

Problem set 5: Library data literacy workshop

Module 5. Total: 100 points. Format: no-code sandbox, written for a general audience with no assumed background in data or finance.

Before you start

This workshop uses the sandbox at sandbox.altfndata.com, a free, browser-based tool. You do not need to install any software, and you do not need any programming experience. You will register with an email address (a work, school, or personal email is fine for a library workshop setting) and your access is approved automatically, so you can start right away.

Problem 1 (10 points): Getting oriented

ALT/FNDATA has three ways to access its data: a sandbox, a documentation site, and a production API for paying business customers. In two to three sentences, explain in plain language which of the three you are using today, and why "no installation required" and "auto-approved" matter for someone using a library or classroom computer.

Problem 2 (15 points): Reading a real record

Open the coverage browser and pick one category that interests you (for example watches, handbags, or fine art). Look at one sample record and, in plain sentences (no jargon), describe what it tells you using each of these pieces of information: `item_title`, `designer`, `sale_date`, `status`, and `usd_price_decimal`. Write your answer as a short paragraph, as if you were explaining the record to a friend who has never seen this kind of data before.

Problem 3 (20 points): Asking the data a simple question

Using the sandbox's SQL editor or a pre-built chart, ask the data one simple question of your own choosing, for example "what are the 10 most expensive sold items from one designer" or "how many items from one auction house sold last year." Write down the question in plain English first, then write the query (or describe the chart settings you used) and report what you found in one or two sentences.

-- your query here, if you used the SQL editor

Problem 4 (20 points): Reading the data critically

Every record has a status of sold or unsold, and every record is tied to the date it actually sold, not the date ALT/FNDATA added it to the dataset. In four to five sentences, written in plain language, explain two things a careful reader should keep in mind before trusting a number pulled from this data: first, why unsold items should not be counted the same way as sold items when looking at prices, and second, why the very newest information in the dataset (the last few months) might look incomplete even though it is not wrong, simply not finished catching up yet.

Problem 5 (15 points): A simple ratio, by hand

Imagine an auction house expected an item to sell for 10,000 dollars at most (its high estimate), and it actually sold for 13,600 dollars. Divide the sale price by the estimate to get a ratio. Is the ratio above or below 1.0, and what does that tell you about how badly people wanted this particular item? Then, in one sentence, explain in

plain language what it would mean if, across hundreds of sales for one brand, the typical (median) version of this same ratio came out below 1.0 instead.

Problem 6 (20 points): Why this data might matter to you

In 150 to 200 words, written in plain language for a general audience, describe one real situation, for yourself, a family member, a small business, or a local reporter, where being able to look up real auction and resale prices for luxury goods might be useful. You do not need financial background to answer this; focus on a concrete, everyday use.

Submission. Turn in this file with your written answers filled in. There is no penalty for choosing simple, everyday examples over technical ones in this problem set.