

Unidad II Modelos de Datos Conceptuales

INF-239, ILI-239 Bases de Datos Profesora Cecilia Reyes Covarrubias – Casa Central Diapositivas realizadas con la colaboración Prof. J.Luis Martí – Campus San Joaquín



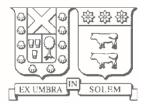
TEMARIO UNIDAD II

- 2.1 Características de los Datos y sus Asociaciones (relaciones)
- 2.2 Tipos de Modelos de Datos
- 2.3 Semántica de los Datos
- 2.4 Notaciones para Modelos de Datos
- 2.5 Ejercicios de Modelos de Datos Conceptual



2.1 CARACTERÍSTICAS DE LOS DATOS Y SUS ASOCIACIONES



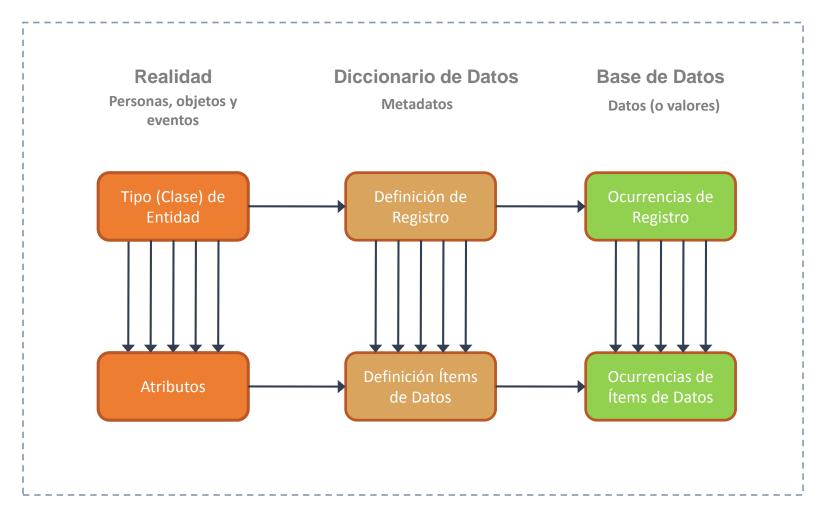


CARACTERÍSTICAS DEL DATO

Para describir qué es un dato, deben considerarse tres niveles de abstracción o estados en que se puede encontrar el dato:

- Realidad constituida por personas, objetos y eventos.
- Diccionario de Datos donde se almacenan los metadatos o definiciones sobre los datos.
- Base de Datos donde se almacenan los valores.





Realidad: comprende el mundo real (sistema, organización), con sus componentes y el medio ambiente en el cual opera.

- Una Entidad es una persona, objeto (cosa, lugar, concepto) o evento, sobre el cual la organización decide coleccionar y almacenar datos.
- Un Tipo (Clase) de Entidades es un conjunto de entidades que poseen características similares, por ejemplo Clientes, Estudiantes, Pacientes.
- Un Atributo es una propiedad de una entidad que se desea registrar. Para cada tipo de entidades, existe un conjunto de atributos de interés para la organización.
- Cada entidad, dentro de una clase de entidades, debe tener al menos un atributo (o una combinación de atributos) que la distinga de otras entidades dentro de su clase, a lo cual se le denomina Clave Primaria (PRIMARY KEY).

Diccionario de Datos: es información acerca de los ítems de datos de una organización, que se almacena y mantiene en el diccionario o repositorio de datos de la misma.

- Un ítem de dato es la unidad de datos más pequeña en una base de datos, y por lo tanto, la unidad de dato más pequeña con significado para el usuario; se refiere a un campo o atributo.
- Por cada ítem de datos, el diccionario de datos guarda:
 - Nombre, largo, tipo, formato.
 - Breve descripción narrativa
 - Rango(s), Dominio, ¿Es valor único (Unicidad)?, ¿Es valor Nulo?
 - Posibles sinónimos
 - Fuente (origen)

Cont. Diccionario de Datos:

- Un Registro es un conjunto de ítemes de datos y/o datos agregados, y corresponde a la definición de una clase de entidades.
- De un registro de datos, el diccionario de datos guarda:
 - Nombre, descripción, tamaño
 - Sus ítems de datos
 - Sus datos agregados

Base de Datos: corresponde al nivel donde físicamente se guardan los datos.

- Existe un registro por cada objeto o entidad del mundo real.
- Se almacenan ocurrencias de datos incluidas sus relaciones o asociaciones.
- Conjuntos de archivos asociados lógicamente entre sí.



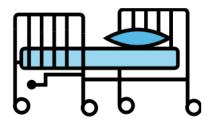
CARACTERÍSTICAS DE LAS RELACIONES

- Una relación es una asociación entre tipos de entidades de una realidad.
- Pueden ser de los siguientes tipos:
 - Una
 - Muchas
 - Condicional
 - En Ambos Sentidos (las más comunes)
 - Recursivas
 - Múltiples asociaciones entre entidades
- Usaremos notación Bachmann para ejemplificar las relaciones o asociaciones.

Asociación Una







Asociación Muchas



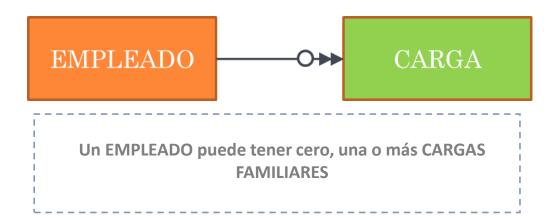
Un EMPLEADO tiene muchas CARGAS.

Pero qué pasa con: EMPLEADOS sin cargas familiares?

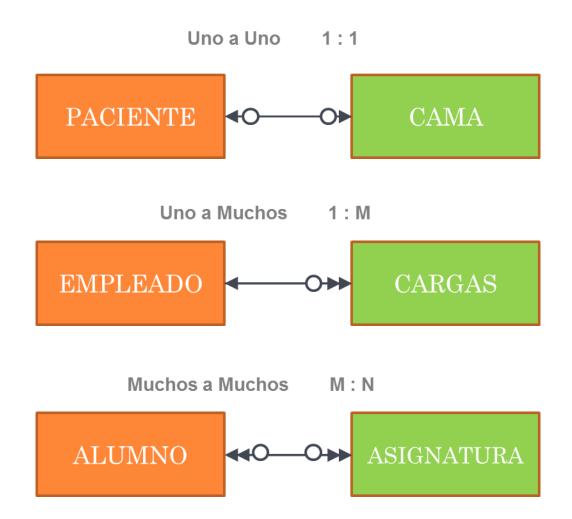
Asociación Condicional



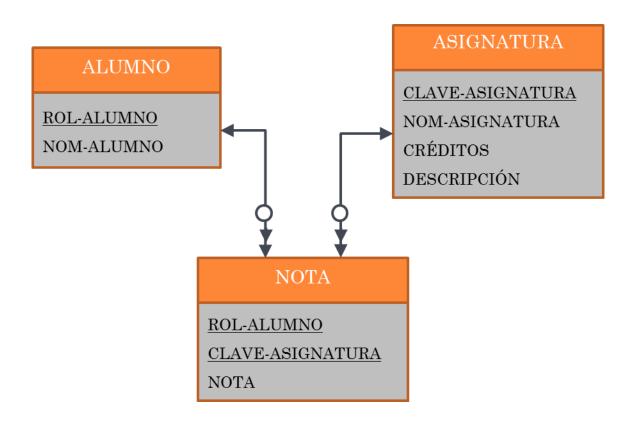
Un PACIENTE puede tener asignada una CAMA (PACIENTE HOSPITALIZADO) o no (PACIENTE AMBULATORIO)



