

Chapter 12 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 12 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 12 Inference and Estimation

Solution to Exercise 12.E1. (a) The standard error:

$$SE = \frac{0.94}{\sqrt{400}} = \frac{0.94}{20} = 0.047$$

The margin of error:

$$ME = 1.97 \times 0.047 = 0.093$$

The interval:

$$3.62 \pm 0.093 \Rightarrow (3.53, 3.71)$$

In APA format:

$$M = 3.62, 95\% \text{ CI } [3.53, 3.71], n = 400$$

(b) Yes — and more than consistent.

The benchmark of 3.50 falls *below* the entire interval, whose lower bound is 3.53. So the data are not merely consistent with the population being at or above 3.50; they are inconsistent with its being *below* 3.50.

The distinction matters. A value inside the interval would mean the data cannot rule the benchmark out. A value below the interval means the data indicate the population exceeds it.

Note the narrowness of the margin: 3.50 misses the lower bound by 0.03, so this is a conclusion the sample supports rather than one it establishes decisively.

(c) Using the sample size formula:

$$n = \left(\frac{1.96 \times 0.94}{0.05} \right)^2 = \left(\frac{1.842}{0.05} \right)^2 = (36.85)^2 = 1,357.8$$

Rounding up, $n = 1,358$.

Worth noticing how expensive this precision is. Going from a margin of error of 0.093 to one of 0.05 — less than halving it — requires increasing the sample from 400 to 1,358, 3.4 times as many respondents.

Solution to Exercise 12.E2. (a) *Team A's* interval is widest, spanning 11.8 points against *Team B's* 5.8 and *Team C's* 3.0.

The reason is sample size. The standard error is $\sigma/\sqrt{n}$, so:

$$n = 25 \Rightarrow SE = \frac{15}{5} = 3.0$$

$$n = 100 \Rightarrow SE = \frac{15}{10} = 1.5$$

$$n = 400 \implies SE = \frac{15}{20} = 0.75$$

Each fourfold increase in n halves the standard error, and the standard errors above do halve exactly: 3.0, 1.5, 0.75.

The interval widths inherit that halving, but only up to rounding. The exact widths are 11.76, 5.88 and 2.94, which halve cleanly; subtract the bounds *as printed* and you get 11.8, 5.8 and 3.0 instead, because each bound is given to one decimal and a width carries two bounds' worth of rounding. The pattern is real; it is the display that blurs it, which is worth knowing before you conclude that a published table has an arithmetic error.

(b) Checking each interval against $\mu = 50$:

Team	Interval	Contains $\mu = 50$?
A	(42.2, 54.0)	Yes
B	(48.9, 54.7)	Yes
C	(47.9, 50.9)	Yes

All three contain the true value. The question invites you to check rather than assume, and the checking is the point.

Had one missed, that would not indicate an error. A 95 per cent procedure fails in about one sample in twenty by construction, and a team whose interval misses has done nothing wrong — they drew an unusual sample, which happens at a known rate. This is exactly what Figure 12.3 shows, where one interval in twenty misses.

What would be an error is concluding from a single missing interval that the method is broken, or that the team was careless.

Notice also that Team C's interval is narrowest and comes closest to excluding μ — its upper bound is 50.9. A larger sample gives a more precise estimate, which also means less room for error before the interval misses. Precision and the risk of a near-miss travel together.

(c) About 5.

A 95 per cent confidence interval is constructed so that 95 per cent of such intervals contain the parameter, so 5 per cent do not. Out of 100 teams:

$$100 \times 0.05 = 5$$

This is an expectation rather than a guarantee. The actual number would vary — three, or seven, or occasionally more — because the count is itself subject to sampling variation. What the 95 per cent guarantees is the long-run rate, not the result in any particular set of 100 studies.

Solution to Exercise 12.E3. The statement contains a defensible instinct and two errors.

What the margin of error means. It is the half-width of a confidence interval around *one* estimate. A margin of ± 3 points on a candidate polling at 50 per cent means the interval for that candidate's support runs from 47 to 53, and that intervals built this way capture the true figure in 95 per cent of samples.

Error 1: applying the margin of error to a difference. The journalist compares the gap between two candidates to a margin of error computed for each separately. That is not the right quantity. A difference between two estimates has its own standard error, larger than either one's, so the threshold for distinguishing a gap from zero is larger than the reported margin — roughly 1.5 times it in typical polling situations.

The journalist's rule is therefore conservative, and conservative for the wrong reason. It happens to demand more evidence than a naive comparison would, while resting on a calculation that does not apply.

Error 2: "statistically tied". An interval that includes zero does not establish that the difference *is* zero. It says the data cannot distinguish the difference from zero — a statement about the evidence, not about the race.

A candidate leading 51–47 is more likely ahead than behind. The poll simply lacks the precision to confirm it. Calling that a tie asserts something the data do not support, in the opposite direction from the error it was meant to avoid.

A defensible rewrite. "A leads by 4 points, which is within the poll's margin of error — so the lead appears in the data but this sample is not precise enough to confirm it."

Sampling error is also only one source of error in a poll. Non-response, coverage, and question wording (§5.4) are not captured by the margin of error at all, and the reported figure therefore understates total uncertainty.

Solution to Exercise 12.E4. (a)

$$n = \left(\frac{1.96 \times 18,000}{2,000} \right)^2 = \left(\frac{35,280}{2,000} \right)^2 = (17.64)^2 = 311.2$$

Rounding up, $n = 312$ households.

