

Solutions

● ● ● In-Class Activities

Activity 1-1: Student Data

- a. Answers will vary.
- b. Answers will vary.
- c. No, every word did not contain the same number of letters.
- d. The variable is the *number of letters in each word*.
- e.
 - *How many hours you slept in the past 24 hours*: quantitative
 - *Whether you slept for at least 7 hours in the past 24 hours*: binary categorical
 - *How many states you have visited*: quantitative
 - *Handedness*: binary categorical, unless you classify “ambidextrous,” in which case it is not binary.
 - *Day of the week on which you were born*: categorical
 - *Gender*: binary categorical
 - *Average study time per week*: quantitative
 - *Score on the first exam in this course*: quantitative
- f. No, neither *average height of students in the class* nor *percentage of students in the class who have used a cell phone today* can legitimately be considered variables when the observational units are the students in your class. Both of these are numbers that provide summary information about the class as a whole. They do not vary from student to student.
- g. If you record the average student height or percentage of student cell phone usage *by class* taught at your school, these would become legitimate variables. Now these numbers would (potentially) take on different values from class to class. The observational units are no longer the students in your class, but rather all classes taught at your school.

Activity 1-2: Variables of State

- a. Binary categorical variable
- b. Not a variable
- c. Quantitative variable
- d. Quantitative variable
- e. Binary categorical variable
- f. Quantitative variable
- g. Not a variable

Activity 1-3: Cell Phone Fraud

- The observational units are the cell phone calls.
- The binary categorical variables are *direction*, *location*, and *whether the call took place on a weekday or weekend*. The non-binary categorical variable is *day of week*.
- The quantitative variables are *duration of the call* and *time of day*.

Activity 1-4: Studies from *Blink*

- | | | |
|-----------|---|----------------------------|
| a. | Observational units: 100 CEOs | |
| | Variable: height | Type: quantitative |
| b. | Observational units: 50 marriage counselors | |
| | Variable: whether the counselor makes the correct prediction about whether a couple will still be married in five years | Type: (binary) categorical |
| c. | Observational units: 200 African-American college students | |
| | Variable 1: whether their version of the exam asks them to indicate race | Type: (binary) categorical |
| | Variable 2: score on SAT-like exam | Type: quantitative |
| d. | Observational units: 10 car dealerships | |
| | Variable 1: gender of customer | Type: (binary) categorical |
| | Variable 2: race of customer | Type: (binary) categorical |
| | Variable 3: price negotiated for the car | Type: quantitative |

Activity 1-5: Student Data

- Many answers are possible, but some examples include these: Do male and female students differ with regard to the number of states they have visited? Are sleeping times associated with the day of the week on which a student was born?
- Answers will vary.

Activity 1-6: A Nurse Accused

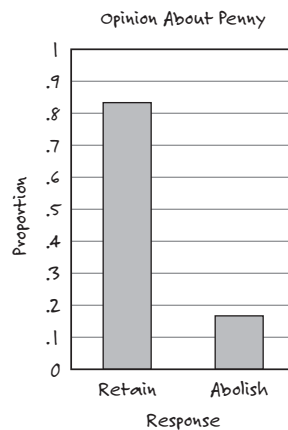
- a. The observational units are the eight-hour shifts.
- b. One variable is *whether Gilbert worked on the shift*. This variable is categorical and binary. The other variable is *whether a patient died on the shift*. This variable is also categorical and binary.

Solutions

● ● ● In-Class Activities

Activity 2-1: Penny Thoughts

- This is a binary categorical variable.
- Answers will vary by class. One example is 15/18 or .833 voted to retain the penny.
- Answers will vary, but for the class in part b, 3/18 or .167 voted to abolish the penny.
- Answers will vary, but for the class in part b, the bar graph is shown here:



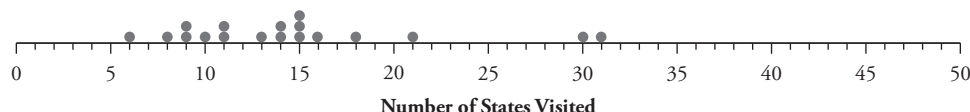
- For the class in part c, more than 80% favors retaining the penny, but the vote is not unanimous. A significant proportion of the class (.167) favors abolishing the penny.