Asociaciones en Ambos Sentidos

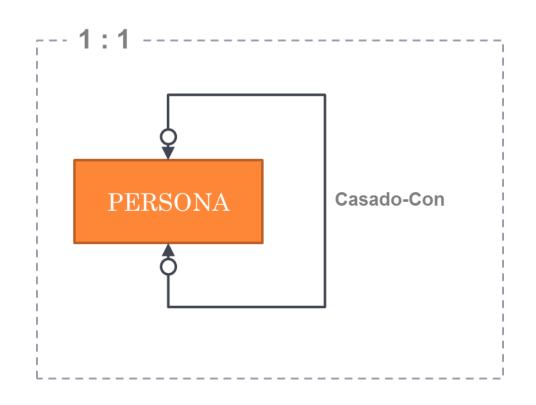


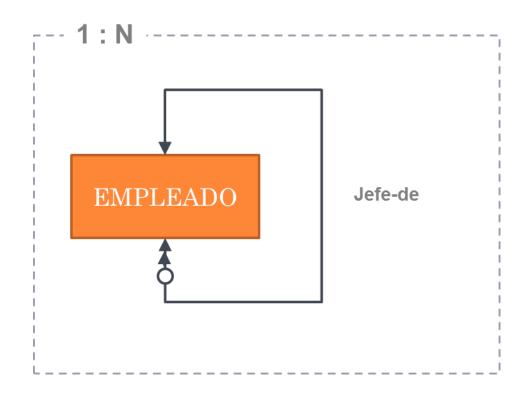
Transformación de Asociación M:N a 1:M



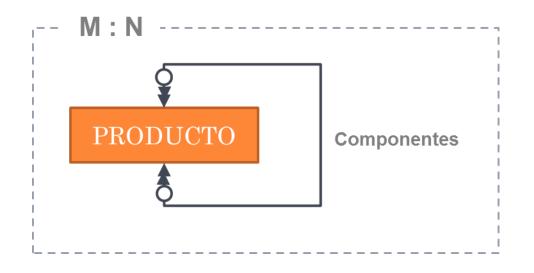
¡¡¡Importancia de la semántica de tiempo!!!

Asociaciones Recursivas (Loop)





Asociaciones Recursivas (Loop)

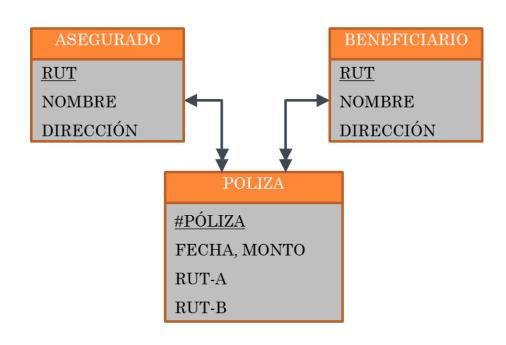


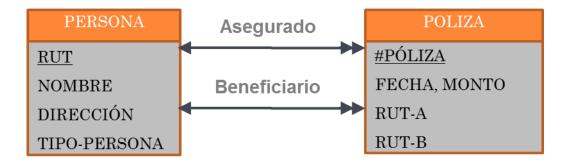
M: N Descompuesta



#PRODUCTO y **#COMPONENTE** son sinónimos

Múltiples Asociaciones entre Entidades





EJERCICIO "Modelando nuestra Sociedad"

- Se pide identificar entidades y asociaciones más relevantes
- Aprendizajes buscados:
 - Saber identificar entidades y asociaciones
 - Importancia del Cliente/Usuario
 - Definición del horizonte de tiempo
 - Importancia del nivel de abstracción del modelo de datos







2.2 TIPOS DE MODELOS DE DATOS





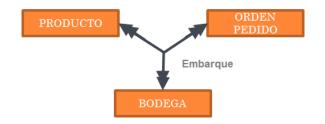
MODELOS DE DATOS

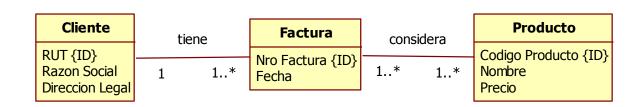
 Un modelo es una representación o abstracción de la realidad: Por ejemplo: modelos de aviones, modelos matemáticos, maquetas, modelos, etc.





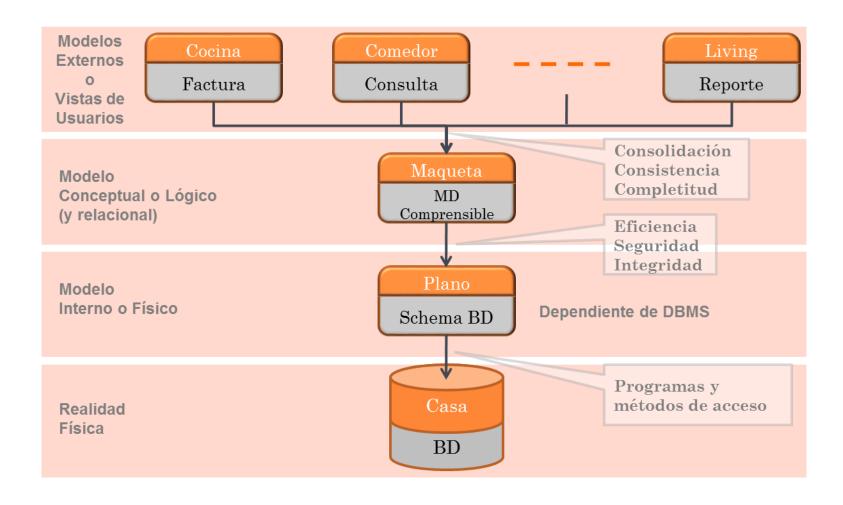
- •Un **modelo de datos** es una representación de entidades del mundo real (personas, objetos y eventos) y sus asociaciones, es decir, es una abstracción de la realidad a partir de sus datos.
- •En un **modelo de datos** se tiene una descripción (abstracta) de todos los datos relacionados con las diversas actividades que en una organización se necesita, para que ésta pueda funcionar adecuadamente.





