

ONLINE TEST BANK AND RESOURCE GUIDE

JARED DERKSEN

INTRO STATS FIFTH EDITION

Richard De Veaux

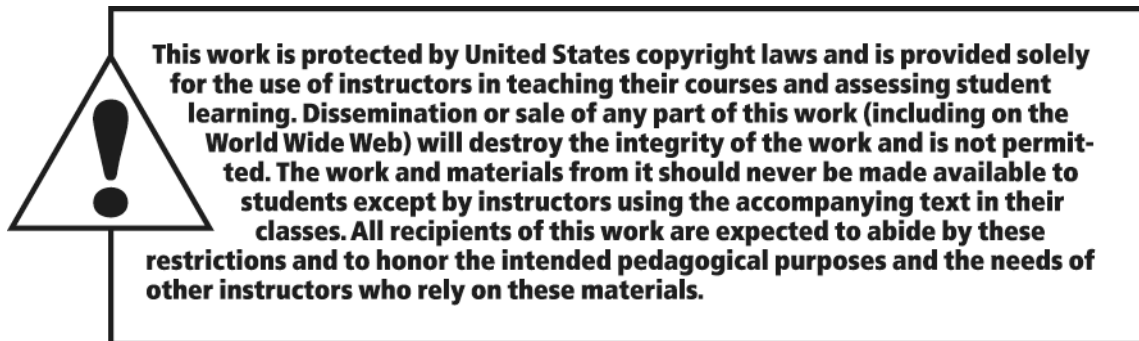
Williams College

Paul Velleman

Cornell University

David Bock

Cornell University



The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, expressed or implied, with regard to these programs or the documentation contained in this book. The author and publisher shall not be liable in any event for incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of these programs.

Reproduced by Pearson from electronic files supplied by the author.

Copyright © 2018, 2014, 2009 Pearson Education, Inc.
Publishing as Pearson, 330 Hudson Street, NY NY 10013

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America.



ISBN-13: 978-0-13-426540-7
ISBN-10: 0-13-426540-8

Contents

Part I	Exploring and Understanding Data	
Chapter 1	Stats Starts Here	1-1
Chapter 2	Displaying and Describing Data	2-1
Chapter 3	Relationships Between Categorical Variables—Contingency Tables	3-1
Chapter 4	Understanding and Comparing Distributions	4-1
Chapter 5	The Standard Deviation as a Ruler and the Normal Model	5-1
	Review of Part I: Exploring and Understanding Data	Part I-1
Part II	Exploring Relationships Between Variables	
Chapter 6	Scatterplots, Association, and Correlation	6-1
Chapter 7	Linear Regression	7-1
Chapter 8	Regression Wisdom	8-1
Chapter 9	Multiple Regression	9-1
	Review of Part II: Exploring Relationships Between Variables	Part II-1
Part III	Gathering Data	
Chapter 10	Sample Surveys	10-1
Chapter 11	Experiments and Observational Studies	11-1
	Review of Part III: Gathering Data	Part III-1
Part IV	From the Data at Hand to the World at Large	
Chapter 12	From Randomness to Probability	12-1
Chapter 13	Sampling Distribution Models and Confidence Intervals for Proportions	13-1
Chapter 14	Confidence Intervals for Means	14-1
Chapter 15	Testing Hypotheses	15-1
Chapter 16	More About Tests and Intervals	16-1
	Review of Part IV: From the Data at Hand to the World at Large	Part IV-1
Part V	Inference for Relationships	
Chapter 17	Comparing Groups	17-1
Chapter 18	Paired Samples and Blocks	18-1
Chapter 19	Comparing Counts	19-1
Chapter 20	Inferences for Regression	20-1
	Review of Part V: Inference for Relationships	Part V-1

How to use this *Test Bank and Resource Guide*

This guide is a supplement to be used in conjunction with the Instructor's Edition of *Intro Stats*, 5th edition by De Veaux, Velleman, and Bock. The authors have integrated many instructor's resources into the text, and these sections precede each chapter. In this *Test Bank and Resource Guide*, all or some of the following features may be found for each chapter and unit.

Solutions to Class Examples

Answers are provided to the chapter examples presented in the Instructor's Edition of the text.

Investigative Tasks

Instead of a quiz, you may choose to have students do a written assignment that applies the major concepts of the chapter. Along with each classroom-tested task, there is a proposed solution to the task and a scoring rubric. Returning the completed rubric to the students will provide them the guidance needed to learn to write clear, complete, and concise statistical analyses.

Chapter Quizzes

You might choose to give a quiz after completing a chapter. For each chapter, there are two or three quizzes that you can choose from, along with solutions. If not used as a quiz, the questions can be used as additional class examples, homework assignments, or extra practice.

Unit Tests

Two or three sample exams (and solutions) are available for you at the end of each of the text's five units. These exams include multiple-choice questions, short questions requiring some calculations or written explanations, and longer questions requiring more in-depth analysis. They are not easy. Understanding Statistics means thinking about the world. All of the problems ask for clear understanding of important statistical concepts, accurate application of statistical techniques, and proper interpretation of the results. Expecting this from the start helps students establish the habit of clear statistical thinking.

Supplemental Resources

We've tried lots of things over the years to help students understand the beauty and power of Statistics. Where applicable, we've included some extra materials. These might be worksheets, group assignments, or class activities.

Chapter 1

Stats Starts Here

Solutions to Class Examples:

Consumer Reports

Who: energy bars

What: brand name, flavor, price, calories, protein, fat

When: not specified

Where: not specified

How: not specified. Are data collected from the label? Are independent tests performed?

Why: information for potential consumers

Categorical variables: brand name, flavor

Quantitative variables: price (US\$), number of calories (calories), protein (grams), fat(grams)

Boston Marathon

Who: Boston Marathon runners

What: gender, country, age, time

When: not specified

Where: Boston

How: not specified. Presumably, the data were collected from registration information.

Why: race result reporting

Categorical variables: gender, country

Quantitative variables: age (years), time (hours, minutes, seconds)

Supplemental Resources:

The following page contains a list of the 50 United States of America. We have found it to be helpful if you collect class data on the number of States visited. On the next page is a potential blank survey that you can pass around on the first day of class to collect some data. Some of the survey questions are left deliberately vague, so that you can discuss potential sources of bias, informally of course.

