ALT/FNDATA ACADEMY · STUDENT STUDY SHEET

# Study sheet: Module 15, art market fundamentals

This session asks a single plain question and answers it with real transaction data: what forms the price of a work of art. Rather than treating price as a mystery or a matter of taste alone, students learn to read it as the outcome of several factors working together, the artist, the medium and format, the size, the period, the rarity, the provenance, the condition, and the timing of the sale. The dataset records what buyers actually paid at more than 100 auction houses with history back to the late 1990s, so students can see the secondary market as it behaves, where a pre-sale estimate meets a real buyer, and where two works by the same artist can sell for very different sums. No finance background and no coding are assumed. The session builds toward a comparable set, the tool every appraiser and dealer uses to judge what a given work should be worth, built here by hand from real sold lots.

## What you should be able to do

1. Name the factors that form an artwork's price, including artist, medium and format, size, period, rarity, provenance, condition, and market timing, and explain how each one moves a price up or down.
2. Distinguish the primary market, where a work first sells through a gallery or dealer, from the secondary market, where a work resells at auction, and state which market this dataset captures.
3. Define pre-sale estimate as the auction house's formalized judgment of comparable value and realized price as what a buyer actually paid, and explain how the two relate.
4. Build a comparables set for a single artist from sold lots, and read it for its typical price, its range, and its outliers.
5. Describe price dispersion within one artist's market, using the minimum, median, and maximum realized price, and explain why two works by the same artist can sell for very different amounts.
6. Identify the limits of auction data for art market fundamentals, including that it captures the secondary market only, that documented fields do not include medium, size, provenance, or condition directly, and that the newest periods are still being ingested and should not be read as a trend.

## Key terms

- **Comparables:** a set of similar sold works used to judge what a given work should be worth, the core tool behind this session's exercises.
- **Realized price:** what a buyer actually paid for a lot, recorded in usd_price_decimal, as distinct from any pre-sale estimate.
- **Pre-sale estimate:** the auction house's formalized judgment of comparable value before a sale, recorded here as sale_estimates_high_usd_price.
- **Primary market:** the market for a work's first sale, typically through a gallery or dealer, not captured in this dataset.
- **Secondary market:** the resale market captured by auction data, where a work changes hands after its first sale.
- **Price dispersion:** the spread between the lowest and highest realized prices within a single artist's market, measured here with minimum, median, and maximum.

- **Attribution:** the identification of a work as by a particular artist, a factor that can swing price sharply when it is uncertain or contested.
- **Provenance:** a work's ownership history, a factor known to move price but not present as a field in this data.
- **Condition:** the physical state of a work, another price-forming factor not present as a field in this data.

## Core sandbox query

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

```
SELECT designer AS artist,
       item_title,
       vendor,
       sale_date,
       usd_price_decimal AS realized_price_usd
FROM all_fine_art_data
WHERE status = 'sold'
  AND designer LIKE '%Warhol%'
ORDER BY usd_price_decimal DESC
LIMIT 50;
```

## Common mistakes to avoid

1. Reading the average realized price for an artist rather than the median, letting a small number of very high sales distort the typical figure.
2. Treating a text search on item_title, such as oil on canvas, as a reliable medium field rather than an imperfect keyword proxy that will miss variant phrasing and catch false matches.
3. Building a comparables set on too few sold lots, producing a range too noisy to support a real judgment about a specific work.
4. Reading the newest quarter's sales as evidence that an artist's market is rising or falling, when the newest periods are still being ingested.
5. Forgetting that this dataset is secondary market only, and drawing a conclusion about an artist's overall market strength without accounting for gallery and dealer activity the data cannot see.
6. Explaining a price outlier with a single factor, such as size alone, when attribution, provenance, condition, and timing are usually acting together.

---

Full lesson, more queries, and the homework are in Module 15 at academy.altfndata.com. Questions: info@altfndata.com.