

Pivot Tables from Basic to Advanced

Summarise thousands of rows in seconds. Build pivot tables, understand the pivot cache and the four areas, change value field settings, show values as % of total, MoM and YoY change and running totals, group dates and numbers into bins, add calculated fields, slicers and timelines, and solve the Video Game Sales case study.

● Live classes 5-7

✦ Excel for Microsoft 365, 2026

Intermediate

🕒 150 min

By the end of this lesson you can

- Build a pivot table and explain the pivot cache, the four areas and Refresh vs Refresh All
- Change the summary function and number format, and show values as % of total, % of row and running totals
- Calculate month-on-month (MoM) and year-on-year (YoY) change
- Group dates by year, quarter and month, group items manually, and group numbers into bins
- Add calculated fields, slicers, timelines and pivot charts, and answer a full case study

Datasets used

superstore_sales.csv

video_game_sales.csv

What a pivot table does

A pivot table **groups** rows by one or more fields and **summarises** a number for each group (sum, count, average...). It is the fastest way to answer "how much, by what?" questions, with no formulas.

Raw data: 6,323 rows

Pivot: sum of sales (₹ lakh)

order_id	region	category	sales
IN-2023-00001	South	Technology	36000.0
IN-2023-00001	South	Office Supplies	3850.0
IN-2023-00001	South	Furniture	62080.0
IN-2023-00002	West	Office Supplies	5216.0
IN-2023-00003	West	Furniture	24591.0
IN-2023-00003	West	Furniture	8288.0
IN-2023-00003	West	Office Supplies	2100.0
IN-2023-00003	West	Technology	69280.0
IN-2023-00004	East	Office Supplies	840.0

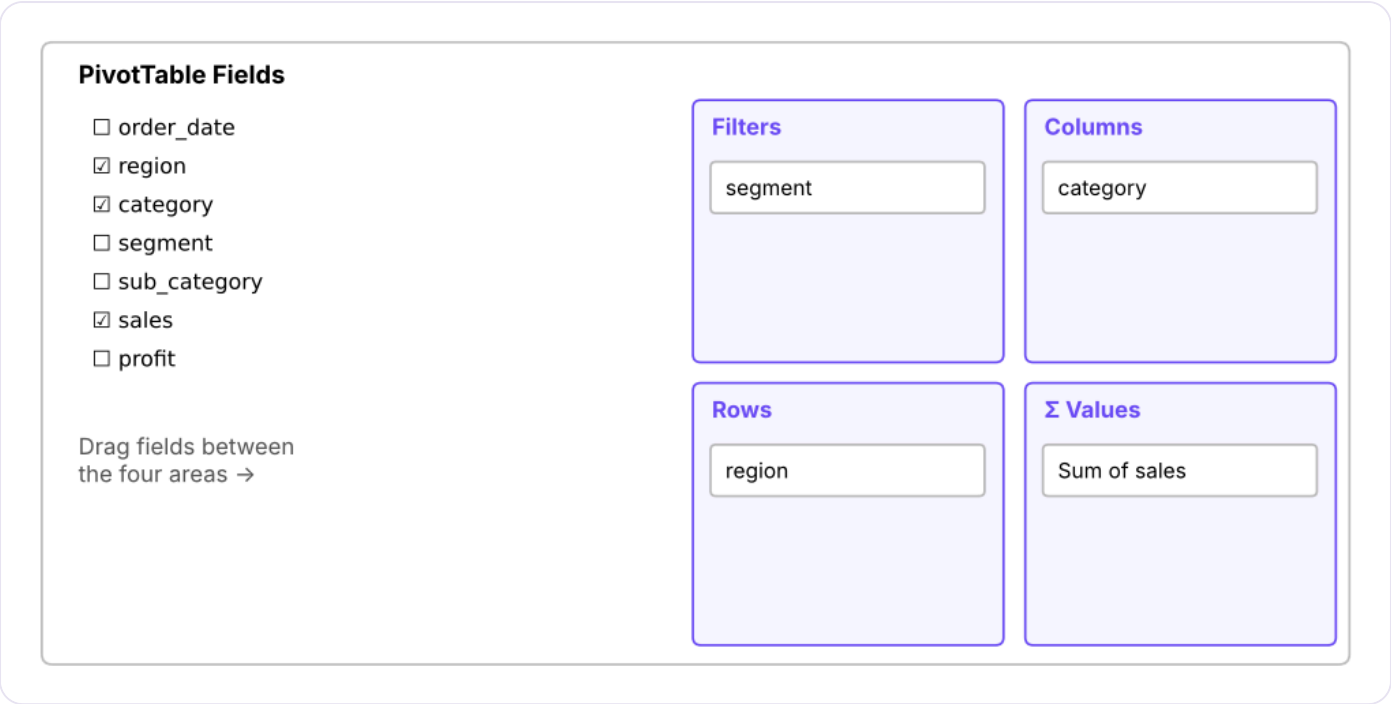
Region	Furniture	Office Supplies	Technology	Total
East	79.4	27.2	358.2	464.8
North	79.8	22.2	388.2	490.2
South	95.9	24.7	418.9	539.5
West	93.3	27.3	399.0	519.6
Grand Total	348.4	101.3	1,564.4	2,014.1

