

Alternative data in finance

This session introduces alternative data as an asset class and situates ALT/FNDATA's auction and resale pricing dataset within it. Students move from the abstract question of what counts as alternative data to a concrete, hands-on look at a real institutional dataset:

What today looks like

- Length: 75 minutes. A 50 minute and a 90 minute variant are noted under Teaching notes.
- Level: undergraduate juniors and seniors, MBA, or graduate finance. No data or coding background required.
- Access: the no-code sandbox at sandbox.altfndata.com only. No API key is needed. Students self-register with a work or school email and are approved automatically.
- Goal of the session: students leave able to define alternative data, name where auction and resale pricing fits, and run their own first query in the sandbox.

OBJECTIVES

By the end of class students can

- Define alternative data and distinguish it from traditional market and fundamental data sources.
- Describe at least three categories of alternative data used by institutional investors, with auction and resale pricing data as one example.
- Explain what a physical luxury asset transaction record contains and why point-in-time correctness (using the actual sale date) matters for research integrity.
- Query a live dataset in a no-code SQL environment to answer a simple pricing question.
- Articulate one hypothesis for how alternative pricing data could complement or contradict a signal from public equity markets.

How the session runs

- 0 to 10 min: What is alternative data. Instructor-led discussion of traditional data (filings, price feeds, analyst estimates) versus alternative data (satellite imagery, web traffic, credit card panels, and transactional pricing data).
- 10 to 20 min: Where auction and resale data fits. Introduce ALT/FNDATA's coverage: watches, jewelry and gems, handbags, fine art, works of art, wine and whisky, automobiles, motorcycles, aircraft, design, coins, and books, with pre-sale estimates, realized prices, and vendor and brand detail on every record.
- 20 to 30 min: Sandbox orientation. Students register and open the sandbox together; instructor walks through the SQL editor, sample datasets, coverage browser, and data dictionary tabs.
- 30 to 55 min: Guided demo. Instructor runs the in-class demo below live, narrating each query and result.
- 55 to 70 min: Student exploration. In pairs, students run one of the demo queries against a different brand or category and note what changes.
- 70 to 75 min: Wrap-up and homework assignment.

The in-class walkthrough

- Open sandbox.altfndata.com and sign in with your registered account.
- Click the coverage browser tab. Type "watches" into the category filter and note the number of brands and vendors covered.
- Click the data dictionary tab and locate the fields designer, model, sale_date, usd_price_decimal, and sale_estimates_high_usd_price. Read each field's description aloud to the class.
- Open the SQL editor tab. Select the watches dataset from the table dropdown.
- Paste the first query from the "Datasets and queries used" section below into the editor and click Run.
- Point out the columns returned: designer, item_title, sale_date, usd_price_decimal. Ask students to identify the highest and lowest realized prices in the result set.

The core query

- Run this in the sandbox SQL editor:
 - `SELECT designer, item_title, sale_date, usd_price_decimal`
 - `FROM all_watches_data`
 - `WHERE designer LIKE '%Rolex%'`
 - `AND status = 'sold'`
 - `ORDER BY usd_price_decimal DESC`
 - `LIMIT 25;`

DISCUSS

Questions for the room

- What distinguishes alternative data from traditional financial data, and why has demand for it grown among institutional investors?
- What are the risks of relying on alternative data that is unevenly collected or has gaps in coverage across time or geography?
- Why does the `sale_date` field matter for point-in-time correctness, and what would go wrong in a research process that used the wrong date?
- Auction and resale pricing data reflects what buyers actually paid for physical goods. What does this tell you that a survey of stated brand preference would not?
- Every record in this dataset carries a status of sold or unsold. Why is it important to capture unsold lots rather than only successful sales?

This week's assignment

- Each student picks one brand covered in the watches, jewelry, or handbags data and writes a two-page memo answering:
- how many sold lots for that brand are in the dataset, what is the highest and lowest realized price observed, and what does the spread between realized prices and high estimates suggest about demand for that brand.
- The memo must include at least one sandbox SQL query (pasted as an appendix) and one exported chart or table as a screenshot.
- Deliverables:
- a PDF memo of no more than two pages, the appended query, and the exported visual.

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