

Data Analyst interview questions

careertips.com.au: common questions and what they're really testing

What to expect

Data analyst interviews usually mix a technical check on SQL and tooling with behavioural questions about how you've worked with stakeholders, plus at least one scenario question that tests judgement when data is messy or incomplete. Because the role sits close to decision-makers, expect at least one question about explaining a technical result to a non-technical audience.

- **Technical:** Direct questions or a short exercise on SQL, Python or BI tool logic, sometimes a live query or take-home case study.
- **Behavioural:** Past experience questions about scoping vague requests, handling stakeholder disagreement, or fixing a reporting error.
- **Scenario:** A hypothetical data problem, such as a number not matching between systems, used to see how you'd triage under time pressure.
- **Client-facing:** Questions about presenting findings to people without a data background, and adjusting detail to the audience.

Most interviews open with background and a walk-through of your resume, move into a technical segment (verbal SQL questions or a shared-screen exercise), then behavioural and scenario questions, and close with your questions for the panel. A second-round interview may include a take-home dataset or a presentation of a prior piece of analysis.

1. Walk me through how you'd clean and validate a dataset before building a dashboard from it.

Why they ask

This checks whether you have a repeatable process rather than ad hoc habits, and whether you think about data quality before visual polish.

How to structure your answer

Answer as a step-by-step walk-through: what you check first, what tools you use at each step, and what you do when you find a problem.

Example answer

"I start by checking the shape of the data: row counts, obvious duplicates, and whether key fields like dates or IDs are the type they should be. In SQL I'll run quick checks for nulls and outliers before I trust any join. If I find something that looks wrong, I go back to the source system owner rather than guessing at a fix, because a silent assumption in the cleaning step ends up wrong in every dashboard downstream. Only once I'm confident in the base table do I move into Power BI to build the actual visuals."

2. Tell me about a time a stakeholder asked for something that wasn't actually what they needed.

Why they ask

Translating business questions into analysis is core to this role, and the panel wants evidence you push back constructively rather than just building what's asked.

How to structure your answer

STAR: situation, task, action, result.

Example answer

"A manager asked me for a dashboard showing total sales by region. When I asked why, it turned out they actually wanted to know why one region had dropped off, which meant a trend and cohort view,

not a static total. I built a quick mock-up of both versions and showed them the difference before committing time to the full build. We agreed on the trend view, and it ended up surfacing a data feed issue in that region that the totals view would have hidden."

3. A dashboard number doesn't match what finance is reporting, and the meeting is in an hour. What do you do?

Why they ask

This is a common real situation for analysts and tests whether you can triage calmly rather than panic or guess.

How to structure your answer

Judgement under pressure: state your immediate priority, your quick diagnostic steps, and how you communicate the risk if you can't fully resolve it in time.

Example answer

"First I'd flag to whoever's chairing the meeting that there's a discrepancy I'm checking, so nobody presents a number with false confidence. Then I'd check the obvious things fast: date filters, whether both reports include the same time period, and whether one side includes adjustments the other doesn't. If I can't resolve it in time, I'd go into the meeting with the raw numbers from both sources and an honest explanation of where the difference likely sits, rather than picking one number and hoping it's right."

4. How would you decide between using SQL, Python or Excel for a given analysis task?

Why they ask

Tests practical judgement about tool choice rather than just tool familiarity.

How to structure your answer

Reasoning question: explain the trade-offs you weigh and give a concrete example of each choice.

Example answer

"It usually comes down to data size, repeatability and audience. If it's a one-off check on a small extract, Excel is often faster and the stakeholder can open it themselves. If I need to pull and join data from source systems, SQL is the right layer, and I'll push as much of the heavy lifting there as possible rather than in the BI tool. Python comes in when I need to automate something that repeats, or when the logic is too complex for a clean SQL query, like a statistical model or a scraping step."

5. How do you explain a statistical result to someone with no data background?

Why they ask

Presenting and data storytelling is a listed skill for this role, and interviewers want to hear you avoid jargon without dumbing down the finding.

How to structure your answer

Communication approach: describe your method for translating technical detail, then give a specific example of doing it.

Example answer

"I lead with the business impact, not the method. Instead of saying the model shows a statistically significant correlation, I'd say something like 'customers who use the app in their first week are twice as likely to stay past three months, and here's what that's worth if we can shift more people into that group.' I keep the technical detail available in an appendix slide for anyone who wants it, but I don't open with it."

6. Describe your approach to building a recurring report that other people will maintain after you.

Why they ask

A lot of analyst work becomes long-running infrastructure, and this checks whether you think about handover and documentation, not just the initial build.

How to structure your answer

Process walk-through: describe your build habits and what you document along the way.

Example answer

"I document the source tables and any assumptions directly in the SQL comments, and I keep a short readme covering what the report is for and who asked for it originally, since that context gets lost fast. In the dashboard itself I try to keep the logic visible rather than buried in calculated fields with cryptic names, so the next person can actually follow it instead of rebuilding it from scratch."