A pivot table turns thousands of order lines into a compact summary: rows grouped by Region, columns by Category, values summed.

Creating a pivot table

STEP BY STEP Step by step

1. Make sure the data is **structured** (Lesson 1): one header row, no blank rows, no merged cells. Ideally convert it to a Table (`Ctrl + T`) so new rows are included automatically.
2. Click any cell in the data → **Insert** → **PivotTable** (or `Alt, N, V`).
3. Choose **New Worksheet** → OK. An empty pivot appears with the **PivotTable Fields** pane on the right.
4. Drag fields into the four areas: drag `region` to **Rows**, `category` to **Columns**, `sales` to **Values**.
5. Right-click any number → **Number Format** → `#,##0` so the format sticks when the pivot refreshes.

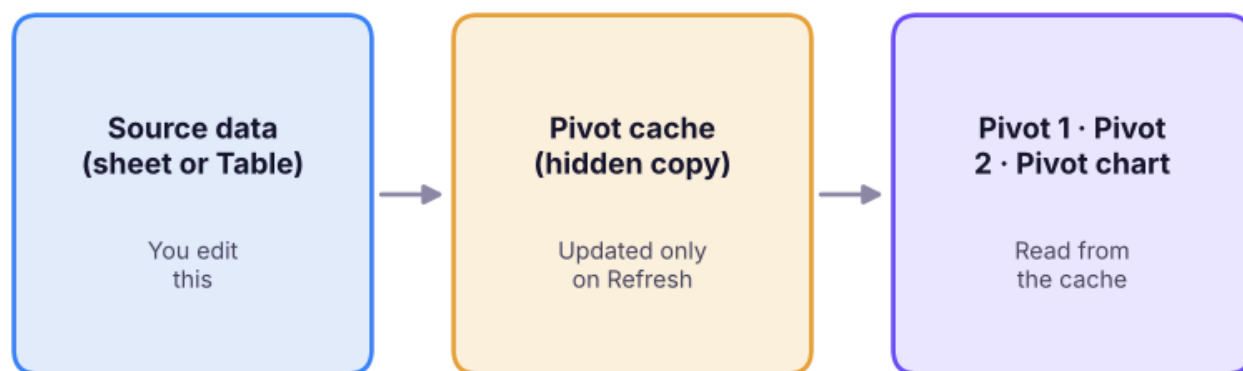


The PivotTable Fields pane: a list of fields at the top, four areas at the bottom.

Area	What it does	Example
Rows	One row per unique value	Region, Category, Order Date
Columns	One column per unique value	Year, Segment
Values	The numbers being summarised	Sum of Sales, Count of Order ID
Filters	Filter the whole pivot by a field	Segment = Corporate

The pivot cache, Refresh and Refresh All

A pivot table does not read your sheet directly. When you create it, Excel copies the data into a hidden **pivot cache**, and the pivot reads from that cache. This is why pivots are fast, and also why they do **not** update automatically when the data changes.



Every pivot reads from a cache. Pivots made from the same data share one cache.

