

FUNDAMENTOS DE BASES DE DATOS

Examen Febrero 2013 - SOLUCION

La duración del examen es de 3 horas.

Presentar la resolución del examen:

- Con las hojas numeradas y escritas de un solo lado.
- Con la cantidad de hojas entregadas en la primer hoja.
- Con cédula de identidad y nombre en cada hoja.
- Escrita a lápiz y en forma prolija.
- Comenzando cada ejercicio en una nueva hoja

Ejercicio 1 (25 puntos)

Se desea modelar la realidad de distintas estancias. Las estancias son identificadas por su dirección, la cual está dada por la localidad, la ruta y el km. Además, interesa el nombre de la estancia y, de todos los animales que ésta tiene, se registra el tipo (ovino, bovino, caprino, etc.) y la cantidad (por ejemplo, del tipo ovino hay 150 ejemplares). Algunas estancias son turísticas y de ellas se conoce la información de las habitaciones, el nombre, la cantidad disponible y capacidad en cantidad de personas de cada habitación (por ejemplo, de la habitación con nombre "simple" hay una cantidad de 10 y cada una tiene capacidad para 1 persona).

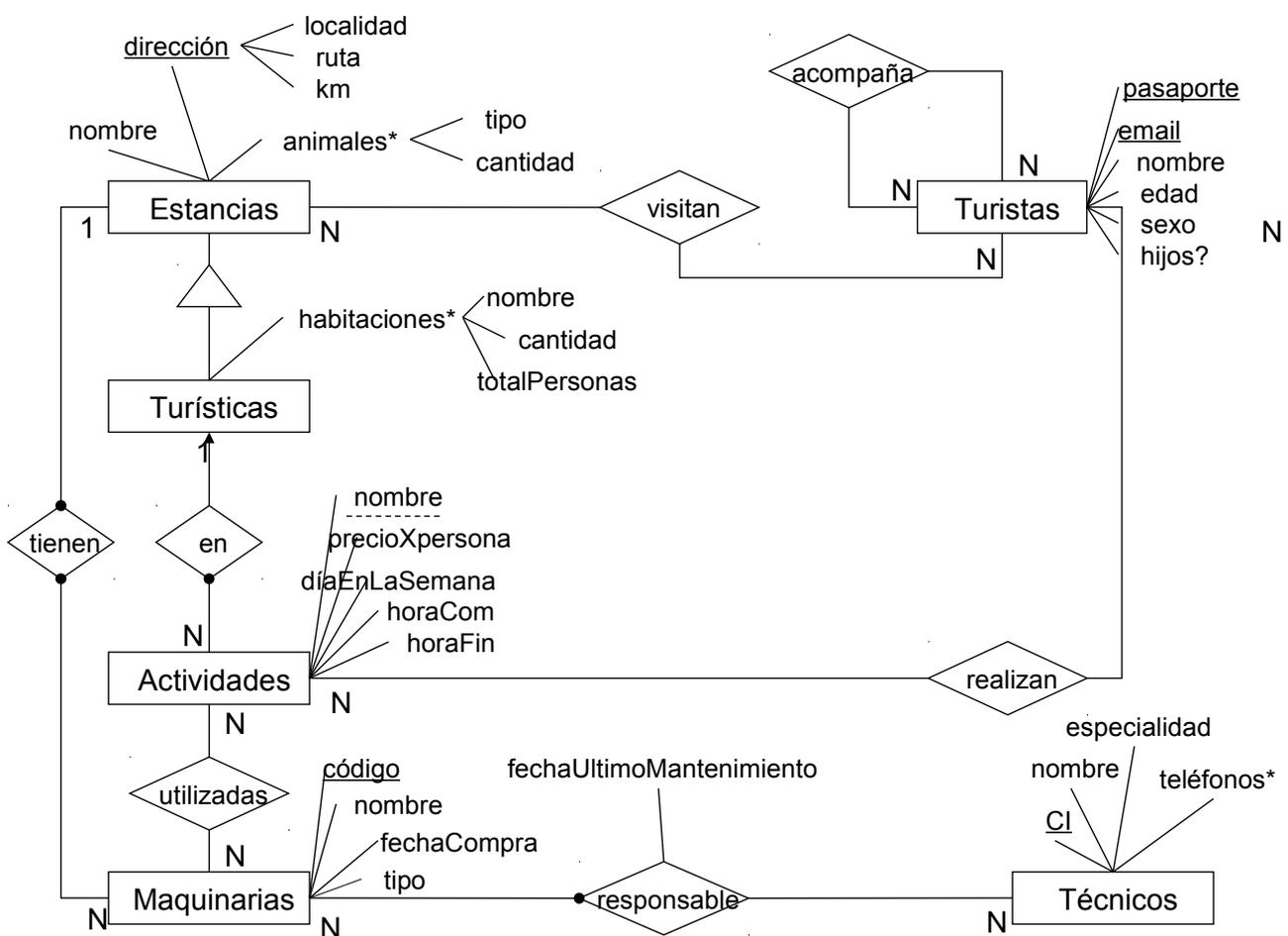
En las estancias turísticas se realizan actividades y el nombre de las mismas las identifica en cada estancia (paseo en tractor, caminatas, avistamientos de aves, cabalgatas, etc.). De cada actividad se conoce el precio por persona, el día de la semana en el cual se realiza, la hora de comienzo y de fin (por ejemplo, lunes de 16:00 a 18:00 hs.). Para el caso de la actividad "paseo en tractor" se utiliza maquinaria de la estancia, por lo que interesa saber qué actividades utilizan qué maquinaria. Las actividades pueden utilizar un único tipo de maquinaria. Es decir, la actividad "paseo en tractor" sólo utiliza tractores y, dependiendo de la cantidad de personas que realicen la actividad, podría utilizarse más de un tractor.

