

Base de Datos (75.15 / 75.28 / 95.05)

Evaluación Integradora - 3 de agosto de 2022

TEMA 20221C4						Padrón: _____
CRT		Proc.		DR		Apellido: _____
SQL		NoSQL		Seg.		Nombre: _____
Nota:						Cantidad de hojas: _____
						<input type="checkbox"/> Aprobado <input type="checkbox"/> Insuficiente

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. (*Cálculo Relacional de Tuplas*) La AASA (*Asociación Argentina de Salto en Alto*) está contactando a deportistas de salto en alto de todo el país para organizar un torneo nacional. La idea de la Asociación es que el torneo sea lo más federal y representativo posible, por eso se buscará que haya al menos un/a saltador/a de cada departamento del país, y que sea el mejor de su departamento.

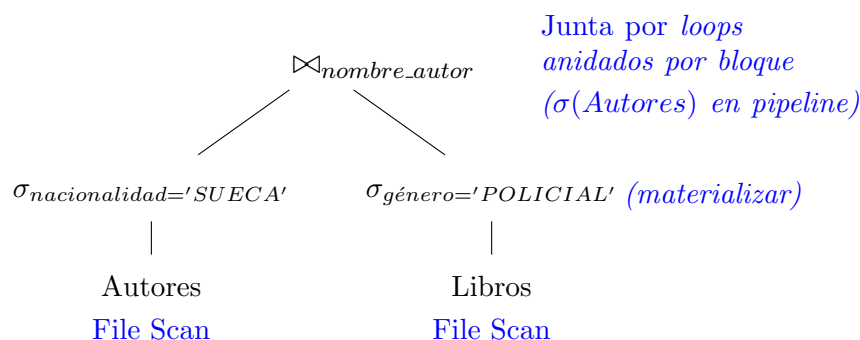
A partir de los siguientes datos sobre los departamentos del país y sus saltadores, escriba una consulta en Cálculo Relacional de Tuplas que encuentre a aquellos/as saltadores/as que tengan la marca más alta de su departamento, devolviendo el nombre del saltador y el nombre del departamento. Si varios saltadores empatan dentro de su departamento, la Asociación quiere invitarlos a todos.

- Departamentos(nombre_dpto, población, superficie)
- Saltadores(nombre_saltador, nombre_dpto, marca)

2. (*Procesamiento de Consultas*) Cándido trabaja en una biblioteca, y utiliza frecuentemente una base de datos relacional que él mismo creó para poder asistir a los lectores. Esta base contiene, entre otras, las siguientes dos tablas:

- Autores(nombre_autor, nacionalidad, fecha_nacimiento)
- Libros(ISBN_libro, nro_ejemplar, género, estante, nombre_autor)

Federica es una ávida lectora que acaba de consultarle a Cándido sobre alguna novela policial de origen sueco, y Cándido escribió una consulta sobre su base de datos que generó el siguiente plan de ejecución:



Dado que no se dispone de índices, el plan construido utiliza *file scans* como método de acceso, y debe materializar el resultado de la selección sobre **Libros** antes de la junta.

Considerando que la computadora de Cándido sólo dispone de $M=30$ bloques de memoria, y utilizando la información de catálogo que se brinda más abajo, se pide:

- a) Estime el costo del plan de ejecución generado, en términos de cantidad de bloques.
- b) Explique por qué la siguiente estimación $n(\bowtie) = \frac{500 \cdot 20000}{\max(500, 20000)} = 500$ para la cardinalidad del resultado no es correcta.
- c) Si tuviera que crear un único índice a efectos de mejorar significativamente el desempeño de esta consulta, indique qué índice crearía.
- d) ¿Cree que una mayor cantidad de memoria podría mejorar significativamente el costo de esta consulta? Justifique su respuesta.

AUTORES	LIBROS
$n(\text{Autores}) = 50.000$	$n(\text{Libros}) = 1.000.000$
$B(\text{Autores}) = 5.000$	$B(\text{Libros}) = 100.000$
$V(\text{nacionalidad}, \text{Autores}) = 100$	$V(\text{nombre_autor}, \text{Libros}) = 50.000$
	$V(\text{género}, \text{Libros}) = 50$

3. (*Diseño relacional*) Considere una relación $R(A, B, C, D, E, G)$ con el siguiente conjunto de dependencias funcionales asociado $F = \{A \rightarrow CE, BC \rightarrow D, DE \rightarrow G\}$. Para cada una de las siguientes afirmaciones, indique si la misma es verdadera o falsa, justificando su respuesta.