Action	Shortcut	Effect
Refresh	Alt + F5	Re-reads the source into this pivot's cache. Every pivot that shares the same cache updates too
Refresh All	Ctrl + Alt + F5	Refreshes every pivot cache, query and data connection in the workbook
Change Data Source	PivotTable Analyze tab	Point the pivot to a bigger range (unnecessary if the source is a Table)
Refresh on open	PivotTable Options → Data	Refreshes automatically when the file opens

TIP Use a Table as the source

If the source is a Table, new rows are included the next time you refresh. If the source is a fixed range like `A1:R6324`, new rows below are ignored until you change the data source.

Value field settings

Click a field in Values → **Value Field Settings** (or right-click a number).

Summarize values by

Sum, Count, Average, Max, Min, Product, Count Numbers, StdDev, Var. Text fields in Values are counted automatically. **Distinct Count** appears only when the pivot is added to the **Data Model** (tick "Add this data to the Data Model" when creating it; Lesson 12).

Show values as

This tab turns a plain sum into an analytical view without writing formulas.

Show values as	Answers	Example
% of Grand Total	What share of the whole?	Furniture = 22% of all sales
% of Column Total	Share within each column	Within 2025, North = 31%
% of Row Total	Mix within each row	In the West, 46% of sales are Technology
% of Parent Row Total	Share within a parent group	Each sub-category's share of its category
Difference From (previous)	Change from the previous item	MoM or YoY change in ₹
% Difference From (previous)	Growth rate	MoM % or YoY %
Running Total In	Cumulative total	Year-to-date sales
Rank Largest to Smallest	Position	Top-selling region = 1
Index	Relative importance	Above 1 = over-represented

Here are three of these views on the Superstore data (computed with pandas to show the exact numbers a pivot would give):

Python

```
import pandas as pd
s = pd.read_csv("superstore_sales.csv", parse_dates=["order_date"])
pv = s.pivot_table(index="region", columns="category", values="sales", aggfunc="sum")
pct_total = (pv / pv.values.sum() * 100).round(1)
pct_row = (pv.div(pv.sum(axis=1), axis=0) * 100).round(1)
print("Sum of sales (₹):"); print(pv.round(0).astype(int).to_string())
print("\n% of grand total:"); print(pct_total.to_string())
print("\n% of row total (category mix inside each region):"); print(pct_row.to_string())
```

OUTPUT

Sum of sales (₹):

category	Furniture	Office Supplies	Technology
region			
East	7942092	2719513	35822514
North	7975868	2219542	38822750
South	9593302	2468888	41890579
West	9330590	2725360	39902570

% of grand total:

category	Furniture	Office Supplies	Technology
region			
East	3.9	1.4	17.8
North	4.0	1.1	19.3
South	4.8	1.2	20.8
West	4.6	1.4	19.8

% of row total (category mix inside each region):

category	Furniture	Office Supplies	Technology
region			
East	17.1	5.9	77.1
North	16.3	4.5	79.2
South	17.8	4.6	77.6
West	18.0	5.2	76.8

MoM and YoY

Month-on-month (MoM) compares each month with the previous month; **year-on-year (YoY)** compares a period with the same period last year (which removes seasonality, such as the festive Q4 peak).

STEP BY STEP MoM % in a pivot

1. Put `order_date` in Rows and group it by **Years** and **Months** (see Grouping below).
2. Add `sales` to Values **twice**.
3. On the second copy: Value Field Settings → Show Values As → % **Difference From** → Base field: `order_date` (Months) → Base item: **(previous)**.
4. Rename the field "MoM %".

