*Data Analysis for Managers*, 2e

Albright, Winston, Zappe

This is the second edition of the book originally entitled Managerial Statistics.

We tried to find all the typos, bad numbers, wrong solutions, and so forth, but with so many details, we missed a few. As we (or you) find errors, we'll list them on this page by chapter. Note that if there is an error in a solution file for some problem, we will try to indicate a way to fix the error without "giving away" the solution to students who might be reading this page. The date of the correction is shown in parentheses.

### Chapter 1

* (4/30/2004) The label toward the bottom of page 24 should be Example 16.2, not 16.3.

### Chapter 2

* (8/1/03) On page 42, the first line after Solution should have 2.5 minutes, not 2.5 seconds.
* (8/1/03) The screenshot in Figure 2.26, page 58, should show counts (18, 48, and 66), not percentages.
* (8/1/03) It's minor, but the third step on page 62 says to format as currency with 2 decimals, and the screenshot below shows 3 decimals. Take your pick!
* (9/23/04) Here is a really interesting one. A user said that they couldn't group on Age in the pivot table requested in problem 43. If you try creating a pivot table from the data in the problem 43 data set, you'll find that the pivot table wizard's guess for the data range extends a couple of rows below the data-- to row 1005 instead of 1003. Evidently, there is some junk (blank text?) in these rows. If you include them in your pivot table range, there will be ages that aren't numbers; hence, grouping on age won't be allowed. The easiest solution is to delete rows 1004 and 1005 before building the pivot table. We didn't introduce these "junk" rows on purpose, but this illustrates how you have to be careful when identifying the range of a data set. A cell that looks empty might not really be empty!

### Chapter 3

* (11/28/2003) One user pointed out that the data file for problem 84 in Chapter 3, page 123, has some errors. Well, maybe. One of its variables has missing data, and the missing entries are signaled by large negative numbers. Unfortunately, this is not at all unrealistic. Real data sets can have missing values that are signaled by weird values. You have to be on the lookout for them! By the way, if the missing values in this text file had been left blank, then you'd have another problem when opening this file in Excel: The columns wouldn't line up correctly. Do you see why?
* (2/10/2004) The solution to problem 27, page 91, isn't correct. Note that the female categories first increment by 10, but then by 20. The solution has to take this into account.
* (9/2/2004) The solution to problem 37 in the Student Solutions Manual is wrong. It's based on the data set for this problem in the first edition of the book. We updated the data set for the second edition. (The solution in the Instructor's Manual is correct.)

### Chapter 4

* (8/27/03) The .mdb file for the EduToys case at the end of the chapter was omitted from the CD-ROM. You can download it from here. (I've renamed it simply as EduToys.mdb.) Download: [EduToys.mdb](file:///C%3A%5CUsers%5Cchris%5CDropbox%5CMy%20Books%5CWeb%20Site%5CDownloads%5CDAM2e%5CEduToys.mdb)

### Chapter 5

* (4/16/2004) The range name used in the Substitutes.xls file for Example 5.4, page 213, is CovarDem, not CovDem. The references to CovDem (in the textbox in the spreadsheet and in the formula toward the bottom of page 216, should be changed to CovarDem.
* (4/16/2004) The solution to part (c) of problem 5.12 on page 202 is wrong. (You can't use the >=4 label as part of the formula.) The fix should be clear.
* (4/16/2004) The solution to problem 5.38 on page 223 refers to equation 4.12. The reference should be to equation 5.13.

### Chapter 7

* (11/28/2003) The wording of problem 36 on page 343 isn't as explicit as it could be. Although it is not stated, the assumption is that the probability of a false detection is 0. That is, an inspection will never detect a problem when there is no problem.
* (1/25/2004) There is a subtle error in the statements of problems 28, 29, 31, and 32 on pages 335-336. If A is an event and B is an information outcome, then the law of total probability requires that the prior probability of A must satisfy P(A) = P(A|B)P(B) + P(A|not B)P(not B). In these problems we've given all of the probabilities on the right-hand side of this equation, but we're also given the prior probability of A, and it does *not* equal the right-hand side of the equation. Therefore, the priors should be changed as follows. When these changes are applied to the solutions we've supplied, the solutions will change accordingly -- and make more sense.
	+ Problem 28: Change the prior probability that the field contains oil from 50% to 48%.
	+ Problem 29: Change the prior probabilities for the maintenance cost of the first machine from 0.35, 0.35, and 0.30 to 0.325, 0.475, and 0.20. (Also, to make this problem more interesting, we suggest that you change the purchase cost of the first machine from $4500 to $3100.)
	+ Problem 31: Change the prior probability that Penn State will win from 0.50 to 0.475. Also, the solution to this problem doesn't reflect the problem statement exactly. The solution implies that Hank's only two options are to bet on Penn State or not bet at all. If you want to allow Hank to bet on Michigan, the solution should be changed accordingly.
	+ Problem 32: Change the prior probability that the new toothpaste will succeed from 0.55 to 0.50.

