The same cleaning, twice

Power Query on the projector, SQL in your hands Names _____

The teacher does each step in the **Power Query Editor** of Power BI Desktop, on the projector, with `video_games.csv`. With your own Windows laptop, do it too. After each step you write the SQL that does the same on the table `games`: Superset → **SQL → SQL Lab**, database `Dades classe`, schema `public`. Pairs: one types, one writes; swap at step 4. Step names and check numbers go here.

Reference · the Power Query Editor

Part	What it is
Ribbon	The tabs with every transformation: **Home**, **Transform**, **Add Column**, **View**.
Queries pane (left)	One query per table you brought in. Today: one, `video_games`.
Data preview (centre)	A preview of the table **after** the selected step. Not the whole file.
Query settings (right)	The query's name and the **Applied Steps** list.

Power Query is a kitchen that writes down the recipe while you cook: every click becomes a line in **Applied Steps**. In SQL you write the whole recipe first, then the kitchen cooks it. Same dish.

The steps

Step	On the projector (Power Query)	Your SQL twin · write
0	**View** tab → tick **Column quality** and **Column distribution**. Then click the message **Column profiling based on top 1000 rows** (bottom left) and switch to the entire data set.	`metascore` was 89% empty in SQL. Power Query says **Empty**: _____ % After the switch: _____ % Rows read the first time: _____
1	Column `price`: **Transform → Replace Values**: `NA` → `null`.	Nothing to write in SQL: the upload already did it. Check with `COUNT(*) - COUNT(price)`. Applied step: _____ Check: _____
2	Right-click the `price` header → **Change Type → Using Locale...** → **Decimal Number**, **English (United States)**.	`price` is already a number in `games`. Applied step: _____ Who decided that `9.99` uses a dot? _____
3	**Add Column → Conditional Column**: name `is_free`; if `price` equals `null` then `true` else `false`. Then **Replace Values** on `price`: `null` → `0`.	The two expressions of 16/11: fill the empty prices, flag the free games. Applied steps: _____ Check, rows with `is_free`: _____
4	Column `owners`: **Transform → Extract → Text Before Delimiter** `..` ; **Replace Values** `,` → nothing; **Replace Values** with **Advanced options → Replace using special characters → Non-breaking Space** → nothing; type **Whole Number**.	`owners_min`, as on 09/11: the text before `..`, digits only, as a number. Applied steps: _____ Check, `MAX(owners_min)`: _____

Step	On the projector (Power Query)	Your SQL twin · write
5	`release_date`: **Change Type → Using Locale… → Date**, **English (United States)**. Look at **Column quality**.	The `CASE` with two `TO_DATE` of 04/11. Any **Error** on the projector? _____ Check, rows that are month-only: _____ that fit neither format: _____
6	**Add Column → Conditional Column**: `price_band`, five branches. Then the filter arrow on `game`: untick `NA`.	The `CASE WHEN` of 11/11 and the filter of 16/11. Applied steps: _____ Check, final `COUNT(*)`: _____
7	**Home → Close & Apply**, then back with **Home → Transform data**. Click an early step in **Applied Steps**, then the last one. Then **View → Advanced Editor**.	What happens to the preview? _____ What is in the Advanced Editor? _____

Q1 Why did Power Query not count `NA` as empty at first?

Q2 Step 3 has two halves. What breaks if you replace `null` with `0` **before** adding `is_free`? Is it the same in SQL?

Q3 Compare the two kitchens. Fill in the table from what you saw today.

	Power Query	SQL (SQL Lab)
How you work		
Where the recipe is		
The original file or table		
Where it runs		
Empty values		
What you need		

Done when Every step has its name and its check number · Q1 to Q3 are answered. **Windows laptop?** Show your Applied Steps list to the teacher.

Stuck?

- **`invalid input syntax for type bigint`**: you split on `' .. '` with normal spaces. Split on `'..'` and keep only digits.
- **On your laptop, `owners` will not turn into a number**: a non-breaking space is left.
- **Your check for step 5 is one number**: it should be two. Still stuck? Ask your neighbours, then the teacher.

At home Finish the handout. **Windows laptop?** Add one more step of your own and its SQL twin.