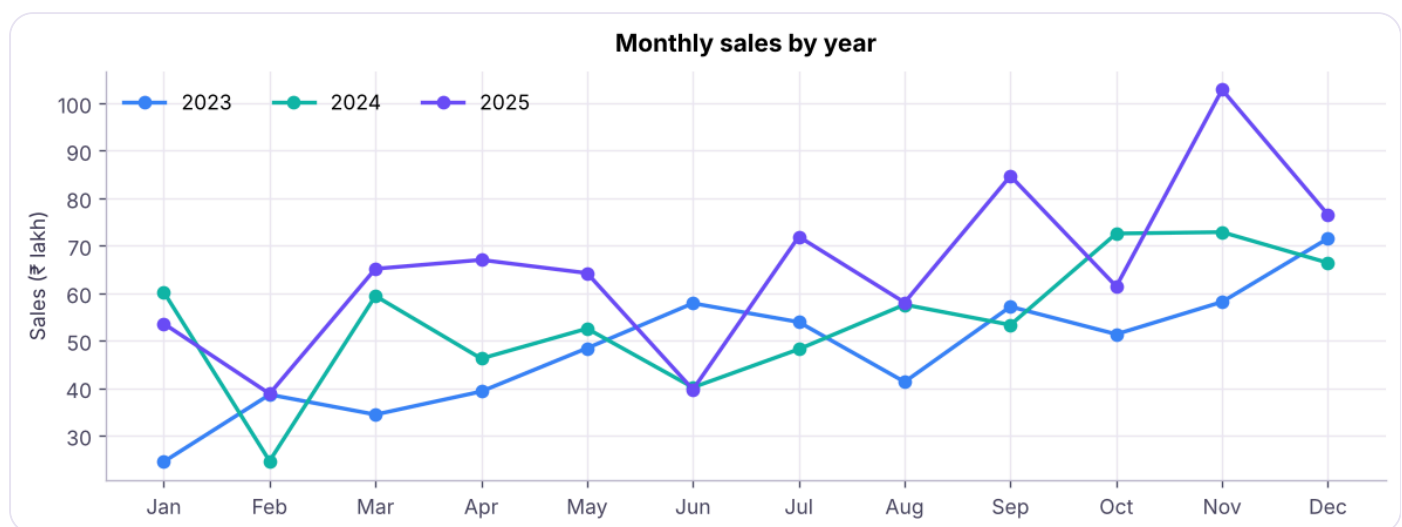
Python

```
m = s.set_index("order_date").resample("MS")["sales"].sum().to_frame("sales")
m["MoM %"] = (m.sales.pct_change() * 100).round(1)
m["YoY %"] = (m.sales.pct_change(12) * 100).round(1)
m.index = m.index.strftime("%b-%Y")
m.tail(8)
```

OUTPUT

	sales	MoM %	YoY %
order_date			
May-2025	6,421,274	-4.2	22.2
Jun-2025	3,968,572	-38.2	-1.1
Jul-2025	7,190,087	81.2	49.2
Aug-2025	5,802,184	-19.3	0.8
Sep-2025	8,468,060	45.9	58.9
Oct-2025	6,140,290	-27.5	-15.4
Nov-2025	10,284,516	67.5	41.2
Dec-2025	7,652,266	-25.6	15.3

8 rows × 3 columns



Monthly sales with YoY growth. The festive season (Oct-Dec) peaks every year, so YoY is a fairer comparison than MoM.

In formulas, the same calculations are: $\text{MoM \%} = (C3 - C2) / C2$, $\text{YoY \%} = (C14 - C2) / C2$ when months are listed one per row.

Grouping

Group dates

Right-click a date in Rows → **Group** → choose Years, Quarters, Months (and Days, Hours if needed). Excel 365 groups dates automatically when you drag a date field to Rows; you can ungroup with

right-click → Ungroup.

TIP

Indian financial year

Pivot grouping uses calendar quarters (Jan–Mar = Q1). For an April–March financial year, add helper columns to the data first: `FY = YEAR(date)+(MONTH(date) ≥ 4)` and `FQ = "Q"&ROUNDUP(MOD(MONTH(date)-4,12)+1)/3,0)`, or use CHOOSE (Lesson 7), and put those in the pivot.

Group items manually

Select several labels (hold Ctrl) → right-click → **Group**. Excel creates a new field such as Product2 with a group called Group1; rename it (for example "Low Sales Products"). This is how you combine the three lowest-selling products into one line, or merge small cities into "Others".

Groups vs bins

	Groups	Bins
What	Segregate items into chunks based on a common behaviour or criteria	Split a numeric field into equal-width intervals
Size	Can be any size	Every bin has the same width
Works on	Text, dates and numbers	Numbers
Example	"Metro" vs "Non-metro" cities	Order value 0–5,000, 5,000–10,000 ...

To bin numbers: put the number field (for example sales) in **Rows** → right-click → Group → Starting at 0, Ending at 50,000, By 5,000. Then put order_id in Values as a Count.

