



## **Modelo Relacional**

### ***Normalización - Diseño***



# Nuestro ejemplo: DB de CDs

Base de Datos de CDs en una sola tabla

Band name	CD Title	Record Label	Songs
Stevie Wonder	Talking Book	Motown	You Are the Sunshine of My Life, Maybe Your Baby, Superstition, ...
Miles Davis Quintet	Miles Smiles	Columbia	Orbits, Circle, ...
Wayne Shorter	Speak No Evil	Blue Note	Witch Hunt, Free-Fi-Fo-Fum, ...
Herbie Hancock	Headhunters	Columbia	Chamaleon, Watermelon Man, ...
Herbie Hancock	Maiden Voyage	Blue Note	Maiden Voyage, ...

## ■ Entidades

- Una entidad es una “cosa” u “objeto de importancia” sobre el cual debemos capturar datos.

## ■ Atributos

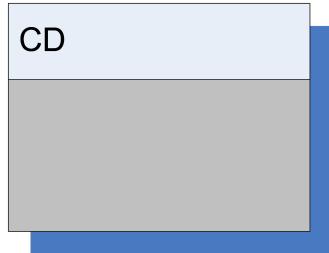
- Un atributo describe información que debemos capturar de una entidad.
- Una entidad puede tener cero ó más atributos que la describen.
- Cada conjunto de atributos describen exactamente a una entidad.
- Cada instancia de una entidad tiene exactamente un valor, posiblemente NULL, para cada uno de sus atributos.
- El valor de un atributo puede ser: Numérico, String de Caracteres, Fecha, Hora, ... algún otro de los tipos básicos.

## ■ NULL

- Nos da la posibilidad de manejar la circunstancia donde nos falta información de algún atributo de una entidad en una forma estándar. Ej. funciones en C no tienen una forma estándar: las que devuelven un puntero -> puntero nulo; las que devuelven un entero positivo -> entero negativo.
- Todos los atributos de una entidad deben tener un valor, o NULL.

# Modelo de Datos (Data Model)

Nuestra entidad CD ...



Por convención los nombres de las entidades van en singular

... y sus atributos según nuestra tabla original



... hasta acá, no hicimos nada.

- Edgar Frank “Ted” Codd propuso el concepto de normalización en diferentes papers durante los '70.
  - Elimina ciertos tipos de redundancia en los datos
  - Evita anomalías
  - Facilita las operaciones de INSERT, UPDATE y DELETE
  - Le aporta claridad al Modelo de Datos



# Primera Forma Normal First Normal Form ó 1NF

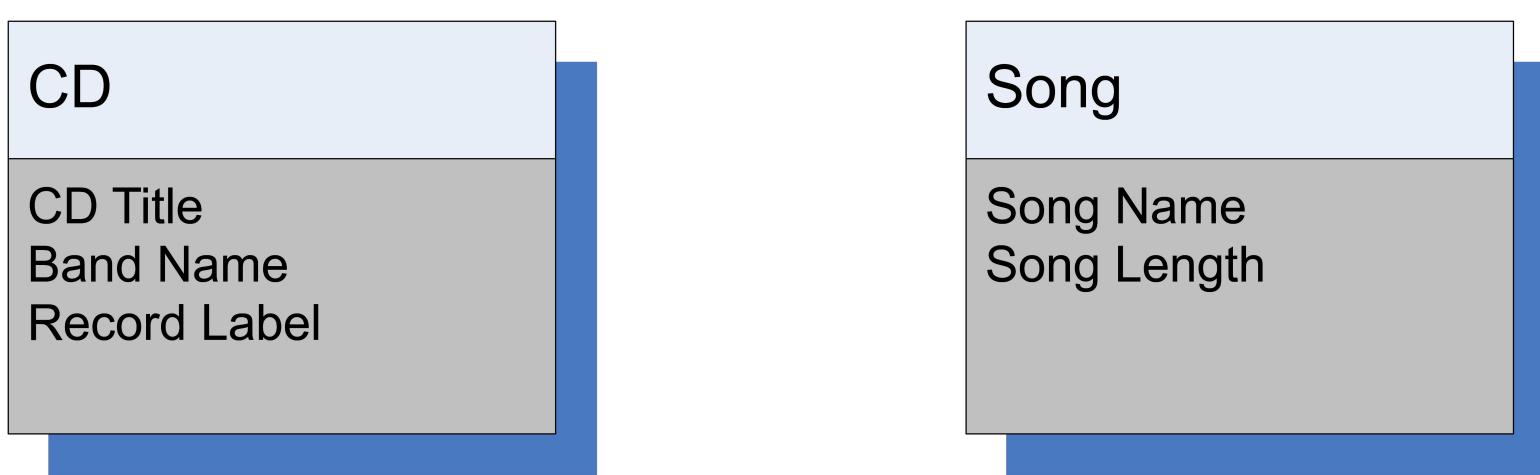
- Una entidad se dice que está en 1NF cuando TODOS LOS ATRIBUTOS TIENEN UN UNICO VALOR
- Para aplicar 1NF a una entidad hay que verificar que todos y cada uno de los atributos de una entidad tengan un único valor para cada una de las instancias de esa entidad