1-2 Part I Exploring and Understanding Data

States – Count the number you have visited

Alabama	Indiana	Nebraska	Rhode Island
Alaska	Iowa	Nevada	South Carolina
Arizona	Kansas	New Hampshire	South Dakota
Arkansas	Kentucky	New Jersey	Tennessee
California	Louisiana	New Mexico	Texas
Colorado	Maine	New York	Utah
Connecticut	Maryland	North Carolina	Vermont
Delaware	Massachusetts	North Dakota	Virginia
Florida	Michigan	Ohio	Washington
Georgia	Minnesota	Oklahoma	West Virginia
Hawaii	Mississippi	Oregon	Wisconsin
Idaho	Missouri	Pennsylvania	Wyoming
Illinois	Montana		

States – Count the number you have visited

Alabama	Indiana	Nebraska	Rhode Island
Alaska	Iowa	Nevada	South Carolina
Arizona	Kansas	New Hampshire	South Dakota
Arkansas	Kentucky	New Jersey	Tennessee
California	Louisiana	New Mexico	Texas
Colorado	Maine	New York	Utah
Connecticut	Maryland	North Carolina	Vermont
Delaware	Massachusetts	North Dakota	Virginia
Florida	Michigan	Ohio	Washington
Georgia	Minnesota	Oklahoma	West Virginia
Hawaii	Mississippi	Oregon	Wisconsin
Idaho	Missouri	Pennsylvania	Wyoming
Illinois	Montana		

States – Count the number you have visited

Alabama	Indiana	Nebraska	Rhode Island
Alaska	Iowa	Nevada	South Carolina
Arizona	Kansas	New Hampshire	South Dakota
Arkansas	Kentucky	New Jersey	Tennessee
California	Louisiana	New Mexico	Texas
Colorado	Maine	New York	Utah
Connecticut	Maryland	North Carolina	Vermont
Delaware	Massachusetts	North Dakota	Virginia
Florida	Michigan	Ohio	Washington
Georgia	Minnesota	Oklahoma	West Virginia
Hawaii	Mississippi	Oregon	Wisconsin
Idaho	Missouri	Pennsylvania	Wyoming
Illinois	Montana		

Statistics – Class Survey

[illegible]

Statistics Quiz A – Chapter 1**Name** _____

1. One of the reasons that the Monitoring the Future (MTF) project was started was “to study changes in the beliefs, attitudes, and behavior of young people in the United States.” Data are collected from 8th, 10th, and 12th graders each year. To get a representative nationwide sample, surveys are given to a randomly selected group of students. In Spring 2016, students were asked about alcohol, illegal drug, and cigarette use. Describe the W’s, if the information is given. If the information is not given, state that it is not specified.

- Who:
- What:
- When:
- Where:
- How:
- Why:

2. Consider the following part of a data set:

Age (years)	Sex	Only child?	Height (inches)	Weight (pounds)	Credit Hours	GPA	Major
21	Female	Yes	67.00	140.0	16	3.60	animal science
20	Female	No	62.00	130.0	18	3.86	biology
28	Female	No	64.00	188.0	21	3.25	psychology
21	Male	No	65.00	140.0	15	2.95	psychology
24	Female	No	67.00	130.0	20	3.00	anthropology
22	Male	Yes	68.00	135.0	15	2.94	journalism

List the variables in the data set. Indicate whether each variable is treated as categorical or quantitative in this data set. If the variable is quantitative, state the units.

Statistics Quiz A – Chapter 1 – Key

1. One of the reasons that the Monitoring the Future (MTF) project was started was “to study changes in the beliefs, attitudes, and behavior of young people in the United States.” Data are collected from 8th, 10th, and 12th graders each year. To get a representative nationwide sample, surveys are given to a randomly selected group of students. In Spring 2016, students were asked about alcohol, illegal drug, and cigarette use. Describe the W’s, if the information is given. If the information is not given, state that it is not specified.

- Who: 8th, 10th, and 12th graders
- What: alcohol, illegal drug, and cigarette use
- When: Spring 2016
- Where: United States
- How: survey
- Why: “to study changes in the beliefs, attitudes, and behavior of young people in the United States”

2. Consider the following part of a data set:

Age (years)	Sex	Only child?	Height (inches)	Weight (pounds)	Credit Hours	GPA	Major
21	Female	Yes	67.00	140.0	16	3.60	animal science
20	Female	No	62.00	130.0	18	3.86	biology
28	Female	No	64.00	188.0	21	3.25	psychology
21	Male	No	65.00	140.0	15	2.95	psychology
24	Female	No	67.00	130.0	20	3.00	anthropology
22	Male	Yes	68.00	135.0	15	2.94	journalism

List the variables in the data set. Indicate whether each variable is treated as categorical or quantitative in this data set. If the variable is quantitative, state the units.

Categorical: sex, only child?, major

Quantitative: age (years), height (inches), weight (pounds), credit hours, GPA

Statistics Quiz B – Chapter 1

Name _____

In November 2003 *Discover* published an article on the colonies of ants. They reported some basic information about many species of ants and the results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida. Information included the scientific name of the ant species, the geographic location, the depth of the nest (in feet), the number of chambers in the nest, and the number of ants in the colony. The article documented how new ant colonies begin, the ant-nest design, and how nests differ in shape, number, size of chambers, and how they are connected, depending on the species. It reported that nest designs include vertical, horizontal, or inclined tunnels for movement and transport of food and ants.

1. Describe the W's, if the information is given:
 - Who:
 - What:
 - When:
 - Where:
 - How:
 - Why:
2. List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units.

Statistics Quiz B – Chapter 1 – Key

In November 2003 *Discover* published an article on the colonies of ants. They reported some basic information about many species of ants and the results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida. Information included the scientific name of the ant species, the geographic location, the depth of the nest (in feet), the number of chambers in the nest, and the number of ants in the colony. The article documented how new ant colonies begin, the ant-nest design, and how nests differ in shape, number, size of chambers, and how they are connected, depending on the species. It reported that nest designs include vertical, horizontal, or inclined tunnels for movement and transport of food and ants.

1. Describe the W's, if the information is given:

- Who: Colonies of ants. "Many species of ants," but no indication of exactly how many.
- What: scientific name, geographic location, average nest depth, average number of chambers, average colony size, how new ant colonies begin, the ant-nest design, and how nests differ in architecture.
- When: November 2003
- Where: not specified
- How: The results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida
- Why: Information of interest to readers of the magazine

2. List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units.

Categorical: species, geographic location, how new ant colonies begin, and nest design.

Quantitative: nest depth (feet), number of chambers (units), and colony size (units).

Statistics Quiz C – Chapter 1

Name _____

In May 2017, Wirecutter published an article entitled “The Best True Wireless Headphones So Far” (<http://thewirecutter.com/reviews/best-true-wireless-headphones/>). They tested 11 “of the most promising true wireless in-ear headphones.” Among other things, the article told the brand of each pair of headphones, its price, battery life, audio quality, ease of setup, and other characteristics. The article provides a number of recommendations including best for the money, best for the gym, best for Apple, and best for Android. The author, Lauren Dragan, describes herself as a voice actor with an audio production degree who has spent hundreds of hours testing headphones for Wirecutter.

1. Describe the W’s, if the information is given:
 - Who:
 - What:
 - When:
 - Where:
 - How:
 - Why:
2. List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units.

Statistics Quiz C – Chapter 1 – Key

In May 2017, Wirecutter published an article entitled “The Best True Wireless Headphones So Far” (<http://thewirecutter.com/reviews/best-true-wireless-headphones/>). They tested 11 “of the most promising true wireless in-ear headphones.” Among other things, the article told the brand of each pair of headphones, its price, battery life, audio quality, ease of setup, and other characteristics. The article provides a number of recommendations including best for the money, best for the gym, best for Apple, and best for Android. The author, Lauren Dragan, describes herself as a voice actor with an audio production degree who has spent hundreds of hours testing headphones for Wirecutter.

