Statistics Learning from Data 2nd edition Peck Solutions Manual

SKU: 9781337558082-SOLUTIONS

Complete Solutions Manual
to Accompany

Statistics: Learning from Data

SECOND EDITION

Roxy Peck
California Polytechnic State University, San Luis Obispo

Tom Short
West Chester University of Pennsylvania

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Prepared by

Stephen Miller
Winchester Thurston School, Pittsburgh, PA



Australia • Brazil • Mexico • Singapore • United Kingdom • United States

Statistics Learning from Data 2nd edition Peck Solutions Manual

© 2019 Cengage Learning

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, Web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher except as may be permitted by the license terms below.

For product information and technology assistance, contact us at **Cengage Learning Customer & Sales Support, 1-800-354-9706**.

For permission to use material from this text or product, submit all requests online at **www.cengage.com/permissions**
Further permissions questions can be emailed to **permissionrequest@cengage.com**.

ISBN-13: 978-133755839-6
ISBN-10: 1-33755839-7

Cengage Learning
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at: **www.cengage.com/global**.

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage Learning Solutions, visit **www.cengage.com**.

Purchase any of our products at your local college store or at our preferred online store **www.cengagebrain.com**.

NOTE: UNDER NO CIRCUMSTANCES MAY THIS MATERIAL OR ANY PORTION THEREOF BE SOLD, LICENSED, AUCTIONED, OR OTHERWISE REDISTRIBUTED EXCEPT AS MAY BE PERMITTED BY THE LICENSE TERMS HEREIN.

READ IMPORTANT LICENSE INFORMATION

Dear Professor or Other Supplement Recipient:

Cengage Learning has provided you with this product (the "Supplement") for your review and, to the extent that you adopt the associated textbook for use in connection with your course (the "Course"), you and your students who purchase the textbook may use the Supplement as described below. Cengage Learning has established these use limitations in response to concerns raised by authors, professors, and other users regarding the pedagogical problems stemming from unlimited distribution of Supplements.

Cengage Learning hereby grants you a nontransferable license to use the Supplement in connection with the Course, subject to the following conditions. The Supplement is for your personal, noncommercial use only and may not be reproduced, or distributed, except that portions of the Supplement may be provided to your students in connection with your instruction of the Course, so long as such students are advised that they may not copy or distribute any portion of the Supplement to any third party. Test banks, and other testing materials may be made available in the classroom and collected at the end of each class session, or posted electronically as described herein. Any material posted electronically must be through a password-protected site, with all copy and download functionality disabled, and accessible solely by your students who have purchased the associated textbook for the Course. You may not sell, license, auction, or otherwise redistribute the Supplement in any form. We ask that you take reasonable steps to protect the Supplement from unauthorized use, reproduction, or distribution. Your use of the Supplement indicates your acceptance of the conditions set forth in this Agreement. If you do not accept these conditions, you must return the Supplement unused within 30 days of receipt.

All rights (including without limitation, copyrights, patents, and trade secrets) in the Supplement are and will remain the sole and exclusive property of Cengage Learning and/or its licensors. The Supplement is furnished by Cengage Learning on an "as is" basis without any warranties, express or implied. This Agreement will be governed by and construed pursuant to the laws of the State of New York, without regard to such State's conflict of law rules.

Thank you for your assistance in helping to safeguard the integrity of the content contained in this Supplement. We trust you find the Supplement a useful teaching tool.

Printed in the United States of America
1 2 3 4 5 6 7 17 16 15 14 13

Statistics Learning from Data 2nd edition Peck Solutions Manual

Contents

Chapter 1: Collecting Data in Reasonable Ways 1

Chapter 2: Graphical Methods for Describing Data Distributions 15

Chapter 3: Numerical Methods for Describing Data Distributions 58

Chapter 4: Describing Bivariate Numerical Data 83

Chapter 5: Probability 119

Chapter 6: Random Variables and Probability Distributions 152

Chapter 7: An Overview of Statistical Inference – Learning from Data 200

Chapter 8: Sampling Variability and Sampling Distributions 207

Chapter 9: Estimating a Population Proportion 227

Chapter 10: Asking and Answering Questions About a Population Proportion 263

Chapter 11: Asking and Answering Questions About the Difference Between Two Proportions 304

Chapter 12: Asking and Answering Questions About a Population Mean 355

Chapter 13: Asking and Answering Questions About the Difference Between Two Means 400

Chapter 14: Learning from Categorical Data 536

Chapter 15: Understanding Relationships – Numerical Data 607

Chapter 16: Asking and Answering Questions About More Than Two Means 640

Online Section 4.6: Logistic Regression 667

Online Section 6.9: Counting Rules and the Poisson Distribution 673

Statistics Learning from Data 2nd edition Peck Solutions Manual

Chapter 1 Collecting Data in Reasonable Ways

Section 1.2 Exercise Set 1

1.1: This is an observational study because the person conducting the study merely recorded (based on a survey) whether or not the boomers sleep with their phones within arm's length, and whether or not people ages 50 to 64 used their phones to take photos.

1.2: This is an observational study because the researchers recorded responses to questions on a survey. No men were assigned to different experimental groups.

1.3: This is an experiment because the researchers assigned different toddlers to experimental conditions (adult played with/talked to the robot or the adult ignored the robot).

1.4: This is an observational study because the researchers surveyed U.S. adults and drew a conclusion from the survey results; there were no experimental treatments assigned.

1.5: This is an experiment because the researchers assigned study participants to one of three treatment groups (meditation, distraction task, or relaxation technique).

Section 1.2 Exercise Set 2

1.6: This is an observational study based on results of a survey (no consumers were assigned to different experimental conditions).

1.7: This is an experiment because the participants (college students) were assigned to different experimental conditions (McDonald's Big Mac coupon or Subway 12-inch Italian BMT coupon).

1.8: This is an observational study because the researchers based their conclusions on the results of a survey. There was no assignment to different experimental conditions.

1.9: This is an experiment because the researchers assigned study participants to different experimental conditions (garlic supplement group or no garlic supplement group).

1.10: This is an experiment because the researchers assigned the children to different experimental groups (ball behaving as expected and ball behaving in an unexpected way).

Additional Exercises for Section 1.2

1.11: This is an observational study because the researchers observed the proportion of patients who got an infection in the two groups (overnight hospitalization or more than one night hospitalization); there was no assignment of subjects to experimental groups.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

1.12: This is an experiment because the researcher assigned three of the statistics discussion sections to receive chocolate, and the remaining three did not receive chocolate (the chocolate or lack of chocolate was the experimental group).

1.13: This is an experiment because the study participants were assigned to one of the two experimental groups (how much would you pay for the mug or how much would you sell the mug for).

1.14: The study described was an experiment because the study participants were asked either the first or second question (the participants were assigned to one of the two experimental groups).

Section 1.3 Exercise Set 1

1.15: (a) The group of 716 bicycle fatalities represents a census of the 2008 bicycle fatalities. (b) Because the group of 716 represents a census, the average age of 41 years is a population characteristic.

1.16: The sample is the 100 San Fernando Valley residents, and the population of interest is all San Fernando Valley residents.

