Problem Solutions

CHAPTER 1

1. Descriptive statistics are procedures used to summarize, organize, and make sense of a set of scores or observations, whereas inferential statistics are procedures that allow researchers to infer or generalize observations made with samples to the larger population from which they were selected.
2. Data describe a set of measurements (made up of raw scores); a raw score describes individual measurements.
3. Samples are selected from populations of interest. Hence, samples consist of a portion of individuals in a population of interest.
4. Experimental, quasi-experimental, and correlational research methods.
5. (a) Independent variable, (b) Dependent variable.
6. The four scales of measurement are nominal, ordinal, interval, and ratio. Ratio scale measurements are the most informative.
7. Yes. Nominal values are often coded (i.e., converted to numeric values) when entered into statistical software.
8. Interval variables DO NOT have a true zero, and ratio variables DO have a true zero.
9. Amount; class.
10. Continuous and discrete numbers.
11. (a) Descriptive statistics. (b) Inferential statistics. (c) Descriptive statistics.
12. (a) False. (b) True. (c) True.
13. No, it is not necessary to make inferences, because ALL individuals in the population were observed.
14. The statistics class has a *population* of 25 students enrolled, but a *sample* of only 23 students attended.
15. By definition, samples will always be smaller than the population. Samples constitute a subset (or a smaller set of values) than those in the population.
16. An experimental research method because the researcher claims to have demonstrated *cause*.
17. A correlational research method because pairs of scores (height, income) were measured and compared for each participant.
18. (a) Quasi-independent variable. (b) Quasi-independent variable. (c) Independent variable. (d) Independent variable. (e) Quasi-independent variable. (f) Independent variable.
19. (a) Sleeping pill (real or fake). (b) Time spent sleeping.

20. (a) Cocaine use (dependent vs. inexperienced). (b) Impulsive behavior.
21. (a) Nominal. (b) Yes, it is appropriate to numerically code "months" because it is a nominal scale variable.
22. Nominal, ordinal, interval, and ratio.
23. The main disadvantage of measuring qualitative data is that the data are on a nominal scale, which limits the types of conclusions that researchers can draw. For this reason, quantitative variables are more often measured because they are more informative in terms of order, differences, and ratios.
24. (a) Qualitative. (b) Quantitative. (c) Quantitative. (d) Qualitative.
25. (a) Continuous. (b) Discrete. (c) Continuous. (d) Discrete.
26.

Variable	Types of Data Qualitative vs. Quantitative	Types of Number Continuous vs. Discrete	Type of Measurement
Gender	Qualitative	Discrete	Nominal
Seasons	Qualitative	Discrete	Nominal
Time of day	Quantitative	Continuous	Ratio
Rating scale score	Quantitative	Discrete	Interval
Movie ratings (1 to 4 stars)	Quantitative	Discrete	Ordinal
Number of students in your class	Quantitative	Discrete	Ratio
Temperature (degrees Fahrenheit)	Quantitative	Continuous	Interval
Time (in minutes) to prepare dinner	Quantitative	Continuous	Ratio
Position standing in line	Quantitative	Discrete	Ordinal

27. (a) Descriptive statistics. (b) Overall, gun ownership really has not changed in the last 40 years with the percent of gun owners in 1972 being identical to the percent of gun owners in 2012.
28. An operational definition.
29. No, an experimental method is not possible because the variable (gender) is a preexisting variable; it is a quasi-independent variable.
30. (a) Continuous. (b) Quantitative. (c) Ratio scale.
31. The 91,373 nineteen-year-old-men likely represent a sample of all 19-year-old men in a much larger population.
32. Equidistant scales. Variables on an interval scale have equidistant scales, meaning that differences on this scale are informative.

