

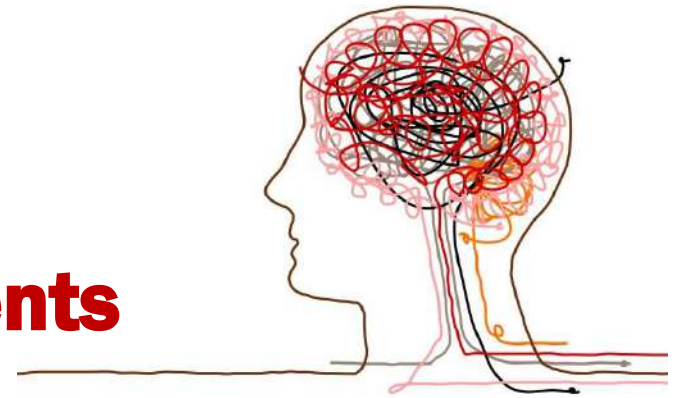
1.3

Statistical Reasoning

Appendix A (pages 605+)



Table of Contents



4	<u>Statistics</u>
5	<u>Descriptive Statistics</u>
17	<u>Inferential Statistics</u>

Fact or Falsehood?

Descriptive statistics are ways of describing data.

true

Two variables are negatively correlated if they change in opposite directions.

true

The median is the most frequently occurring score in a distribution.

false

The third variable problem is that all scientific studies must involve at least three variables.

false

Testing positive for measles when you don't actually have measles is an example of a Type I error.

true

Statistics

A branch of mathematics used by researchers to organize, summarize, and interpret data

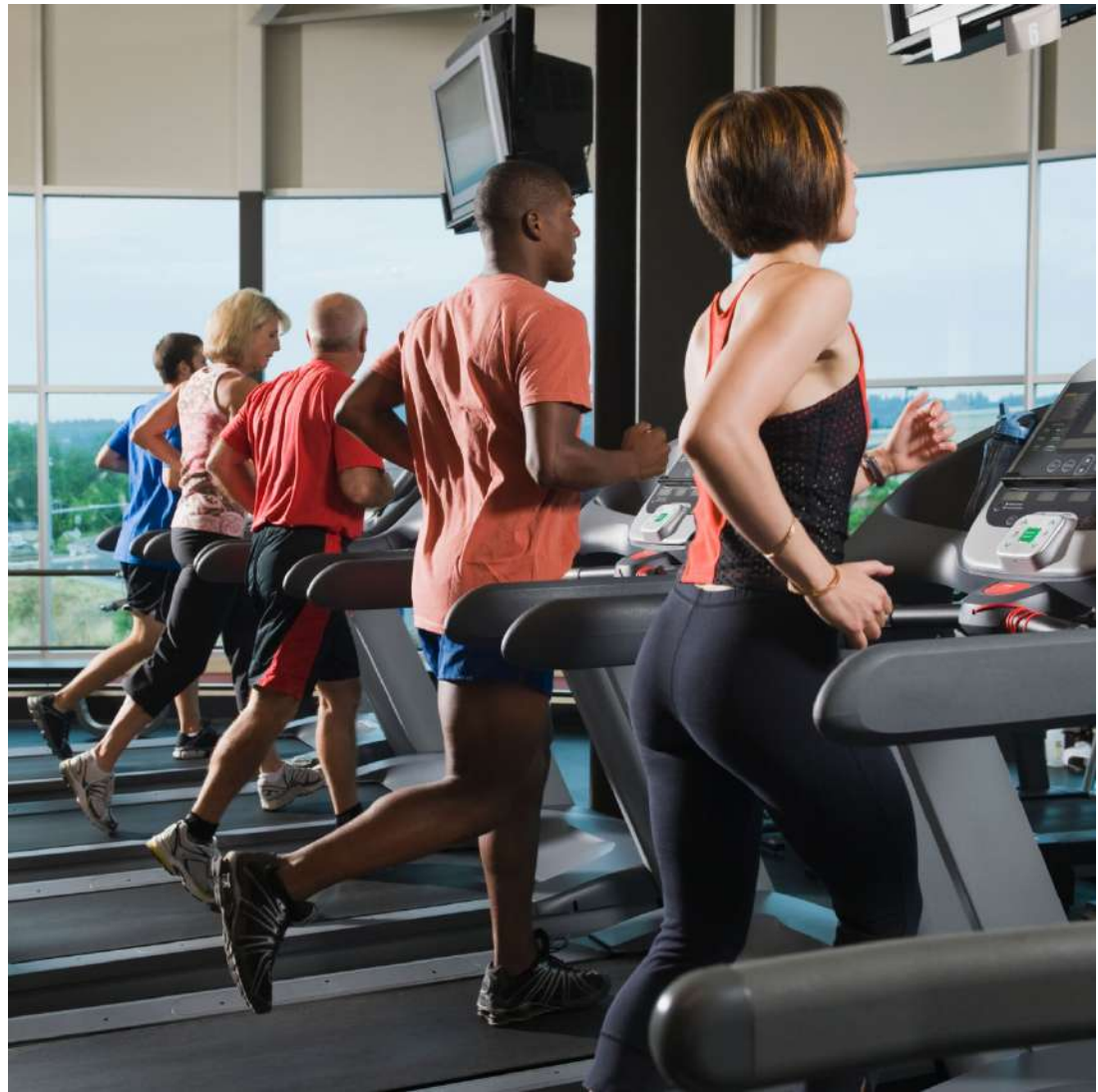
- We use statistics both to *describe* data and as a basis for *inferring* information from data.
- Statistics are often presented in a graphic or visual form.



