

## **4° PARCIALITO**

## **BASE DE DATOS**

**(TA044)**

**Cuatrimestre:** 1º Cuatrimestre 2025

**Curso:** 1 - Román

Brian Céspedes	108219	20-05-2025
----------------	--------	------------

## 1. Estimación de Costos

Podemos estimar el costo de la junta en base a los 3 algoritmos conocidos y vistos en clase: loops anidados por bloque, sort-merge y junta hash grace.

### a) Loops anidados por bloque

$$\begin{aligned} \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= \\ &= B(\text{listadoempresas}) + \lceil \frac{B(\text{listadoempresas})}{M-2} \rceil \cdot B(\text{vendedores}) \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 1000 + \lceil \frac{1000}{12-2} \rceil \cdot 2000 \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 1000 + \lceil \frac{1000}{10} \rceil \cdot 2000 \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 1000 + \lceil 100 \rceil \cdot 2000 \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 1000 + 100 \cdot 2000 \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 201000 \end{aligned}$$

### b) Sort-Merge

Calculamos el costo de ordenar listadoempresas con M bloques de memoria disponibles:

$$\begin{aligned} \text{Cost}(\text{Ord}_M(\text{listadoempresas})) &= 2 \cdot B(\text{listadoempresas}) \cdot \lceil \log_{M-1}(B(\text{listadoempresas})) \rceil \\ \text{Cost}(\text{Ord}_M(\text{listadoempresas})) &= 2 \cdot 1000 \cdot \lceil \log_{12-1}(1000) \rceil \\ \text{Cost}(\text{Ord}_M(\text{listadoempresas})) &= 2 \cdot 1000 \cdot \lceil \log_{11}(1000) \rceil \\ \text{Cost}(\text{Ord}_M(\text{listadoempresas})) &= 2 \cdot 1000 \cdot \lceil 2,88 \rceil \\ \text{Cost}(\text{Ord}_M(\text{listadoempresas})) &= 2 \cdot 1000 \cdot 3 \\ \text{Cost}(\text{Ord}_M(\text{listadoempresas})) &= 6000 \end{aligned}$$

Calculamos el costo de ordenar vendedores con M bloques de memoria disponibles:

$$\begin{aligned} \text{Cost}(\text{Ord}_M(\text{vendedores})) &= 2 \cdot B(\text{vendedores}) \cdot \lceil \log_{M-1}(B(\text{vendedores})) \rceil \\ \text{Cost}(\text{Ord}_M(\text{vendedores})) &= 2 \cdot 2000 \cdot \lceil \log_{12-1}(2000) \rceil \\ \text{Cost}(\text{Ord}_M(\text{vendedores})) &= 2 \cdot 2000 \cdot \lceil \log_{11}(2000) \rceil \\ \text{Cost}(\text{Ord}_M(\text{vendedores})) &= 2 \cdot 2000 \cdot \lceil 3,17 \rceil \\ \text{Cost}(\text{Ord}_M(\text{vendedores})) &= 2 \cdot 2000 \cdot 4 \\ \text{Cost}(\text{Ord}_M(\text{vendedores})) &= 16000 \end{aligned}$$

Calculamos el costo total:

$$\begin{aligned} \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= \\ &= \text{Cost}(\text{Ord}_M(\text{listadoempresas})) + \text{Cost}(\text{Ord}_M(\text{vendedores})) + \\ &\quad + B(\text{listadoempresas}) + B(\text{vendedores}) \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 6000 + 16000 + 1000 + 2000 \\ \text{Cost}(\text{listadoempresas} \bowtie \text{vendedores}) &= 25000 \end{aligned}$$

Es importante notar que en las clases teóricas para este método se mencionó el limitante de memoria, pero en este caso podemos asumir que dicha restricción no se aplica.

### c) Junta hash grace

Para aplicar el método, es necesario definir la cantidad de particiones  $P$  a realizar, esta cantidad está limitada por la memoria disponible  $M$ .

- **Primer límite:**

$$P \leq M - 1$$

$$P \leq 12 - 1$$

$$P \leq 11$$

- **Segundo límite:**

$$\min\left(\left\lceil \frac{B(\text{listadoempresas})}{P} \right\rceil, \left\lceil \frac{B(\text{vendedores})}{P} \right\rceil\right) \leq M - 2$$

$$\min\left(\left\lceil \frac{1000}{P} \right\rceil, \left\lceil \frac{2000}{P} \right\rceil\right) \leq 12 - 2$$

$$\min\left(\left\lceil \frac{1000}{P} \right\rceil, \left\lceil \frac{2000}{P} \right\rceil\right) \leq 10$$

- **Tercer límite:**

$$\min(V(\text{ciudad}, \text{listadoempresas}), V(\text{ciudad}, \text{vendedores})) \geq P$$

$$\min(100, 50) \geq P$$

$$50 \geq P$$

$$P \leq 50$$

Del primer límite, supongamos que  $P = 11$  (se seguiría cumpliendo el tercer límite), luego, en el segundo límite:

$$\min\left(\left\lceil \frac{1000}{11} \right\rceil, \left\lceil \frac{2000}{11} \right\rceil\right) \leq 10$$

$$\min([90, 91], [181, 82]) \leq 10$$

$$\min(91, 182) \leq 10$$

$91 \leq 10$ , lo cual nos lleva a un absurdo. No podemos aplicar este método para esta junta, notemos que para cualquier valor de  $P < 11$  seguiríamos llegando a algo inconsistente.

## 2. Cantidad de Filas Devueltas

Podemos hallar la cantidad de filas devueltas de la siguiente manera:

$$n(\text{listadoempresas} \bowtie \text{vendedores}) = \frac{n(\text{listadoempresas}) \cdot n(\text{vendedores})}{\max(V(\text{ciudad}, \text{listadoempresas}), V(\text{ciudad}, \text{vendedores}))}$$

$$n(\text{listadoempresas} \bowtie \text{vendedores}) = \frac{10000 \cdot 50000}{\max(100, 50)}$$

$$n(\text{listadoempresas} \bowtie \text{vendedores}) = \frac{500000000}{100}$$

$$n(\text{listadoempresas} \bowtie \text{vendedores}) = 5000000$$

### 3. Conclusión

El mejor método para realizar esta junta es el de **Sort-Merge**. Se ha estimado un costo de 25000. Se devolverá una cantidad de 5000000 de filas.