### Chapter 8

* (4/16/2004) The reference to cell B9 on page 390 (just above the last formula) should be changed to E9.

### Chapter 9

* (10/22/03) In the formula for sp at the top of page 442, the whole right-hand-side should be under a square root sign.
* (10/22/03) In the formula for SE(p1-p2) at the top of page 454, the right-hand-side should be the square root of the sum, not the sum of the square roots. (This *is* implemented correctly in the example that follows.)
* (10/22/03) There is a bug in StatPro on the p-value for the test of equality of variances used in the two-sample procedure. (See Figure 9.13, page 442, e.g.) The formula is based on a one-tailed test, but it should be based on a two-tailed test. In general, the correct formula should be =2\*MIN(FDIST(R,SS1-1,SS2-1),1-FDIST(R,SS1-1,SS2-1)), where R is the ratio of variances (with the variance from sample 1 in the numerator) and SS1 and SS2 are the sample sizes. E.g., the formula in cell F24 for Figure 9.13 should be =2\*MIN(FDIST(F23,F9-1,G9-1),1-FDIST(F23,F9-1,G9-1)).
* (11/12/2003) Right above Figure 9.9 on page 435 there is an error. In Goal Seek, the set cell should be D10, not B10.
* (4/16/2004) In the solution to problem 1, parts d and e, on page 422, change 1.8 to 0.8 to match the numbers in the problem statement.

### Chapter 10

* (11/3/2003) The powerpoint slides that accompany the book have an error. Slides 10 and 11 for this chapter refer to sigma. They should refer to alpha.
* (11/5.2003) The solution to problem 10.3 (page 494) uses a one-tailed test. Given the wording of the problem, it seems more appropriate to use a two-tailed test.

### Chapter 11

* (11/11/02) I have Professor Krish Muralidhar from University of Kentucky to thank for this one. In our validation section, Section 11.7, we calculate a "standard error of estimate" for the new data. This could be misleading if there is a "mean shift" in the new data. In this case the standard deviation of the residuals, basically the measure we calculated, could still be low even though the fitted values are shifted from the actual values. (To see what I mean, add 1000 to all of the new overhead values. You'll see that the two validation measures in the book don't change at all.) The problem is that the new residuals no longer have mean 0. The fix is to calculate a different "standard error" measure for the new data. It should be =SQRT(SUMSQ(residuals)/(n-k-1)), where "residuals" is a column of the new residuals, and k is the number of explanatory variables. If you calculate this measure when the overhead values are shifted by 1000, you'll get a huge value, indicating that the old equation doesn't work well at all for the new data.

### Chapter 12

* (10/22/03) The CD-ROM has two versions of the Catalogs.xls file used in Example 12.3, one "data only" and one "finished". Unfortunately, they have *different data* for the PrevSpent variable from customer 251 down. The statistics quoted at the bottom of page 642 are based on the "finished" data, but you'll get different statistics if you start with the "data only" data. I suggest you use the following "data only" file: [Catalogs.xls](file:///C%3A%5CUsers%5Cchris%5CDropbox%5CMy%20Books%5CWeb%20Site%5CDownloads%5CDAM2e%5CCatalogs.xls). It matches the "finished" version.
* (5/28/2005) The Dupree case on page 685 states that there are 67 observations. There are actually only 40 observations in the file.

### Chapter 13

* (11/28/2003) The formula for RMSE on page 697, equation 13.3, is missing a square root sign. The whole expression on the right-hand side should be under a square root.
* (11/28/2003) This is not an error in the text, but rather in the cell comment accompanying the StatPro runs test. It states: "If this [p-] value is sufficiently small, say, less than .05, you can reject the null hypothesis of no randomness and conclude that the series does not alternate enough (too few runs) or alternates too much (too many runs)." The italicized no should be deleted. The null hypothesis is that the series is random.

Visit the [Cengage](http://www.cengage.com/decisionsciences/albright) site for our books.

Send e-mail to albright@indiana.edu

Albright is retired from the [Kelley School of Business, Indiana University, Bloomington](http://www.kelley.indiana.edu/) and now works as a consultant for [Palisade Corp](http://www.palisade.com/).

Back to [home page](file:///C%3A%5CUsers%5Cchris%5CDropbox%5CMy%20Books%5CWeb%20Site%5Cdefault.htm)

Updated: 1/13/2015