CHAPTER 7— ANSWERS TO EXERCISES

1.

a. The estimate at the 90% confidence level is 22.82% to 23.18%. This means that there are 90 chances out of 100 that the confidence interval will contain the true population percentage of victims in the American population.

Due to the large sample size, we converted the proportions to percentages, subtracting from 100, rather than 1.

$$\text{Standard error} = \sqrt{\frac{(23)\,(100-23)}{160{,}040}} = 0.105 = 0.11$$

$$\begin{aligned}\text{Confidence interval} &= 23 \pm 1.65(0.11)\\ &= 23 \pm 0.18\\ &= 22.82 \text{ to } 23.18\end{aligned}$$

b. The true percentage of crime victims in the American population is somewhere between 22.72% and 23.28% based on the 99% confidence interval. There are 90 chances out of 100 that the confidence interval will contain the true population percentage of crime victims.

$$\begin{aligned}\text{Confidence interval} &= 23 \pm 2.58(0.11)\\ &= 23 \pm 0.28\\ &= 22.72 \text{ to } 23.28\end{aligned}$$

2.

a. For lower-class respondents:

$$s_{\bar{Y}} = \frac{s}{\sqrt{N}} = \frac{2.83}{\sqrt{122}} = 0.256 = 0.26$$

$$\begin{aligned}\text{Confidence interval} &= 12.11 \pm 1.96(0.26)\\ &= 12.11 \pm 0.51\\ &= 11.60 \text{ to } 12.62\end{aligned}$$

For working-class respondents:

$$s_{\bar{Y}} = \frac{s}{\sqrt{N}} = \frac{2.91}{\sqrt{541}} = 0.125 = 0.13$$

$$\begin{aligned}\text{Confidence interval} &= 13.01 \pm 1.96(0.13)\\ &= 13.01 \pm 0.25\\ &= 12.76 \text{ to } 13.25\end{aligned}$$

b.

For lower-class respondents:

$$s_{\bar{Y}} = \frac{s}{\sqrt{N}} = \frac{2.83}{\sqrt{122}} = 0.26$$

$$\begin{aligned}\text{Confidence interval} &= 12.11 \pm 2.58(0.26)\\ &= 12.11 \pm 0.67\\ &= 11.44 \text{ to } 12.78\end{aligned}$$

For middle-class respondents:

$$s_{\bar{Y}} = \frac{s}{\sqrt{N}} = \frac{2.93}{\sqrt{475}} = 0.13$$

$$\begin{aligned}\text{Confidence interval} &= 14.99 \pm 2.58(0.13)\\ &= 14.99 \pm 0.34\\ &= 14.65 \text{ to } 15.33\end{aligned}$$

c. As our confidence level increases, the confidence interval gets wider, not narrower. This is because a wider interval is needed to increase the probability that our calculated interval includes the true population value. Thus, increasing confidence leads to less precise intervals.

3.

a. For Canadians

$$s_p = \sqrt{\frac{(0.51)(1-0.51)}{1{,}004}} = 0.02$$

$$\begin{aligned}\text{Confidence interval} &= 0.51 \pm 1.96(0.02)\\ &= 0.51 \pm 0.04\\ &= 0.47 \text{ to } .55\end{aligned}$$

b. For Americans

$$\begin{aligned}\text{Confidence interval} &= 0.45 \pm 1.96(0.02)\\ &= 0.45 \pm 0.04\\ &= 0.39 \text{ to } 0.49\end{aligned}$$

c. Based on the calculated 95% confidence interval, the majority of Americans do not believe climate change is a serious problem. The true percentage of Americans who believe climate change is a serious problem is under 50%, somewhere between 39% and 49%, based on this Pew Research Center sample. On the other hand, it is possible that the majority of Canadians believe climate change is a serious problem. We can be 95% confident that the true percentage of Canadians is somewhere between 47% and 55%.

