

Base de Datos (75.15 / 75.28 / 95.05)

Evaluación Integradora - 18 de diciembre de 2019

TEMA 20192C2						Padrón: _____ Apellido: _____ Nombre: _____ Cantidad de hojas: _____ <input type="checkbox"/> Aprobado <input type="checkbox"/> Insuficiente
SQL		Proc.		DR		
CyT		Seg.		NoSQL		
Corrigió: Nota:						

Criterio de aprobación: El examen está compuesto por 6 ítems, cada uno de los cuales se corrige como B/B-/Reg/Reg-/M. Se aprueba con nota mayor o igual a 4(cuatro), equivalente a desarrollar el 60 % del examen correctamente.

1. (*SQL*) Una encuestadora mide desde hace muchos años cuáles son los programas de televisión favoritos en la Argentina. Para ello cuenta con una base de datos de espectadores, a los que mensualmente contacta por teléfono para realizarles la siguiente pregunta: ¿Cuál es su programa favorito en la televisión abierta en Argentina?

Las siguientes tablas contienen información sobre los espectadores encuestados, los programas registrados y el programa votado en cada mes por cada espectador que respondió la pregunta.

- Espectadores(cod_espectador, localidad, año_nacimiento)
- Programas(nombre_programa, fecha_creación)
- ProgramasFavoritos(mes, año, cod_espectador, nombre_programa)

Nos interesa concentrarnos en la tabla **ProgramasFavoritos**, con el objetivo de calcular la moda para cada mes (es decir, cuál fue el programa más votado para ese mes). Escriba una consulta SQL que encuentre el programa más votado en cada mes, devolviendo (**mes**, **año**, **programa_más_votado**).

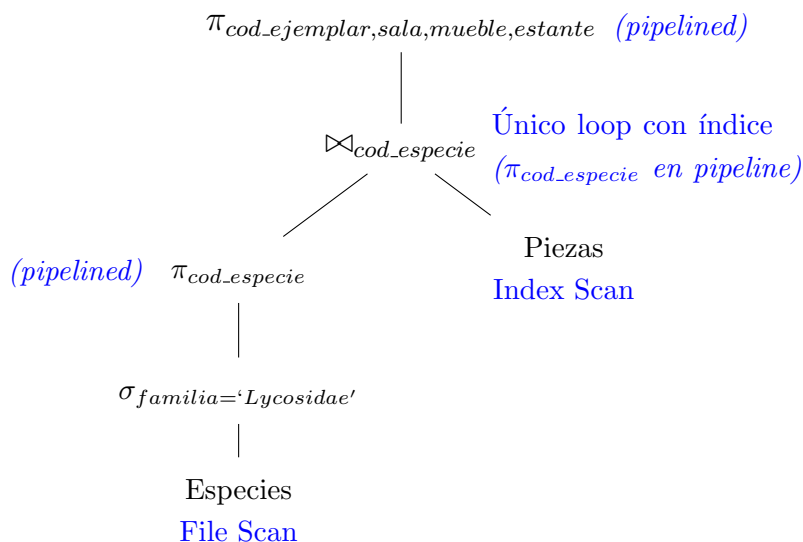
Nota: Considere que es posible que en algún mes haya más de un programa que sea el más votado.

2. (*Procesamiento de consultas*) El Museo de Ciencias Naturales posee un catálogo de 1 millón de piezas de distintas especies. Cada pieza corresponde a una especie en particular, que a su vez pertenece a una determinada familia de la taxonomía. Esta información se encuentra almacenada en las siguientes dos tablas:

- `Especies(cod_especie, nombre_especie, familia)`
- `Piezas(cod_ejemplar, _cod_especie, sala, mueble, estante)`

Adicionalmente, se dispone de un índice de *clustering* por el atributo `cod_especie` para acceder a la tabla `Piezas`.

El director del museo está interesado en identificar todas las piezas que el museo posee correspondientes a la familia *Lycosidae* (conocida comúnmente como ‘tarántula’). Para obtenerlas se arma el siguiente plan de ejecución:



Se pide:

- Estime el costo del plan de ejecución propuesto. Indique cuál es el mínimo de memoria que requerirá para poder llevar adelante el plan con el costo que indicó.
- Estime la cardinalidad del resultado en términos de cantidad de tuplas.

Considere la siguiente información de catálogo:

ESPECIES	PIEZAS
$n(\text{Especies}) = 40.000$	$n(\text{Piezas}) = 1.000.000$
$B(\text{Especies}) = 4.000$	$B(\text{Piezas}) = 200.000$
$V(\text{familia}, \text{Especies}) = 100$	$V(\text{cod_especie}, \text{Piezas}) = 40.000$
	$H(I(\text{cod_especie}, \text{Piezas})) = 2$

Considere que los atributos `cod_especie` y `cod_ejemplar` ocupan 8 bytes, mientras que los restantes atributos ocupan 16 bytes.

