

Segundo Parcial – BD 2025

Nombre: _____

Ejercicio 1 (4 puntos): Sea la BD de una pizzería con las siguientes tablas:

cliente (cid, nombre, teléfono, dirección, edad)
pizzas (pid, znombre, tamaño, precio)
pedido (cid, pid, phora, paño, pmes, pdia, cantidad)

Se tiene la siguiente información de las tablas:

cliente	pizzas	pedido
1200 tuplas	40 tuplas	1.600.000 tuplas
	V(pizzas,tamaño) = 4	V(pedido,Zid) = 40
	V(pizzas,znombre) = 10	V(pedido,cid) = 1200
		V(pedido,paño) = 40

Asumir que un bloque tiene tamaño de 4096 bytes. Todos los años se piden todas las pizzas.

$\Pi_{\text{z.nombre}} (\text{cliente} \bowtie \sigma_{\text{paño}=2025} (\text{pedido}))$

Puede asumir que el último año es el 2025. Asumir que pedido tiene un índice primario en paño. Sea la consulta:

Responder:

1. Estimar la cantidad de tuplas del resultado intermedio de la selección (calcular en no más de 2 líneas) – 30%
2. ¿Cuántos bloques tiene el resultado intermedio de la selección? (calcular en no más de 3 líneas) – 30%
3. ¿Cuántas claves de búsqueda entran en un nodo del índice sobre pedido? (calcular en no más de 5 líneas) – 40%

Ejercicio 2 (3 puntos): (optimización de consultas) Considerar la reunión natural de tres tablas: S, T y U. Por simplicidad asumimos que cada una tiene 1000 tuplas. Asumimos:

$S(b, c)$	$T(c, d)$	$U(d, a)$	
$V(U, a) = 50$			
$V(S, b) = 100$			
$V(S, c) = 500$	$V(T, c) = 20$		
	$V(T, d) = 50$	$V(U, d) = 1000$	

Se trabaja con el algoritmo de programación dinámica para encontrar orden de aplicación de reuniones naturales. Usar lo siguiente:

	{S}	{T}	{U}		{S,T}	{S,U}	{T,U}
Size	1000	1000	1000	Size	2000	1,000,000	1000
Cost	0	0	0	Cost	0	0	0
Best plan	S	T	U	Best plan	$S \bowtie T$	$S \bowtie U$	$T \bowtie U$

1. Considerando las tres tablas S, T y U. ¿cuál es la partición ganadora que se escoge y cuál es el mejor plan? (solo dar el resultado en 2 líneas). 50%
2. Calcular el tamaño de $((S \bowtie T) \bowtie U)$ en no más de 5 líneas. 50%

Ejercicio 3 (3 puntos): (retorno de la información) Responder:

1. Supongamos que tenemos el lenguaje de consultas booleanas. Asumir que se usan índices invertidos. Expresar la definición del operador $t_1 \text{ XOR } t_2$ (t_1 y t_2 términos). Debe estar en términos de operadores de conjuntos. (Usar no más de dos líneas).
2. Supongamos que tenemos una colección de documentos de solo campeonatos de fútbol y hay que extraer el vocabulario de la colección. Listar tres tipos de términos a extraer (de los cubiertos en el capítulo); sea bien específico (responder en una linea).