CHAPTER 4— ANSWERS TO EXERCISES

1.

a. The table reveals seven response categories for political views.

b. The sum of the squared percentages, ΣPct^2, is equal to 2,301.52.

Political Views	Percentage (%)	Percentage Squared (%²)
Extremely liberal	3.6	12.96
Liberal	12.7	161.29
Slightly liberal	11.1	123.21
Moderate	39.5	1560.25
Slightly conservative	14.4	207.36
Conservative	14.9	222.01
Extremely conservative	3.8	14.44
Total	100.0	$\Sigma = 2{,}301.52$

c. Using the formula, we calculate the IQV as follows:

$$\text{IQV} = \frac{K(100^2 - \sum Pct^2)}{100^2(K-1)} = \frac{7(100^2 - 2{,}301.52)}{100^2(7-1)} = \frac{53{,}889.36}{60{,}000} = 0.90$$

The calculated IQV is close to 1 and suggests that Americans are fairly diverse in their political views.

2. The table indicates five response categories for highest degree attained. This is an ordinal measure.

To calculate the IQR, we rely on three calculations: $Q1 = (N)(.25)$, $Q3 = (N)(.75)$, and $\text{IQR} = Q3 - Q1$.

For males, $N = 525$. $Q1$ = the 131.25th case or high school and $Q3$ = 393.75th case or bachelor's degree. The IQR is 2 (high school) to 4 (bachelors).

For females, $N = 654$. $Q1$ = the 163.5th case or high school and $Q3$ = 490.50th case or bachelor's degree. The IQR is identical to males, 2 (high school) to 4 (bachelors).

Based on the IQR calculation, we know that 50% of all cases for males and females lies between high school and bachelor's degree, but doesn't tell us much about the variability of the distribution. A better measure would be IQV (as we calculated in Exercise 1) or if interval-ratio measures (actual years of education) are available, the variance and standard deviation would provide more information.

3.

a. The range of convictions in 1990 is (583 – 79) = 504. The range of convictions in 2009 is (426 – 102) = 324. The range of convictions is larger in 1990 than in 2009.

b. The mean number of convictions is 295.67 in 1990 and 261.67 in 2009.

c.

1990

Govt. Level	No. of Convictions	$(Y - \bar{Y})$	$(Y - \bar{Y})^2$
Federal	583	287.33	82,558.53
State	79	–216.67	46,945.89
Local	225	–70.67	4,994.25
Total	887		134,498.67
	$\bar{Y} = 295.67$		

$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma\left(Y - \bar{Y}\right)^2}{N-1}} = \sqrt{\frac{134,498.67}{2}} = 259.32$$

2009

Govt. Level	No. of Convictions	$(Y - \bar{Y})$	$(Y - \bar{Y})^2$
Federal	426	164.33	27,004.35
State	102	–159.67	25,494.50
Local	257	–4.67	21.80
Total	785		52,520.65
	$\bar{Y} = 261.67$		

$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma\left(Y - \bar{Y}\right)^2}{N-1}} = \sqrt{\frac{52,520.65}{2}} = 162.05$$

d. The standard deviation is larger in 1990 than in 2009, thus indicating more variability in number of convictions in 1990 than in 2009. This supports our results from 3a.

4. Type of offenses is a nominal measure. This limits what types of measures we can use. The mode is the best measure of central tendency, identifying the most frequent or the category with the most cases. For 2009 and 2012, the mode is immigration. The number of suspects for this category declined from 2009 to 2012 by 2,855 cases.

   The IQV is the best measure of variability. It will tell us if criminal cases are dispersed throughout the different categories or concentrated in a few. There is more variability across the categories for 2009 (IQV = 0.86) than for 2012 (IQV = 0.72). As we review the frequencies, we can see that most of the 2012 cases are concentrated in immigration and drug offenses, whereas for 2009 the cases are dispersed across more categories (immigration, drug, property, and public order).

5.

   a. The range of projected increase in the elderly population for the Western states is 36.2%. The range of percent increase for the Midwestern states is 9.8%. The Western states have a much larger range.

   b. The IQR for the Western states is 17.3%. The IQR for the Midwestern states is 3.7%. Again, the value for the Western states is greater.

   c. There is great variability in the projected increase in the elderly population in Western states, chiefly caused by the large increases in Nevada, Arizona, Wyoming, and Alaska, as measured by either the range or the IQR.