1.17: The headline referenced in the previous exercise is not appropriate because the Los Angeles County Metropolitan Transportation Authority only surveyed 100 residents from those who attended a community forum in Van Nuys on a Monday. Residents who could not attend the forum were not included in the survey. The residents who attended the forum are likely those who feel strongly about the issue. As such, it is not appropriate to generalize the survey results to all San Fernando Valley residents. The sample is biased, and we cannot be sure that the results of this survey are representative of the population as a whole. If results from this study are to be reported at all, a more appropriate headline might be "Over two-thirds of those attending a community forum OK with 1-cent transit tax."

1.18: There are several reasonable approaches. One is described here. Using the list of all students at the school, write their names on identical slips of paper. Thoroughly mix the slips of paper, and select 150 slips. The individuals whose names are on the slips of paper constitute the sample.

1.19: (a) The population of interest is all U.S. women. (b) Although the details of the sampling scheme are not presented, the sample size is large (which is generally desirable). However, not all states were represented in the sample; only women from Maryland, Minnesota, Oregon and Pennsylvania were included in the sample. As such, it might be difficult to generalize beyond the population of women in those four states. (c) Given that only women from four states were included in the sample, the sample is not likely to be

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

representative of the population of interest. (d) Selection bias is present because the selection method excluded women from all states other than Maryland, Minnesota, Oregon and Pennsylvania.

Section 1.3 Exercise Set 2

1.20: The given percentages are statistics, because they were computed from sample data obtained by a survey.

1.21: The group of soldiers surveyed represents a census. The reported percentage is a population characteristic, because it was computed using data from the entire population of soldiers.

1.22: (a) This is a convenience sample because there was no random selection of a top medical school or an expert within the school. (b) Because there was no random selection of the infectious disease expert, generalization to all doctors is not reasonable.

1.23: There are several reasonable approaches. One is described here. Write the names of all students enrolled at the college on identical slips of paper. Thoroughly mix the slips of paper and select 100 of the slips. The students whose names are on the 100 slips of paper constitute the simple random sample.

1.24: (a) The population of interest is all Arizona drivers. (b) Although the details of the sampling scheme are not presented, there was random selection of AAA Arizona members. Therefore, the sample was selected in a reasonable way. (c) This sample is not representative of the population of interest because only those people who are AAA Arizona members were included in the sample. (d) Selection bias is present because those drivers who are not AAA Arizona members have been excluded from the sample. Nonresponse bias could also be present because there was no indication of how many surveys were not returned.

Additional Exercises for Section 1.3

1.25: (a) One example of a leading question is "Knowing that there are health problems associated with consuming too much sugar, and that soft drinks contain large amounts of added sugar, should there be a tax on soft drinks to encourage people to consume less?" (b) A double-barreled question is "How satisfied are you with the food and service at this restaurant?" It would be better to split the bad question into two questions, such as "How satisfied are you with the food at this restaurant" and "How satisfied are you with the service at this restaurant."

1.26: The population is all 7000 property owners in this particular rural county. The sample is the 500 property owners selected at random from the 7000 total owners in the county.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

1.27: The population is the 5000 bricks in the lot available at the auction. The sample is the 100 bricks chosen for inspection.

1.28: The chairman does not understand the power of random selection. Random samples tend to reflect the distribution of voters in the population. Although it is possible to obtain a random sample that is not representative of the population, the risk of getting a sample that is not representative of the population does not depend on what fraction of the population is sampled. The random selection process allows us to be confident that the resulting sample will adequately reflect the population, even when the sample consists of only a relatively small fraction of the population.

1.29: Bias introduced through the two different sampling methods may have contributed to the different results. The online sample could suffer from voluntary response bias in that perhaps only those who feel very strongly would take the time to go to the website and register their vote. In addition, younger people might be more technologically savvy, and therefore the website might represent the views of younger people (particularly students) who support the parade. The telephone survey telephone responses might represent the view of permanent residents (as students might only use cell phones and not have a local phone number).

Section 1.4 Exercise Set 1

1.30: Random assignment allows the researcher to create groups that are equivalent, so that the subjects in each experimental group are as much alike as possible. This ensures that the experiment does not favor one experimental condition (playing Unreal Tournament 2004 or Tetris) over another.

1.31: (a) Allowing subject participants to choose which group they want to be in could introduce systematic differences between the two experimental conditions (compression socks group or regular athletic socks group), resulting in potential confounding. Those who would choose to wear compression socks might, in some way, be different from those who would choose the regular athletic socks. We would not know if differences in the measured variables from the blood samples between the two groups were due to the compression socks, or due to some inherent differences in the subjects who chose their experimental groups. (b) It would be good to have the runners be blind to the type of socks they were given to eliminate the possible psychological advantage the runners might have if they knew they were wearing compression socks. That way, we could better determine if the compression socks were the cause of any performance gains.

1.32: (a) The attending nurse was responsible for administering medication after judging the degree of pain and nausea, so the researchers did not want the nurse's personal beliefs about the different surgical procedures to influence measurements. (b) Because the

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

children who had the surgery could easily determine whether the surgical procedure was laparoscopic repair or open repair based on the type of incision.

1.33: There are several possible approaches. One is described here. Write each subject's name on identical slips of paper. Mix the slips of paper thoroughly and draw out slips one at a time. The names on the first 15 slips are assigned to the experimental condition of listening to a Mozart piano sonata for 24 minutes. The names on the next 15 slips are assigned to the experimental condition of listening to popular music for the same length of time. The remaining 15 names are assigned to the relaxation with no music experimental condition.

1.34: (1) Does a dietary supplement consisting of Omega 3, Omega 6, and antioxidants reduce hair loss in women with stage 1 hair loss? (2) The experimental conditions are the supplement and control. (3) The response variable is the change in hair density. (4) The experimental units are the 120 women who volunteered for the study. (5) Yes, the design incorporates random assignment of women to either the supplement group or the control group. (6) Yes, there was a control group. (7) Yes, there was blinding. The expert who determined the change in hair density did not know which of the women had taken the supplement. There is no indication that the women were blinded as to which treatment they received, although this could be incorporated.

Section 1.4 Exercise Set 2

1.35: Random assignment allows the researcher to create groups that are equivalent, so that the subjects in each experimental group are as much alike as possible. This ensures that the experiment does not favor one experimental condition (distraction or no distraction) over another.

1.36: (a) Allowing subject participants to choose which group they want to be in could introduce systematic differences between the two experimental conditions (knee replacement surgery with exercise and exercise therapy alone), resulting in potential confounding. Those who chose knee replacement surgery plus exercise might, in some way, be different from those who chose exercise therapy alone. We would not know if differences in pain relief between the two groups were due to the knee replacement surgery with exercise, or due to some inherent differences in the subjects who chose their experimental groups. (b) Because the purpose of this experiment is to determine whether knee replacement surgery with exercise provided more pain relief than exercise therapy alone, a control group could be used to provide a baseline against which the treatment groups can be compared to determine if the treatments had an effect.

1.37: Blinding of both the dog handlers and experimental observers is important so that the dogs are not intentionally or otherwise guided in determining which patients have cancer. The blinding guarantees that the dogs do not rely on any information other than the patient's breath.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

1.38: There are several possible approaches. One is described here. Write each subject's name on identical slips of paper. Mix the slips of paper thoroughly and draw out slips one at a time. The names on the first 20 slips are assigned to one type of keyboard (experimental condition), and the remaining 20 are assigned to the other type of keyboard (the other experimental condition).