(b) With $n = 200$:

$$SE = \frac{18,000}{\sqrt{200}} = \frac{18,000}{14.14} = 1,273$$

$$ME = 1.96 \times 1,273 = 2,495$$

About \$2,500 — noticeably worse than the \$2,000 she wanted. The shortfall in sample size is 36 per cent and the loss of precision is about 25 per cent, which is the square root relationship in action.

(c) At 99 per cent confidence the critical value rises to 2.576:

$$n = \left(\frac{2.576 \times 18,000}{2,000} \right)^2 = (23.18)^2 = 537.5$$

Rounding up, $n = 538$ households.

The comparison across the three parts is the lesson. Holding the margin of error at \$2,000, raising confidence from 95 to 99 per cent raises the required sample from 312 to 538 — a 72 per cent increase in fieldwork for 4 percentage points of confidence.

With a budget of 200 she cannot achieve either target. Her realistic options are to accept a margin of error near \$2,500 and report it plainly, or to seek funding for the larger sample. What she should not do is report the \$2,000 precision she planned for.

Solution to Exercise 12.E5. *The connection.* A confidence interval contains exactly those values that are consistent with the data at the stated level. A hypothesis test asks whether one specific value is consistent with the data. So the interval answers, for every possible value at once, the question a test answers for one.

Concretely: if a proposed value falls outside a 95 per cent interval, a two-tailed test of that value at $\alpha = .05$ would reject it. If it falls inside, the test would not reject.

Given the interval (42.3, 57.8):

$\mu_0 = 40$. This falls *outside* the interval, below its lower bound of 42.3. The data are therefore inconsistent with a population mean of 40, and a test of that null hypothesis would reject it at $\alpha = .05$.

$\mu_0 = 50$. This falls *inside* the interval, and close to its centre. The data are consistent with a population mean of 50, and a test would not reject that null hypothesis.

Why no formal test is needed. Because the interval already encodes the answer. It was constructed from the same sample mean and standard error a test would use, and the boundary of the interval is precisely the point at which a test would change its verdict.

Two cautions worth adding.

Failing to reject $\mu_0 = 50$ does not establish that μ equals 50. The interval also contains 45, 52 and 57, and the data are equally consistent with all of them. This is the asymmetry Chapter 13 develops: evidence can rule values out and cannot confirm one.

And this correspondence holds for a two-tailed test at the matching level. A one-tailed test, or a different α , corresponds to a differently constructed interval.

The interval offers something a test does not: a test returns a verdict on the one

value the researcher happened to nominate, while the interval lets a reader check any value they care about — which is the argument for estimation-first reporting in §12.4.5.

Solution to Exercise 12.E6. Your answer will use a different variable from the one below; the reasoning is the same whichever you pick. The worked version uses `socsurvey`'s `income`, in thousands of dollars.

(a) Three figures, with three different meanings.

$$\bar{x} = 90.4 \quad s = 48.2 \quad SE = \frac{48.2}{\sqrt{477}} = 2.21$$

The *mean* estimates the population mean. The *standard deviation* describes how much individual respondents vary. The *standard error* describes how much the sample mean would vary across repeated samples of 477.

Notice the size of the third against the second: the standard error is roughly $1/\sqrt{477}$ — about one 22th — of the standard deviation. Estimates of the mean are far more stable than the variable they summarise.

(b) With $t^* = 1.96$:

$$90.4 \pm 1.96 \times 2.21 \Rightarrow (86.1, 94.8)$$

Reported in the format of §12.4.1, that is $M = 90.4$, 95% CI [86.1, 94.8], $n = 477$. All four elements are required: an interval without n , or without its confidence level, cannot be judged by a reader.

(c) Expect roughly nine or ten of the 10 to contain the full-sample mean, since a 95 per cent procedure misses about one time in twenty.

The instructive outcomes are the unexpected ones. Ten out of ten is common — with only 10 intervals, seeing zero misses is more likely than not. Eight out of ten happens too, and does not indicate an error.

The point is that 95 per cent is a long-run rate, and 10 trials is far too few to observe it reliably. It is the same point as Exercise 12.E2(c), arrived at in data you drew yourself.

(d) The intervals become *narrower* — by a factor of about $\sqrt{100/30} = 1.83$ — because the standard error shrinks with the square root of sample size.

What does *not* change is the coverage rate. Roughly the same proportion of intervals will contain the full-sample mean, because the confidence level governs how often the procedure succeeds regardless of n . A larger sample buys precision, not reliability.

That distinction is the most useful thing in the exercise, and it is the one most often expected to go the other way.

(e) Statistics Canada’s tax-filer programme puts average after-tax income for Canadian families at \$82.6k in 2023 (Statistics Canada 2025). The interval in (b) runs from \$86.1k upward, so the published figure falls *outside* it, by \$3.5k.

What that does not show. It is not evidence that Statistics Canada is wrong, and it is not evidence that the sample is unrepresentative in any way you could fix. The interval is a statement about the population *socsurvey* was drawn from — and *socsurvey* was not drawn from Canada. It was simulated. A confidence interval can only ever be about the process that generated the data, and here that process was a random number generator.

What it does show. Exactly how easy this mistake is to make. Every step was correct: the arithmetic, the format, the interpretation. The interval is a valid 95 per cent interval. It simply does not refer to Canada, and nothing inside the calculation says so — which is why the question of what population a sample came from is settled before any of this arithmetic begins (§5.1.2).

There is a second gap worth naming: the published figure is for *families*, while *socsurvey*’s unit is an individual respondent. Two figures called “average income” need not be averages of the same thing.