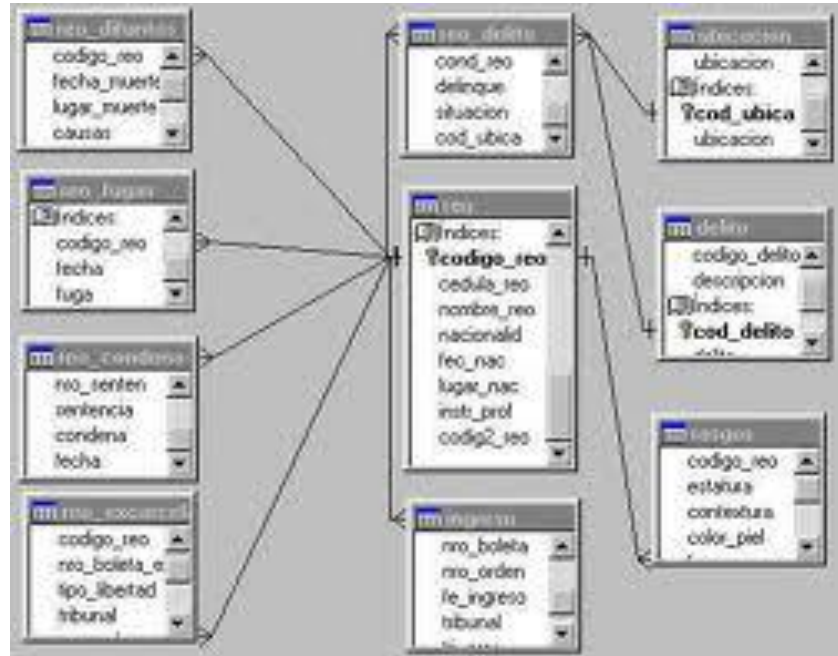




NORMALIZACIÓN DE DATOS



NORMALIZACIÓN DE DATOS



Reglas bien definidas sobre las entidades que presentan anomalías en los atributos que las describen. Se pretende eliminar o disminuir al máximo la redundancia de datos y de datos innecesarios.

TERMINOLOGÍA RELACIONAL

Modelo Relacional	Programador	Usuario
Afinidad	Archivo	Tabla
Tupla (Fila)	Registro	Fila
Atributo	Campo	Columna

EJEMPLO

Atributos



	NumEmpleado	Edad	Sexo	Nombre
T	010110	21	F	Andres
U	010100	22	M	David
P	101000	22	M	Grover
L	201100	21	F	Josue
A	111100	19	M	Mario
S	111101	20	F	Nancy
	111111	19	M	Eduardo

Estructura de la Afinidad:

EMPLEADO(NumEmpleado, Edad, Sexo, Nombre)

DEPENDENCIAS FUNCIONALES

- Es una relación entre uno o más atributos; es decir, si se da el valor de un atributo se puede obtener o buscar el valor de otro.
- En general:
 - El atributo Y depende del atributo X; siempre y cuando el valor de X determine el valor de Y.
- Notación:
 $X \rightarrow Y$ (X se llama Determinante!!)

EJEMPLO

- $\text{IdCliente} \rightarrow \text{NombreCliente}$
 - El atributo IdCliente determina funcionalmente al atributo NombreCliente .
- **Otro ejemplo:**
 - Sea: $\text{Importe} = \text{Cantidad} * \text{Precio}$
 - Entonces:
 $(\text{Cantidad}, \text{Precio}) \rightarrow \text{Importe}$

DEPENDENCIAS FUNCIONALES

- Las dependencias funcionales pueden involucrar grupos de atributos.
- **Ejemplo:**
 - Sea: NOTAS(IdAlumno, IdCurso, Puntaje)
 - La combinación: IdAlumno + IdCurso determina un Puntaje.
 - Entonces:
 - $(\text{IdAlumno}, \text{IdCurso}) \rightarrow \text{Puntaje}$

PARTICULARIDAD

- Si $X \rightarrow (Y, Z)$; Entonces: $X \rightarrow Y \wedge X \rightarrow Z$
- **Ejemplo:**
 - Si $\text{IdAlumno} \rightarrow (\text{NombreAlumno}, \text{Especialidad})$
 - Entonces:
 - $\text{IdAlumno} \rightarrow \text{NombreAlumno}$ y
 - $\text{IdAlumno} \rightarrow \text{Especialidad}$
 - Al Contrario:
 - Si $(X, Y) \rightarrow Z$
 - No es cierto que: $X \rightarrow Z$ ó $Y \rightarrow Z$

PARTICULARIDAD ...