Keeweeboy/Dreamstime.com

Descriptive Statistics

Mathematical methods used to organize and summarize, or *describe* data



Tetera images/Alamy

Q: How many hours do you exercise each week?

Descriptive Statistics

Frequency Distribution

A summary of how often various scores occur in a sample of scores. Score values are arranged in order of magnitude, and the number of occurrences is recorded.

A frequency polygon is a graph that represents a frequency distribution. It is constructed by connecting the midpoints of the tops of the bars in a bar chart. It is also known as a line graph.

Q: How many hours do you exercise each week?

Table A.1

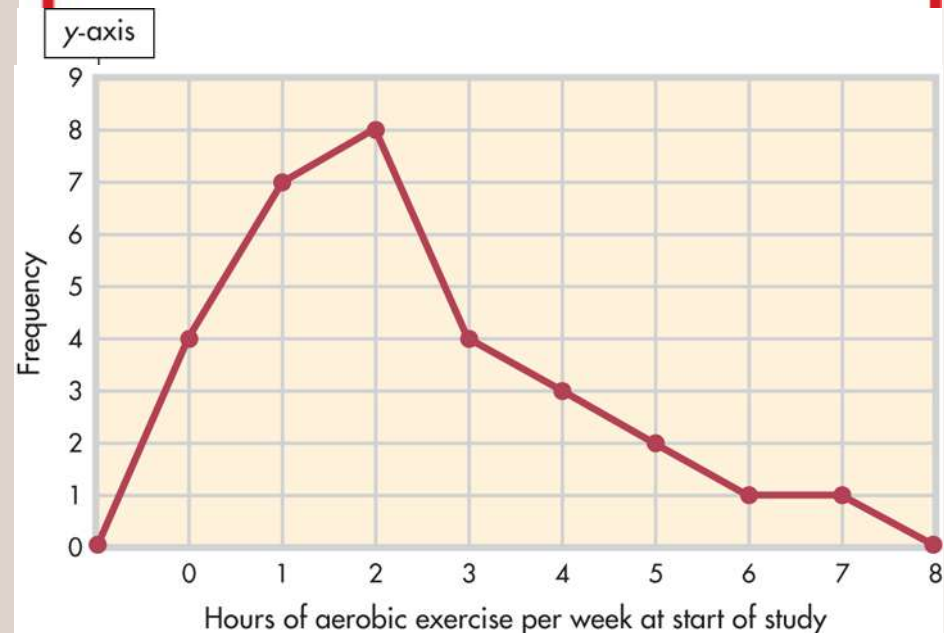
A Frequency Distribution Table

Hours of
Aerobic Exercise
per Week

Frequency

0

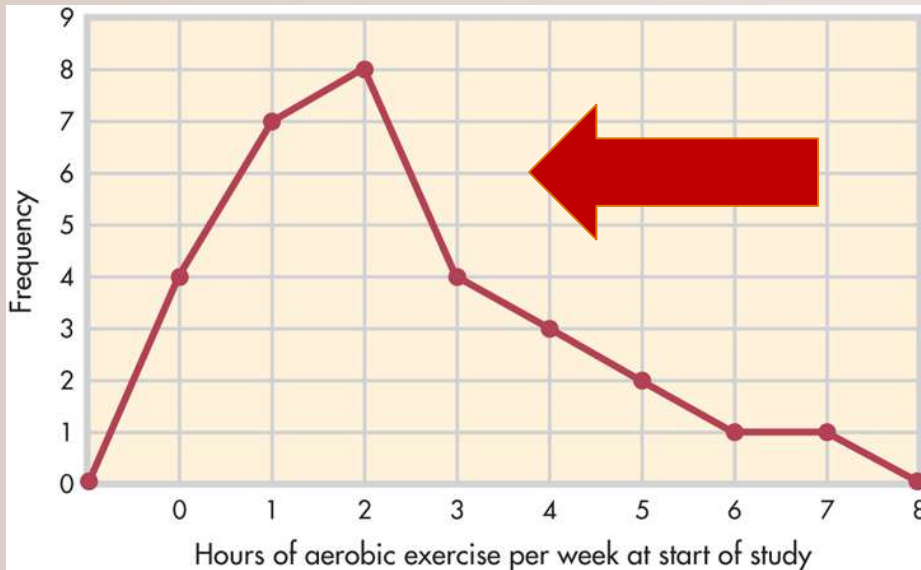
4



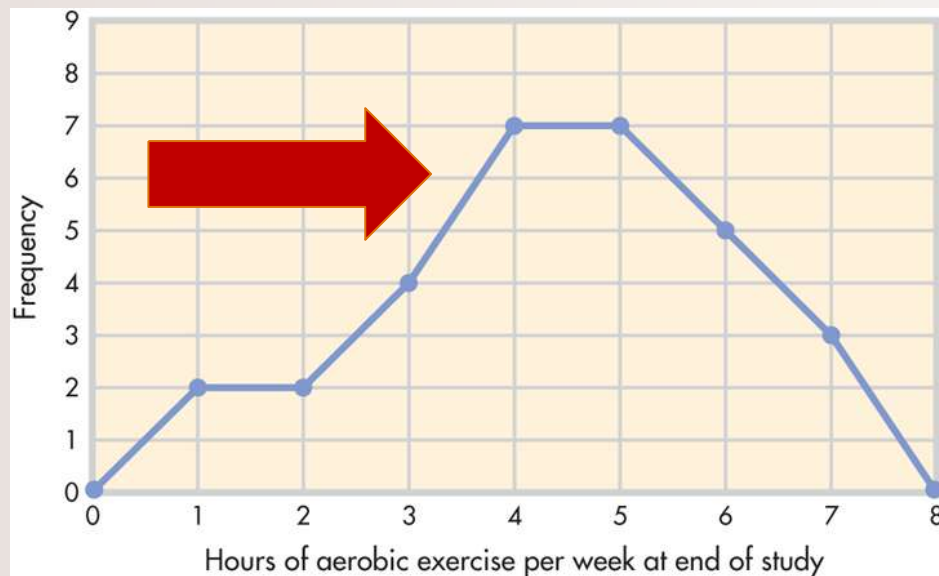
Descriptive Statistics

Frequency polygons quickly show if the shape of a distribution is skewed, or asymmetrical. In a skewed distribution, more scores occur on one side of the distribution than the other.

This polygon shows a positive skew: more scores are at the lower end of the scale.



This polygon shows a negative skew: more scores are at the higher end of the scale.



