

# Answers to Odd-Numbered "Practice Using Statistical Software (IBM SPSS)" Questions: Full Version

## Chapter 1

1. There are 27,534 cases in the dataset.
3.
  - a. The value that is recorded for the variable "Marital status of respondent" [MARSTAT] is "1".
  - b. The attribute that is associated with the value "1" is "Married." The person is married.
5.
  - a. The value that is recorded for the variable "Total household income - 2012" [INCMHSD] is "9".
  - b. The attribute that is associated with the value "9" is "\$50,000 to \$59,999." The combined income of all of the people in that person's household in 2012 was \$50,000 to \$59,999.
7.
  - a. The variable VBR\_10 captures whether or not people voted in the last federal election.
  - b. The variable has the following attributes: "Yes," "No," "Valid skip," "Don't know," "Refusal," and "Not stated." The value "1" is associated with the attribute "Yes," the value "2" is associated with the attribute "No," the value "6" is associated with the attribute "Valid skip," the value "7" is associated with the attribute "Don't know," the value "8" is associated with the attribute "Refusal," and the value "9" is associated with the attribute "Not stated."
  - c. The attributes/values that are designated as missing are: "Valid skip" (6), "Don't know" (7), "Refusal" (8), and "Not stated" (9).
  - d. It is a dichotomous variable.
9.
  - a. The variable WHW\_120C captures the number of hours that people work at their jobs each week.
  - b. The values on this variable represent quantities.
  - c. The attributes/values that are designated as missing are: "Valid skip" (999.6), "Don't know" (999.7), "Refusal" (999.8), and "Not stated" (999.9).
  - d. It is a ratio-level variable.

## Chapter 2

### 1. Frequencies

**Statistics**

SEX Sex of respondent

N	Valid	27534
	Missing	0

**SEX Sex of respondent**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	12613	45.8	45.8	45.8
	2 Female	14921	54.2	54.2	100.0
Total		27534	100.0	100.0	

- Overall, 12,613 men answered the survey; 45.8 per cent of survey respondents are men.
- Overall, 14,921 women answered the survey; 54.2 per cent of survey respondents are women.

### 3. Frequencies

**Statistics**

SVR\_10 Canadian shared values - Human rights

N	Valid	26924
	Missing	610

**SVR\_10 Canadian shared values - Human rights**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 To a great extent	14891	54.1	55.3	55.3
	2 To a moderate extent	10483	38.1	38.9	94.2
	3 To a small extent	1250	4.5	4.6	98.9
	4 Not at all	300	1.1	1.1	100.0
	Total	26924	97.8	100.0	
Missing	7 Don't know	514	1.9		
	8 Refusal	96	.3		
	Total	610	2.2		
Total		27534	100.0		

- Among the people who gave a valid answer to the question, 55.3 per cent say that Canadians share the value of human rights to a great extent.
- Among the people who gave a valid answer to the question, 38.9 per cent say that Canadians share the value of human rights to a moderate extent.
- Among the people who gave a valid answer to the question, 94.2 per cent say that Canadians share the value of human rights to either a great or a moderate extent.

5. Frequencies

Statistics

PRD\_10 Pride - Being Canadian

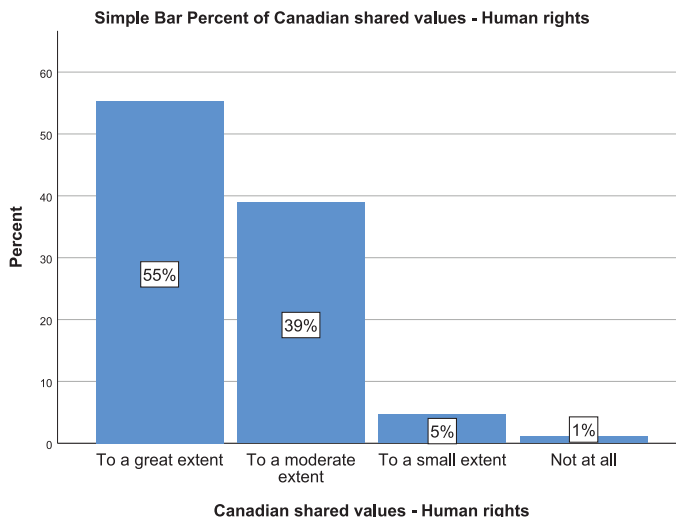
N	Valid	25404
	Missing	2130

PRD\_10 Pride - Being Canadian

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Very proud	15987	58.1	62.9	62.9
	2 Proud	6811	24.7	26.8	89.7
	3 Somewhat proud	1971	7.2	7.8	97.5
	4 Not very proud	461	1.7	1.8	99.3
	5 Not proud at all	174	.6	.7	100.0
	Total	25404	92.3	100.0	
Missing	6 No opinion	372	1.4		
	7 Not a Canadian citizen	1686	6.1		
	97 Don't know	24	.1		
	98 Refusal	48	.2		
	Total	2130	7.7		
Total		27534	100.0		

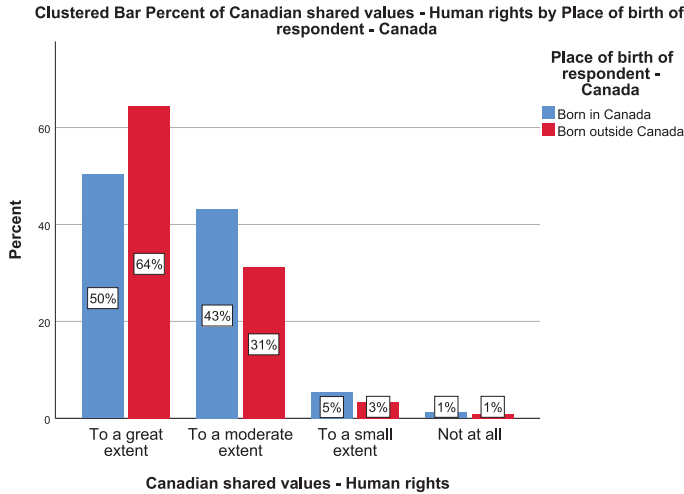
- a. Among the people who gave a valid answer to the question, 62.9 per cent say that they are very proud to be Canadian.
- b. The percentage from question 4(a) is different than the percentage from question 5(a) because a different denominator is used to calculate it. In question 4(a) people who had “No opinion” and who are “Not a Canadian citizen” are included in the denominator used to calculate the percentage, whereas in question 5(a) people with these two attributes are excluded from the denominator used to calculate the percentage.

7. a. GGraph



- b. This bar graph displays the information more effectively than the pie graph from question 6, because the bar graph makes it easy to see that the percentage of people who gave each answer gets smaller as you move from the highest category to the lowest category.

9. a. **GGraph**



- b. This graph shows that people born outside Canada are more likely than people born in Canada to say that Canadians share the value of human rights to a great extent. In contrast, people born in Canada are more likely than people born outside Canada to say that Canadians share the value of human rights to a moderate or a small extent.

## Chapter 3

### 1. Frequencies

**Statistics**

LIVARR06 Living arrangement of respondent's household (6 categories)

N	Valid	27534
	Missing	0

**LIVARR06 Living arrangement of respondent's household (6 categories)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 Spouse only	7629	27.7	27.7	27.7
	3 Spouse and single/non-single child(ren)	7044	25.6	25.6	53.3
	1 Alone	6423	23.3	23.3	76.6
	5 Living with one or two parents	3733	13.6	13.6	90.2
	4 Single/non-single child(ren) only	1431	5.2	5.2	95.4
	6 Other living arrangement	1274	4.6	4.6	100.0
Total		27534	100.0	100.0	

The mode is “Spouse only.” This shows that the largest number of people report living with only their spouse.

3. **Frequencies****Statistics**

INCM Annual personal income of the respondent - 2012

N	Valid	20541
	Missing	6993

**INCM Annual personal income of the respondent - 2012**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No income	1855	6.7	9.0	9.0
	2 Less than \$ 5,000	741	2.7	3.6	12.6
	3 \$ 5,000 to \$ 9,999	827	3.0	4.0	16.7
	4 \$ 10,000 to \$ 14,999	1403	5.1	6.8	23.5
	5 \$ 15,000 to \$ 19,999	1228	4.5	6.0	29.5
	6 \$ 20,000 to \$29,999	2530	9.2	12.3	41.8
	7 \$ 30,000 to \$ 39,999	2624	9.5	12.8	54.6
	8 \$ 40,000 to \$ 49,999	2186	7.9	10.6	65.2
	9 \$ 50,000 to \$ 59,999	1756	6.4	8.5	73.8
	10 \$ 60,000 to \$ 79,999	2419	8.8	11.8	85.5
	11 \$ 80,000 to \$ 99,999	1238	4.5	6.0	91.6
	12 \$ 100,000 or more	1734	6.3	8.4	100.0
		Total	20541	74.6	100.0
Missing	97 Don't know	4732	17.2		
	98 Refusal	1783	6.5		
	99 Not stated	478	1.7		
	Total	6993	25.4		
	Total	27534	100.0		

The mode is “\$30,000 to \$39,999.” This shows that it is most common for people to report having an annual income of \$30,000 to \$39,999.

The median is also “\$30,000 to \$39,999.” This shows that half of people report an annual income of \$30,000 to \$39,999 or less, and half of people report an annual income of \$30,000 to \$39,999 or more. (Or, half of people report an annual income of \$39,999 or less and half of people report an annual income of \$30,000 or more.)

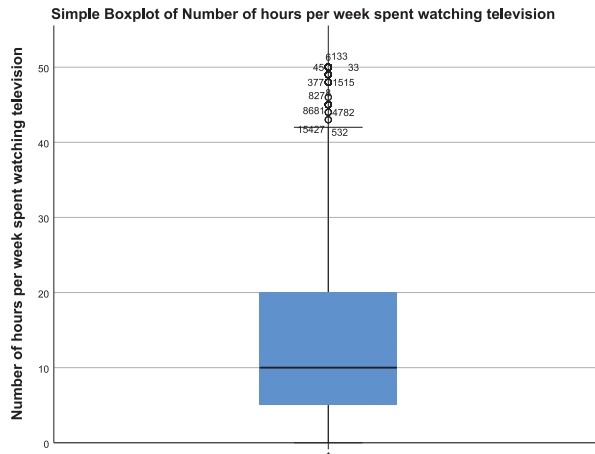
5. **Frequencies****Statistics**

INCM Annual personal income of the respondent - 2012

N	Valid	20541
	Missing	6993
Percentiles	25	5.00
	50	7.00
	75	10.00

The interquartile range of this variable is from “\$15,000 to \$19,999” (the attribute associated with the value “5”) to “\$60,000 to \$79,999” (the attribute associated with the value “10”). This shows that the middle 50 per cent of people have an annual income between \$15,000 and \$79,999. In other words, the 50 per cent of people in the middle of the income distribution have annual incomes between \$15,000 and \$79,999.

