Collecting Data Part II

- Up to this point we have explored observational studies, these cannot be used to draw cause-and-effect relationships between variables. They instead give us a 'snap-shot' of the population at some time.

Establishing Relationships between variables.

- After collecting data we may explore relationships between variables.

Ex:
- does a medication help treat high blood pressure.
- Is the ammount of time studying associated with higher grades.
- Are midterm grades associated with final exam grades.

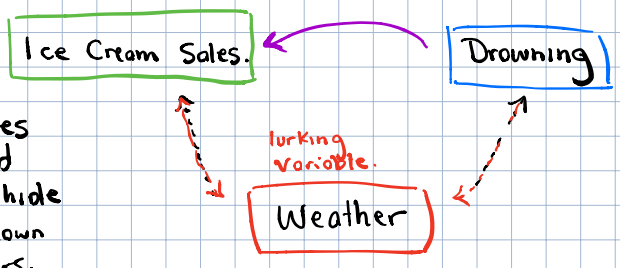
- we are interested in seeing if changes in one variable (in-part) may affect the other variable. We consider one variable 'changer', and the other the 'responder'.

Explanatory variable: variable that explains or causes changes in another.

Responding variable: The variable that responds to or changes due to changes in explanatory variable.

A Lurking Variable: Is a variable that can affect the response variable, but is not accounted for in the studies design.

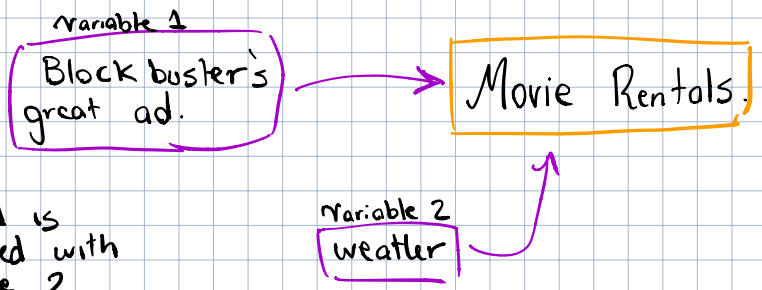
Example.



- lurking variables are not included in study, they hide or 'lurk' unknown to researchers.

Confounding variable: This is a variable that is related to the explanatory variable of an experiment, that influences the response variable. This creates a false perception of association.

Examples



variable 1 is confounded with variable 2.

Confounding arises when the response we see in an experiment is at least partially attributable to uncontrolled variables.

↳ extraneous variables

When Exploring relationships we 2 Options

1. **Observational Studies.**
   - cannot establish cause-and-effect relationships.

2. **Experiments.**
   - used to establish cause and effect relationships between variables.
   - we look at several 'experimental designs'.

Experiments

- In experiments we impose treatments on experimental units to measure 'response'.

Some Important Vocabulary

Experimental units: individuals that are involved in experiment. When humans we call them 'subjects'.

Factors: a common term to denote explanatory variables.

↳ **extraneous factors**: thought to have an effect but aren't involved in study, out of researchers control.

↳ **confounded factors**: are factors whose response effect cannot be distinguished from one another.

↳ Factors may have multiple levels.
Ex: Age: 1-4, 5-8, ...
Color: red, green, blue, ...

The combination of specific levels from all factors that one unit undergoes is known as a **treatment**.

A **Placebo**: is a treatment that resembles another in appearance, but has no active ingredients at play.

↳ - Give a drug as one treatment.
- Give tablet resembling drug as another.

A **control group** is a group that recieves no treatment or a placebo treatment.

What must a well-designed Experiment Include?