1.39: Was there a control group in which there were identical sheets of paper with no words written on them? Was there any random assignment of experimental units to treatment groups? How were the experimental units selected? How many water bottles were used in the study? Were the water bottles identical? How many bottles of water were used? What measurements were made on the water? Were measurements made both before and after the words were applied to the bottle? Who took the measurements? Was the person taking the measurements blinded to the presence or absence of words on the pieces of paper?

Additional Exercises for Section 1.4

1.40: The experimental conditions were the presence or absence of music with a vocal component. The response is the time required to complete the surgical procedure.

1.41: (a) Some surgical procedures are more complex and require a greater degree of concentration; music with a vocal component might be more distracting when the surgical procedure is more complex. (b) The temperature of the room might affect the comfort of the surgeon; if the surgeon is too hot or too cold, she or he might be uncomfortable, and therefore more easily distracted by the vocal component. (c) If the music is too loud, the surgeon might be distracted and unable to focus, regardless of the presence or absence of the vocal component. If the music is too soft, the surgeon might try to concentrate on listening to the vocal component, and therefore pay more attention to the music rather than the surgical procedure.

1.42: Random assignment of surgeons to music condition is important because there might be something inherently different about surgeons who want no vocals versus those who do want vocals. Random assignment ensures that the experiment does not favor one experimental condition over another.

1.43: This experiment could not have been double-blind because the surgeon would know whether or not there was a vocal component to the music.

1.44: Yes, the random assignment of subjects to experimental groups has been successful in creating groups that are similar in age. Both the LR and OR groups have similar maximum ages, and the LR group does have a few children with slightly lower ages than the OR group. Overall, however, the LR and OR groups are quite similar with respect to ages.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

1.45: (a) Probably not, because the judges might not believe that Denny's food is as good as other restaurants. (b) Experiments are often blinded in this way to eliminate preconceptions about particular experimental treatments.

Section 1.5 Exercise Set 1

1.46: (a) This was most likely an observational study. (b) It is not reasonable to conclude that pushing a shopping cart causes people to be less likely to purchase junk food because the results of observational studies cannot be used to draw cause-and-effect conclusions.

1.47: (a) It is not reasonable to conclude that owning a dog causes higher heart rate variability. This was an observational study, so cause-and-effect conclusions cannot be drawn. (b) It is not reasonable to generalize the results of this survey to all adult Americans because the study participants were not randomly selected from the population of all adult Americans.

1.48: The researcher would have had to assign the nine cyclists at random to one of the three experimental conditions (chocolate milk, Gatorade, or Endurox).

1.49:

Study 1:

Question 1: This is an observational study.

Question 2: Yes, random selection was used.

Question 3: No, this was not an experiment so there were no experimental groups.

Question 5: It is reasonable to generalize to the population of students at this particular large college.

Study 2:

Question 1: This study was an experiment.

Question 2: Random selection was not used.

Question 3: There was no random assignment to experimental conditions (the grouping was based on gender).

Question 4: No, the conclusion is not appropriate because of confounding of gender and treatment (women ate pecans, and men did not eat pecans).

Question 5: It is not reasonable to generalize to a larger population.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

Study 3:

Question 1: This is an observational study.

Question 2: There was no random selection.

Question 3: There was no random assignment to experimental groups.

Question 4: No, the conclusion is not appropriate because this was an observational study, and therefore cause-and-effect conclusions cannot be drawn.

Question 5: We cannot generalize to any larger population beyond the 200 volunteers.

Study 4:

Question 1: This is an experiment.

Question 2: There was no random selection from some population.

Question 3: Yes, there was random assignment to experimental groups.

Question 4: Yes, because this was a simple comparative experiment with random assignment of subjects to experimental groups. We can draw cause-and-effect conclusions.

Question 5: We cannot generalize to a larger population because there was no random selection from some population.

Study 5:

Question 1: This is an experiment.

Question 2: Yes, there was random selection from students enrolled at a large college.

Question 3: Yes, random assignment of subjects to experimental groups was used.

Question 4: Yes, because this was a simple comparative experiment with random assignment of subjects to experimental groups. We can draw cause-and-effect conclusions.

Question 5: Due to the random selection of students, we can generalize conclusions from this study to the population of all students enrolled at the large college.

Section 1.5 Exercise Set 2

1.50: (a) Random selection from the population of affluent Americans is required. (b) No, because the population sampled from was affluent Americans.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

1.51: In order to determine if the conclusions implied by the headline are appropriate, I would need to know if dieters were randomly assigned to the experimental conditions (large fork or small fork). In order to generalize to the population of dieters, I would also want to know if the study participants were randomly selected from the population of dieters.

1.52: Random assignment ensures that the experiment does not favor one experimental condition (talking on the phone, not talking on the phone) over another. If the person crossing the virtual street was on the phone the first 10 crossings, and not on the phone the last 10 crossings, we wouldn't know if any difference between the treatments was due to the phone use or due to the person being either more or less aware of the surroundings for the last 10 crossings, for example.

1.53:

Study 1:

Question 1: This is an observational study.

Question 2: No, there was no random selection from a population.

Question 3: No, there was no random assignment to experimental groups.

Question 4: No, the conclusion that you can "strengthen your marriage with prayer" is not appropriate. There was no experiment conducted, so a cause-and-effect conclusion cannot be drawn.

Question 5: No, it is not reasonable to generalize conclusions from this study to some larger population because this was a voluntary response sample.

Study 2:

Question 1: This is an observational study.

Question 2: Yes, there was random selection from the population of AAUW members.

Question 3: There was no random assignment to experimental groups (this is not an experiment).

Question 4: No, the conclusion that you can "strengthen your marriage with prayer" is not appropriate. There was no experiment conducted, so a cause-and-effect conclusion cannot be drawn.

Question 5: Due to random selection, it is reasonable to generalize the conclusions from this study to the population of AAUW members.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Study 3:

Question 1: This was an observational study.

Question 2: No, there was no random selection from a population.

Question 3: No, there was no random assignment to experimental groups (this was an observational study, not an experiment).

Question 4: No, the conclusion that you can "strengthen your marriage with prayer" is not appropriate. Since this was an observational study, a cause-and-effect conclusion cannot be drawn.

Question 5: It is not reasonable to generalize conclusions from this study to a larger population because there was no random selection of study participants.

Additional Exercises for Section 1.5

1.54: It is not reasonable to conclude that being raised with two or more animals is the cause of the observed lower allergy rate. This was an observational study, so cause-and-effect conclusions cannot be drawn.

1.55: It might be that people who live in the South have a less healthy diet and exercise less than those in other parts of the country. As a result, the higher percentage of Southerners with high blood pressure might have nothing to do with living in the South.

1.56: This is an experiment.

1.57: There was no random selection from some population.

1.58: Yes, there was random assignment to experimental groups (portrait orientation or landscape orientation).

1.59: Yes, it is reasonable to draw the conclusion that reasoning using information displayed on a small screen is improved by turning the screen to landscape orientation because this was an experiment in which there was random assignment of subjects to experimental groups.

1.60: No, it is not reasonable to generalize the conclusions from this study to some larger population because there was no random selection of study participants from a population.

Chapter 1: Are You Ready to Move On? Chapter 1 Review Exercises

1.61: (a) This is an experiment due to the random assignment of subjects to experimental conditions (the five different rooms). (b) This is an observational study because there was no random assignment of subjects to experimental conditions; the researchers merely recorded what they observed on the MySpace pages. (c) This is an observational study

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

because there was no random assignment of subjects to experimental conditions; the researchers merely recorded the responses of the survey participants. (d) This is an experiment because of the random assignment of study participants (the adults with back pain) to experimental conditions (the four different treatments).