7. GGraph



- a. The twenty-fifth percentile is 5 hours, the fiftieth percentile is 10 hours, and the seventy-fifth percentile is 20 hours.
- b. The interquartile range is from 5 to 20 hours. Excluding outliers, the range is from 0 to 42 hours.

9. Frequencies

		Statistics	
		INCM Annual personal income of the respondent - 2012	INCM_RECODED Annual personal income of the respondent - 2012 (recoded)
N	Valid	20541	20541
	Missing	6993	6993

Frequency Table

INCM Annual personal income of the respondent - 2012					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No income	1855	6.7	9.0	9.0
	2 Less than \$ 5,000	741	2.7	3.6	12.6
	3 \$ 5,000 to \$ 9,999	827	3.0	4.0	16.7
	4 \$ 10,000 to \$ 14,999	1403	5.1	6.8	23.5
	5 \$ 15,000 to \$ 19,999	1228	4.5	6.0	29.5
	6 \$ 20,000 to \$29,999	2530	9.2	12.3	41.8
	7 \$ 30,000 to \$ 39,999	2624	9.5	12.8	54.6
	8 \$ 40,000 to \$ 49,999	2186	7.9	10.6	65.2
	9 \$ 50,000 to \$ 59,999	1756	6.4	8.5	73.8
	10 \$ 60,000 to \$ 79,999	2419	8.8	11.8	85.5
	11 \$ 80,000 to \$ 99,999	1238	4.5	6.0	91.6
	12 \$ 100,000 or more	1734	6.3	8.4	100.0
Total		20541	74.6	100.0	
Missing	97 Don't know	4732	17.2		
	98 Refusal	1783	6.5		
	99 Not stated	478	1.7		
	Total	6993	25.4		
Total		27534	100.0		

**INCM\_RECoded Annual personal income of the respondent - 2012  
(recoded)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00 People with no income	1855	6.7	9.0	9.0
	2.00 People with an income from \$1 to \$19,999	4199	15.3	20.4	29.5
	3.00 People with an income from \$20,000 to \$39,999	5154	18.7	25.1	54.6
	4.00 People with an income from \$40,000 to \$59,999	3942	14.3	19.2	73.8
	5.00 People with an income from \$60,000 to \$79,999	2419	8.8	11.8	85.5
	6.00 People with an income of \$80,000 or more	2972	10.8	14.5	100.0
	Total	20541	74.6	100.0	
Missing	9.00 People with a 'Missing' answer	6993	25.4		
Total		27534	100.0		

The new, recoded variable shows that 9.0 per cent of people have no annual personal income. About one in five people (20.4 per cent) have an annual income between \$1 and \$19,999. A slightly higher percentage of people—25.1 per cent—have an annual income between \$20,000 and \$39,999. Another 19.2 per cent of people have an annual income from \$40,000 to \$59,999 and the remaining 26.2 per cent have higher annual incomes. When the variable is recoded this way, it shows that the most common annual personal income is \$20,000 to \$39,999; this is also the median annual personal income.

## Chapter 4

### 1. Means

**Case Processing Summary**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
WKWEHRC Number of paid hours worked per week - All jobs	18181	66.0%	9353	34.0%	27534	100.0%

### Report

WKWEHRC Number of paid hours worked per week - All jobs

Mean	N	Std. Deviation
37.800	18181	14.3795

The mean is 37.80. This shows that, on average, people work at their jobs for 37.80 paid hours per week (when people who do not work for pay are excluded). The standard deviation is 14.38. Since the standard deviation is relatively small compared to the mean, it shows that the distribution of the “Number of paid hours worked per week” variable isn’t widely spread out.

3. **Frequencies**

**Statistics**

WKWEHRC Number of paid hours worked per week - All jobs

N	Valid	18181
	Missing	9353
Mean		37.800
Median		40.000
Skewness		-.113
Std. Error of Skewness		.018
Kurtosis		.713
Std. Error of Kurtosis		.036

- a. The median is 40.00. This shows that half of people work 40 paid hours or more at their jobs per week, and half of people work 40 paid hours or fewer at their jobs per week. The median is slightly higher than the mean, but not by much.
- b. The kurtosis is 0.71. This shows that the distribution of the “Number of paid hours worked per week” variable is moderately leptokurtic. In other words, it’s slightly more peaked than a normal distribution.
- c. The skew is  $-0.11$ . This shows that the distribution of the “Number of paid hours worked per week” variable is very slightly left-skewed compared to a normal distribution. In other words, it has a tail that extends very slightly to the left, because some people work a very low number of paid hours at their jobs each week. But, since the skew is between  $-0.5$  and  $+0.5$ , the distribution of this variable is considered approximately normal in terms of its skew.

5. **Means**

**Case Processing Summary**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
SCF_100C Number of close friends	27112	98.5%	422	1.5%	27534	100.0%

**Report**

SCF\_100C Number of close friends

Mean	N	Std. Deviation
6.48	27112	9.552

The mean is 6.48. This shows that, on average, people have 6.48 close friends. The standard deviation is 9.55. Since the standard deviation is substantially larger than the mean, it shows that the “Number of close friends” variable has a wide spread.



7. **Frequencies**

Statistics		
SCF_100C Number of close friends		
N	Valid	27112
	Missing	422
Mean		6.48
Median		5.00
Mode		5
Std. Deviation		9.552
Skewness		9.255
Std. Error of Skewness		.015
Kurtosis		141.209
Std. Error of Kurtosis		.030
Range		200

- a. The mean is 6.48. The median and the mode are both 5.00. The mean shows that, on average, people have 6.48 close friends. The median shows that half of people have 5 close friends or more, and half of people have 5 close friends or fewer. The mode shows that it is most common for people to have 5 close friends.
- b. The standard deviation is 9.55 and the range is 200. Both of these statistics show that the “Number of close friends” variable is widely dispersed.
- c. The kurtosis is 141.21. This shows that the distribution of the “Number of close friends” variable is highly leptokurtic; in other words, the distribution is very peaked compared to a normal distribution.
- d. The skew is 9.26. It shows that the distribution of the “Number of close friends” variable is highly right-skewed compared to a normal distribution. In other words, the distribution has a tail that extends far to the right because some people report very high numbers of close friends.

## Chapter 5

1. a. **Means**

	Case Processing Summary					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
WGHT_PER Person weight	27534	100.0%	0	0.0%	27534	100.0%

Report		
WGHT_PER Person weight		
Mean	N	Std. Deviation
1058.034174	27534	1044.350335

b. <No output>

c. **Means**

**Case Processing Summary**

	Included		Cases Excluded		Total	
	N	Percent	N	Percent	N	Percent
STD_WGHT Standardized person weight	27534	100.0%	0	0.0%	27534	100.0%

**Report**

STD_WGHT Standardized person weight		
Mean	N	Std. Deviation
1.0000	27534	.98707

3. a. **Frequencies**

**Statistics**

SEX Sex of respondent		
N	Valid	27534
	Missing	0

**SEX Sex of respondent**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	12613	45.8	45.8	45.8
	2 Female	14921	54.2	54.2	100.0
	Total	27534	100.0	100.0	

b. **Frequencies**

**Statistics**

SEX Sex of respondent		
N	Valid	29131913
	Missing	0

**SEX Sex of respondent**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	14388232	49.4	49.4	49.4
	2 Female	14743681	50.6	50.6	100.0
	Total	29131913	100.0	100.0	

c. **Frequencies****Statistics**

SEX Sex of respondent

N	Valid	27534
	Missing	0

**SEX Sex of respondent**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	13599	49.4	49.4	49.4
	2 Female	13935	50.6	50.6	100.0
Total		27534	100.0	100.0	

- d. The frequency distributions in (a) and (c) have the same total number of cases, which is the same as the number of cases in the dataset, whereas the frequency distribution in (b) has 29 million cases that represent the Canadian population. The frequency distributions in (b) and (c) have the same percentages of men and women, whereas the frequency distribution in (a) has different percentages of men and women.

**Chapter 6**1. **Explore****Case Processing Summary**

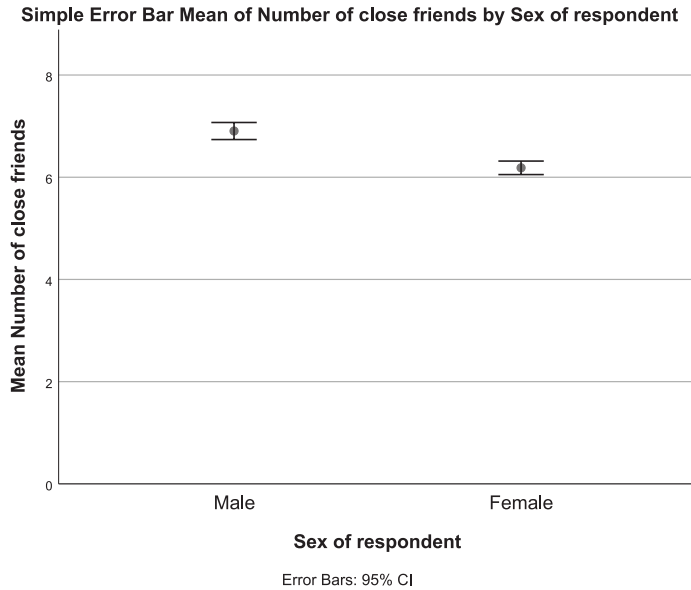
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SCF_100C Number of close friends	27166	98.7%	368	1.3%	27534	100.0%

**Descriptives**

		Statistic	Std. Error	
SCF_100C Number of close friends	Mean	6.54	.054	
	95% Confidence Interval for Mean	Lower Bound	6.43	
		Upper Bound	6.65	
	5% Trimmed Mean	5.46		
	Median	5.00		
	Variance	79.824		
	Std. Deviation	8.934		
	Minimum	0		
	Maximum	200		
	Range	200		
	Interquartile Range	5		
	Skewness	8.641	.015	
	Kurtosis	133.857	.030	

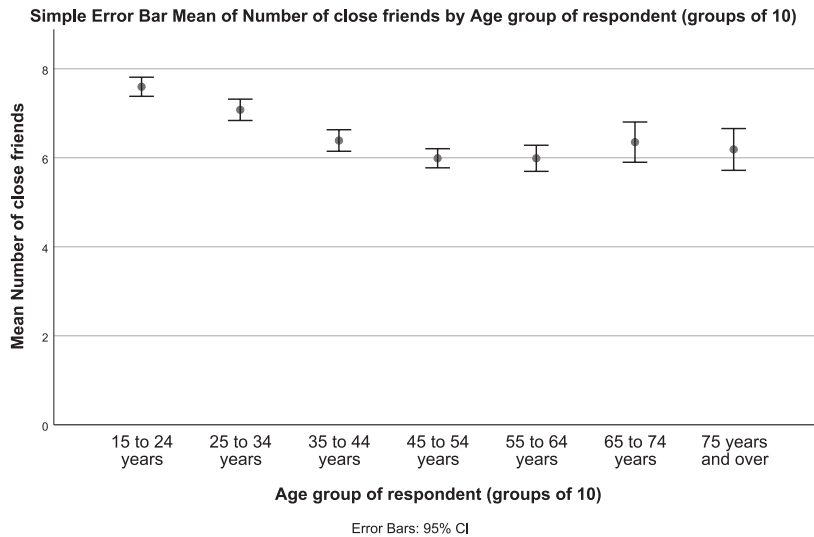
- a. The mean is 6.54. In the sample, on average, people have 6.54 close friends.  
 b. The 95 per cent confidence interval for the mean is 6.43 to 6.65. In the population, the average number of close friends is likely between 6.43 and 6.65.

