**Tercer trabajo de bases de datos 2 (20%)**

**Primer punto.**

Una metodología interesante para la **formalización** del proceso de generación de esquemas de documentos a partir de una base de datos relacional se presenta en “*Creating Collections with Embedded Documents for Document Databases Taking into Account the Queries*” (Ver artículo adjunto en el comprimido). Lea todo el artículo. Note que dicho artículo tiene algunos comentarios o correcciones menores. Por favor leerlos también.

Note que en este trabajo se evalúa no solo la capacidad de hacer un programa, sino la comprensión y aplicación de una metodología (la del artículo).

Luego, proceda a resolver los siguientes puntos.

1. **(10%)** Realice un formulario en Java para implementar la regla 1 (*Rule 1*) de la página 8, Ecuación (1).

De acuerdo con dicha metodología, defina los parámetros necesarios y su formato de entrada para lograr obtener **Q’**. Por simplicidad, suponga que **Q** contiene **todos** los atributos de las dos tablas.

Ejemplo

Parámetros de entrada:

- Solicite los nombres de las **dos** tablas: Por ejemplo, PAÍS y DEPTO; donde PAÍS 1-M DEPTO.

- Solicite los atributos de cada tabla: Por ejemplo,

para PAÍS: *idp* y *nomp*

para DEPTO: *codd*, *ndep*, *idp* y *capital*

Por supuesto cada tabla puede tener 1 o muchos atributos.

- Solicite los datos para S2, por ejemplo, para S2 supongamos que solo se va a considerar (para este ejemplo en particular) la consulta “SELECT nomp, codd, ndep FROM PAIS NATURAL JOIN DEPTO”.

- Solicite los datos para S1, por ejemplo, para S1 en este ejemplo supongamos que solo se va a considerar la consulta “SELECT codd, capital FROM DEPTO”.

Por supuesto, tanto para S1 como para s2 habrá usualmente más de una consulta...(S1 y S2 conformes a como se definen en la metodología).

De acuerdo con lo anterior, para este caso **particular**, el formulario luego del ingreso de los datos (parámetros) debe generar (imprimir) el siguiente esquema **Q’**:

Q’= {idp, nomp, DEPTO\_of\_PAÍS: {codd, ndep, capital}}

**Nota**: Solamente implemente la regla 1 para dos tablas (de acuerdo a como la metodología lo pide), **no** se pide generalizarla para *n* tablas, pero la regla debe quedar implementada en su totalidad.

Analice (según la metodología) si su formulario requiere parámetros adicionales aparte de los mostrados en este ejemplo.

1. **(15%)** Haga lo mismo e implemente la regla 4 (*Rule 4*) de la página 10, Ecuación (2).

**Nota**: Solamente implemente la regla 4 para tres tablas (de acuerdo a como la metodología lo pide), **no** se pide generalizarla para *n* tablas, pero la regla debe quedar implementada en su totalidad. Por simplicidad, suponga que **Q** contiene **todos** los atributos de las tres tablas.

1. **(15%)** Haga lo mismo e implemente la regla 5 (*Rule 5*) de la página 11, Ecuación (3).

**Nota**: Solamente implemente la regla 5 para tres tablas (de acuerdo a como la metodología lo pide), **no** se pide generalizarla para *n* tablas, pero la regla debe quedar implementada en su totalidad. Por simplicidad, suponga que **Q** contiene **todos** los atributos de las tres tablas.

**Nota**: Los programas (formularios) deben funcionar para **todos** los casos que la regla correspondiente **atiende**, **no solo** para ejemplos **particulares**.

1. **(10%)** Luego de leer todo el artículo, ¿cuáles extensiones o mejoras (**aparte** de las propuestas por los autores en la Sección 8, *Further research* y de los trabajos descritos en la Sección 2, *Related works*) se le ocurren? Proponga **al menos dos** y explique claramente porque sus propuestas (las de usted) no son consideradas actualmente por dicha metodología. No implemente sus propuestas, solo descríbalas y ejemplifiquelas.

e) **(10%)** En la Nota 2 de la página 8, ver líneas resaltadas en azul en el artículo los autores afirman lo siguiente: “*If the relational database schema initially satisfied the normal form of BCNF, 3NF, or 4NF, then there can be no other relationships between tables T1 and T2.* ”

¿es correcta esta afirmación?