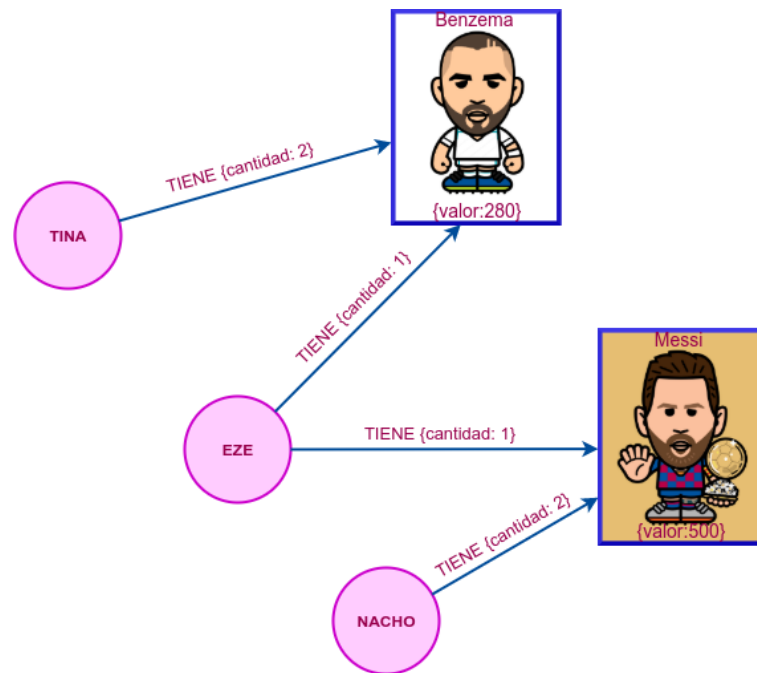
3. (*Diseño relacional*) Dada la relación $R(A, B, C)$ con el siguiente conjunto de dependencias funcionales $F = \{A \rightarrow B, B \rightarrow C\}$, indique si las siguientes afirmaciones son verdaderas o falsas, justificando su respuesta.
- La relación R se encuentra en $2FN$ (Segunda Forma Normal).
 - La descomposición $\rho_1 = \{R_1(A, B), R_2(B, C)\}$ de R preserva todas sus dependencias funcionales.
 - La descomposición $\rho_2 = \{R_1(A, B), R_2(A, C)\}$ de R preserva todas sus dependencias funcionales.
 - La descomposición $\rho_2 = \{R_1(A, B), R_2(A, C)\}$ preserva la información de R .
 - La descomposición $\rho_3 = \{R_1(A, C), R_2(B, C)\}$ preserva la información de R .
4. (*Concurrencia y transacciones*) Explique el funcionamiento básico del método de control de concurrencia conocido como *Snapshot Isolation*, mencionando sus ventajas y desventajas respecto al método más pesimista basado en *locks*. Indique en particular si bajo dicho esquema es posible que sea necesario abortar una transacción, y si es posible garantizar un nivel de aislamiento *Serializable*.
5. (*Seguridad*) Explique en qué consiste el principio del *menor privilegio posible*, e indique de qué manera el método *RBAC* de control de acceso puede ayudar a implementarlo.
6. (*NoSQL*) Tina está llenando su Álbum de la *Champions League* y está ansiosa por conseguir la figurita más difícil, correspondiente a Lionel Messi. Dado que en su escuela los alumnos disponen de una base en *Neo4j* para registrar todas las figuritas que cada uno tiene, Tina va a aprovecharla para encontrar a alguien a quien le sobre una figurita de Messi y quiera intercambiarla con ella. En dicha base se define a las personas y a las figuritas como nodos del grafo, mientras que las aristas indican qué personas poseen qué figuritas, registrando también la cantidad de unidades que poseen de la misma (por ejemplo, si una persona posee dos ejemplares, esto quiere decir que un ejemplar está pegado en su álbum y el otro se encuentra en su pilón de figuritas sobrantes). A continuación se ilustra la estructura de los nodos y aristas del grafo:

```

1 CREATE (a:Persona {nombre: 'Eze'})
2 CREATE (f:Figurita {nombre_jugador: 'Lionel Messi', valor: 500})

1 MATCH (eze: Persona {nombre: 'Eze'}),
2       (messi: Figurita {nombre_jugador: 'Lionel Messi'})
3 CREATE (eze)-[:TIENE {cantidad: 1}]- (messi)

```



Con el fin de encontrar amigos que tienen la figurita de Messi pero no tienen alguna figurita que a ella le sobra, escribió la siguiente consulta en *Neo4j*:

```

1 MATCH (a: Persona)-[t1:TIENE]->(m:Figurita),
2     (yo: Persona {nombre: 'Tina'})-[t2:TIENE]->(sobra:Figurita)
3 WHERE m.nombre_jugador='Lionel Messi'
4 AND t1.cantidad>1
5 AND t2.cantidad>1
6 AND NOT (a)-[:TIENE]->(sobra)
7 RETURN a.nombre, sobra.nombre_jugador

```

El problema es que la figurita de Messi es la que mayor valor tiene en la colección, y todos los amigos a los que les preguntó no quisieron cambiarla por ninguna figurita de menor valor. Por lo tanto, Tina decidió hacer una oferta más razonable contemplando el valor de las figuritas, a partir de las siguientes dos ideas:

- Ofrecer dos figuritas a cambio de la de Messi, en vez de una.
- Intentar que el valor conjunto de las dos figuritas que ofrece sea lo más pequeño posible, pero garantizando que sea al menos superior al valor de la figurita de Messi.

Escriba una consulta en *Cypher* que devuelva el listado de los 10 ofrecimientos de figuritas más convenientes que podría hacer, indicando el nombre del amigo, el nombre de los dos jugadores cuyas figuritas entregaría y la suma de los valores de dichas figuritas, mostrando primero a los pares de figuritas de menor valor conjunto.