

Chapter 14 Solutions

Making Sense of Social Data

Mark Paradis, Ph.D.

September 2, 2026

These are the worked solutions to the *end-of-chapter* exercises for Chapter 14 of *Making Sense of Social Data*. Solutions to the exercises inside each section are in the book itself, printed immediately after the exercises they answer.

Do the exercise before you read the solution. A solution read first is just more text to read; a solution read second is a check on your reasoning, and only the second one teaches you anything.

The complete set for every chapter is at <https://www.markparadispolitics.com/books/making-sense-of-social-data/online/solutions/>.

Chapter 14 The Chi-Square Test

Solution to Exercise 14.E1. (a) The marginal totals are: rows 140, 140, 88, 112; columns 214, 192, 74; $N = 480$.

Applying $E = (\text{row} \times \text{column})/N$:

<i>Expected</i>	Support	Oppose	No opinion
Party A	62.42	56.00	21.58
Party B	62.42	56.00	21.58
Party C	39.23	35.20	13.57
Independent	49.93	44.80	17.27

The smallest expected frequency is 13.57, comfortably above 5. All twelve cells pass, so the assumption holds.

(b) Summing $(O - E)^2/E$ across the twelve cells:

$$\chi^2 = 19.70 \quad df = (4 - 1)(3 - 1) = 6$$

(c) The critical value at $\alpha = .05$ with $df = 6$ is 12.59. Since $19.70 > 12.59$, reject H_0 . The p -value is approximately .003.

Party affiliation and support for electoral reform are associated.

(d) For a 4×3 table, $\min(r - 1, c - 1) = \min(3, 2) = 2$:

$$V = \sqrt{\frac{19.70}{480 \times 2}} = \sqrt{0.0205} = .143$$

With $\min(r - 1, c - 1) = 2$, the benchmarks are .07 small, .21 medium, .35 large. A value of .14 falls between small and medium — a modest association.

(e) *A complete report*:

“A chi-square test of independence found a significant association between party affiliation and support for electoral reform, $\chi^2(6, N = 480) = 19.70$, $p = .003$, Cramér’s $V = .14$. Party A supporters favoured the reform (55.7 per cent support, 30.0 per cent oppose) while Party B supporters opposed it (32.9 per cent support, 52.9 per cent oppose). Party C voters and independents fell between, and the proportion expressing no opinion was similar across all four groups.”

(f) The largest contributions come from Parties A and B:

<i>Contributions</i>	Support	Oppose	No opinion
Party A	3.89	3.50	0.12
Party B	4.32	5.79	0.12
Party C	0.04	0.29	1.45
Independent	0.09	0.01	0.09

Parties A and B together contribute 17.73 of the 19.70 — nearly 90 per cent.

Party C and independents contribute almost nothing; their opinion distributions are close to what independence predicts.

Substantively: the association is driven almost entirely by the two major parties taking opposite positions. Party A supporters back the reform at above the overall rate, Party B supporters oppose it, and the smaller party and independents are unremarkable.

This is a partisan split on a specific policy rather than a general relationship between party identity and reform attitudes — which is a more precise finding than the significant chi-square alone conveys, and it is available only from the decomposition.

Solution to Exercise 14.E2. (a) For a 3×4 table, $\min(r - 1, c - 1) = \min(2, 3) = 2$:

$$V = \sqrt{\frac{18.4}{600 \times 2}} = \sqrt{\frac{18.4}{1,200}} = \sqrt{0.01533} = .124$$

(b) With $\min(r - 1, c - 1) = 2$, the benchmarks are .07 small, .21 medium, .35 large.

A value of .12 is small — above the “small” threshold and well below “medium”. Employment status and self-rated health are associated, weakly.

Note the importance of using the right benchmark row. Judged against the $\min = 1$ benchmarks (.10, .30, .50), .12 would look similarly small; but for a larger table the thresholds fall further, and a mechanical comparison against the wrong row can misclassify an effect (§14.4).

(c) The conclusion is not supported, on three grounds.

The test establishes association, not causation. These are the three structures of §3.5, and the chi-square test distinguishes none of them.

Reverse causation is highly plausible here — more so than in most examples. Poor health causes unemployment: people who are ill lose jobs, struggle to find work, and leave the labour force. The “out of labour force” category in particular contains many people whose health prevents employment.

Confounders are abundant. Age, education, income, region, and prior chronic illness all plausibly affect both employment status and health.

Also worth noting: the association is weak ($V = .12$), so even if causal it would explain little of the variation in health.

What would be needed. Longitudinal data establishing temporal order — health measured before job loss, then again after. Ideally a natural experiment: plant closures and mass layoffs are the standard design here, because they make job loss independent of the individual’s prior health. And measurement of the candidate confounders, so their contribution can be examined.

The plant-closure design, or something like it, is what the literature actually uses — worth naming for that reason.

Solution to Exercise 14.E3. (a) The marginal totals are: rows 74, 68, 23; columns 80, 72, 13; $N = 165$.

