en otras referencias será posible encontrar hasta 6.



PRIMERA FORMA NORMAL, 1NF

- No debe tener ninguna de sus filas de datos repetida
- Cada columna contiene un valor único
- La tabla posee una llave primaria o PK
- Crea una tabla separada para cada conjunto de datos relacionados



PRIMERA FORMA NORMAL, 1NF

Hagamos un análisis de esta entidad.

- ¿Se cumple la primera forma normal?
- ¿Qué problemas identificas?
- ¿Qué ajustes harías para que se cumpla la primera forma normal?

PRODUCTOS			
ID	nombre_producto	precio	categoria
123	Shampoo	20000	1,2
124	Jabon	30000	1,3
123	Shampoo	20000	1
125	Desodorante	5000	2
111	Crema	18000	3
126	Limpiador	7000	1



SEGUNDA FORMA NORMAL, 2NF

- Se debe cumplir la primera forma normal
- Toda columna que no es clave primaria es dependiente de la clave primaria entera
- Las tablas deben relacionarse con una clave foránea
- Crear tablas separadas para conjuntos de valores que se aplican a varios registros



SEGUNDA FORMA NORMAL, 2NF

Hagamos un análisis de esta entidad.

- ¿Se cumple la primera forma normal?
- ¿Qué problemas identificas?
- ¿Qué ajustes harías para que se cumpla la segunda forma normal?, recuerda tener en cuenta la creación de nuevas tablas y relaciones para hacer que se cumpla esta forma normal

PRODUCTOS				
ID	nombre_producto	precio	categoria	descripcion_cat
123	Shampoo	20000	1	Aseo
124	Jabon	30000	1	Aseo
123	Shampoo	20000	1	Aseo
125	Desodorante	5000	2	Belleza
111	Crema	18000	3	Cuidado piel
126	Limpiador	7000	1	Aseo



TERCERA FORMA NORMAL, 3NF

- Debe cumplirse la segunda forma normal
- Toda columna no primaria no depende de otra columna no primaria
- Los campos que no dependen de la clave pueden ser eliminados o separados por tablas



TERCERA FORMA NORMAL, 3NF

Hagamos un análisis de esta entidad.

- ¿Se cumple la primera forma normal?
- ¿Se cumple la segunda forma normal?
- ¿Qué ajustes harías para que se cumpla la tercera forma normal?, recuerda tener en cuenta la creación de nuevas tablas y relaciones para hacer que se cumpla esta forma normal

PRODUCTOS					
ID	nombre_producto	precio	categoria	Proveedor	Ciudad_Proveedor
123	Shampoo	20000	1	Palmolive	Cali
124	Jabon	30000	1	Palmolive	Cali
125	Desodorante	5000	2	Ésika	Medellin
111	Crema	18000	3	Natura	Bogotá
126	Limpiador	7000	1	Palmolive	Cali



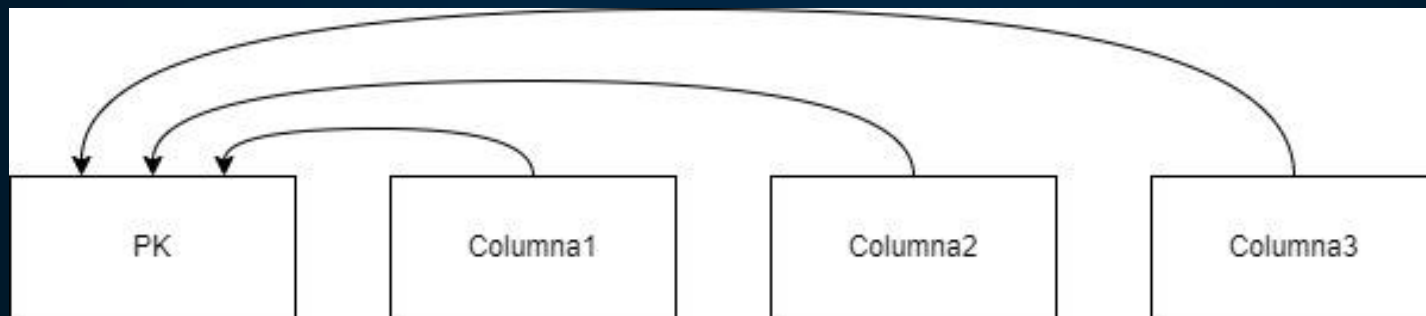
TERCERA FORMA NORMAL, 3NF

¡EXCEPCIÓN!: Cumplir con la tercera forma normal no siempre será posible, ni práctico ya que puede degradar el rendimiento o exceder las capacidades de memoria si hay muchas tablas pequeñas en nuestra base de datos.