3. **GGraph**



The round dots for men and women show the mean number of close friends for men and women in the sample, as reported in the statistics from question 2: 6.90 and 6.18, respectively. The whiskers below and above each dot show the distance between the lower bound and the upper bound of the 95 per cent confidence interval for the mean for men and women, as reported in the statistics from question 2. So, for men, the whiskers extend from 6.74 to 7.07, and for women the whiskers extend from 6.05 to 6.32.

5. **GGraph**



The round dots for each age group show the mean number of close friends for people in each age group in the sample, as reported in the statistics from question 4. For example, the lowest average number of close friends is among people aged 45 to 54 and among people aged 55 to 64 (5.99 for both groups), and the highest average number of close friends is among people aged 15 to 24 (7.60). The whiskers below and above each dot show the distance between the lower bound and the upper bound of the 95 per cent confidence interval for the mean for each age group, as reported in the statistics from question 4. So, for people aged 15 to 24, the whiskers extend from 7.38 to 7.81, whereas for people aged 75 and over, the whiskers extend from 5.72 to 6.66.

## 7. Frequencies

### Statistics

VCG\_300\_RECODED Volunteer work - 12 months (recoded)

N	Valid	27479
	Missing	55

### VCG\_300\_RECODED Volunteer work - 12 months (recoded)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	17713	64.3	64.5	64.5
	1.00	9766	35.5	35.5	100.0
	Total	27479	99.8	100.0	
Missing	System	55	.2		
Total		27534	100.0		

## Means

### Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
VCG_300_RECODED Volunteer work - 12 months (recoded)	27479	99.8%	55	0.2%	27534	100.0%

### Report

VCG\_300\_RECODED Volunteer work - 12 months (recoded)

Mean	N	Std. Deviation
.3554	27479	.47864

Overall, 35.5 per cent of people volunteered in the past 12 months. The mean of the recoded variable is 0.355. When 35.5 per cent is converted into a proportion (by dividing it by 100), the result is 0.355, which corresponds to the mean of the recoded variable.

9. Explore

SEX Sex of respondent

Case Processing Summary

	SEX Sex of respondent	Valid		Cases Missing		Total	
		N	Percent	N	Percent	N	Percent
VCG_300_RECODED	1 Male	13572	99.8%	27	0.2%	13599	100%
Volunteer work - 12 months (recoded)	2 Female	13907	99.8%	28	0.2%	13935	100%

Descriptives

	SEX Sex of respondent		Statistic	Std. Error
VCG_300_RECODED Volunteer work - 12 months (recoded)	1 Male	Mean	.3248	.00402
		95% Confidence Interval for Mean	Lower Bound	.3169
	Upper Bound		.3326	
	5% Trimmed Mean		.3053	
	Median	.0000		
	Variance	.219		
	Std. Deviation	.46830		
	Minimum	.00		
	Maximum	1.00		
	Range	1.00		
	Interquartile Range	1.00		
	Skewness	.749	.021	
	Kurtosis	-1.440	.042	
	2 Female	Mean		.3853
95% Confidence Interval for Mean			Lower Bound	.3772
		Upper Bound	.3934	
		5% Trimmed Mean	.3726	
Median		.0000		
Variance		.237		
Std. Deviation		.48669		
Minimum		.00		
Maximum		1.00		
Range		1.00		
Interquartile Range		1.00		
Skewness		.471	.021	
Kurtosis		-1.778	.042	

- a. The mean shows that the proportion of men in the sample who volunteered in the past 12 months is 0.3248, or 32.48 per cent. Similarly, the proportion of women in the sample who volunteered in the past 12 months is 0.3853, or 38.53 per cent.
- b. The 95 per cent confidence interval for the mean shows that the proportion of men in the population who volunteered in the past 12 months is likely to be between 0.3169 and 0.3326 (or 31.69 and 33.26 per cent). Similarly, the 95 per cent confidence interval for the mean shows that the proportion of women in the population who volunteered in the past 12 months is likely to be between 0.3772 and 0.3934 (or 37.72 and 39.34 per cent). Since these 95 confidence intervals do not overlap, in the population it is likely that a higher proportion of women than men volunteered in the past 12 months.

## Chapter 7

### 1. Means

#### Case Processing Summary

	Included		Cases Excluded		Total	
	N	Percent	N	Percent	N	Percent
SCF_100C Number of close friends * SEX Sex of respondent	27166	98.7%	368	1.3%	27534	100.0%

#### Report

SCF_100C Number of close friends			
SEX Sex of respondent	Mean	N	Std. Deviation
1 Male	6.90	13396	9.860
2 Female	6.18	13769	7.915
Total	6.54	27166	8.934

- a. In the sample, the difference between the mean number of close friends for men and for women is 0.72, or less than 1 close friend. This isn't a very large difference.
  - b. Cohen's  $d$  is 0.08. Since Cohen's  $d$  is less than 0.1, the effect size is very small. (In other words, the relationship between people's sex/gender and their number of close friends is very weak.)
3. a. The output from all three procedures display the mean, the number of cases, and the standard deviation for each group. The output from the Explore procedure and the Independent Samples T-Test procedure both display the standard error of the mean for each group.
  - b. The output from the Explore procedure displays additional statistics about the distribution of the variable within each group, and includes the median, the variance, the minimum, the maximum, the range, the interquartile range, the skew, and the kurtosis.

The output from the Means procedure is the only one to show the mean, the standard deviation, and the number of cases for the sample overall (not divided by group).

The output from the Independent Samples T-test procedure shows the  $t$ -statistic, the degrees of freedom, and the significance test associated with those results (for both versions of the  $t$ -test). It also shows the difference between the means, the standard error of the difference, and the 95 per cent confidence interval for the difference between means.

- c. Yes, the answers to the two questions correspond. In question 2(c) in this chapter, the  $t$ -test results show that there is likely a difference between the two group means in the population. In question 2(b) in Chapter 6, the 95 per cent confidence intervals for the mean do not overlap, suggesting that there is likely a difference between the two group means in the population.

5. T-Test

**Group Statistics**

PCT_10 Trust people in general		N	Mean	Std. Deviation	Std. Error Mean
SCF_100C Number of close friends	1 Most people can be trusted	14286	7.33	9.527	.080
	2 You cannot be too careful in dealing with people	12380	5.61	8.112	.073

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means				95% Confidence Interval of the Difference		
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SCF_100C Number of close friends	Equal variances assumed	106,360	.000	15,683	26664	.000	1,714	.109	1,499	1,928
	Equal variances not assumed			15,863	26656,259	.000	1,714	.108	1,502	1,925

- a. A non-directional research hypothesis for this relationship is this: “In the population, there is a relationship between people’s general orientation toward trusting people and their number of close friends.” (An alternative non-directional hypothesis is this: “In the population, those who think that most people can be trusted have a different number of close friends, on average, than those who think you cannot be too careful in dealing with people.”)
- b. The null hypothesis associated with this research hypothesis is this: “In the population, there is no relationship between people’s general orientation toward trusting people and their number of close friends.” (An alternative null hypothesis is this: “In the population, those who think that most people can be trusted have the same number of close friends, on average, as those who think you cannot be too careful in dealing with people.”)
- c. The t-statistic of 15.86 has a p-value that is less than 0.05, so the null hypothesis is rejected. In the population, there is likely a relationship between people’s general orientation toward trusting people and their number of close friends.

7. T-Test

**Group Statistics**

DH1GED Education - Highest degree (4 categories)		N	Mean	Std. Deviation	Std. Error Mean
SCF_100C Number of close friends	>= 3	15779	6.78	9.057	.072
	< 3	11213	6.19	8.624	.081

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means				95% Confidence Interval of the Difference		
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SCF_100C Number of close friends	Equal variances assumed	5,363	.021	5,380	26990	.000	.590	.110	.375	.805
	Equal variances not assumed			5,425	24834,615	.000	.590	.109	.377	.803



- a. A non-directional research hypothesis for this relationship is this: “In the population, there is a relationship between having a post-secondary education (or not) and the number of close friends that people have.” (An alternative non-directional hypothesis is this: “In the population, people with a post-secondary education have a different number of close friends, on average, than people who do not have a post-secondary education.”)
- b. The null hypothesis associated with this research hypothesis is this: “In the population, there is no relationship between having a post-secondary education (or not) and the number of close friends that people have.” (An alternative null hypothesis is this: “In the population, people with a post-secondary education have the same number of close friends, on average, as people who do not have a post-secondary education.”)
- c. The t-statistic of 5.43 has a p-value that is less than 0.05 so the null hypothesis is rejected. In the population, there is likely a relationship between having a post-secondary education (or not) and the number of close friends that people have.

## Chapter 8

### 1. Oneway

**Descriptives**

SCF\_100C Number of close friends

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1 Less than High School	4011	6.19	9.831	.155	5.88	6.49	0	200
2 Graduated from High School	7202	6.19	7.873	.093	6.01	6.37	0	200
3 Post-secondary diploma	8600	6.44	9.262	.100	6.25	6.64	0	200
4 University degree	7179	7.18	8.789	.104	6.98	7.39	0	200
Total	26992	6.53	8.884	.054	6.43	6.64	0	200

**ANOVA**

SCF\_100C Number of close friends

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4423.794	3	1474.598	18.719	.000
Within Groups	2125962.559	26988	78.774		
Total	2130386.353	26991			

- a. In the sample, the average number of close friends among people with less than a high school education and among people with only a high school education is the same: 6.19. Among people with a post-secondary diploma, the average number of close friends is 0.25 higher. People with a university degree have one more close friend (0.99), on average, than people with only a high school education or less than a high school education.

- b. For people with the three lowest levels of education (less than high school, high school only, and post-secondary diploma), the 95 per cent confidence intervals for the mean number of close friends all overlap. As a result, we cannot be confident that, in the population, the average number of close friends among people with each of these three levels of education is different. However, among people with a university degree, the lower bound of the 95 per cent confidence interval for the mean is higher than the upper bound of the 95 per cent confidence interval for the mean for the other three levels of education. Thus, it is likely that, in the population, people with a university degree have more close friends, on average, than people with lower levels of education.