De las maquinarias se conoce el código identificador, el tipo y la fecha de compra. Todas las estancias tienen maquinarias, sean o no turísticas.

Por otro lado, se almacenan datos de diferentes técnicos, cédula de identidad, nombre, especialidad y todos sus teléfonos. A todas las máquinas se les realiza un mantenimiento, dado por algunos de los técnicos, por esta razón interesa registrar la fecha del último mantenimiento dado a una máquina por cierto técnico.

Las estancias realizan publicidad de las mismas vía e-mail, por esto, se registra información de todos los turistas que las visitan. El pasaporte o el e-mail identifican a un turista. También registran su nombre, edad, sexo y si tienen hijos. Para saber qué información enviarle a un turista, cuando éste visita una estancia turística se registra cuáles fueron las actividades que realizó, y para no enviar la misma información a cada integrante de una familia, interesa modelar los turistas que llegaron juntos.

Se pide: Modelo Entidad-Relación completo del problema.



RNEs:

- La relación “acompaña” es simétrica.
- La relación “acompaña” es irreflexiva.
- Las actividades utilizan maquinarias pertenecientes a las estancias turísticas en las cuales se realizan dichas actividades.
- Los turistas realizan actividades de las estancias turísticas que ellos visitan.
- Cada actividad utiliza un único tipo de maquinaria.

Ejercicio 2 (20 puntos)

El siguiente esquema representa parte del esquema relacional de una inmobiliaria determinada:

CASAS (cas-cod, cas-calle, cas-nro_puerta, ci-propietario)

Representa las casas ofrecidas por la inmobiliaria, su calle y número de puerta, y el propietario de la misma.

PERSONAS (ci, nombre, apellido, teléfono, mail)

Representa las personas que han celebrado contratos a través de la inmobiliaria (propietarios e inquilinos).

CONTRATOS (cas-cod, ci-inquilino, fecha, monto-total, cant-meses)

Representa los contratos correspondientes a una casa realizados por un inquilino en una determinada fecha, y el monto total del alquiler por la cantidad de meses acordada.

Sea la siguiente consulta sobre dicho esquema:

```
SELECT P.nombre, P.apellido, Cts.monto-total, Cs.cas-cod
FROM Contratos Cts, Personas P, Casas Cs
WHERE Cts.cas-cod = Cs.cas-cod AND Cts.ci-inquilino = P.ci AND
      Cts.monto-total > 110.000 AND Cs.ci-propietario='1.234.567-8';
```

Se pide:

- Enunciar las heurísticas vistas en el teórico para generar planes lógicos.
- Construir un plan lógico para la consulta utilizando las heurísticas. Justifique debidamente los pasos que se siguen.
- Calcular todos los tamaños intermedios, justificando.
- Construir un plan físico correspondiente al plan lógico de la parte anterior aprovechando la mayor cantidad de índices que sea posible. Justifique debidamente las decisiones tomadas.

Datos:

		CASAS	PERSONAS	CONTRATOS
Cantidad Tuplas		400	40	260
Índices primarios (niveles: 1)			- ci	- cas-cod, ci-inquilino, fecha
Índices secundarios	Hash	- ci-propietario, cas-cod		- monto-total
	B+		- mail	
Cantidad de tuplas por bloque		50	30	20

Observaciones:

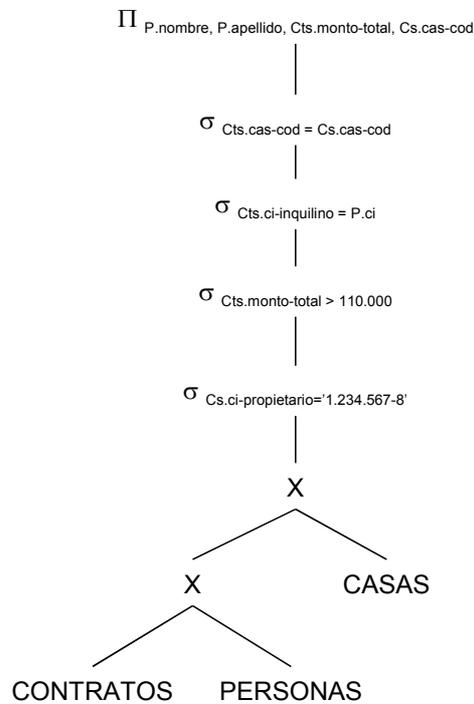
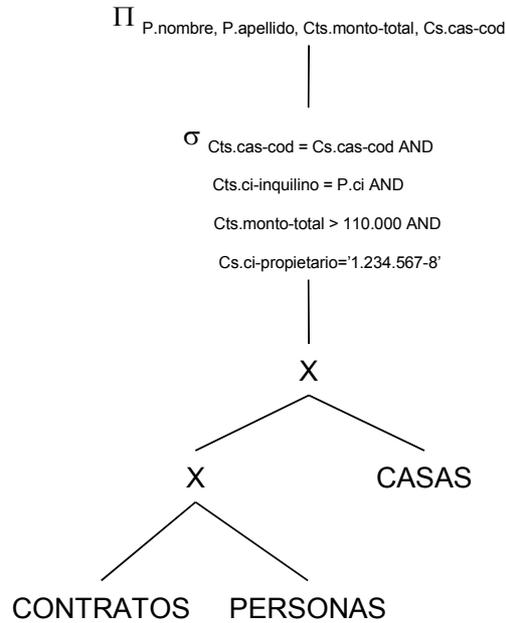
- el individuo con cédula '1.234.567-8' posee 7 casas en alquiler.
- un 25% de los contratos celebrados es de a lo sumo \$U 110.000.
- la máxima cantidad contratos realizados por un mismo inquilino es igual a 4.

