**CHAPTER 4 SPSS PROBLEMS SOLUTIONS**

[*GSS10SSDS*]

1.

a. Each of the appropriate measures of variability (standard deviation, variance, and range) is greater for respondent age than are those for respondents’ age when their first child was born. Specifically, the standard deviation (17.557 versus 5.933), variance (308.248 versus 35.199), and range (71 versus 40) for respondent age exceed the values for age when first child was born. Greater values indicate greater variability, so respondent age has more variability.

b. Age varies widely across society, yet social and biological factors often determine the age range for having a child. That is, people have their first children according to societal norms as well as when their reproductive systems allow for procreation and child bearing. Because the window is shorter for when people first have a child, we would expect the variability in this age range to be less than respondent age overall.

2. According to our output, on average, men tend to be older than women when first having a child (25.64 years compared to 22.47). The median values also indicate this same trend, with men having a value of 25 and women having a value of 21. Moreover, certain measures of variability also depict more variation in the ages of men than women when first having a child. Specifically, the values for the variance (41.387 versus 27.267), standard deviation (6.433 versus 5.222), and range (40 versus 30) are larger for men than for women, indicative of greater variability.

3. Students complete on their own.

4.

a. Looking at the response options for each variable (which are the same), we can stratify responses based on respondents’ confidence in either institution. Therefore, we know that both variables are measured at the ordinal level.

b. Appropriate measures of central tendency for both variables would be the mode and median (the mean cannot be calculated because the variables are not measured at the interval-ratio level). Even though these variables are ordinal, the only appropriate measure of variability would be the IQV because of the relatively few categories respondents have to choose from (only three). We could use the range and IQR, but those estimates wouldn’t really provide much detail regarding the variability.

c. CONARMY

|  |  |  |
| --- | --- | --- |
| ***Confidence Level*** | ***%*** | ***%2*** |
| A Great Deal | 53.9 | 2,905.21 |
| Only Some | 39.4 | 1,552.36 |
| Hardly Any | 6.7 | 44.89 |
| Total | 100.0 | 4,502.46 |

Appropriate measures of central tendency would be the median and mode, since this is a categorical variable. Both the median and mode are “A Great Deal.”

CONPRESS

|  |  |  |
| --- | --- | --- |
| ***Confidence Level*** | ***%*** | ***%2*** |
| A Great Deal | 8.5 | 72.25 |
| Only Some | 45.4 | 2,061.16 |
| Hardly Any | 46.0 | 2,116.00 |
| Total | 100.0 | 4,249.41 |

Appropriate measures of central tendency would be the median and mode, since this is a categorical variable. The median is “Only Some.” Since “Only Some” and “Hardly Any” have relatively similar frequencies/percentages, we could consider this variable being bimodal.

d. Looking at the categorical percentages for each variable, respondents have more confidence in the military than the press.

e. CONARMY, Men

|  |  |
| --- | --- |
| ***Confidence Level*** | ***%*** |
| A Great Deal | 57.5 |
| Only Some | 36.8 |
| Hardly Any | 5.7 |
| Total | 100.0 |

CONPRESS, Men

|  |  |
| --- | --- |
| ***Confidence Level*** | ***%*** |
| A Great Deal | 8.7 |
| Only Some | 44.0 |
| Hardly Any | 47.3 |
| Total | 100.0 |

CONARMY, Women

|  |  |
| --- | --- |
| ***Confidence Level*** | ***%*** |
| A Great Deal | 51.2 |
| Only Some | 41.3 |
| Hardly Any | 7.4 |
| Total | 100.0 |

CONPRESS, Women

|  |  |
| --- | --- |
| ***Confidence Level*** | ***%*** |
| A Great Deal | 8.3 |
| Only Some | 46.5 |
| Hardly Any | 45.1 |
| Total | 100.0 |

According to the frequency distributions, men are slightly more confident in the military than are women, though both men and women reported similar levels of confidence with respect to the press.

5.

a. With respect to hours worked last week, Whites have a standard deviation of 15.415, a variance of 237.636, a range of 88, and an IQR of 15. Conversely, Blacks have a standard deviation of 15.764, a variance of 248.516, a range of 88, and an IQR of 14.5. Given these numbers, there is not much variability between Whites and Blacks in terms of hours worked.

b. Students complete on their own.

**CHAPTER 4 EXERCISE SOLUTIONS**

1.

1. 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 2009.
2. The mean number of convictions is 295.67 in 1990 and 261.67 in 2009.