3. Post Hoc Tests

Multiple Comparisons

Dependent Variable: SCF\_100C Number of close friends

LSD

(I) DH1GED Education - Highest degree (4 categories)	(J) DH1GED Education - Highest degree (4 categories)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1 Less than High School	2 Graduated from High School	-.005	.175	.975	-.35	.34
	3 Post-secondary diploma	-.257	.170	.130	-.59	.08
	4 University degree	-.997*	.175	.000	-1.34	-.65
2 Graduated from High School	1 Less than High School	.005	.175	.975	-.34	.35
	3 Post-secondary diploma	-.252	.142	.076	-.53	.03
	4 University degree	-.991*	.148	.000	-1.28	-.70
3 Post-secondary diploma	1 Less than High School	.257	.170	.130	-.08	.59
	2 Graduated from High School	.252	.142	.076	-.03	.53
	4 University degree	-.740*	.142	.000	-1.02	-.46
4 University degree	1 Less than High School	.997*	.175	.000	.65	1.34
	2 Graduated from High School	.991*	.148	.000	.70	1.28
	3 Post-secondary diploma	.740*	.142	.000	.46	1.02

\*. The mean difference is significant at the 0.05 level.

- a. The post-hoc tests shows that the average number of close friends among people with a university degree is significantly different than the average number of close friends among people with lower levels of education. The p-values of all of the significance tests that include the university degree group are less than 0.05, whereas the others are greater than 0.05.
- b. Yes, the answers to the two questions correspond. In (a) the post-hoc tests show that, in the population, the average number of close friends among people with a university degree is likely different than the average number of close friends among people with lower levels of education. In question

1(b), the 95 per cent confidence intervals for the mean suggest that, in the population, people with a university degree have more close friends, on average, than people with lower levels of education.

### 5. Oneway

#### Descriptives

WKWEHRC Number of paid hours worked per week - All jobs

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1 Less than High School	1768	33.118	18.1116	.4307	32.273	33.963	.1	75.0
2 Graduated from High School	5075	36.824	14.4894	.2034	36.425	37.223	.1	75.0
3 Post-secondary diploma	6592	38.931	13.1932	.1625	38.613	39.250	.1	75.0
4 University degree	5793	39.598	13.2020	.1735	39.258	39.938	.1	75.0
Total	19228	38.041	14.1896	.1023	37.841	38.242	.1	75.0

#### ANOVA

WKWEHRC Number of paid hours worked per week - All jobs

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	69633.368	3	23211.123	117.372	.000
Within Groups	3801684.940	19224	197.757		
Total	3871318.307	19227			

- a. In the overall sample, people spend an average of 38.04 hours per week doing paid work. In the overall population, the average number of hours that people spend doing paid work each week is likely between 37.84 and 38.24.
- b. In the sample, people with less than a high school education spend an average of 33.12 hours per week doing paid work. People with only a high school education spend an average of 36.82 hours per week doing paid work—3.70 hours more than the average of people with less than a high school education. People with a post-secondary diploma spend an average of 38.93 hours per week doing paid work, compared to 39.60 hours among people with a university degree; there is only a small difference (0.67 hours) between these two averages in the sample.
- c. For the four educational groups, none of the 95 per cent confidence intervals for the mean number of hours spent doing paid work overlap. Thus, it is likely that, in the population, people with less than a high school education spend the lowest number of hours doing paid work each week, on average. People with only a high school education spend slightly more hours doing paid work per week, on average, followed by people with a post-secondary diploma. People with a university degree are likely to

spend the highest number of hours doing paid work each week, on average. Thus, in the population, people with higher levels of education are likely to spend more time doing paid work each week, on average, than people with lower levels of education.

## 7. Post Hoc Tests

### Multiple Comparisons

Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

LSD

(I) DH1GED Education - Highest degree (4 categories)	(J) DH1GED Education - Highest degree (4 categories)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1 Less than High School	2 Graduated from High School	-3.7064*	.3884	.000	-4.468	-2.945
	3 Post-secondary diploma	-5.8135*	.3766	.000	-6.552	-5.075
	4 University degree	-6.4801*	.3821	.000	-7.229	-5.731
2 Graduated from High School	1 Less than High School	3.7064*	.3884	.000	2.945	4.468
	3 Post-secondary diploma	-2.1071*	.2626	.000	-2.622	-1.592
	4 University degree	-2.7736*	.2704	.000	-3.304	-2.244
3 Post-secondary diploma	1 Less than High School	5.8135*	.3766	.000	5.075	6.552
	2 Graduated from High School	2.1071*	.2626	.000	1.592	2.622
	4 University degree	-.6665*	.2533	.008	-1.163	-.170
4 University degree	1 Less than High School	6.4801*	.3821	.000	5.731	7.229
	2 Graduated from High School	2.7736*	.2704	.000	2.244	3.304
	3 Post-secondary diploma	.6665*	.2533	.008	.170	1.163

\*. The mean difference is significant at the 0.05 level.

- The post-hoc tests show that the average number of hours spent doing paid work each week among people in each educational group is significantly different than the average number of hours spent doing paid work each week in every other educational group. The p-values of all of the significance tests are less than 0.05.
- Yes, the answers to the two questions correspond. In (a), the post-hoc tests show that, in the population, people in each educational group likely spend a different number of hours doing paid work each week, on average, than people in every other educational group. In question 5(c), the 95 per cent confidence intervals for the mean suggest that, in the population, people with higher levels of education are likely to spend more time doing paid work each week, on average, than people with lower levels of education.

## Chapter 9

### 1. Crosstabs

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VCG_340 Donated money or goods - 12 months * SEX Sex of respondent	27428	99.6%	106.000	0.4%	27534.000	100.0%

#### VCG\_340 Donated money or goods - 12 months \* SEX Sex of respondent Crosstabulation

			SEX Sex of respondent		
			1 Male	2 Female	Total
VCG_340 Donated money or goods - 12 months	1 Yes	Count	9525	11101	20626
		Expected Count	10179.9	10446.1	20626.0
		% within SEX Sex of respondent	70.4%	79.9%	75.2%
	2 No	Count	4012	2790	6802
		Expected Count	3357.1	3444.9	6802.0
		% within SEX Sex of respondent	29.6%	20.1%	24.8%
Total	Count	13537	13891	27428	
	Expected Count	13537.0	13891.0	27428.0	
	% within SEX Sex of respondent	100.0%	100.0%	100.0%	

In the sample, 79.9 per cent of women donated money or goods in the past 12 months, compared to only 70.4 per cent of men, a difference of 9.5 percentage points.

### 3. Crosstabs

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VCG_340 Donated money or goods - 12 months * SEX Sex of respondent	27428	99.6%	106.000	0.4%	27534.000	100.0%

#### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	335.443 <sup>a</sup>	1	.000		
Continuity Correction <sup>b</sup>	334.931	1	.000		
Likelihood Ratio	336.701	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	335.430	1	.000		
N of Valid Cases	27428				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 3357.10.

b. Computed only for a 2x2 table

- a. A non-directional research hypothesis for this relationship is this: "In the population, people's sex/gender is related to whether or not they donated money or goods in the past 12 months."

- b. The null hypothesis associated with this research hypothesis is this: “There is no relationship in the population between people’s sex/gender and whether or not they donated money or goods in the past 12 months.”
- c. The chi-square statistic of 335.44 has a p-value that is less than 0.05 so the null hypothesis is rejected. In the population, there is likely a relationship between people’s sex/gender and whether or not they donated money or goods in the past 12 months.

5. **Crosstabs**

**VISMIN Visible minority status of the respondent. = 1 Visible minority**

**Case Processing Summary<sup>a</sup>**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VCG_340 Donated money or goods - 12 months * SEX Sex of respondent	4496	99,7%	14,350	0,3%	4510,350	100,0%

a. VISMIN Visible minority status of the respondent. = 1 Visible minority

**VCG\_340 Donated money or goods - 12 months \* SEX  
Sex of respondent Crosstabulation<sup>a</sup>**

% within SEX Sex of respondent

		SEX Sex of respondent		
		1 Male	2 Female	Total
		VCG_340 Donated money or goods - 12 months	1 Yes	64,1%
	2 No	35,9%	25,0%	30,5%
<b>Total</b>		<b>100,0%</b>	<b>100,0%</b>	<b>100,0%</b>

a. VISMIN Visible minority status of the respondent. = 1 Visible minority

**VISMIN Visible minority status of the respondent. = 2 Not a visible minority**

**Case Processing Summary<sup>a</sup>**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VCG_340 Donated money or goods - 12 months * SEX Sex of respondent	22590	99,6%	81,912	0,4%	22671,912	100,0%

a. VISMIN Visible minority status of the respondent. = 2 Not a visible minority

**VCG\_340 Donated money or goods - 12 months \* SEX  
Sex of respondent Crosstabulation<sup>a</sup>**

% within SEX Sex of respondent

		SEX Sex of respondent		
		1 Male	2 Female	Total
		VCG_340 Donated money or goods - 12 months	1 Yes	71,7%
	2 No	28,3%	19,1%	23,6%
<b>Total</b>		<b>100,0%</b>	<b>100,0%</b>	<b>100,0%</b>

a. VISMIN Visible minority status of the respondent. = 2 Not a visible minority

- a. The relationship between sex/gender and making a donation does not change that much when visible minority status is taken into account. Among people in the sample who are visible minorities, 64.1 per cent of men donated money or goods in the past 12 months, compared to 75.0 per cent of women, for a difference of 10.9 percentage points. Among people

in the sample who are not visible minorities, 71.7 per cent of men donated goods or services in the past 12 months, compared to 80.9 per cent of women, for a difference of 9.2 percentage points.

- b. In the zero-order relationship shown in question 1, there is a 9.5 percentage point difference between men and women. The two partial relationships are fairly similar to the zero-order relationship: 10.9 and 9.2. So, this is a weak example of specification: the relationship between sex/gender and donating money or goods in the past 12 months is slightly stronger among people who are visible minorities than among people who are not visible minorities. (Alternatively, you could argue that this is an example of replication, since the difference between the partial relationships and the zero-order relationship is relatively small.)

## 7. Crosstabs

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VBR_25 Federal election - Vote in next election * REP_05 Interest in politics	26061	94.7%	1473.000	5.3%	27534.000	100.0%

### Symmetric Measures

		Value	Asymptotic Standard Error <sup>a</sup>	Approximate T <sup>b</sup>	Approximate Significance
Ordinal by Ordinal	Gamma	.611	.007	62.682	.000
N of Valid Cases		26061			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

The gamma of 0.611 shows that the error in predicting how likely people are to vote in the next federal election can be reduced by 61.1 per cent if we know how interested they are in politics.

## 9. Crosstabs

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VCG_340 Donated money or goods - 12 months * SEX Sex of respondent	27428	99.6%	106.000	0.4%	27534.000	100.0%

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-.111	.000
	Cramer's V	.111	.000
N of Valid Cases		27428	

Cramér’s V is 0.111. Since Cramér’s V is close to 0.1, the effect size is small. (In other words, the relationship between people’s sex/gender and whether or not they donated money or goods in the past 12 months is weak.)