4.

a. 90% confidence interval for males

$$s_p = \sqrt{\frac{(0.18)(1-0.18)}{435}} = 0.02$$

$$\begin{aligned}\text{Confidence interval} &= 0.18 \pm 1.65(.02)\\ &= 0.18 \pm 03\\ &= 0.15 \text{ to } .21\end{aligned}$$

b. 90% confidence interval for females

$$s_p = \sqrt{\frac{(0.40)(1-.40)}{566}} = 0.02$$

$$\begin{aligned}\text{Confidence interval} &= 0.40 \pm 1.65(.02)\\ &= 0.40 \pm .03\\ &= 0.37 \text{ to } 43\end{aligned}$$

5.

Due to the large sample size, we converted the proportion to full percentages, subtracting from 100 (rather than 1).

$$\begin{aligned}\text{Confidence interval} &= 51 \pm 1.96(0.67)\\ &= 49.69\% \text{ to } 52.31\%\end{aligned}$$

We set the interval at the 95% confidence level. However, no matter whether the 90%, 95%, or 99% confidence level is chosen, the calculated interval includes values below 50% for the vote for a Republican candidate. Therefore, you should tell your supervisors that it would not be possible to declare a Republican candidate the likely winner of the votes coming from men if there was an election today because it seems quite possible that less than a majority of male voters would support her or him.

6.

a.

$$s_{\overline{Y}} = \frac{s}{\sqrt{N}} = \frac{0.80}{\sqrt{914}} = 0.03$$

$$\begin{aligned}\text{Confidence interval} &= 1.27 \pm 1.96(0.03)\\ &= 1.27 \pm 0.06\\ &= 1.21 \text{ to } 1.33\end{aligned}$$

b. The calculation of a confidence interval is still appropriate. For large enough samples, which 914 certainly is, the distribution of the sample means will be normal, no matter what the shape of the actual distribution of severe binge drinking. That being the case, we can confidently calculate confidence intervals based on normal distributions to get, in this instance, the 95% confidence interval.

7.

a.

$$s_p = \sqrt{\frac{(0.64)(1-0.64)}{1{,}403}} = 0.01$$

$$\begin{aligned}\text{Confidence interval} &= 0.64 \pm 1.96(0.01)\\ &= 0.64 \pm 0.02\\ &= 0.62 \text{ to } 0.66\end{aligned}$$

b. Based on our answer in 7a, we know that a 90% confidence interval will be more precise than a 95% confidence interval that has a lower bound of 62% and an upper bound of 66%. Accordingly, a 90% confidence interval will have a lower bound that is greater than 62% and an upper bound that is less than 66%. Additionally, we know that a 99% confidence interval will be less precise than what we calculated in 7a. Thus, the lower bound for a 99% confidence interval will be less than 62% and the upper bound will be greater than 66%.

8.

a. No estimate of the mean is needed. The error in a sample is related to the standard deviation and sample size, not to the mean.

b. Reducing sampling error to ±$500 means reducing this quantity to ±$500:

1.96(Standard Error) or

$$1.96\left(\frac{5{,}000}{\sqrt{N}}\right)$$

So we solve this equation:

$$500 = 1.96\left(\frac{5{,}000}{\sqrt{N}}\right)$$

So

$$\sqrt{N} = 19.6\,,\ N \cong 384$$

c. Here we solve this equation

$$500 = 2.58\left(\frac{5{,}000}{\sqrt{N}}\right)$$

So

$$\sqrt{N} = 25.8\,,\ N \cong 666$$

9.

Country	Mean	Standard Error	Confidence Interval
France	14.12	$5.73/\sqrt{975} = 0.18$	$14.12 + 0.18(1.65) = 14.42$ $14.12 - 0.18(1.65) = 13.82$
Japan	12.48	$2.53/\sqrt{528} = 0.11$	$12.48 + 0.11(1.65) = 12.66$ $12.48 - 0.11(1.65) = 12.30$
Croatia	12.18	$2.71/\sqrt{480} = 0.12$	$12.18 + 0.12\,(1.65) = 12.38$ $12.18 - 0.12\,(1.65) = 11.98$
Turkey	9.15	$11.98/\sqrt{783} = 0.43$	$9.15 + 0.43\,(1.65) = 9.86$ $9.15 - 0.43\,(1.65) = 8.44$

10.

