

Sampling for Program Assessment

How much evidence you need, and how to choose it fairly

You don't always need to score every piece of student work or review every record. A well-chosen sample can give you trustworthy results with far less effort. What matters is that the sample **fairly represents** the people or work you're drawing conclusions about, and that your report explains how you chose it.

Census or sample?

Situation	Suggested approach
Small program (roughly 30 or fewer students or records in the year)	Use all of them (a census). Sampling saves little and weakens results.
Larger program	Draw a random sample . Many programs find that 20 to 50 pieces of work give a clear picture for one outcome.
Work spread across several course sections	Use a stratified sample : draw from each section in proportion to its size, so no single section dominates.
Very small numbers every year	Combine two or more years of work, and say so in your report.

The 20 to 50 range is a common rule of thumb for program-level assessment, not a statistical requirement. Larger samples give more stable results, especially when you plan to compare groups.

How to draw a random sample

1. List every eligible piece of work or record (for example, all capstone papers from the last two sections).
2. Number the list.
3. Use a random number generator or a spreadsheet's random function to pick the sample.
4. Remove names and identifying information before scoring.

What to say in your report

In **Part B**, describe:

- the full group (for example, all 34 papers from PSY 424 in fall and spring),
- how many you assessed and how you chose them,
- who scored the work, and whether scorers practiced on shared examples first.

Avoid convenience samples. Using "the papers I happened to have" or "the best examples" makes results hard to trust. If a random sample isn't possible, explain why and what that might mean for the findings.

Created in collaboration with Claude AI.