CHAPTER 2

1. Step 1: Find the real range. Step 2: Find the interval width. Step 3: Construct the frequency distribution.
2. Grouped data are distributed in intervals; ungrouped data are not.
3. (a) No. (b) Yes.
4. To ensure that a single score cannot be counted in more than one interval.
5. A percentile rank.
6. Ungrouped data sets with only a few different scores, and qualitative or categorical variables.
7. Rule 1: A vertical rectangle represents each interval, and the height of the rectangle equals the frequency recorded for each interval. Rule 2: The base of each rectangle begins and ends at the upper and lower boundaries of each interval. Rule 3: Each rectangle touches adjacent rectangles at the boundaries of each interval.
8. Midpoint; Upper boundary.
9. When the data are discrete. Histograms are only used with continuous data.
10. The *x*-axis.
11. (a)

Intervals	*f(x)*
18–22	5
13–17	3
8–12	7
3–7	5

(b) Interval: 8–12.

12. (a)

Classes	*f(x)*
L	9
C	16
R	5

(b) Yes, the rat did press the center lever the most.

13. The intervals overlap at the upper and lower boundaries, which might lead to some scores being counted in more than one interval.
14. Three errors are (1) the intervals overlap, (2) the class width for each interval is not equal, and (3) the distribution includes an open interval.

15. The lower class boundaries are 35, 46, 57, and 68.
16. The upper class boundaries are 3, 6, 9, 12, 15, and 18.
17. No, the data should remain ungrouped because the data are categorical.
18. The interval width for each interval is 3.
19. (a)

Number of Dreams	Percentile Rank
4	100%
3	88%
2	60%
1	24%
0	10%

(b) Two Dreams

20. Sixty children qualify for the new cognitive-behavioral therapy.
21. (a)

11	0	4	4	5	7	9	
15	2	3	3	5	8	9	9
22	0	1	3	3	5	6	7

(b)

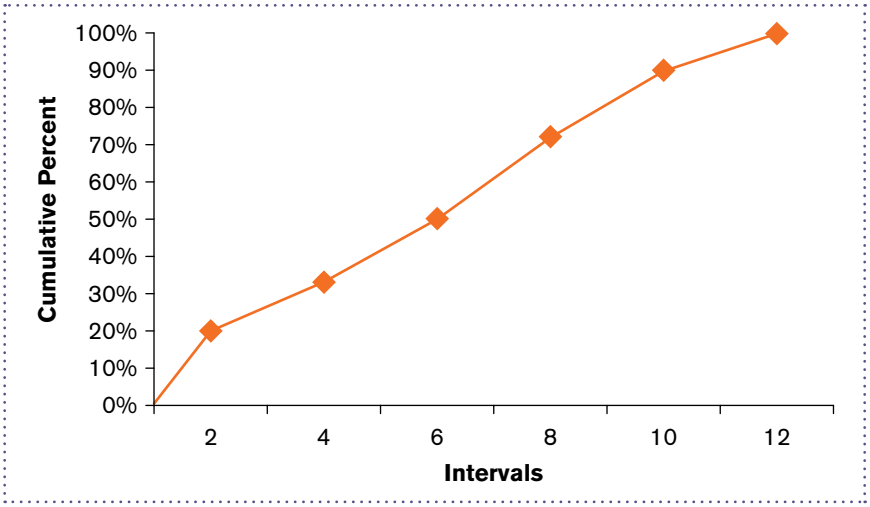
1	10	14	14	15	17	19	52	53	53	55	58	59	59
2	20	21	23	23	25	26	27						

22. (a) 13, 17, 32, 32, 32, 40, 41, 49, 66, and 68.