Descriptive Statistics

Measure of Central Tendency

A single number that presents information about the “center” of a frequency distribution.

Useful for putting information about a distribution into a more compact form, summarizing the distribution with a single score that is typical in some way



Descriptive Statistics: Measure of Central Tendency

Q: How many hours do you exercise each week?

20 people respond: 0, 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4 (hours)

Mode—The most frequently occurring score in the distribution.

0, 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4

mode = 2

Median—The score that divides a frequency distribution exactly in half.
The same number of scores fall above and below the median.

0, 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4



median = 2

Mean—The sum of a set of scores in a distribution divided by the number of scores.
The mean is also known as the average.

0, 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4 (hours)

mean = 1.5

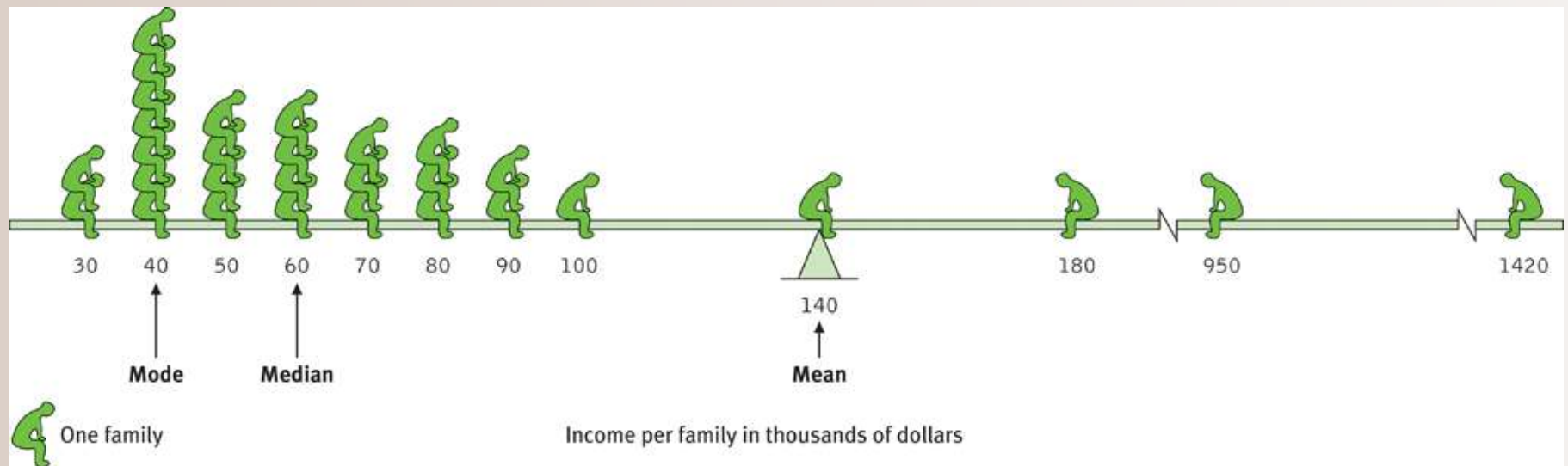
20 respondents

Descriptive Statistics:

Measure of Central Tendency

Analyze Carefully!

