

## Problem set 14: The auction business model

---

**Module 14.** Total: 100 points. Format: no-code sandbox, with an optional API extension in Problem 4 using a shared class API key from your instructor.

### Before you start

This problem set works with the fine art data (documented fields used here: `designer`, `model`, `item_title`, `sale_date`, `usd_price_decimal`, `sale_estimates_high_usd_price`, `status`, `vendor`, `stock_ticker`). In that table the `designer` field holds the artist or maker name and the `vendor` field holds the auction house. You will build the two measures a business analyst can actually compute from public auction records: how well a house's pre-sale estimates track what buyers pay, and how concentrated or spread out a house's revenue is across its lots. Choose one major auction house to use throughout, and confirm the exact spelling in the coverage browser before you write any filter, since exact-match filters are spelling-sensitive. Remember that no field in this table records buyer's premium, hammer price, seller's commission, or guarantee terms; `usd_price_decimal` is the realized price the buyer paid, nothing more.

---

### Problem 1 (15 points): The two revenue streams and why estimate accuracy matters

In four to five sentences, explain the buyer's premium and the seller's commission as the two revenue streams that fund an auction house, and state plainly that neither figure is a field in the data. Then explain why estimate accuracy, the ratio of realized price to high estimate, is the computable business signal an analyst can use instead, and why it matters commercially to a house beyond making its specialists look good on sale night.

### Problem 2 (20 points): Estimate accuracy for your chosen house

Write a SQL query against the fine art data that computes estimate accuracy for your chosen house: the median (or another percentile) of `usd_price_decimal` divided by `sale_estimates_high_usd_price`, restricted to sold records, with a guard against zero or null high estimates, alongside a count of the sold lots included. State in one sentence why the guard is required.

```
-- your query here
```

### Problem 3 (20 points): Average lot value versus lot volume, and sell-through, by house

Write SQL that computes, across houses, both average lot value (total realized value divided by count of sold lots, restricted to sold records and grouped by vendor) and sell-through rate (the share of offered lots that sold, with both sold and unsold in the denominator, grouped by vendor). You may use one query or two; state which figure each query produces, and in two to three sentences explain what it would mean for one house to show a high average lot value alongside a low sell-through rate.

```
-- your query (or queries) here
```

### Problem 4 (20 points): One house via the API (optional code extension)

Using the shared class API key your instructor has provided, write the endpoint and JSON request body you would send to `POST /v1/tables/{name}/query` to retrieve the raw fields needed to compute your chosen house's average lot value or estimate accuracy (`vendor`, `designer`, `item_title`, `sale_date`, `usd_price_decimal`, `sale_estimates_high_usd_price`), filtered to that house and to sold records. Then, in two to three sentences,

explain why the query endpoint cannot return the averaged or ratio figure directly, and how you would paginate on offset and aggregate the rows yourself in the tutorials notebook or with the Python client.

```
POST /v1/tables/<table_name>/query
Header: X-API-Key: <class key>
Body:
{
  ...
}
```

### **Problem 5 (25 points): Value concentration and a saleroom business brief**

Write a SQL query that computes value concentration among your chosen house's top lots: the share of the house's total realized value held by its top 20 sold lots by `usd_price_decimal`, restricted to sold records for that house. Then write a short business brief, 200 to 300 words, in the voice of an analyst preparing a competitive read on the house for a rival saleroom's management. The brief must report the house's estimate accuracy, its average lot value against its sell-through, and its value concentration among its own top lots, and must argue, using those figures together, what kind of auction business the house runs and where its consignment strategy appears to be working or under strain. The brief must include an explicit limitations section stating that the data captures the secondary market only, that no field records buyer's premium, hammer price, commission, or guarantees so none of those can be computed, and that the newest periods are still being ingested, so the brief must not read any recent-period total as evidence of a trend.

```
-- your query here
```

---

**Submission.** Turn in this file with your SQL, JSON body, and written answers filled in.