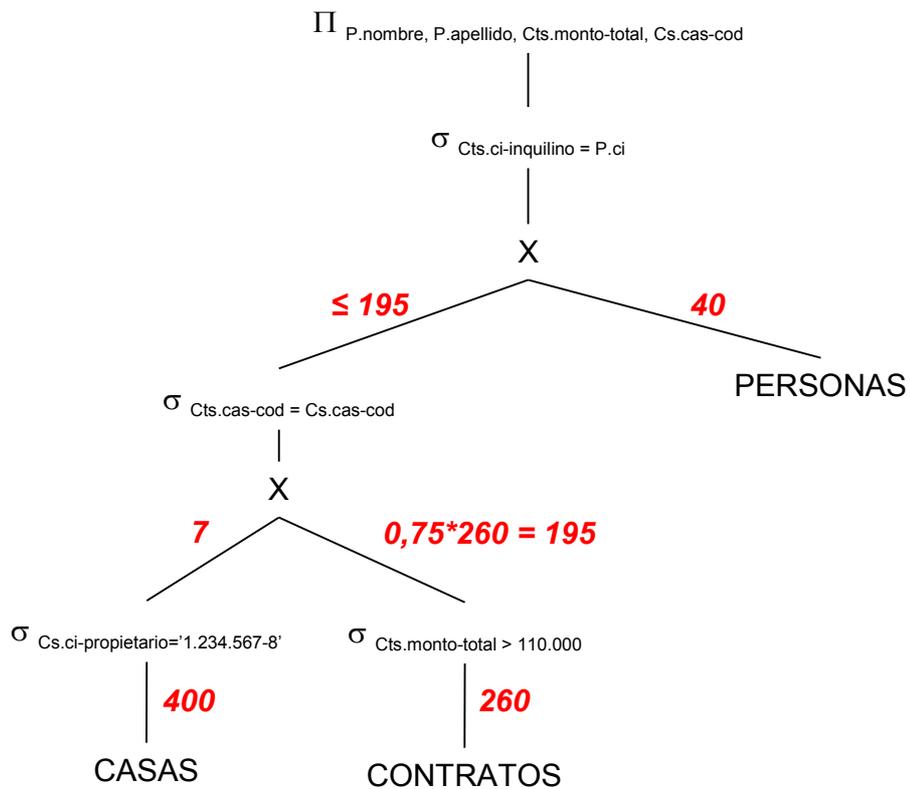
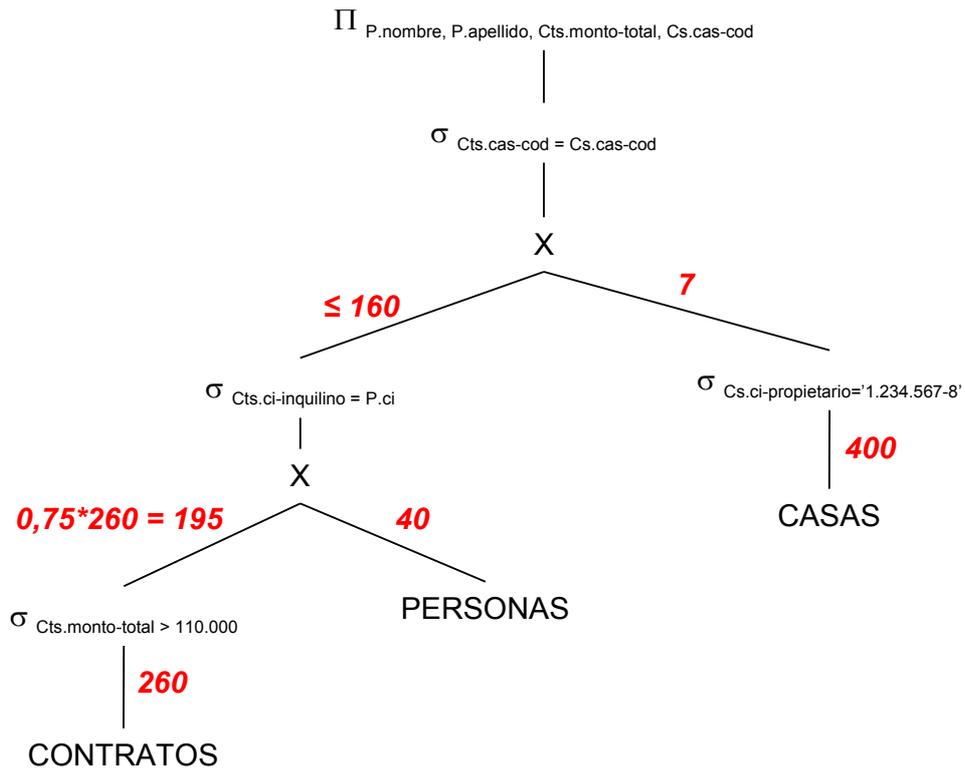
a)