1.62: The population of interest is the 15,000 students at the college. The 200 students who were interviewed constitute the sample.

1.63: (a) 84% is a population characteristic. (b) 24.1 years is a statistic. (c) 22% is a population characteristic. (d) 6.4 days is a statistic. (e) 63 hours is a statistic.

1.64: (1) The study participants were volunteers and were not randomly selected. (2) The study participants were all students at Texas Women's University. (3) The study participants were all women (because they are students at a Women's university).

1.65: The council president should assign a unique identifying number to each of the names on the petition, numbered from 1 to 500. On identical slips of paper, write the numbers 1 to 500, with each number on a single slip of paper. Thoroughly mix the slips of paper and select 30 numbers. The 30 numbers correspond to the unique numbers assigned to names on the petition. These 30 names constitute the sample.

1.66: (a) (1) The patients are the population of interest. (2) The study description indicates no random selection of participants, so it does not appear as if the sample was selected in a reasonable way. (3) No, the sample is not likely to be representative of the population of interest. The sample consisted of only undergraduate students, so even if there was random selection of participants, the study results could not be generalized to the population of all patients. (4) It is likely that this study design is affected by selection bias because only undergraduate students were included in the study, thus systematically excluding all non-undergraduate students from the population of interest. (b) No, the stated conclusions are not reasonable because there was no random selection of study participants, and the study suffers from selection bias.

1.67: Without random assignment of the study participants to experimental condition, confounding could impact the conclusions of the study. For example, people who would choose an attractive avatar might be more outgoing and willing to engage than someone who would choose an unattractive avatar.

1.68: (a) Yes, by randomly selecting the 852 children to be in one experimental group (the book group), the remaining children, by default, are in the control group. (b) The control group allows the experimenter to assess how the response variable behaves when the treatment is not used. This provides a baseline against which the treatment groups can be compared to determine if the treatment has an effect. In this case, the researcher can determine whether

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

children given the reading books have better school performance, as measured by a reading test.

1.69: (a) It seems as if the alternate assignment to the experimental groups (large serving bowls, small serving bowls) would tend to produce groups that are similar. People who arrive to the party at approximately the same time might, in some way, be similar to each other, so dividing them into the different experimental groups as described would tend to make the two groups similar to each other. (b) Blinding ensures that individuals do not let personal beliefs influence their measurements. The research assistant that weighed the plates and estimated the calorie content of the food might (intentionally or not) have let her or his personal beliefs influence the estimate of the calorie content of the food on the plate.

1.70: There are several possible approaches. One is described here. Write each subject's name on identical slips of paper. Mix the slips of paper thoroughly and draw out slips one at a time. The names on the first 10 slips are assigned to the first hand drying method. The names on the next 10 slips are assigned to the second hand drying method. The remaining 10 names are assigned to the third hand drying method.

1.71: (a) (1) The experiment is designed to answer the question "Does using hand gestures help children learn math?" (2) The two experimental conditions are using hand gestures and not using hand gestures. (3) The response variable is the number correct on the six-problem test. (4) The experimental units are the 128 children in the study; they were selected because they were the children who answered all six questions on the pretest incorrectly. (5) Yes, the children were assigned randomly to one of the two experimental groups. (6) Yes, the control group is the experimental condition of not using any hand gestures. (7) There was no blinding and, indeed, it would not be possible to include blinding of subjects in this experiment (the children would know whether or not they were using hand gestures), and there is no need to blind the person recording the response because the test was graded with each answer correct or incorrect, so there is no subjectivity in recording the responses. (b) It seems as if the conclusions are reasonable because the subjects were assigned to the treatment groups at random.

1.72: (a) Yes, it is reasonable to generalize the stated conclusion to all 18-year-olds with a publically accessible MySpace web profile because the profiles were selected at random from all MySpace web profiles of 18-year-olds. (b) No, it is not reasonable to generalize the stated conclusion to all 18-year-old MySpace users because those users without publically accessible profiles were not included in the random selection process. (c) No, it is not reasonable to generalize the stated conclusion to all MySpace users because the study only included 18-year-old MySpace users.

1.73: (a) No, the 60 games selected were the 20 most popular (by sales) for each of three different gaming systems. The study excluded the games that were not in the top 20 most

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

popular (by sales). (b) It is not reasonable to generalize the researcher's conclusions to all video games due to selection bias (there was a systematic exclusion of those games not in the top 20 most popular (by sales)).

1.74: (a) The study described is not an experiment because there were no experimental conditions to which study participants were randomly assigned. (b) No, it is not reasonable to conclude that physical activity is the cause of the observed difference in body fat percentage. This was an observational study, and cause-and-effect conclusions cannot be drawn.

1.75:

Study 1:

Question 1: The study described is an observational study.

Question 2: No, there was no random selection from a population.

Question 3: No, there was no random assignment to experimental groups.

Question 4: No, it is not reasonable to conclude that taking calcium supplements is the cause of the increased heart attack risk.

Question 5: No, it is not reasonable to generalize conclusions from this study to a larger population because there was no random selection from a larger population.

Study 2:

Question 1: The study described is an observational study.

Question 2: Yes, there was random selection from the population of people living in Minneapolis who receive Social Security.

Question 3: No, there was no random assignment of subjects to experimental groups.

Question 4: No, it is not reasonable to conclude that taking calcium supplements is the cause of the increased heart attack risk.

Question 5: Yes, it is reasonable to generalize the results of this study to the population of people living in Minneapolis who receive Social Security.

Study 3:

Question 1: The study described is an experiment.

Question 2: Yes, there was random selection from the population of people living in Minneapolis who receive Social Security.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

Question 3: No, there was no random assignment of subjects to experimental groups.

Question 4: No, it is not reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack due to confounding and the lack of random assignment of subjects to experimental conditions. The participants in this study who did not have a previous history of heart problems were given the calcium supplement, and those with a history of heart problems were not given the supplement. It is not possible to determine the role of the calcium supplement because only those study participants who did not have a history of heart problems were given the supplement.

Question 5: It is possible to generalize the results from this study to the population of all people living in Minneapolis who receive Social Security. However, it is unclear (due to the confounding described in Question (4) what the conclusion would be.

Study 4:

Question 1: The study described is an experiment because there was random assignment of subjects to experimental conditions.

Question 2: No, there was no random selection from some larger population.

Question 3: Yes, there was random assignment of study participants to experimental groups.

Question 4: Yes, it is reasonable to conclude that taking calcium supplements is the cause of the increased risk of heart attack.

Question 5: No, it is not reasonable to generalize conclusions from this study to some larger population because of the lack of random selection of study participants from a population.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

Chapter 2 Graphical Methods for Describing Data Distributions

Section 2.1 Exercise Set 1

2.1: (a) numerical, discrete (b) categorical (c) numerical, continuous (d) numerical, continuous (e) categorical

2.2: (a) discrete (b) continuous (c) discrete (d) discrete

2.3:

Data Set 1:

Question 1: There is one variable in the data set.

Question 2: The variable is categorical.

Question 3: The purpose of the graphical display is to summarize the data distribution.

