

Chapter 2 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 2 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 2 Fundamental Statistical Concepts

Solution to Exercise 2.E1. (a) Two problems, and both are vocabulary problems.

Median is not average. “Average” in ordinary usage suggests the mean, and for income the two differ substantially because the distribution is right-skewed — a long tail of high earners pulls the mean above the median. Reporting a median as an average invites the reader to imagine a symmetric distribution that does not exist.

“Earns” is not “total income.” The census figure covers total income, which includes government transfers, pensions, and investment income. Earnings are only part of it, and for many people — retirees, students, those on benefits — much the smaller part.

A third problem is easy to miss: the population is Canadians aged 15 and over, which includes secondary school students. “The average Canadian” implies working adults.

(b) The census figure is a statistic: a summary computed from the people who returned a form. The true median income of all Canadians is a parameter — the value that would result from observing everyone, which nobody has done. Even a census is a sample in practice, because coverage is incomplete and the long-form questionnaire goes to a fraction of households. The reported figure estimates the parameter; it is not the parameter.

(c) Several things. The mean alongside the median, which would reveal the skew. The margin of error. Whether the figure is before or after tax. Whether it is per person or per household — a household figure is much larger and the two are frequently confused. And the year, since income figures are revised and inflation makes cross-year comparison of nominal dollars misleading.

Solution to Exercise 2.E2. 1. *Self-reported health.* Categorical, ordinal; ordinal scale; median. The categories rank clearly but the distance from “good” to “very good” is undefined.

2. *Number of children under 18.* Numerical, discrete; ratio scale; mean, though the median is often more useful because the distribution is skewed and bounded below at zero.

3. *Country of birth.* Categorical, nominal; nominal scale; mode. The only average available.

4. *V-Dem Liberal Democracy Index.* Numerical, continuous; interval in practice, and this deserves comment. It runs 0 to 1 and looks ratio, but it is a composite of rescaled expert judgements, so a country at 0.8 is not twice as democratic as one at 0.4. Mean is conventional and defensible; treating it as ratio is not.

5. *Voted in the last election.* Categorical, nominal and binary; nominal scale; mode. Note the useful special case: coded 0/1, the mean equals the proportion who voted, so numerical tools can be applied without violating anything.

6. *Annual personal income*. Numerical, continuous; ratio scale; median, because income is right-skewed and the mean is pulled by the upper tail. This is exercise E1's point arriving from the other direction.

7. *Religiosity on a 1–10 scale*. Categorical, ordinal strictly; ordinal scale; median strictly, mean commonly and with the equal-interval assumption stated.

The pattern worth drawing out: items 4 and 7 are both scales running over a numeric range, and both are treated as interval in practice while being ordinal in principle. Items 5 and 6 show that the appropriate average depends on the distribution as well as the scale.

Solution to Exercise 2.E3. (a) A defensible conceptual definition: social capital is the set of resources available to people through their social networks — the connections, norms of reciprocity, and trust that allow individuals and groups to act together effectively. It is worth distinguishing *bonding* capital, within tight-knit groups, from *bridging* capital, across them: the two can vary independently.

(b) Three operationalisations at three scales. *Nominal*: whether a resident belongs to any voluntary association (yes/no). *Ordinal*: frequency of contact with neighbours on a five-point scale from never to daily. *Ratio*: number of voluntary associations per thousand residents, or the count of people a respondent could ask for a substantial favour.

(c) Threats to validity. Membership counts formal organisations and misses informal networks — which are the main form of social capital in some communities, so the measure will systematically understate them. Self-reported contact frequency is subject to recall error and social desirability. Organisational density measures supply rather than use: a neighbourhood may have many associations that few residents join.

(d) Organisational density is the most feasible from secondary data. Canadian sources include the Canada Revenue Agency's registered charities list, aggregated by postal area, and Statistics Canada's General Social Survey cycles on social identity and giving, which carry association membership and volunteering. Any specific, real, accessible source will do; "government data" is not one.

