# Chapter 1
## *Introduction to Data Analysis in an Evidence-Based Practice Environment*

1.1. Statistical skills can play an important role in nursing because they help nurses to:
a. Calculate appropriate doses and clinical measurements
b. Generate clinical questions
*c. Evaluate and generate research evidence for nursing practice
d. Make better use of computers and the Internet

1.2. In the context of a quantitative study, a concept is called a(n):
a. Operational definition
*b. Variable
c. Statistic
d. Parameter

1.3. An example of a variable is:
*a. Systolic blood pressure
b. Pi (π)
c. 52.5 kilograms
d. Number of seconds in a minute

1.4. An example of a datum is:
a. Systolic blood pressure
b. Pi (π)
*c. 52.5 kilograms
d. Number of seconds in a minute

1.5. Which of the following is *not* a component of a research question?
a. An independent variable
b. A population
*c. A sample
d. A dependent variable

1.6. Identify the dependent variable in the following: In elderly men, what is the effect of chronic fatigue on level of depression?
a. Age
b. Sex
c. Chronic fatigue
*d. Depression

1.7. Which of the following is a continuous variable?
a. Number of pages in a book
*b. Age at death
c. Falls during hospitalization
d. Number of times married

1.8. Measurement is the assignment of numbers to characteristics of people or objects according to specified _____ . (Fill in the blank.)

*a. Rules
b. Definitions
c. Concepts
d. Parameters

1.9. The measurement level that classifies attributes, indicates magnitude, and has equal intervals between values, but does not have a rational zero, is:

a. Nominal
b. Ordinal
*c. Interval
d. Ratio

1.10. The measurement level that is sometimes called *categorical* or *qualitative* is:

*a. Nominal
b. Ordinal
c. Interval
d. Ratio

1.11. It is not meaningful to calculate an arithmetic average with data from which of the following?

a. Nominal measures
b. Ordinal measures
*c. Nominal and ordinal measures
d. All measures can be meaningfully averaged.

1.12. Degree of pain measured as *none, a little*, or *a lot* is measured on which of the following scales?

a. Nominal
*b. Ordinal
c. Interval
d. Ratio

1.13. Body temperature is measured on which of the following scales?

a. Nominal
b. Ordinal
*c. Interval
d. Ratio

1.14. Type of birth (vaginal or cesarean) is measured on the:

*a. Nominal scale
b. Ordinal scale
c. Interval scale
d. Ratio scale

1.15. Which of the following is a ratio-level measure?
*a. Dietary cholesterol intake (mg)
b. Cognitive impairment on a 50-item scale
c. Pain on a 10-point scale
d. Military rank

1.16. Ratio-level measures are different than any other level by virtue of which property?
a. Classification
b. Equal intervals between values
*c. A true, rational zero
d. Indication of magnitude

1.17. Which level of measurement communicates the most information?
a. Nominal
b. Ordinal
c. Interval
*d. Ratio

1.18. Researchers typically collect data from a _____ and hope to generalize their results to a _____. (Fill in the blanks.)
a. Population, sample
b. Statistic, parameter
c. Sample, statistic
*d. Sample, population

1.19. If the average amount of sleep for all people in the United States was 7.6 hours per night, this average would be a(n) _____ of the population of U.S. residents. (Fill in the blank.)
a. Variable
*b. Parameter
c. Statistic
d. Datum

1.20. If a nurse researcher measured the anxiety level of 100 hospitalized children, the children's average score on an anxiety scale would be a(n):
a. Variable
b. Parameter
*c. Statistic
d. Operational definition

1.21. Statistical methods that are used to draw conclusions about a population are called:
*a. Inferential statistics
b. Descriptive statistics
c. Univariate statistics
d. Multivariate statistics

# Chapter 2
## *Frequency Distributions: Tabulating and Displaying Data*

2.1. A major purpose of constructing a frequency distribution with sample data is to:
a. Estimate a population parameter
b. Test a research hypothesis
*c. Get an organized view of an entire set of scores
d. Get experience with statistical software