Appropriate Graphical Display: Bar chart

Data Set 2:

Question 1: There is one variable in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of the graphical display is to compare groups (full-time students or part-time students).

Appropriate Graphical Display: Comparative dotplot, comparative stem-and-leaf displays, and comparative histograms are all appropriate.

Data Set 3:

Question 1: There are two variables in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of the graphical display is to investigate the relationship between two numerical variables.

Appropriate Graphical Display: Scatterplot

Data Set 4:

Question 1: There is one variable in the data set.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

Question 2: The variable is categorical.

Question 3: The purpose of a graphical display is to compare groups (faculty, students)

Appropriate Graphical Display: Comparative bar chart

Data Set 5:

Question 1: There is one variable in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of a graphical display is to summarize a data distribution.

Appropriate Graphical Display: Dotplot, stem-and-leaf, and histogram are all appropriate graphical displays.

Section 2.1 Exercise Set 2

2.4: (a) categorical

(b) categorical

(c) numerical – discrete

(d) numerical – continuous

(e) categorical

(f) numerical – continuous

2.5: (a) continuous (b) continuous (c) continuous (d) discrete

2.6: **Data Set 1:**

Question 1: There is one variable in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of a graphical display is to summarize a data distribution.

Appropriate Graphical Display: Dotplot, stem-and-leaf, and histogram are all appropriate graphical displays.

Data Set 2:

Question 1: There is one variable in the data set.

Question 2: The variable is categorical.

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

Question 3: The purpose of a graphical display is to compare two groups (male and female color choice).

Appropriate Graphical Display: Comparative bar chart

Data Set 3:

Question 1: There are two variables in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of a graphical display is to investigate the relationship between two numerical variables.

Appropriate Graphical Display: Scatterplot

Data Set 4:

Question 1: There is one variable in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of a graphical display is to compare two groups.

Appropriate Graphical Display: Comparative dotplot, comparative stem-and-leaf displays, and comparative histograms are all appropriate.

Data Set 5:

Question 1: There is one variable in the data set.

Question 2: The variable is categorical.

Question 3: The purpose of a graphical display is to summarize a data distribution.

Appropriate Graphical Display: Bar chart.

Additional Exercises for Section 2.1

2.7: (a) numerical (b) numerical (c) categorical (d) numerical (e) categorical

2.8: Discrete: (b); Continuous: (a) and (d)

2.9: (a) categorical (b) numerical (c) numerical (d) categorical

2.10: Discrete: (b); Continuous: (c)

2.11: (a) numerical (b) numerical (c) numerical (d) categorical (e) categorical (f) numerical (g) categorical

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

2.12:

Question 1: There is one variable in the data set.

Question 2: The variable is categorical.

Question 3: The purpose is to summarize the data distribution.

Appropriate graphical display: Bar chart

2.13:

Question 1: There is one variable in the data set.

Question 2: The variable is numerical.

Question 3: The purpose is to compare groups (male students, female students).

Appropriate graphical display: Comparative dotplot, comparative stem-and-leaf, or histograms are all appropriate.

2.14:

Question 1: There are two variables in the data set.

Question 2: The variable is numerical.

Question 3: The purpose is to investigate the relationship between two numerical variables.

Appropriate graphical display: Scatterplot

2.15:

Question 1: There is one variable in the data set.

Question 2: The variable is numerical.

Question 3: The purpose of a graphical display is to summarize the data distribution.

Appropriate graphical display: Dotplots, stem-and-leaf plots, and histograms are all appropriate graphical displays.

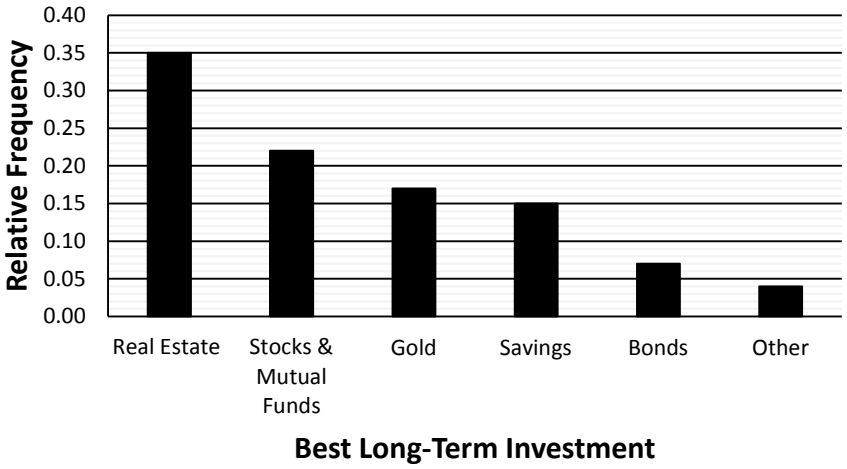
© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

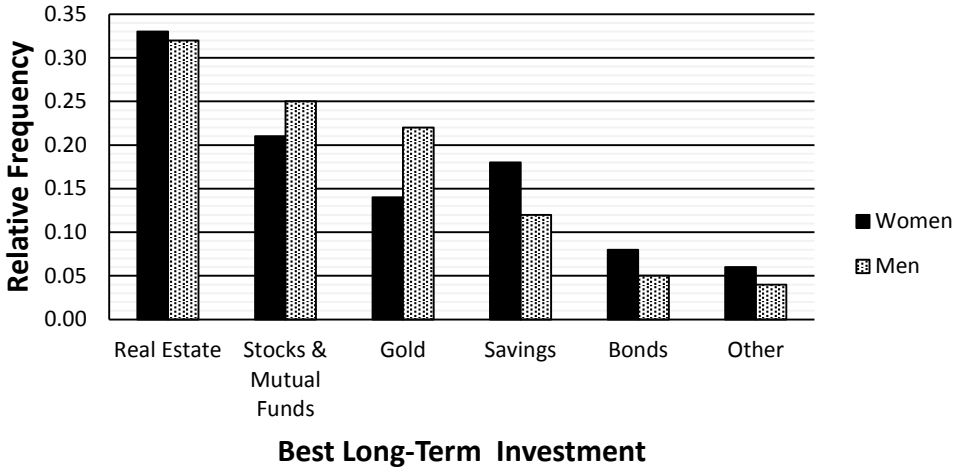
Section 2.2 Exercise Set 1

2.16: (a)



(b) Over half of the responses (57%) were from people who indicated that the best long-term investments were real estate (35%) and stocks & mutual funds (22%). The remaining 43% of respondents indicated that gold (17%), savings (15%), bonds (7%), and other (4%) were the best long-term investments.

2.17: (a)



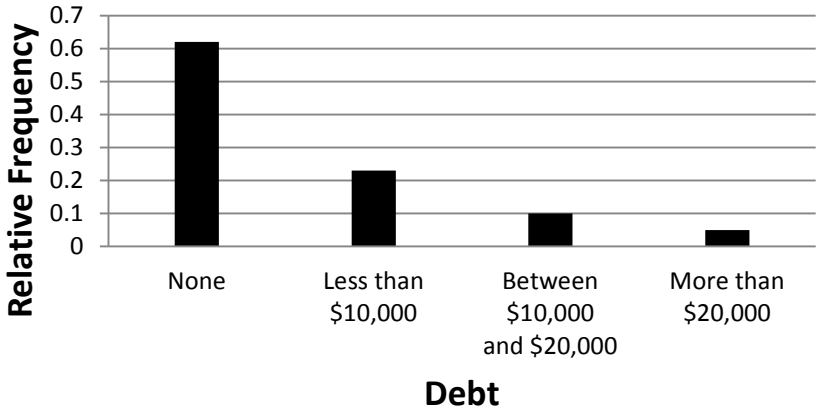
(b) In general, women and men have similar rankings for the best long-term investments. However, one notable difference is that women rank saving higher than gold, and men rank gold higher than savings.