1. Describe the W’s, if the information is given:
 - Who: 11 wireless in-ear headphones.
 - What: brand, price, battery life, audio quality, ease of setup, and other characteristics.
 - When: May 2017
 - Where: not specified, probably the United States
 - How: presumably lab tests of the 11 models
 - Why: information for potential consumers
2. List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units.

Categorical: brand, audio quality, ease of setup

Quantitative: price (US\$), battery life (probably hours)

Statistics Quiz D – Chapter 1

Name _____

1. In the fall of 2007, the *Pew Internet & Life Project* conducted telephone interviews with a sample of American adults aged 18 and older about online shopping. American adults aged 18 and older constitute the _____ of the study.
 - A. Who
 - B. What
 - C. When
 - D. Where
 - E. How

2. A few of the variables for which data were collected in the *Pew Internet & Life Project* study about online shopping include age, gender, income, and number of hours spent shopping online per month. Which of the variables is categorical?
 - A. Age
 - B. Gender
 - C. Income
 - D. Number of hours spent shopping online
 - E. None

3. The *Pew Internet & Life Project* study about online shopping asked respondents to indicate their education level on the following scale: *Less than High School, High School, Some College, College +*. Which of the following statements is (are) true?
 - A. Education level is a categorical variable.
 - B. Education level is nominal scaled.
 - C. Education level is ordinal scaled.
 - D. Both A and B
 - E. Both A and C

4. *Consumer Reports Health* routinely compares drugs in terms of effectiveness and safety. In summer 2008 they reviewed drugs used to treat arthritis. Among the information reported was convenience of use (how many pills required each day) and possible side effects (e.g., dizziness, stomach upset). Convenience of use and possible side effects constitute the _____ of the study.
 - A. Who
 - B. What
 - C. When
 - D. Where
 - E. How

5. What is the “Who” in a *Consumer Reports Health* study on the effectiveness and safety of drugs used to treat arthritis?
 - A. drugs to treat arthritis currently on the market
 - B. convenience of use and possible side effects
 - C. summer 2008
 - D. the United States
 - E. testing on drugs
6. A *Consumer Reports Health* study on the effectiveness and safety of arthritis drugs collected data on possible side effects. This is what kind of variable?
 - A. Quantitative
 - B. Categorical
 - C. Nominal
 - D. Both A and C
 - E. Both B and C
7. A *Consumer Reports Health* study on arthritis drugs takes into consideration cost. Cost is
 - A. is a nominal variable.
 - B. is a categorical variable.
 - C. is a quantitative variable.
 - D. is an ordinal variable.
 - E. is an irrelevant variable.
8. The Human Resources Department of a large corporation maintains records on its employees. Data are maintained of the following variables: *Age, Employment Category, Education, Whether or not the employee participates in a wellness program, and Paycheck benefit deductions*. Which of these variables are categorical?
 - A. *Age, Employment Category, and Education*
 - B. *Employment Category, Education, and Whether or not the employee participates in a wellness program*
 - C. *Education, Whether or not the employee participates in a wellness program, and Paycheck benefit deductions*
 - D. All of the variables
 - E. None of the variables

1-12 Part I Exploring and Understanding Data

7. A *Consumer Reports* study on tipping takes into consideration median amount of tipping for service providers. Tipping is
- A. is a nominal variable.
 - B. is a categorical variable.
 - C. is a quantitative variable.
 - D. is an ordinal variable.
 - E. is an irrelevant variable.
8. The Human Resources Department of a large corporation maintains records on its employees. Data are maintained of the following variables: *Age*, *Employment Category*, *Education*, and *Whether or not the employee has an advanced degree*. Which of these variables are categorical?
- A. *Age*, *Employment Category*, and *Education*
 - B. *Employment Category* and *Education*
 - C. *Education* and *Whether or not the employee has an advanced degree*
 - D. All of the variables
 - E. None of the variables

Statistics Quiz D – Chapter 1 – Key

1. A
2. B
3. E
4. B
5. A
6. E
7. C
8. B
9. C
10. B

Statistics Quiz E – Chapter 1

Name _____

1. A university is interested in gauging student satisfaction in its online MBA program. A survey is designed and administered via the Internet to a sample of students currently active in the program. Which of the following would best describe the cases?
 - A. Participants
 - B. Respondents
 - C. Experimental Units
 - D. Subjects
 - E. Variables

2. In a survey undertaken by a university to gauge student satisfaction in its online MBA program, one question asked students to indicate their employment status (unemployed, employed part-time, employed full-time). Which of the following is true?
 - A. This variable is categorical.
 - B. This variable is quantitative.
 - C. This is an identifier variable.
 - D. Both A and C.
 - E. Both B and C.

3. In a survey undertaken by a university to gauge student satisfaction in its online MBA program, one question asked students to indicate the number of credits they had transferred into the program. Which of the following is true?
 - A. This variable is categorical.
 - B. This variable is transactional.
 - C. This variable is quantitative.
 - D. This is an identifier variable.
 - E. This variable is nominal.

4. Researchers in e-commerce design an experiment to determine what factors are most important to online consumers when completing a transaction via the Internet. Individuals perform tasks on a set of Web sites and record their impressions about various attributes. Which of the following would best describe the cases?
 - A. Participants
 - B. Respondents
 - C. Experimental Units
 - D. Identifiers
 - E. Variables

5. A popular travel magazine regularly reviews hotels worldwide. In a recent issue, it focused on hotels in Hawaii. Among the variables for which it provided data was whether or not the hotel included a spa. This is best described as a
- A. quantitative variable.
 - B. identifier variable.
 - C. ordinal variable.
 - D. categorical variable.
 - E. nominal variable.
6. A popular travel magazine regularly reviews hotels worldwide. In a recent issue, it focused on hotels in Hawaii. Among the variables for which it provided data was the price range for rooms with an ocean view. Which of the following statements is true?
- A. This variable is quantitative and has no units.
 - B. This variable is quantitative and the units are \$.
 - C. This variable is quantitative and the units are number of rooms.
 - D. This variable is qualitative and ordinal.
 - E. This variable is qualitative and nominal.
7. A mid-priced chain of hotels, *Hometown Suites*, strives to make its guests “feel at home” by providing amenities such as microwaves in every room. Comment cards are used to get feedback on the importance of such amenities by asking guests to rate them using the scale: ___ Essential ___ Important ___ Not Important. These data are
- A. qualitative.
 - B. nominal.
 - C. ordinal.
 - D. both A and B.
 - E. both A and C.
8. *Businesses* are interested in the work experience of recent graduates from a local business school. Whether or not the graduates have work experience constitutes the _____ of the study.
- A. Who
 - B. What
 - C. When
 - D. Where
 - E. How

1-16 Part I Exploring and Understanding Data

9. What is the “What” in a *Consumer Reports Tipping* study on the level of tipping during the current holiday season compared to the last holiday season?
- A. whether or not tipped
 - B. amount of tip compared to last year
 - C. the type of tip
 - D. the United States
10. A *Consumer Reports* survey on the level of tipping for service providers. This is what kind of variable?
- A. Quantitative
 - B. Ordinal
 - C. Nominal
 - D. Both A and C
 - E. Both B and C

Statistics Quiz E – Chapter 1 – Key

1. B
2. A
3. C
4. A
5. D
6. B
7. E
8. B
9. B
10. B

Chapter 2

Displaying and Describing Data

Solutions to Class Examples:

1. Answers will vary according to your data.
2. Answers will vary according to your data.
3. See Class Example 3.
4. Answers will vary according to your data.
5. See Class Example 5.