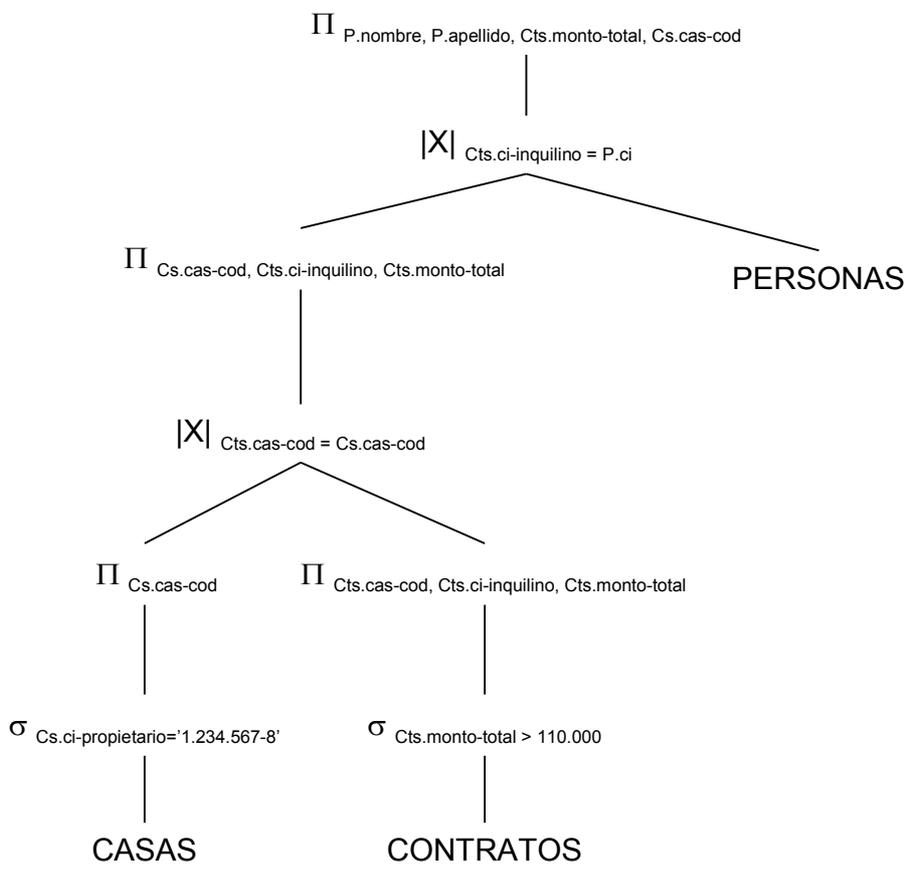
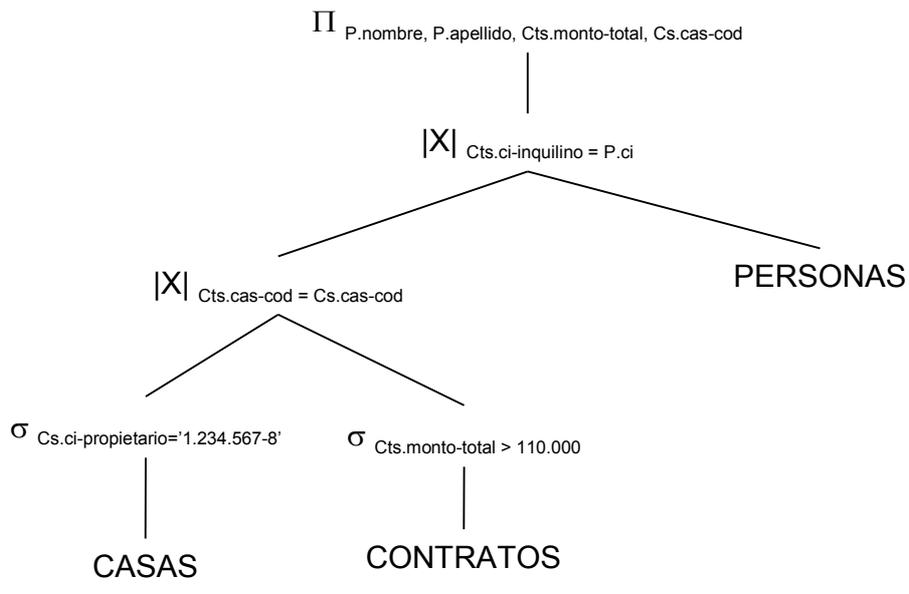
- Construir el árbol canónico.
- Cambiar las selecciones conjuntivas por una "cascada" de selecciones simples.