Python

```
bins = pd.cut(s.sales, range(0, 55000, 5000), right=False)
dist = s.groupby(bins, observed=False).size().to_frame("Count of order lines")
dist.index = [f"{int(i.left):,}-{int(i.right - 1):,}" for i in dist.index]
dist
```

OUTPUT

Count of order lines	
0–4,999	2530
5,000–9,999	829
10,000–14,999	473
15,000–19,999	423
20,000–24,999	238
25,000–29,999	172
30,000–34,999	177
35,000–39,999	202
40,000–44,999	61
45,000–49,999	92

10 rows × 1 columns

Most order lines are small; a few very large ones (machines, copiers) form the long tail. This is a frequency distribution, the pivot-table version of a histogram (Lesson 10).

Layout, subtotals and grand totals

The **Design** tab controls how the pivot looks:

Setting	Options	Tip
Report Layout	Compact (default), Outline, Tabular	Tabular + Repeat All Item Labels gives a flat table you can copy and reuse
Subtotals	Don't show, at bottom, at top	Turn off for exports
Grand Totals	On for rows/columns, off	Off when sharing a percentage view
Blank Rows	Insert after each item	Easier to read printed reports
Expand / Collapse	+/- buttons, or right-click → Expand/Collapse entire field	Drill from Category down to Sub-category
Show Details	Double-click any number	Opens a new sheet with the rows behind that number: great for auditing

Parent total: when two fields are in Rows (Category → Sub-category), "% of Parent Row Total" shows each sub-category's share of its own category.

Sorting and filtering inside a pivot

- **Sort:** right-click a value → Sort → Largest to Smallest.
- **Label filter:** Row Labels drop-down → Label Filters → Begins with, contains ...
- **Value filter / Top 10:** Row Labels drop-down → Value Filters → **Top 10** → change to Top 5 Items by Sum of sales. This is how you answer "Find the top 5 categories".
- **Slicers:** PivotTable Analyze → **Insert Slicer** → pick fields (Region, Segment). Slicers are clickable buttons that filter the pivot and show what is selected. Right-click a slicer → **Report Connections** to control several pivots with one slicer (they must share the same cache).
- **Timeline:** Insert Timeline on a date field to filter by years, quarters, months or days with a slider.

Calculated fields

A **calculated field** creates a new value field from other fields in the pivot: PivotTable Analyze → Fields, Items & Sets → **Calculated Field**.

Example: Profit Margin = profit / sales, formatted as a percentage.

WATCH OUT Calculated fields work on totals

Excel calculates the field on the **sums**: Profit Margin for the North = total North profit ÷ total North sales. That is correct for ratios. But a calculated field like =quantity * price would multiply the *sum* of quantity by the *sum* of price, which is wrong. Add such row-level calculations as a new column in the source data instead.

Python

```
pm = s.groupby("category")[["sales", "profit"]].sum()
pm["Profit Margin"] = (pm.profit / pm.sales * 100).round(1).astype(str) + "%"
pm.round(0)
```

OUTPUT

	sales	profit	Profit Margin
category			
Furniture	34,841,853	-818,153	-2.3%
Office Supplies	10,133,304	1,514,794	14.9%
Technology	156,438,412	17,655,524	11.3%

3 rows × 3 columns

GETPIVOTDATA and pivot charts

When you type = and click a cell inside a pivot, Excel writes GETPIVOTDATA, for example =GETPIVOTDATA("sales", \$A\$3, "region", "North"). It keeps returning North's sales even if the pivot layout changes. Turn it off in PivotTable Analyze → Options → Generate GetPivotData if you prefer normal references.

PivotChart (PivotTable Analyze → PivotChart) creates a chart linked to the pivot: filtering the pivot or clicking a slicer updates the chart. This is the basis of the dashboard in Lesson 11.

Case study: Video Game Sales (regional sales trends)

You are an analyst at a game publisher. The file `video_game_sales.csv` has 2,400 fictional titles with platform, release year, genre, publisher and sales in millions of units by region: North America (`na_sales`), Europe (`eu_sales`), Japan (`jp_sales`), other regions and global. Answer the questions with pivot tables. The answers below were computed from the file so you can check your pivots.