Statistics Learning from Data 2nd edition Peck Solutions Manual

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Section 2.2 Exercise Set 2

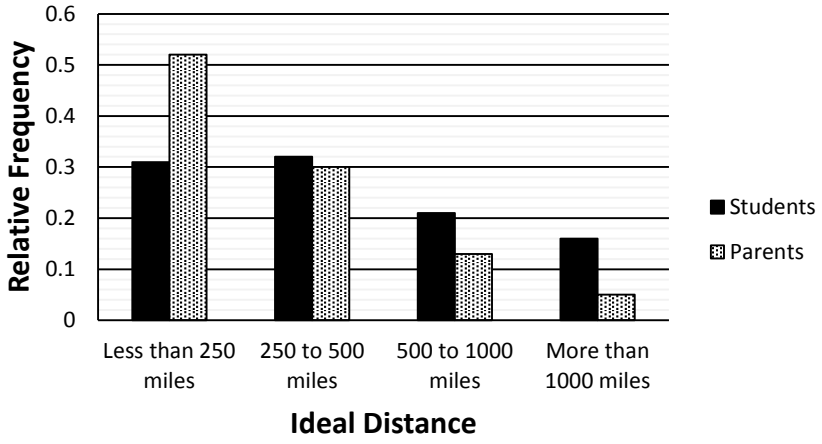
2.18: (a)



(b) Over 60% of students graduating with an AA degree from a public community college in 2008 graduate with no debt. As the amount of debt increases, fewer students reach that debt level. Twenty-three percent of students have less than $10,000 in debt, 10% have between $10,000 and $20,000 in debt, and only 5% have over $20,000 in debt.

2.19: (a) One would want to use relative frequencies when constructing a comparative bar chart to compare ideal distance for students and parents because the number of students (8,347) and number of parents (2,087) are not equal to each other.

(b)



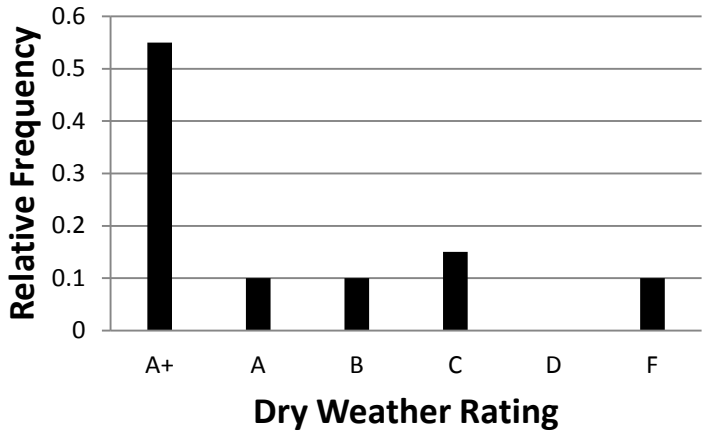
(c) Students tend to want to be farther away from home than their parents would like them to be. In fact, 37% of students want to be over 500 miles from home, whereas only 18% of parents want their children to be over 500 miles from home.

Statistics Learning from Data 2nd edition Peck Solutions Manual

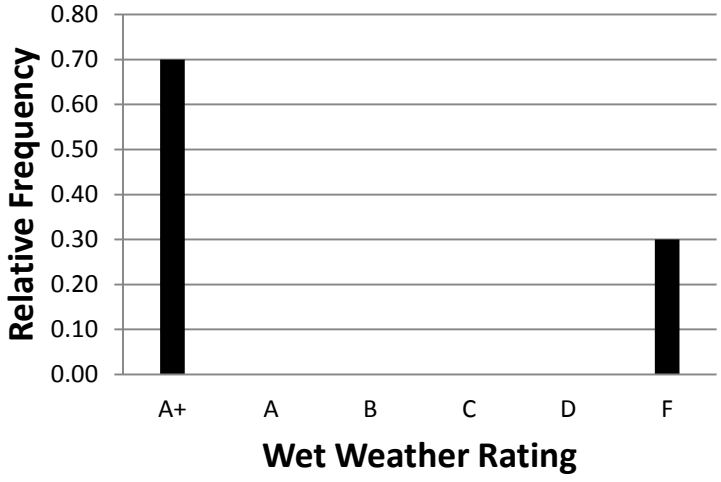
Additional Exercises for Section 2.2

2.20: (a)

Rating	Relative Frequency
A+	11/20 = 0.55
A	2/20 = 0.10
B	2/20 = 0.10
C	3/20 = 0.15
D	0/20 = 0.00
F	2/20 = 0.10



(b)



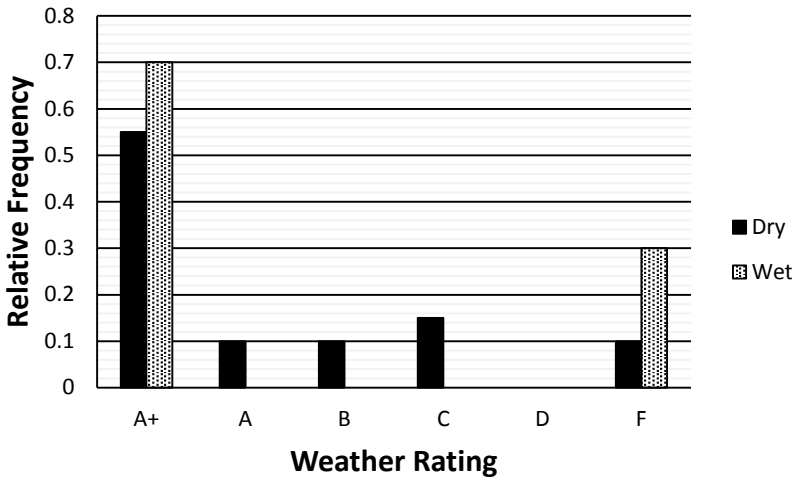
(c) Seventy-five percent (75%) of the dry weather ratings are B or higher, and 70% of wet weather ratings are B or higher, indicating that dry weather ratings are higher than wet weather ratings. Note that the wet weather ratings are only A+ or F, so the wet weather