MODELOS DE DATOS

- El propósito de un modelo de datos es doble:
 - Representar una realidad a través de sus datos.
 - Hacer que los datos sean comprensibles a través de una **notación** estándar.
- Se pueden identificar distintos tipos de modelos de datos según su nivel de abstracción.
- Cada modelo de datos tiene sus restricciones en cuanto a su capacidad de representar una realidad, dependiendo de los tipos de semánticas que puede abarcar.







- Se refiere al significado que tienen los datos en si mismos y entre sí.
- Se representa generalmente por símbolos definidos como convención o norma; veremos la simbología o notación de:
 - Charles Bachman (en dos formatos para asociaciones)
 - Peter Chen (E-R y E-R-E)
 - Grady Booch, Ivar Jacobson y Jim Rumbaugh (UML)
- Se identifican los siguientes tipos de semánticas...

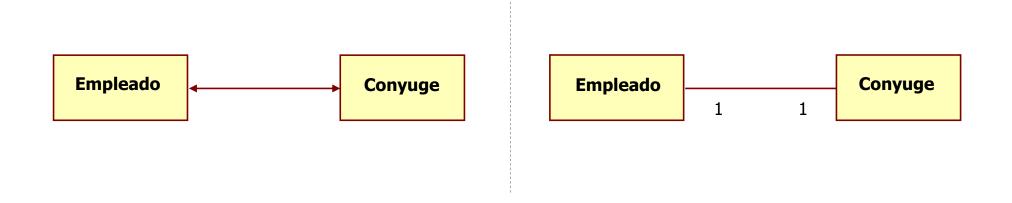
- Cardinalidad (o conectividad o multiciplidad)
- ii. Grado
- iii. Dependencia existencial
- iv. Tiempo
- v. Unicidad
- vi. Herencia (clase-subclase o generalización)
- vii. Categorización (interfaces, herencia selectiva)
- viii. Agregación

i. Cardinalidad

- Se refiere al número de entidades con que otra entidad se relaciona.
- Tipos de multiplicidades:
 - 1:1 (Uno a Uno)
 - 1 : N (Uno a Muchos)
 - M: N (Muchos a Muchos)

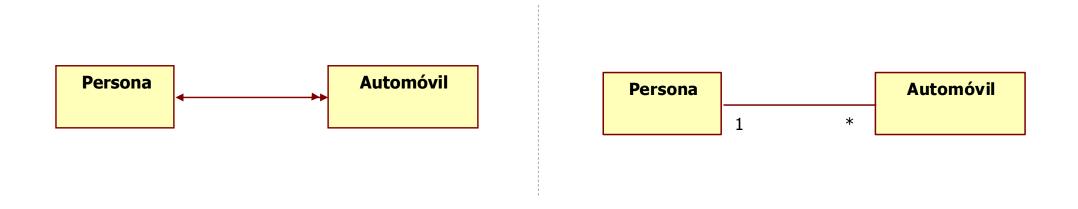
i. Cardinalidad

Asociación 1:1: sucede cuando una entidad X se relaciona sólo con otra entidad, bajo determinada asociación, y esta última sólo lo hace con X.



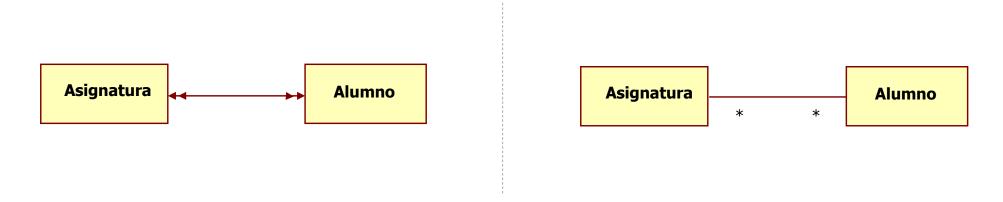
i. Cardinalidad

Asociación 1:M: se presenta cuando una entidad X se relaciona con varias entidades, bajo determinada asociación, pero cada una de éstas sólo lo hace con X.



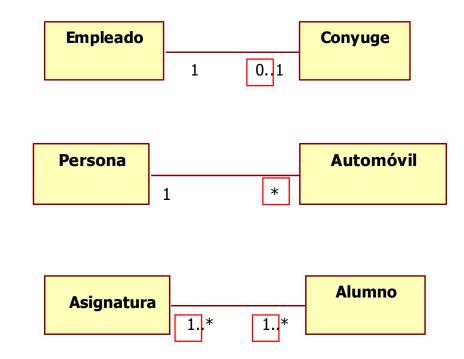
i. Cardinalidad

Asociación M:N: ocurre cuando una entidad X se relaciona con varias entidades, bajo determinada asociación, y cada una de éstas, a su vez, se relaciona con X y, probablemente, otras entidades más del mismo tipo de X.



i. Cardinalidad

Propiedad adicional: obligatoriedad (opcionalidad)

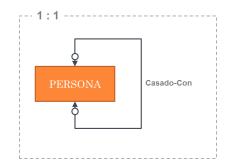


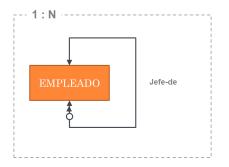
ii. Grado

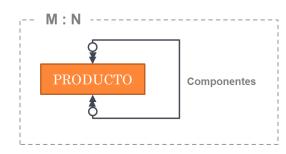
- Se define como el número de **tipos o clases de entidades** que participan en una asociación.
- Tipos de asociaciones según el grado:
 - Unaria
 - Binaria
 - Ternaria
 - N-aria

ii. Grado

Asociación Unaria: asociación que considera un solo tipo de entidades que se relaciona consigo misma.



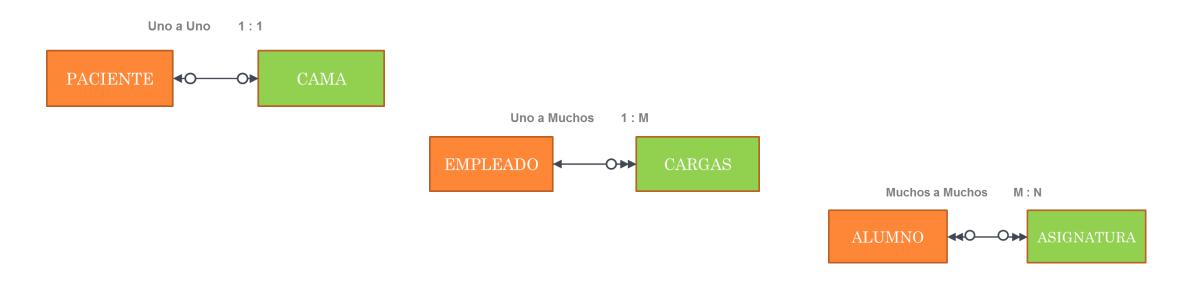




En asociaciones unarias es conveniente aclarar los roles de los vínculos para evitar ambigüedades.