2. Mover las selecciones lo más abajo que se pueda en el árbol.
3. Poner a la izquierda de los productos las hojas que generen menos tuplas.
4. Cambiar secuencias de selecciones y productos por join's.
5. Mover las proyecciones lo más abajo posible en el árbol, agregando las proyecciones que sean necesarias.

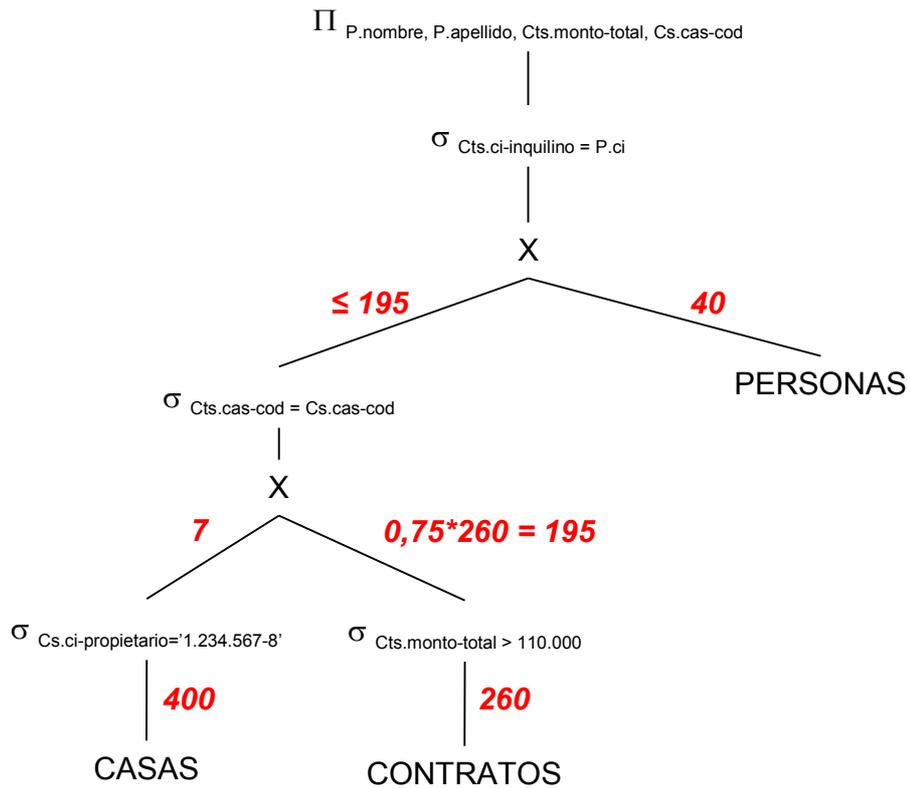
b)



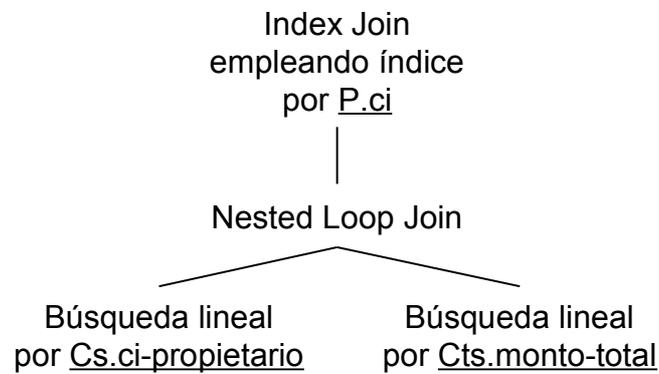




c)



d)



Ejercicio 3 (25 puntos)

Dada la siguiente realidad:

En el sector de maternidad de un sanatorio se lleva el registro de los nacimientos. Para esto se almacenan los siguientes datos. Sobre el recién nacido se registra el nombre, el peso al nacer, la fecha de nacimiento, la hora de nacimiento, la CI (cédula de identidad) de la madre. Sobre la madre, se registra su CI, su nombre, su dirección y su teléfono. Del médico se registra su CI, su nombre, su fecha de ingreso al sanatorio, su especialidad y su teléfono. De los enfermeros se registra su CI, su nombre, su fecha de ingreso al sanatorio y su teléfono. Cada médico puede tener más de una especialidad. El recién nacido se identifica con su nombre y la CI de la madre. En cada nacimiento participan varios médicos y varios enfermeros (es decir, que a cada recién nacido corresponden varios médicos y varios enfermeros). Cada médico atiende como máximo un parto por día, o sea, a una sola madre pero no a un solo niño, ya que podría ser un parto múltiple.