... en nuestro ejemplo

Band name	CD Title	Record Label	Songs
Stevie Wonder	Talking Book	Motown	You Are the Sunshine of My Life, Maybe Your Baby, Superstition, ...
Miles Davis Quintet	Miles Smiles	Columbia	Orbits, Circle, ...
Wayne Shorter	Speak No Evil	Blue Note	Witch Hunt, Free-Fi-Fo-Fum, ...
Herbie Hancock	Headhunters	Columbia	Chamaleon, Watermelon Man, ...
Herbie Hancock	Maiden Voyage	Blue Note	Maiden Voyage, ...



# Primera Forma Normal First Normal Form ó 1NF



... todavía no mostramos cómo se relacionan

# El Identificador Único

## Clave Primaria / Primary Key (PK)

- Antes de discutir relaciones necesitamos ponerle una regla a nuestras entidades.
- Cada entidad debe tener un “Identificador Único” que cumpla con las siguientes reglas:
  - Es ÚNICO para todas las instancias de una entidad
  - Tiene un valor NO NULO para cada instancia y durante todo el tiempo de vida de esa instancia
  - Tiene un valor que NUNCA CAMBIA durante el tiempo de vida de esa instancia



... el problema más crítico es saber elegir el Identificador Único

- El Identificador Único puede ser uno o más atributos de una entidad.
- Errores típicos: elegir atributos como Identificador que puede no ser único.
- Clásico: Nombre de una persona.
- Peor todavía: un ejemplo a continuación ...

## **CURP - Clave Única de Registro de Población (México)**

Código alfanumérico único de identidad de 18 caracteres utilizado para identificar oficialmente tanto a residentes como a ciudadanos mexicanos de todo el país

Los 18 elementos que componen la CURP son los siguientes:

- primera letra y la primera vocal del primer apellido
- primera letra del segundo apellido
- primera letra del nombre de pila; si el primer nombre es José o María, se tomará en cuenta el segundo nombre (en caso de no tener segundo nombre se toma el único)
- fecha de nacimiento (2 últimos dígitos del año, 2 del mes y 2 del día de nacimiento)
- letra del sexo (H o M)
- dos letras correspondientes a la entidad de nacimiento; en el caso de extranjeros, se marca como NE (Nacido Extranjero)
- primera consonante interna (no inicial) del primer apellido
- primera consonante interna (no inicial) del segundo apellido
- primera consonante interna (no inicial) del nombre, y
- dos dígitos, para evitar duplicaciones !!!

Por ejemplo, si una persona se llamará Luis Raúl Bello Mena, con sexo masculino, y hubiera nacido el 13 de Marzo de 1992 en el estado de Colima, su CURP podría ser  
**BEML920313HCMNS09**



## **La Solución**

Agregar un atributo inventado y que no tenga nada que ver con los atributos que describen a la entidad. Con esto podemos tener total control y garantizar que cumplan con las reglas de los Identificadores Únicos.

