



# Instructor Solutions Manual for Essential Statistics, Regression, and Econometrics

Second Edition

Gary Smith

Pomona College



Amsterdam • Boston • Heidelberg • London • New York • Oxford  
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo  
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier  
225 Wyman Street, Waltham, MA 02451, USA  
525 B Street, Suite 1800, San Diego, California 92101-4495, USA  
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

© 2015 Gary Smith. Published by Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: [www.elsevier.com/permissions](http://www.elsevier.com/permissions)

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

### Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

For information on all Academic Press publications visit our website: [www.elsevierdirect.com](http://www.elsevierdirect.com)

Working together to grow  
libraries in developing countries

[www.elsevier.com](http://www.elsevier.com) | [www.bookaid.org](http://www.bookaid.org) | [www.sabre.org](http://www.sabre.org)

ELSEVIER

BOOK AID  
International

Sabre Foundation



# Contents

|                                   |     |
|-----------------------------------|-----|
| 1. Data, Data, Data               | 1   |
| 2. Displaying Data                | 9   |
| 3. Descriptive Statistics         | 29  |
| 4. Probability                    | 41  |
| 5. Sampling                       | 51  |
| 6. Estimation                     | 59  |
| 7. Hypothesis Testing             | 69  |
| 8. Simple Regression              | 79  |
| 9. The Art of Regression Analysis | 101 |
| 10. Multiple Regression           | 121 |
| 11. Modeling (Optional)           | 139 |

This page intentionally left blank

# Data, Data, Data

- 1.1** Quantitative data have natural numerical values; qualitative (or categorical) data do not.
  - a.** Quantitative.
  - b.** Quantitative.
  - c.** Qualitative.
  - d.** Quantitative.
- 1.2** Unlike numerical data, categorical data do not have natural numerical values.
  - a.** Categorical.
  - b.** Numerical.
  - c.** Numerical.
  - d.** Categorical.
- 1.3** Unlike numerical data, categorical data do not have natural numerical values.
  - a.** Categorical.
  - b.** Numerical.
  - c.** Numerical.
  - d.** Categorical.
- 1.4** Cross-sectional data are at the same point in time. Time series data are at different points in time. Longitudinal (or panel) data are the same things at different points in time.
  - a.** Cross-sectional data.
  - b.** Cross-sectional data.
  - c.** Time series data.
  - d.** Panel data.
- 1.5** This depends on current Big Mac prices and exchange rates.
- 1.6** Cross-sectional data are at the same point in time. Time series data are at different points in time. Longitudinal (or panel) data are the same things at different points in time.
  - a.** Time series data.
  - b.** Time series data.
  - c.** Cross-sectional data.
  - d.** Cross-sectional data.

## 2 ESSENTIAL STATISTICS, REGRESSION, AND ECONOMETRICS

- 1.7** Cross-sectional data are at the same point in time. Time series data are at different points in time. Longitudinal (or panel) data are the same things at different points in time.
- a. Time series data.
  - b. Time series data.
  - c. Cross-sectional data.
  - d. Cross-sectional data.
- 1.8** Cross-sectional data are at the same point in time. Time series data are at different points in time. Longitudinal (or panel) data are the same things at different points in time.
- a. Time series data.
  - b. Cross-sectional data.
  - c. Panel data.
  - c. Panel data.
- 1.9** The level of a price index does not mean anything, but percentage changes in a price index measure percentage changes in prices.
- a. No, index numbers cannot be used to compare prices in a single year.
  - b. No, index numbers cannot be used to compare prices in a single year.
  - c. Yes, comparing the 2000 and 2010 index values, the percentage increase in the price of housing was larger than the percentage increase in the price of food.
- 1.10** The level of a price index does not mean anything, but percentage changes in a price index measure percentage changes in prices.
- a. No, index numbers cannot be used to compare prices in a single year.
  - b. Yes, the index was lower in 2010 than in 2000.
  - c. Yes. Food:  $100(219.6 - 168.4)/168.4 = 30.4$  percent. CPI:  $100(218.1 - 172.2)/172.2 = 26.7$  percent.
- 1.11** Food,  $100(219.6 - 168.4)/168.4 = 30.4$  percent; apparel  $100(119.5 - 129.6)/129.6 = -7.8$  percent; and housing  $(100)(216.3 - 169.6)/169.6 = 27.5$  percent.
- 1.12** The level of a price index does not mean anything, but percentage changes in a price index measure percentage changes in prices.
- a. No, index numbers cannot be used to compare prices in a single year.
  - b. No, index numbers cannot be used to compare prices in a single year.
  - c. Yes, the index declined between 2000 and 2010.
- 1.13** The level of a price index does not mean anything, but percentage changes in a price index measure percentage changes in prices.
- a. No, index numbers cannot be used to compare prices in a single year.
  - b. Yes, the index declined between 2000 and 2010.