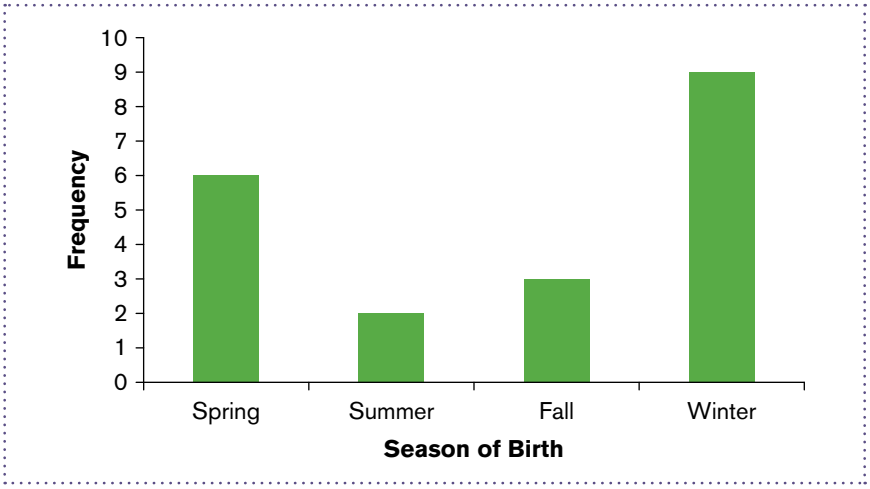
(b) Ten scores.

23. (a) A bar chart or pie chart if distributing the letters grades along the *x*-axis. A histogram or frequency polygon if distributing the numeric values on a 4.0 grading scale. (b) A bar chart or pie chart. (c) A bar chart or pie chart distributing the frequencies for each behavioral therapy. (d) A histogram or frequency polygon.
24. (a) Histogram. (b) Bar chart. (c) Histogram. (d) Bar chart.
25. A = 8, B = 3, C = 12
26. (a) A = 78, B = 86, C = 68, D = 13. (b) Yes, this was a difficult test because half the class would fail.
27. (a) 35 students. (b) 6 students. (c) 15 students.

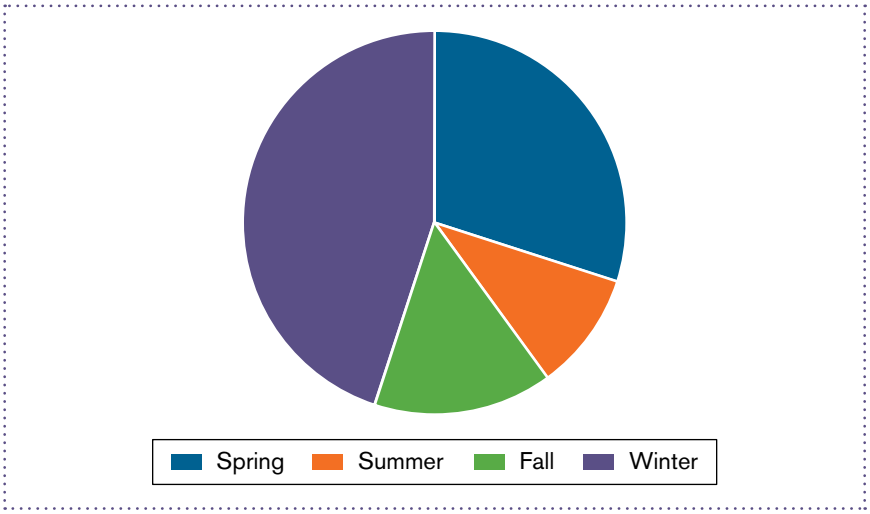
28.


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Cumulative Percent
2
4
6
8
10
12
Intervals

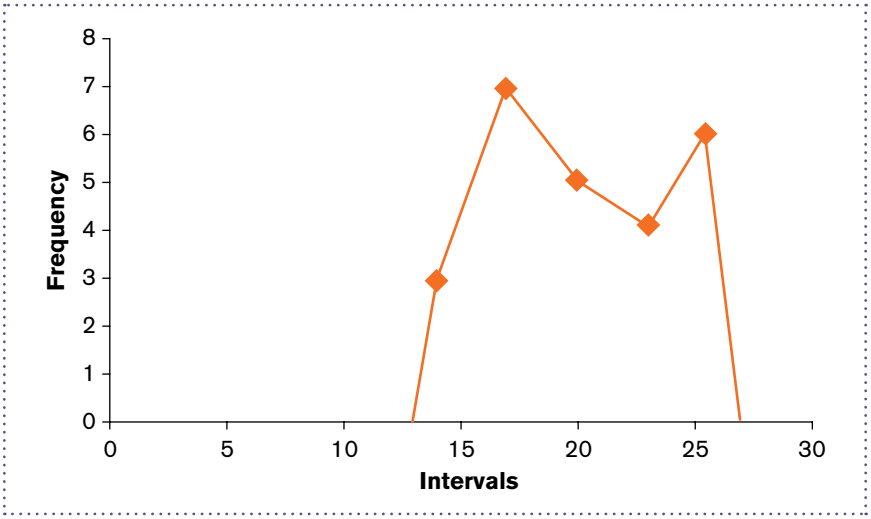
29. (a)