## Chapter 10

### 1. Correlations

**Correlations**

		RFE_10C Number of relatives respondent feels close to	SCF_100C Number of close friends
RFE_10C Number of relatives respondent feels close to	Pearson Correlation	1	.258**
	Sig. (2-tailed)		.000
	N	26920	26693
SCF_100C Number of close friends	Pearson Correlation	.258**	1
	Sig. (2-tailed)	.000	
	N	26693	27166

\*\* . Correlation is significant at the 0.01 level (2-tailed).

- a. The Pearson’s correlation coefficient for the relationship between these two variables is 0.26. Since the Pearson’s correlation coefficient is less than 0.3, the relationship between the number of relatives people feel close to and their number of close friends is weak.
- b. The direction of the relationship is positive. In the context of these two variables, this indicates that people who feel close to more relatives tend to have more close friends. Conversely, people who feel close to fewer relatives tend to have fewer close friends.

### 3. Nonparametric Correlations

**Correlations**

		RFE_10C Number of relatives respondent feels close to	SCF_100C Number of close friends
Spearman’s rho	RFE_10C Number of relatives respondent feels close to	Correlation Coefficient	1.000
		Sig. (2-tailed)	.
		N	26501
	SCF_100C Number of close friends	Correlation Coefficient	.360**
		Sig. (2-tailed)	.000
		N	26277

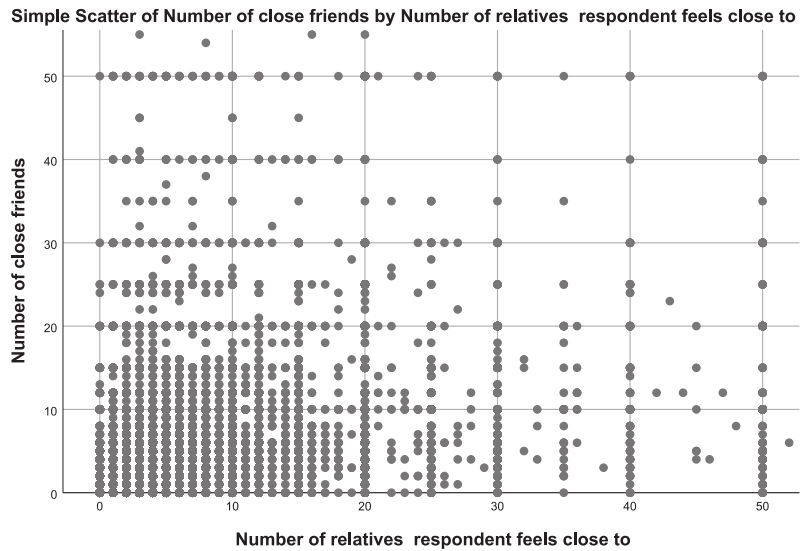
\*\* . Correlation is significant at the 0.01 level (2-tailed).

- a. The Spearman’s correlation coefficient for the relationship between these two variables is 0.36. Since the Spearman’s correlation coefficient is between 0.3 and 0.5, the rank-order relationship between the number of relatives people feel close to and their number of close friends is weak to moderate.



- b. The direction of the rank-order relationship is positive. In the context of these two variables, this indicates that people who ranked higher in terms of the number of relatives they feel close to tend to be ranked higher in terms of their number of close friends. Conversely, people who ranked lower in terms of the number of relatives they feel close to tend to be ranked lower in terms of their number of close friends.

5. a. GGraph



- b. The scatterplot shows that many people have low values on both variables; that is, they have relatively few relatives that they feel close to and relatively few close friends. But it's hard to discern the overall pattern of the relationship between the two variables since there are many overlapping cases.

7. Correlations

SEX Sex of respondent = 1 Male

Correlations<sup>a</sup>

		RFE_10C Number of relatives respondent feels close to	SCF_100C Number of close friends
RFE_10C Number of relatives respondent feels close to	Pearson Correlation	1	.226**
	Sig. (2-tailed)		.000
	N	13274	13131
SCF_100C Number of close friends	Pearson Correlation	.226**	1
	Sig. (2-tailed)	.000	
	N	13131	13396

\*\* . Correlation is significant at the 0.01 level (2-tailed).

a. SEX Sex of respondent = 1 Male

**SEX Sex of respondent = 2 Female**

**Correlations<sup>a</sup>**

		RFE_10C Number of relatives respondent feels close to	SCF_100C Number of close friends
RFE_10C Number of relatives respondent feels close to	Pearson Correlation	1	.303**
	Sig. (2-tailed)		.000
	N	13646	13561
SCF_100C Number of close friends	Pearson Correlation	.303**	1
	Sig. (2-tailed)	.000	
	N	13561	13769

\*\* . Correlation is significant at the 0.01 level (2-tailed).

a. SEX Sex of respondent = 2 Female

- a. When sex/gender is taken into account, the relationship between the number of relatives people feel close to and the number of close friends they have becomes weaker among men and stronger among women. Among men, the correlation between the number of relatives people feel close to and their number of close friends is 0.23. Among women, the correlation between the same two variables is 0.30.
- b. The zero-order correlation in question 1 is 0.26. The partial correlation among men (0.23) is smaller than the zero-order correlation, and the partial correlation among women (0.30) is larger than the zero-order correlation. So, this is an example of specification. The relationship between the number of relatives people feel close to and the number of close friends they have is weaker among men and stronger among women.

## Chapter 11

1. a-b. <No output>
- c. **Frequencies**

**Statistics**

		GRP_10C Number of groups - 12 months	GRP_10C_RECODED Number of groups - 12 months (recoded)
N	Valid	17902	27395
	Missing	9632	139

**Frequency Table**

**GRP\_10C Number of groups - 12 months**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	7393	26.8	41.3	41.3
	2	4487	16.3	25.1	66.4
	3	2638	9.6	14.7	81.1

	4	1534	5.6	8.6	89.7
	5	860	3.1	4.8	94.5
	6	420	1.5	2.3	96.8
	7	149	.5	.8	97.6
	8	113	.4	.6	98.3
	9 9 groups or more	308	1.1	1.7	100.0
	Total	17902	65.0	100.0	
Missing	96 Valid skip	9493	34.5		
	97 Don't know	62	.2		
	98 Refusal	70	.3		
	99 Not stated	7	.0		
	Total	9632	35.0		
Total		27534	100.0		

**GRP\_10C\_RECODED Number of groups - 12 months  
(recoded)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	9493	34.5	34.7	34.7
	1.00	7393	26.8	27.0	61.6
	2.00	4487	16.3	16.4	78.0
	3.00	2638	9.6	9.6	87.6
	4.00	1534	5.6	5.6	93.2
	5.00	860	3.1	3.1	96.4
	6.00	420	1.5	1.5	97.9
	7.00	149	.5	.5	98.5
	8.00	113	.4	.4	98.9
	9.00	308	1.1	1.1	100.0
	Total	27395	99.5	100.0	
Missing	97.00	62	.2		
	98.00	70	.3		
	99.00	7	.0		
	Total	139	.5		
Total		27534	100.0		

### 3. Regression

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	5.252	.071		74.076	.000	5.113	5.391
	GRP_10C_RECODED Number of groups - 12 months (recoded)	.834	.030	.165	27.557	.000	.775	.894

a. Dependent Variable: SCF\_100C Number of close friends

The 95 per cent confidence interval for the slope coefficient shows that, in the population, for each additional group that people participated in during the past 12 months, the regression line capturing the relationship with the number of close friends is predicted to rise between 0.78 and 0.89. In other words, the

general pattern of the relationship between the variables suggests that participating in one additional group during the past year is associated with an increase in the number of close friends that is between 0.78 and 0.89.

The 95 per cent confidence interval for the constant coefficient shows that, in the population, the regression line capturing the relationship between the number of groups people participated in during the past 12 months and the number of close friends is predicted to cross the vertical axis between 5.11 and 5.39. In other words, the general pattern of the relationship between the variables suggests that participating in no groups in the past year is associated with having a number of close friends that is between 5.11 and 5.39.

5. Regression

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	SCP_110 Number of new people met - Past month <sup>b</sup>	.	Enter

a. Dependent Variable: SCF\_100C Number of close friends

b. All requested variables entered.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.165 <sup>a</sup>	.027	.027	8.744

a. Predictors: (Constant), SCP\_110 Number of new people met - Past month

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	56636.836	1	56636.836	740.795	.000 <sup>b</sup>
	Residual	2033615.768	26599	76.454		
	Total	2090252.603	26600			

a. Dependent Variable: SCF\_100C Number of close friends

b. Predictors: (Constant), SCP\_110 Number of new people met - Past month

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.901	.058		101.687	.000
	SCP_110 Number of new people met - Past month	.161	.006	.165	27.218	.000

a. Dependent Variable: SCF\_100C Number of close friends

- a. The constant coefficient shows that those who met no new people in the past month are predicted to have 5.90 close friends. The slope coefficient shows that each additional new person met in the past month is associated with having an additional 0.16 close friends; in other words, meeting

approximately six new people in the past month is associated with having one additional close friend.

- b. The t-statistic of 27.22 has a p-value that is less than 0.05, so there is likely a relationship in the population between the number of new people met in the past month and the number of close friends that people have.

## 7. Regression

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	5.901	.058		101.687	.000	5.787	6.015
	SCP_110 Number of new people met - Past month	.161	.006	.165	27.218	.000	.149	.173

a. Dependent Variable: SCF\_100C Number of close friends

The 95 per cent confidence interval for the slope coefficient shows that, in the population, for each additional new person met in the past month, the regression line capturing the relationship with the number of close friends is predicted to rise between 0.15 and 0.17. In other words, the general pattern of the relationship between the variables suggests that meeting an additional new person in the past month is associated with an increase in the number of close friends that is between 0.15 and 0.17.

The 95 per cent confidence interval for the constant coefficient shows that, in the population, the regression line capturing the relationship between the number of new people met in the past month and the number of close friends is predicted to cross the vertical axis between 5.79 and 6.02. In other words, the general pattern of the relationship between the variables suggests that meeting no new people in the past month is associated with having a number of close friends that is between 5.79 and 6.02.

## 9. Regression

Model	Variables Entered	Variables Removed	Method
1	SCP_110_CENTRED Number of new people met - Past month (centred) <sup>b</sup>	.	Enter

a. Dependent Variable: SCF\_100C Number of close friends

b. All requested variables entered.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.165 <sup>a</sup>	.027	.027	8.744

a. Predictors: (Constant), SCP\_110\_CENTRED Number of new people met - Past month (centred)

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	56636.836	1	56636.836	740.795	.000 <sup>b</sup>
	Residual	2033615.768	26599	76.454		
	Total	2090252.603	26600			

a. Dependent Variable: SCF\_100C Number of close friends

b. Predictors: (Constant), SCP\_110\_CENTRED Number of new people met - Past month (centred)

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.545	.054		122.041	.000
	SCP_110_CENTRED Number of new people met - Past month (centred)	.161	.006	.165	27.218	.000

a. Dependent Variable: SCF\_100C Number of close friends

- a. The “Model Summary” and the “ANOVA” tables are identical to the regression produced in question 5. As well, the independent variable rows of the “Coefficients” table are identical (the rows for SCP\_110 and SCP\_110\_CENTRED). The only part of the output that is different is the constant row of the “Coefficients” table.
- b. The constant coefficient shows that those who met four new people in the past month are predicted to have 6.55 close friends. Meeting one additional new person in the past month is associated with a 0.16 increase in people’s number of close friends; similarly, meeting one less new person in the past month is associated with a 0.16 decrease in people’s number of close friends.