Investigative Task

The task, Dollars for Students, examines education spending in the United States.

Statistics Quiz A – Chapter 2**Name** _____

1. Students in a statistics course were asked to answer a variety of questions. One question asked for students to pick a favorite Superpower. Results are listed to the right.

Superpower	Num of Students
Fly	19
Freeze Time	11
Invisibility	7
Super strength	4
Telepathy	12

- a. Create a visual display for these data.

- b. Describe the distribution of the students' choices.

2. Which of the following variables would be appropriate to graph using a pie or bar graph?

- | | | | |
|--|-----|----|----|
| a. Annual income for 20 employees | Yes | or | No |
| b. The favorite baseball team of 30 students | Yes | or | No |
| c. The number of pages in 15 textbooks | Yes | or | No |
| d. The country of origin of 25 immigrants | Yes | or | No |

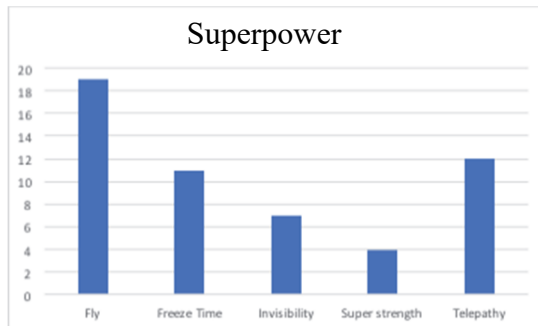
3. Describe why the area principle is important in making a bar graph. It might be fun to ask an artist to liven up a bar graph by turning the bars into images. But include in your explanation why this might be risky.

Statistics Quiz A – Chapter 2 KEY

1. Students in a statistics course were asked to answer a variety of questions. One question asked for students to pick a favorite Superpower. Results are listed to the right.

Superpower	Num of Students
Fly	19
Freeze Time	11
Invisibility	7
Super strength	4
Telepathy	12

- a. Create a visual display for these data.



- b. Describe the distribution of the students' choices.

The ability to fly is the students' most popular choice, followed by telepathy and freeze time. Super strength is the least popular, with invisibility also not very popular.

2. Which of the following variables would be appropriate to graph using a pie or bar graph?

- a. Annual income for 20 employees
 b. The favorite baseball team of 30 students
 c. The number of pages in 15 textbooks
 d. The country of origin of 25 immigrants

Yes or ☒ No
☒ Yes or No
 Yes or ☒ No
☒ Yes or No

3. Describe why the area principle is important in making a bar graph. It might be fun to ask an artist to liven up a bar graph by turning the bars into images. But include in your explanation why this might be risky.

If the reader of the graphs is going to understand the relationship between the groups, it is important that the area allotted to each group match the proportion of group in the sample. If an artist turns the bars into some kind of image, he might be not make each picture the proper area, and thus distort the actual percentage breakdown of each group.

Statistics Quiz B – Chapter 2**Name** _____

1. A survey conducted in a college intro stats class asked students about the number of credit hours they were taking that quarter. The number of credit hours for a random sample of 16 students is given in the table.

10	10	12	14	15	15	15	15
17	17	19	20	20	20	20	22

- Sketch a histogram of these data
- Find the mean and standard deviation for the number of credit hours.
- Find the median and IQR for the number of credit hours.
- Is it more appropriate to use the mean and standard deviation or the median and IQR to summarize these data? Explain.

2. Suppose that the student taking 22 credit hours in the data set in the previous question was actually taking 28 credit hours instead of 22 (so we would replace the 22 in the data set with 28). Indicate whether changing the number of credit hours for that student would make each of the following summary statistics increase, decrease, or stay about the same:

- mean _____
- median _____
- range _____
- IQR _____
- standard deviation _____

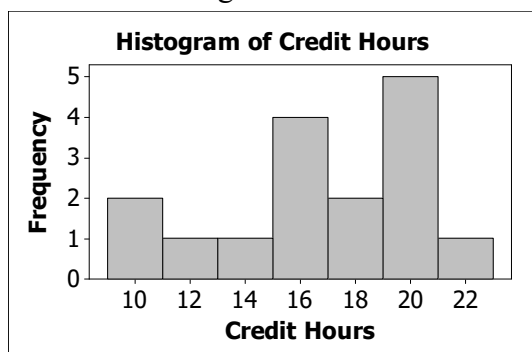
Statistics Quiz B – Chapter 2 - KEY

Name _____

1. A survey conducted in a college intro stats class asked students about the number of credit hours they were taking that quarter. The number of credit hours for a random sample of 16 students is given in the table.

10	10	12	14	15	15	15	15
17	17	19	20	20	20	20	22

- a. Sketch a histogram of these data



- b. Find the mean and standard deviation for the number of credit hours.

$$\bar{x} = 16.3 \text{ credit hours}; s = 3.7 \text{ credit hours}$$

- c. Find the median and IQR for the number of credit hours.

The median is 16.0 credit hours.
 $IQR = Q3 - Q1 = 20 - 14.5 = 5.5$ credit hours

- d. Is it more appropriate to use the mean and standard deviation or the median and IQR to summarize these data? Explain.

It is more appropriate to use the median and IQR to summarize these data, because these data are not unimodal and symmetric.

2. Suppose that the student taking 22 credit hours in the data set in the previous question was actually taking 28 credit hours instead of 22 (so we would replace the 22 in the data set with 28). Indicate whether changing the number of credit hours for that student would make each of the following summary statistics increase, decrease, or stay about the same:

- | | |
|-----------------------|----------------------------|
| a. mean | <u>increase</u> |
| b. median | <u>stay about the same</u> |
| c. range | <u>increase</u> |
| d. IQR | <u>stay about the same</u> |
| e. standard deviation | <u>increase</u> |