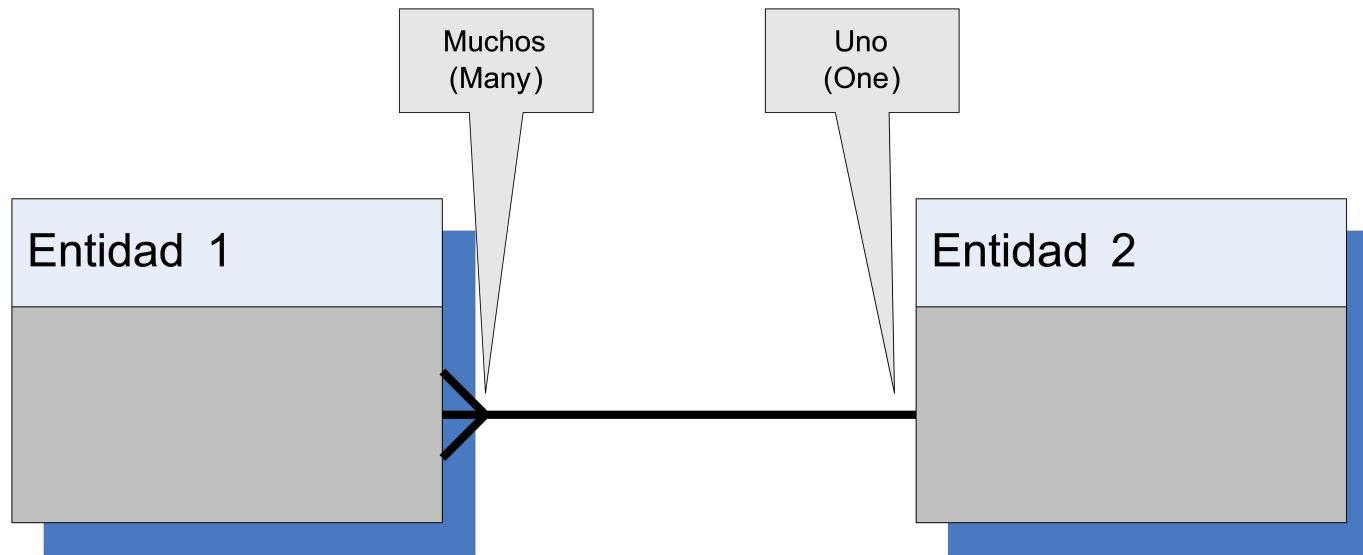


# El Identificador Único Clave Primaria / Primary Key (PK)

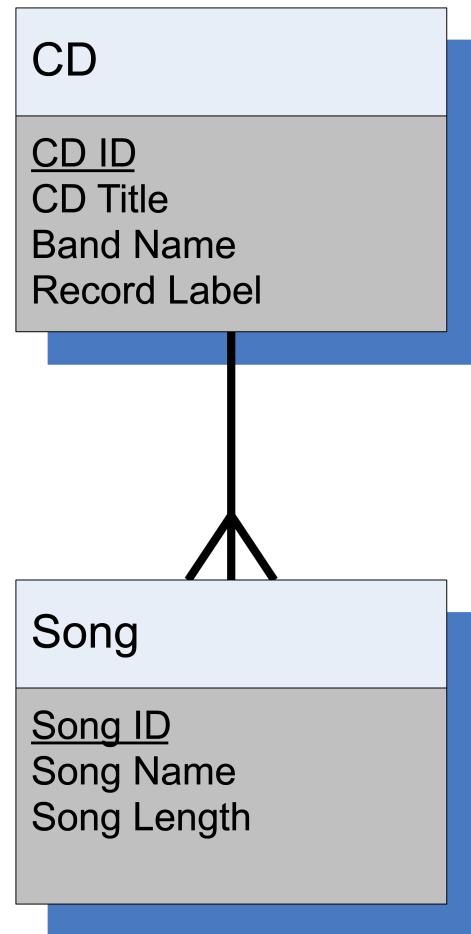
En nuestro ejemplo



- Una relación describe una asociación binaria entre dos entidades
- Puede existir entre una entidad y sí misma (relación recursiva)
- Cada lado de la relación tiene dos componentes: el nombre y el grado o cardinalidad



En nuestro ejemplo





## Segunda Forma Normal Second Normal Form (2NF)

- Se dice que una entidad está en 2NF si ya está en 1NF y dado un identificador único, cualquier atributo que no pertenezca al identificador único depende totalmente del identificador único. Quéeeee???

Habilidades de los empleados

Empleado	Habilidad	Lugar de Trabajo
Jones	Mecanografía	114 Main Street
Jones	Taquigrafía	114 Main Street
Jones	Tallado	114 Main Street
Bravo	Limpieza ligera	77 Industrial Way
Ellis	Alquimia	77 Industrial Way
Ellis	Malabarismo	77 Industrial Way
Harrison	Limpieza ligera	77 Industrial Way

El IU candidato de la tabla es {Empleado, Habilidad}.

El atributo restante, Lugar de Trabajo, es dependiente solo en parte del IU candidato, llamada Empleado. Por lo tanto la tabla no está en 2NF. Observe la redundancia de la manera en que son representadas los Lugares de Trabajo: nos dicen tres veces que Jones trabaja en la 114 Main Street, y dos veces que Ellis trabaja en 73 Industrial Way. Esta redundancia hace a la tabla vulnerable a anomalías de actualización: por ejemplo, es posible actualizar el lugar del trabajo de Jones en sus registros "Mecanografía" y "Taquigrafía" y no actualizar su registro "Tallado". Los datos resultantes implicarían respuestas contradictorias a la pregunta "¿Cuál es el lugar actual de trabajo de Jones?".