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

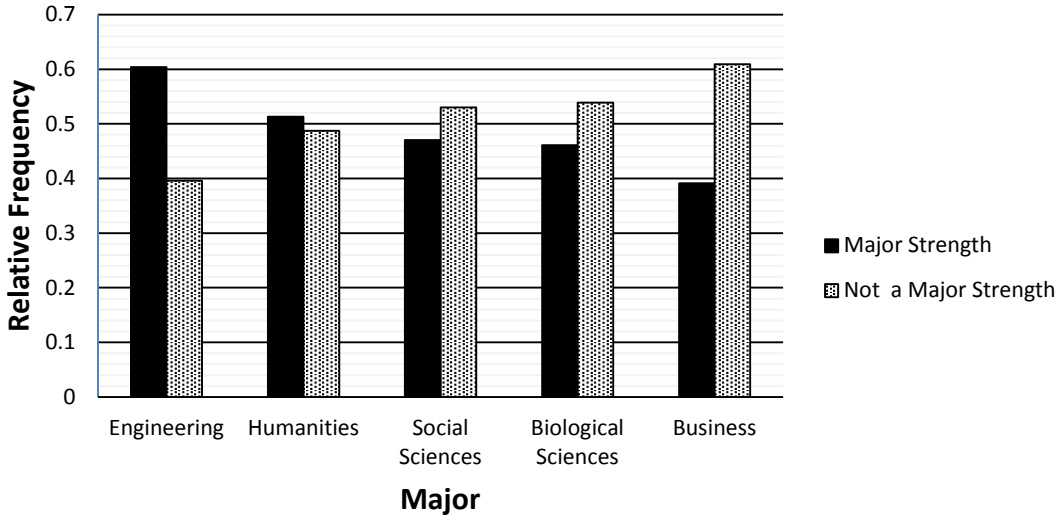
ratings are more extreme than dry weather ratings. If we only consider A+ ratings, then the wet weather ratings tend to be better than dry weather ratings because only 55% of dry weather ratings are A+, compared with 70% of wet weather ratings being A+.

(d)



The wet weather ratings are more extreme than the dry weather ratings (only ratings of A+ and F are found in the wet weather ratings). Dry weather ratings include all rating levels except for D.

2.21:



© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

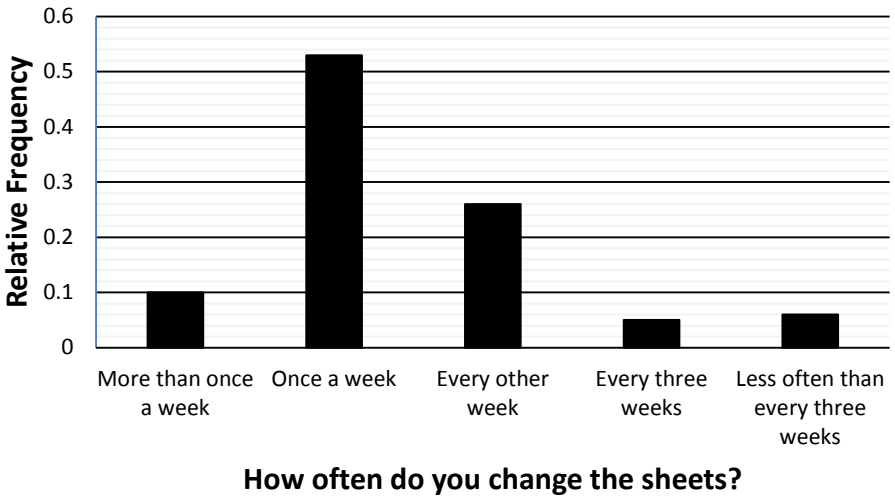
Statistics Learning from Data 2nd edition Peck Solutions Manual

The majors in which a majority of students indicated that critical thinking is "a major strength" are engineering and humanities. A majority of students in other majors (social sciences, biological sciences, and business) indicated that critical thinking is "not a major strength." The greatest disparities between "a major strength" and "not a major strength" occurred with engineering and business majors, and humanities majors had the smallest difference between the two responses.

2.22: (a)

How Often	Relative Frequency
More than once a week	0.10
Once a week	0.53
Every other week	0.26
Every three weeks	0.05
Less often than every three weeks	0.06

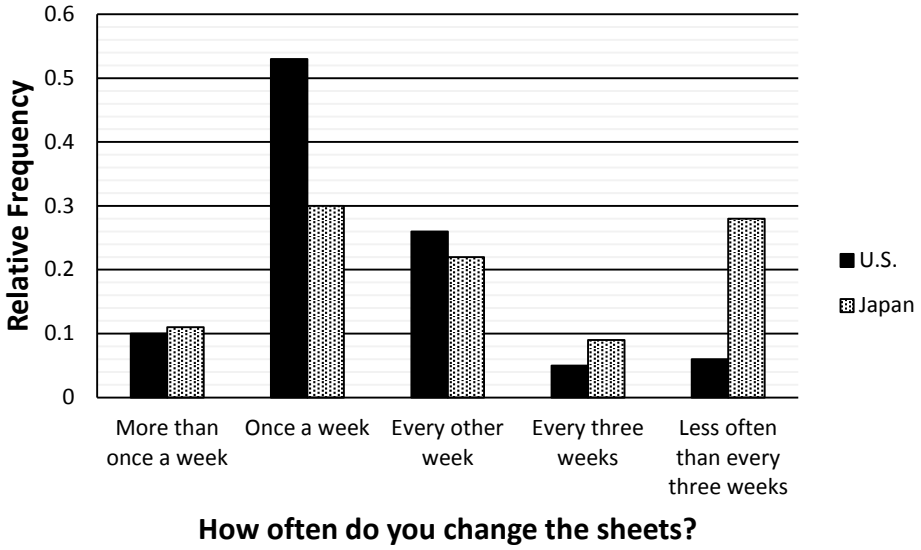
(b)



© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

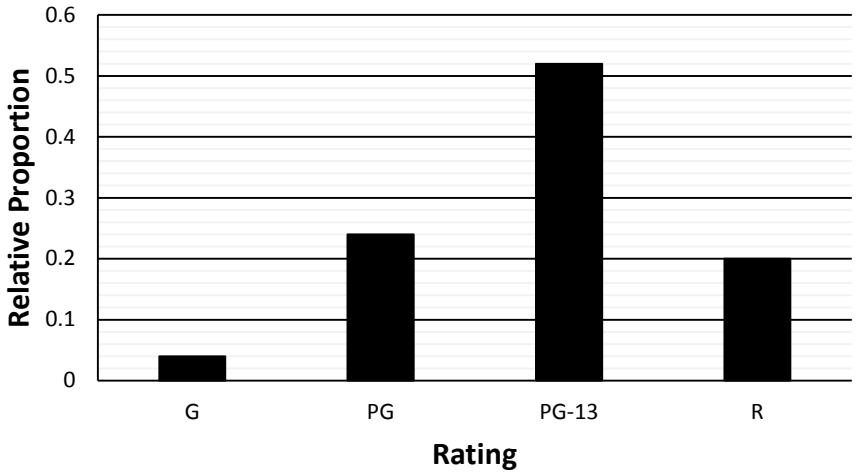
(c)



The biggest differences between the sheet-changing habits of Americans and Japanese are in the "once a week" and the "less often than every three weeks" categories. A much larger percentage of Americans change their sheets once a week than do the Japanese, and a much larger percentage of Japanese change their sheets less often than every three weeks. The remaining categories are fairly close to each other.