- c. Yes, comparing the 2000 and 2010 index values, the percentage decrease in the cost of boys' apparel was larger than the percentage decrease in the cost of men's apparel.

**1.14** Men  $(100)(117.5 - 133.1)/117.5 = -11.7$  percent; women  $(100)(109.5 - 121.9)/121.9 = -10.1$  percent; boys  $(100)(91.5 - 116.2)/116.2 = -21.3$  percent; and girls  $(100)(95.4 - 119.7)/119.7 = -20.3$  percent.

**1.15** It makes no sense to deflate a country's population by its price level because the population is not measured in dollars.

**1.16** Using hamburger prices, real wealth is:

$$\frac{\$100,000}{\$2/\text{hamburger}} = 50,000 \text{ hamburgers}$$

This measure of real wealth tells us that \$100,000 is enough to buy 50,000 hamburgers. Using hourly wages, real wealth is:

$$\frac{\$100,000}{\$10/\text{hour}} = 10,000 \text{ hours}$$

This measure of real wealth tells us that this person would have to work for 10,000 hours to earn \$100,000.

A price deflator measures wealth by what the wealth can buy; a wage deflator measures wealth by how much a person would have to work to obtain that wealth.

**1.17** The cost of the market basket in 2000 is  $10(\$2.50) + 6(\$4.00) + 3(\$5.00) = \$64.00$ ; the cost of the market basket in 2010 is  $10(\$3.00) + 6(\$5.00) + 3(\$7.00) = \$81.00$ , a percentage increase of  $100(\$81 - \$64)/\$64 = 26.6$  percent.

a. We can scale the index to equal 100 in 2000 by multiplying the 2000 and 2010 costs by  $100/64$ : 2010  $(100/64)(64) = 100$ ; 2010  $(100/64)(81) = 126.6$ . The index increased by  $(100)(126.6 - 100)/100 = 26.6$  percent, as did food prices.

b. We can scale the index to equal 100 in 2010 by multiplying the 2000 and 2010 costs by  $100/81$ : 2010  $(100/81)(64) = 79.0$ ; 2010  $(100/81)(81) = 100$ . The index increased by  $(100)(100 - 79.0)/79.0 = 26.6$  percent, as did food prices.

**1.18** The cost of the market basket in 2000 is  $12(\$2.50) + 6(\$4.00) + 2(\$5.00) = \$64.00$ ; the cost of the market basket in 2010 is  $12(\$3.00) + 6(\$5.00) + 2(\$7.00) = \$80.00$ , a percentage increase of  $100(\$80 - \$64)/\$64 = 25.0$  percent.

a. We can scale the index to equal 100 in 2000 by multiplying the 2000 and 2010 costs by  $100/64$ : 2010  $(100/64)(64) = 100$ ; 2010  $(100/64)(80) = 125.0$ . The index increased by  $(100)(125.0 - 100)/100 = 25.0$  percent, as did food prices.

b. We can scale the index to equal 100 in 2010 by multiplying the 2000 and 2010 costs by  $100/80$ : 2010  $(100/80)(64) = 80.0$ ; 2010  $(100/80)(80) = 100$ . The index increased by  $(100)(100 - 80.0)/80.0 = 25.0$  percent, as did food prices.