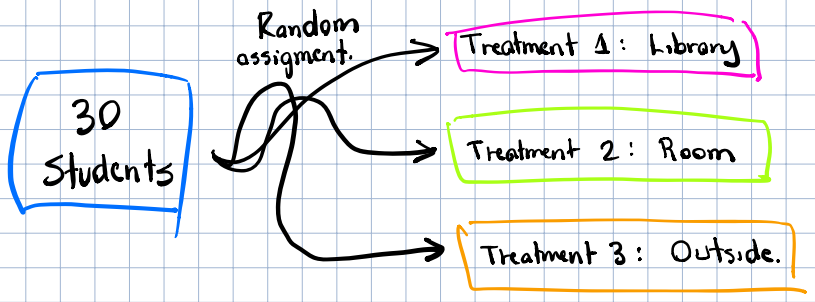
Control
Replication
Randomize

1. Comparison of **at least** two treatment groups, one of which might be a control.
2. **Random assignment/allocation** of treatments to experimental units.
3. **Replication**: require more than one experimental unit for each treatment group.
4. Control of potential confounding variables where appropriate.

We look at 3 types of experimental Design

1. **Completely Randomized Design:** All experimental units are assigned at random to treatment groups.

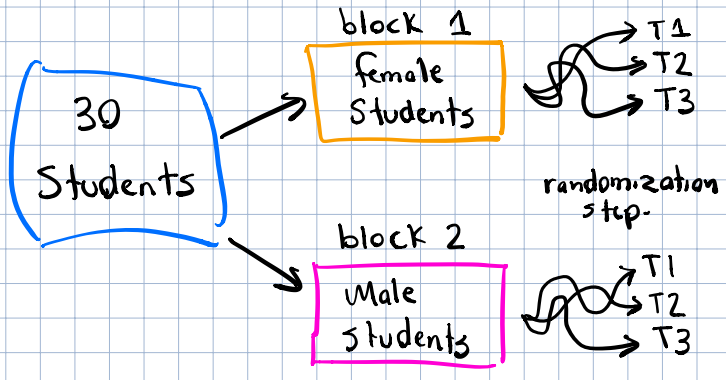
Ex: A researcher wants to conduct an experiment to determine which environment is bet suited for studying. A library, ones own room, outside. 30 university students participate in the study



2. **Randomized Block Design**

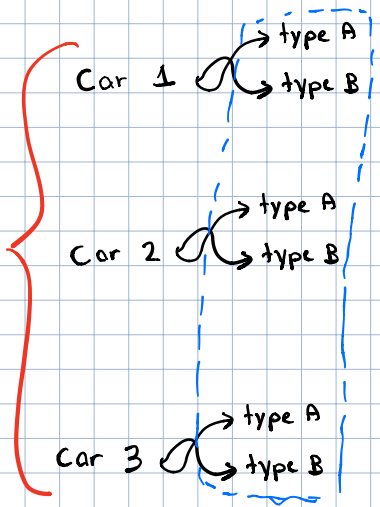
- experimental units are divided into blocks that are expected to have an outcome on response.
- we block and then randomize

Ex: Same example, but the researcher thinks that gender has an effect on study habits.



Matched Pair Design: A special type of randomized block design that compares 2-treatment groups using the same or similar experimental units.

Ex: What type of gasoline is more efficient, Gas A, Gas B, 3 cars are used in this experint.



we would compare effiancy.

Ex: Sometimes we are unable to use the same units in a matched pairs design, so we use similar units.

- A researcher wants to determine wether or not sleep deprivation has an effect on test scores?
- pair students with similar GPAs for experiment.

high GPAs.

pair1 { • → normal ammount of sleep; • → sleep deprivation

pair { • → normal ammount of sleep; • → sleep deprivation

⋮

pair { • → normal ammount of sleep; • → sleep deprivation

Compare results for each pair of people across all experimental units.

Additional Precautions

- **Single-blind experiment**: Experimental units do not know what treatment they are assigned to.

- **double-blind experiment**: neither research team, or units know which treatment is which.

confirmation bias: occurs when someone finds results because they want to confirm their own beliefs.

- Another serious weakness of experiments is a lack of **realism**, meaning treatments/units/design does not realistically replicate the conditions we would like to study.