Pie, heatmap and friends

The stats screen of `games_clean` Name _____

Recall · From memory (4 min)

Laptop closed, alone. You correct it yourself later: open **Charts**, search `UT3`, and check against your December charts.

N.	Question	Your answer
1	Which chart for “how many games in each price band”? And why not a line?	
2	Which chart for “releases over time”? Which control changes year into month?	
3	Which chart shows how prices spread? Why was our first one a single giant bar?	
4	Which chart relates price and metascore? On how many games, roughly?	
5	Name two preattentive attributes, and the author of the data-ink ratio.	

Reference

Superset, dataset `games_clean` (database `Dades classe`, schema `public`). The stats screen at the end of a match: a **pie** for how your damage split between weapons, a **heatmap** for where the fights happened, one **big number** for your score, a **table** for the full scoreboard.
A pie is fine only if 1) the slices are parts of **one whole** and add up to 100 % · 2) there are **few** slices, five at most · 3) the differences are **obvious**. Fail one: use a bar chart.
A heatmap is two dimensions making a grid and a metric painting each cell: for seeing patterns, not for reading values.

Part A · The pie, and when not to (10 min)

Step	Do this	You should see / write
1	New chart on `games_clean`, chart type **Pie Chart**. **Dimensions**: `owners_tier`. **Metric**: `COUNT(*)`. **Create chart**.	Slices: _____
2	**Customize** → **Label Type**: choose one that shows the percentage.	Biggest slice: _____ Its share: _____ %
3	**Save** → `UT3 · Owners tier share`.	**Charts** lists it.
4	Do not save from here on. Change **Dimensions** to `publisher` and **Row limit** to `10`. **Update chart**.	Share of Big Fish Games in the pie: _____ %
5	In SQL Lab: `SELECT ROUND(100.0 * COUNT(*) FILTER (WHERE publisher = 'Big Fish Games') / COUNT(*), 2) FROM games_clean;`	Its real share of all games: _____ %

A1 Which condition did the pie of step 4 break? Its slices still add up to 100 %: 100 % of what?

_____

Part B · The heatmap (14 min)

Releases by month and year. Is there a season for releasing games on Steam?

Step	Do this	You should see / write
1	**Datasets** → `games_clean` → edit (pencil) → **Calculated columns** → **Add item**. Name it `release_month`.	A new row with a **SQL expression** box.
2	The box holds the sample text `<enter SQL expression here>`: select it all and delete it. Type `EXTRACT(MONTH FROM release_date)::int` → **Save** → **OK** in **Confirm save**.	`release_month` in the dataset panel of your next chart.
3	New chart on `games_clean`, chart type **Heatmap**. **X-axis**: `release_year`. **Y-Axis**: `release_month`. **Metric**: `COUNT(*)`.	The three controls filled in.
4	**Filters** → `release_year` → **Greater or equal (>=)** → `2014` → **Save**.	A chip under **Filters**.
5	**Sort X Axis** and **Sort Y Axis**: **Axis ascending**. **Create chart**.	Grid: _____ columns × _____ rows Darkest: _____ Month 1: _____ row
6	Set **Normalize Across** to **x**, **Update chart**. Then **Customize** → tick **Normalized**. **Save** → `UT3 · Releases by month and year`.	Palest month in every column: _____

B1 After step 5 the colour only repeats what one of your December charts said: which one? After step 6 the counts are the same numbers: what changed?

_____

Part C · One number, one table (8 min)

Step	Do this	You should see / write
1	New chart, chart type **Big Number**, **Metric** `COUNT(*)`. Then **Customize** → **Number format** → `,d`. **Save** → `UT3 · Total games`.	Before `,d`: _____ After: _____
2	New chart, chart type **Table**. **Query mode**: **Raw records**. **Columns**: `game`, `publisher`, `release_year`, `price`, `owners_min`. **Filters**: `owners_min` **Greater or equal (>=)** `20000000`. **Ordering**: `owners_min [desc]`. **Save** → `UT3 · Biggest games table`.	Rows: _____ First: _____ How many are free? _____

C1 Write one question about `games_clean` that a table answers better than any chart.

_____

Done when Every blank is filled · four charts are saved with the prefix `UT3 ·` (the lying pie is **not** one of them).

Stuck?

- **Database error as soon as `release_month` is used**: the sample text is still in front of the expression. Edit the dataset, open the column with the + at its left, and delete it.
- **`release_month` is not in the dataset panel**: save the dataset edit and reload the chart page. **Months in the order 1, 10, 11, 12, 2**: the expression must end in `::int`.
- **Pie with a hundred thin slices**: **Row limit** must be 10. **The table shows 1,000 rows**: count the zeros in 20000000.

At home For each saved `UT3 ·` chart, write in one line the question it answers: your cheat sheet for Wednesday's drill.