- 1.19** The Dow. The percentage increase in the Dow was  $100(16,804.71 - 240.01)/240.01 = 6901.67$  percent. The percentage increase in the CPI was  $100(712.3 - 51.3)/51.3 = 1288.50$  percent.
- 1.20** The Dow Jones Average is calculated by dividing the sum of 30 prices by a divisor that is adjusted for a variety of reasons to maintain a consistent index over time. The average price is equal to  $(\text{index})(\text{divisor})/30$ , since  $(\text{index})(\text{divisor})$  is equal to the sum of the 30 prices.
- a. No. The index value of 17,295.92 is obtained by dividing the sum of the prices by the 0.15571590501117 divisor, not by 30.
  - b. No. The index value divided by the divisor is equal to the sum of the prices, divided by the divisor squared.
  - c. No. The index value multiplied by the divisor is equal to the sum of the prices.
- 1.21** The CPI was 37.7 in 1969 and 126.1 in 1989, for a percent increase of  $100(126.1 - 37.7)/37.7 = 234.5$  percent. The Dow was 800.36 on December 31, 1969 and 2753.20 on December 29, 1989 (the market was not open on December 31), for a percent increase of  $100(2753.20 - 800.36)/800.36 = 244.0$  percent. Stock prices increased slightly more than the CPI, so the real value of stocks increased slightly.
- 1.22** The CPI was 126.1 in 1989 and 215.9 in 2009, for a percent increase of  $100(215.9 - 126.1)/126.1 = 71.2$  percent. The Dow was 2753.20 on December 29, 1989 (the market was not open on December 31) and 10,428.05 on December 31, 2009, for a percent increase of  $100(10,428.05 - 2753.20)/2753.20 = 278.8$  percent. Stock prices increased much more than the CPI, so the real value of stocks increased greatly.
- 1.23** The real stock price is equal to the nominal stock price divided by the CPI. If real and nominal stock prices were equal in 1864–1865, then the CPI equaled 1. If real stock prices increased more than nominal prices between 1865 and 1918, the CPI must have fallen below 1 (deflation). If real stock prices were again equal to nominal prices during World War II, the CPI must have risen back to 1 (inflation between 1918 and World War II). If real stock prices have fallen below nominal prices since then, the CPI must have risen above 1 (inflation). Their comparison tells us nothing about stock prices, only that the CPI fell between 1865 and 1918 and has generally risen since then, with the CPI in World War II roughly equal to its level in 1864–1865, and higher since World War II.
- 1.24** For nominal stock prices to again equal real stock prices, the CPI would have to fall back to 1, its level in 1864–1865.

**1.25** Using the data in Tables 1.8 and 1.9:

**a.** We multiple the 1970 wage by the increase in the CPI between 1970 and 2010:

$$\text{1970 wage in 2010 dollars} = (\text{1970 wage}) \left( \frac{\text{2010 CPI}}{\text{1970 CPI}} \right) = \$3.40 \left( \frac{653.2}{116.3} \right) = \$19.10$$

**b.** The percentage change in real wages between 1970 and 2010 can be calculated by comparing the actual 2010 wage with the 1970 wage in 2010 dollars:  $100(19.07 - 19.10)/19.10 = -0.15$  percent, a decline of less than 1 percent. Real wages did not increase over this 40-year period.

**1.26** Using the data in Tables 1.8 and 1.9:

**a.** We multiple the 1980 wage by the increase in the CPI between 1980 and 2000:

$$\text{1980 wage in 2000 dollars} = (\text{1980 wage}) \left( \frac{\text{2000 CPI}}{\text{1980 CPI}} \right) = \$6.85 \left( \frac{515.8}{246.8} \right) = \$14.32$$

**b.** The percentage change in real wages between 1980 and 2000 can be calculated by comparing the actual 2000 wage with the 1980 wage in 2000 dollars:  $100(14.02 - 14.32)/14.32 = -2.1$  percent. Real wages decreased over this 20-year period.