Note how describing the mean, median, and mode in isolation in the distribution below can lead to misleading conclusions!



Descriptive Statistics

Measure of Variability

- **Range** – the highest score in the distribution minus the lowest score, a simple measure of variability.
- **Standard Deviation** – the square root of the sum of the squared deviations from the mean divided by the number of scores in the distribution.

The larger the standard deviation, the more spread out are the scores in a distribution.

Table A.2

Calculating the Standard Deviation

Weight X	Mean $\bar{X}$	Weight – Mean $X - \bar{X}$	(Weight – Mean) Squared $(X - \bar{X})^2$
155	124	31	961
149	124	25	625
142	124	18	324
138	124	14	196
134	124	10	100
131	124	7	49
127	124	3	9
125	124	1	1
120	124	-4	16
115	124	-9	81
112	124	-12	144
110	124	-14	196
105	124	-19	361
102	124	-22	484
95	124	-29	841
Sum (Σ) = 1,860		$\Sigma = 0$	$\Sigma = 4,388$
Mean ($\bar{X}$) = 124			

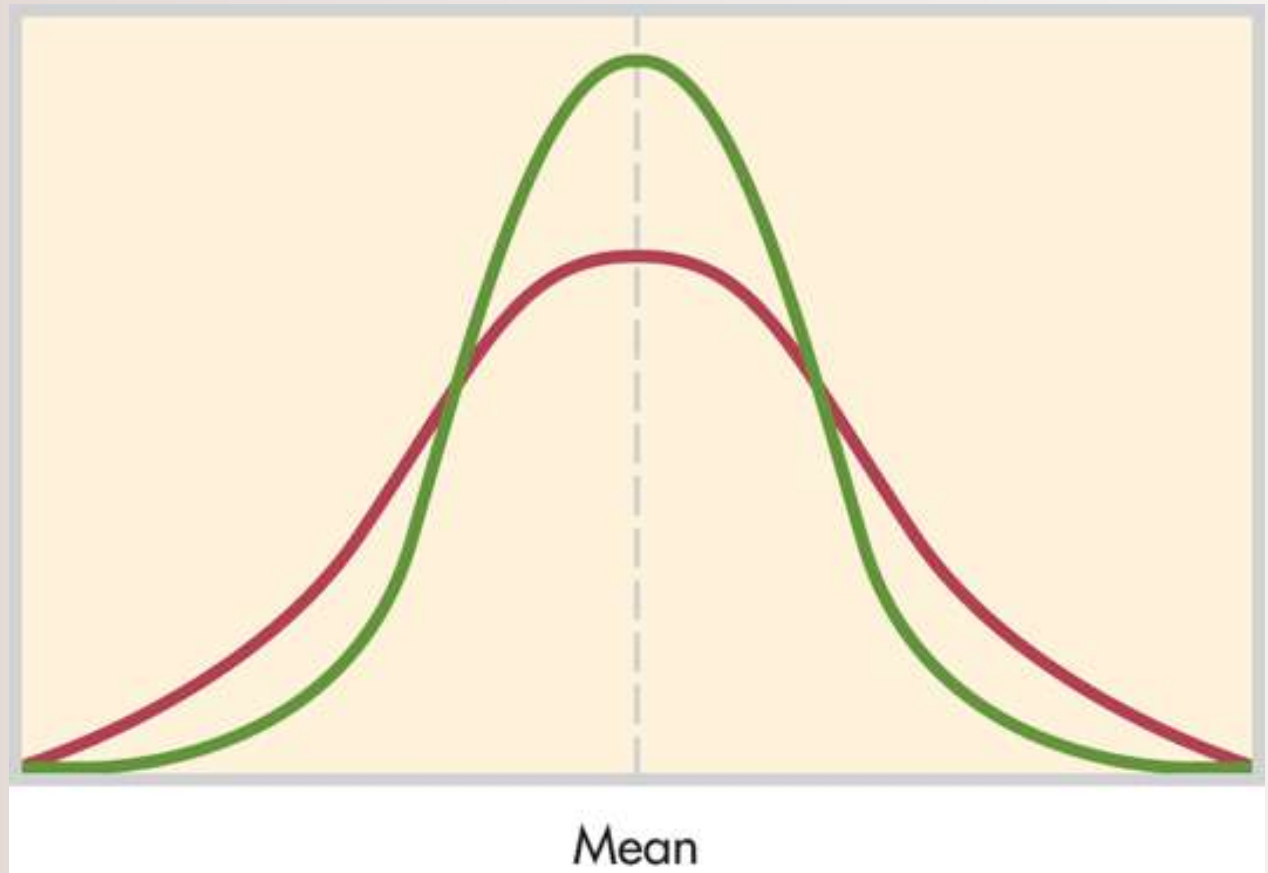
$$SD = \sqrt{\frac{\Sigma(X - \bar{X})^2}{N}} = \sqrt{\frac{4,388}{15}} = 17.10$$