For Bernie Sanders

$$s_p = \sqrt{\frac{(0.55)(1-0.55)}{1{,}754}} = 0.01$$

$$\begin{aligned}\text{Confidence interval} &= 0.55 \pm 1.65(0.01)\\ &= 0.55 \pm 0.02\\ &= 0.53 \text{ to } 0.57\end{aligned}$$

For Hillary Clinton

$$s_p = \sqrt{\frac{(0.38)(1-0.38)}{1,754}} = 0.01$$

$$\begin{aligned} \text{Confidence interval} &= 0.18 \pm 1.65(0.01) \\ &= 0.38 \pm 0.02 \\ &= 0.36 \text{ to } 0.40 \end{aligned}$$

11.

For Republicans

$$s_p = \sqrt{\frac{(0.18)(1-0.18)}{446}} = 0.02$$

$$\begin{aligned} \text{Confidence interval} &= 0.18 \pm 1.96(0.02) \\ &= 0.18 \pm 0.04 \\ &= 0.14 \text{ to } 0.22 \end{aligned}$$

For Democrats

$$s_p = \sqrt{\frac{(0.15)(1-0.15)}{522}} = 0.02$$

$$\begin{aligned} \text{Confidence interval} &= 0.15 \pm 1.96(0.02) \\ &= 0.15 \pm 0.04 \\ &= 0.11 \text{ to } 0.19 \end{aligned}$$

12.

a.

$$s_p = \sqrt{\frac{(61)(100-61)}{527}} = 2.12$$

$$\begin{aligned} \text{Confidence interval} &= 61 \pm 1.96(2.12) \\ &= 61 \pm 4.16 \\ &= 56.84 \text{ to } 65.16 \end{aligned}$$

b.

$$s_p = \sqrt{\frac{(61)(100-61)}{527}} = 2.12$$

$$\begin{aligned} \text{Confidence interval} &= 61 \pm 2.58(2.12) \\ &= 61 \pm 5.47 \\ &= 55.53 \text{ to } 66.47 \end{aligned}$$

c. The 95% and 99% confidence intervals only include values above 50% (i.e., the majority). Since we are estimating whether the majority of Millennials (>50%) believe that their generation has a unique and distinctive identity, intervals are compatible with the idea that more than 50% of Millennials hold this view.

13.

a. For those who thought that homosexual relations were always wrong:

$$s_p = \sqrt{\frac{(0.40)(1-.40)}{950}} = 0.02$$

$$
\begin{aligned}
\text{Confidence interval} &= 0.40 \pm 1.96(.02) \\
&= 0.40 \pm 0.04 \\
&= 0.36 \text{ to } 0.44
\end{aligned}
$$

For those who thought that homosexual relations were not wrong at all:

$$s_p = \sqrt{\frac{(0.49)(1-0.49)}{950}} = 0.02$$

$$
\begin{aligned}
\text{Confidence interval} &= 0.49 \pm 1.96(.02) \\
&= 0.49 \pm 0.04 \\
&= 0.45 \text{ to } 0.53
\end{aligned}
$$

b.