6.

   a. Because the mean and median values for occupational prestige are relatively close to similar (mean = 38.82, median = 38.00 for high school degree; mean = 50.02, median = 50.00 for bachelor's degree), these statistics do not suggest any significant skew between either of the types of graduates.

   b. Using the variance, standard deviation, range, and interquartile range to compare the variability for the two groups, those with bachelor's degrees have variance (172.360), standard deviation (13.129), and interquartile range (21) values that exceed those of the high school group (126.725, 11.257, and 16, respectively). Range scores are identical. Overall, this suggests that there is more variability of occupational prestige in the bachelor's degree group.

7.

   a. The range is 3.6 (6.5 – 2.9). The 25th percentile, 3.05, means that 25% of cases fall below 3.05 divorce rate per 1,000 population. Likewise, the 75th percentile means that 75% of all cases fall below 4.6 divorce rate per 1,000 population.

25th percentile	10(0.25) = 2.5th case	So (3.0 + 3.1)/2 = 3.05
75th percentile	10(0.75) = 7.5th case	So (4.5 + 4.7)/2 = 4.6

   The IQR is thus 4.6 – 3.05 = 1.55. Both measures of variability are appropriate, but the range is somewhat better, as the value for the IQR is fairly small. In other words, the range gives us a better picture of the variability of divorce rates for all states in our sample.

b.

State	Divorce Rate per 1,000 Population	$Y - \bar{Y}$	$(Y - \bar{Y})^2$
Alaska	4.3	0.2	0.04
Florida	4.7	0.6	0.36
Idaho	4.9	0.8	0.64
Maine	4.5	0.4	0.16
Maryland	3.1	–1.0	1.00
Nevada	6.5	2.4	5.76
New Jersey	3.0	–1.1	1.21
Texas	3.3	–0.8	0.64
Vermont	3.8	–0.3	0.09
Wisconsin	2.9	–1.2	1.44
Total	41	0.00	11.34

$$\bar{Y} = \frac{\Sigma Y}{N} = \frac{41}{10} = 4.1$$

$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma\left(Y - \bar{Y}\right)^2}{N-1}} = \sqrt{\frac{11.34}{9}} = 1.12$$

c. Divorce rates may vary by state due to factors such as variation in religiosity, state policy (i.e., no fault divorce laws), or employment opportunities.

8.

a. Degree is an ordinal measure; age when first child was born is an interval-ratio measure.

b. As degree increases, the age when first child was born increases. This indicates a positive relationship. The youngest first-time parents are those with less than a high school degree, while the oldest first-time parents are those with graduate degrees. The difference is 8.27 years (28.99 – 20.72). The variability in age when first child was born is larger as educational attainment increases. The standard deviation for the graduate degree group is largest at 6.205 years, the smallest is for the less than high school group, 4.745.

9.

a. The mean number of crimes is 3,038.9 and the standard deviation is 583.004. The mean amount of dollars (in millions) spent on police protection is $1,703.95 and the standard deviation is $1,895.21.

b. Because the number of crimes and police protection expenditures are measured according to different scales, it isn't appropriate to directly compare the mean and standard

deviation for one variable with the other. But we can talk about each distribution separately. We know from examining the mean (3,038.90) and standard deviation (583.00) for the number of crimes that the standard deviation is large, indicating a wide dispersion of scores from the mean. For the number of crimes, states such as Missouri and South Dakota contribute more to its variability because they have values far from the mean (both above and below). With respect to police protection expenditures, we can see that there is a large dispersion from the mean of $1,703.95, as the standard deviation is $1,895.21. States such as New York and North Dakota contribute more to its variability because they have values far from the mean (both above and below).

c. Among other considerations, we need to consider the economic conditions in each state. A downturn in the local and state economy may play a part in the number of crimes and police expenditures per capita.

10. We should be cautious when making generalized statements about the relationship between education and ideal number of children because we only have statistics for two groups. We would need data from additional ethnic groups to make specific statements about this relationship. Therefore, we must restrict our discussion to Chinese Americans and Filipino Americans. On average, Chinese Americans are more educated than Filipino Americans (15.55 years vs. 13.42 years), and both groups have about the same standard deviation (3.643 for Chinese Americans and 3.704 for Filipino Americans). Additionally, Chinese Americans report a lower number of ideal children (2.88) than Filipino Americans (4.00). Both groups have about the same standard deviation (2.167 for Chinese Americans and 2.098 for Filipino Americans). Based on these findings, we might suggest that as level of education increases, the ideal number of children decreases (but we can't be certain this is the case for all Americans).

11.

a. Type of paid work is a nominal variable. The appropriate measure of variability would be the index of qualitative variation (IQV).

b.

