

Study sheet: Module 8, 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: a pricing power ratio, a sell-through rate, a count of sold lots by quarter, and each of those summaries can be charted honestly or dishonestly with the same underlying numbers. Students use the sandbox's pre-built charts and export tools to build two figures, a bar chart comparing pricing power across brands and a line chart of sold-lot counts over time, and then practice the editorial judgment that separates a chart that informs from one that misleads: choosing the right chart type for the question, deciding where an axis should start, and deciding which window of time is fair to show. The session closes on a caution that runs through the rest of the course, that the newest quarters in this dataset are still being ingested, and a line chart that lets that under-ingestion masquerade as a market decline is a dishonest chart regardless of how the axis is drawn.

What you should be able to do

1. 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.
2. 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.
3. Identify axis truncation and cherry-picked time windows as two common techniques that make a chart misleading without altering the underlying numbers.
4. 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.
5. Critique a chart built by a classmate against a short honesty checklist covering axis choice, time window, and labeling.
6. (Extension) Export a query result via the production API and reproduce one of the two chart types in a Python plotting library.

Key terms

- **Pricing power:** the median of realized price over high estimate for sold lots, a demand signal above or below 1.0.
- **Sell-through:** the share of offered lots, sold and unsold combined, that found a buyer.
- **Axis truncation:** starting a numeric axis above zero to exaggerate the apparent size of a difference.
- **Cherry-picked window:** selecting a start and end date for a time series specifically because it supports a preferred conclusion.
- **Recency under-ingestion:** the newest quarters in the dataset have fewer records because ingestion is still catching up, not because activity fell.
- **Data-ink ratio:** the design principle that a chart should devote as much of its visual weight as possible to the data itself, not decoration.
- **Small multiples:** a series of similar small charts, one per category, used to compare many groups without cluttering a single chart.

- **Caption honesty:** the practice of writing a chart's title or footnote so that a reader who reads only that text still reaches an accurate conclusion.

Core sandbox query

The query the module is built around. Run it, then change the brand or category and compare.

```
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  
ORDER BY pricing_power DESC;
```

Full lesson, more queries, and the homework are in Module 8 at academy.altfndata.com. Questions: info@altfndata.com.