- Ejemplo:
- Si (IdAlumno, IdCurso) \rightarrow Puntaje
- Entonces:
 - IdAlumno por sí mismo NO puede determinar Puntaje, del mismo modo que IdCurso por si solo no puede determinar puntaje.

CLAVES

- Son grupos de uno o más atributos que identifican de modo único a una hilera.
- **Ejemplo:**
 - **DOCENTE**(CodigoID, Nombre, DNI, Telefono, Especialidad)
 - La Clave primaria, se encuentra subrayada!!

CodigoID
Nombre
DNI
Telefono
Especialidad

CLAVE CONCATENADA

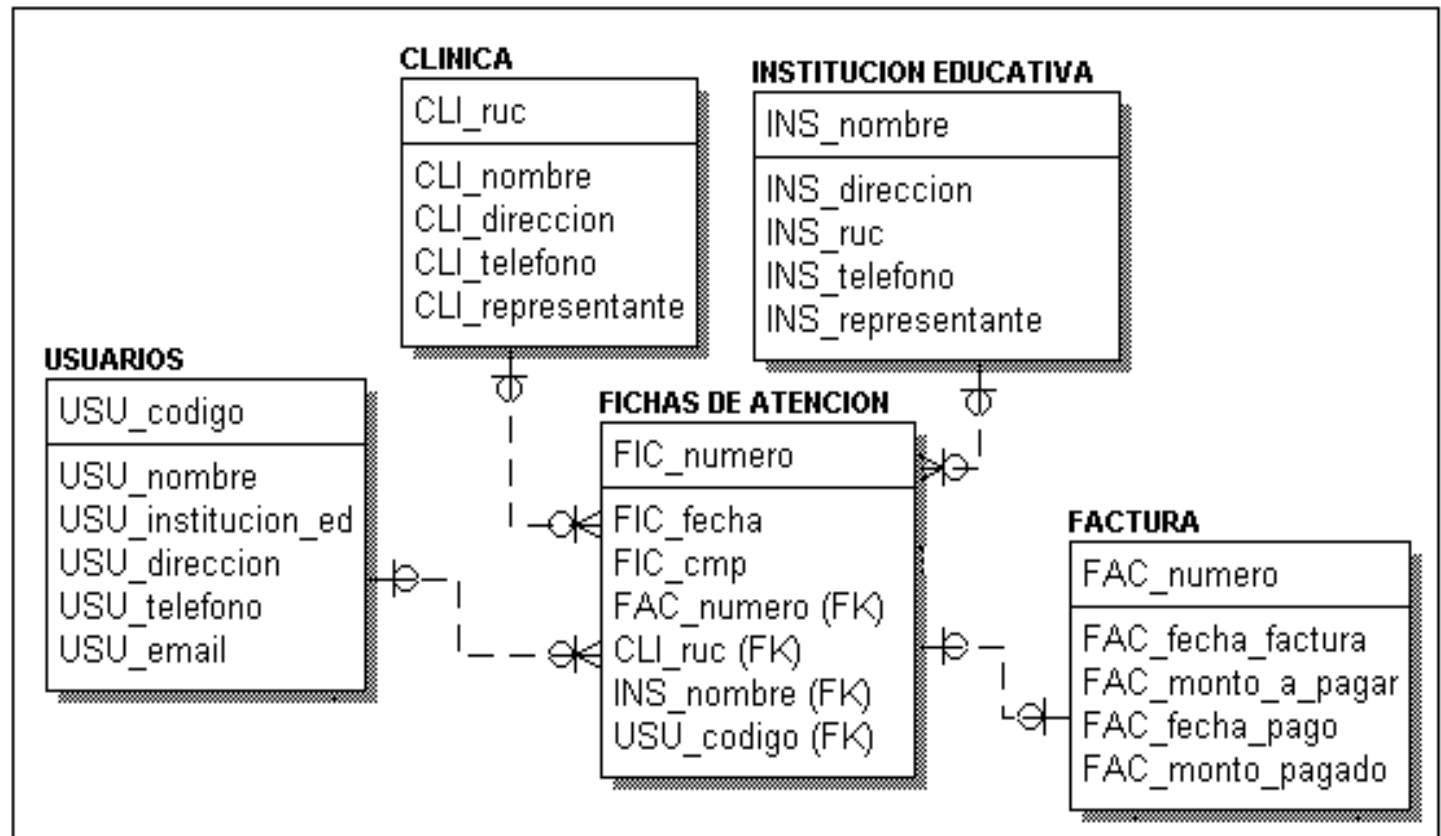
- Son claves formadas por dos ó más atributos.
- Ejemplos:
 - Usaremos los atributos:
NumFactura+CodServicio para registrar el PAGO por un servicio atendido.