**1.27** Poverty thresholds can be converted to 1960 dollars by deflating by the CPI for that year:

|      | Poverty  | CPI   | 100(Poverty)/CPI |
|------|----------|-------|------------------|
| 1960 | \$3022   | 100.0 | \$3022           |
| 1970 | \$3968   | 116.3 | \$3412           |
| 1980 | \$8414   | 246.8 | \$3409           |
| 1990 | \$13,359 | 391.4 | \$3413           |
| 2000 | \$17,604 | 515.8 | \$3413           |
| 2010 | \$22,050 | 653.2 | \$3376           |

Since 1970, the thresholds have increased at roughly the rate of inflation, so that the real value has been constant.

**1.28** Because prices increased by a factor of  $653.2/45.6$ , the value of his contract in 2010 dollars is  $(653.2/45.6)\$80,000 = \$1,145,965$ . This is much smaller than average Major League Baseball salaries in 2010.

**1.29** Because prices increased by a factor of  $653.2/27$ , the value of his one-game salary in 2010 dollars is  $(653.2/27)\$500 = \$12,096$ , which is far less than the average NFL player was paid in 2010.

**1.30** Instead of dividing the number of applications each year by 1.04, he should have divided by an amount that increases by 4 percent a year, for example, leaving the 1987 value as is, dividing the 1988 value by 1.04, and dividing the 1989 value by  $1.04^2$ .

- 1.31** The real value of the tooth fairy payments fell by almost half, from  $\$0.12/25 = 0.0048$  in 1900 to  $\$1/340.4 = 0.0029$  in 1987. If tooth fairy payments had kept up with inflation between 1900 and 2010, the 2010 payment would have been  $(653.2/25)\$0.12 = \$3.14$ .
- 1.32** We need to deflate the \$90,516 benefit by the projected increase in the CPI over this 46-year period:  $\$90,516/1.035^{46} = \$18,598$  (less than half the 2008 salary).
- 1.33** This number does not take into account the fact that New Jersey is a very small state, with relatively little air in which to release chemicals. When the amount of toxic chemicals is divided by New Jersey's area (in square miles), giving the pounds released per square mile, there were only three states that polluted more than New Jersey.
- 1.34** In order to determine what the value of the Dow would have been if it had increased by the same amount as consumer prices between 1970 and 2010, we need to inflate the 1970 Dow value of 838.92 by the increase in the CPI over this 40-year period:  $838.92(219.2/39.8) = 4620.38$ . The actual value of the Dow on December 31, 2010 was 11,577.51; therefore, the Dow increased much more than the CPI over this 40-year period.
- 1.35** They received more German currency for returning the bottles, but this currency presumably was worth less because of the hyperinflation.
- 1.36** Because the price level increased by more than the amount of currency in circulation, the real value of the money supply decreased.
- 1.37** There is no guarantee that the same stuff and nonsense are put in the bags. Perhaps 10 years ago she had a dog and young children who ate bulky, but inexpensive, dog food and cereal, but now it is just her and her husband eating steak and shrimp. Her inflation calculation is flawed by a neglect of time. Her 153 percent price increase is over a 10-year period, which is about 9.7 percent a year, which is not inconsistent with government data.
- 1.38** We inflate the 1971 cost by the increase in prices between 1971 and 2008:  $\$0.08(100/19.0) = \$0.421$ , slightly higher than the actual 2008 cost.
- 1.39** We inflate the 1991 cost by the increase in prices between 1991 and 2008:  $\$0.29(100/64.3) = \$0.451$ , higher than the actual 2008 cost.
- 1.40** The percentage increase in the CPI was  $100(100.0 - 97.1)/97.1 = 2.99$  percent.
- 1.41** The percentage increase in the CPI was  $100(42.8 - 30.7)/30.7 = 39.41$  percent.
- 1.42** The CPI in 1958 was 13.6. To set this equal to 100, we multiply by  $(100/13.6)$ :  $(100/13.6)(13.6) = 100.0$ . To keep the relative values constant, we consequently

