

# Collecting Data: Sampling Plans and Experiments

Grade 12 – AP Statistics math · Lesson · Standards: AP Statistics Unit 1 (2026-27)

**Quick answer:** Good data collection starts with a statistical question and a sampling plan that uses chance, such as a simple random, stratified or cluster sample. Experiments use random assignment, control and replication to show cause and effect; observational studies cannot rule out confounding.

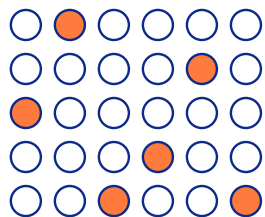
**Standards:** AP Statistics Unit 1 (2026-27) · **Level:** Grade 12 AP Statistics

## What makes a good statistical question?

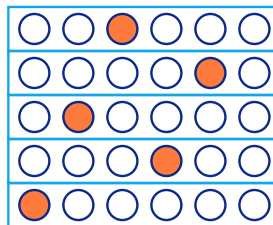
A statistical question expects **variability** in the answer: "How long do students at our school spend on homework each night?" has many answers, so data are needed. "How many students attend our school?" has one answer.

## How do sampling methods differ?

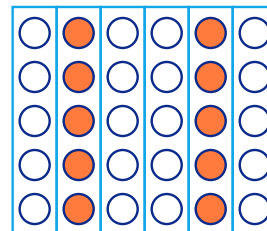
Random sample



Stratified



Cluster



Three sampling plans: simple random, stratified (one group per row) and cluster (whole columns).

### Sampling plans

- **Simple random sample (SRS):** every group of  $n$  individuals has an equal chance of being chosen.
- **Stratified:** split into similar groups (strata), then take an SRS from every stratum.
- **Cluster:** split into groups that each mirror the population, choose whole clusters at random, and survey everyone in them.
- **Systematic:** choose every  $k$ th individual after a random start.

Convenience samples and voluntary-response samples are biased because the sampler or the respondents decide who is included. Other sources of bias are **undercoverage** (part of the population has no chance of selection), **nonresponse**, and **response bias** from leading wording

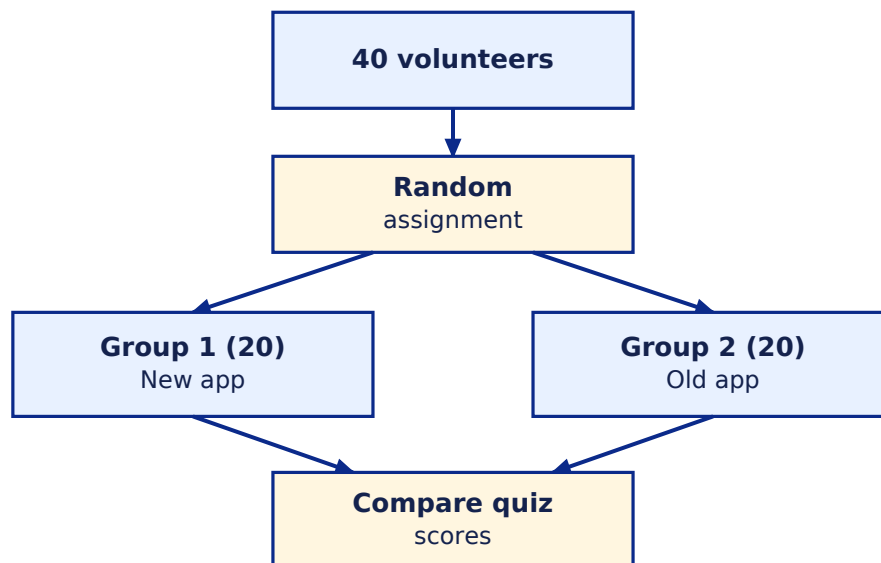
or untruthful answers.

## What separates an experiment from an observational study?

An **observational study** records data without imposing treatments, so lurking variables can be **confounded** with the explanatory variable. An **experiment** imposes treatments on experimental units.

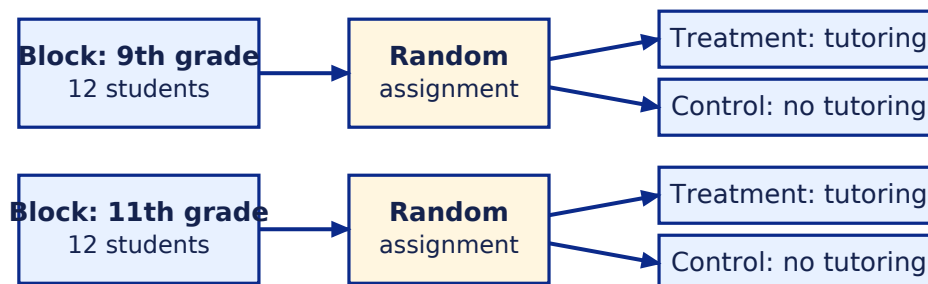
### Principles of experimental design

- **Comparison:** use a control group or compare treatments.
- **Random assignment:** chance decides who gets each treatment, spreading other variables evenly.
- **Replication:** enough units in every group to see a real difference.
- **Control:** keep other conditions the same, and use blinding or a placebo when expectations could matter.



*Completely randomized design.*

In a **randomized block design**, units are first grouped by a variable that affects the response, and random assignment is done within each block. A **matched-pairs design** uses pairs of similar units, or two measurements on one unit, with treatments randomized within each pair.



## What conclusions are allowed?

|                  | Random assignment                                 | No random assignment                              |
|------------------|---------------------------------------------------|---------------------------------------------------|
| Random sample    | Cause and effect, generalize to the population    | Generalize to the population, no cause and effect |
| No random sample | Cause and effect, only for the units in the study | Neither                                           |

### Method: describe a randomized experiment

1. Name the experimental units, treatments and response variable.
2. Say how you will assign treatments at random (labels and a random number generator).
3. Mention control, blinding and replication.
4. State how the responses will be compared.

### Mathbot's tip

Random sampling lets you generalize to a population. Random assignment lets you claim a cause. Say which one the study used before you write any conclusion.

### Key points to remember

- A statistical question anticipates variability in the answers.
- Random sampling supports generalizing; random assignment supports cause and effect.
- Convenience and voluntary-response samples are biased.
- Experiments need comparison, random assignment, replication and control.
- Blocking and matched pairs reduce variability from a known variable.

## Collecting Data: Sampling Plans and Experiments: frequently asked questions

### What is the difference between random sampling and random assignment?

Random sampling chooses who is in the study and supports generalizing to the population. Random assignment decides which treatment each participant receives and supports cause-and-effect conclusions. A study can use one, both or neither.

### Why is a voluntary-response sample biased?

People with strong opinions are most likely to respond, so the sample over-represents extreme views. Online polls and call-in shows are typical examples. The bias does not shrink by collecting more responses.

### What does confounding mean?

Two variables are confounded when their effects on the response cannot be separated. In an observational study, a lurking variable may be tied to the explanatory variable and also affect the response, so no cause can be claimed.

### **When should an experiment use blocks?**

Use blocks when a known variable, such as age or soil type, strongly affects the response. Randomizing within each block removes that variable's variation from the comparison, making treatment differences easier to detect.

### **What is the purpose of a placebo and blinding?**

A placebo gives the control group an inactive treatment that looks real, so expectations are equal in all groups. Blinding keeps subjects, evaluators or both from knowing the treatment, which prevents bias in the responses and in their measurement.