## Segunda Forma Normal Second Normal Form (2NF)

Una alternativa 2NF a este diseño representaría la misma información en dos tablas

Empleados

Empleado	Lugar de Trabajo
Jones	114 Main Street
Bravo	77 Industrial Way
Ellis	77 Industrial Way
Harrison	77 Industrial Way

Habilidades de los empleados

Empleado	Habilidad
Jones	Mecanografía
Jones	Taquigrafía
Jones	Tallado
Bravo	Limpieza ligera
Ellis	Alquimia
Ellis	Malabarismo
Harrison	Limpieza ligera



## Segunda Forma Normal Second Normal Form (2NF)

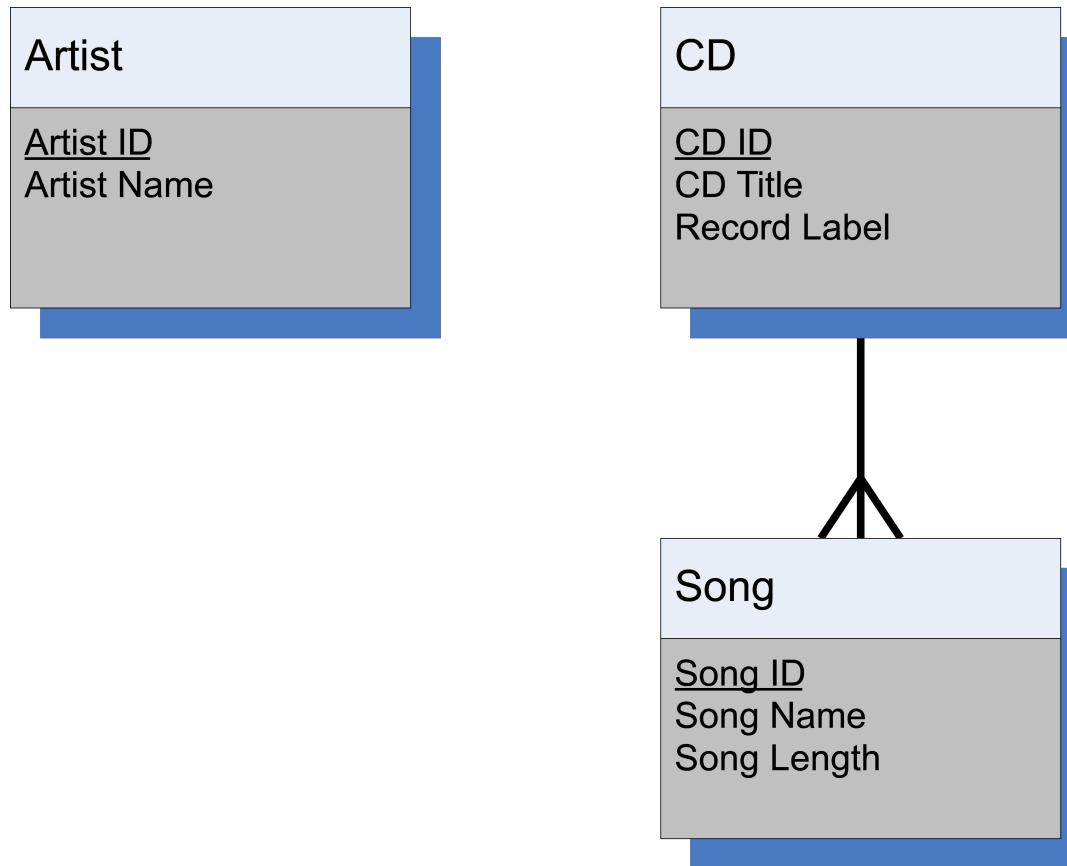
En nuestro ejemplo

Band name	CD Title	Record Label	Songs
Stevie Wonder	Talking Book	Motown	You Are the Sunshine of My Life Maybe Your Baby, Superstition, ...
Miles Davis Quintet	Miles Smiles	Columbia	Orbits, Circle, ...
Wayne Shorter	Speak No Evil	Blue Note	Witch Hunt Free-Fi-Fo-Fum, ...
Herbie Hancock	Headhunters	Columbia	Chameleon, Watermelon Man, ...
Herbie Hancock	Maiden Voyage	Blue Note	Maiden Voyage, ...



