

Chapter 2 Summary

Data

What did you learn?

Data is information in context

- Who, what, why, where, when, and how.
- We must know at least who (cases), what (variables), and why (how do we treat variables) to be able to say anything useful about the data.

Variables are separated in two basic ways: categorical or quantitative

- Categorical variables group cases into categories and counts them.
- Quantitative variables record measurements or amounts in units.
- Thinking needs to be done as to how to treat the variable.

Statistics are used to help companies market to their customers (*Amazon.com*)

Data needs context: who, what, when, where, why, and how

Context	The context ideally tells <i>Who</i> was measured, <i>What</i> was measured (in units), <i>How</i> the data were collected, <i>Where</i> the data were collected, and <i>When</i> and <i>Why</i> the study was performed.
Data	Systematically recorded information, whether numbers, or labels, together with its context.

Data needs to be organized

Data table	An arrangement of data in which each row represents a case and each column represents a variable.
------------	---------------------------------------------------------------------------------------------------

Who

Case	A case is an individual about whom or which we have data.
Respondents	Individuals who answer a survey
Subjects/ Participants	People on whom we experiment
Experimental units	Animals, plants, web sites, and other inanimate objects upon whom we experiment
Records	The rows in a database, usually identifying cases

What and Why

Area codes were originally designated so that easy numbers to dial on a rotary dial were in high population areas (New York City was 212, L.A. was 213, etc.)

Variable	A variable holds information about the same characteristic for many cases.
Units	A quantity or amount adopted as a standard of measurement, such as dollars, hours, or grams.
Categorical variable	A variable that names categories (whether with words or numerals) is called categorical.
Quantitative variable	A variable in which the numbers act as numerical values is called quantitative. Quantitative variables always have units.

When, Where, and How also can help understand the data

What can go wrong?

- Don't label a variable as categorical or quantitative without thinking about the question you want it to answer.
- Just because your variable's values are numbers, don't assume that it's quantitative.
- Always be skeptical.