Python

```
g = pd.read_csv("video_game_sales.csv")
print(g.shape); g.head(4)
```

OUTPUT

```
(2400, 11)
```

	rank	name	platform	year	genre	publisher	na_sales	eu_sales	jp_sales
0	1	Hidden Kingdom	XSX	2023	Sports	Red Lantern Studios	3.69	2.62	0.29
1	2	Galactic Kingdom	PS5	2020	Adventure	Saffron Works	3.49	1.79	0.74
2	3	Galactic League	Mobile	2017	Strategy	Red Lantern Studios	3.08	2.44	0.51
3	4	Shadow League	Mobile	2025	Role-Playing	Northwind Play	1.66	2.12	2.31

4 rows × 11 columns

CASE STUDY Q1. Which genres sell the most worldwide?

Pivot: Rows = genre, Values = Sum of global_sales, sorted largest to smallest.

Python

```
q1 = g.pivot_table(index="genre", values="global_sales",
aggfunc="sum").sort_values("global_sales", ascending=False)
q1["% of total"] = (q1.global_sales / q1.global_sales.sum() * 100).round(1)
q1.round(1).head(6)
```

OUTPUT

	global_sales	% of total
genre		
Adventure	136.7	10.6
Sports	133.8	10.4
Action	116.3	9.0
Racing	113.6	8.8
Strategy	111.5	8.6
Simulation	108.4	8.4

6 rows × 2 columns

CASE STUDY Q2. Who are the top 5 publishers, and how many titles did each release?

Pivot: Rows = publisher, Values = Sum of global_sales and Count of name; Value Filter → Top 5 by global sales.

Python

```
q2 = g.groupby("publisher").agg(global_sales=("global_sales", "sum"), titles=("name",
"count")).sort_values("global_sales", ascending=False).head(5)
q2["sales per title"] = (q2.global_sales / q2.titles).round(2)
q2.round(1)
```

OUTPUT

	global_sales	titles	sales per title
publisher			
Starforge Interactive	168.0	295	0.6
Blue Comet Games	161.4	316	0.5
Northwind Play	144.6	266	0.5
Red Lantern Studios	144.5	268	0.5
Pixel Harbor	143.1	267	0.5

5 rows × 3 columns

CASE STUDY Q3. How does regional preference differ by genre?

Pivot: Rows = genre, Values = Sum of na_sales, eu_sales, jp_sales, other_sales; Show Values As → % of Row Total.