## Segunda Forma Normal Second Normal Form (2NF)

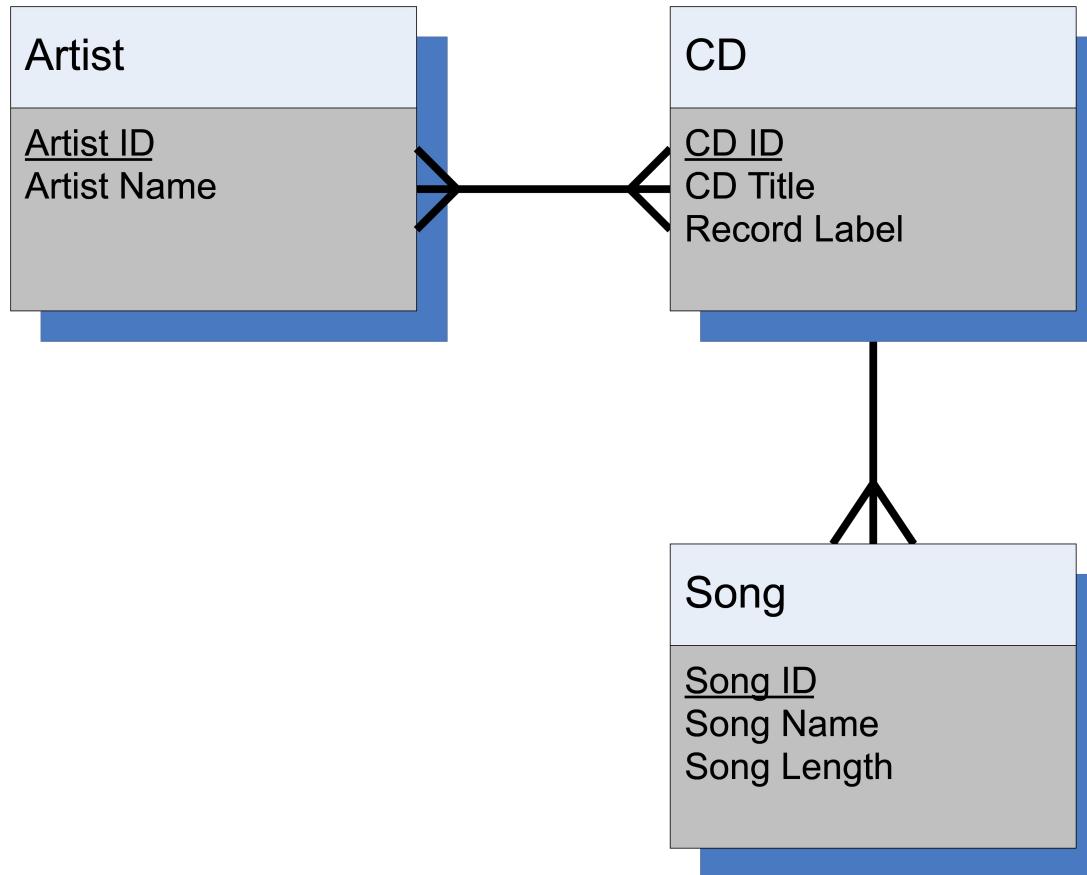
En nuestro ejemplo





# Segunda Forma Normal Second Normal Form (2NF)

En nuestro ejemplo



## ■ **1-a-1 / One-to-One / 1-to-1**

- Poco frecuentes. Por lo general si entre dos entidades existe una relación 1-to-1 estamos hablando de la misma entidad. Por lo tanto los atributos de una de ellas se pueden agregar a la otra. Caso típico: entidades que tiene atributos que rara vez cambian (configuración) y otros que cambian frecuentemente (estado).

## ■ **1-a-Muchos / One-to-Many / 1-to-M**

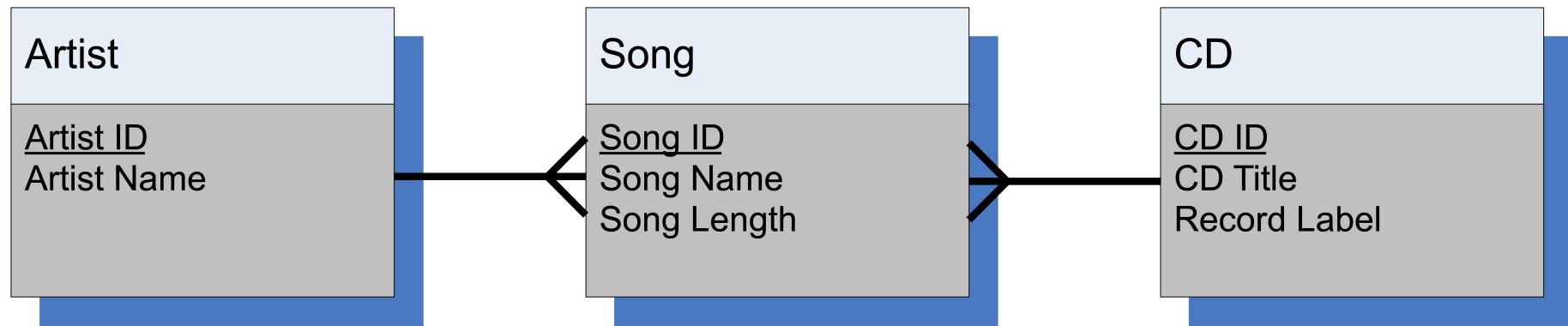
- Son las más comunes de encontrar. “Casi todo” el modelo relacional se basa en este tipo de relaciones. Más adelante veremos cómo se modelan físicamente.