PAGO

NumFactura CodServicio
Tratamiento Importe

CLAVE FORÁNEA (FK)

- Se usan para establecer una relación con otra tabla, en la cuál estos atributos se han definido como Clave Primaria.



FORMAS NORMALES

Primera Forma Normal (1NF)

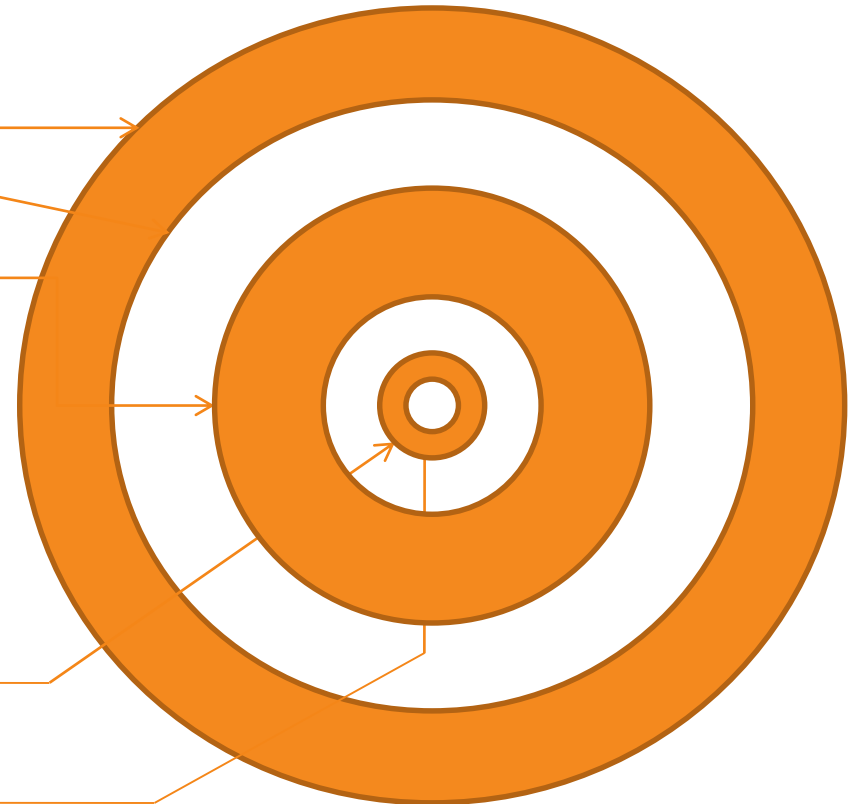
Segunda Forma Normal (2NF)

Tercera Forma Normal (3NF)

Forma Normal de Boyce-Codd (BNF)