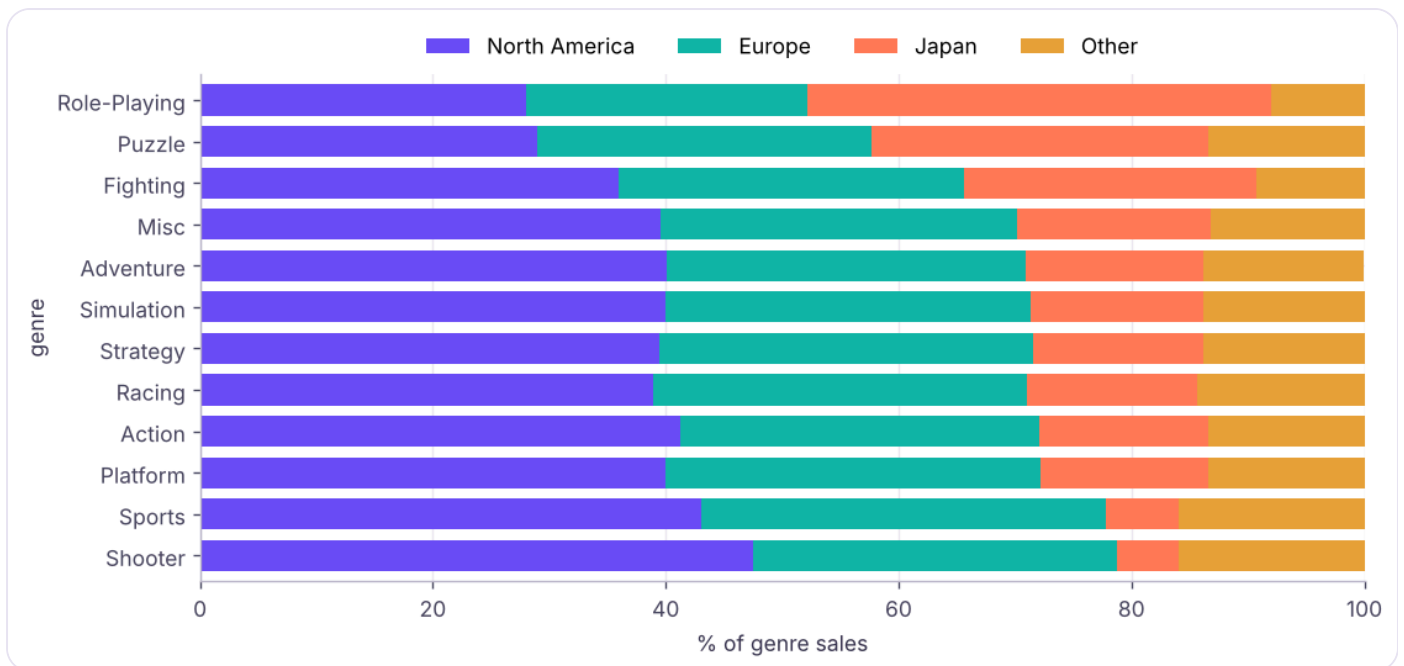
Python

```
reg = ["na_sales", "eu_sales", "jp_sales", "other_sales"]
q3 = g.groupby("genre")[reg].sum()
q3 = (q3.div(q3.sum(axis=1), axis=0) * 100).round(1).sort_values("jp_sales",
ascending=False)
q3.head(6)
```

OUTPUT

	na_sales	eu_sales	jp_sales	other_sales
genre				
Role-Playing	28.0	24.2	39.8	8.1
Puzzle	29.0	28.7	28.9	13.4
Fighting	36.0	29.6	25.1	9.3
Misc	39.6	30.6	16.6	13.2
Adventure	40.1	30.8	15.3	13.7
Simulation	40.0	31.3	14.9	13.8

6 rows × 4 columns



Regional mix by genre (% of each genre's sales). Role-Playing and Puzzle games lean much more on Japan than Shooter or Sports titles do.

CASE STUDY Q4. How have sales changed over the years?

Pivot: Rows = year, Values = Sum of global_sales, plus a second copy shown as % Difference From (previous) for YoY growth. Insert a PivotChart (line).

Python

```
q4 = g.pivot_table(index="year", values="global_sales", aggfunc="sum")
q4["YoY %"] = (q4.global_sales.pct_change() * 100).round(1)
q4.round(1).tail(8)
```

OUTPUT

	global_sales	YoY %
year		
2018	87.2	19.0
2019	59.4	-31.9
2020	89.2	50.2
2021	78.0	-12.6
2022	71.9	-7.9
2023	76.4	6.3
2024	80.9	5.9
2025	87.0	7.4

8 rows × 2 columns

CASE STUDY Q5. Which platform leads in each region?

Pivot: Rows = platform, Values = the four regional sales columns; sort each column, or use conditional formatting → Top 1 per column.

Python

```
q5 = g.groupby("platform")[reg].sum().round(1)
q5.loc["Leader"] = q5.idxmax()
q5
```

OUTPUT

	na_sales	eu_sales	jp_sales	other_sales
platform				
3DS	45.90	40.40	37.60	17.20
Mobile	63.30	50	22.60	21.10
PC	43.40	34.20	14.90	14.90
PS3	48.10	38.50	18.80	17
PS4	55.40	43.10	16.90	19.20
PS5	58.50	45.40	20.80	19.70
Switch	40	34.40	38	13
X360	42.20	32.60	15.20	14.60
XOne	50.10	39.20	17.70	17.40
XSX	53.80	40.40	18	17.20
Leader	Mobile	Mobile	Switch	Mobile

11 rows × 4 columns

KEY TAKEAWAYS Insights to present

Write 3–5 sentences of insight, not just tables. For example: which genres to prioritise for the Japanese market, which publishers are efficient (high sales per title), and how platform transitions show in the year trend. Keep every number traceable to a pivot.

