Las filas que desaparecen

Dos tablas, una respuesta: `pokemon` + `generations` Nombres _____

Superset → **SQL → SQL Lab**, base de datos `Dades classe`, esquema `public`. Hoy, dos cajas de zapatos: `pokemon` (949 cartas) y una segunda, diminuta, que subís vosotros, `generations` (cartas de región). Por parejas, un solo teclado; cambiad quién escribe al acabar la Parte A. Anotad aquí los resultados.

Referencia

Escribís	Con las dos cajas
`FROM pokemon AS p`	`p` es un **alias**: el nombre corto de la caja. `p.pokemon` es la línea `pokemon` de una carta de `p`.
`ON p.x = g.x`	La regla del clip: qué línea debe ser igual en las dos cartas. Esa columna compartida es la **clave**.
`INNER JOIN`	Conservad solo las **parejas con clip**. Una carta sin pareja se queda sobre la mesa.
`LEFT JOIN`	Conservad **todas las cartas de la caja izquierda** (la tabla escrita primero, después de `FROM`). ¿Sin pareja? Recibe una carta de región en blanco: `NULL` en todas las columnas `g.`.

Parte A · Una segunda caja, con clip (14 min)

A1 Descargad `generations.csv` de la página Recursos del sitio web del curso y subidlo: **+ (arriba a la derecha) → Data → Upload CSV to database**. **CSV file** `generations.csv` · **Database** `Dades classe` · **Schema** `public` (nunca vacío) · **Table name** `generations` · **Delimiter** , → **Upload**. Luego ejecutad `SELECT * FROM generations;`

Filas _____ Columnas _____ ¿Qué columna está impresa en las cartas de **las dos** cajas? _____

A2. Sujetad las cartas con el clip. Ejecutad esto y leed la primera fila.

```
SELECT p.pokemon, p.type_1, g.region, g.release_year
FROM pokemon AS p
INNER JOIN generations AS g ON p.generation_id = g.generation_id
ORDER BY p.id;
```

Primera fila: _____

`pokemon` tiene 949 filas. **Antes de ejecutar nada más**, haced una predicción: ¿cuántas filas tiene este join? _____
Ahora contadlas: sustituid la primera línea por `SELECT COUNT(*)` y borrad la línea del `ORDER BY`.

Filas _____ ¿Ha habido algún error o aviso? _____

A3 En vuestro recuento, cambiad una palabra: `INNER` → `LEFT`.

Filas _____ Diferencia con A2 _____ ¿Dónde visteis ese número la semana pasada? _____

A4. Esta falla **a propósito**: el error es la lección.

```
SELECT pokemon, region FROM pokemon
INNER JOIN generations ON generation_id = generation_id;
```

El error dice: _____

¿Por qué la base de datos no puede responder? _____

La línea `ON` que funciona: _____

¿Atascados?

- **`generations` tiene 16 filas**: el CSV se subió dos veces. Repetid con **If table already exists → Replace**.
- **El join da 7592 filas**: el `ON` está mal o falta: cada Pokémon se ha unido a cada región (949 × 8).
- **`is ambiguous`** fuera de A4: una columna que está en las dos cajas necesita su prefijo, `p.` o `g.`.
- Una consulta cada vez. ¿Atascados tras dos intentos? Preguntad a los vecinos y luego al profesor.

Parte B · Las filas que desaparecen (20 min)

Una consulta por pregunta, siempre con los alias `p` y `g` y `ON p.generation_id = g.generation_id`. Para cada join, apuntad cuántas filas salen.

B1 Región y primer juego de `pikachu`, `lucario` y `greninja`. Una sola consulta.

`pikachu` _____ `lucario` _____ `greninja` _____

B2 ¿Cuántos Pokémon hay por región? `INNER JOIN` + `GROUP BY g.region`.

Filas _____ Más: _____ con _____ Menos: _____ con _____

B3 Lo mismo con `LEFT JOIN`.

Filas _____ ¿Cuál es la fila nueva? _____

B4 Buscad `p.pokemon LIKE 'charizard%'` con cada join.

`INNER` _____ filas `LEFT` _____ filas ¿Quién falta? _____

B5 Listad los Pokémon que el `INNER JOIN` deja fuera: `LEFT JOIN` más `WHERE g.region IS NULL`. Leed veinte nombres.

¿Cuántos? _____ ¿Qué clase de Pokémon son? _____

¿Por qué desaparecen? _____

B6 Dad la vuelta al join: `FROM generations AS g LEFT JOIN pokemon AS p`, agrupad por región, usad `COUNT(p.id)`.

Filas _____ Región que aparece por primera vez: _____ con _____ Pokémon

¿Por qué no había aparecido nunca antes? _____

B7 Los cinco Pokémon con más ataque, con su región: una vez con `INNER`, otra con `LEFT`.

El primero con `INNER` _____ El primero con `LEFT` _____

¿Qué top 5 publicaríais, y por qué? _____

_____

B8 Ataque medio por región, la región más fuerte primero (redondeado, `INNER JOIN`).

Las tres primeras: _____

Tres recuentos, una tabla

Rellenadla al final de la clase, con todo el grupo.

Join	Filas	Quién falta
`pokemon INNER JOIN generations`		
`pokemon LEFT JOIN generations`		
`generations LEFT JOIN pokemon`		

Habéis acabado cuando Todos los huecos están rellenos · B5, B6 y B7 tienen su frase. **¿Pareja rápida?** Con vuestro conjunto de datos abiertos: ¿qué columna es un código que necesita una segunda tabla para entenderse (código de municipio, isla, número de categoría)? Decid cuál es la clave y qué tres columnas debería tener esa tabla auxiliar.

En casa Una tienda tiene `customers(customer_id, name)` y `orders(order_id, customer_id)`. Escribid la consulta que lista los clientes que nunca han hecho un pedido. ¿Qué join, qué tabla a la izquierda, qué filtro?