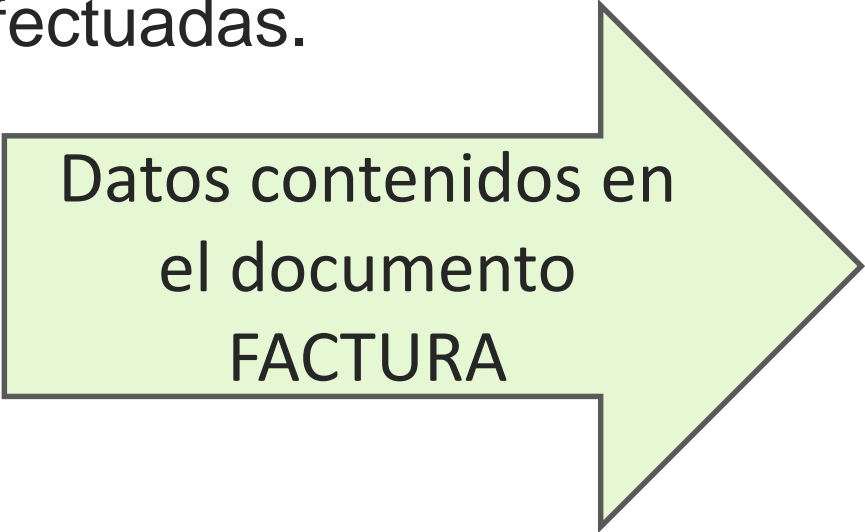
Cuarta Forma Normal (4NF)

Quinta Forma Normal (5NF)



RELACIÓN NO NORMALIZADA

- En un Sistema de Ventas se registra la **FACTURA**; para llevar la cuenta de las ventas efectuadas.



Datos contenidos en
el documento
FACTURA

FACTURA

Numero de Factura
Fecha de Factura
RUC del Cliente
Nombre del Cliente
Direccion del Cliente
Telefono del Cliente
Descripcion del Producto
Cantidad del Pedido
Precio del Producto
Unidad de Medida
Importe
Total Bruto
Descuento
Impuesto de Venta
Total Neto

PRIMERA ANOMALÍA

- En una misma Factura, hay un grupo de Productos distintos que conforman la Línea de Venta.
- Entonces, Aplicaremos la siguiente regla para reducir esta anomalía:

PRIMERA FORMA NORMAL

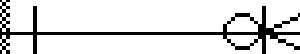
- “Eliminar los grupos repetidos”.
 - Estrategia: *Dividir la tabla FACTURA en dos tablas, de la siguiente manera:*

FACTURA

Numero de Factura
Fecha de Factura
RUC del Cliente
Nombre del Cliente
Direccion del Cliente
Telefono del Cliente
Total Bruto
Descuento
Impuesto de Venta
Total Neto

LINEA DE VENTA

Numero de Factura (FK)
Codigo del Producto
Descripcion del Producto
Cantidad del Pedido
Precio del Producto
Unidad de Medida
Importe



SEGUNDA ANOMALÍA

- No todos los atributos de la tabla LINEA DE VENTA dependen de la clave concatenada!

LINEA DE VENTA

Numero de Factura (FK)
Codigo del Producto

Descripcion del Producto
Cantidad del Pedido
Precio del Producto
Unidad de Medida
Importe

SEGUNDA FORMA NORMAL

- “Los atributos de una tabla serán funcionalmente dependientes de la clave primaria”.

• Esta regla se aplica a las tablas que se identifican a través de atributos de claves concatenadas.

FACTURA

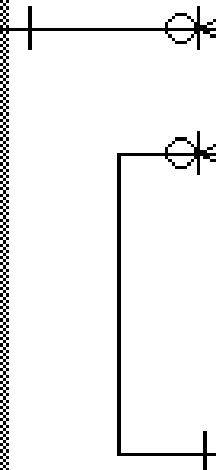
Numero de Factura
Fecha de Factura
RUC del Cliente
Nombre del Cliente
Direccion del Cliente
Telefono del Cliente
Total Bruto
Descuento
Impuesto de Venta
Total Neto

LINEA DE VENTA

Numero de Factura (FK)
Codigo del Producto (FK)
Cantidad del Pedido
Importe

PRODUCTOS

Codigo del Producto
Descripcion del Producto
Precio del Producto
Unidad de Medida



TERCERA ANOMALÍA

- En FACTURA hay atributos dependientes de otros atributos no claves.