## Chapter 12

### 1. Regression

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	SCP_110 Number of new people met - Past month, GRP_10C_RECODED Number of groups - 12 months (recoded) <sup>b</sup>	.	Enter

a. Dependent Variable: SCF\_100C Number of close friends

b. All requested variables entered.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.215 <sup>a</sup>	.046	.046	8.637

a. Predictors: (Constant), SCP\_110 Number of new people met - Past month, GRP\_10C\_RECODED Number of groups - 12 months (recoded)

ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	95880.165	2	47940.083	642.683	.000 <sup>b</sup>
	Residual	1975225.331	26480	74.594		
	Total	2071105.497	26482			

a. Dependent Variable: SCF\_100C Number of close friends

b. Predictors: (Constant), SCP\_110 Number of new people met - Past month, GRP\_10C\_RECoded Number of groups - 12 months (recoded)

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.890	.072		68.093	.000
	GRP_10C_RECoded Number of groups - 12 months (recoded)	.718	.030	.144	23.593	.000
	SCP_110 Number of new people met - Past month	.135	.006	.138	22.612	.000

a. Dependent Variable: SCF\_100C Number of close friends

- a. The constant coefficient shows that those who participated in no groups in the past 12 months, and who met no new people in the past month, are predicted to have 4.89 close friends.
- b. The unstandardized slope coefficient of the “Number of groups” variable shows that each additional group that people participated in during the past 12 months is associated with a 0.72 increase in their number of close friends, controlling for the number of new people they met in the past month.

The unstandardized slope coefficient of the “Number of new people met” variable shows that each additional new person met in the past month is associated with a 0.14 increase in the number of close friends that people have, controlling for the number of groups they participated in during the past 12 months.

- c. The standardized slope coefficients show that the number of groups people participated in during the past 12 months has a stronger relationship with the dependent variable (“Number of close friends”) than the number of new people they met in the past month.

## 3. Frequencies

## Statistics

N		SEX Sex of respondent	WOMEN Women
		Valid	27534
	Missing	0	0

**Frequency Table**

**SEX Sex of respondent**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	13599	49.4	49.4	49.4
	2 Female	13935	50.6	50.6	100.0
Total		27534	100.0	100.0	

**WOMEN Women**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	13599	49.4	49.4	49.4
	1.00	13935	50.6	50.6	100.0
Total		27534	100.0	100.0	

**5. Regression**

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	WOMEN Women, GRP_10C_RECODED Number of groups - 12 months (recoded), SCP_110 Number of new people met - Past month <sup>b</sup>	.	Enter

a. Dependent Variable: SCF\_100C Number of close friends

b. All requested variables entered.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.218 <sup>a</sup>	.048	.047	8.631

a. Predictors: (Constant), WOMEN Women, GRP\_10C\_RECODED Number of groups - 12 months (recoded), SCP\_110 Number of new people met - Past month

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	98470.053	3	32823.351	440.590	.000 <sup>b</sup>
	Residual	1972635.444	26479	74.499		
Total		2071105.497	26482			

a. Dependent Variable: SCF\_100C Number of close friends

b. Predictors: (Constant), WOMEN Women, GRP\_10C\_RECODED Number of groups - 12 months (recoded), SCP\_110 Number of new people met - Past month

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.206	.090		58.125	.000
	GRP_10C_RECODED Number of groups - 12 months (recoded)	.723	.030	.145	23.776	.000
	SCP_110 Number of new people met - Past month	.133	.006	.136	22.284	.000
	WOMEN Women	-.627	.106	-.035	-5.896	.000

a. Dependent Variable: SCF\_100C Number of close friends



- a. The constant coefficient shows that men who did not participate in any groups in the past 12 months, and who did not meet any new people in the past month, are predicted to have 5.21 close friends.
- b. The unstandardized slope coefficient of the “Number of groups” variable shows that each additional group that people participated in during the past 12 months is associated with a 0.72 increase in their number of close friends, controlling for sex/gender and the number of new people they met in the past month.

The unstandardized slope coefficient of the “Number of new people met” variable shows that each additional new person met in the past month is associated with a 0.13 increase in the number of close friends that people have, controlling for sex/gender and the number of groups they participated in during the past 12 months.

The unstandardized slope coefficient of the “Women” dummy variable shows that women are predicted to have 0.63 fewer close friends than men, controlling for the number of new people they met in the past month and the number of groups they participated in during the past 12 months.

- c. The standardized slope coefficients show that the number of groups people participated in during the past 12 months has a stronger relationship with the dependent variable (“Number of close friends”) than sex/gender or the number of new people met in the past month.

## 7. Regression

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	SINGLE Single, PREVIOUS_RELATIONSHIP Previous relationship <sup>b</sup>	.	Enter

a. Dependent Variable: SCF\_100C Number of close friends

b. All requested variables entered.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.053 <sup>a</sup>	.003	.003	8.923

a. Predictors: (Constant), SINGLE Single, PREVIOUS\_RELATIONSHIP Previous relationship

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6134.186	2	3067.093	38.519	.000 <sup>b</sup>
	Residual	2161164.217	27142	79.625		
	Total	2167298.403	27144			

a. Dependent Variable: SCF\_100C Number of close friends

b. Predictors: (Constant), SINGLE Single, PREVIOUS\_RELATIONSHIP Previous relationship

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.397	.069		92.345	.000
	PREVIOUS_RELATIONSHIP Previous relationship	-.734	.175	-.026	-4.200	.000
	SINGLE Single	.820	.124	.041	6.590	.000

a. Dependent Variable: SCF\_100C Number of close friends

- a. The constant coefficient shows that people who are currently in a relationship (married or common-law) are predicted to have 6.40 close friends. (People who are currently in a relationship have a “0” value on the “Single” dummy variable and the “Previous relationship” dummy variable.)
- b. The unstandardized slope coefficient of the “Previous relationship” dummy variable shows that people who were previously in a long-term relationship (widowed, separated, or divorced) are predicted to have 0.73 fewer close friends than people who are currently in a relationship (married or common-law).

The unstandardized slope coefficient of the “Single” dummy variable shows that people who are single (never married) are predicted to have 0.82 more close friends than people who are currently in a relationship (married or common-law).

**9. Means**

**Case Processing Summary**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
SCP_110 Number of new people met - Past month	26850	97.5%	684	2.5%	27534	100.0%
SCP_110_RESCALED Number of new people met - Past month (scaled to 10)	26850	97.5%	684	2.5%	27534	100.0%

**Report**

	SCP_110 Number of new people met - Past month	SCP_110_RESCALED Number of new people met - Past month (scaled to 10)
Mean	3.75	.3751
N	26850	26850
Std. Deviation	9.076	.90760

## Chapter 13

### 1. a. Frequencies

Statistics			
		AGEGR10 Age group of respondent (groups of 10)	AGE Age
N	Valid	27534	27534
	Missing	0	0

### Frequency Table

AGEGR10 Age group of respondent (groups of 10)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 15 to 24 years	4289	15.6	15.6	15.6
	2 25 to 34 years	4629	16.8	16.8	32.4
	3 35 to 44 years	4411	16.0	16.0	48.4
	4 45 to 54 years	4985	18.1	18.1	66.5
	5 55 to 64 years	4337	15.8	15.8	82.3
	6 65 to 74 years	2817	10.2	10.2	92.5
	7 75 years and over	2066	7.5	7.5	100.0
	Total	27534	100.0	100.0	

AGE Age					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	19.50	4289	15.6	15.6	15.6
	29.50	4629	16.8	16.8	32.4
	39.50	4411	16.0	16.0	48.4
	49.50	4985	18.1	18.1	66.5
	59.50	4337	15.8	15.8	82.3
	69.50	2817	10.2	10.2	92.5
	79.50	2066	7.5	7.5	100.0
	Total	27534	100.0	100.0	

### b. Means

Case Processing Summary						
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
AGE Age	27534	100.0%	0	0.0%	27534	100.0%
AGE_CENTRED Age (centred)	27534	100.0%	0	0.0%	27534	100.0%

Report		
	AGE Age	AGE_CENTRED Age (centred)
Mean	45.7352	.7352
N	27534	27534
Std. Deviation	18.26165	18.26165

3. a-d. <No output>

e. **Frequencies**

		Statistics				
		DH1GED Education - Highest degree (4 categories)	LTHS Less than high school	HS High school only	DIPLOMA Post-secondary diploma	UNI University degree
N	Valid	27342	27342	27342	27342	27342
	Missing	192	192	192	192	192

**Frequency Table**

**DH1GED Education - Highest degree (4 categories)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Less than High School	4063	14.8	14.9	14.9
	2 Graduated from High School	7322	26.6	26.8	41.6
	3 Post-secondary diploma	8697	31.6	31.8	73.4
	4 University degree	7260	26.4	26.6	100.0
	Total	27342	99.3	100.0	
Missing	9 Not stated	192	.7		
Total		27534	100.0		

**LTHS Less than high school**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	23279	84.5	85.1	85.1
	1.00	4063	14.8	14.9	100.0
	Total	27342	99.3	100.0	
Missing	System	192	.7		
Total		27534	100.0		

**HS High school only**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	20020	72.7	73.2	73.2
	1.00	7322	26.6	26.8	100.0
	Total	27342	99.3	100.0	
Missing	System	192	.7		
Total		27534	100.0		

**DIPLOMA Post-secondary diploma**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	18645	67.7	68.2	68.2
	1.00	8697	31.6	31.8	100.0
	Total	27342	99.3	100.0	
Missing	System	192	.7		
Total		27534	100.0		

## UNI University degree

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	20083	72.9	73.4	73.4
	1.00	7260	26.4	26.6	100.0
	Total	27342	99.3	100.0	
Missing	System	192	.7		
Total		27534	100.0		

## 5. Regression

Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	WOMEN Women, IS_VISMIN Visible minority, AGE_CENTRED Age (centred) <sup>b</sup>	.	Enter
2	LTHS Less than high school, PREVIOUS_RELATIONSHIP Previous relationship, DIPLOMA Post-secondary diploma, SINGLE Single, UNI University degree <sup>b</sup>	.	Enter

a. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

b. All requested variables entered.