A tener en cuenta

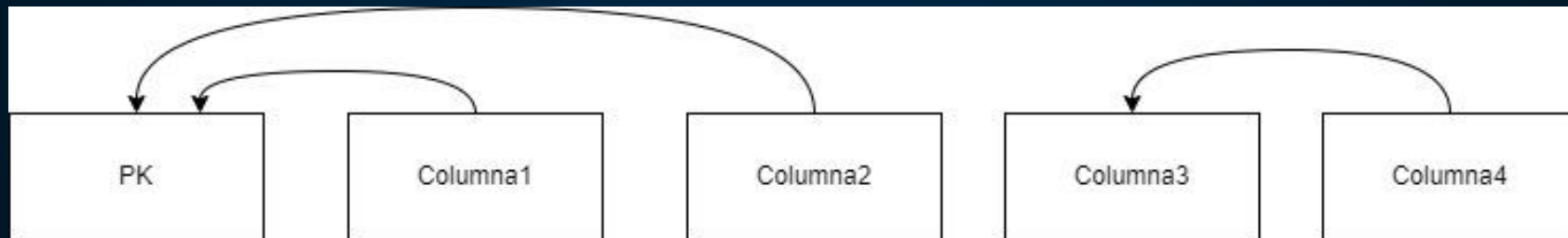
Dependencias funcionales: Son aquellos atributos que dependen de la clave primaria





A tener en cuenta

Dependencias transitivas: Un atributo no depende de la clave primaria pero su clave primaria se encuentra en la misma tabla





Ejercicio práctico

Ahora, teniendo en cuenta las primeras formas de normalización (1NF, 2NF, 3NF), vamos a tomar el modelo del trabajo (Movies Eafit) y vamos a solucionar algunos problemas asociados a la normalización



OTRAS FORMAS DE NORMALIZAR

Existen otras formas de normalizar, las cuales agregan complejidad al dominio pero mayor consistencia a los datos si se aplica de la manera correcta, estas formas de normalizar son:

- Forma normal Boyce-Codd (Todos los campos dependen de la clave primaria, aplica la 1NF, 2NF y 3NF)

FACTURA			
id_compra	id_vendedor	id_cliente	id_tienda
33456	1	11	1
23457	1	12	1
22123	2	15	2
446799	3	11	1



OTRAS FORMAS DE NORMALIZAR

Existen otras formas de normalizar, las cuales agregan complejidad al dominio pero mayor consistencia a los datos si se aplica de la manera correcta, estas formas de normalizar son:

- Forma normal Boyce-Codd (Todos los campos dependen de la clave primaria, aplica la 1NF, 2NF y 3NF)

FACTURA			
id_compra	id_vendedor	id_cliente	id_tienda
33456	1	11	1
23457	1	12	1
22123	2	15	2
446799	3	11	1



OTRAS FORMAS DE NORMALIZAR

- Forma normal 4NF: depende de la forma normal Boyce-Coddly elimina todas las dependencias multivaluadas

DIRECCION_PROVEEDORES	
id_proveedor	direccion
1	calle52 #73A-43
2	calle70 #73BB-01
2	cra 91c #45-12
1	calle 80 #33-56



OTRAS FORMAS DE NORMALIZAR

Para las dependencias transitivas se tiene en cuenta el atributo multivalor, este no solo almacena múltiple información en un campo asociada a una dirección un teléfono o correo, sino que puede contener la siguiente estructura

PROVEEDORES			
id_proveedor	proveedor	ciudad_Proveedor	dirreccion
1	Palmolive	Cali	calle52 #73A-43
2	Ésika	Medellin	calle70 #73BB-01
3	Natura	Bogotá	calle 80 #33-56
1	Palmolive	Cali	cra 91c #45-12
1	Palmolive	Cali	cra 96A #45-10



OTRAS FORMAS DE NORMALIZAR

- Forma normal 5NF: debe poseer la 4 forma normal, y se enfoca principalmente en separar en pequeños conjuntos tablas de gran tamaño que es inmanejable, también aplica para mitigar información redundante que puede presentarse en estos casos

PRODUCTO_CATEGORIA_TIPO		
producto	categoria	tipo
jabon liquido	aseo	disponible domicilio
jabon liquido	limpieza	agotado
desodorante	belleza	recoger en sitio
desodorante	limpieza	disponible domicilio

Original

PRODUCTO_CATEGORIA	
id_categoria	id_producto
jabon liquido	aseo
jabon liquido	limpieza
desodorante	belleza
desodorante	limpieza

PRODUCTO_TIPO	
producto	tipo
jabon liquido	aseo
jabon liquido	limpieza
desodorante	belleza
desodorante	limpieza

Normalizado

CATEGORIA_TIPO	
producto	tipo
aseo	aseo
belleza	belleza
limpieza	limpieza



OTRAS FORMAS DE NORMALIZAR

Para aplicar la quinta forma normal se debe tener en cuenta el concepto de proyección y de unión.

