

**Base de Datos (75.15 / 75.28 / 95.05)**

Evaluación Integradora - 18 de julio de 2018

24° aniversario del atentado a la AMIA

|                     |  |              |  |               |  |   |
|---------------------|--|--------------|--|---------------|--|---|
| <b>TEMA 20181C3</b> |  |              |  |               |  | <b>Padrón:</b> _____  |
| <b>SQL</b>          |  | <b>Proc.</b> |  | <b>CyT</b>    |  | <b>Apellido:</b> _____  |
| <b>CyT2</b>         |  | <b>NoSQL</b> |  | <b>NoSQL2</b> |  | <b>Nombre:</b> _____  |
| Corrigió:           |  |              |  |               |  | <b>Cantidad de hojas:</b> _____   |
| <b>Nota:</b>        |  |              |  |               |  | <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. (*SQL*) Una base de datos guarda dos matrices cuadradas ralas  $A$  y  $B$  de grandes dimensiones a través de sendas tablas  $A(\underline{i}, \underline{j}, v)$  y  $B(\underline{i}, \underline{j}, v)$ , en donde los atributos  $i$  y  $j$  identifican a los elementos  $(i, j)$  de las matrices, mientras que  $v$  representa su valor,  $a_{ij}$  ó  $b_{ij}$  según el caso. Cada tabla almacena únicamente aquellas posiciones en las que el valor de la matriz es distinto de 0.

Escriba una consulta en lenguaje SQL que devuelva como resultado una tabla  $R(\underline{i}, \underline{j}, v)$  que represente la suma de ambas matrices,  $R = A + B$ . Sólo almacene aquellas posiciones para las cuales el valor del resultado es distinto de 0. En el Cuadro 1 encontrará un pequeño ejemplo ilustrativo.

Nota: En algunas de las posibles soluciones le será necesario reemplazar valores nulos. Puede utilizar para ello la función estándar `COALESCE(v, w)`, que devuelve como resultado el valor de  $v$  cuando  $v$  no es NULL, y el valor de  $w$  cuando  $v$  es NULL.

| $A = \begin{bmatrix} 3 & 0 & -7 \\ -1 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$   | $B = \begin{bmatrix} 5 & 1 & 7 \\ 0 & 2 & 0 \\ 0 & -3 & -1 \end{bmatrix}$ | $R = \begin{bmatrix} 8 & 1 & 0 \\ -1 & 6 & 0 \\ 0 & -3 & 0 \end{bmatrix}$ |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
|---|---|---|---|---|---|---|---|---|----|---|---|----|---|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|----|---|---|---|---|---|----|
| <b>A</b>  | <b>B</b>  | <b>R</b>  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| <table><tr><th>i</th><th>j</th><th>v</th></tr><tr><td>1</td><td>1</td><td>3</td></tr><tr><td>1</td><td>3</td><td>-7</td></tr><tr><td>2</td><td>1</td><td>-1</td></tr><tr><td>2</td><td>2</td><td>4</td></tr><tr><td>3</td><td>3</td><td>1</td></tr></table> | i   | j   | v | 1 | 1 | 3 | 1 | 3 | -7 | 2 | 1 | -1 | 2 | 2 | 4 | 3 | 3 | 1 | <table><tr><th>i</th><th>j</th><th>v</th></tr><tr><td>1</td><td>1</td><td>5</td></tr><tr><td>1</td><td>2</td><td>1</td></tr><tr><td>1</td><td>3</td><td>7</td></tr><tr><td>2</td><td>2</td><td>2</td></tr><tr><td>3</td><td>2</td><td>-3</td></tr><tr><td>3</td><td>3</td><td>-1</td></tr></table> | i | j | v | 1 | 1 | 5 | 1 | 2 | 1 | 1 | 3 | 7 | 2 | 2 | 2 | 3 | 2 | -3 | 3 | 3 | -1 | <table><tr><th>i</th><th>j</th><th>v</th></tr><tr><td>1</td><td>1</td><td>8</td></tr><tr><td>1</td><td>2</td><td>1</td></tr><tr><td>2</td><td>1</td><td>-1</td></tr><tr><td>2</td><td>2</td><td>6</td></tr><tr><td>3</td><td>2</td><td>-3</td></tr></table> | i | j | v | 1 | 1 | 8 | 1 | 2 | 1 | 2 | 1 | -1 | 2 | 2 | 6 | 3 | 2 | -3 |
| i   | j   | v   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 1   | 3   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 3   | -7  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 2   | 1   | -1  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 2   | 2   | 4   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 3   | 3   | 1   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| i   | j   | v   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 1   | 5   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 2   | 1   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 3   | 7   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 2   | 2   | 2   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 3   | 2   | -3  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 3   | 3   | -1  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| i   | j   | v   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 1   | 8   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 1   | 2   | 1   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 2   | 1   | -1  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 2   | 2   | 6   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |
| 3   | 2   | -3  |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |