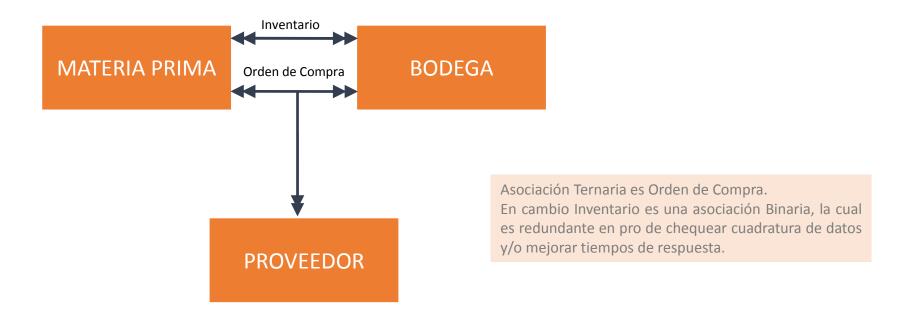
ii. Grado

Asociación Binaria: relación conformada por dos tipos de entidades asociadas entre sí.



ii. Grado

Asociación Ternaria: relación conformada por tres tipos de entidades asociadas entre sí.

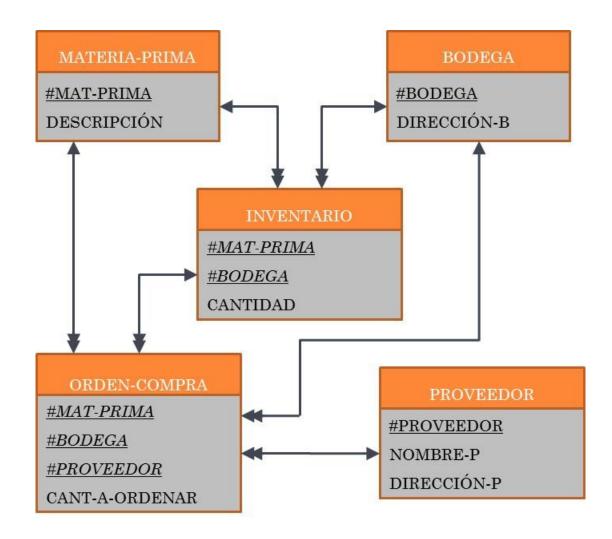


ii. Grado

Ejemplo de la descomposición de las asociaciones ternaria y binaria anteriores que eran MUCHOS: MUCHOS.

PK = Atributos subrayados

FK = Atributos en cursiva



iii. Dependencia existencial

- Se refiere al hecho de que la ocurrencia de una entidad depende de la presencia de otra entidad.
- Dentro de este contexto aparece el concepto de integridad referencial.
- Una entidad no puede existir sin que previamente exista otra (ej.: una Factura no puede existir sin que el Cliente aún no existe).
- Una entidad no puede existir si el que le da sentido es eliminado (ej.: si un Cliente es borrado, deben eliminarse todas las Facturas asociadas a él).

iii. Dependencia existencial

- Beneficios:
 - Integridad referencial asegura calidad de los datos.
 - Acceso fácil a la entidad dependiente.
- En un modelo relacional está presente gracias a:
 - Claves foráneas
 - Claves primarias compuestas dentro de "tipos de entidades débiles" (o entidades de intersección, también conocidas como NUB)

iii. Dependencia existencial

Ejemplo con clave foránea

CLIENTE

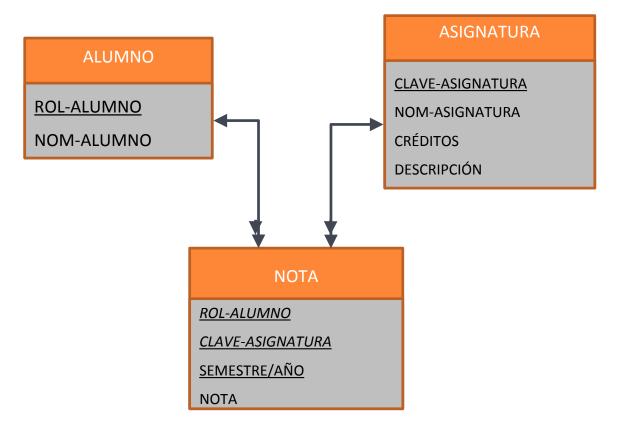
RUT-CLTE	NOM-CLTE	
7.434.000-1	Mario Aldunate	
8.434.900-6	Cecilia Bustos	
8.674.008-1	Felipe Miranda	
9.484.567-9	Andrea Gómez	

FACTURA

NRO- FACTURA	FECHA FACTURA	RUT-CLTE
25678-1	12/12/2008	7.434.000-1
28978-3	15/12/2008	8.434.900-6
29077-5	15/12/2008	7.434.000-1
29979-1	22/12/2008	9.484.567-9

iii. Dependencia existencial

Ejemplo con clave primaria compuesta



EJERCICIO "Semántica de Dependencia – Integridad Referencial"

DEPARTAMENTO

NRO-DEPTO	NOMBRE
100	Producción
150	Ventas
200	Compras

EMPLEADO

NRO- EMP	NOM-EMP	NRO-DEPTO
4340	Mario Aldunate	100
4980	Cecilia Bustos	150
6065	Felipe Miranda	150
6670	Andrea Gómez	120

¿Qué problema de Integridad Referencial existe en estas tablas?

¿Por qué se pudo haber generado este problema?

¿Cómo se soluciona el problema encontrado?

9 www.129rf.com

iv. Tiempo

- Se refiere a cómo el contenido de una BD varía o no en el tiempo (interés por registrar cómo un dato varía en el tiempo versus sólo registrar el valor actual).
- Se tienen las siguientes posibilidades:
 - Estampillas de Tiempo (registrar "la historia")
 - Restricción de Inserción (registrar cuándo se inserta)
 - Restricción de Retención (registrar cuándo se retiene o elimina en forma lógica)

iv. Tiempo

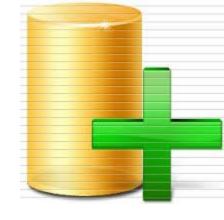
FOR DIE WOHLFAHKTSPFLEGE 100 WOHLFAHKTSPFLEGE 110 WOHLFAHKTSPFLEGE TO BE WOHLFAHT TO BE WOHLFAHT

Estampillas de Tiempo (timestamps)

- Una estampilla o marca de tiempo es un atributo que se incluye dentro de la descripción de una entidad para indicar momento o intervalo de validez de los valores almacenados.
- Ejemplo: Fecha de Pago, Fecha de Contrato (ambos para una validez inmediata, retroactiva, futura).
- Un problema típico de semántica de tiempo es el Manejo de Precio de los Productos.
- Como solución suelen diseñarse dos entidades: una con el valor actual y otra donde se registren los cambios en el tiempo (archivo diferencial).