2.23: The relative frequency distribution is:

Rating	Relative Frequency
G	1/25 = 0.04
PG	6/25 = 0.24
PG-13	13/25 = 0.52
R	5/25 = 0.20

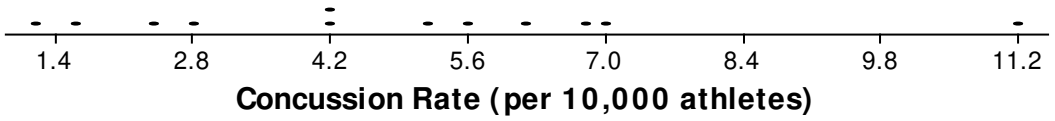


© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

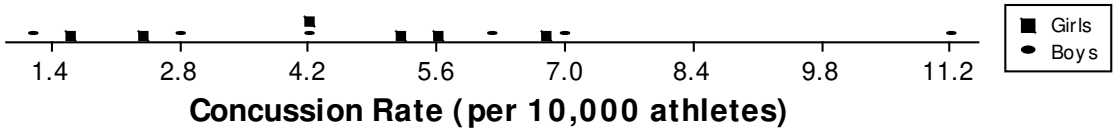
PG-13 is the rating with the highest relative proportion (0.52), followed by PG (0.24), R (0.20), and G (0.04). Seventy-two percent (72%) of the top 25 movies of 2015 are PG-13 or R, and the remaining 28% are rated G or PG.

Section 2.3 Exercise Set 1

2.24: (a) The dotplot below shows the concussion rate (concussions per 10,000 athletes).

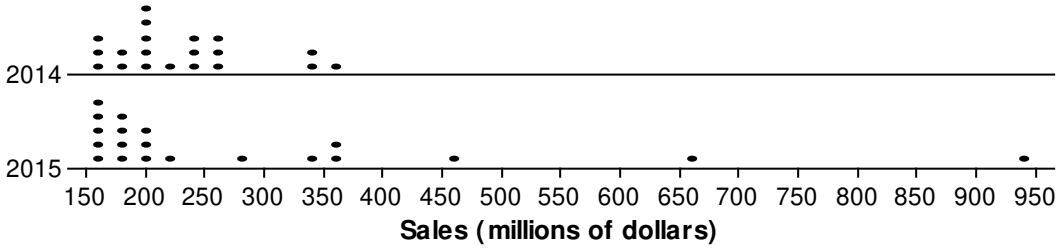


(b) The dotplot below shows the concussion rate (concussions per 10,000 athletes), with different symbols for boys and girls.



The sport with an unusually high (compared to all the other sports) concussion rate is football. Without considering football, the concussion rates for girls' sports is essentially the same as the concussion rate for boys' sports. However, if we consider football, the concussion rate for girls' sports tends to be lower than that for boys' sports.

2.25: (a)



(b) The shape of both distributions is positively skewed. The distribution of 2015 ticket sales is centered at about $280 million, which is higher than the center of the distribution of 2014 ticket sales, which is centered around $230 million. The lowest ticket sales for both 2014 and 2015 are approximately $150 million. Ticket sales for 2014 has a maximum value of approximately $350 million, which is much lower than the highest ticket sales for 2015 of $937 million. The spread between the lowest and highest values for 2014 is

© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

Statistics Learning from Data 2nd edition Peck Solutions Manual

approximately $200 million, which is less than the spread for 2015 of a little under $800 million.

2.26:

```
28 | 8
29 |
30 |
31 |
32 | 8
33 | 0
34 | 178
35 | 00145678899
36 | 238999
37 | 0034566777
38 | 01124558
39 | 00259
40 | 045
41 | 2
42 | 2

Legend: 34|1 = 34.1 years
```

The distribution of median ages is centered at approximately 37 years old, with values ranging from 28.8 to 42.2 years. The distribution is approximately symmetric (possibly slightly negatively skewed), with one possible outlier of 28.8 years.

2.27: (a)

Very Large Urban Area		Large Urban Area
	1	023478
	2	369
8	3	0033589
99	4	0366
8711	5	012355
9730	6	
2	7	
	8	
3	9	

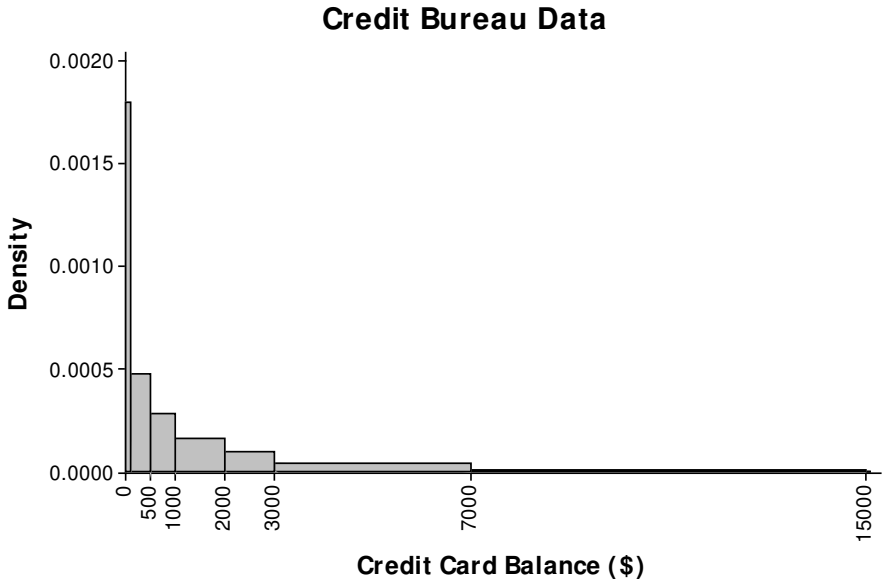
Legend: 4|6 = 46 extra hours per year

(b) The statement "The larger the urban areas, the greater the extra travel time during peak period travel" is generally consistent with the data. Although there is overlap between the times for the very large and large urban areas, the back-to-back stem-and-leaf plot shows that for the very large urban areas, the extra travel time during peak period is generally

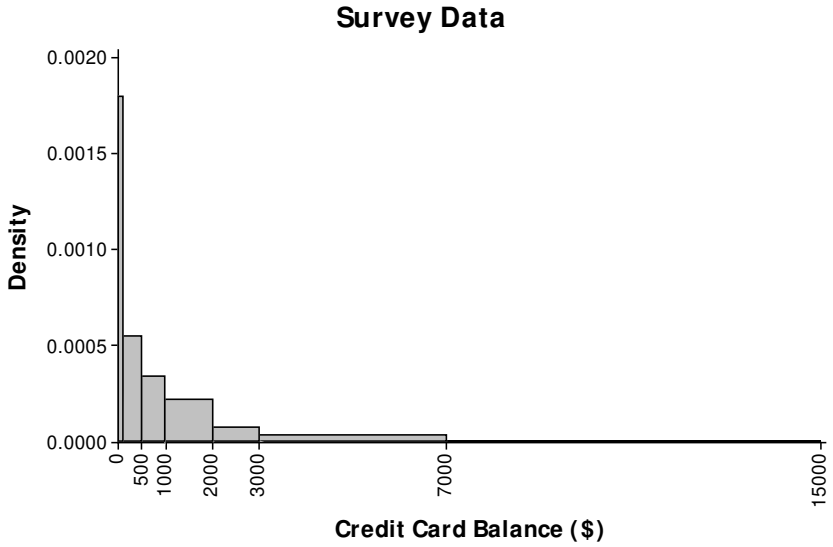
© Cengage Learning. All rights reserved. No distribution allowed without express authorization.

longer than for the large urban areas. The extra travel time during peak period is centered at approximately 58 hours for the very large urban areas, which is higher than the center of approximately 34 hours for the large urban areas.

2.28: (a)



(b)



© Cengage Learning. All rights reserved. No distribution allowed without express authorization.