CHAPTER 3— ANSWERS TO EXERCISES

1.

a. Mode = Routine ($f = 379$)

b. Median = Routine

c. Based on the mode and median for this variable, most respondents indicate that their lives are "routine."

d. A mean score could not be interpreted for this variable. A mean would have no meaning for a nominal measurement.

2.

a. Ordinal.

b. The mode can be found in two ways: by looking for the (1) highest frequency ($f = 251$) or (2) highest percentage (32.5%)—"strongly agree".

c. The median can be found in two ways: by using either the frequencies column or the cumulative percentages.

Using Frequencies	Using Cumulative Percentages
$\frac{N+1}{2} = \frac{773+1}{2} = 387$th case	Notice that 50% of the observations fall in the "agree" cumulative percentage category.
Starting with the frequency in the first category (251), add up the frequencies until you find where the 387th case falls. The 387th case falls at the "agree" category, which is the median.	The 50% mark, or the median, is located in the "agree" category.

The data indicate strong support for same-sex marriage. The largest single category is "strongly agree"— 32.5%. Combined with the "agree" category (the second largest category), more than half of the GSS sample agree that homosexuals should have the right to marry (32.5 + 25.2 = 57.7).

3.

a. Interval ratio. The mode can be found in two ways: by looking either for the highest frequency (14) or the highest percentage (43.8%). The mode is the category that corresponds to the value "40 hours worked last week." The median can be found in two ways: by using either the frequencies column or the cumulative percentages.

Using Frequencies	Using Cumulative Percentages
$\frac{N+1}{2} = \frac{32+1}{2} = 16.5$th case	Notice that 34.4% of the observations fall in or below the "32 hours worked last week" category; 78.1% fall in or below the "40 hours worked last week" category.
Starting with the frequency in the first category (1), add up the frequencies until you find where the 16th and 17th cases fall. Both these cases correspond to the category "40 hours worked last week," which is the median.	The 50% mark, or the median, is located somewhere within the "40 hours worked last week" category. So the median is "40 hours worked last week."

b. Since the median is merely a synonym for the 50th percentile, we already know that its value is "40 hours worked last week."

25th percentile = (32 × 0.25) = 8th case = 30 hours worked last week.

75th percentile = (32 × 0.75) = 24th case = 40 hours worked last week

4.

HOW IMPORTANT TO BE A CHRISTIAN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	VERY IMPORTANT	136	11.5	34.7	34.7
	FAIRLY IMPORTANT	49	4.2	12.5	47.2
	NOT VERY IMPORTANT	89	7.5	22.7	69.9
	NOT IMPORTANT AT ALL	118	10.0	30.1	100.0
	Total	392	33.2	100.0	
Missing	IAP	771	65.4		
	CANT CHOOSE	14	1.2		
	NA	2	.2		
	Total	787	66.8		
Total		1179	100.0		

a. The mode category is "very important", $f = 136$.

b. The median can be found in two ways: either by using the frequencies column or by calculating the cumulative percentages. The median score is 3—not very important.

c. 20th percentile = (392 × 0.20) = 78th case = "Very important"

80th percentile = (392 × 0.80) = 314th case = "Not important at all"

5.

breakfastV8526 2014 T02 OFTN EAT BRKFST F2 * raceV1070 2014 RACE--B/W/H F1234 Crosstabulation

