

Random Sampling

AP Statistics | Collecting Data | SRS, stratified, cluster, systematic

1. Population, sample, census

- **Population:** the entire group we want to know about (all 100 students at a school).
- **Sample:** the part of the population we actually collect data from.
- **Census:** data from *every* individual. Slow, expensive, often impossible.

A spoonful tells you how the whole pot of soup tastes, but only if the pot was stirred. In statistics the stirring is **random selection**, and it matters far more than the size of the spoon.

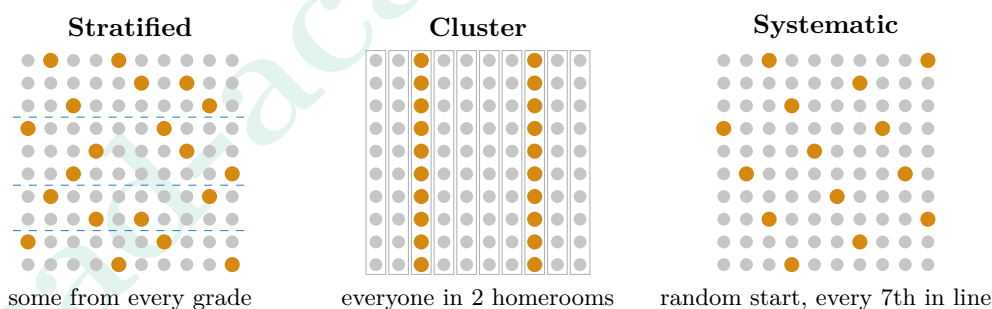
2. The simple random sample (SRS)

Definition and how to describe one

An SRS of size n is chosen so that **every group of n individuals is equally likely** to be the sample.

To earn the point, say *how*: (1) number the students 1 to 100; (2) use a random number generator to pick 20 **different** numbers from 1 to 100; (3) survey the students with those numbers. *Label, randomize, select.*

3. Stratified, cluster, systematic



- **Stratified random sample:** split the population into groups of *similar* individuals (strata), then take an SRS from **every** stratum. No group is left out by bad luck, and estimates tend to be more precise than an SRS of the same size.
- **Cluster sample:** split the population into groups, usually by location (clusters), randomly choose some clusters and take **everyone** in them. Cheap and convenient.
- **Systematic random sample:** pick a *random* starting point, then take every k -th individual. Careful if the list has a pattern that repeats every k .

Stratified vs cluster

Stratified: *some* individuals from *every* group (groups similar inside; goal: precision).

Cluster: *all* individuals from *some* groups (groups mixed inside; goal: convenience).

A company picks 5 of its 40 stores at random and surveys every employee in them: all from some groups, so it is a **cluster** sample.

4. Why random at all?

Random selection is what lets you **generalize**: what you see in the sample is a fair estimate of the population. Without random selection you cannot generalize, however big the sample is.

Where students lose points

1. Writing “choose randomly” without saying *how*.
2. Mixing up stratified and cluster.
3. Forgetting to say the random numbers must be *different* (no repeats).
4. Thinking a bigger sample fixes a non-random one.
5. Defining an SRS as “everyone has an equal chance”; it is every *group* of size n .

Now practise it

Theory sticks once you use it. On **sad-academy.com** you’ll find practice questions for this exact subtopic in the AP format, and an AI that grades your free-response answers against the scoring guidelines, the way an AP reader would.

sad-academy.com