10
9
8
7
6
5
4
3
2
1
0
Frequency
Spring
Summer
Fall
Winter
Season of Birth

(b)


Spring
Summer
Fall
Winter

30.



31. The interval of 60-79 has the largest portion of students (20% of students fall in this interval).

32. The percentile point for the 50th percentile is 74.5 minutes.

33. (a) Real range = 56. (b) 30%.

34. (a) While more men earned a bachelor's degree in psychology in 1970–1971, women earn more than three times the number of bachelor's degrees in psychology as of 2005–2006. (b) Ungrouped data, because years are not distributed consecutively.

35. (a) Women. (b) Women.

36. (a) 31–50 age group. (b) Two problems with this table are that the interval width is not equal, and the table includes an open class (51+).

37. (a) A relative percent distribution. (b) About 576 Americans.

38. (a) A pictorial frequency distribution. (b) Yes, these data reflect the distribution of wealth by country, so it is meaningful to "map" the distribution of these data because the data are distributed by country.

CHAPTER 3

1. (a) N. (b) n. (c) μ. (d) M or $\bar{X}$.

2. Measures of central tendency are statistical measures used to locate a single score that is most representative or descriptive of all scores in a distribution.

3. The median is the score in the middle of a distribution. It is always at the center of a distribution.

4. A population mean is the mean for a set of scores in an entire population, whereas the sample mean is the mean for a sample, or subset of scores from a population.

5. Five characteristics of the mean are as follows: (1) Changing an existing score will change the mean; (2) adding a new score or completely removing an existing score will change the mean, unless that value equals the mean; (3) adding, subtracting, multiplying, or dividing each score in a distribution by a constant will cause the mean to change by that constant; (4) the sum of the differences of scores from their mean is zero; and (5) the sum of the squared differences of scores from their mean is minimal.
6. The weighted mean equals the arithmetic mean when the sample sizes or "weights" for a set of scores are the same or equal.
7. Data that are normally distributed on an interval or ratio scale of measurement.
8. Data that are skewed and ordinal data.
9. The mode is used with other measures of central tendency for any modal distribution and for nominal data.
10. (a) Median. (b) Mean. (c) Mean.
11. Mean = 4, median = 2, mode = 0.
12. (a) College students: mean = 25, median = 18, mode = 21. Parents: mean = 14, median = 18, mode = 21. (b) Because both distributions are skewed, the median would be the appropriate measure of central tendency. This might be misleading, though, because the median indicates that texting was the same between groups (the median was 18 in both samples), even though differences exist in regard to the mean.
13. The mean because the data are distributed normally.
14. The mean because the data are normally distributed, and the duration (in hours) is a ratio scale measure.
15. (a) Positively skewed distribution. (b) The median.
16. The median because the data are negatively skewed.
17. The mode because the data are nominal.
18. Bimodal distribution.
19. Weighted mean = 14.69.
20. Weighted mean = 3.65.
21. (a) The mean, because time is a ratio measurement. (b) The mode, because blood types are nominal data. (c) The median, because college rankings are ordinal data.
22. (a) M = 22 points. (b) M = 10 points. (c) M = 24 points. (d) M = 6 points.
23. M = 6 new friends.
24. The sum of the differences of scores from the mean is 0.
25. The mean will decrease.
26. The mean will increase.
27. (a) The mean will increase. (b) The mean will not change. (c) The mean will decrease.

