

Data visualization and storytelling

This session teaches students to turn market data into a figure that tells the truth, and to recognize when a figure does not. Auction and resale pricing data can be summarized many ways:

What today looks like

- Length: 90 minutes. A 60 minute and a 120 minute variant are noted under Teaching notes.
- Level: undergraduate or graduate students in a data visualization, business communication, or data journalism course. No finance background or prior charting tool is assumed.
- Access: the no-code sandbox at sandbox.altfndata.com is the primary path, and is all that is needed for the core session. No API key is required. An optional API and Python extension, for exporting a query result and reproducing a chart in a plotting library, requires an instructor class API key from info@altfndata.com.
- Goal: students leave able to build a pricing power bar chart and a sold-lot-count line chart in the sandbox, name at least two techniques that make an accurate chart misleading, and state the recency caveat as a chart-honesty obligation, not just a data footnote.

OBJECTIVES

By the end of class students can

- Select an appropriate chart type for a given market data question, distinguishing when a bar chart, a line chart, or a table best serves the reader.
- Build a pricing power comparison chart and a sold-lot-count time series using the sandbox's pre-built charts and export tools, without writing code.
- Identify axis truncation and cherry-picked time windows as two common techniques that make a chart misleading without altering the underlying numbers.
- Explain why a time series ending in the most recent quarter requires a caveat about ongoing data ingestion, and revise a chart's title or footnote to state that caveat plainly.
- Critique a chart built by a classmate against a short honesty checklist covering axis choice, time window, and labeling.
- (Extension) Export a query result via the production API and reproduce one of the two chart types in a Python plotting library.

How the session runs

- 0 to 15 min: Introduce the session's premise, that the same numbers can produce an honest or a dishonest chart, and preview the two figures the class will build.
- 15 to 25 min: Sandbox orientation. Confirm students can open a data table, run a query, and reach the pre-built charts and export tabs.
- 25 to 45 min: Guided demo, build a pricing power comparison bar chart across a handful of brands.
- 45 to 60 min: Guided demo, build a sold-lot-count line chart by quarter for a single category, and walk through the recency caveat as a visualization problem, not just a data problem.
- 60 to 75 min: Small-group exercise, each group builds one chart, then deliberately produces a misleading version of the same chart using truncation or a cherry-picked window, and prepares to explain the trick.
- 75 to 85 min: Class discussion, groups present their honest and misleading pairs and the class applies the honesty checklist.
- 85 to 90 min: Wrap-up and homework assignment.

The in-class walkthrough

- Open sandbox.altfndata.com, sign in, and select the handbags data table (or another category the instructor prefers) from the SQL editor dropdown.
- Open the coverage browser tab and confirm how two or three well-known brands are written in the designer field before filtering.
- Return to the SQL editor and run the first guided query below, which computes pricing power for a small set of brands.
- Open the pre-built charts tab and render the result as a bar chart. Ask students what chart type would have served the same data poorly, for example a line chart implying a trend across brands that have no natural order.
- Run the second guided query, sold-lot count by quarter for a single category, and render it as a line chart.
- Point to the final one or two quarters on the chart and ask students what they notice about the shape of the line there. Introduce the recency caveat directly: the newest quarters are still being ingested, so a downturn at the very end of the line is very likely an ingestion artifact, not a demand signal, and a chart that omits this caveat misleads its reader even though every number in it is correct.

The core query

- Run this in the sandbox SQL editor:
 - `SELECT designer,`
 - `approx_percentile(usd_price_decimal / sale_estimates_high_usd_price, 0.5) AS pricing_power,`
 - `COUNT(*) AS sold_lots`
 - `FROM all_handbags_data`
 - `WHERE status = 'sold'`
 - `AND sale_estimates_high_usd_price > 0`
 - `AND designer IN ('Hermès', 'Chanel', 'Louis Vuitton')`
 - `GROUP BY designer`

DISCUSS

Questions for the room

- What makes a bar chart the right choice for comparing pricing power across brands, and what would be lost or distorted by showing the same comparison as a line chart?
- When does truncating a y-axis serve the reader by making a real difference visible, and when does it cross into misleading exaggeration? Where is the line?
- A line chart of sold-lot count dips in its final quarter because of ingestion lag, not falling demand. What obligation does the chart's creator have to the reader in that situation?
- How would you design a caveat or footnote so that a reader who only glances at the chart still absorbs the recency warning?
- What is the difference between choosing a favorable time window because it is genuinely the most relevant period, and cherry-picking a window because it tells a better story?

This week's assignment

- Each student selects one category table and one comparison question (for example, pricing power across three brands, or sell-through across two auction houses, or sold-lot count over time for a single brand), and produces two charts from the same underlying query:
- one built to inform as honestly as possible, and one built to mislead using a specific, named technique such as axis truncation or a cherry-picked window.
- The submission is a short memo, no more than two pages, that includes both charts, the SQL query used, and a paragraph explaining exactly how the misleading version distorts the reader's takeaway and what an honest caption or footnote would need to say to correct it, including the recency caveat if the chart involves a time series.
- Grading criteria:
- correct chart-type choice for the honest version (25 percent), a clearly identifiable and well-explained distortion technique in the misleading version (30 percent), quality of the written critique connecting the two (30 percent), and clarity and professionalism of the memo (15 percent).

NEXT

Wrapping up

- Restate the one-sentence takeaway before students leave.
- Point to the facilitator guide for discussion guidance and teaching notes.
- Questions or a class API key: info@altfndata.com