<i>Expected</i>	Daily	Weekly	Never
Urban	35.88	32.29	5.83
Suburban	32.97	29.67	5.36
Rural	11.15	10.04	1.81

1 cell violates the strict rule: rural–never, with an expected frequency of 1.81.

Note that the observed count in that cell is 2, but the assumption concerns *expected* rather than observed frequencies — a distinction worth checking, since students often test the wrong table.

(b) *Yes, narrowly.* The lenient rule permits up to 20 per cent of cells below 5, with none below 1.

Here $1/9 = 11.1$ per cent of cells fall below 5, which is under the 20 per cent allowance. The minimum expected frequency is 1.81, which exceeds 1.

So the table passes the lenient rule and fails the strict one. A chi-square test would be defensible, and the result should be reported with the caveat noted.

(c) Two options, with a recommendation.

Collapse categories. The “never” column is the problem, containing only 13 of 165 respondents. Merging it with “weekly” to give a “less than daily” category would produce a 3×2 table with all expected frequencies above 10.

This is defensible on substantive grounds — the meaningful contrast may well be daily versus not — and must be reported as a decision taken.

Use Fisher’s exact test. Extensions beyond 2×2 exist and are available in standard software. This requires no change to the data.

Recommendation: collapse the columns, provided the daily-versus-not contrast answers the research question. The category is sparse because few people never use transit at all, and a three-way distinction that one level barely populates is not serving the analysis.

If the “never” group is substantively the point of the study — if the question concerns transit non-users specifically — then collapsing destroys the research question and Fisher’s exact test is the right answer instead.

Note also that with $N = 165$ this study has limited power, so a non-significant result would establish little whichever route is taken.

Solution to Exercise 14.E4. Your article will differ from any other reader’s. What follows is what a complete answer covers, in the order the exercise asks for it.

Is the inventory accurate and complete? Most published chi-square reports give the statistic, degrees of freedom, and p -value. Many omit N from the parentheses,

and a substantial proportion omit the effect size entirely. The exercise is about what is *missing*; listing only what is present skips it.

Do the degrees of freedom check out? This is the most useful verification available to a reader, and it catches real errors. If the article describes a 3×4 table and reports $df = 11$, something is wrong — $(r - 1)(c - 1)$ gives 6, and $rc - 1$ gives 11, which is Pearson's original error (§14.2.4).

Was the effect size computed correctly? From χ^2 and N you can recover Cramér's V yourself, using the benchmark row that matches the table's dimensions.

Is the conclusion proportionate? Two failure modes to watch for: causal verbs where the design is cross-sectional, and a small V described as though it were substantial. Check also whether the article reports the percentages; without them the direction of the association is unavailable.

One pattern is worth noticing across several articles: how often a chi-square result is reported without an effect size, and whether the ones that omit it tend to have large samples. If they do, that is §13.6.1 operating in a literature rather than in a textbook.

Solution to Exercise 14.E5. Your answer will use different variables from anyone else's; the reasoning is the same whichever pair you pick.

(a) State the hypotheses before computing anything, and in the form appropriate to chi-square — about independence rather than about a parameter value. Writing $H_0 : \mu = \dots$ means the wrong template has been imported.

(b) Both counts and row percentages are required, with the percentaging direction stated (§10.2.2). The expected frequency check should report the minimum expected value rather than merely asserting that the assumption holds.

(c) A complete report includes the statistic with df and N , the exact p -value, Cramér's V with the correct benchmark row, and the percentages showing direction.

(d) The three outcomes are all instructive, and which one occurs depends on the variables chosen.

Persists: the control variable does not explain the association. Most attempts will land here.

Vanishes: the control variable was producing it. This is the most instructive outcome and takes some searching to find — in *socsurvey*, education is the most likely candidate to produce it, since it relates to a great many other variables.

Differs across strata: interaction. Worth reporting carefully, since a combined analysis would have concealed it entirely.

One thing to watch for in your own work: stratifying repeatedly until an interesting result appears, then reporting only that one. This exercise invites it. Noticing the temptation, and reporting every attempt, is the point of §14.7's ethics box — and it is easier to see in yourself here than in a published paper.

(e) *Real against simulated.* Statistics Canada's Labour Force Survey voting supplement publishes turnout by education, and the pattern runs the same way as socsurvey's: turnout rises with education, and the gap between the least and most educated is large.

Would a chi-square test settle whether the two agree? No, and the reason is worth being precise about. A test of independence asks whether two variables are associated *within one table*. It cannot compare a table against an external set of percentages — that is a goodness-of-fit question (§14.6), and even then the answer would be uninformative here.

Run as goodness-of-fit against the real figures, the test would almost certainly reject, for two reasons that have nothing to do with Canada. socsurvey is simulated, so its percentages were chosen rather than observed. And self-reported turnout in surveys runs above official turnout everywhere it has been measured, so even a real survey would fail this comparison.

What the comparison is good for is the direction and the rough size of the gradient. Those match, which tells you the simulation is a reasonable place to practise — and nothing at all about Canada.