multiply all the CPI values in Table 1.10 by  $(100/13.6)$ , including the 2008 value:  $(100/13.6)(100) = 735.3$ . The percentage change in the CPI does not depend on the base year. If  $\text{CPI} = 100$  in 2008, the percentage change between 1958 and 2008 is  $100(100 - 13.6)/13.6 = 635.3$  percent. If  $\text{CPI} = 100$  in 1958, the percent change between 1958 and 2007 is  $100(735.3 - 100)/100 = 635.3$  percent.

- 1.43** The CPI in 1985 was 50.7. To set this equal to 100, we multiply by  $(100/50.7)$ :  $(100/50.7)(50.7) = 100.0$ . To keep the relative values constant, we consequently multiply all the CPI values in Table 1.10 by  $(100/50.7)$ , including the 2008 value:  $(100/50.7)(100) = 197.2$ . The percentage change in the CPI does not depend on the base year. If  $\text{CPI} = 100$  in 2008, the percentage change between 1985 and 2008 is  $100(100 - 50.7)/50.7 = 97.2$  percent. If  $\text{CPI} = 100$  in 1985, the percentage change between 1985 and 2007 is  $100(197.2 - 100)/100 = 97.2$  percent.

- 1.44** In 1960 and 1990 cigarette consumption per person over the age of 18 was:

$$1960: \frac{484.4 \text{ billion}}{116.1 \text{ million}} = 4172$$

$$1990: \frac{525 \text{ billion}}{185.7 \text{ million}} = 2827$$

which represents a 32 percent decline.

- 1.45** The costs are:

$$C_{2000} = \$2.00(1000) + \$1.00(500) = \$2500$$

$$C_{2010} = \$3.00(1000) + \$2.00(500) = \$4000$$

This is a 60.0 percent increase in his cost of living:  $100(\$4000 - \$2500)/\$2500 = 60.0$  percent. If the BPI is scaled to equal 100 in 2000, the 2010 BPI would be 60.0 percent higher: 160.0.

- 1.46** The percentage increase in the price of beer is  $100(\$3.00 - \$2.00)/\$2.00 = 50.0$  percent. The percentage increase in the price of pretzels is  $100(\$2.00 - \$1.00)/\$1.00 = 100.0$  percent. Of the total \$2500 spent on beer and pretzels, a fraction  $\$2000/\$2500 = 0.80$  was spent on beer and a fraction  $\$500/\$2500 = 0.20$  was spent on pretzels. Thus, the overall percent increase in spending is  $(50.0 \text{ percent})(0.80) + (100.0 \text{ percent})(0.20) = 60.0$  percent, which is the same answer reached in Exercise 1.45.

- 1.47** The costs are:

$$C_{2000} = \$2.00(1000) + \$3.00(400) = \$3200$$

$$C_{2010} = \$2.50(1000) + \$5.00(400) = \$4500$$

This is a 40.6 percent increase in her cost of living:  $100(\$4500 - \$3200)/\$3200 = 40.6$  percent. If the VPI is scaled to equal 100 in 2000, the 2010 VPI is 40.6 percent higher: 140.6.

**1.48** The costs are:

$$C_{2000} = \$2.00(1200) + \$3.00(250) = \$3150$$

$$C_{2010} = \$2.50(1200) + \$5.00(250) = \$4250$$

This is a 34.9 percent increase in her cost of living:  $100(\$4250 - \$3150)/\$3150 = 34.9$  percent. The percentage increase in her cost of living is lower than calculated in the preceding exercise because she shifted her spending away from tofu, which had a larger price increase than carrot juice.

**1.49** The rate of inflation slowed, but prices were still increasing.

**1.50** We need to adjust the data for the number of drivers (or the number of miles driven) in each age group.