Bars and lines

The leaderboard and the timeline of `games_clean`

Name _____

Superset → **Charts**, dataset `games_clean` (database `Dades classe`, schema `public`): 26,685 Steam games. After a match you get two screens: the **leaderboard** ranks the players, the **timeline** replays the match. Today you build both. Work alone; write what you see here.

Reference

Zone of Explore	Use it for
Dataset panel (left)	Dragging a column into a control.
Data tab	Deciding **what** is computed.
Customize tab	Deciding **how it looks**.
Chart + results	The picture and the rows behind it: check the numbers.

Word	In SQL	In Explore
Dimension the player names	The column in `GROUP BY`.	**X-axis**, **Dimensions**.
Metric the score	The aggregate: `COUNT`, `AVG`, `SUM`.	**Metrics**.

Monday's chart was this query: `SELECT price_band, COUNT(*) FROM games_clean GROUP BY price_band;`

Part A · The leaderboard (16 min)

Top 10 publishers. Which publishers have released the most games on Steam?

Step	Do this	You should see / write
1	**Charts → + Chart** → dataset `games_clean`, chart type **Bar Chart** → **Create new chart**.	An empty chart in the Explore view.
2	**X-axis**: `publisher`. **Metrics**: `COUNT(*)`. **Row limit**: `10`. **Create chart**.	Ten bars. In which order? _____
3	**X-Axis Sort By** → `COUNT(*)`, untick **X-Axis Sort Ascending**. **Update chart**.	Tallest on the left. First: _____ Fifth: _____
4	Chart menu (···, top right) → **View query**.	`GROUP BY` _____ (the dimension) Aggregate: _____ (the metric) Last line: `LIMIT` _____
5	**Filters** → `publisher` → operator **Not equal to (≠)** → value `Unknown` → **Save**. **Update chart**.	No bar called `Unknown`. Who entered the top 10? _____
6	**Customize** tab → **Bar orientation** → **Horizontal**. The sort controls are now called **Y-Axis Sort By** and **Y-Axis Sort Ascending**.	Names readable without tilting your head.
7	**Save** → `UT3 · Top 10 publishers`. Then read **Results**, under the chart.	Top publisher: _____ Its games: _____

A1 Is `Unknown` a publisher? What is it, and where does it come from?

A2 **Row limit** `10`: ten random publishers, or the top ten? Which step told you? And why could a new one enter in step 5?

Part B · The timeline (15 min)

Games released per year. How has the number of releases on Steam changed between 2004 and 2018?

Step	Do this	You should see / write
1	New chart on `games_clean`, chart type **Line Chart**.	An empty chart in the Explore view.
2	**X-axis**: `release_date`. **Time Grain**: **Year**. **Metrics**: `COUNT(*)`. **Create chart**.	One line. Its shape, in three words: _____
3	Hover over the point of 2018. Then look at **Results**, under the chart.	Hover box: _____ Exact number: _____
4	**Save** → `UT3 · Games released per year`.	**Charts** lists it.
5	Change **Time Grain** to **Month**. **Update chart**. Hover over the highest point.	Peak month: _____ Games that month: _____
6	Do not overwrite: **Save** → **Save as...** → `UT3 · Games released per month`.	Two saved line charts.
7	Optional. Back on **Year** grain: **Metrics** → **Simple** tab → column `price`, aggregate `AVG`. Do not save.	Did games get more expensive as they got more numerous? _____

B1 The year chart and the month chart come from the same rows. What question does each one answer?

Year _____ Month _____

Part C · Bar or line? (4 min)

No Superset: decide, and give the reason in a few words.

Question	Bar or line?	Why
Games per `owners_tier` (`Niche`, `Mid`, `Hit`)?		
Average price per release year?		
Games per publisher, top 10?		

Done when Every blank is filled · three charts are saved with the prefix `UT3 ·` (top 10 publishers, per year, per month).

Stuck?

- **Time Grain is missing or greyed out**: the X-axis is not a date column. Use `release_date`, not `release_year`.
- **A bar chart with thousands of thin bars**: **Row limit** was left at its default (10,000). Set it to 10.
- **The line is a single dot**: `release_date` went to **Dimensions** instead of **X-axis**.
- **Filter value `Unknown` finds nothing**: it is case-sensitive, capital U.
- **"Save" overwrote the year chart with the month version**: set the grain back to **Year** and save again; then **Save as...** for the month one.
- **Superset does not open**: pair up with a neighbour and tell the teacher.

At home Think for Monday: how would you show whether most games are cheap or expensive? A bar per exact price would have hundreds of bars. Optional, useful for the final project: on your own open dataset, one bar chart and, if it has a date or a year, one line chart, saved with the prefix `UT3 ·` too.