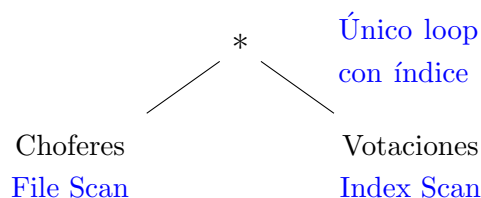
Cuadro 1: Ejemplo correspondiente al Ej. 1. Matrices  $A$ ,  $B$  y  $R$  (*arriba*) y su representación como tabla (*abajo*).

2. (*Procesamiento de Consultas*) El concurso *Grandes Choferes* ha sido un éxito: cientos de miles de usuarios del transporte público votaron a sus choferes de colectivo preferidos, y el Gobierno de la Ciudad ha volcado la información en una base de datos relacional con las siguientes tablas:

- Usuarios(código\_SUBE, apellido, nombre, dirección\_correo)
- Votaciones(código\_SUBE, CUIL\_chofer, fecha\_hora)
- Choferes(CUIL\_chofer, apellido, nombre, línea\_colectivo)

Adicionalmente se cuenta con un índice de *clustering* por el atributo CUIL\_chofer en la tabla *Votaciones*.

El Gobierno de la Ciudad quiere analizar cuáles fueron las líneas de colectivos más votadas, y para ello deberá primero calcular la junta de las tablas *Choferes* y *Votaciones*, para lo cual construye el siguiente plan de ejecución:



Se pide:

- a) Estime el costo del plan de ejecución en términos de cantidad de accesos a disco, considerando que la junta se realiza por el método de único loop con índice.
- b) Estime la cardinalidad del resultado de la junta, tanto en términos de cantidad de tuplas como de cantidad de bloques en disco.

Utilice para sus cálculos la siguiente información de catálogo:

| CHOFERES  | VOTACIONES   |
|---|--|
| $n(\text{Choferes}) = 20000$                        | $n(\text{Votaciones}) = 400000$                    |
| $B(\text{Choferes}) = 2000$                         | $B(\text{Votaciones}) = 40000$                     |
| $V(\text{línea\_colectivo}, \text{Choferes}) = 200$ | $V(\text{CUIL\_chofer}, \text{Votaciones}) = 2000$ |
|   | $H(I(\text{CUIL\_chofer}, \text{Votaciones})) = 3$ |

Nota: No considere el costo de almacenar el resultado final de la consulta en disco.

3. (*Concurrencia y Transacciones*) Dado el siguiente par de transacciones:

$$T_1 : b_{T_1}; R_{T_1}(Y); W_{T_1}(Z); W_{T_1}(X); c_{T_1};$$

$$T_2 : b_{T_2}; R_{T_2}(X); R_{T_2}(Y); W_{T_2}(Y); W_{T_2}(X); c_{T_2};$$

Proponga un solapamiento de las mismas que sea serializable y recuperable, pero que no evite *rollbacks* en cascada en caso de que alguna de las transacciones aborte en algún momento de su ejecución.