			raceV1070 2014 RACE--B/W/H F1234			
			1 BLACK:(1)	2 WHITE:(2)	3 HISPANIC:(3)	Total
breakfastV8526 2014 T02 OFTN EAT BRKFST F2	1 NEVER:(1)	Count	4	14	5	23
		% within raceV1070 2014 RACE--B/W/H F1234	9.3%	6.4%	7.0%	6.9%
	2 SELDOM:(2)	Count	11	22	9	42
		% within raceV1070 2014 RACE--B/W/H F1234	25.6%	10.1%	12.7%	12.7%
	3 SOMETIMES:(3)	Count	10	43	21	74
		% within raceV1070 2014 RACE--B/W/H F1234	23.3%	19.7%	29.6%	22.3%
	4 MOST DAYS:(4)	Count	2	25	9	36
		% within raceV1070 2014 RACE--B/W/H F1234	4.7%	11.5%	12.7%	10.8%
	5 NEARLY EVERY DAY:(5)	Count	7	30	2	39
		% within raceV1070 2014 RACE--B/W/H F1234	16.3%	13.8%	2.8%	11.7%
	6 EVERYDAY:(6)	Count	9	84	25	118
		% within raceV1070 2014 RACE--B/W/H F1234	20.9%	38.5%	35.2%	35.5%
Total		Count	43	218	71	332
		% within raceV1070 2014 RACE--B/W/H F1234	100.0%	100.0%	100.0%	100.0%

a.

	Mode	Median
Black	Seldom ($f = 11$)	Sometimes
White	Everyday ($f = 84$)	Nearly everyday
Hispanic	Everday ($f = 25$)	Most days

b. Teens' breakfast habits vary by race/ethnicity. Out of the three racial/ethnic groups, black students were more likely to report seldom or sometimes eating breakfast. On the other hand, white and Hispanic students eat breakfast more frequently. The mode for white and Hispanic students is everyday.

6. The mode can be found by looking for the highest frequency in each column; the mode for each group is listed below:

Males: Working full time

Females: Working full time

The median can be found in two ways: either by using the frequencies column or by using the cumulative percentages. However, since the problem only gives the frequencies, we'll use those to solve for the median.

Males	Females
$\frac{N+1}{2}=\frac{513+1}{2}=257$th case	$\frac{N+1}{2}=\frac{503+1}{2}=252$nd case
Starting with the frequency in the first category (303), add up the frequencies until you find where the 257th case falls. We actually don't need to do any adding, as both these cases correspond with the first category, "Working Full Time."	Starting with the frequency in the first category (263), add up the frequencies until you find where the 252nd case falls. We actually don't need to do any adding, as both these cases correspond with the first category, "Working Full Time."

When using both the mode and the median to estimate participation in the labor force, it appears that there are no substantial differences between males and females.

7. We begin by multiplying each household size by its frequency.

Household Size	Frequency	Frequency × Y(*f*Y)
1	381	381
2	526	1,052
3	227	681
4	200	800
5	96	480
6	42	252
7	19	133
8	5	40
9	2	18
10	2	20
Total	$N = 1{,}500$	$\Sigma fY = 3{,}857$

$$\overline{Y}=\frac{\Sigma f\,Y}{N}=\frac{3{,}857}{1{,}500}=2.57$$

The mean number of people per household is 2.57.

8.

a. Based on the table, the 90th percentile is in response category "4."

b. The median is "2."

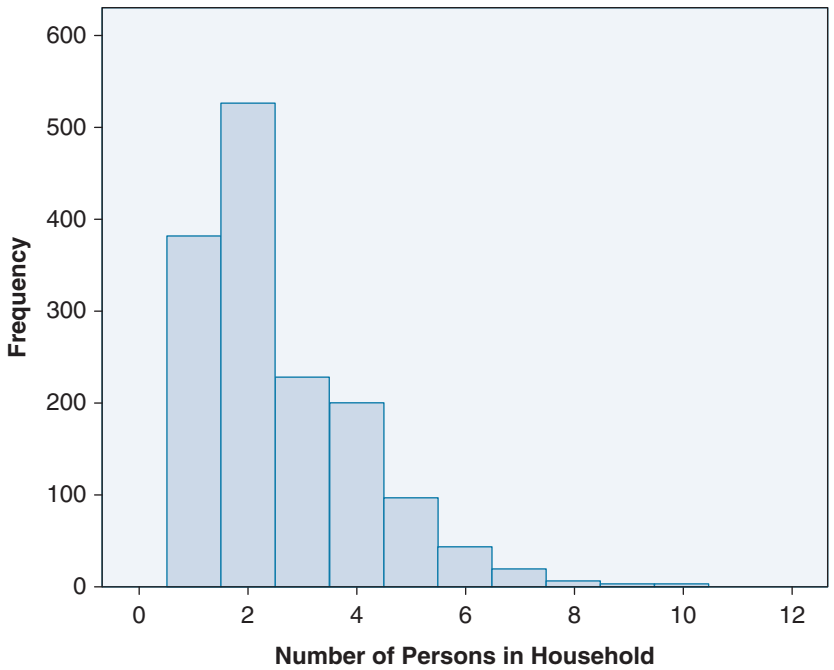
c. The mean is 2.46 = 1,673/679.

