Section 1-1

Statistical and Critical Thinking

WHAT IS STATISTICS?

- Actual and specific numbers derived from data.
- **Definition:** The science of planning studies and experiments; obtaining data; and then organizing, summarizing, presenting, analyzing, and interpreting those data and then drawing conclusions based on them.

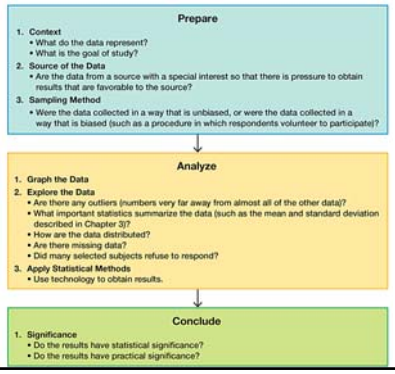
SOME IMPORTANT DEFINITIONS

- **Data:** collections of observations, such as measurements, genders, or survey responses.
- **Population:** the entire group of subjects or complete collection of measurements to be studied.
- **Census:** the collection of data from every element in a population.
- **Sample:** a subset of, or sub-collection from, a population.

STATISTICAL THINKING

This section introduces basic principles of statistical thinking used throughout this book. Whether conducting statistical analysis of data that we have collected, or analyzing a statistical analysis done by someone else, we should not rely on blind acceptance of mathematical calculation. We should consider the following process and its related factors:

THE STATISTICAL PROCESS



PREPARE: CONTEXT

- What do the values represent?
- Where did the data come from?
- Why were they collected?
- An understanding of the context will directly affect the statistical procedure used.

PREPARE: SOURCE OF THE DATA

- Is the source objective?
- Is the source biased?
- Is there some incentive to distort or spin results to support some self-serving position?
- Is there something to gain or lose by distorting results?
- Be vigilant and skeptical of studies from sources that may be biased.

PREPARE: SAMPLING METHOD

- Does the method chosen greatly influence the validity of the conclusion?
- A **voluntary response sample** (or **self-selected sample**) is one in which the respondents themselves decide whether to be in the study. Voluntary response samples often have bias (those with special interest are more likely to participate). These samples' results are not necessarily valid.
- Other methods are more likely to produce good results.

ANALYZE

1. Graph and explore. This will be discussed in detail in Chapter 2.
2. Apply Statistical Methods. This will be discussed in detail in Chapters 6, 7, 8 and 10.

CONCLUDE

- Make statements that are clear to those without an understanding of statistics and its terminology.
- Avoid making statements not justified by the statistical analysis.

CONCLUDE: STATISTICAL SIGNIFICANCE

- Consider the likelihood of getting the results by chance.
- If results could easily occur by chance, then they are ***not statistically significant***.
- If the likelihood of getting the results is so small, then the results are ***statistically significant***.

PRACTICAL IMPLICATIONS

- State practical implications of the results.
- There may exist some ***statistical significance*** yet there may be ***NO practical significance***.
- Common sense might suggest that the finding does not make enough of a difference to justify its use or to be practical.

ANALYZING DATA: POTENTIAL PITFALLS

- Bad Samples
- Small Samples
- Misleading Conclusions
- Sample Data Reported Instead of Measured
- Loaded Questions
- Order of Questions
- Misleading graphs and pictographs
- Percentages

BAD SAMPLES

1. **Voluntary Response Sample** (or **Self-Selected Sample**): a sample in which the respondents themselves decide whether to be included. For example, call in television and radio polls.

2. Asking Emory University students which brand of soft drink they prefer. (This is an example of a biased sample.)

SMALL SAMPLES

EXAMPLE: Two-thirds of students were suspended more than three times. What was not disclosed was that only 3 students were in the sample!

QUESTIONS

Loaded Questions

EXAMPLE: Don't you think that changing the HOPE scholarship program is unfair?

Order of Questions

EXAMPLE:

- Would you say that traffic contributes more or less to air pollution than industry?
- Would you say that industry contributes more or less to air pollution than traffic?

MISLEADING CONCLUSIONS

- When forming a conclusion, make clear statements understandable to those without an understanding of statistics or its terminology.
- Avoid making statements not justified by statistical analysis.

CORRELATION AND CAUSATION

Just because there is an association between two variables (correlation) ***DO NOT*** concluded that one variable caused the other to occur (causation).

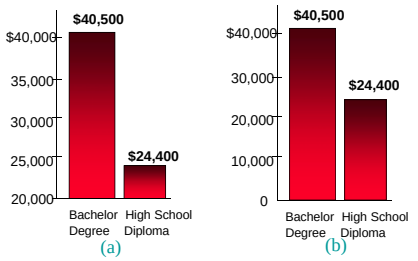
EXAMPLE: A study showed that truck drivers weigh more than adults who do not drive trucks. Conclusion: Trucks cause people to gain weight.

SAMPLE DATA REPORTED INSTEAD OF MEASURED

When collecting data from people, it is better to take measurements instead asking people to *report* results.

For example, if you ask people their weight, you will get their *desired* weight instead of their actual weight.

MISLEADING GRAPHS



Salaries of People with Bachelor's Degrees and with High School Diplomas

PICTOGRAPHS

Double the length, width, and height of a cube, and the volume increases by a factor of eight.

PERCENTAGES

Remember a 100% increase in a quantity means that it has doubled.

EXAMPLE: Suppose you are considering purchasing an item that originally cost $100. Store A reduced the price by 5% last week and has reduced it by 5% again this week. Store B has reduced the price by 10% this week. Which store has the better deal?

SELF-INTEREST STUDY

A **self-interest study** is a study which is sponsored by parties with an interest to promote.

EXAMPLE: Texas Instruments does a study and concludes that students prefer to use graphing calculators over computer software for mathematics classes.

PARTIAL PICTURES

A **partial picture** is where all the information is not presented.

EXAMPLE: "Ninety percent of all our cars sold in this country in the last 10 years are still on the road."

It was not disclosed that 90% of cars the advertiser sold were sold actually sold in the *last three years*.