Utilizando los siguientes nombres para los atributos:

nom-rn	dir-madre	esp
peso	tel-madre	ci-enf
fnac	ci-med	nom-enf
hnac	nom-med	fing-enf
ci-madre	fing-med	tel-enf
nom-madre	tel-med	

Se pide:

a) Deducir todas las dependencias funcionales que se cumplen.

nom-rn, ci-madre \rightarrow fnac, hnac, peso

ci-madre \rightarrow nom-madre, dir-madre, tel-madre

ci-med \rightarrow nom-med, fing-med, tel-med

ci-med, fnac \rightarrow ci-madre

ci-enf \rightarrow nom-enf, fing-enf, tel-enf

b) Para la relación universal R (tabla que contiene todos los atributos), y el conjunto de dependencias funcionales hallado en la parte a), hallar todas las claves. Justificar.

R (nom-rn, peso, fnac, hnac, ci-madre, nom-madre, dir-madre, tel-madre, ci-med, nom-med, fing-med, tel-med, esp, ci-enf, nom-enf, fing-enf, tel-enf)

Atributos que no pertenecen a ninguna df: { esp }, deben estar en todas las claves.

Atributos que no se encuentran a la derecha de ninguna df: { ci-med, ci-enf, nom-rn }, deben estar en todas las claves.

Atributos que no se encuentran a la izquierda de ninguna df: { hnac, peso, nom-madre, dir-madre, tel-madre, nom-med, fing-med, tel-med, nom-enf, fing-enf, tel-enf }, no pueden estar en ninguna clave.

$(ci-med, ci-enf, nom-rn, esp)^+ = \{ ci-med, ci-enf, nom-rn, esp, nom-med, fing-med, tel-med, nom-enf, fing-enf, tel-enf \}$

Probamos agregando atributos:

$(ci-med, ci-enf, nom-rn, esp, ci-madre)^+ = R \Rightarrow$ es clave

$(ci-med, ci-enf, nom-rn, esp, fnac)^+ = R \Rightarrow$ es clave

Son las únicas claves porque el resto de los atributos no pueden estar en ninguna clave.

c) Dar una descomposición de R en 3NF sin pérdida de dependencias y con JSP. Aplicar el algoritmo dado en el curso (mostrando los pasos).

Dados R y

$F = \{ \text{nom-rn, ci-madre} \rightarrow \text{fnac, hnac, peso}$
 $\text{ci-madre} \rightarrow \text{nom-madre, dir-madre, tel-madre}$
 $\text{ci-med} \rightarrow \text{nom-med, fing-med, tel-med}$
 $\text{ci-med, fnac} \rightarrow \text{ci-madre}$
 $\text{ci-enf} \rightarrow \text{nom-enf, fing-enf, tel-enf} \}$

Paso 1) Hallamos F_{minimal} .

$F_{\text{minimal}} = \{ \text{nom-rn, ci-madre} \rightarrow \text{fnac}$
 $\text{nom-rn, ci-madre} \rightarrow \text{hnac}$
 $\text{nom-rn, ci-madre} \rightarrow \text{peso}$
 $\text{ci-madre} \rightarrow \text{nom-madre}$
 $\text{ci-madre} \rightarrow \text{dir-madre}$
 $\text{ci-madre} \rightarrow \text{tel-madre}$
 $\text{ci-med} \rightarrow \text{nom-med}$
 $\text{ci-med} \rightarrow \text{fing-med}$
 $\text{ci-med} \rightarrow \text{tel-med}$
 $\text{ci-med, fnac} \rightarrow \text{ci-madre}$
 $\text{ci-enf} \rightarrow \text{nom-enf}$
 $\text{ci-enf} \rightarrow \text{fing-enf}$
 $\text{ci-enf} \rightarrow \text{tel-enf} \}$

Se puede verificar fácilmente que ninguna de las dfs tiene atributos redundantes a la izquierda y ninguna de las dfs es redundante.

Paso 2) Generamos una relación para cada grupo de dfs con el mismo lado izquierdo

R1 (nom-rn, ci-madre, fnac, hnac, peso)
R2 (ci-madre, nom-madre, dir-madre, tel-madre)
R3 (ci-med, nom-med, fing-med, tel-med)
R4 (ci-med, fnac, ci-madre)
R5 (ci-enf, nom-enf, fing-enf, tel-enf)

Paso 3) Generamos una relación, cuyos atributos sean los de una de las claves, para asegurar la propiedad de JSP.

R6 (ci-madre, ci-med, ci-enf, nom-rn, esp)

Aclaración: Al hacer el paso 3, saltamos el paso de incluir una relación con los atributos que no pertenecen a ninguna df, ya que la misma quedaría incluida en la que agregamos después y se eliminaría.

Descomposición $\rho = \{ \mathbf{R1, R2, R3, R4, R5, R6} \}$

d) Decir si la descomposición anterior está en BCNF, justificando. Si no está en BCNF, llevarla a dicha forma normal, aplicando el algoritmo dado en el curso (mostrando los pasos).

Para que la descomposición r esté en BCNF, cada relación debe estar en BCNF:

R1 (nom-rn, ci-madre, fnac, hnac, peso)

$F_{R1} = \{ \text{nom-rn, ci-madre} \rightarrow \text{fnac, hnac, peso} \}$

=> clave única: (nom-rn, ci-madre)

R1 está en BCNF, ya que todas sus dfs cumplen que su lado izquierdo es superclave.

R2 (ci-madre, nom-madre, dir-madre, tel-madre)

$F_{R2} = \{ \text{ci-madre} \rightarrow \text{nom-madre, dir-madre, tel-madre} \}$

=> clave única: ci-madre

R2 está en BCNF, ya que todas sus dfs cumplen que su lado izquierdo es superclave.