FACTURA

Numero de Factura

Fecha de Factura

RUC del Cliente

Nombre del Cliente

Direccion del Cliente

Telefono del Cliente

Total Bruto

Descuento

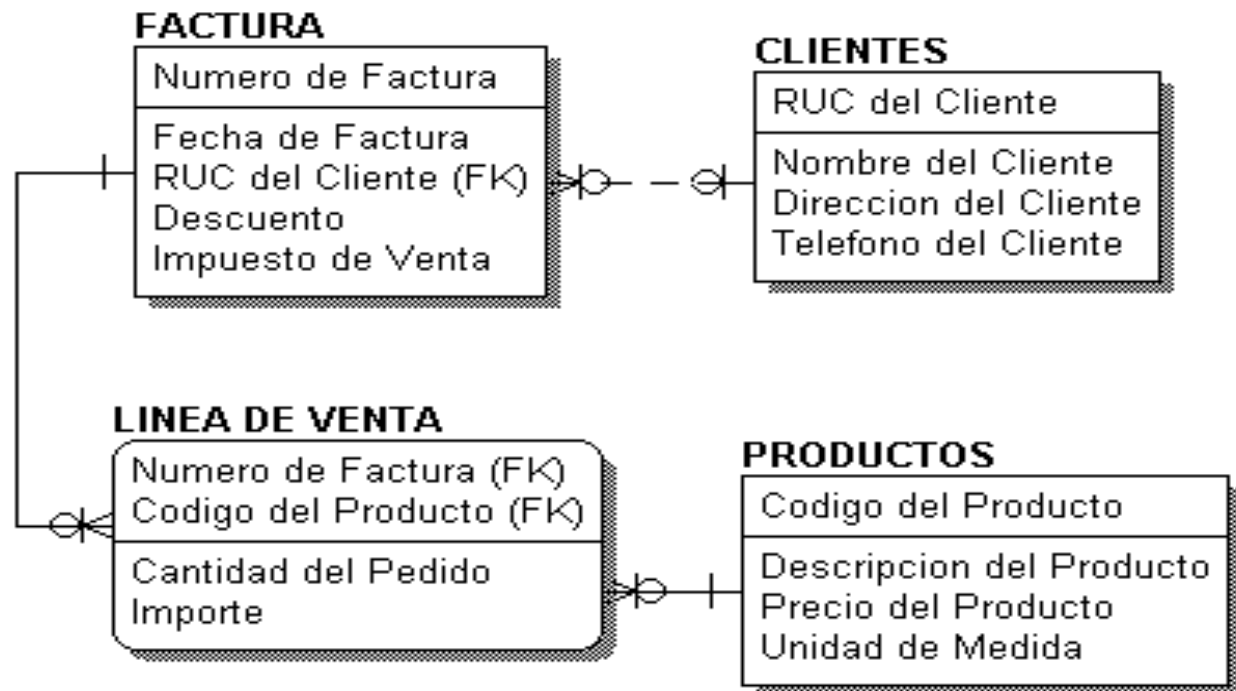
Impuesto de Venta

Total Neto

TERCERA FORMA NORMAL

- “Todos los atributos deben ser funcionalmente dependientes de su clave principal”.

MODELO DE DATOS DE VENTAS



- Además; se pueden eliminar los atributos calculables.