Statistics Quiz C – Chapter 2**Name** _____

1. The students in a biology class kept a record of the height (in centimeters) of plants for a class experiment.

49	67	38	55	62
54	36	41	56	43
48	75	44	60	48
52	48	53	59	32

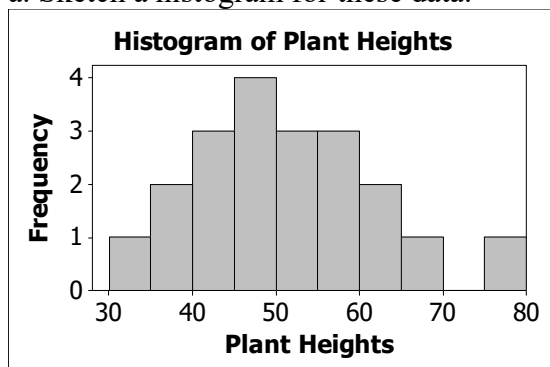
- a. Sketch a histogram for these data.
 - b. Find the mean and standard deviation of the plant heights.
 - c. Is it appropriate to use the mean and standard deviation to summarize these data? Explain.
 - d. Describe the distribution of plant heights.
2. All students in a physical education class completed a basketball free-throw shooting event and the highest number of shots made was 32. The next day a student who had just transferred into the school completed the event, making 35 shots. Indicate whether adding the new student's score to the rest of the data made each of these summary statistics increase, decrease, or stay about the same:
- a. mean _____
 - b. median _____
 - c. range _____
 - d. IQR _____
 - e. standard deviation _____

Statistics Quiz C – Chapter 2 – KEY

1. The students in a biology class kept a record of the height (in centimeters) of plants for a class experiment.

49	67	38	55	62
54	36	41	56	43
48	75	44	60	48
52	48	53	59	32

- a. Sketch a histogram for these data.



- b. Find the mean and standard deviation of the plant heights.

$$\bar{x} = 51.0 \text{ cm}; s = 10.6 \text{ cm}$$

- c. Is it appropriate to use the mean and standard deviation to summarize these data? Explain.

Yes, the data are roughly unimodal and symmetric with no outliers.

- d. Describe the distribution of plant heights.

The data are roughly symmetric with no outliers; however there is a small gap from 70 to 75 cm. The average plant height is 51.0 centimeters, with a standard deviation of 10.6 centimeters. The range of plant heights is 43 centimeters. The distribution of plant heights has a mode between 45 and 49 centimeters.

2. All students in a physical education class completed a basketball free-throw shooting event and the highest number of shots made was 32. The next day a student who had just transferred into the school completed the event, making 35 shots. Indicate whether adding the new student's score to the rest of the data made each of these summary statistics increase, decrease, or stay about the same:

- | | |
|-----------------------|----------------------------|
| a. mean | <u>increase</u> |
| b. median | <u>stay about the same</u> |
| c. range | <u>increase</u> |
| d. IQR | <u>stay about the same</u> |
| e. standard deviation | <u>increase</u> |

Statistics Quiz D – Chapter 2**Name** _____

1. A brake and muffler shop reported the repair bills, in dollars, for their customers yesterday.

88	283	312	290	172
154	400	381	346	181
203	118	143	252	227
56	192	292	213	422

- Sketch a histogram for these data.
- Find the mean and standard deviation of the repair costs.
- Is it appropriate to use the mean and standard deviation to summarize these data? Explain.
- Describe the distribution of repair costs.

2. In a survey of 1,500 millennials (ages 18-34), *Business Insider* asked which services were used to consume video content. Use the data on the right to construct an appropriate display **and** describe your graph.

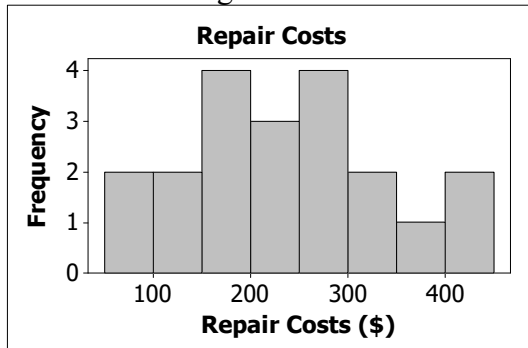
Service	Percent
YouTube	81%
Netflix	79%
Hulu	37%
Amazon Prime	34%
HBO	21%
Crackle	12%
None	3%
Other	2%

Statistics Quiz D – Chapter 2 – KEY

1. A brake and muffler shop reported the repair bills, in dollars, for their customers yesterday.

88	283	312	290	172
154	400	381	346	181
203	118	143	252	227
56	192	292	213	422

- a. Sketch a histogram for these data.



- b. Find the mean and standard deviation of the repair costs.

$$\bar{x} = \$236.25; s = \$103.43$$

- c. Is it appropriate to use the mean and standard deviation to summarize these data? Explain.

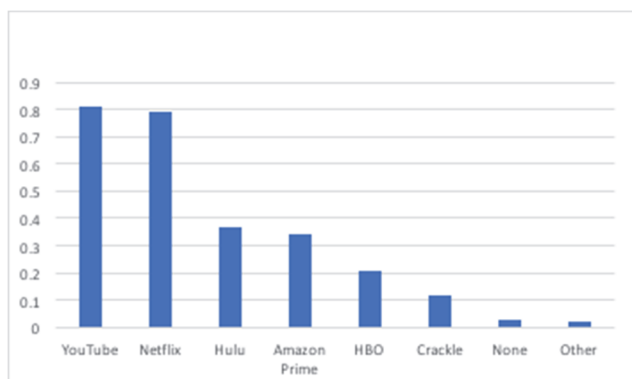
Yes, the data are roughly unimodal and symmetric with no outliers.

- d. Describe the distribution of repair costs.

The repair costs averaged \$236.25, ranging from \$56 to \$422 with a standard deviation of \$103.43. The distribution was approximately symmetric, with typical repair costs clustered between \$150 and \$300.

2. In a survey of 1,500 millennials (ages 18-34), *Business Insider* asked which services were used to consume video content. Use the data on the right to construct an appropriate display **and** describe your graph.

Service	Percent
YouTube	81%
Netflix	79%
Hulu	37%
Amazon Prime	34%
HBO	21%
Crackle	12%
None	3%
Other	2%



YouTube and Netflix are clearly the dominant choices. And as these percentages add to more than 100%, it is clear that most millennials use more than one service, making the pie chart an inappropriate choice. Hulu and Amazon are used by about one-third of the millennials, while Crackle and HBO are clearly the least popular.

Dollars for Students

In 2008 the U.S. Census Bureau published *Public Education Finances*, reporting the average amount (dollars per student) spent by public schools in each state during the 2006 school year. (The table seen here divides states according to whether they lie east or west of the Mississippi River).

Write a report describing the amounts states spent to educate their children.

Eastern State	\$ per Student	Western State	\$ per Student
Alabama	7,646	Alaska	11,460
Connecticut	12,323	Arizona	6,472
Delaware	11,633	Arkansas	7,927
D.C.	13,446	California	8,486
Florida	7,759	Colorado	8,057
Georgia	8,565	Hawaii	9,876
Illinois	9,149	Idaho	6,440
Indiana	8,793	Iowa	8,360
Kentucky	7,662	Kansas	8,392
Maine	10,586	Louisiana	8,402
Maryland	10,670	Minnesota	9,138
Massachusetts	11,981	Missouri	8,107
Michigan	9,572	Montana	8,581
Mississippi	7,221	Nebraska	8,736
New Hampshire	10,079	Nevada	7,345
New Jersey	14,630	New Mexico	8,086
New York	14,884	North Dakota	8,603
North Carolina	7,388	Oklahoma	6,961
Ohio	9,598	Oregon	8,545
Pennsylvania	11,028	South Dakota	7,651
Rhode Island	11,769	Texas	7,561
South Carolina	8,091	Utah	5,437
Tennessee	6,883	Washington	7,830
Vermont	12,614	Wyoming	11,197
Virginia	9,447		
West Virginia	9,352		
Wisconsin	9,970		

A complete report will include a visual display (stem-and-leaf plot), appropriate statistics, and a well-written description of the expenditures (in context, of course).

