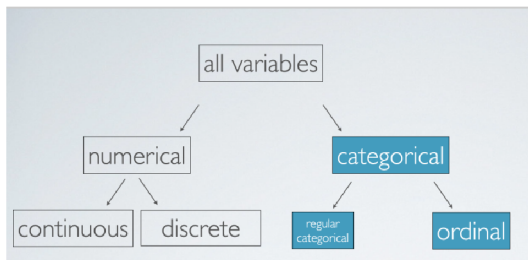


# Course 1 - Week 1 - Intro

1. Key components of a statistical study
  - a. Research question, population, sample and where to generalize
2. Explain data matrix, observation and variable
3. Types of variables



For example, “height” is a numerical continuous variable while “number of family members” is a numerical discrete variable.

Customer satisfaction rating “very good”, “good”, etc should be treated as ordinal categorical variable.

4. What's the difference between observational studies and experiments?
  - a. In observational studies, we collect **existing** data and record their values (like street interviews). We do **not produce** any data. We can only draw **association** instead of causation in this type of study. There might be **confounding variables** that affect both **explanatory variables** and **response variables** and make it seem like there is a correlation.
  - b. In experiments, we random assign objects to treatment and control groups. In this way, we are actually **creating** data for the study. We can establish **causal** relationship.
5. List some sources of sampling bias
  - a. Convenience sample
  - b. Non-response
  - c. Voluntary response: such as a war survey at the bus stop. only those with a strong opinion participate
6. List some sampling methods
  - a. *Simple Random Sample*: Each case is equally likely to be selected
  - b. *Stratified Sample*: Divide population into homogeneous strata (e.g., male and female) and random sample from each stratum
  - c. *Cluster Sample*: first divide population into heterogeneous clusters and randomly sample a few clusters (use ALL examples)
  - d. *Multistage Sample*: first divide population into heterogeneous clusters, randomly sample a few clusters, and continue randomly sampling from these clusters (two random process)
7. Four principles of experimental design?
  - a. Control: treatment group & control group
  - b. Randomize: random assign objects to treatment
  - c. Replicate: collect sufficient samples or replicate the experiment
  - d. Block: variables (other than explanatory variables) that might affect the outcome. For example, we want to investigate the effect of light and air condition on exam performance; however, we also know that these two factors affect exam performance differently for males and females. Therefore, we separate our experiment on gender.
8. Difference between the concept of blocking and stratifying
  - a. Stratifying happens in random sampling process

- b. Blocking happens in random assignment process
- 9. Explain placebo, placebo effect, blinding, and double-blind
  - a. placebo: fake treatment
  - b. placebo effect: tell control group objects that they are being treated (actually not) and there is an improvement
  - c. blinding: experiment objects don't know which group they are in
  - d. double-blind: both the researchers and the experiment objects don't know which group they are in
- 10. Difference between random sampling and random assignment

	Random assignment	No random assignment	
Random sampling	causal and generalizable	not causal, but generalizable	Generalizability
No random sampling	causal, but not generalizable	neither causal nor generalizable	No generalizability
	Causation	Association	