R3 (ci-med, nom-med, fing-med, tel-med)

$F_{R3} = \{ \text{ci-med} \rightarrow \text{nom-med, fing-med, tel-med} \}$

=> clave única: ci-med

R3 está en BCNF, ya que todas sus dfs cumplen que su lado izquierdo es superclave.

R4 (ci-med, fnac, ci-madre)

$F_{R4} = \{ \text{ci-med, fnac} \rightarrow \text{ci-madre} \}$

=> clave única: (ci-med, fnac)

R4 está en BCNF, ya que todas sus dfs cumplen que su lado izquierdo es superclave.

R5 (ci-enf, nom-enf, fing-enf, tel-enf)

$F_5 = \{ \text{ci-enf} \rightarrow \text{nom-enf, fing-enf, tel-enf} \}$

=> clave única: ci-enf

R5 está en BCNF, ya que todas sus dfs cumplen que su lado izquierdo es superclave.

R6 (ci-madre, ci-med, ci-enf, nom-rn, esp)

$F_6 = \{ \}$

=> clave única: (ci-madre, ci-med, ci-enf, nom-rn, esp)

R6 está en BCNF, ya que no tiene dfs por lo tanto ninguna viola la definición de BCNF.

Finalmente, **p está en BCNF.**

e) Deducir todas las dependencias multivaluadas (no triviales) que se cumplen en cada tabla de la última descomposición obtenida. Para cada multivaluada que proponga, mostrar qué parte de la letra lo induce a pensar que ésta se cumple.

Existen dependencias multivaluadas solamente en R6 (ci-madre, ci-med, ci-enf, nom-rn, esp), y son el siguiente conjunto:

$D = \{ \text{ci-med} \twoheadrightarrow \text{esp}$
 $\text{nom-rn, ci-madre} \twoheadrightarrow \text{ci-med, esp} \}$

ci-med \twoheadrightarrow esp, por "Cada médico puede tener más de una especialidad"

nom-rn, ci-madre \twoheadrightarrow ci-med, esp

por "En cada nacimiento participan varios médicos y varios enfermeros", y considerando que el médico y la especialidad tienen que quedar juntos porque no son independientes entre sí.

f) Llevar la última descomposición obtenida a 4NF, aplicando el algoritmo dado en el curso (mostrando los pasos).

R1, R2, R3, R4 y R5, ya se encuentran en 4NF, porque no tienen dependencias multivaluadas no triviales.

Debemos llevar R6 a 4NF, aplicando el algoritmo visto en el curso:

R6 (ci-madre, ci-med, ci-enf, nom-rn, esp)

$D_{R6} = \{ \text{ci-med} \twoheadrightarrow \text{esp}$
 $\text{nom-rn, ci-madre} \twoheadrightarrow \text{ci-med, esp} \}$

clave: (ci-madre, ci-med, ci-enf, nom-rn, esp)

ci-med \twoheadrightarrow esp viola 4NF porque el lado izquierdo no es superclave.

Tomamos la mvd ci-med \twoheadrightarrow esp, que viola 4NF, y generamos:

R61 (ci-med, esp), $D_{R61} = \{ \}$, está en 4NF

R62 (ci-madre, ci-med, ci-enf, nom-rn), $D_{R62} = \{ \text{nom-rn, ci-madre} \twoheadrightarrow \text{ci-med}$
 $\text{nom-rn, ci-madre} \twoheadrightarrow \text{ci-enf} \}$

clave: (ci-madre, ci-med, ci-enf, nom-rn)

La mvd nom-rn, ci-madre \twoheadrightarrow ci-med viola 4NF porque el lado izquierdo no es superclave. R62 no está en 4NF.

Tomamos la mvd nom-rn, ci-madre \twoheadrightarrow ci-med, que viola 4NF, y generamos:

R621 (nom-rn, ci-madre, ci-med), $D_{R621} = \{ \}$, está en 4NF

R622 (nom-rn, ci-madre, ci-enf), $D_{R622} = \{ \}$, está en 4NF

Descomposición final, en 4NF:

{R1, R2, R3, R4, R5, R61, R621, R622}

Ejercicio 4 (30 puntos)

El siguiente esquema de base de datos representa parte de la información de las ventas de una determinada automotora:

AUTOMOVILES (nroChasis, matricula, marca, modelo, año, disponible?, fechaIngreso)

Contiene la información de los automóviles manejados por la automotora, el número de chasis que lo identifica, su matrícula, marca, modelo, año, un atributo que indica si el automóvil puede ser vendido y la fecha de ingreso del auto a la automotora.

CLIENTES (ciCliente, nombre, direccion, telefono)

Contiene, para cada cliente registrado en la automotora, su cédula de identidad, nombre completo, dirección y teléfono.

VENEDORES (ciVendedor, nombre, direccion, telefono)

Contiene, para cada vendedor la automotora, su cédula de identidad, nombre completo, dirección y teléfono.

VENTAS (nroChasis, fecha, ciCliente, ciVendedor, precio)

Contiene la información de las ventas realizadas por la automotora. De cada venta se registra el número de chasis del automóvil vendido, el cliente comprador, el vendedor, la fecha y el precio en dólares por el cual se realizó la venta.