Statistical Package for the Social Sciences

Applied Assessment Workbook

Answer Key

For instructor's use with:

Statistics for the Behavioral Sciences, 2e

SPSS in Focus: Entering and Defining Variables
Exercise 1.1 Answer Key

Enter data by column:

With regard to the SPSS exercise,, answer the following questions:

State whether you used the **data view** or **variable view** to complete each of the following steps:

Naming variables	variable view
Entering the values for each variable	data view

State the following values for the data you entered in SPSS:

The number of values entered (overall)	20
The number of values entered in each group	10
The number of groups	2

Enter data by row:

With regard to the SPSS exercise, answer the following questions:

State whether you used the **data view** or **variable view** to complete each of the following steps:

Naming variables	variable view
Coding variables	variable view
Entering the values for each variable	data view

State the following values for the data you entered in SPSS:

The number of values entered (overall)	20
The number of values entered in each group	10
The number of groups	2

SPSS in Focus: Frequency Distributions for Quantitative Data
Exercise 2.1 Answer Key

With regard to the SPSS exercise, answer the following questions:

State the dependent variable: Number of absences

State the following values (you can find these in the SPSS output table):

Total number of scores entered	100
The score at the 50th percentile	5
The frequency at or above the 80th percentile	22
The frequency at or below the 80th percentile	80

Explain why the frequencies at or above and at or below the 80th percentile do not sum to 100.

The frequencies at or above and at or below the 80th percentile do not sum to the total number of scores entered (N = 100), because the frequency at the 80th percentile (2) is counted in both percentile ranges.

Remember that a main goal for using frequency distributions is to simplify large data sets, which makes it easier to interpret research data. That being said, how would you characterize or interpret the data displayed in the SPSS output table you created?

The data indicate that half the students at this school were absent from school 5 or fewer times in a given school year, and one-fourth missed 0 or 1 day of school (25 out of 100 students missed only 0 or 1 day of school).

SPSS in Focus: Frequency Distributions for Categorical Data
Exercise 2.2 Answer Key

With regard to the SPSS exercise, answer the following questions:

Answer each of the following questions about coding data in SPSS.

Are categories displayed as numbers or words in data view? Numbers
Are categories displayed as numbers or words in the SPSS output table? Words

State the dependent variable: Texting efficiency (Always, Sometimes, Never)

State the following values (you can find these in the SPSS output table):

Total number of scores entered	45
The number of categories	3
The category in the 50th percentile	N

Remember that a main goal for using frequency distributions is to simplify large data sets, which makes it easier to interpret research data. That being said, how would you characterize or interpret the data displayed in the SPSS output table you created?

The largest frequency of texting in class was among those who never (N) looked at the keys. So those who efficiently text, were more often observed texting during class.

SPSS in Focus: Histograms, Bar Charts, and Pie Charts
Exercise 2.3 Answer Key

With regard to the SPSS exercise, answer the following questions:

For the histogram, state the:

Dependent variable	Amount of calories per meal
Scale of measurement of the dependent variable	Ratio

For the bar chart and pie chart, state the:

Dependent variable	Frequency in each health category
Scale of measurement of the dependent variable	Ratio

Remember that a main goal for using graphs is to simplify large data sets. This goal makes it easier to interpret research data. That being said, how would you characterize or interpret the data displayed in the SPSS output graphs you created?

Most meals (as measured by number of calories) were unhealthy. This is evident both in terms of the number of calories (shown in the histogram) and the frequency of meals in each health category (shown in the bar chart and pie chart).

For your own reference, state which display is easiest for you to read. Please pick only one. There is no right or wrong answer here. It is simply worth recognizing the type of graphical display that makes the most sense to you.

SUGGESTION FOR USING THIS QUESTION IN CLASS: You can record responses of students for this question and distribute it in a frequency distribution for categorical data, and then create a bar chart from it. When you show students the results in class, this often leads to rather interesting conversation and opinions with regards to the “best” way for displaying data. So this little exercise can reinforce previous lectures, as well as gauge student interest. This is something to consider for your classes if you find the time. This exercise typically requires about 10 to 15 minutes of class time; maybe more time when student participation is high.