¡No confundir con los atributos de tiempo definidos por el usuario!

iv. Tiempo



Restricción de Inserción

- Instante del tiempo en que el dato puede ser ingresado a la BD, antes o después de la presencia de otro dato relacionado.
- ¿Cuándo se inserta? Antes o después de un cierto evento en el tiempo.
- Ejemplo: pago de dividendo antes o después de la fecha de vencimiento.

¡Es diferente a la semántica de Dependencia (Integridad Referencial)!

iv. Tiempo

Restricción de Retención

- Periodo de tiempo en el cual tiene sentido o vigencia un dato.
- ¿Hasta cuándo lo retengo en la BD? ó ¿cuándo lo elimino porque ya no aporta en la operación cotidiana? ó ¿cuándo le cambio su vigencia?. Se puede definir la fecha en que queda retenido o conservado para la historia.
- ¿Cuándo se elimina? Antes o después de un cierto evento en el tiempo.
- Ejemplo: un alumno cuándo se titula, ¿debe ser eliminado de la base de datos o retenido?

¡ Importancia de la eliminación lógica!



v. Unicidad

Se refiere a la presencia única (o exclusiva) de un dato, entidad o asociación.

Unicidad por Identificador

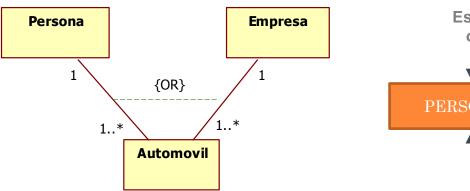
Persona

RUT {ID}

Nombre

Domicilio

Exclusividad









vi. Herencia

 Permite ver un conjunto de entidades, con propiedades similares (atributos, operaciones), como un mismo tipo de entidades.

Conceptos asociados:

- Supertipo(clase): tipo de entidad más general (generalización)
- **Subtipo(clase):** subconjunto de un (super)tipo que se ha dividido. Hereda (reusa) las propiedades definidas en el supertipo del cual se origina.
- Asociación del tipo "IS-A" "Es una(a)".





vi. Herencia

Propiedades de la herencia:

- Cobertura: si todas las entidades del supertipo pueden ser clasificadas por, al menos, uno de los subtipos, la jerarquía de herencia tiene cobertura completa (total, exhaustiva).
- Exclusividad: si una entidad sólo puede pertenecer a uno de los subtipos a la vez, la jerarquía de herencia es exclusiva.
- **Dinamicidad:** aunque una entidad pertenezca a uno de los subtipos, es posible que también sea de otro.

ABUELO

vi. Herencia

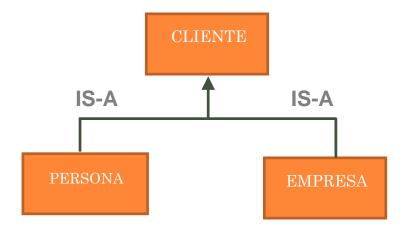
Cobertura Total o Exhaustiva: en el Banco sólo existen 2 tipos de CLIENTES, y un CLIENTE pertenece a uno de esos tipos.

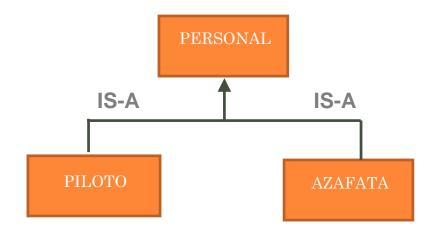
Exclusividad: un CLIENTE es del tipo PERSONA ó EMPRESA, pero no ambos.

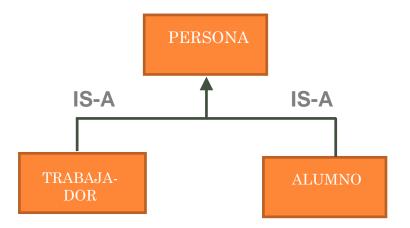
Cobertura No Exhaustiva: en una Línea Aérea hay 3 tipo de PERSONAL (Piloto, Azafata, Tierra), pero sólo de 2 de ellos interesa tener atributos propios.

Exclusividad: Se es PILOTO ó AZAFATA, pero no ambos.

Dinamicidad (o no exclusividad): una PERSONA puede trabajar y estudiar a la vez, es decir, ser del tipo TRABAJADOR Y ALUMNO.

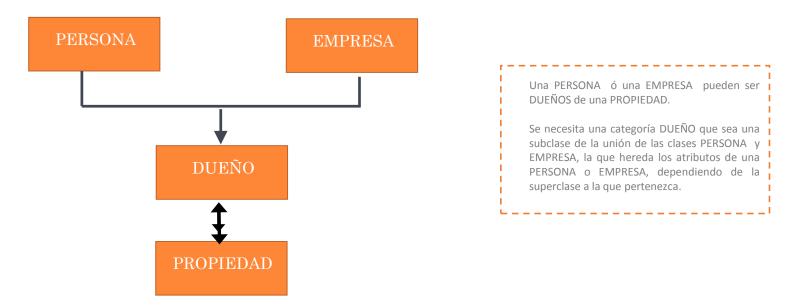






vii. Categorización (Herencia Selectiva)

- Modela una relación de supertipo(clase) subtipo(clase), donde los supertipos (clases) tienen diferentes tipos de datos (incluso pueden tener diferente identificador).
- Se le llama Categoría al subtipo(clase).



viii. Agregación



- Es una colección de entidades, normalmente, de diferentes tipos.
- Conceptos asociados:
 - Objeto compuesto: que agrupa a la colección (o entidad agregada, que está compuesta de otras entidades)
 - Objeto componente: contenido en el anterior (o entidad que es parte de la anterior).
 - Asociación del tipo "IS-PART-OF "Es parte de", o "(con)tiene".
 - Ejemplos: cuerpo humano, computador, automóvil, factura.