d. The distribution is slightly positively skewed; the mean is larger than the median.

Ideal Number of Children	Frequency	Frequency × Y(fY)
0	11	0
1	14	14
2	395	790
3	188	564
4	56	224
5	11	55
6	2	12
7	2	14
Total	$N = 679$	$\sum fY = 1{,}673$

9.

a. There appear to be a few outliers (i.e., extremely high values); this leads us to believe that the distribution is skewed in the positive direction.



b. The median can be found in two ways: by using either the frequencies column or the cumulative percentages. The data are in frequencies; we'll use those to solve the median. Because the median (2) is less than the mean (2.57), we can conclude that the distribution is skewed in a positive direction. Our answer to Question 9a is further supported.

Using Frequencies

$$\frac{N+1}{2} = \frac{1{,}500+1}{2} = 750.5\text{th case}$$

Starting with the frequency in the first category (381), add up the frequencies until you find where the 750th and 751st cases fall. Both these cases correspond to the category "2," which is the median.

10.

a. We begin by multiplying each category by its frequency.

Hours Worked Last Week	Frequency	Frequency × Y
20	3	60
25	2	50
28	1	28
29	1	29
30	3	90
32	1	32
40	14	560
50	2	100
52	1	52
55	1	55
60	1	60
64	1	64
70	1	70
Total	$N = 32$	$\Sigma fY = 1{,}250$

$$\overline{Y} = \frac{\Sigma f Y}{N} = \frac{1{,}250}{32} = 39.06$$

b. The median was 40 hours worked last week. This distribution is slightly skewed in a positive direction (i. e., the value of the median is greater than the mean).

11. The mean and the median represent a precise statistical middle. The mean is often referred to as the "arithmetic middle," by definition, summing everyone's income and dividing the total by the total number of people. The mean is sensitive to extremes, very low or high values, and so when we consider income, the preferred measure is the median. The median is the midpoint of all collected incomes, representing the exact point where 50% of all cases are either above or below. Because Clinton and Sanders' middle-class income amount is higher than the U.S. Census estimated mean or median, their definition of middle class is not based on the statistical middle. Are they operationalizing a middle-class life style, one that includes home and car ownership, occupational status, and wealth?

12. The data are reordered to calculate the median and the mean for each country group.

The mean homicide rate is similar in both groups, though slightly higher for the second most populated countries (6.94 vs. 6.5). For both country groups, the median is lower than the mean, indicating a positively skewed distribution. The difference between the median and mean is greater for the most populated country group (6.5 – 3.55 vs. 6.94 – 4.85).

#1–10 Most Populated	#11–20 Most Populated
0.3	0.7
0.6	1.5
0.8	3.4
2.6	4.3
3.3	4.8
3.8	4.9
7.8	8.1
9.0	9.3
10.3	13.5
26.5	18.9
Median = 3.55	Median = 4.85
Mean = 65/10 = 6.50	Mean = 69.4/10 = 6.94

13.

a. The data are reordered to calculate the median.

Infant Mortality Rates
2.52
3.43
4.65
5.87
10.41
14.58
15.61
18.87
26.11
58.19
115.08
Median = 14.58, 6th case
Mean = 275.32/11 = 25.03

b. The mean is greater than median, indicating a positively skewed distribution.