Solution to Exercise 2.E4. (a) Relative change is the difference divided by the starting value.

$$\frac{6,000 - 2,000}{2,000} \times 100 = \frac{4,000}{2,000} \times 100 = 200\%$$

The claim checks out. Note that a 200% increase means the figure tripled, not doubled — a common slip.

(b) Both are right, for different questions, and an answer choosing one without saying which question is incomplete. The absolute figure of 4,000 additional deaths

is what determines the burden on coroners, treatment services, and families — it is the policy-relevant magnitude. The rate per 100,000 is what makes the change comparable over time and across jurisdictions, because it removes the effect of population growth. A researcher asking whether the epidemic worsened needs the rate; an agency budgeting for response needs the count.

(c) Rates per 100,000:

$$2015 : \frac{2,000}{35,000,000} \times 100,000 = 5.71$$

$$2020 : \frac{6,000}{38,000,000} \times 100,000 = 15.79$$

The rate rose from 5.71 to 15.79 per 100,000, an increase of about 176% rather than 200%.

So population growth accounts for part of the headline increase, but only a small part. The interpretation is therefore *qualified rather than overturned* — and that is the honest answer. Concluding that the rate calculation “debunks” the 200% figure would be overcorrecting. The epidemic worsened substantially on either measure.

Solution to Exercise 2.E5. (a) The variable is ordinal in principle: nothing guarantees that the step from 2 to 3 represents the same increment of trust as the step from 7 to 8. Using the mean assumes equal intervals.

Whether it is appropriate depends on what is being claimed. For a difference this small — 5.4 against 5.1, three tenths of a point on a ten-point scale — the assumption is doing real work, and the answer would also depend on the sample size and the spread, neither of which is given. A defensible answer says: the mean is conventional and acceptable as a summary, but a difference of 0.3 should not be reported as a finding without a measure of uncertainty, and the distributions should be compared directly.

(b) Validity threats include social desirability, since expressing distrust of government may feel improper to some respondents; ambiguity about what “the federal government” refers to — the current administration, the institution, or the civil service — with different respondents answering about different objects; and the possibility that men and women use the scale differently, which would produce a measured difference with no underlying difference in trust. That last one is the most serious here, because it would manufacture exactly the finding reported.

(c) No. “60% more” is a ratio claim, and ratios require a true zero. On this scale, 0 means “no trust at all” as a *label*, not as an absence of a measurable quantity, and the scale is at best interval. The defensible statement is that the provincial means differ by 2.6 points on a ten-point scale — a difference, not a multiple.

Solution to Exercise 2.E6. (a) Numerical and continuous in appearance, though

built from discrete components. The scale is interval at best. It runs 0 to 100 with an apparent zero, but that zero is the bottom of a constructed scoring scheme rather than an absence of freedom, and the underlying components are expert judgements.

(b) No. A ratio claim requires a true zero, and a score of 0 on this index does not mean the complete absence of a measurable quantity called freedom — it means the lowest score the instrument can assign. There is also no guarantee that a ten-point difference means the same thing at the top of the scale as at the bottom.

(c) Several choices embed judgements, and naming any one of them answers the question. Summing 25 equally weighted questions asserts that each contributes equally to freedom — that a given restriction on press freedom is exactly as important as an equivalent restriction on assembly. The selection of the 25 questions determines what counts as freedom at all, and civil and political liberties are included while economic security is not. And the 0–4 scoring of each question is assigned by analysts applying criteria that require interpretation.

(d) The concern is construct validity, and specifically that freedom and democracy are not the same construct. Freedom House measures civil liberties and political rights; a country could conduct genuinely competitive elections while restricting press freedom, or protect liberties under a government nobody elected. Using one as a measure of the other assumes they travel together, which is an empirical claim rather than a definitional one. A critic might add that the index is produced by a single organisation with a known institutional perspective, so it should not be the only measure a study relies on.

Solution to Exercise 2.E7. Classifications for the Chapter 1 table.