Summary

KEY TAKEAWAYS

- Pivot tables summarise structured data by dragging fields into **Rows, Columns, Values and Filters**.
- They read from a **pivot cache**: Refresh updates one cache; Refresh All updates everything. Use a Table as the source.
- Value Field Settings: change the function (Sum, Count, Average...) and **Show Values As** (% of total, % of row, difference from previous, running total, rank).
- MoM = change from the previous month; YoY = change from the same period last year.
- Group dates (years/quarters/months), group items manually, and group numbers into equal **bins**.
- Calculated fields work on totals, so use them for ratios like margin; add row-level maths to the source data.
- Slicers and timelines make pivots interactive; double-click a number to see the rows behind it.

Quick quiz

6 questions · tap an answer to check it

1. You add 500 rows to the source range, but the pivot does not change. Why?

- A** Pivot tables only read the first 1,000 rows
- B** The pivot reads from its cache, which updates only when you refresh (and the range must include the new rows)
- C** The new rows need a pivot formula
- D** Pivot tables cannot be updated

Answer: B. Refresh re-reads the source. If the source is a fixed range, also change the data source, or use a Table.

2. Which "Show Values As" option gives month-on-month growth?

- A % of Grand Total
- B Running Total In
- C % Difference From → (previous)
- D Rank Largest to Smallest

Answer: C. % Difference From the previous item compares each month with the one before it.

3. What is the main difference between groups and bins?

- A Bins are equal-width numeric intervals; groups can be any size and any data type
- B Groups only work with dates
- C Bins only work with text
- D There is no difference

Answer: A. Bins split numbers into equal intervals. Groups are custom collections of items.

4. A calculated field Revenue = Qty * Price gives inflated numbers. Why?

- A Calculated fields cannot multiply
- B It multiplies the sum of Qty by the sum of Price for each group
- C The pivot needs refreshing
- D Price must be in Rows

Answer: B. Calculated fields work on aggregated values. Compute Qty × Price in the source data instead.

5. How do you see the individual rows behind a number in a pivot table?

A Press F9

B Double-click the number (Show Details)

C Use GETPIVOTDATA

D Insert a slicer

Answer: B. Double-clicking opens a new sheet with the underlying records.

6. Two pivots were built from the same data. You click Refresh on one. What happens to the other?

A It also updates, because they share the same pivot cache

B Nothing; you must refresh it separately

C It is deleted

D It shows an error

Answer: A. Pivots from the same source share a cache; refreshing the cache updates all of them.

EXERCISE Superstore pivots

Build pivot tables from `superstore_sales.csv` to answer: (1) sales and profit by category and sub-category, with profit margin; (2) yearly sales by region with YoY %; (3) the 5 cities with the highest profit; (4) the number of order lines by discount (0%, 10%, ... 40%) and their total profit.

Show solution (worked answer)

(1) Rows: category, sub_category; Values: Sum of sales, Sum of profit; Calculated Field Margin = profit/sales formatted as %. (2) Rows: region; Columns: order_date grouped by Years; Values: Sum of sales, plus a second copy shown as % Difference From (previous) Years. (3) Rows: city; Values: Sum of profit; Value Filters → Top 10 → Top 5 Items by Sum of profit. (4) Rows: discount; Values: Count of order_id and Sum of profit. You will see profit turn negative at the higher discount levels, matching the Lesson 5 crosstab.

INTERVIEW CORNER Questions you may be asked

Q: What is a pivot cache and why does it matter? A hidden copy of the source data that pivots read from. It makes pivots fast and lets several pivots share data, but pivots only update on refresh, and many separate caches make files large.

Q: How do you show each item's share of the total in a pivot? Value Field Settings → Show Values As → % of Grand Total (or % of Column/Row Total).

Q: How would you count unique customers in a pivot? Create the pivot with "Add this data to the Data Model" ticked, then choose Distinct Count in Value Field Settings.

Q: When would you use formulas (SUMIFS) instead of a pivot? When the report needs a fixed layout, mixes data from several sources, or must update without refreshing; pivots are better for exploration and flexible summaries.

Practise more online

Free, trusted places to test yourself on this topic. Links open the original site in a new tab.