Statistics - Investigative Task**Chapter 2**

	Components	Comments
Think	Demonstrates clear understanding of statistical concepts, vocabulary, and procedures in analyzing and describing these data.	
Show	Visual/Numerical: ○ stem-and-leaf plot ○ plot well-labeled ○ plot correctly constructed ○ summary statistics correct	
Tell	Verbal: Describes the distribution of expenditures <i>in context</i> , including... ○ shape (skewed right by higher spending in the East) ○ center (median) ○ spread (IQR) The written analysis... ○ also interprets at least one quartile, or the max or min <i>in context</i> ○ identifies the W's ○ uses statistical vocabulary correctly ○ avoids speculation	

Components are scored as **Essentially correct**, **Partially correct**, or **Incorrect**

1: Visual/Numerical

E – Has all 4 features

P – Has only 3 of the 4 features, but attempts an appropriate graph (ex: histogram)

I – Graph is not appropriate (ex: bar chart), has many errors, or is missing

2: Shape

E – Identifies skewness to the right, attributed to generally higher spending in the East.

P – Mis-identifies skewness OR overlooks East/West spending difference

I – Description of shape is missing or incorrect

3: Center and Spread

E – Correctly interprets median and IQR in context

P – Correctly interprets only one of median/IQR OR lacks context OR uses mean and SD

I – Has more than one of the three shortcomings described in “P”

4: Written Analysis

E – Has all 4 listed properties.

P – Has only 2 or 3 of the listed properties.

I – Has fewer than of the properties.

Scoring

- E's count 1 point, P's are 1/2
- Grade: A = 4, B = 3, etc., with +/- based on rounding (ex: 3.5 rounded to 3 is a B+)

Name _____ **Grade** _____

NOTE: We present a model solution with some trepidation. This is not a scoring key, just an example. Many other approaches could fully satisfy the requirements outlined in the scoring rubric. That (not this) is the standard by which student responses should be evaluated.

Model Solution – Investigative Task – Dollars for Students

The U.S. Census published *Public Education Finances*, a collection of educational information in 2008. Among the data reported was information on average educational spending by state during the 2006 school year, in dollars per student.

Spending by many states was near the nationwide median of \$8581 per student. The middle 50% of states spent between \$7759 and \$10586 per student, an interquartile range of \$2827. The distribution of the average number of education dollars spent per student was skewed to the right. Many Eastern States spent in excess of \$12,000 per student, including the maximum of \$14884 per student in New York. Utah spent the least, \$5437 per student.

Dollars Spent per Student

5		4
6		4489
7		2335666789
8		00013344555677
9		11345589
10		056
11		014679
12		36
13		4
14		68

Key : 6|7 = 6000 - 6999
dollars spent per student

Statistics Quiz E – Chapter 2**Name** _____

1. A automobile marketing firm conducts a study to see what types of cars people owned before buying an American car. The results are shown below.

Previous Ownership	Frequency
American	760
Japanese	375
Korean	72
German	37
Other	24
Total	1268

The relative frequency of those who owned Japanese cars previously who now bought American cars is

- A. 59.9 %
 B. 29.6%
 C. 5.7%
 D. 14.9%
 E. 2.9%
2. A restaurant uses comment cards to get feedback from its customers about newly added items to the menu. It recently introduced homemade organic veggie burgers. Customers who tried the new burger were asked if they would order it again. The data are summarized in the table below. What percentage of customers would definitely order the veggie burger again?

Response	Frequency
<i>Definitely would.</i>	10
<i>Most likely would.</i>	40
<i>Maybe</i>	12
<i>Definitely would not.</i>	3

- A. 10%
 B. 15%
 C. 20%
 D. 40%
 E. 77%

2-14 Part I Exploring and Understanding Data

3. A restaurant uses comment cards to get feedback from its customers about newly added items to the menu. It recently introduced homemade organic veggie burgers. Customers who tried the new burger were asked if they would order it again. The data are summarized in the table below. What percentage of customers would most likely or definitely order the veggie burger again?

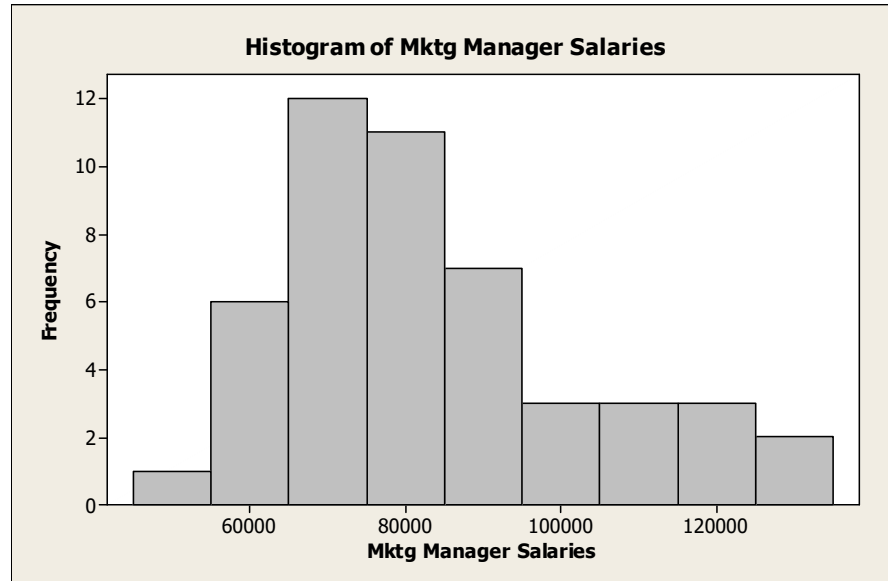
<i>Response</i>	<i>Frequency</i>
<i>Definitely would.</i>	10
<i>Most likely would.</i>	40
<i>Maybe</i>	12
<i>Definitely would not.</i>	3

- A. 10%
 B. 15%
 C. 40%
 D. 50%
 E. 77%
4. A restaurant uses comment cards to get feedback from its customers about newly added items to the menu. It recently introduced homemade organic veggie burgers. Customers who tried the new burger were asked if they would order it again. Which of the following would be an appropriate method for displaying the data shown in the table?

<i>Response</i>	<i>Frequency</i>
<i>Definitely would.</i>	10
<i>Most likely would.</i>	40
<i>Maybe</i>	12
<i>Definitely would not.</i>	3

- A. Histogram.
 B. Dotplot.
 C. Pie chart.
 D. Stem-and-leaf display.
 E. Both C and D.