Si lo es explique porque lo es;

Si no lo es presente un contraejemplo (¿quizás los autores pasaron por alto algo?).

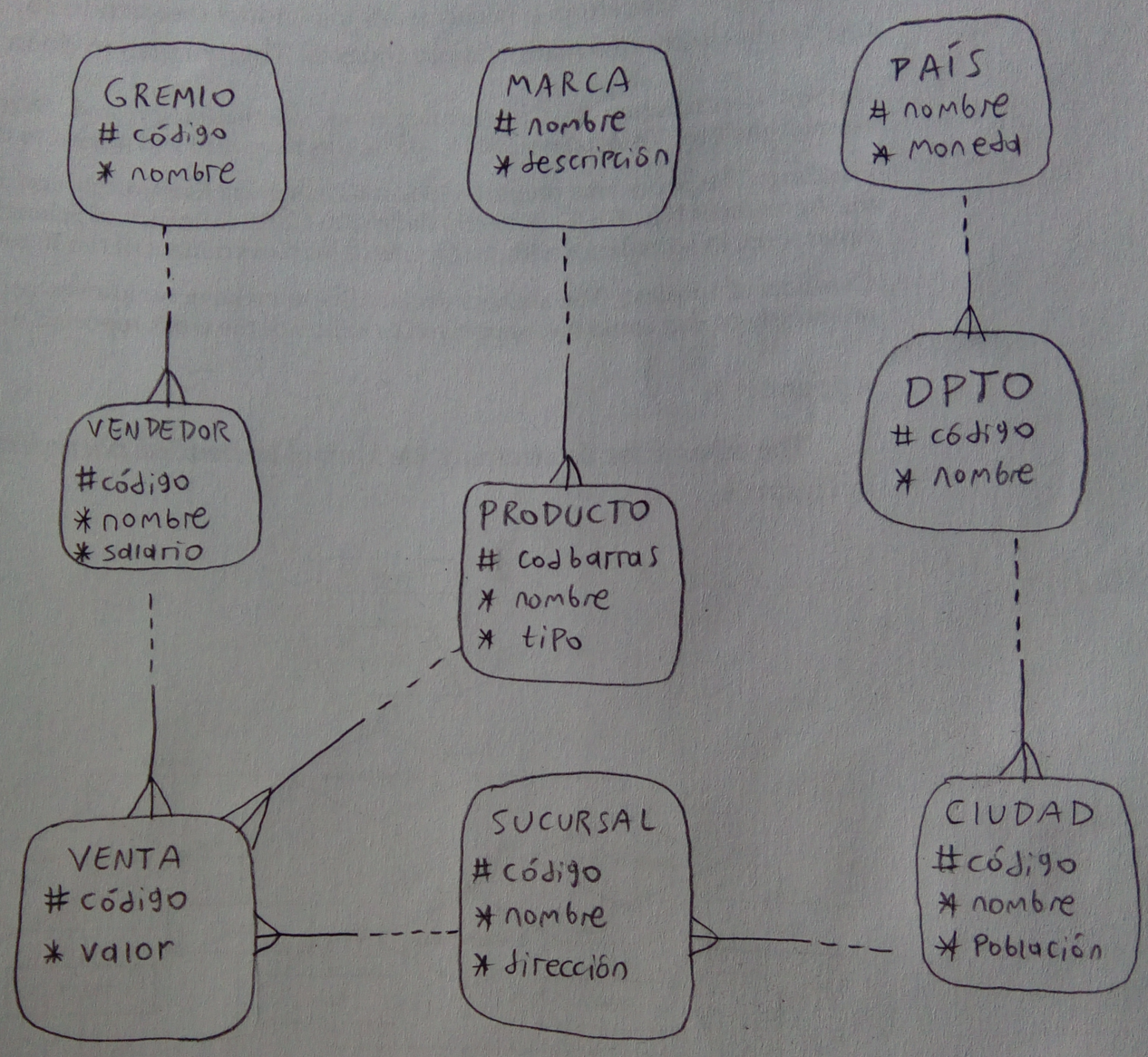
Para responder esta pregunta, implica que usted repase o estudie y tenga claros los conceptos de las formas normales mencionadas.