SE_INTERESA (nroChasis, ciCliente, ciVendedor)

Contiene la información de aquellos clientes que demostraron interés por algún automóvil y el vendedor que los atendió.

En este esquema no existen tablas vacías y se cumplen las siguientes dependencias de inclusión:

$$\begin{aligned}\prod_{ciCliente}(VENTAS) &\subseteq \prod_{ciCliente}(CLIENTES) \\ \prod_{ciVendedor}(VENTAS) &\subseteq \prod_{ciVendedor}(VENEDORES) \\ \prod_{nroChasis}(VENTAS) &\subseteq \prod_{nroChasis}(AUTOMOVILES) \\ \prod_{ciCliente}(SE_INTERESA) &\subseteq \prod_{ciCliente}(CLIENTES) \\ \prod_{nroChasis}(SE_INTERESA) &\subseteq \prod_{nroChasis}(AUTOMOVILES) \\ \prod_{ciVendedor}(SE_INTERESA) &\subseteq \prod_{ciVendedor}(VENEDORES)\end{aligned}$$

Resolver las siguientes consultas en Álgebra Relacional:

1. Cédula de identidad y nombre de los clientes que se interesaron por todos los automóviles del año 2008 disponibles, pero no compraron ningún automóvil.

$$\begin{aligned}\text{Autos disponibles del 2008 } A &= (\prod_{nroChasis} (\sigma_{\text{año}='2008' \wedge \text{disponible?}='SI'} AUTOMOVILES)) \\ \text{Clientes interesados por todos los de } A & B = \prod_{ciCliente, nroChasis} (SE_INTERESA) \% A \\ R &= \prod_{ciCliente, nombre} (CLIENTES * (B - \prod_{ciCliente}(VENTAS)))\end{aligned}$$

2. Marca y modelo de los autos vendidos a mayor precio.

$$\begin{aligned}\text{Mayor precio de venta } A &= (\prod_{precio} VENTAS \mid_{\$1 > \$2} \mid \prod_{precio} VENTAS) \% (\prod_{precio} VENTAS) \\ R &= \prod_{marca, modelo} (A \mid_{\$1 = \$6} \mid VENTAS)\end{aligned}$$

Resolver las siguientes consultas en Cálculo Relacional:

3. Nombre de los vendedores que mostraron a clientes interesados todos los automóviles marca CHEVROLET y que lograron vender al menos uno de éstos.

$$\begin{aligned}\{v.nombre / VENEDORES (v) \wedge \\ (\forall a1)(AUTOMOVILES (a1) \wedge a1.marca = \text{"Chevrolet"} \rightarrow \\ (\exists i)(SE_INTERESA (i) \wedge a1.nroChasis = i.nroChasis \wedge v.ciVendedor = i.ciVendedor) \wedge \\ (\exists v1)(VENTAS (v1) \wedge v1.ciVendedor = v.ciVendedor \wedge \\ (\exists a3)(AUTOMOVILES (a3) \wedge \\ a3.nroChasis = v1.nroChasis \wedge a3.marca = a1.marca))\}\end{aligned}$$

4. Cédula de cliente y marca tal que dicho cliente sólo ha comprado automóviles de esa marca pero no demostró interés por ninguno ellos.

$$\{c.ciCliente, s.marca / CLIENTES(c) \wedge AUTOMOVILES(s) \wedge$$
$$(\forall v)(VENTAS(v) \wedge v.ciCliente = c.ciCliente \rightarrow$$
$$(\exists a)(AUTOMOVILES(a) \wedge a.nroChasis = v.nroChasis \wedge a.marca = s.marca \wedge$$
$$\neg(\exists i)(SE_INTERESA(i) \wedge i.ciCliente = c.ciCliente \wedge i.nroChasis = v.nroChasis))) \wedge$$
$$(\exists v1)(VENTAS(v1) \wedge v1.ciCliente = c.ciCliente)\}$$

Resolver las siguientes consultas en SQL, sin utilizar vistas, ni subconsultas dentro del FROM:

5. Cédula de identidad y premio a cobrar (5% sobre el total de las ventas realizadas por ese vendedor) para aquellos vendedores que vendieron más de 10 autos en diciembre de 2012. Para calcular el premio sólo deben considerarse las ventas realizadas por ese vendedor en diciembre de 2012.

```
SELECT v.ciVendedor, 0,05*sum(v.precio) as premio
FROM ventas v
WHERE MONTH(v.fecha) = 'diciembre' AND YEAR(v.fecha)='2012'
GROUP BY v.ciVendedor
HAVING count(*)>10
```

6. Número de chasis de automóviles del año 2005 por los cuales han demostrado interés al menos 5 clientes pero que no han sido vendidos.

```
SELECT a.nroChasis
FROM AUTOMOVILES a, SE_INTERESA si
WHERE a.anio = "2005" AND
      a.nroChasis = si.nroChasis AND
      NOT EXISTS (SELECT * FROM VENTAS v WHERE v.nroChasis = a.nroChasis)
GROUP BY a.nroChasis
HAVING COUNT(*) >= 5;
```