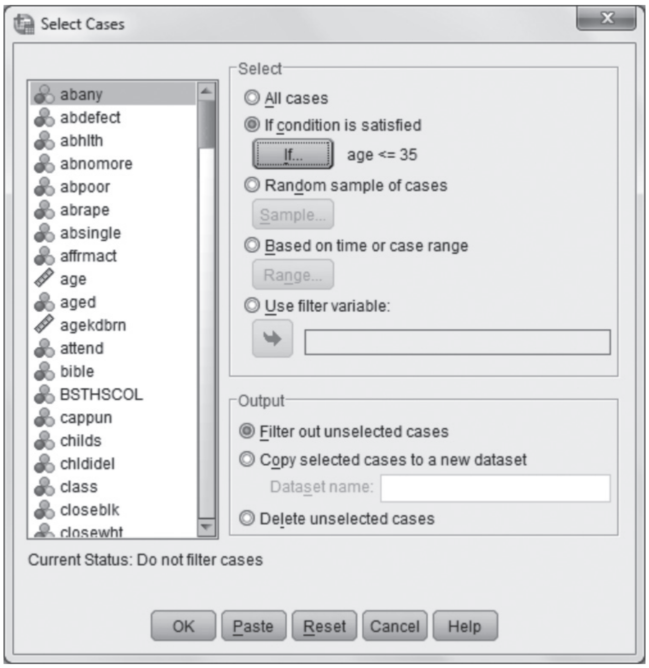
$$s_p = \sqrt{\frac{(0.10)(1-0.10)}{950}} = 0.01$$

$$
\begin{aligned}
\text{Confidence interval} &= 0.10 \pm 1.96(.01) \\
&= 0.10 \pm 0.02 \\
&= 0.08 \text{ to } 0.12
\end{aligned}
$$

SPSS SOLUTIONS

1.

a. Instructors may want to advise students to set 9, –1, and 8 as "MISSING" before running EXPLORE procedure. Results presented here excluded 98 (DK) and 99 (NA), –1 (IAP), or 9 (DK, NA) for analysis. In this problem, students should follow the example of the SPSS demonstration. The SPSS dialog box that selects cases based on respondent age should appear as follows:



b.

Case Processing Summary

	sex RESPONDENTS SEX	Cases Valid N	Valid Percent	Missing N	Missing Percent	Total N	Total Percent
hrsrelax HOURS PER DAY R HAVE TO RELAX	1 MALE	78	49.1%	81	50.9%	159	100.0%
	2 FEMALE	79	39.7%	120	60.3%	199	100.0%

Descriptives

	sex RESPONDENTS SEX			Statistic	Std. Error
hrsrelax HOURS PER DAY R HAVE TO RELAX	1 MALE	Mean		3.77	.275
		99% Confidence Interval for Mean	Lower Bound	3.04	
			Upper Bound	4.50	
		5% Trimmed Mean		3.64	
		Median		4.00	
		Variance		5.894	
		Std. Deviation		2.428	
		Minimum		0	
		Maximum		12	
		Range		12	
		Interquartile Range		3	
		Skewness		.707	.272
		Kurtosis		.941	.538
	2 FEMALE	Mean		3.00	.264
		99% Confidence Interval for Mean	Lower Bound	2.30	
			Upper Bound	3.70	
		5% Trimmed Mean		2.82	
		Median		3.00	
		Variance		5.513	
		Std. Deviation		2.348	
		Minimum		0	
		Maximum		14	
		Range		14	
		Interquartile Range		2	
		Skewness		1.591	.271
		Kurtosis		5.086	.535

The samples for males and females were much smaller than the entire sample (78 males, 79 females). In this sample, males reported more hours to relax per day (3.77) compared with females (3.00).

For females in this younger sample, the mean number of relaxation hours is lower than the mean hours in the full sample (3.00 vs. 3.18). For males, mean hours are higher for the younger sample (3.77) versus the complete sample (3.68).

The width of the confidence intervals was much larger in the younger sample than the complete sample. At the 99% confidence level, the values for the lower and upper bounds for males in the younger sample were 3.04 and 4.50, respectively (width of 1.46). For females in the younger sample, the lower bound of the 99% confidence interval was 2.30 and the upper bound was 3.70 (width of 1.40). When compared with the confidence intervals in the SPSS Demonstration, the widths of these intervals are wider. As sample size decreases, the width of the confidence interval increases even at the same confidence level, making our estimates less precise.

