

Data Scientist interview questions

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

What to expect

Data scientist interviews in Australia typically mix technical assessment with judgement questions about how you handle ambiguity, communicate results and work with engineering. Because the role has split into engineering-heavy and analysis-heavy tracks, expect the panel to probe which side you lean toward and whether that matches what they actually need.

- **Technical:** Questions or live exercises on model building, statistics, SQL or Python, sometimes a take-home task or whiteboard problem.
- **Behavioural:** Past-experience questions about handling messy data, disagreements over model results, or working with non-technical stakeholders.
- **Scenario:** Hypothetical situations testing judgement, such as a model showing signs of drift or a stakeholder pushing to ship despite weak validation.
- **Process:** Walk-throughs of how you'd approach a project end to end, from problem framing to deployment.

Most processes run two to three stages: an initial screen covering background and motivation, a technical round (often a take-home exercise reviewed live or a case study), and a final panel round with the hiring manager and a business stakeholder to assess communication and fit.

1. Walk me through how you'd approach building a model to solve [a stated business problem], from data to deployment.

Why they ask

Tests whether you think in a structured way across the full lifecycle: framing, feature engineering, validation and deployment, rather than jumping straight to algorithms.

How to structure your answer

Process walk-through: state how you'd clarify the business question first, then move through data assessment, feature engineering, model selection, validation and handover to engineering, naming trade-offs at each step.

Example answer

"I'd start by clarifying what decision the model needs to support, because that shapes the target variable and the acceptable error type. Then I'd assess what data is actually available versus what I'd want, and do feature engineering on the raw data, being careful about leakage. I'd try a simple baseline model first before anything more complex, validate against a holdout set and check performance across relevant subgroups, not just the aggregate. Once I'm confident it's stable, I'd work with the engineering team on deployment, agreeing on monitoring thresholds so we catch drift early rather than after it's affected decisions."

2. Tell me about a time you designed an experiment or test that didn't go the way you expected. What did you do?

Why they ask

Experiment design is a core task for this role, and the panel wants evidence you understand statistical rigour and can respond to inconvenient results without cutting corners.

How to structure your answer

STAR: situation, task, action, result, with emphasis on what the unexpected outcome revealed and how you adjusted.

Example answer

"I set up a test to measure the effect of a proposed change, but the early results showed no significant difference, which wasn't what the team expected. Rather than extend the test until something significant appeared, I checked the sample size calculation and confirmed we hadn't hit it yet, then reviewed the randomisation for any imbalance. Once we reached the required sample, the effect was there but smaller than hoped. I presented both the result and the reasoning behind waiting for proper power, which meant the team trusted the conclusion instead of second-guessing it."

3. A model you deployed six months ago is now underperforming. How do you diagnose and respond?

Why they ask

Model drift is a live operational issue and shows whether you think about production models as ongoing responsibilities rather than one-off deliverables.

How to structure your answer

Scenario, judgement under pressure: outline immediate triage, root-cause investigation, and how you'd balance a quick fix against a proper retrain.

Example answer

"First I'd check whether the input data distribution has shifted, since that's the most common cause of drift, and compare recent predictions against actuals to quantify how much performance has degraded. If it's a genuine data shift, I'd look at whether a quick recalibration buys time or whether a full retrain is needed. I'd flag the issue to stakeholders early with a rough timeline rather than waiting until I have a complete fix, because a model quietly underperforming in production is worse than one flagged as degraded."

4. How would you explain a model's limitations to a stakeholder who wants to use it for a decision it wasn't built for?

Why they ask

Communicating model behaviour to the business is an explicit task for this role, and the panel wants to see you can push back without being obstructive.

How to structure your answer

Behavioural/STAR with a client-facing lens: focus on how you translated technical limitations into terms the stakeholder could act on.

Example answer

"I had a stakeholder who wanted to use a churn model to prioritise a customer segment it hadn't been trained on. I explained, without jargon, that the model's accuracy for that segment was unproven and could mislead the decision, then offered a practical alternative: a smaller validation exercise on that segment before rolling it into the main process. That let them get useful information quickly without treating an untested output as reliable."

5. Given the skills overlap between data science and roles like machine learning engineering or data analysis, how do you see this role, and which track are you more suited to?

Why they ask

The role data explicitly notes the split into engineering-heavy and analysis-heavy tracks, so the panel wants to know your self-assessment matches what they're actually hiring for.

How to structure your answer

Direct reflective answer: state your genuine leaning, back it with concrete examples, and acknowledge the trade-off.

Example answer

"I lean toward the analysis-heavy side of the work: I enjoy the experiment design and communicating findings more than the infrastructure side of deployment, though I'm capable of working closely with engineering to get models live. If this role is weighted more toward MLOps and production engineering, I'd want to be upfront that my strength is in the modelling and measurement rather than the platform work."

6. SQL and Python question: how would you identify and handle a feature with a lot of missing values before feeding it into a model?

Why they ask

Feature engineering from raw data is a listed task, and this tests practical technical judgement rather than theory.

How to structure your answer

Technical walk-through: name the diagnostic steps first, then the decision criteria for how to handle the missingness.

Example answer

"I'd first check whether the missingness is random or systematic, since that changes the right approach. I'd write a quick SQL query to see if missing values correlate with another variable, like a time period or customer segment, which would suggest it's not random. Depending on that, I'd either impute using a reasonable method, such as median or a model-based imputation, or create a missingness indicator as its own feature so the model can use the pattern itself, rather than just dropping rows and losing signal."