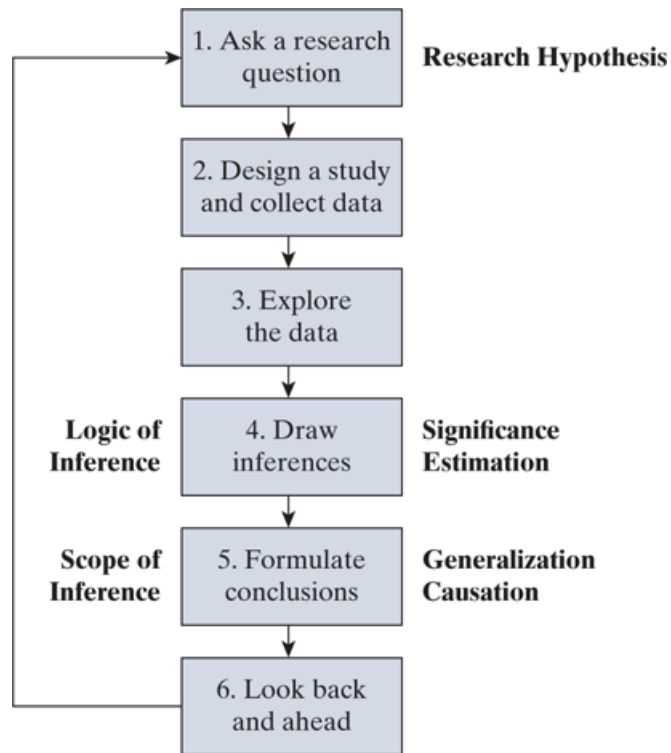


Understand the 6-Step Process



Understand the 4 Pillars of Statistical Inference

Significance

How **Strong** is the evidence of an effect?

Estimation

What is the **size** of the effect?

Generalization

How **broadly** do the conclusions apply?

Causation

Can we say what **caused** the observed difference?

Understand the following terminology

Observational Units

These individual entities on which data are recorded

Categorical Variable

Variables with non-numeric values or with which arithmetic cannot be done

Quantitative Variable

Variables with numeric values with which arithmetic can be done

Distribution

The pattern of outcomes of a variable

1) When surveys are administered, it is hoped that the respondents give accurate answers. Does the mode of survey delivery affect this? American researchers investigated this question (Schober et al., 2015). They had 634 people agree to be interviewed on an iPhone and they were randomly assigned one of the four interview modes (Human Voice, Human Text, Automated Voice, Automated Text). One question that was asked was whether or not they exercise less than once per week on a typical week. The table below shows the counts for the number of people who answered “yes” and the number who answered “no,” separated by the interview mode.

	Human Text	Automated Text	Human Voice	Automated Voice	Total
Yes	34	46	21	20	121
No	124	111	139	139	513
Total	158	157	160	159	634

a) Identify the two variables in this study.

Interview Method and whether the respondent answered “yes” or “no” to the question about weekly exercise.

b) Are the two variables in this study categorical or quantitative?

Both are categorical variables.

c) What are the observational units in this study?

Each of the 634 survey respondents.

2) Student researchers at Hope College conducted an experiment to determine whether there is a difference in memorization ability of students when they take notes on paper using handwriting versus taking notes on a computer. They randomly assigned 20 students to the paper-based notetaking group and 20 students to the computer-based notetaking group. They showed all their subjects a 12-minute video about the sun, asking students to take notes using the method assigned. After the video was over, the notes were collected and the students were given a 10-question quiz on information about the sun given in the video, with the instructor recording the number correct out of 10.

a) Identify the two variables in the study.

Whether the student hand-writes notes or types on a computer, and the student’s score on a 10 question quiz.

b) Are the two variables in this study categorical or quantitative?

Note-taking method is categorical, the students’ score is quantitative

c) What are the observational units in this study?

Each of the 20 students.

3) Suppose that the observational units in a statistical study are purchases made on a particular website (i.e. Amazon) for one month.

- a) Identify two possible categorical variables that could be recorded for these purchases.

Answers will vary. Possibilities could be “what is the item purchased”, “is it a gift”, “is it next-day shipping”, “is this a recurring purchase”

- b) Identify two possible quantitative variables that could be recorded for these purchases.

Answers will vary. Possibilities could be “how many items purchases”, “total cost”, “calculated tax”

- c) State two research questions that the company executives might want to investigate with this data.

Answers will vary. Examples include “What items sell the most” and “are recurring purchases, on average, more profitable than one-time purchases?”

4) Start thinking about possible project ideas or research questions. The project dataset must be looking for an association with a quantitative variable (response) and must include a minimum of 3 other variables (explanatory or confounding). At least one of these other variables must be categorical and at least one must be quantitative.