Descriptive Statistics

Variability

Information about the *spread* of the scores in a distribution

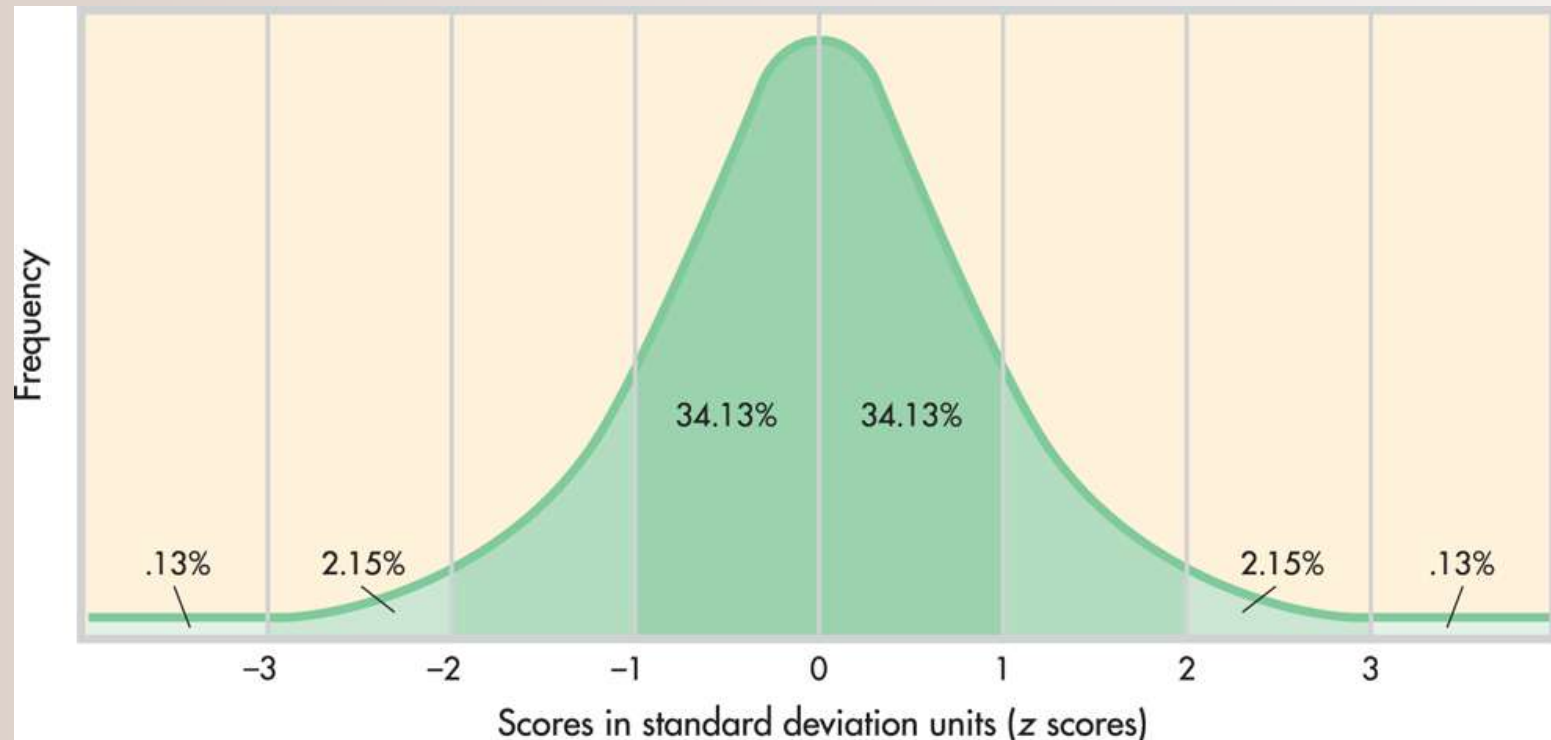
These distributions have the same mean but different variability—the scores are spread out differently.