2.2. In a frequency distribution, the two key informational components are:
*a. Score values (*X*), frequencies (*f*)
b. A horizontal (X) axis, a vertical (Y) axis
c. Frequencies (*f*), percentages (%)
d. Participant ID number (*id)*, score values (*X)*

2.3. In a frequency distribution, which of the following is true?
a. $\Sigma N = \%$
b. $\Sigma N = f$
c. $\Sigma f = \%$
*d. $\Sigma f = N$

2.4. In the equation $\Sigma \% = 100.0$, the symbol $\Sigma$ signifies:
a. A percentage
*b. The sum of
c. A data value
d. A frequency

2.5. In a frequency distribution, percentages are sometimes called:
a. Proportions
b. Relative proportions
*c. Relative frequencies
d. Cumulative proportions

2.6. Data for which of the following variables is most likely to be presented in a grouped frequency distribution?
a. Nursing specialty area
*b. Daily cholesterol intake
c. Number of abortions
d. Number of pets owned

2.7. The level of measurement for data appropriately presented in a bar graph is:
a. Interval or ratio
b. Nominal only
c. Interval only
*d. Nominal or ordinal

2.8. In a frequency distribution graph, frequencies are typically presented on the ____ and data values are presented on the _____. (Fill in the blanks.)

*a. *Y* axis, *X* axis
b. *X* axis, *Y* axis
c. *f* axis, *N* axis
d. *N* axis, *f* axis

2.9. Which of the following sets of data is *not* unimodal?

*a. 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 4, 5, 5, 5, 5, 5, 5, 5, 5
b. 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4
c. 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 3, 3, 4, 5
d. 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 5, 5, 5, 6, 6, 7, 7, 8, 8, 9, 9

2.10. Which of the following variables is most likely to be negatively skewed in a general population?

a. Number of times arrested
*b. Age at retirement
c. Number of times married
d. Age at birth

2.11. A normal distribution is *not:*

a. Skewed
b. Leptokurtic
c. Platykurtic
*d. All of the above

2.12. A wild code is*:*

*a. A value that is impossible given the coding scheme
b. An outlier or high value
c. A code for which there is a very low frequency
d. A code for which there is a very high frequency

The next eight questions pertain to the following table (Table 2):

**Table 2**

| Number of Pregnancies of Study Participants | Frequency | Percentage | Cumulative Percentage |
|---|---|---|---|
| 0 | 24 | 11.1 | 11.1 |
| 1 | 29 | 13.5 | 24.6 |
| 2 | 78 | 36.3 | 60.9 |
| 3 | 46 | 21.4 | 82.3 |
| 4 | 22 | 10.2 | 92.5 |
| 5 | 11 | 5.1 | 97.6 |
| 6 | 4 | 1.9 | 99.5 |
| 7 | 1 | 0.4 | 100.0 |
| Total | 215 | 100.0 | |

2.13 In Table 2, the variable is _____ and the measurement level is _____. (Fill in the blanks.)

a. Discrete, interval
*b. Discrete, ratio
c. Continuous, interval
d. Continuous, ratio

2.14. Table 2 is an example of a:

*a. Frequency distribution
b. Grouped frequency distribution
c. Class interval
d. Data matrix

2.15. In Table 2, the value of *N* is:

a. 24
b. 100.0
*c. 215
d. 7

2.16. In Table 2, the cumulative relative frequency for five or fewer pregnancies is:

a. 210
b. 199
c. 92.5
*d. 97.6

2.17. The best way to graph information in Table 2 would be to construct:

*a. A histogram
b. A pie chart
c. A bar graph
d. Either a pie chart or a bar graph

2.18. In Table 2, the distribution of data would be described as:

a. Symmetric
*b. Positively skewed
c. Negatively skewed
d. It cannot be determined.

2.19. In Table 2, the distribution of data would be described as:

*a. Unimodal
b. Bimodal
c. Multimodal
d. It cannot be determined.