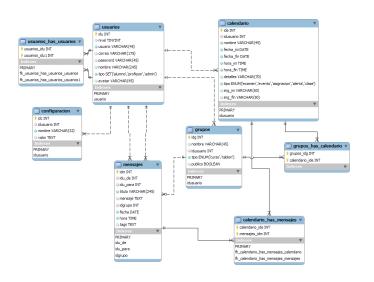
2.4 NOTACIONES PARA MODELOS DE DATOS





NOTACIONES MÁS COMUNES

- Diagrama de Bachmann (ya vistos)
- Diagramas Entidad Relación básico (E-R) y extendido (E-R-E)
- Diagrama de Clases de UML (Unified Modeling Languague)



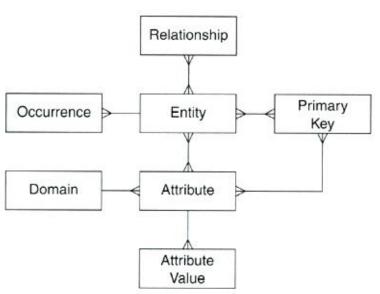
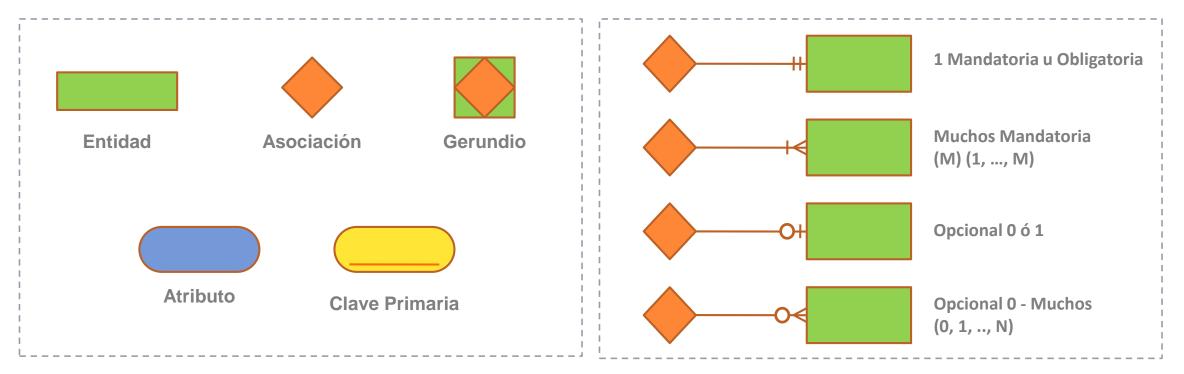
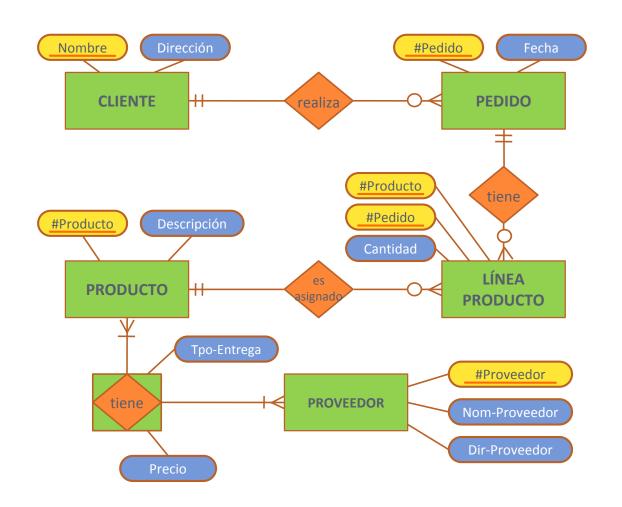


DIAGRAMA ENTIDAD - RELACIONAMIENTO

- Entity-Relationship Model (ER Model), propuesto por Peter Chen (1976).
- Existen extensiones al modelo original.
- La notación original o básica es la siguiente:



EJEMPLO DIAGRAMA E-R



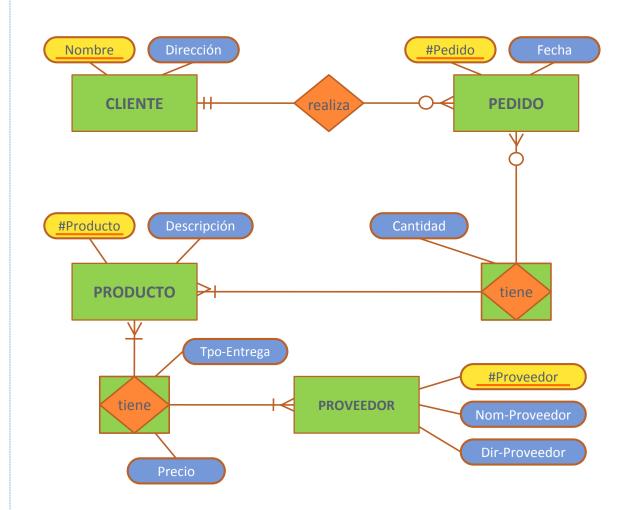
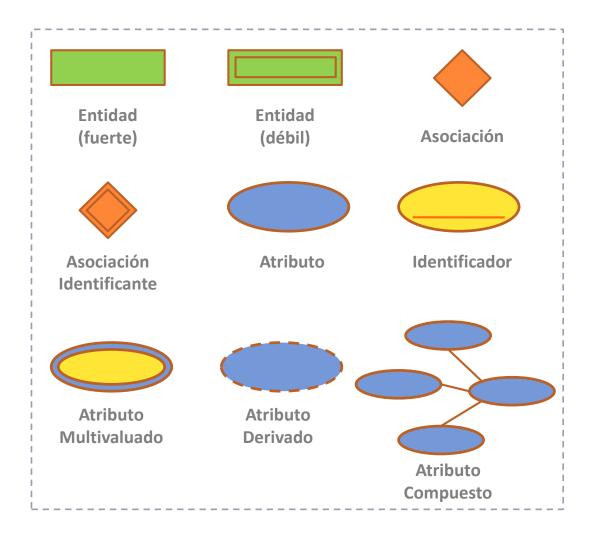
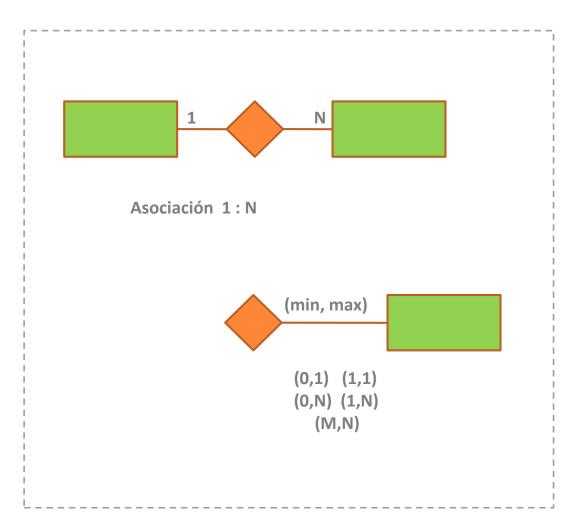