Descriptive Statistics:

Measuring Variability – The Normal Curve

- A theoretical distribution that is symmetrical with the mean, median, and mode all falling at the exact middle of the distribution
- Sometimes called a “bell curve” or a standard normal curve/distribution



Descriptive Statistics

Correlation: The Relationship Between Two Variables

Positive Correlation:

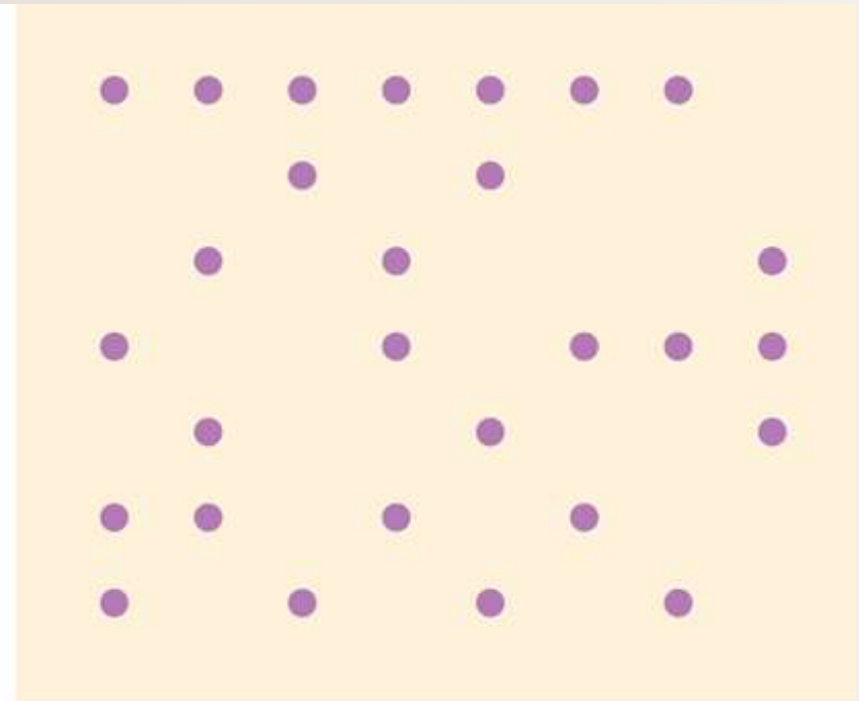
Two variables change systematically in the same direction, either increasing or decreasing together.

Negative Correlation:

Two variables change systematically in the **opposite** directions, one increasing as the other decreases.

A scatter plot with points that seem to fall randomly indicates no correlation between the variables.

Level of coping



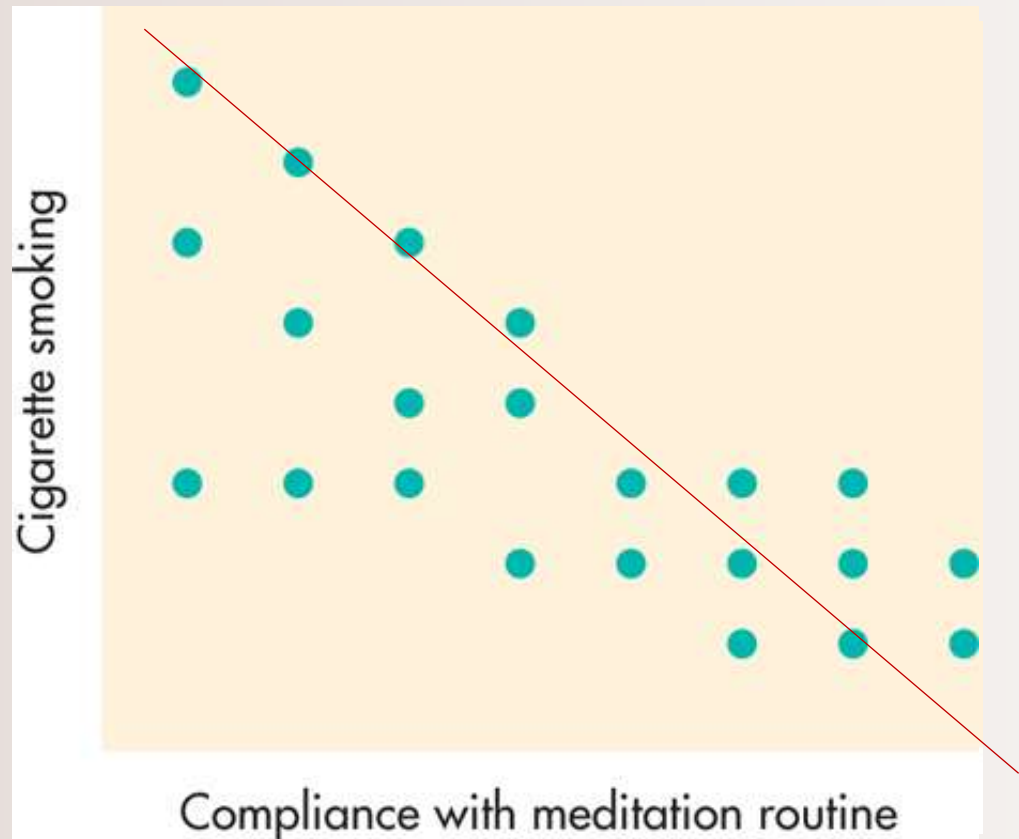
Compliance with
aerobic exercise program

Descriptive Statistics

The Correlation Coefficient

The statistical measure of the relationship between two variables. The coefficient ranges from +1.00 to -1.00.