5. Below is a histogram of salaries (in \$) for a sample of U.S. marketing managers.



The shape of this distribution is

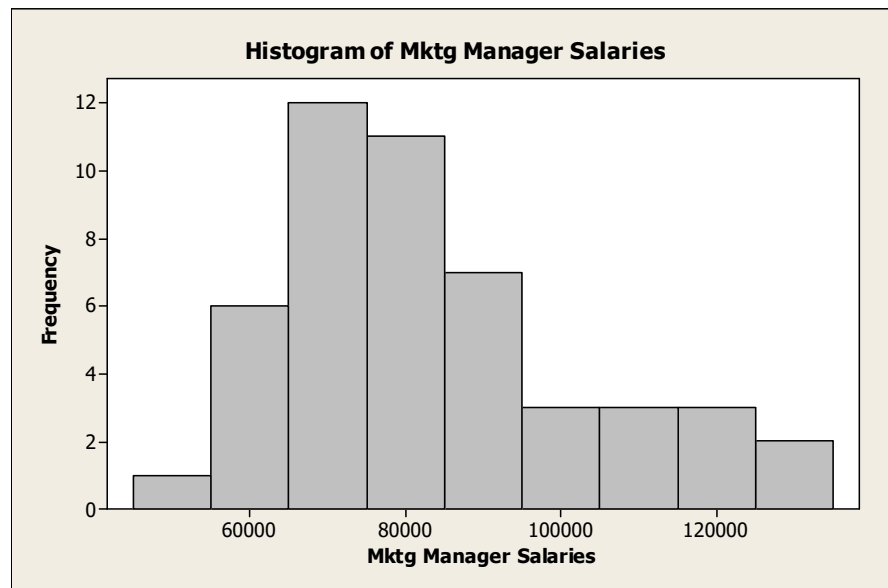
- A. symmetric.
 - B. bimodal.
 - C. right skewed.
 - D. left skewed.
 - E. normal.
6. Here is the five number summary for salaries of U.S. marketing managers.

Min	Q1	Median	Q3	Max
46360	69693	77020	91750	129420

The IQR is

- A. \$83,060.
- B. \$22,057.
- C. \$69,693.
- D. \$77,020.
- E. \$14,566.

7. Below is a histogram of salaries (in \$) for a sample of U.S. marketing managers.



The most appropriate measure of central tendency for these data is the

- A. median.
 - B. mean.
 - C. mode.
 - D. range.
 - E. standard deviation.
8. Consider the five number summary for salaries of U.S. marketing managers.

Min	Q1	Median	Q3	Max
46360	69693	77020	91750	129420

Suppose the marketing manager who was earning \$129,420 got a raise and is now earning \$140,000. Which of the following statement is true?

- A. The mean would increase.
- B. The median would increase.
- C. The range would increase.
- D. Both A and C.
- E. All of the above.

9. The following table shows data for total assets (\$ billion) for a small sample of U.S. banks (late 2013).

<i>BANK</i>	<i>ASSETS (\$ billion)</i>
State Street Bank and Trust	160.5
Discover Bank	63.9
BancWest	72.8
Citizens Bank	130.0
Northern Trust	83.8
Huntington Bank	53.8
Key Bank	91.8
People's United	27.9

The mean for the total assets data (\$ billion) is

- A. \$78.3.
 - B. \$56.3.
 - C. \$85.6.
 - D. \$120.5.
 - E. \$42.4.
10. The following table shows representative recent closing share prices for a small sample of companies based in India in late 2013.

<i>COMPANY</i>	<i>CLOSING SHARE PRICE</i>
20 Microns	30.95
ABC Paper	24.30
Bank of MA	36.20
Photoquip	37.00
Saksoft	67.20
Marg LTD	13.99
Galaxy ENT	10.40
Sonatasoft	68.75
EDynamics	49.95
DB Corp.	287.95

The standard deviation in closing share prices is

- A. \$81.6.
- B. \$25.8.
- C. \$36.6.
- D. \$62.7.
- E. \$67.6.

Statistics Quiz E – Chapter 2 – KEY

1. B
2. B
3. E
4. C
5. C
6. B
7. A
8. D
9. C
10. A

Statistics Quiz F – Chapter 2**Name** _____

1. A clothing store uses comment cards to get feedback from its customers about newly added items. It recently introduced plus size fashion wear. Customers who purchased the items were asked to fill out an online comment survey giving 10% off the next purchase. The data are summarized in the table below. What percentage of customers were at least satisfied with the item(s) purchased (Satisfied or Very satisfied)?

<i>Response</i>	<i>Frequency</i>
<i>Very satisfied.</i>	15
<i>Satisfied.</i>	30
<i>Less than fully satisfied.</i>	12
<i>Not satisfied.</i>	4

- A. 49.2%
 B. 73.8%
 C. 24.6%
 D. 26.2%
 E. 68.9%
2. A clothing store uses comment cards to get feedback from its customers about newly added items. It recently introduced plus size fashion wear. Customers who purchased the items were asked to fill out an online comment survey giving 10% off the next purchase. The data are summarized in the table below. What percentage of customers would be less likely to purchase another item (Less or Not fully satisfied)?

<i>Response</i>	<i>Frequency</i>
<i>Very satisfied.</i>	15
<i>Satisfied.</i>	30
<i>Less than fully satisfied.</i>	12
<i>Not satisfied.</i>	4

- A. 10%
 B. 15%
 C. 40%
 D. 50%
 E. 77%

2-20 Part I Exploring and Understanding Data

3. A clothing store uses comment cards to get feedback from its customers about newly added items. It recently introduced plus size fashion wear. Customers who purchased the items were asked to fill out an online comment survey giving 10% off the next purchase. The data are summarized in the table below. Which of the following would be an appropriate method for displaying the data shown in the table?

<i>Response</i>	<i>Frequency</i>
<i>Very satisfied.</i>	15
<i>Satisfied.</i>	30
<i>Less than fully satisfied.</i>	12
<i>Not satisfied.</i>	4

- A. Histogram.
 B. Dotplot.
 C. Pie chart.
 D. Stem-and-leaf display.
 E. Both C and D.
4. The following table shows total assets (\$ billion) for a small sample of U.S. banks.

<i>BANK</i>	<i>ASSETS (\$ billion)</i>
Bank of New York	88
Regions Financial	80
Fifth Third Bank	58
State Street Bank and Trust	92
Branch Banking and Trust Company	81
Chase Bank	70
Key Bank	89
PNC Bank	84

The mean for these data is

- A. \$ 80.25 billion.
 B. \$ 100.35 billion.
 C. \$ 75.68 billion.
 D. \$ 84 billion.
 E. \$ 89 billion.

5. The following table shows total assets (\$ billion) for a small sample of U.S. banks.

<i>BANK</i>	<i>ASSETS (\$ billion)</i>
Bank of New York	88
Regions Financial	80
Fifth Third Bank	58
State Street Bank and Trust	92
Branch Banking and Trust Company	81
Chase Bank	70
Key Bank	89
PNC Bank	84

The standard deviation for these data is

- A. \$12.78 billion.
 - B. \$ 11.27 billion.
 - C. \$ 127.01 billion.
 - D. \$ 21.67 billion.
 - E. \$ 34 billion.
6. Consider the five number summary of hourly wages (\$) for a sample of sales managers.