- $\{A, C, D, E\}$ es superclave de R .
- $\{A, B, E\}$ es clave candidata de R .
- R se encuentra en Segunda Forma Normal (2FN).
- La descomposición de R en $\{R_1(A, C, E), R_2(B, C, D), R_3(D, E, G)\}$ preserva la información.
- La descomposición de R en $\{R_1(A, C, E), R_2(A, B, D), R_3(D, E, G)\}$ preserva todas las dependencias funcionales.

4. (*SQL*) José Capablanca fue un ajedrecista cubano considerado uno de los mejores de todos los tiempos. Gran Maestro del Ajedrez y campeón mundial, se le atribuye haber permanecido invicto durante 8 años, desde el 10 de febrero de 1916 hasta el 21 de marzo de 1924.

Las siguientes dos tablas almacenan información sobre las partidas de ajedrez en que participó Capablanca, y el detalle de cada una de las movidas ocurridas en cada partida.

- `PartidasCapablanca(cod_partida, nombre_jugador_blancas, nombre_jugador_negras, fecha, lugar)`
`//(7, 'José Capablanca', 'Alexander Alekhine', 1927-07-10, 'Buenos Aires, Argentina')`
- `Movidas(cod_partida, nro_movimiento, color_piezas, notación_movida)`
`//(7, 26, 'BLANCAS', 'Cg7++')`

Una partida de ajedrez se compone de una serie de movidas alternadas entre un jugador que utiliza las piezas blancas y otro que utiliza las piezas negras, hasta que uno de ellos logra hacer un “jaque mate” a la pieza que representa al rey del equipo opuesto, ganando así la partida.

Las movidas se describen mediante una notación estándar. Por ejemplo, la notación `Cc3` indica que el jugador *mueve el caballo a la posición c3 del tablero*. El estudio de esta notación excede al presente curso, y nos alcanzará con saber que los caracteres `++` contenidos al final de la notación de una movida indican que con esa movida el jugador hizo un jaque mate al rival, ganando la partida.

Por ejemplo, la tupla `(7, 26, 'BLANCAS', 'Cg7++)` indica que en la movida número 26 de la partida 7, el jugador que juega con las blancas movió el caballo a la posición g7, haciendo un jaque mate sobre las negras y ganando la partida.

Escriba una consulta en SQL que verifique que después del 10 de febrero de 1916 y hasta el 21 de marzo de 1924 José Capablanca no perdió ninguna partida (en otras palabras, que no le realizaron ningún *jaque mate*).

Ayuda: Considere utilizar el operador de comparación `LIKE` para encontrar las movidas que representan un jaque mate. En este contexto, el carácter `'%'` dentro del patrón de comparación representa una cantidad arbitraria de caracteres.

5. (*NoSQL*) La diosa griega *Afrodita* tuvo 16 descendientes, algunos de ellos mortales y otros inmortales (dioses), fruto de su relación con distintos dioses y mortales. Alejandro, que es fanático de la mitología, volcó esta información como parte de una base de datos en Neo4j para construir una base de conocimiento (*knowledge base*) de la mitología griega.

En esta base, los dioses y los mortales se representan como:

```
1      (afrodita:Dios:Ser { nombre: 'Afrodita', género: 'femenino'})
2      (eneas:Mortal:Ser { nombre: 'Eneas', género: 'masculino'})
3      ...
```

Mientras que las relaciones de filiación se describen de la siguiente forma:

```
1      (eneas)-[:HIJO_DE]->(afrodita)
2      (eneas)-[:HIJO_DE]->(anquises)
```

Escriba una consulta en Neo4j que encuentre a los dioses con los que Afrodita tuvo hijos, indicando para cada uno el nombre del dios y la cantidad de hijos que tuvo con él, y ordenando el resultado de manera de comenzar por los dioses con los que más hijos tuvo.

6. (*Seguridad*) El modelo de seguridad RBAC tiene la posibilidad de expresar jerarquías de roles. ¿Cuál/es de la/s siguiente/s utilidad/es cree que posee el uso de jerarquías de roles en la definición de permisos de seguridad? Justifique su respuesta en cada caso.

- Permiten cumplir con el criterio del menor privilegio posible.
- Permiten minimizar la redundancia en la definición de las reglas de seguridad.
- Facilitan el mantenimiento de las reglas de seguridad.
- Permiten dividir las responsabilidades.
- Permiten asignar a un mismo usuario más de un rol.
- Evitan cometer errores en la asignación de permisos.