Statistician interview questions

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

What to expect

Statistician interviews test three things at once: whether your methods are sound, whether you can explain them to someone who will never read your code, and whether you can work out what question is actually being asked. Expect the technical discussion to be grounded in real problems rather than textbook definitions.

- **Methods and technical:** Questions about study design, model choice, assumptions, sampling and inference, usually built around a problem the team has faced.
- **Software and coding:** Discussion of your work in R, Python, SAS, SQL or Stata, sometimes with a short practical task or a walk-through of code you have written.
- **Case or scenario:** A loosely specified problem where you are expected to ask clarifying questions before proposing an approach.
- **Behavioural and collaboration:** How you have handled disagreement, shifting scope, or a stakeholder who wanted a particular answer.
- **Communication:** An explanation task, often to a non-statistical panellist, to see how you handle a lay audience.
- **Ethics, privacy and data governance:** How you handle identifiable data, consent, linkage and the limits of what should be reported, particularly in health and government settings.

Most processes start with a short screen with the hiring manager covering your background and one or two methods questions. The main interview is usually a panel of two or three people, often including a senior statistician and a client-side stakeholder, and runs through a case discussion, technical questions and behavioural questions in that order. Some employers add a take-home analysis task or ask you to present a past piece of work to a non-technical audience, and a final conversation with the wider team is common. Panel interviews are the norm; be ready to write on a whiteboard or share your screen.

1. Walk us through how you would design a survey to estimate a health measure across a state, when some groups are much less likely to respond than others.

Why they ask

Survey design is core to the role and the panel wants to see whether you think about the whole process, from the target population through to the final estimate, rather than jumping straight to analysis.

How to structure your answer

Define the target population and the estimand first, then the sampling frame and design, then the fieldwork plan including non-response follow-up, then weighting and adjustment, then how you would validate the result against independent benchmarks. Finish with the limitations you would disclose.

Example answer

“I would start by pinning down exactly what we are estimating and for whom, because a state-level average and a regional estimate need different sample sizes. From there I would build the frame, stratify by region and by the demographic groups we expect to differ, and oversample the groups that are hardest to reach so their estimates are usable in their own right. For non-response I would plan follow-up waves and a mixed-mode approach, since a single mail-out will always leave some groups under-represented. When the data come in, I would weight back to known population benchmarks, check how far the weighted and unweighted estimates diverge, and compare the sample against independent sources such as ABS figures. I would report the design effect and the margins of error alongside the estimates so users know what the numbers can and cannot support.”

2. Tell me about a time a stakeholder disagreed with your analysis or wanted a different result. How did you handle it?

Why they ask

Statisticians regularly deliver inconvenient findings. The panel is testing whether you hold the line on method while keeping the working relationship intact.

How to structure your answer

Use STAR. Set the situation and what the stakeholder wanted, describe the specific action you took to understand their concern and re-check your work, then give the resolution and what changed afterwards.

> **Example answer**
>
> *"A program manager in a previous role was expecting an evaluation to show their intervention had improved outcomes, and my first cut of the analysis showed almost no movement once I adjusted for the fact that the groups differed going in. I sat down with them before sending anything in writing, went through how the adjustment worked, and asked what they had seen on the ground that the data might be missing. That conversation surfaced a subgroup we had not pre-specified, so I ran it as an exploratory analysis and flagged it clearly as exploratory in the report. Their concerns also led me to check the missing data pattern, which turned out to explain part of the initial result. In the end the headline finding stood, but the report was more useful to them, and they brought me into the design of the next evaluation at the start rather than the end."*

3. You have a dataset where a simple model and a more complex model give different answers, and the more complex one fits better. What do you do?

Why they ask

Model choice under uncertainty is the everyday judgement call in this role, and the panel wants to see whether you treat a better fit as the end of the story or the start of the questions.