## ■ **Muchos-a-Muchos / Many-to-Many / M-to-M**

- Menos comunes que las 1-to-M. A veces se puede redefinir el modelo para evitarlas. Pero si no ...

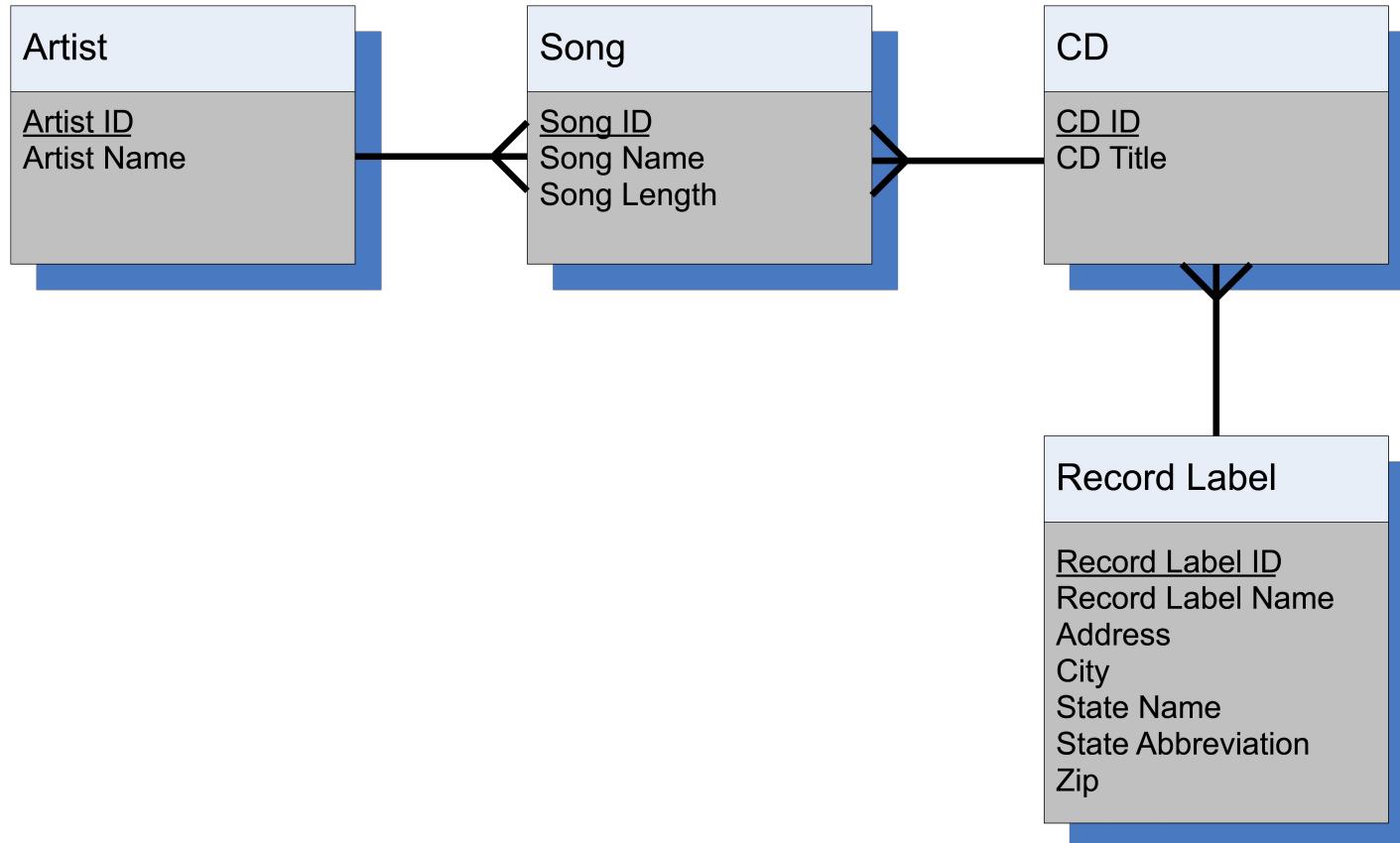
- Se pueden resolver con la siguiente técnica:
  1. Crear una nueva entidad (a veces se la llama Junction Entity). Nombrarla adecuadamente, xEj ArtistCD. En nuestro ejemplo Song puede ser la Junction Entity entre Artist y CD.
  2. Relacionar la nueva entidad con las dos entidades originales. Cada una debe tener una relación 1-to-M con la Junction Entity.
  3. Si la nueva entidad no tiene un Identificador Único obvio, heredar los identificadores de las entidades originales en la Junction Entity y hacer ambos atributos el Identificador único de la nueva entidad.