**1990**

|  |  |  |  |
| --- | --- | --- | --- |
| ***Government Level*** | ***# of Convictions*** | ***( Y - )*** | ***( Y - )2*** |
| Federal | 583 | 287.33 | 82,560.44 |
| State | 79 | -216.67 | 46,944.44 |
| Local | 225 | -70.67 | 4,993.78 |
| Total | 887 | -0.01 | 134,498.67 |
|  | = 295.67 |  |  |
|  | | | |

**2009**

|  |  |  |  |
| --- | --- | --- | --- |
| ***Government Level*** | ***# of Convictions*** | ***( Y - )*** | ***( Y - )2*** |
| Federal | 426 | 164.33 | 27,005.44 |
| State | 102 | -159.67 | 25,493.44 |
| Local | 257 | -4.67 | 21.78 |
| Total | 785 | -0.01 | 52,520.67 |
|  | = 261.67 |  |  |
|  | | | |

d. The standard deviation is larger in 1990 than in 2009, thus indicating more variability in number of convections in 1990 than in 2009. This supports our results from 1a.

2.

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. The standard deviation for Western states is 10.6948. The standard deviation for Midwestern is 2.7831. Again, the value for the western states is greater.

d. 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.

3.

a. Because the mean and median values for occupational prestige are relatively close to similar (mean = 40.59, median = 40.00 for high school degree; mean = 50.95; median = 51.00 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 (167.185), standard deviation (12.930), and interquartile range (23) values than exceed those of the high school group (130.393, 11.419, and 17, respectively). Only with the range does the high school group exceed those in the college group (58 compared to 52). Overall, this suggests that there is more variability of prestige in the college graduate group.

4.

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** |  | **()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 | |
|  | | | |

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.

5.

a. For those diagnosed with cancer:

Then, the standard deviation is , or 4.06

For those not diagnosed with cancer:

Then, the standard deviation is , or 4.58

b. Although there is a slightly more variability in psychological distress score among those not diagnosed with cancer than those diagnosed with cancer, the standard deviations of the two groups are very close (4.58 vs. 4.06). An interesting finding is that the respondents diagnosed with cancer are not only slightly less diverse in terms of the distress level, but their average psychological distress score is also less than those not diagnosed with cancer (3.9 vs. 4.87).

c. No, we didn’t need the mean of Y to calculate the variance or standard deviation.

6.

a. The mean numbers 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 $1,704.0 and the standard deviation is $1,895.214.

b. Because the number of crimes and police protection expenditures is 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. 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.

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.

7.

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 European countries tended to exhibit more variability; the standard deviation for European countries is 3.5 years, while for non-European countries it is 2.6 years. The IQRs also attest to more variability in the distribution of life expectancy for European countries (IQR = 4.3) compared with non-European countries (IQR = 2.5).

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. Perhaps, we would find different results if more countries were incorporated into the analyses.

A table of results is shown below:

|  |  |  |
| --- | --- | --- |
| ***Life Expectancy*** |  | ***Statistic*** |
| European countries | Mean | 82.2 |
|  | Median | 82.7 |
|  | Variance | 12.1 |
|  | Standard deviation | 3.5 |
|  | Minimum | 76.3 |
|  | Maximum | 84.8 |
|  | Range | 8.5 |
|  | Interquartile range | 4.3 |
| Non-European countries | Mean | 83.3 |
|  | Median | 83.5 |
|  | Variance | 6.7 |
|  | Standard deviation | 2.6 |
|  | Minimum | 79.2 |
|  | Maximum | 86.4 |
|  | Range | 7.2 |
|  | Interquartile range | 2.5 |

8. Of the 70 countries in our sample, males had a higher rate of labor force participation than females, according to the mean values. That is, 79.44 percent of men ages 15-64 were formally employed while 57.02 women abided by the same conditions. Women had greater variation in labor force participation, such that the standard deviation for women (18.54) was approximately three times the standard deviation for men (6.20). One possible reason for the lower rates of labor force participation for females would be their gender role as homemakers, in that some countries socialize their women to stay at home in order to take care of children. However, the standard deviation for women also suggests that labor force participation varies widely throughout the world. In some countries, women are encouraged to work in the formal labor force as much as men (e.g., the United States), while others strictly forbid women from working (e.g., Saudi Arabia).

9. 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 more data from a number of groups in order 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 versus 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). Again, for this variable, 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 remember: we can’t be certain this is the case for all Americans!).

10.

1. Age

Range = 89 – 18 = 71; IQR = 62 – 35 = 27; standard deviation = 17.557

Age when 1st child born

Range = 53 – 13 = 40; IQR = 27 – 20 = 7; standard deviation = 5.933

1. On all three measures, respondents’ age has more variability than their age when their first child was born. The range is 71, compared to only 40; the IQR is 27, compared to only 7; and the standard deviation is 17.557, compared to only 5.933. This is due in large part to biological constraints on the years one can bear children and, to a lesser extent, societal norms about what is the appropriate age to become a parent.