CONSIDERACIONES PARA LA BCNF

- Sea la tabla ASESOR; con las características:
- Un IdAlumno puede tener varias especialidades
 - $\text{IdAlumno} \rightarrow \rightarrow \text{Especialidad}$
- Un IdAlumno puede tener varios Asesores.
 - $\text{IdAlumno} \rightarrow \rightarrow \text{Asesor}$
- IdAlumno por sí mismo NO puede ser una CLAVE !!
 - $(\text{IdAlumno}, \text{Especialidad}) \rightarrow \text{Asesor}$ y
 - $(\text{IdAlumno}, \text{Asesor}) \rightarrow \text{Especialidad}$
 - Cualquiera de estas combinaciones puede ser la Clave.
 - Aparece el término: **Clave Candidata**

ASESOR

IdAlumno
Especialidad
Asesor

OTRAS CONSIDERACIONES ...

- Un Especialidad puede tener varios Asesores
 - Especialidad Asesor
- Un Asesor sólo imparte asesoría en una Especialidad.
 - Asesor Especialidad
- Luego; Asesor es un **DETERMINANTE !!**

ASESOR

IdAlumno
Especialidad
Asesor

ANOMALÍAS

- **De Eliminación**

- Si un Estudiante deja la Escuela; al eliminar la fila perderíamos el hecho de que un Asesor imparte Asesoría en una Especialidad determinada.

- **De Inserción**

- ¿Cómo almacenaremos el hecho de que un Asesor asesora en una Especialidad determinada? Esto no será posible hasta que el Estudiante se inscriba en una materia!!

FORMA NORMAL DE BOYCE y CODD

- Una Afinidad está en BCNF si cada determinante es una clave candidata.
 - En el ejemplo anterior; ASESOR no está en BCNF porque tiene un determinante Asesor que no es una clave candidata.
- Así:
 - ESTUD_ASESOR(IdAlumno + Asesor)
 - ASESOR_ESPECIALIDAD(Asesor, Especialidad)

CONSIDERACIONES PARA LA 4NF

- Sea la tabla ESTUDIANTE con las características:
 - Un estudiante puede inscribirse en varias Especialidades y participar en diversas Actividades.
 - Existe una dependencia multivaluada entre IdAlumno y Especialidad; y, Actividades.
 - IdAlumno $\rightarrow\rightarrow$ Especialidad
 - IdAlumno $\rightarrow\rightarrow$ Actividad

ESTUDIANTE

IdAlumno
Especialidad
Actividad

OTRAS CONSIDERACIONES

ESTUDIANTE

IdAlumno
Especialidad
Actividad

- La Tabla **ESTUDIANTE** está en 2NF porque todo es Clave.
- Está en 3NF porque no tiene dependencias transitivas; y
- Está en BCNF porque no tiene determinantes que no son claves.

ANOMALÍAS

- De Inserción
 - Si un estudiante toma otra especialidad; se debe ingresar una fila para la nueva especialidad, y juntarla con cada una de las actividades del estudiante.
 - Sucede lo mismo si un estudiante se inscribe en una nueva actividad.
- De Eliminación
 - Si un estudiante deja una especialidad se deben eliminar cada uno de los registros que contienen tal materia.

CUARTA FORMA NORMAL

- Una Afinidad está en 4NF si está en BCNF y no tiene dependencias de valores múltiples.
 - Para evitar tales anomalías, se deben eliminar las dependencias de valores múltiples. Esto se hace construyendo dos Afinidades, donde cada una almacena datos para solamente uno de los atributos de valores múltiples.
 - Las afinidades resultantes, no tienen anomalías; así:
- ESTU-ESPECIALIDAD(IdAlumno, Especialidad)
- ESTU-ACTIVIDAD(IdAlumno, Actividad)

QUINTA FORMA NORMAL

- **Dependencias de Unión**
 - Ocurre cuando una tabla tiene dependencia de unión con varias de sus proyecciones y se puede obtener la tabla por medio de la unión de dichas proyecciones.
- **PROYECCION**
 - Creación de una tabla cuyos elementos forman un subconjunto de una tabla dada. Se incluyen todas las filas y algunas columnas.
- **UNIÓN**
 - Formar a partir de 2 tablas, una nueva con todos los campos de una de ellas y los registros de ambas; excepto los repetidos. Ambas tablas deben tener el mismo grado y las mismas columnas.