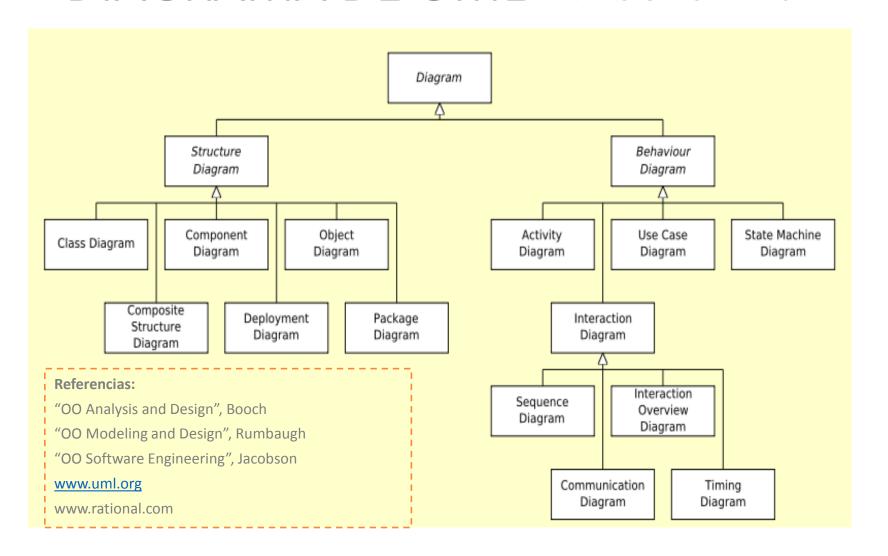
DIAGRAMA E - R - E





- UML (Unified Modeling Languague) permite modelar, construir y documentar un sistema
 Orientado a Objetos (OO).
- UML ha sido impulsado por los 3 autores más importantes de la OO: Grady Booch, Ivar Jacobson y Jim Rumbaugh.
- Se reconocen dos versiones: UML 1.0 (1994) y UML 2.0 (2005).
- El Diagrama de Clases es uno de los tantos diagramas de UML.
- Es un tipo de diagrama estático que describe la estructura de un sistema mostrando sus clases, atributos y asociaciones entre ellos.
- Se utilizan durante el proceso de análisis y diseño de sistemas, donde se crea el diseño conceptual de la información que se manejará, y los componentes que se encargarán del funcionamiento y la relación entre ellos.

DIAGRAMA DE UML - VERSIÓN UML 2.0



CLASE

Atributos (Variables)

Métodos

(Funciones)

STACK

datos : Array

tope : Integer

push(nro)

pop() : Integer

[♦]largo() : Integer

Visibilidad nombre : tipo = valor_inicial {Propiedad}

Visibilidad	Descripción	Atributo	Método
+ pública	visible desde todos lados.	◆	♦
# protegida	no accessible desde fuera de la clase, si desde métodos de la clase y de las subclase derivadas (herencia).	%	P
- privada	solo será accessible desde dentro de la clase (solo sus métodos lo pueden accesar).	>	

Propiedad

{unique}

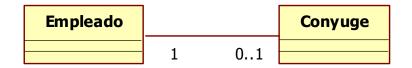
{not null}

...

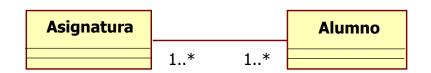
Ejemplo:

+ Apellido : String = "Unknown" {not null}

Cardinalidad o multiplicidad







1 : uno obligatorio

* : cero o muchos

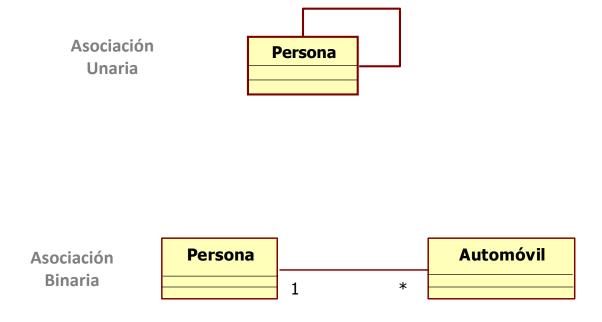
10..*: 10 o más

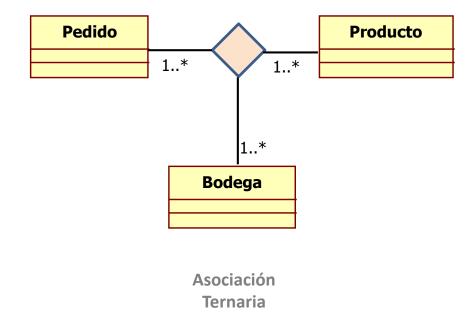
1..40 : de 1 a 40 ocurrencias

3, 5, 8 : 3 ó 5 u 8 ocurrencias

Si no va un número o símbolo equivale a 1

Grado



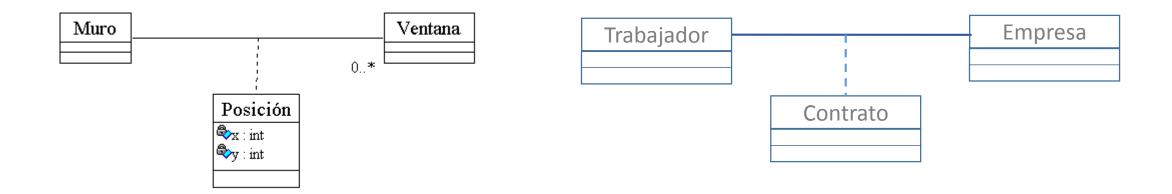


Nombres y Roles



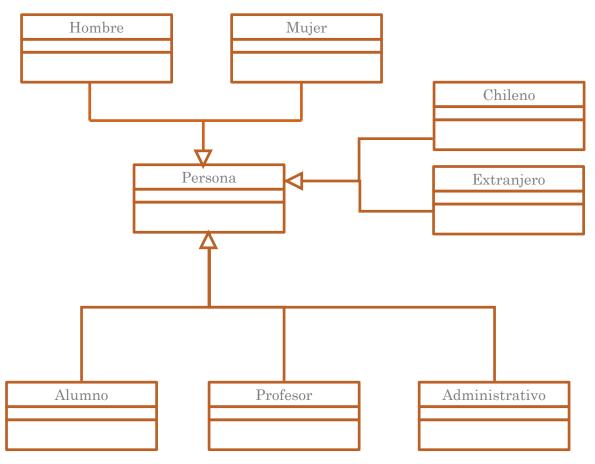


Clase de Asociación



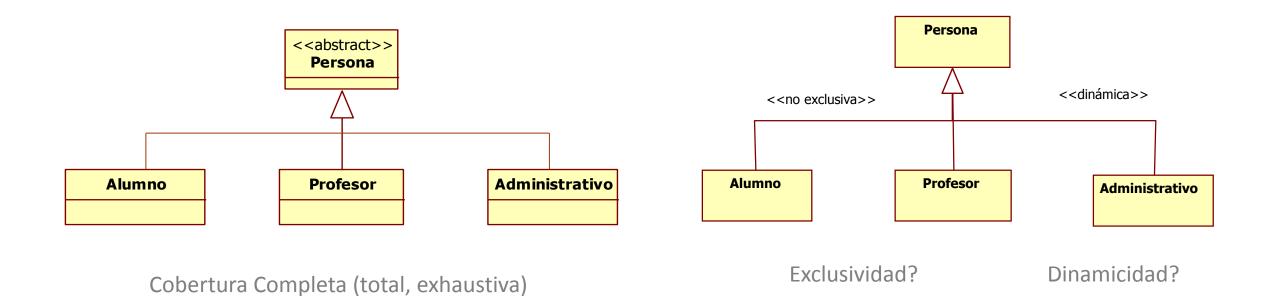
Una clase de asociación se caracteriza por ser a la vez una asociación y una clase, lo que permite añadir atributos y métodos a la asociación

Herencia



Herencia: se sugiere unir todas las subclases obtenidas a partir de un mismo criterio, hacia la superclase en cuestión.

Herencia

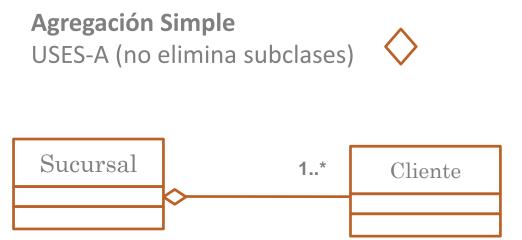


Usuario Herencia # login: String + pública # password: String # protegida # nombre: String - claveDB privada + getNombre() + login() + load() Empleado Alumno # sueldo: Integer # curso: String # titulo: String - promedio : Single + pagarSueldo() nuevaAnotacion() nuevaNota() + registrarAsistencia() rendirPrueba() Profesor Administrativo - sueldo : Integer - cargo: String - titulo: String + marcarTarjeta() + fijarPrueba() Secretaria Director Portero + limpiarPasillo() +matricularAlumno() + contratarProfesor()

Agregación

Agregación Compuesta (composición)
HAS-A (eliminación en cascada)





Restricciones

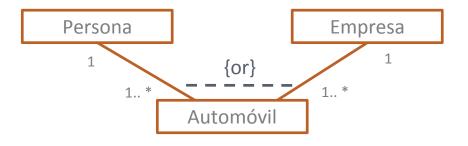
Restricción a nivel de atributos

Empleado ... edad sexo

$$\{\text{sexo} = 'M' ==> \text{edad in } [18..65]\}$$

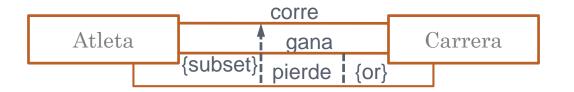
$$\{\text{sexo} = 'F' ==> \text{edad in } [18..60]\}$$

Restricción a nivel de asociaciones





Ejemplo Restricciones







2.5 EJERCICIOS MODELOS DE DATOS CONCEPTUALES





LINEA AEREA VOLARE



La aerolínea →VOLARE→ tiene graves problemas para manejar sus datos, dada la diversidad de ellos y el gran volumen que representan. Por ello necesita se le diseñe un modelo de datos con notación Bachmann que le permita tener una visión preliminar del contenido que tendría una base de datos que los soporte.

→VOLARE→ posee 200 aviones de distinto tipo y capacidad, al comenzar el año se realiza una programación de los aviones, asignándolos a distintos vuelos. Un vuelo tiene definido un número que lo identifica en forma única, la fecha y hora de partida; y las posibles escalas; además del personal que le corresponderá viajar en él (pilotos y azafatas).

Los pasajeros pueden reservar o comprar sus pasajes para un determinado vuelo, en la misma línea aérea o en alguna agencia de viajes. Si lo reservan deben abonar un 50% del valor, devolviéndosele el dinero sólo si avisa con más de 48 horas de anticipación que no lo usará. Tanto al reservarlo como al comprarlo se le asigna inmediatamente el asiento que ocuparán y sus aeropuertos de origen y destino. Los pasajes pueden ser de clase turista o clase ejecutiva, existiendo un costo diferente.

La línea aérea →VOLARE→ lleva un control de las ventas realizadas por las agencias de viajes, ya que para motivarlos les ofrece pasajes gratis en función del monto semestral de ventas que registre la agencia.

EL CENSO



El Censo se lleva a cabo cada 10 años y es el operativo estadístico más amplio que se realiza en Chile. Está a cargo del INE (Instituto Nacional de Estadística) donde se mantiene un registro histórico para hacer las comparaciones de rigor que permitan ver la evolución de las principales preguntas. Las preguntas suelen ser las mismas o se agregan nuevas de un censo a otro.

Por cada hogar encuestado existe un formulario único (cuestionario censal) que tiene un número de folio, la fecha de la entrevista, el RUT del encuestador y la dirección del hogar. Y de ahí el conjunto de datos a recolectar, los cuales están separados en secciones: vivienda (tipo, material, etc.), hogar (en una vivienda pueden haber varios hogares o familias), personas que residen en ese hogar (RUT, nombre, parentesco con jefe de hogar, sexo, nivel educacional, etc.); y en el caso de las mujeres de 15 años o más, hay un conjunto de datos adicionales solicitados para ellas.

Para el polémico último Censo del 2012, el INE contrató y capacitó a casi 14 mil personas como encuestadores. Por cada uno de ellos, se llenó una ficha con sus datos personales, a los cuales cualquier ciudadano podía acceder a través del sitio web del INE para verificar la identidad del censista que estaba entrando a una casa. Además el INE para organizar a los encuestadores en sus visitas, dividió a las comunas del país en distritos, y éstos en manzanas, donde cada manzana está compuesta por un número de viviendas a encuestar.

Suponiendo que lo anterior es una interpretación del censo, se les pide ayudar al INE a obtener un modelo de datos conceptual usando simbología E-R.

LOS BOMBEROS

Suponer que el Cuerpo de Bomberos de Viña del Mar tiene el siguiente Diagrama de Clases para registrar los incendios que suceden en la ciudad, pero no lo entienden. Explicarles los conceptos de:

- Asociación condicional
- Herencia
- Clase de Asociación
- Agregación