Min	Q1	Median	Q3	Max
20.94	37.64	44.77	49.34	67.11

The range for these data is

- A. \$11.70
 - B. \$46.17
 - C. \$67.11
 - D. \$20.94
 - E. \$44.77
7. Consider the five number summary of hourly wages (\$) for a sample of sales managers.

Min	Q1	Median	Q3	Max
20.94	37.64	44.77	49.34	67.11

The IQR for these data is

- A. \$11.70
- B. \$46.17
- C. \$67.11
- D. \$20.94
- E. \$44.77

2-22 Part I Exploring and Understanding Data

8. Consider the five number summary of hourly wages (\$) for a sample of sales managers. Suppose the mean hourly wage is \$38.50. What can we say about the shape of the distribution?

Min	Q1	Median	Q3	Max
20.94	37.64	44.77	49.34	67.11

- A. The distribution of hourly wages for sales managers is symmetric.
 B. The distribution of hourly wages for sales managers is skewed right.
 C. The distribution of hourly wages for sales managers is skewed left.
 D. The distribution of hourly wages for sales managers is bimodal.
 E. None of the above.
9. Consider the five number summary of hourly wages (\$) for a sample of advertising / promotion managers.

Min	Q1	Median	Q3	Max
19.64	29.36	34.18	40.86	57.26

Suppose there had been an error and that the lowest hourly wage was \$15.50 instead of \$19.64. This would result in

- A. an increase in the median.
 B. an increase in the standard deviation.
 C. a decrease in the range.
 D. a decrease in the IQR.
 E. an increase in the mean.
10. Of the following stem-and-leave plots of four data sets each containing 11 observations, which represents the set of data that has the greatest standard deviation?

Variable: Set A

Decimal point is at the colon.
Leaf unit = 0.1

0 : 0
 1 : 0
 2 : 0
 3 : 0
 4 : 0
 5 : 0
 6 : 0
 7 : 0
 8 : 0
 9 : 0
 10 : 0

Variable: Set B

Decimal point is at the colon.
Leaf unit = 0.1

0 : 012345
 1 :
 2 :
 3 :
 4 :
 5 :
 6 :
 7 :
 8 :
 9 : 6789
 10 : 0

Variable: Set C

Decimal point is at the colon.
Leaf unit = 0.1

0 : 0123
 1 :
 2 :
 3 :
 4 : 9
 5 : 01
 6 :
 7 :
 8 :
 9 : 789
 10 : 0

Variable: Set D

Decimal point is at the colon.
Leaf unit = 0.1

0 : 012
 1 :
 2 :
 3 :
 4 : 89
 5 : 012
 6 :
 7 :
 8 :
 9 : 89
 10 : 0

- A. Set A.
 B. Set B.
 C. Set C.
 D. Set D.
 E. both Set C and Set D.

Statistics Quiz F – Chapter 2 – KEY

1. B
2. E
3. C
4. A
5. B
6. B
7. A
8. C
9. B
10. B

Chapter 3

Relationships Between Categorical Variables—Contingency Tables

Solutions to Class Examples:

1. See Class Example 1.
2. Answers will vary according to your data.
3. See Class Example 3.
4. See Class Example 4.

Investigative Task

Race and the Death Penalty uses a three-way contingency table and requires comparing marginal and conditional distributions.

Supplemental Resources

After the Investigative Task, there is a worksheet on the relationship between smoking and education level.

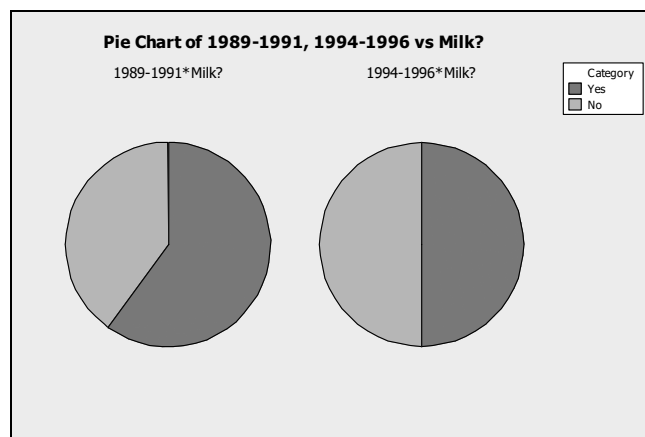
Statistics Quiz A – Chapter 3

Name _____

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from “Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes” (Shanthy A. Bowman, *Journal of the American Dietetic Association*, 102(9), pp. 1234-1239):

		Nationwide Food Survey Years			
		1987-1988	1989-1991	1994-1996	Total
Drinks Fluid Milk	Yes	354	502	366	1222
	No	226	335	366	927
	Total	580	837	732	2149

- Find the following:
 - What percent of the young girls reported that they drink milk? _____
 - What percent of the young girls were in the 1989-1991 survey? _____
 - What percent of the young girls who reported that they drink milk were in the 1989-1991 survey? _____
 - What percent of the young girls in 1989-1991 reported they drink milk? _____
- What is the marginal distribution of milk consumption?
- Do you think that milk consumption by young girls is independent of the nationwide survey year? Use statistics to justify your reasoning.
- Consider the following pie charts of a subset of the data above:



Do the pie charts above indicate that milk consumption by young girls is independent of the nationwide survey year? Explain.

Statistics Quiz A – Chapter 3 - Key

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from “Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes” (Shanthy A. Bowman, *Journal of the American Dietetic Association*, 102(9), pp. 1234-1239):

		Nationwide Food Survey Years			
		1987-1988	1989-1991	1994-1996	Total
Drinks Fluid Milk	Yes	354	502	366	1222
	No	226	335	366	927
	Total	580	837	732	2149

- Find the following:
 - What percent of the young girls reported that they drink milk? 56.9%
 - What percent of the young girls were in the 1989-1991 survey? 38.9%
 - What percent of the young girls who reported that they drink milk were in the 1989-1991 survey? 41.1%
 - What percent of the young girls in 1989-1991 reported they drink milk? 60.0%
- What is the marginal distribution of milk consumption?
Yes: 56.9%; No: 43.1%
- Do you think that milk consumption by young girls is independent of the nationwide survey year? Use statistics to justify your reasoning.
No. 56.9% of all young girls surveyed reported drinking milk, but 60% of the young girls reported drinking milk in the 1989-1991 survey. Since these percentages differ, milk consumption and year are not independent.
- Consider the following pie charts of a subset of the data above:

Do the pie charts above indicate that milk consumption by young girls is independent of the nationwide survey year? Explain.

No. It looks like there is some sort of relationship between milk consumption and nationwide survey year, since the percentage of young girls who reported drinking milk is a larger slice of the pie chart for the 1989-1991 survey than the same response for the 1994-1996 survey.