QUINTA FORMA NORMAL

- Conocida como FN de Proyección-Unión; es un nivel de normalización designado para reducir redundancias que guardan hechos multivalores aislando semánticamente relaciones múltiples.
- Una Tabla se dice que está en 5NF si y sólo si; está en 4NF y las únicas dependencias que existen son las dependencias de Unión de una tabla con sus proyecciones; relacionándose entre las distintas proyecciones mediante la clave primaria o cualquier clave candidata.

QUINTA FORMA NORMAL

- Este es el caso de una Empresa que guarda información de sus empleados con sus datos Personales, Datos Profesionales, Datos Clínicos y Datos de su vivienda.
- Si esta tabla ya está en 4NF; entonces podremos tener las tablas:
 - Empleados-Personales
 - Empleados-Profesionales
 - Empleados-Clinicos
 - Empleados-Vivienda.

EMPLEADOS

EmpleadoID
NumSeguro
Nombre
Direccion
EstadoCivil
FechaNacimiento
Profesion
Universidad
AñoEgreso
TrabajoActual
AñoIngreso
EnfermCronicas
TipoSangre
Estatura
Peso
TipoVivienda
TipoConstruccion
NumHabitaciones
NumBaños
NumPisos

TABLAS EN 5NF

EMPLEADO_PERSONAL

EmpleadoID
NumeroSeguro
Nombre
Direccion
EstadoCivil
FechaNacimiento

EMPLEADO_PROFESIONAL

EmpleadoID
Profesion
Universidad
AñoEgreso
TrabajoActual
AñoIngreso

EMPLEADO_CLINICO

EmpleadoID
EnfermCronicas
TipoSangre
Estatura
Peso

EMPLEADO_VIVIENDA

EmpleadoID
TipoVivienda
TipoConstruccion
NumHabitaciones
NumBaños
NumPisos

RESUMEN

Forma	Características que la define
1NF	Cualquier Afinidad
2NF	Todos los atributos que no son clave dependen por completo de las claves
3NF	No hay dependencias transitivas
BCNF	Cada determinante es una candidata para clave
4NF	No hay dependencias de valores múltiples
5NF	Proyección-Unión
DK/NF	Todas las restricciones en las actividades son consecuencias lógicas de los dominios y las claves

RESUMEN DE LAS AFINIDADES

Referente a las Afinidades uno-a-uno

- Los atributos que tienen una relación uno-a-uno deben aparecer juntos en cuando menos una afinidad. Llame a esta afinidad R y a los atributos A y B.
- A o B deben ser la clave de R
- Un atributo puede agregarse a R si está determinado funcionalmente por A o B.
- Un Atributo que no está determinado funcionalmente por A o B no puede agregarse a R.
- A y B deben aparecer juntos en R, pero no deberán aparecer juntos en otras afinidades.
- A o B deben usarse consistentemente para representar el par en las afinidades diferentes a R.

RESUMEN DE LAS AFINIDADES

Referente a Relaciones Muchos-a-Uno

- Los atributos que tienen una relación muchos-a-uno pueden existir juntos en una afinidad.
- Supongamos que C determina D en una Afinidad S.
- C debe ser la clave de S
- Un atributo puede agregarse a S si está determinado por C
- Un atributo que no está determinado por C no puede agregarse a S.

RESUMEN DE LAS AFINIDADES

Referente a las relaciones Muchos-a-Muchos

- Los atributos que tienen una relación muchos-a-muchos pueden existir juntos en una afinidad.
- Supongamos que los atributos E y f, residen juntos en la afinidad T.
- La clave de T debe ser (E, F)
- Un atributo puede agregarse a t si está determinado por la combinación (E, F)
- Un atributo no puede agregarse a T, si no está determinado por la combinación (E, F)
- Si agregar un nuevo atributo G, expande la clave a (E, F, G); entonces el tema de la afinidad ha sido cambiado. G no pertenece a T o el nombre de T debe cambiarse para reflejar el nuevo tema.