2.20. In Table 2, the most likely number to be an outlier is:

a. 0
b. 1
*c. 7

d. 24

# Chapter 3
## *Central Tendency, Variability, and Relative Standing*

3.1. A distribution of data values can be described in terms of all of the following characteristics *except:*
a. Central tendency
b. Variability
*c. Relative standing
d. Shape

3.2. Central tendency indexes are all of the following *except* which of the following statements?
a. They are descriptive statistics.
*b. They summarize how dispersed a set of scores is.
c. They provide information about a value around which scores cluster.
d. They are appropriate for interval- and ratio-level measures.

3.3. In the following distribution (10 11 12 13 14 15 15 15 15) the mode is:
a. 11
b. 12
c. 14
*d. 15

3.4. In the following distribution (10 11 12 13 14 15 15 15 15) the median is:
a. 11
b. 12
*c. 14
d. 15

3.5. The median is all of the following *except*:
a. The $50^{th}$ percentile
b. The point that divides a distribution in half
c. $Q_2$
*d. The most popular score in the distribution

3.6. For which of the following set of numbers are the mean, median, and mode the same value?
*a. 1 2 3 3 4 4 4 4 4 5 5 6 7
b. 1 1 2 2 3 3 4 4 5 5 6 6 7 7
c. 1 1 1 2 3 3 4 4 5 5 6 7 7 7
d. All of the above

3.7. In which type of distribution is the mean a higher value than the median or mode?
a. A leptokurtic distribution
*b. A positively skewed distribution

c. A negatively skewed distribution
d. A normal distribution

3.8. If there are outliers at either end of a distribution that is symmetric, a researcher might:
*a. Calculate a trimmed mean
b. Report the median rather than the mean
c. Report the mode rather than the mean
d. Omit the variable from further analyses

3.9. Which of the following indexes of dispersion is *not* in the original units of measurement of the variable?
a. Range
b. Interquartile range
c. Standard deviation
*d. Variance

3.10. Which of the following indexes of dispersion tends to be least stable—most likely to fluctuate from one sample to another from the same population?
*a. Range
b. *IQR*
c. Standard deviation
d. Variance

3.11. Which of the following indexes involves the calculation of deviation scores ($x$)?
a. Range
b. *IQR*
*c. *SD*
d. *M*

3.12. Which of the following indexes involves the calculation of percentiles?
a. *z*
*b. *IQR*
c. *SD*
d. *M*

3.13. Which of the following statistical symbols does not belong with the others?
a. *SD*
b. *IQR*
c. *M*
*d. μ

3.14. What percentage of cases for a normally distributed variable lies within 1 *SD* above and below the mean?
a. 34%
b. 50%
*c. 68%

d. 95%

3.15. In calculating standard scores, which two descriptive statistics are needed?
a. Median, *IQR*
b. Median, percentiles
c. Mean, Range
*d. Mean, *SD*

3.16. A *z* score of 0.00 corresponds to an original score that:
a. Could not be used in the calculation of the mean
*b. Is the same as the mean in the original distribution
c. Is the lowest score in the original distribution
d. Is an outlier

3.17. A *z* score of -1.00 corresponds approximately to a score for a normally distributed variable that is at the:
a. 1st percentile
b. 10th percentile
*c. 16th percentile
d. 84th percentile

3.18. An extreme outlier is:
a. More than 3 *SD*s above the mean
b. Equivalent to a *z* score of -3.0 or lower, or +3.0 or higher
c. More than three times the value of the mean
*d. More than 3 times the *IQR*, below $Q_1$ or above $Q_3$

3.19. In a boxplot, information about a distribution is depicted in terms of:
*a. Percentiles
b. Standard deviation units
c. *z* scores
d. *T* scores

3.20. The number 100 can always be thought of as:
a. A mean of a distribution when the *SD* is 15
b. A value equivalent to the 10th percentile
*c. A number whose real limits are 99.5 and 100.5
d. An outlier