2. We ran all variables in a single Explore procedure. In a single Explore procedure, the group N's are based on common valid responses. Running the variables separately will produce different results from the ones shown here.

a.

Class Identification				Statistic
Hours per day to relax	LOWER	Mean		4.20
		90% Confidence Interval	Lower Bound	2.89
			Upper Bound	5.51
		Median		3.00
		Std. Deviation		3.819

(Continued)

(Continued)

Class Identification				Statistic
	WORKING	Mean		3.10
		90% Confidence Interval	Lower Bound	2.87
			Upper Bound	3.33
		Median		3.00
		Std. Deviation		2.091
	MIDDLE	Mean		3.75
		90% Confidence Interval	Lower Bound	3.43
			Upper Bound	4.08
		Median		3.00
		Std. Deviation		2.660
	UPPER	Mean		3.36
		90% Confidence Interval	Lower Bound	2.51
			Upper Bound	4.20
		Median		3.00
		Std. Deviation		1.781

b.

Class Identification				Statistic
Highest year of school completed	LOWER	Mean		12.04
		90% Confidence Interval	Lower Bound	11.05
			Upper Bound	13.03
		Median		12.00
		Std. Deviation		2.879
	WORKING	Mean		13.03
		90% Confidence Interval	Lower Bound	12.68
			Upper Bound	13.37
		Median		13.00
		Std. Deviation		3.175
	MIDDLE	Mean		15.36
		90% Confidence Interval	Lower Bound	15.03
			Upper Bound	15.69
		Median		16.00
		Std. Deviation		2.726
	UPPER	Mean		16.86
		90% Confidence Interval	Lower Bound	15.69
			Upper Bound	18.03
		Median		16.00
		Std. Deviation		2.476

c.

Class Identification				Statistic
Number of hours worked last week	LOWER	Mean		35.72
		90% Confidence Interval	Lower Bound	31.48
			Upper Bound	39.96
		Median		40.00
		Std. Deviation		12.398
	WORKING	Mean		41.80
		90% Confidence Interval	Lower Bound	40.32
			Upper Bound	43.29
		Median		40.00
		Std. Deviation		13.633
	MIDDLE	Mean		42.85
		90% Confidence Interval	Lower Bound	40.98
			Upper Bound	44.72
		Median		40.00
		Std. Deviation		15.426
	UPPER	Mean		42.86
		90% Confidence Interval	Lower Bound	34.66
			Upper Bound	51.05
		Median		40.00
		Std. Deviation		17.320

d.

Class Identification				Statistic
Highest year of school completed, mother	LOWER0	Mean		9.48
		90% Confidence Interval	Lower Bound	7.89
			Upper Bound	11.07
		Median		12.00
		Std. Deviation		4.656
	WORKING	Mean		11.31
		90% Confidence Interval	Lower Bound	10.84
			Upper Bound	11.79
		Median		12.00
		Std. Deviation		4.324

(Continued)

(Continued)

Class Identification				Statistic
	MIDDLE	Mean		13.06
		90% Confidence Interval	Lower Bound	12.60
			Upper Bound	13.52
		Median		12.00
		Std. Deviation		3.806
	UPPER	Mean		14.50
		90% Confidence Interval	Lower Bound	12.38
			Upper Bound	16.62
		Median		14.50
		Std. Deviation		4.485

The mean number of relaxation hours is highest for respondents in the lower-class group (4.20) and lowest for the working-class group (3.10). Hours worked is highest for the upper class group (42.86) and lowest for the lowest class group (35.72). The amount of working hours may be negatively related to relaxation, but cannot be determined based on the current analysis.

For both educational variables—respondent's and mother's—the ranking of the mean scores are as expected. Upper class respondents have the highest educational year average (16.86 years), mother's educational year average (14.50 years). The lowest averages for both variables were reported for lower class respondents—for respondent (12.04) and mother (9.48).