**Segundo punto**

Interacción MongoDB - Oracle - Java.

En este punto se simulará una pequeña base de datos operativa (*write-heav*y) en Oracle y una base de datos para análisis de estadísticos (*red-heavy*) en MongoDB.

Considere el siguiente modelo entidad-relación (E-R). Por simplicidad, se han omitido las leyendas.



Cree en Oracle las tablas (modelo relacional, con claves primarias, foráneas, etc.) correspondientes a este modelo E-R. Para las tablas **use los nombres** de las entidades y de los atributos del modelo E-R. Por supuesto, puede dar el nombre que desee a las claves foráneas. Use números de mínimo 8 dígitos y cadenas de caracteres de mínimo 20 caracteres para declarar sus atributos.

Las tablas serán cargadas con datos por medio de SQL\*Plus, NO por medio de Java.

Haga un formulario en Java que tiene los siguientes botones:

**- (20%) Botón Generar y Cargar Estadísticas**: Al presionar este botón se conecta a la base de datos Oracle y se extraen los siguientes estadísticos:

* Total de ventas (suma del atributo valor) por sucursal.
* Total de ventas (suma del atributo valor) por ciudad.
* Total de ventas (suma del atributo valor) por departamento.
* Total de ventas (suma del atributo valor) por país.
* Total de ventas (suma del atributo valor) por vendedor.
* Total de ventas (suma del atributo valor) por gremio.
* Total de ventas (suma del atributo valor) por producto.
* Total de ventas (suma del atributo valor) por marca.

Los anteriores datos insértelos en una base de datos MongoDB (para ello debe investigar cómo conectar Java con MongoDB). **Usted define la estructura de su base de datos MongoDB, lo importante es que los datos anteriores queden insertados allí.**

- **(20%) Botón Visualizar Estadísticas**: Al presionar este botón se conecta a la base de datos MongoDB y se deben mostrar en el formulario los siguientes datos.

* Las 3 sucursales (código y nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada una de ellas).
* Las 3 ciudades (código y nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada una de ellas).
* Los 3 departamentos (código y nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada uno de ellos).
* Los 3 países (nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada uno de ellos).
* Los 3 vendedores (código y nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada uno de ellos).
* Los 3 gremios (código y nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada uno de ellos).
* Los 3 productos (codbarras y nombre) con mayor total de ventas (se debe mostrar el total de ventas de cada uno de ellos).
* Las 3 marcas (nombre y descripción) con mayor total de ventas (se debe mostrar el total de ventas de cada una de ellas).

Usted define el formato de mostrado, lo importante es que se muestren los datos solicitados en el formulario.

**Notas adicionales:**

* Sus soluciones deben funcionar para cualquier cantidad de filas que tengan las tablas.
* Puede usar todas las estructuras de datos y todas las tablas auxiliares que desee.
* Para entregar por email a *[fjmoreno@unal.edu.co](mailto:fjmoreno@unal.edu.co)*, el viernes 23 de octubre **hasta las 6 pm**. Solo se califican trabajos enviados a ese correo.
* **No se reciben trabajos en hora posterior**. No se reciben versiones “mejoradas”. No se califican trabajos enviados “por accidente” a otros correos.
* **Se debe incluir un informe donde se describa cómo se solucionó cada punto.** Este informe hace parte de la calificación del trabajo (máximo 4 hojas). **No enviar los datos de prueba que usted usó para probar sus códigos.**
* Grupos de **tres** personas.
* Los trabajos deben ser independientes entre los grupos. Trabajos copiados **así sea en un SOLO punto** se califican con 0 (cero) en su totalidad para todos los integrantes. Las soluciones presentadas deben ser originales. El trabajo debe ser desarrollado por los integrantes del grupo no por personas ajenas a él.
* El monitor les puede ayudar con aspectos técnicos pero su función **no** es hacerles la práctica **ni está autorizado** para **cambiar las condiciones del trabajo**.
* **Si trabaja con otras herramientas**, así su trabajo funcione y sea “espectacular”, el trabajo **NO** será calificado.
* **No** se califica la “belleza” de los formularios, se califica la funcionalidad.
* Cualquier duda consultarla con el profesor.