How to structure your answer

State your reasoning out loud: what the models are assuming, how they differ, what the purpose of the analysis is, then how you would test the difference, then how you would report it. Show that you weigh interpretability and intended use, not just fit.

> **Example answer**
>
> *"First I would ask what the analysis is for. If the goal is prediction, I would compare the models on held-out data and look at how stable the results are across folds, because a complex model that wins by a small margin on one split is not a reliable answer. If the goal is explanation, I would be more cautious, because a model with interactions and nonlinear terms can fit better while telling a story nobody can act on. I would check whether the added complexity is doing real work or absorbing a handful of influential points, look at the residuals, and test whether the simpler model's assumptions are actually violated. If the two models lead to different conclusions, I would report both, say plainly which one I trust and why, and describe what data would settle it. What I would not do is quietly pick whichever one gave the more convenient answer."*

4. How do you satisfy yourself that a model is valid and that your conclusions are defensible?

Why they ask

The panel needs to know your analysis will survive review, whether that is internal peer review, an ethics committee or a journal.

How to structure your answer

Give a clear process rather than a list of tests: assumptions, diagnostics, sensitivity analysis, validation against external sources, and documentation. Mention who reviews your work and how you handle disagreement at that stage.

Example answer

“I start by writing down what the model assumes and why those assumptions are reasonable for this data, because that is the part most likely to be challenged later. Then I run diagnostics that actually speak to those assumptions, not a standard printout I ignore, and I look at influential observations individually to understand what they are doing. I always run a sensitivity analysis: if the conclusion flips under a reasonable alternative specification, that is the finding, not the original result. Where an external benchmark exists, such as a known population total or a previous estimate, I check against it. I keep the analysis in a script so anyone can rerun it, and I ask a colleague to review the code and the methods before results go anywhere. If a reviewer disagrees, I would rather document the disagreement in the methods note than let it disappear.”

5. What is your working setup in R or Python, and how do you use SQL or SAS day to day?

Why they ask

This is a practical check on whether you can slot into the team's existing toolchain and whether your code would be usable by someone else.

How to structure your answer

A structured walk-through of your actual setup: how you get data, how you clean and analyse it, how you document and version it, and how you hand it on. Be honest about which tools you use deeply and which you would need to pick up.

Example answer

“Most of my analysis work is in R, usually tidyverse for data handling and Quarto for reporting, with models run through base R, lme4 or brms depending on the problem. I pull data with SQL rather than exporting spreadsheets, and I keep everything in Git so there is a record of what changed and why. I write functions for anything I do more than twice, and I leave a short README so someone else can run the pipeline without asking me. I have used Python for machine learning work and for anything that needs to sit alongside a production system, and SAS mainly when a client's existing reporting sits on it. Stata I would describe as working knowledge rather than deep. If a team is standardised on a tool I have used less, I would expect to be productive in it within the first few weeks.”

6. You are asked to analyse a linked dataset containing identifiable health information, and a colleague suggests including a variable that would make small groups identifiable in the results. How do you respond?

Why they ask

Statisticians in health and government settings work under privacy and ethics obligations, and the panel wants evidence that you treat these as part of the method, not an administrative hurdle.

How to structure your answer

Show judgement: identify the issue, check the governing framework, propose a workable alternative, and describe how you would escalate if the answer was not yours to make.

Example answer

“I would raise it straight away rather than after the results were drafted, because re-identifiability in small cells is a real risk and it is much harder to unwind later. I would check what the ethics approval and the data custodian's conditions actually permit, since those documents usually say something specific about small cell reporting and whether the variable was in scope when the application was made. In most cases there is a way to answer the underlying question: collapsing categories, grouping over a longer time period, or running the model without that variable and testing how much it changes the conclusion. If the question genuinely needs the variable, that is a conversation with the data custodian and possibly a variation to the ethics approval, not something I would decide on my own. I would also record the decision and the reasoning so the next person working on the dataset is not starting from scratch.”