Grade 8

Type of Work	Percentage (%)	Percentage Squared (%²)
Lawn work	28	784
Food service	3	9
Babysitting	37	1369
Other	32	1,024
Total	100.0%	$\Sigma = 3{,}186$

$$IQV = \frac{K(100^2 - \sum Pct^2)}{100^2(K-1)} = \frac{4(100^2 - 3{,}186)}{100^2(4-1)} = \frac{27{,}256}{30{,}000} = 0.91$$

The IQV for 8th graders is 0.91.

Grade 10		
Type of Work	**Percentage (%)**	**Percentage Squared (%²)**
Lawn work	20	400
Food service	10	100
Babysitting	28	784
Other	42	1,764
Total	100.0%	$\Sigma = 3{,}048$

$$IQV = \frac{K(100^2 - \sum Pct^2)}{100^2(K-1)} = \frac{4(100^2 - 3{,}048)}{100^2(4-1)} = \frac{27{,}808}{30{,}000} = 0.93$$

The IQV for 10th graders is 0.93.

c. Though both IQVs are more than 0.90, there is slightly more variation among 10th graders than 8th graders in the type of jobs they hold. The difference could be attributed to more employment options for older students. Younger students may be limited to the kind of work they can do (due to age, experience, and time), leading to more informal jobs, such as lawn work and babysitting.

12. Since the variable is interval-ratio, we should use variance (or standard deviation), range, or IQR. Among these three measures, variance and/or standard deviation is preferred. For measurements of central tendency, as discussed in Chapter 3, if we are looking for the average life expectancy for these 10 countries, we should rely on the mean.

On average, non-European countries have a slightly higher life expectancy at birth. Both the mean and median are higher for non-European countries than for European countries. Also, the distribution of non-European countries exhibits more variability; the standard deviation for European countries is 3.03 years, while for non-European countries it is 3.43 years.

These differences might be explained by access and availability of health care and/or diet. However, the difference might simply be random due to the small number of countries presented in this example. We are likely to find different results if more countries were incorporated into the analyses.

A table of results is shown below:

Life Expectancy		Statistic
European countries	Mean	79.96
	Median	81.20
	Variance	9.19
	Standard deviation	3.03
	Minimum	74.60
	Maximum	81.80

Life Expectancy		Statistic
	Range	7.20
	Interquartile range	4.10
Non-European countries	Mean	81.56
	Median	82.30
	Variance	11.78
	Standard deviation	3.43
	Minimum	75.70
	Maximum	84.70
	Range	9
	Interquartile range	4.95

13. Overall, Obama voters were younger, more educated, and attended religious services less than McCain voters. The youngest voters were female Obama voters at 50.99 years (s = 16.62), followed by male Obama voters, 51.71 years (s = 15.59). For education, males who voted for Obama had the highest mean of 14.84 (s = 3.07). Males who voted for McCain had 14.60 years of education (s = 2.41). McCain voters, both males and females, attended religious services more often than Obama voters. Mean scores were 3.93 for males (s = 2.80) and 4.64 for females (s = 2.76), indicating church attendance about once a month to 2 × 3 times per month. The standard deviations indicate a consistency in the distributions of education, age, and religious service attendance across all four groups. The largest standard deviations are for age, ranging from 15.61 to 16.62 years. These wide standard deviations indicate more dispersion around the mean age scores.

SPSS SOLUTIONS

1. Age is an interval measure. The best measures of variability are variance and standard deviation. The standard deviation indicates that there are 17.073 years, a wide dispersion around the mean of 50.12. Assuming a normal distribution of age, we can expect about 68% of cases to be between 33.05 and 67.19 years. The GSS appears to be an older population.

Statistics

age AGE OF RESPONDENT

N	Valid	1490
	Missing	10
Mean		50.12
Std. Deviation		17.073
Variance		291.495
Range		71
Percentiles	25	36.00
	50	50.00
	75	62.00

2. Women have higher average prestige scores and years of education than men in the GSS 2014 sample. For PRESTG10, the variability for both samples are about the same—the standard deviation for men is 13.533 and 13.636 for women. (Both range scores are identical, though the IQR for women is larger: 20 vs. 16 for men.) For EDUC, the standard deviation for men is slightly higher than the standard deviation for women (3.142 vs. 2.967).
3. Students to complete on their own.
4.
    a. With respect to hours worked last week, whites have a standard deviation of 15.239, a variance of 232.235, a range of 88, and an IQR of 15. Conversely, blacks have a standard deviation of 15.200, a variance of 231.051, a range of 86, and an IQR of 8. Given these numbers, white and black respondents have the same amount of variability in hours worked last week. One distribution is not more or less dispersed than the other.
    b. Students to complete on their own.