En nuestro ejemplo





Lo que nos falta en nuestro ejemplo (desagregar Record Label)



# Tercera Forma Normal

## Third Normal Form (3NF)

- Una entidad se dice que está en 3NF si ya está en 2NF y ningún atributo no primario es dependiente “transitivamente” de un Identificador Único. ...Uf!, otra vez?

Ganadores del Torneo

Torneo	Año	Ganador	Fecha de Nacimiento
Indiana Invitational	1998	Al Frederickson	21-Jul-75
Cleveland Open	1999	Bob Albertson	28-Sep-68
Des Moines Masters	1999	Al Frederickson	21-Jul-75
Indiana Invitational	1999	Chip Masterson	14-03-77

La única clave candidata es {Torneo, Año}.

La violación de la 3NF ocurre porque el atributo no primario Fecha de Nacimiento del ganador es dependiente transitivamente de {Torneo, Año} vía el atributo no primario Ganador. El hecho de que la Fecha de Nacimiento del ganador es funcionalmente dependiente en el Ganador hace la tabla vulnerable a inconsistencias lógicas, pues no hay nada que impida a la misma persona ser mostrada con diferentes fechas de nacimiento en diversos registros.



## Tercera Forma Normal Third Normal Form (3NF)

- Para expresar los mismos hechos sin violar la 3NF, es necesario dividir la tabla en dos:

Ganadores del Torneo

Torneo	Año	Ganador
Indiana Invitational	1998	Al Frederickson
Cleveland Open	1999	Bob Albertson
Des Moines Masters	1999	Al Frederickson
Indiana Invitational	1999	Chip Masterson