PROYECCIÓN: Para el caso de la proyección esta consiste en sacar subconjuntos de la tabla original que son necesarios, tienen mayor relevancia y son los más utilizados. Se crea una tabla nueva de una de gran tamaño.

UNIÓN: Maneja los subconjuntos que previamente han sido creados a partir de una proyección, por medio de las relaciones se unen las tablas y se obtiene mayor detalle y definición de la información.

Importante:

Cuando se aplican estos conceptos sobre SQL para extraer información no se hacen modificaciones a la tabla, el concepto llevado a la 5 forma normal modifica el modelo relacional, o la estructura de base de datos.



RESUMEN

1NF: Identificar grupos de datos repetidos sobre la misma entidad y registro

2NF: Debe contener 1NF y se debe identificar las dependencias transitivas y funcionales

3NF: Debe contener 2NF y se deben eliminar las dependencias transitivas y todo atributo no principal debe depender de la clave primaria

NFBC: Debe contener 3NF y se cumple si las dependencias funcionales dependen de una clave primaria

4NF: Debe contener NFBC y esta consiste en eliminar todas las dependencias multivaluadas

5NF: Debe contener 4NF y se utiliza para dividir tablas de gran tamaño y mejorar la lectura y manipulación de los datos



EJERCICIO PRACTICO

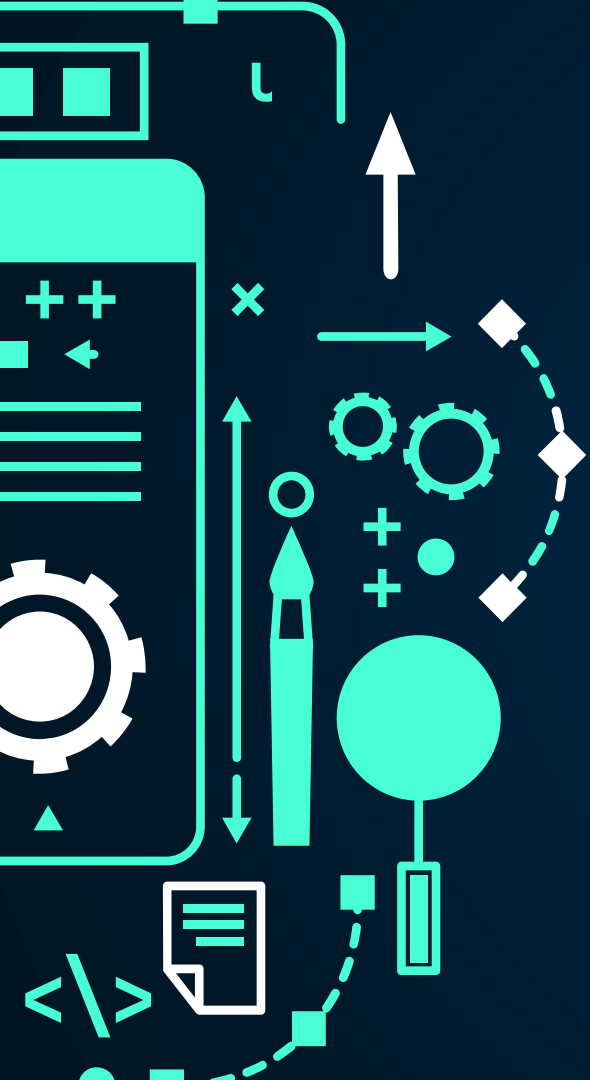
De acuerdo al Excel entregado, identifica los problemas de los datos, aplica las formas de normalización aplicaremos máximo hasta la 4NF y mínimo hasta la 3NF.

Crea el modelo relacional a partir de los datos entregados luego de aplicar la normalización, para garantizar que la estructura de almacenamiento cumple con lo esperado, se evitan datos duplicados, los datos son consistentes, no hay redundancia, y proporcionará mayor velocidad en las consultas.



REFERENCIAS

- Descripción conceptos básicos de normalización: <https://docs.microsoft.com/en-us/office/troubleshoot/access/database-normalization-description>
- Libro: Fundamentos de base de datos 5 edición, Silberschatz



Gracias!