Model Summary<sup>c</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.260 <sup>a</sup>	.068	.068	13.7069
2	.330 <sup>b</sup>	.109	.109	13.4007

a. Predictors: (Constant), WOMEN Women, IS\_VISMIN Visible minority, AGE\_CENTRED Age (centred)

b. Predictors: (Constant), WOMEN Women, IS\_VISMIN Visible minority, AGE\_CENTRED Age (centred), LTHS Less than high school, PREVIOUS\_RELATIONSHIP Previous relationship, DIPLOMA Post-secondary diploma, SINGLE Single, UNI University degree

c. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	259402.499	3	86467.500	460.230	.000 <sup>b</sup>
	Residual	3567500.695	18988	187.879		
	Total	3826903.194	18991			
2	Regression	417892.335	8	52236.542	290.882	.000 <sup>c</sup>
	Residual	3409010.859	18983	179.580		
	Total	3826903.194	18991			

a. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

b. Predictors: (Constant), WOMEN Women, IS\_VISMIN Visible minority, AGE\_CENTRED Age (centred)

c. Predictors: (Constant), WOMEN Women, IS\_VISMIN Visible minority, AGE\_CENTRED Age (centred), LTHS Less than high school, PREVIOUS\_RELATIONSHIP Previous relationship, DIPLOMA Post-secondary diploma, SINGLE Single, UNI University degree

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.	Collinearity Statistics	
		B	Std. Error	Beta	t		Tolerance	VIF
1	(Constant)	41.826	.146		285.843	.000		
	AGE_CENTRED Age (centred)	.073	.007	.075	10.743	.000	.994	1.006
	IS_VISMIN Visible minority	-1.100	.261	-.030	-4.210	.000	.995	1.005
	WOMEN Women	-6.980	.199	-.245	-35.015	.000	1.000	1.000
2	(Constant)	42.348	.234		180.936	.000		
	AGE_CENTRED Age (centred)	-.034	.008	-.035	-4.206	.000	.665	1.504
	IS_VISMIN Visible minority	-1.616	.259	-.043	-6.241	.000	.968	1.034
	WOMEN Women	-7.370	.196	-.259	-37.584	.000	.987	1.013
	LTHS Less than high school	-4.242	.373	-.086	-11.361	.000	.814	1.228
	DIPLOMA Post-secondary diploma	1.486	.255	.050	5.826	.000	.645	1.550
	UNI University degree	2.292	.265	.074	8.637	.000	.638	1.568
	PREVIOUS_RELATIONSHIP Previous relationship	.674	.387	.012	1.741	.082	.937	1.067
	SINGLE Single	-5.215	.261	-.169	-19.977	.000	.657	1.521

a. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

**Excluded Variables<sup>a</sup>**

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics		
						Tolerance	VIF	Minimum Tolerance
1	LTHS Less than high school	-.127 <sup>b</sup>	-18.164	.000	-.131	.992	1.008	.992
	DIPLOMA Post-secondary diploma	.045 <sup>b</sup>	6.436	.000	.047	.989	1.011	.989
	UNI University degree	.081 <sup>b</sup>	11.406	.000	.082	.969	1.032	.968
	PREVIOUS_RELATIONSHIP Previous relationship	.019 <sup>b</sup>	2.648	.008	.019	.945	1.058	.945
	SINGLE Single	-.194 <sup>b</sup>	-23.168	.000	-.166	.682	1.466	.679

a. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

b. Predictors in the Model: (Constant), WOMEN Women, IS\_VISMIN Visible minority, AGE\_CENTRED Age (centred)

**Collinearity Diagnostics<sup>a</sup>**

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions								
				(Constant)	AGE_CENTRED Age (centred)	IS_VISMIN Visible minority	WOMEN Women	LTHS Less than high school	DIPLOMA Post-secondary diploma	UNI University degree	PREVIOUS_RELATIONSHIP Previous relationship	SINGLE Single
1	1	2.094	1.000	.08	.05	.08	.08					
	2	.859	1.561	.02	.93	.02	.05					
	3	.753	1.668	.02	.00	.83	.14					
	4	.294	2.667	.88	.02	.08	.73					
2	1	3.243	1.000	.01	.01	.02	.03	.01	.01	.01	.01	.02
	2	1.379	1.533	.00	.16	.00	.01	.02	.01	.01	.18	.06
	3	1.076	1.736	.00	.00	.12	.00	.10	.08	.17	.05	.00
	4	.987	1.812	.00	.01	.00	.00	.45	.12	.01	.06	.00

5	.740	2.093	.00	.08	.40	.01	.10	.04	.03	.27	.03
6	.688	2.172	.01	.06	.38	.09	.02	.00	.06	.41	.01
7	.433	2.737	.02	.04	.05	.75	.00	.09	.12	.00	.06
8	.340	3.088	.00	.62	.01	.03	.07	.06	.06	.00	.68
9	.113	5.359	.95	.00	.01	.09	.23	.58	.53	.02	.14

a. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

**Residuals Statistics<sup>a</sup>**

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	23.406	45.847	38.035	4.6909	18992
Residual	-43.5795	49.9774	.0000	13.3979	18992
Std. Predicted Value	-3.118	1.665	.000	1.000	18992
Std. Residual	-3.252	3.729	.000	1.000	18992

a. Dependent Variable: WKWEHRC Number of paid hours worked per week - All jobs

- a. The constant coefficient of the second model shows that 45-year-old men who are not visible minorities, who have only a high school education, and who are currently in a relationship are predicted to spend 42.35 hours doing paid work per week.
- b. The unstandardized slope coefficient of the “Less than high school” dummy variable shows that people with less than a high school education are predicted to spend 4.24 fewer hours doing paid work per week than people with only a high school education, controlling for age, visible minority status, sex/gender, and relationship status.

The unstandardized slope coefficient of the “Post-secondary diploma” dummy variable shows that people with a post-secondary diploma are predicted to spend 1.49 more hours doing paid work per week than people with only a high school education, controlling for age, visible minority status, sex/gender, and relationship status.

The unstandardized slope coefficient of the “University degree” dummy variable shows that people with a university degree are predicted to spend 2.29 more hours doing paid work per week than people with only a high school education, controlling for age, visible minority status, sex/gender, and relationship status.

- c. The unstandardized slope coefficient of the “Previous relationship” dummy variable shows that people who were previously in a long-term relationship (widowed, separated, or divorced) are predicted to spend 0.67 more hours doing paid work per week than people who are currently in a relationship (married or common-law), controlling for age, visible minority status, sex/gender, and highest level of education. But, since the p-value associated with this slope coefficient is greater than 0.05, we cannot be confident that, in the population, there is any difference between people who are currently in a relationship and people who were previously in a long-term relationship in terms of the number of hours they spend doing paid work each week.

The unstandardized slope coefficient of the “Single” dummy variable shows that people who are single (never married) are predicted to spend 5.22 fewer hours doing paid work per week than people who are currently in a relationship (married or common-law), controlling for age, visible minority status, sex/gender, and highest level of education.

- d. The standardized slope coefficients show that among the independent variables used in regression, sex/gender has the strongest relationship with the number of hours spent doing paid work each week.
- e. The  $R^2$  shows that, overall, 10.9 per cent of the variation in the number of hours spent doing paid work each week can be explained by age, visible minority status, sex/gender, highest level of education, and marital status. Since the  $R^2$  and the adjusted  $R^2$  are the same, it suggests that all of the independent variables in this model are good predictors of the dependent variable.
- f. The tolerances and variance inflation factors indicate that there are no collinearity problems among the independent variables in this regression: all of the tolerances are above 0.1 and all of the variance inflation factors are below 10. The three variables with the lowest tolerances (and thus highest variance inflation factors) are the “University degree” dummy variable (0.64), the “Post-secondary diploma” dummy variable (0.65), and the “Single” dummy variable (0.66).

9. a. **GGraph**

Simple Scatter of Unstandardized Residual by Number of paid hours worked per week - All jobs

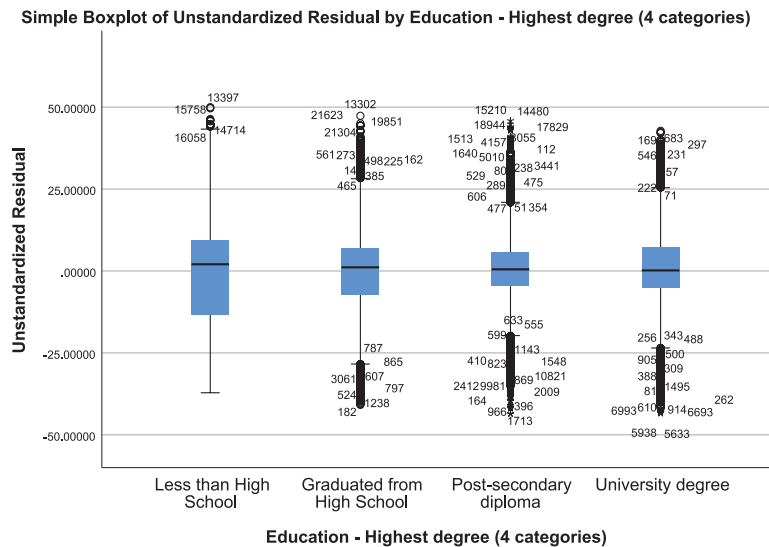




These results show that there is a relationship between the regression residuals and the dependent variable. Ideally, there should be no relationship between the regression residuals and the dependent variable. (The only remaining variation should be random.)

The model systematically over-predicts the number of paid hours worked for people working fewer than 40 hours per week, and systematically under-predicts the number of paid hours worked for people working more than 40 hours per week. This suggests that the model still needs improvement. Independent variables that help to explain why people work very low and very high numbers of hours each week should be added.

b. **GGraph**



These results show that there is no relationship between the regression residuals and people’s highest level of education. The median of each box plot is near 0, the ideal value of a residual. The box plots also show that there is slightly more variation in the residuals—a wider range and interquartile range—for people who have less than a high school education than for people with higher levels of education.