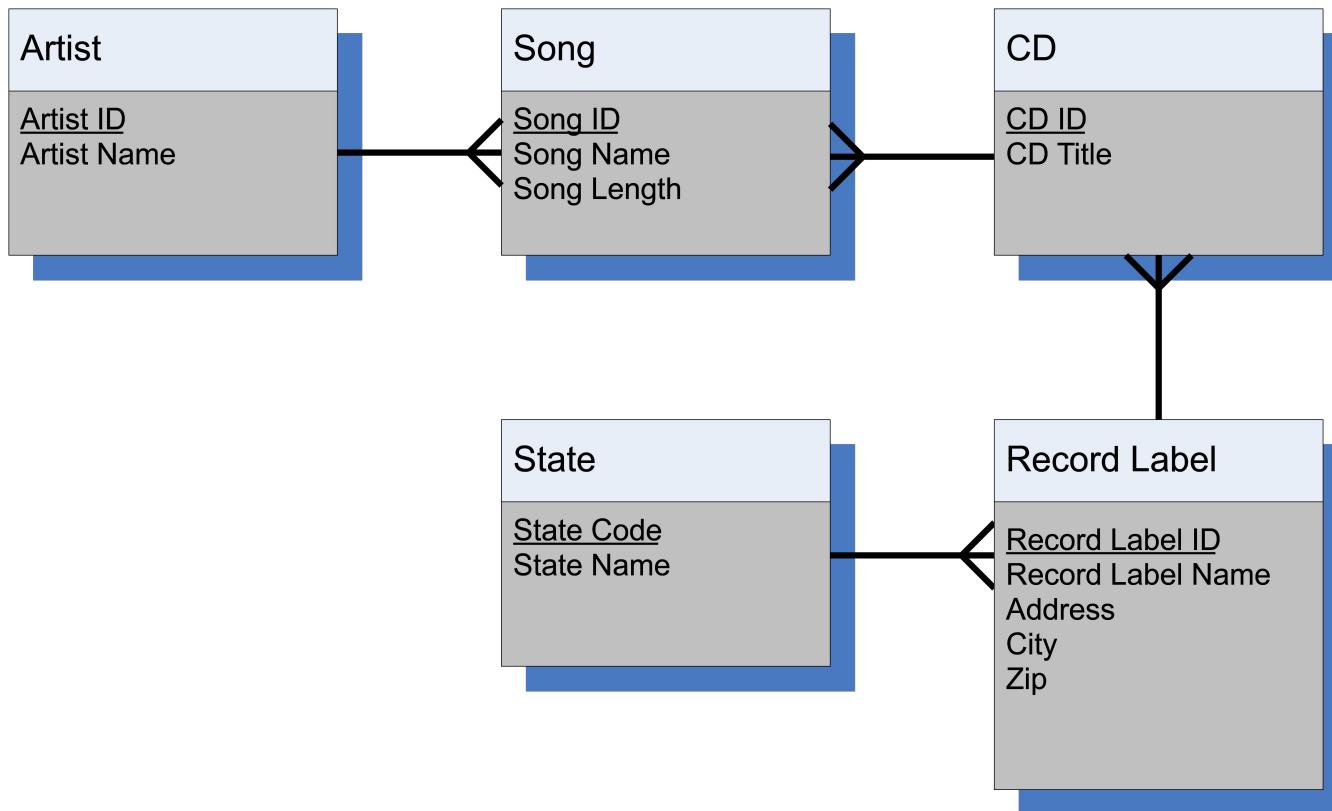
Fecha de Nacimiento del Ganador

Ganador	Fecha de Nacimiento
Al Frederickson	21-Jul-75
Bob Albertson	28-Sep-68
Al Frederickson	21-Jul-75
Chip Masterson	14-03-77



## Tercera Forma Normal Third Normal Form (3NF)

En nuestro ejemplo, esto sucede con los atributos de State incluidos en la entidad Record Label





# Metodología para el Modelado Lógico de Datos

1. Identificar y Modelar las Entidades.
2. Identificar y Modelar las Relaciones entre las Entidades.
3. Identificar y Modelar los Atributos.
4. Identificar los Identificadores Únicos.
5. Normalizar

Este proceso raramente es lineal.

1. Las Entidades se convierten en Tablas.
2. Los Atributos se convierten en Columnas. Hay que elegir el tipo de datos (DataType) apropiado para cada atributo.
3. Los Identificadores Únicos se convierten en Columnas, las cuales no pueden valer NULL. A estas columnas se las llama Primary Key (PK).
4. Las Relaciones se modelan como Claves Foráneas (Foreign Key ó FK)

... para nuestro ejemplo, primero las Entidades, Atributos y PKs

Tabla	Columna	TipoDato	Notas
CD	CDId	INT	Primary Key
	CDTitle	TEXT(50)	
Artist	ArtistId	INT	Primary Key
	ArtistName	TEXT(50)	
Song	SongId	INT	Primary Key
	SongName	TEXT(50)	
RecordLabel	RecordLabelId	INT	Primary Key
	RecordLabelName	TEXT(50)	

## ... ahora las Relaciones

- Para modelar las Relaciones hay que agregar las Foreign Keys (FK).
- Una FK es el Identificador Único de la Tabla del otro lado de la relación.
- La relación más común es la 1-to-M, así que para diseña la se aplica la siguiente Regla:

*Se pone la PK de la Entidad que está del lado 1(One) de la relación dentro de la entidad que está del lado M(Many) como FK.*

En nuestro ejemplo ...

- Poner una columna RecordLabelID en la tabla CD
- Poner una columna CdID en la tabla Song
- Poner una columna ArtistID en la tabla Song

Tabla	Columna	TipoDato	Notas
CD	CDId	INT	Primary Key
	CDTitle	TEXT (50)	
	RecordLabelId	INT	Foreign Key
Artist	ArtistId	INT	Primary Key
	ArtistName	TEXT (50)	
Song	SongId	INT	Primary Key
	SongName	TEXT (50)	
	CDId	INT	Foreign Key
	ArtistId	INT	Foreign Key
RecordLabel	RecordLabelId	INT	Primary Key
	RecordLabelName	TEXT (50)	

... ahora el código simplificado (faltan “forzar” las FK)

```
CREATE TABLE CD (      CD_ID INT NOT NULL,  
                      RECORD_LABEL_ID INT,  
                      CD_TITLE TEXT(50),  
                      PRIMARY KEY (CD_ID)  
                );  
CREATE TABLE Artist (ARTIST_ID INT NOT NULL,  
                     ARTIST_NAME TEXT(50),  
                     PRIMARY KEY (ARTIST_ID)  
                );  
CREATE TABLE Song (   SONG_ID INT NOT NULL,  
                     CD_ID INT,  
                     ARTIST_ID INT,  
                     SONG_NAME TEXT(50),  
                     PRIMARY KEY (SONG_ID)  
                );  
CREATE TABLE RecordLabel (  
                      RECORD_LABEL_ID INT NOT NULL,  
                      RECORD_LABEL_NAME TEXT(50),  
                      PRIMARY KEY (RECORD_LABEL_ID)  
                );
```



**¿Dudas? ¿Consultas?**  
***Muchas gracias***