

# Chapter 17 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 17 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 17 Analysis of Variance

**Solution to Exercise 17.E1.** (a)  $df_{\text{between}} = k - 1 = 3$  and  $df_{\text{within}} = N - k = 100 - 4 = 96$ .

They add to  $3 + 96 = 99$ , which is  $N - 1$ . The check passes.

|     | Source         | $df$ | Sum of squares | Mean square | $F$  |
|-----|----------------|------|----------------|-------------|------|
| (b) | Between groups | 3    | 180            | 60          | 6.26 |
|     | Within groups  | 96   | 920            | 9.583       |      |
|     | Total          | 99   | 1100           |             |      |

Each mean square is its sum of squares over its degrees of freedom, and  $F = 60/9.583 = 6.26$ , with  $p < .001$ .

(c)  $\eta^2 = 180/1100 = .164$ .

(d) *Mean scores differed across the four groups*,  $F(3, 96) = 6.26$ ,  $p < .001$ ,  $\eta^2 = .164$ . The grouping accounts for about sixteen per cent of the variation, and which groups differ is not established by this test.

**Solution to Exercise 17.E2.** (a) *Independent-samples  $t$ -test* (§15.3). Two groups, one continuous outcome, and random assignment — so the difference is an effect rather than an association.

(b) *One-way analysis of variance* (this chapter). 5 groups rules out the  $t$ -test, and running the ten pairwise comparisons instead would carry a familywise rate of about forty per cent. Note that education level is *ordered*, which analysis of variance ignores; a regression treating it as a numeric predictor would use that information, at the cost of assuming the steps are equal.

(c) *Simple linear regression* (Chapter 16). Two continuous variables, and the question is how much, not whether the groups differ.

(d) *Chi-square test of independence* (Chapter 14). Two categorical variables and no continuous outcome at all.

(e) *Paired-samples  $t$ -test* (§15.4). The same twenty people appear twice, so the two columns are linked case by case and treating them as independent groups would be the error of §15.4.3.

**Solution to Exercise 17.E3.** (a)  $df_1 = 3$  means  $k = 4$  groups.  $df_2 = N - k$ , so  $N = 1,996 + 4 = 2,000$ .

(b)

$$\eta^2 = \frac{45.2 \times 3}{45.2 \times 3 + 1,996} = .064$$

(c) An  $F$  of 45.2 is enormous and would be reported as decisive. The effect it corresponds to is 6.4 per cent of the variation — real, and small. The two are compatible because  $F$  rises with sample size and  $\eta^2$  does not, and this study has 2,000 cases.

(d) That they have reported the effect size whether they meant to or not.  $F$  and its two degrees of freedom determine  $\eta^2$  exactly, so a reader can always recover it. The omission is therefore not a shortage of information but a choice about what to put in front of the reader — and the quantity left out is the one that speaks to how much the finding matters.

**Solution to Exercise 17.E4.** (a) Equal variances, and badly. The largest standard deviation is 9.8 against a smallest of 2.1 — a ratio of 4.7, so the variances — which are the squares — differ by a factor of  $4.7^2 = 22$ . The rule of thumb in §17.4.1 asks for a standard deviation ratio below about 2.

(b) The group means themselves are unaffected: each is computed from its own group and nothing pools them. What to distrust is  $MS_{\text{within}}$ , and therefore  $F$ , the  $p$ -value and any interval built from the pooled standard error. The estimate survives; the stated confidence does not.

(c) Use Welch's correction (§17.4.3). It will cost degrees of freedom — with one group of six and one group far more variable than the rest, probably a great many — and that reduction is the honest accounting of how much independent information is actually present.

(d) Yes, and it makes the finding more interesting rather than less. A waiting-list control that varies four times as much as the treatment groups suggests the treatments did something the comparison of means alone does not capture: they made people more alike. A difference in variance is a result, not only an inconvenience, and it is worth reporting as one.

**Solution to Exercise 17.E5.** (a) First, the three  $t$ -tests are uncorrected, so their familywise error rate is 14.3 per cent rather than 5. Second — and worse — only the two significant ones were reported. Selecting which comparisons to publish on the basis of their results means the reported  $p$ -values describe neither the tests that were run nor the ones that were not.

(b) It would fix the second problem and not the first. Reporting all three uncorrected tests is honest about what was done and still understates the error rate.

(c) The omnibus test, which they already have, followed by a post-hoc procedure reporting *all* 3 comparisons with adjusted  $p$ -values and intervals — the analysis of §17.5.3. If the comparisons of interest had been specified before the data were collected, they could instead run those as planned comparisons and say so (§17.5.4).

**Solution to Exercise 17.E6.** Your figures depend on which download you use, but the shape of the reasoning does not. What to expect, and why:

(a) Income group is strongly related to life expectancy in *worldsim*, so  $F$  is very large and  $\eta^2$  is high — far higher than anything else in this chapter. Report both degrees of freedom, which will be  $k - 1$  and  $N - k$  for the 52 countries of wave 1.

(b) More reasonable than for `region`, though not perfect: the income groups are of uneven size and the poorest group has the widest spread in life expectancy. Compare the standard deviations rather than testing them.

(c) Little. Welch will cost some degrees of freedom and the conclusion will not move, which is the case §17.4.3 describes as the usual one — and the reason to use the correction by default rather than only when it matters.

(d) `region` gives a smaller  $\eta^2$  than `income_grp` because regions are internally heterogeneous in exactly the way that inflates  $SS_{\text{within}}$ : a region contains countries at very different income levels, and income is what drives life expectancy here. Grouping by something closer to the cause puts more of the variation between groups and less inside them. The grouping variable is a choice, and  $\eta^2$  is partly a measure of how well it was chosen.

**Solution to Exercise 17.E7.** Your figures depend on the census year and on which provinces you choose, but what you should be able to say about them does not. What to expect, and why:

(a)–(b) Report  $F$  with both degrees of freedom,  $p$ , and  $\eta^2$ . With three provinces and a handful of metropolitan areas each,  $N$  will be small — perhaps fifteen or twenty — so the degrees of freedom will be small and the critical value correspondingly large.

(c) You would not treat  $\eta^2$  as an estimate of a population value: at this sample size it is badly biased upward (§17.5.2), and a large one is the expected result whether or not provinces really differ. You would report it as a description of these areas and say so.

(d) Two reasons, and they pull in opposite directions. A metropolitan area's median income is an average over hundreds of thousands of people, so the noise that makes individual-level  $\eta^2$  small has been averaged away — which inflates  $\eta^2$  and makes area-level groupings look much tidier. But  $F$  also depends on sample size, and 39,085 individuals give an  $F$  of 420 where twenty areas cannot approach it. So the area-level analysis will typically show a *larger*  $\eta^2$  and a *smaller*  $F$  — which is the clearest demonstration in this book that the two answer different questions.