## Chapter 14

### 1. Crosstabs

**Case Processing Summary**

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
WOMEN Women *	27182	98.7%	352.000	1.3%	27534.000	100.0%
IS_VISMIN Visible minority						

**WOMEN Women \* IS\_VISMIN Visible minority  
Crosstabulation**

Count		IS_VISMIN Visible minority		Total
		.00	1.00	
WOMEN Women	.00	11126	2292	13418
	1.00	11546	2218	13764
Total		22672	4510	27182

**Frequencies**

**Statistics**

VISMIN\_WOMEN Visible minority women

N	Valid	27182
	Missing	352

**VISMIN\_WOMEN Visible minority women**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	24964	90.7	91.8	91.8
	1.00	2218	8.1	8.2	100.0
	Total	27182	98.7	100.0	
Missing	System	352	1.3		
Total		27534	100.0		

3. a. **Frequencies**

**Statistics**

DH1GED Education - Highest degree (4 categories)      POSTSECONDARY Has a postsecondary education

N	Valid	27342	27342
	Missing	192	192

**Frequency Table**

**DH1GED Education - Highest degree (4 categories)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Less than High School	4063	14.8	14.9	14.9
	2 Graduated from High School	7322	26.6	26.8	41.6
	3 Post-secondary diploma	8697	31.6	31.8	73.4
	4 University degree	7260	26.4	26.6	100.0
	Total	27342	99.3	100.0	
Missing	9 Not stated	192	.7		
Total		27534	100.0		

**POSTSECONDARY Has a postsecondary education**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	11385	41.4	41.6	41.6
	1.00	15957	58.0	58.4	100.0
	Total	27342	99.3	100.0	
Missing	System	192	.7		
Total		27534	100.0		

## b. Means

## Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
AGE_POSTSEC Age x Postsecondary education	27342	99.3%	192	0.7%	27534	100.0%

## Report

AGE\_POSTSEC Age x Postsecondary education

Mean	N	Std. Deviation
.5254	27342	12.11230

## 5. Means

## Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
AGE_SQUARED Age (squared)	27534	100.0%	0	0.0%	27534	100.0%

## Report

AGE\_SQUARED Age (squared)

Mean	N	Std. Deviation
334.0164	27534	338.14768

## 7. a. &lt;No output&gt;

## b. Means

## Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
SCP_110_LOG2 Number of new people met - past month (log base 2)	12724	46.2%	14810	53.8%	27534	100.0%

## Report

SCP\_110\_LOG2 Number of new people met - past month (log base 2)

Mean	N	Std. Deviation
2.1822	12724	1.41718

## Chapter 15

### 1. Frequencies

		Statistics	
		DISCRIM Victim of discrimination - 5 years	DISCRIM_ RECODED Experienced discrimination
N	Valid	26574	26574
	Missing	960	960

#### Frequency Table

DISCRIM Victim of discrimination - 5 years					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	8211	29.8	30.9	30.9
	2 No	18363	66.7	69.1	100.0
	Total	26574	96.5	100.0	
Missing	7 Don't know	95	.3		
	8 Refusal	554	2.0		
	9 Not stated	311	1.1		
	Total	960	3.5		
Total		27534	100.0		

DISCRIM_RECODED Experienced discrimination					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	18363	66.7	69.1	69.1
	1.00	8211	29.8	30.9	100.0
	Total	26574	96.5	100.0	
Missing	System	960	3.5		
Total		27534	100.0		

### 3. Logistic Regression

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 <sup>a</sup>	Visible minority	.694	.034	419.661	1	.000	2.003	1.874	2.140
	Constant	-.933	.015	3886.190	1	.000	.393		

a. Variable(s) entered on step 1: Visible minority.

The 95 per cent confidence interval for the odds ratio of the “Visible minority” dummy variable shows that, in the population, a regression capturing the relationship between visible minority status and experiencing discrimination is likely to show that people who are visible minorities have between 87 per cent and 114 per cent higher odds of experiencing discrimination. In other words, the general pattern of the relationship between the variables suggests that being a visible minority is associated with odds of experiencing discrimination that are 87 per cent to 114 per cent higher than for people who are not visible minorities.

## 5. Logistic Regression

### Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	26411	95.9
	Missing Cases	1123	4.1
	Total	27534	100.0
Unselected Cases		0	.0
Total		27534	100.0

a. If weight is in effect, see classification table for the total number of cases.

### Dependent Variable Encoding

Original Value	Internal Value
.00	0
1.00	1

## Block 0: Beginning Block

### Classification Table<sup>a,b</sup>

Observed		DISCRIM_RECODED Experienced discrimination	Predicted		Percentage Correct
			.00	1.00	
Step 0	DISCRIM_RECODED Experienced discrimination	.00	18257	0	100.0
		1.00	8146	0	.0
Overall Percentage					69.1

a. Constant is included in the model.

b. The cut value is .500

### Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-.807	.013	3668.566	1	.000	.446

**Variables not in the Equation**

			Score	df	Sig.
Step 0	Variables	Age	889.972	1	.000
		Visible minority	429.692	1	.000
		Women	148.370	1	.000
	Overall Statistics	1331.322	3	.000	

**Block 1: Method = Enter**

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	1349.904	3	.000
	Block	1349.904	3	.000
	Model	1349.904	3	.000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	31280.684 <sup>a</sup>	.050	.070

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

**Classification Table<sup>a</sup>**

Observed		Predicted		Percentage Correct
		DISCRIM_RECODED Experienced discrimination .00	DISCRIM_RECODED Experienced discrimination 1.00	
Step 1	DISCRIM_RECODED Experienced discrimination .00	17803	454	97.5
	DISCRIM_RECODED Experienced discrimination 1.00	7658	488	6.0
Overall Percentage				69.3

a. The cut value is .500

**Variables in the Equation**

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	Age	-.021	.001	750.057	1	.000	.979
	Visible minority	.564	.035	262.770	1	.000	1.757
	Women	.373	.027	184.358	1	.000	1.452
	Constant	-.157	.040	15.602	1	.000	.855

a. Variable(s) entered on step 1: Age, Visible minority, Women.

The odds ratio of the “Age” variable shows that each one-year increase in age is associated with 2 per cent lower odds of experiencing discrimination, controlling for visible minority status and sex/gender.

The odds ratio of the “Visible minority” dummy variable shows that people who are visible minorities are predicted to have 76 per cent higher odds

of experiencing discrimination than people who are not visible minorities, controlling for age and sex/gender.

The odds ratio of the “Women” dummy variable shows that women are predicted to have 45 per cent higher odds of experiencing discrimination than men, controlling for age and visible minority status.

7. a. **Means**

**Case Processing Summary**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
DISCRIM_RECoded Experienced discrimination	26404	100.0%	0	0.0%	26404	100.0%

**Report**

DISCRIM_RECoded Experienced discrimination		
Mean	N	Std. Deviation
.3085	26404	.46189

b. **Means**

**Case Processing Summary**

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
AGE Age	26404	100.0%	0	0.0%	26404	100.0%
IS_VISMIN Visible minority	26404	100.0%	0	0.0%	26404	100.0%
WOMEN Women	26404	100.0%	0	0.0%	26404	100.0%

**Report**

	AGE Age	IS_VISMIN Visible minority	WOMEN Women
Mean	45.5891	.1661	.5062
N	26404	26404	26404
Std. Deviation	18.30166	.37220	.49997

c. The standardized slope coefficient of the “Age” variable is  $-0.832$ .

The standardized slope coefficient of the “Visible minority” variable is  $0.454$ .

The standardized slope coefficient of the “Women” variable is  $0.404$ .

d. The standardized slope coefficients show that age has a stronger relationship with the dependent variable (“Experienced discrimination”) than visible minority status or sex/gender.

9. Logistic Regression

**Case Processing Summary**

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	25512	92.7
	Missing Cases	2022	7.3
	Total	27534	100.0
Unselected Cases		0	.0
Total		27534	100.0

a. If weight is in effect, see classification table for the total number of cases.

**Dependent Variable  
Encoding**

Original Value	Internal Value
.00	0
1.00	1

**Block 0: Beginning Block**

**Classification Table<sup>a,b</sup>**

Observed		Predicted		Percentage Correct
		DISCRIM_RECODED Experienced discrimination .00	1.00	
Step 0	DISCRIM_RECODED Experienced discrimination .00	17757	0	100.0
	1.00	7872	0	.0
Overall Percentage				69.3

a. Constant is included in the model.

b. The cut value is .500

**Variables in the Equation**

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-.813	.014	3608.942	1	.000	.443

**Variables not in the Equation**

		Score	df	Sig.	
Step 0	Variables	Age	888.460	1	.000
	Visible minority	423.998	1	.000	
	Women	151.909	1	.000	
Overall Statistics		1329.732	3	.000	

**Block 1: Method = Enter**

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	1349.350	3	.000
	Block	1349.350	3	.000
	Model	1349.350	3	.000



**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	30268.105 <sup>a</sup>	.051	.072

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

**Classification Table<sup>a</sup>**

Observed		Predicted		Percentage Correct	
		DISCRIM_RECODED Experienced discrimination .00	1.00		
Step 1	DISCRIM_RECODED Experienced discrimination	.00	17330	427	97.6
		1.00	7400	472	6.0
Overall Percentage					69.5

a. The cut value is .500

**Variables in the Equation**

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	Age	-.022	.001	751.420	1	.000	.978
	Visible minority	.573	.036	259.665	1	.000	1.774
	Women	.383	.028	187.202	1	.000	1.466
	Constant	-.152	.040	14.239	1	.000	.859

a. Variable(s) entered on step 1: Age, Visible minority, Women.

**Block 2: Method = Enter****Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	36.627	2	.000
	Block	36.627	2	.000
	Model	1385.977	5	.000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	30231.478 <sup>a</sup>	.053	.074

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

**Classification Table<sup>a</sup>**

Observed		Predicted		Percentage Correct	
		DISCRIM_RECODED Experienced discrimination .00	1.00		
Step 1	DISCRIM_RECODED Experienced discrimination	.00	17063	695	96.1
		1.00	7219	653	8.3
Overall Percentage					69.1

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	Age	-.021	.001	653.789	1	.000	.979
	Visible minority	.528	.039	182.421	1	.000	1.696
	Women	.400	.028	201.420	1	.000	1.491
	Christian	-.188	.034	30.029	1	.000	.829
	Other religion	.008	.060	.020	1	.887	1.008
	Constant	-.067	.044	2.379	1	.123	.935

a. Variable(s) entered on step 1: Christian, Other religion.

- a. The odds ratio of the “Christian” dummy variable shows that people who are Christians are predicted to have 17 per cent lower odds of experiencing discrimination than people who have no religious affiliation, controlling for age, visible minority status, and sex/gender.

The odds ratio of the “Other religion” dummy variable shows that people who are affiliated with another (non-Christian) religion have 1 per cent higher odds of experiencing discrimination than people who have no religious affiliation, controlling for age, visible minority status, and sex/gender. But, since the p-value associated with this odds ratio is greater than 0.05, we cannot be confident that, in the population, there is any difference between people with no religious affiliation and people affiliated with another (non-Christian) religion in terms of their odds of experiencing discrimination.

- b. The odds ratio of the “Age” variable does not change substantially once religious affiliation is controlled for (0.98 in both models).

The odds ratio of the “Visible minority” dummy variable becomes smaller once religious affiliation is controlled for (changing from 1.77 to 1.70). When religious affiliation is not taken into account, people who are visible minorities are predicted to have 77 per cent higher odds of experiencing discrimination than people who are not visible minorities (controlling for age and sex/gender). Once religious affiliation is accounted for, people who are visible minorities are only predicted to have 70 per cent higher odds of experiencing discrimination than people who are not visible minorities (also controlling for age and sex/gender).

In contrast, the odds ratio of the “Women” dummy variable becomes slightly larger once religious affiliation is controlled for (changing from 1.47 to 1.49). When religious affiliation is not taken into account, women are predicted to have 47 per cent higher odds of experiencing discrimination than men (controlling for age and visible minority status). Once religious affiliation is accounted for, women are predicted to have

- 49 per cent higher odds of experiencing discrimination than men (also controlling for age and visible minority status).
- c. The Nagelkerke  $R^2$  of the second model is 0.074, whereas the Nagelkerke  $R^2$  of the first model is 0.072. Since the difference between them is small, it suggests that accounting for religious affiliation does not substantially improve the fit of the logistic regression model.