Two approaches to collecting data in statistics

study a study in which all data, including observations and measurements, are recorded in a way that the subject that is being measured or studied.

An study allows information to be gathered or impacting the subject(s) at all

is a process or action that has results called . Researchers gather data by a treatment and observing responses.

example 1

a team of zoologists might want to study the habits of an endangered bird species, but to disturb or interact with the birds may cause the animals to behave differently than they normally would. Therefore, the team may choose to observe the birds from a safe distance using binoculars.

Key steps involved in designing an **observational study**.

- Determine the of the study.
  - What is the of interest?
  - What is needed to answer the main question of interest?
- Develop a to collect data.
  - How will subjects be ?
- Determine the most appropriate method and select the .
- the data.
- Describe and the data using appropriate statistical procedures and graphs.
- the findings of the study.

In order to conduct a experiment, at least a portion of the population studied subjected to the , , or being evaluated

A is a group of study participants who are subjected to the treatment, action, or process being studied in the experiment. By using a , researchers can the outcomes of the experiment between this group and the group actually the treatment, and understand the of what is being studied.

Control has several different uses in design.

- First, an experiment is controlled because we as experimenters assign to . Otherwise, we would have an .
- Second, a control treatment is a treatment that is used as a or basis of comparison for the other treatments. This control treatment might be the treatment in common use, or it might be a treatment (at all).
- For example, a study of new pain killing drugs could use a as a control treatment

units are the things to which we the treatments. These could be plots of land receiving fertilizer, groups of customers receiving different rate structures, or batches of feedstock processing at different temperatures. One way to determine the experimental unit is by the consideration that an experimental unit should be able to .

units (or units) are the actual on which the response is measured. These may differ from the units. For example, consider the effect of different fertilizers on the nitrogen content of corn plants. Different field plots are the units, but the measurement units might be a of the corn plants on the field plot, or a of leaves, stalks, and roots from the field plot.

Design and Analysis of Experiments
Gary W. Oehlert University of Minnesota

[ ] are the different procedures we want to compare. These could be [ ] kinds or amounts of fertilizer in agronomy, [ ] long distance rate structures in marketing, or [ ] temperatures in a reactor vessel in chemical engineering.

There are many situations where a [ ] is applied to [ ] of objects, some of which are later measured for a [ ]

Example 1

Fertilizer is applied to a plot of land containing corn plants, some of which will be harvested and measured. The plot is the [ ] and the plants are the [ ]

Example 2

Ingots of steel are given different heat treatments, and each ingot is punched in four locations to measure its hardness. Ingots are the [ ] units and locations on the ingot are [ ] units.

A common source of difficulty is the distinction between [ ] units and [ ] units. Consider an educational study, where six classrooms of 25 first graders each are assigned at random to two different reading programs, with all the first graders evaluated via a common reading exam at the end of the school year. Are there six experimental units (the classrooms) or 150 (the students)?

If students were the units, we could see than one reading program in each classroom. However, the nature of the makes it clear that all the students in the classroom the same program, so the as a whole is the experimental unit. We don't how a classroom , though; we how read.
Thus students are the units for this experiment

are outcomes that we observe after applying a treatment to an experimental unit. That is, the is what we , to judge what happened in the experiment; we often have response.

is the use of a , probabilistic mechanism for the assignment of treatments to units. Other aspects of an experiment can also be randomized: for example, the order in which units are evaluated for their responses.

Error is the random variation present in all experimental results. experimental units will give responses to the treatment, and it is often true that applying the same treatment over and over again to the same unit will result in responses in different trials. Experimental error does refer to conducting the wrong experiment or dropping test tubes.

Design and Analysis of Experiments

Gary W. Oehlert University of Minnesota

Not all experimental designs are created equal. A good experimental design must

- Avoid error
- Be
- Allow of error
- Have validity.

Design and Analysis of Experiments

Gary W. Oehlert University of Minnesota

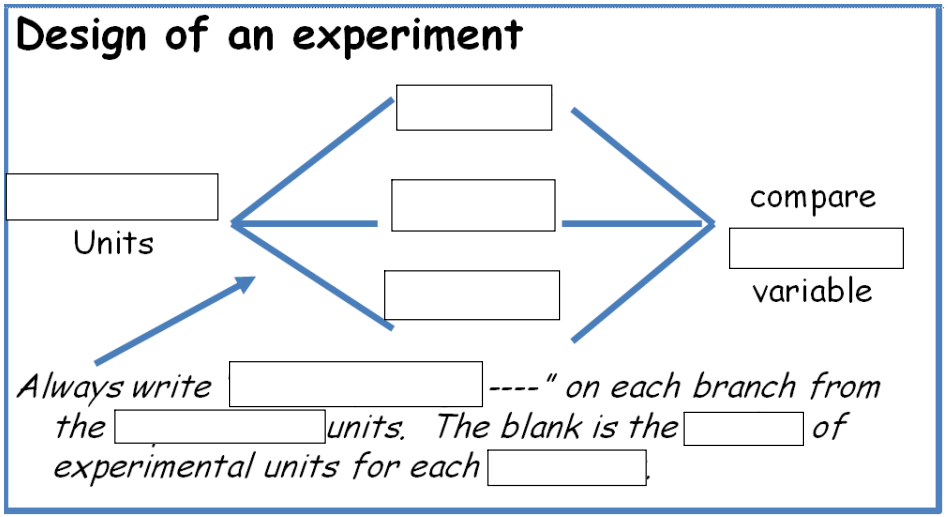
The basic principles of [ ] design are

1. [ ] – Experimental units/subjects should be [ ] assigned to treatment groups

2. [ ] - Experimenters need to control any [ ] variables, generally by comparing [ ] treatment groups

3. [ ] – The experiment should involve [ ] experimental units/subjects.

An experiment is characterized by the treatments and experimental units to be used, the way treatments are assigned to units, and the responses that are measured.



A local community has just installed red light cameras at its busiest intersection. The police department hopes that the cameras will encourage drivers to be more careful and that incidents of drivers running red lights at this intersection will decrease. Design an observational study that the police department could use to determine if the installation of the traffic light has had the deserved effect.

a. What is the focus of the study?

b. What is the variable of interest?

c. Determine the data collection plan.

d. Funds are limited and there are only a few days to conduct the study. What is the most appropriate sampling method?

e. The police chief also wonders if there is a difference in driver behavior at different times of day. How would you incorporate this concern into your sampling method?