4. (*Concurrencia y Transacciones II*) Para las transacciones  $T_1$  y  $T_2$  definidas en el Ejercicio 3, introduzca *locks* en la ejecución de cada una respetando el *Protocolo de Lock de 2 Fases (2PL)*, y luego proponga un solapamiento que produzca un *deadlock*.

Nota: Escriba el solapamiento únicamente hasta el punto en que las transacciones queden en *deadlock*.

Nota: Utilice la notación  $L_{T_i}(V)$  y  $U_{T_i}(V)$  para indicar que una transacción  $T_i$  adquiere ó libera un *lock* sobre un ítem  $V$ .

5. (*NoSQL*) Explique en qué consiste el modelo de bases de datos *orientadas a grafos* y cuáles son sus ventajas respecto al modelo relacional de bases de datos. Mencione un ejemplo de consulta que pueda resolverse en forma más simple en una base de datos orientada a grafos respecto a una base de datos relacional.
6. (*NoSQL II*) Una cadena de supermercados almacena información sobre las ventas que realiza en una base de datos MongoDB. El siguiente documento JSON ejemplifica la estructura almacenada por cada venta, que incluye los siguientes campos:

- El número de ticket, que se utiliza como *\_id*.
- La fecha y hora de la venta.
- El DNI del comprador.
- El nombre del comprador.
- Un vector de documentos embebidos conteniendo cada uno un código de producto, nombre de producto, cantidad de unidades y precio por dicha cantidad de unidades.
- El monto total de la venta.

```

1  {
2    "_id": "0057-46290015",
3    "fecha_venta": Date("2018-02-06 11:53:28"),
4    "DNI_comprador": 38115218,
5    "nombre_comprador": "BENITEZ, JUAN ANTONIO",
6    "productos": [
7      { "codigo_producto": 108431,
8        "nombre_producto": "pañales makako 20 un.",
9        "cantidad_producto": 3,
10       "precio_por_cantidad": 241.50
11     },
12     { "codigo_producto": 210160,
13       "nombre_producto": "papas fritas boom 150 gr.",
14       "cantidad_producto": 1,
15       "precio_por_cantidad": 41.00
16     }
17   ],
18   "monto_total": 282.50

```

La gerencia identifica a las familias que tienen bebés como aquellas que adquieren pañales “*Makako*” en sus compras. Recientemente se ejecutó la siguiente consulta para analizar la evolución mensual de la cantidad de compradores únicos que han adquirido dichos pañales:

```
[
  { $match: { "productos.codigo_producto": { $in: [108431] } } },
  { $project: { mes: { $month: "$fecha_venta" },
                año: { $year: "$fecha_venta" },
                comprador: "$DNI_comprador" } } },
  { $group: { _id: { "año": "$año", "mes": "$mes" },
              compradores: { $addToSet: "$comprador" } } } },
  { $project: { cant_compradores: { $size: "$compradores" } } },
  { $sort: { "_id.año": -1 , "_id.mes": -1 } },
  { $limit: 10 }
]
```

Lamentablemente, esta consulta mostró que en los últimos 10 meses la cantidad de compradores de pañales ha ido en disminución, y la gerencia lo atribuye a que las familias con bebés están frecuentando menos el supermercado.

Se quiere entonces proponer ofertas interesantes para las familias que tienen bebés, y para ello se desea saber cuáles son los productos más frecuentemente adquiridos junto con pañales *Makako*. Escriba una consulta en MongoDB que encuentre los cinco productos que aparecen más frecuentemente en tickets de compra que incluyen pañales *Makako*, indicando el código del producto y la cantidad de tickets de compra distintos en los que aparece combinado con dichos pañales. El formato de los documentos resultantes deberá ser:

```
1 {
2   "_id": 210160,
3   "cantidad_tickets_compra": 3804
4 }
5 ...
```