SPSS SOLUTIONS

1.

a. On average, women were older by 0.88 years (50.52 – 49.64). Men used their e-mail more often (6.62 vs. 6.27) and watched more television (3.14 vs. 2.79).

b. Ranking the degree categories by average hours of e- mail:

Graduate	Bachelor	Jr. College	High School	LT High School
12.27	8.49	5.81	5.28	1.59

2.

a. Nominal, mode

b. Ordinal, mode, or median

c. Interval-ratio, mean

d. Nominal, mode

e. Interval-ratio, mean

f. Nominal, mode

3.

Statistics

premarsx SEX BEFORE MARRIAGE

N	Valid	982
	Missing	518
Mean		3.14
Median		4.00
Mode		4

premarsx SEX BEFORE MARRIAGE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 ALWAYS WRONG	184	12.3	18.7	18.7
	2 ALMST ALWAYS WRG	60	4.0	6.1	24.8
	3 SOMETIMES WRONG	168	11.2	17.1	42.0
	4 NOT WRONG AT ALL	570	38.0	58.0	100.0
	Total	982	65.5	100.0	
Missing	0 IAP	495	33.0		
	8 DK	19	1.3		
	9 NA	4	.3		
	Total	518	34.5		
Total		1500	100.0		

a. For an ordinal measure, the mode or median would be the most appropriate measures of central tendency. The mode and median are identical—Category 4. The largest number of respondents indicated that premarital sex was "not wrong at all," and the median is located in the same category.

b. We present statistics tables for three of the BIBLE categories.

Statistics[a]

premarsx SEX BEFORE MARRIAGE

N	Valid	203
	Missing	122
Mean		3.73
Median		4.00
Mode		4

a. bible FEELINGS ABOUT THE BIBLE = 3 BOOK OF FABLES

Statistics[a]

premarsx SEX BEFORE MARRIAGE

N	Valid	445
	Missing	237
Mean		3.32
Median		4.00
Mode		4

a. bible FEELINGS ABOUT THE BIBLE = 2 INSPIRED WORD

Statistics[a]

premarsx SEX BEFORE MARRIAGE

N	Valid	316
	Missing	140
Mean		2.48
Median		2.50
Mode		1

a. bible FEELINGS ABOUT THE BIBLE = 1 WORD OF GOD

Measures of central tendency are similar for Bible respondents in Category 2 (inspired word) and Category 3 (book of fables). The median and modes for attitudes about premarital sex are Category 4—"not wrong at all."

The Bible respondents who disapprove most of premarital sex are those from Category 1 (word of God). The mode category is "always wrong," and the median is somewhere between "almost always wrong" and "sometimes wrong."

4.

a. EDUC is an interval-ratio measure. Mean would be the appropriate measure of central tendency.

b. On average, men and women have the same years of educational attainment. There are 672 males with an average of 13.77 years of education and 828 women with an average of 13.78 years of education.

5.

a. The best measure of central tendency for CHILDS is mean.

b. Ranking number of children by CLASS:

Lower, 2.27

Upper, 2.03

Working, 1.90

Middle, 1.83

Statistics[a]

NUMBER OF CHILDREN

N	Valid	688
	Missing	3
Mean		1.90
Median		2.00
Mode		2

a. SUBJECTIVE CLASS IDENTIFICATION = WORKING CLASS

Statistics[a]

NUMBER OF CHILDREN

N	Valid	615
	Missing	0
Mean		1.83
Median		2.00
Mode		2

a. SUBJECTIVE CLASS IDENTIFICATION = MIDDLE CLASS

Statistics[a]

NUMBER OF CHILDREN

N	Valid	39
	Missing	0
Mean		2.03
Median		2.00
Mode		2

a. SUBJECTIVE CLASS IDENTIFICATION = UPPER CLASS

Statistics[a]

NUMBER OF CHILDREN

N	Valid	146
	Missing	2
Mean		2.27
Median		2.00
Mode		2

a. SUBJECTIVE CLASS IDENTIFICATION = LOWER CLASS

c. Ranking ideal number of children by CLASS:

Lower, 3.38

Upper, 3.33

Working, 3.25

Middle, 3.08