A -1.00 correlation, or a perfect negative correlation, would indicate that the frequency of the yoga sessions were always accompanied by high levels of energy, and vice versa. A +1.00 correlation, or a perfect positive correlation, would indicate that the frequency of the yoga sessions were always accompanied by high levels of energy, and vice versa.

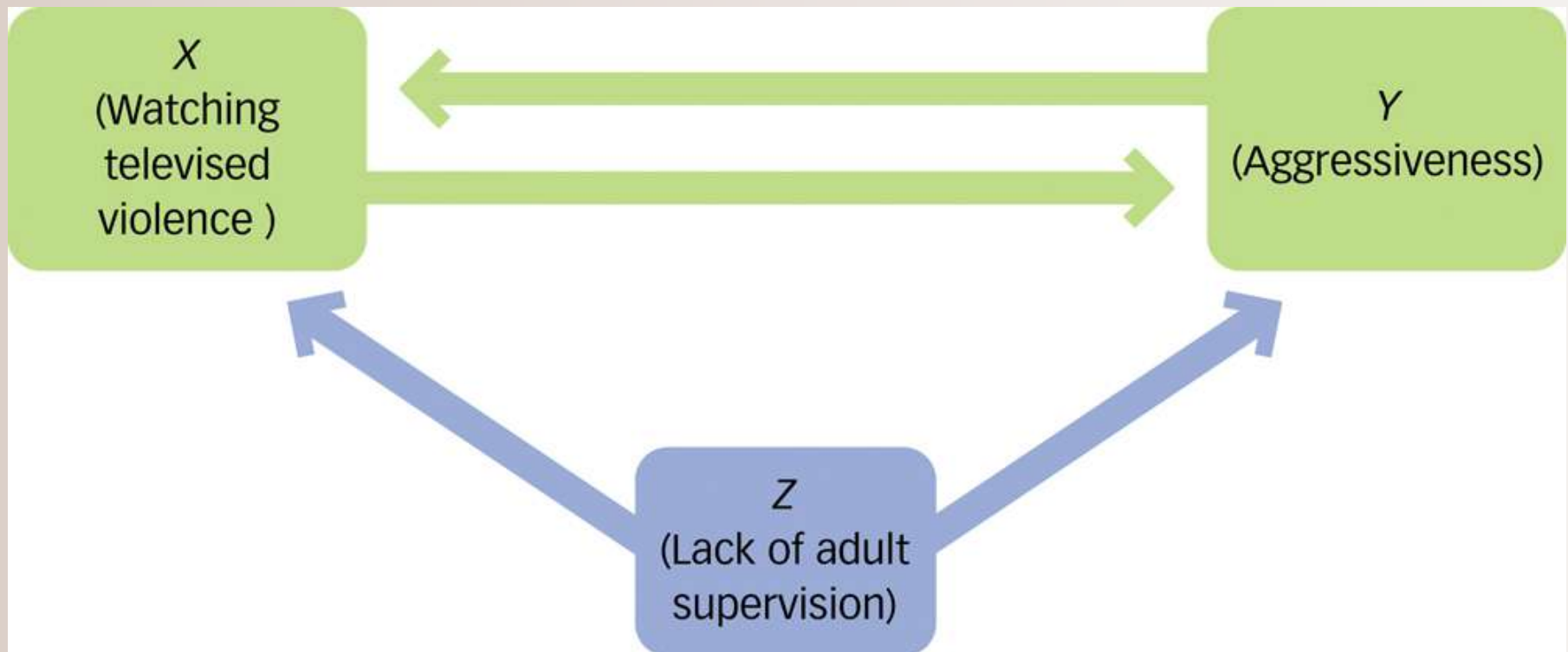


Descriptive Statistics

Correlation and Causality

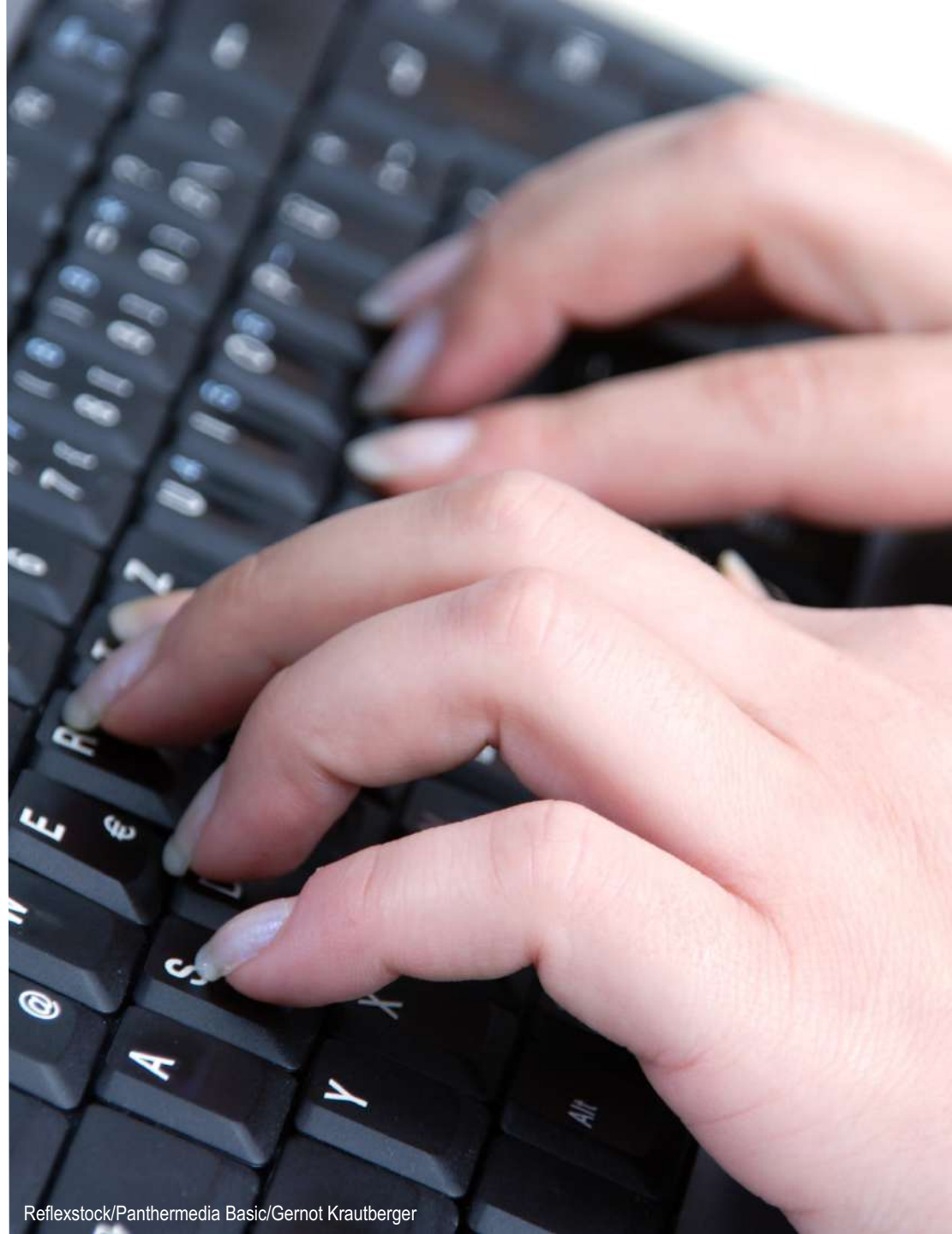
**Correlation does not
equal causation!**

Q: What third variable might cause both aggression and the watching of violent programs?



Inferential Statistics

- Mathematical methods used to draw conclusions about a population based on a sample of data
- In psychology, inferential statistics are used to determine the likelihood of a study's outcome being due to chance or to real differences.



Inferential Statistics

The ability to make inferential statistics and the scope of the inferences depend on the **population** (a complete set of something: people, nonhuman animals, objects, events) and **sample** (a subset of the population) of the subject under study.

The students in your class form the population of students studying psychology in this class, but they only form a sample of all the students at your institution.

All the doctors in the United States form the population of the doctors in the United States, but they only form a sample of the doctors in the world.

Information about a sample can be used to make inferences about the population.



Inferential Statistics

Inferential Errors

Type I error: Erroneously concluding that study results are significant



The data from a lie detector test indicate that a suspect is lying when he is actually telling the truth.

Type II error: Failing to find a significant effect that does, in fact, exist



The data from a lie detector test indicate that a suspect is telling the truth when he is actually lying.

Inferential Statistics

Generalizing from Samples

- Representative samples are better than biased samples.
- Less-variable observations are more reliable than those that are more variable.
- More cases are better than fewer.
- Look for statistical significance. The difference between observations is probably not due to chance variation between the samples.

How to Avoid Inferential Errors



Inferential Statistics

Paying Attention to Methodology

- In cross-sectional studies, different groups are studied at the same time.
- In longitudinal studies, the same group is studied over time.
- Cross-sectional studies show the effects of time and environment on many different individuals.
- Longitudinal studies show these effects on the same people.

How to Avoid Inferential Errors