Income bracket (7 categories). Categorical, ordinal; ordinal scale; median.

Trust in government (0–10). Ordinal in principle, treated as interval in practice; median strictly, mean conventionally.

Country of birth. Categorical, nominal; nominal scale; mode.

Number of children. Numerical, discrete; ratio scale; mean or median, with the median preferable given the skew.

Self-reported health (5-point). Categorical, ordinal; ordinal scale; median.

Party identification. Categorical, nominal; nominal scale; mode.

The exercise is designed to be revisited, so the useful question is what changed. In Chapter 1 students identified *limitations* of each measure — what it leaves out or blurs. Here they identify what may legitimately be *computed* from it. The two are connected: the income bracket's limitation (distances undefined, top bracket open) is exactly what makes it ordinal and rules out a mean.

Solution to Exercise 2.E8. 1. The population is residents of the municipality — or, more defensibly given the design, users of the library. The sample is the 600 residents surveyed. The statistic is the 32% figure. The parameter it purports

to estimate is the true change in mean wellbeing among the population, which is unknown.

2. Any defensible operationalisation works, provided it is specific and its scale is stated. A validated multi-item wellbeing scale administered before and after, producing an ordinal or quasi-interval score. Or a ratio measure: hours per week spent in social activity outside the home. Or nominal: whether the respondent reports having someone to call in an emergency.

3. “Rose 32%” applies a ratio operation to a scale that does not support it. A move from 5.0 to 6.6 on a 0–10 scale is a rise of 1.6 points; calling it 32% implies the scale has a true zero, which it does not. A better report states the before and after values and the difference in points, and gives a measure of uncertainty — for example, “mean wellbeing rose from 5.0 to 6.6 points on a ten-point scale.”

4. Reliability: a single-item measure administered once before and once after is vulnerable to mood, weather, and question order, so an apparent change may be noise. Validity: wellbeing scales are subject to social desirability, and respondents who know the programme is being evaluated may report improvement they do not feel — particularly if surveyed by the people running it.

5. Distributing the survey inside the library restricts the sample to library users, and to the more frequent users at that. The statistic can then say nothing about residents generally, and cannot support a claim that the programme improved community wellbeing — only, at most, that wellbeing rose among people already using the service. This is a selection problem, and no increase in sample size corrects it. Chapter 5 treats it fully.

The thing to notice is that item 5 undermines item 1: the population named in the claim and the population the sample can speak about are different.

Solution to Exercise 2.E9. The answer depends on the wave you look at and on current StatCan figures. The reasoning is what carries here, not the particular cities.

(a) In `cma_panel`, the highest simulated crime rate belongs to Winnipeg, and the lowest to one of the Quebec CMAs. Read this from the table rather than assuming it.

(b) On the real Crime Severity Index, the highest-scoring CMAs are consistently western — Kelowna, Lethbridge, Winnipeg, Regina, Saskatoon — and the lowest are the Quebec CMAs, with Toronto also near the bottom.

(c) Broadly yes at the extremes, and imperfectly in the middle. Winnipeg being high and the Quebec CMAs low matches. Finding a discrepancy is not an error: the Preface promises that orderings track reality *broadly*, not exactly, and the simulated levels are deliberately wrong throughout.

(d) The Crime Severity Index weights each offence by the average sentence handed down for it, so a homicide contributes far more than a shoplifting incident. What this *adds* is a distinction a simple rate cannot make: two cities with identical

incident rates but different offence mixes are correctly separated. What it *obscures* is the weighting itself — the index is a composite (§2.6.6), and its weights encode sentencing practice, which reflects prosecutorial and judicial decisions rather than harm alone. A change in sentencing law changes the index without any change in crime.

Comparing it to a simple rate is therefore not straightforward, because the two are different constructs on different scales. The CSI is an index number normalised to a base year, not a count per capita; a city can rank higher on one and lower on the other. This is a measurement-validity question rather than a data problem, and naming it as one is the point of the exercise.