Activity 2-2: Hand Washing

- The proportion of men who washed their hands is 2393/3206 or .746. The proportion of women who washed their hands is 2802/3130 or .895.
- Yes, these proportions are consistent with the bar graphs. The heights of the "washed hands" bars are about .75 and .90.
- It does appear that women are a little more likely to wash their hands after using a public restroom than men are. About 90% of the women in this sample did, whereas only 75% of the men did, indicating that women are about 1.2 times more likely to wash their hands after using a public restroom.
- Atlanta: .73 Chicago: .88 New York: .79 San Francisco: .88
- There does not appear to be much difference between these cities in terms of hand washing. Chicago and San Francisco appear to have identical proportions of people who washed their hands after using a public restroom, whereas New York seems to have about 10% fewer hand washers. Atlanta seems to have the smallest percentage of hand washers among these four cities, with just under 75% of the sample from Atlanta washing their hands.

Activity 2-3: Student Travels

- This is a quantitative variable.
- Answers will vary. Here is one example:



- Answers will vary.
- Answers will vary.
- The number of states visited by students in this class varies from a minimum of 6 to about 21. There are two high outliers of 30 and 31 states. The typical number of states visited by a student in this class seems to be about 14.

Activity 2-4: Buckle Up!

- The states are the observational units for these data.
- The *primary or secondary law* variable is categorical and binary. The *percentage usage* variable is quantitative.
- Answers will vary. The typical usage percentage for a primary-type seatbelt law state is about 86%; for states with a secondary-type law, the typical usage percentage appears to be about 77%.
- No, a state with a primary law does not always have a higher usage percentage than a state with a secondary law; for example, Tennessee (p, 74.4%) and Virginia (s, 8.4%).
- Yes, states with a primary law *tend to* have higher usage percentages than states with a secondary law. You see this in the dotplot because most of the dots for the primary law states are clustered at the high percentage values (from 80–95%), whereas most of the secondary law states have percentages that fall between 65%–85%.
- Yes, the data seem to support the contention that tougher laws lead to more seatbelt usage, but you cannot draw a definite cause-and-effect conclusion. There might be hidden or “confounding” variables that explain the association between primary seatbelt laws and increased seatbelt usage.

Activity 2-5: February Temperatures

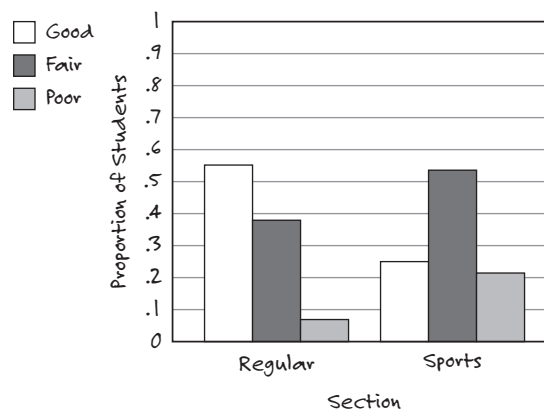
- San Luis Obispo tended to have the highest temperatures that month (its temperatures are all clustered from 50–90°F, with a significant chunk above 75°F), and Lincoln tended to have the lowest temperatures, with many temperatures below 45°F, whereas Sedona’s temperatures all ranged from about 48–68°F.
- Sedona had the most day-to-day consistency in its high temperatures that month (the temperatures stayed between 48–68°F all month), whereas Lincoln had the least consistency because its temperatures ranged from about 10°F all the way to about 75°F.

Activity 2-6: Sporting Examples

- a. The observational units are the students enrolled in one or the other of these sections.
- b. The variables are *section* (categorical, binary), *grade* (categorical, not binary), and *total points earned* (quantitative).
- c. Students in the regular section tended to score more points than those in the sports section. Scores in the regular section appear to be centered around 340 (85% of the possible points), whereas those in the sports section are centered around 310–320 points (a bit less than 80% of the possible points). Scores in the sports section are more spread out than those in the regular section. Students in the sports section had the six lowest scores, all less than 260 points, but that section also had the highest overall score, greater than 390 points.
- d. Students in the regular section tended to score more points than those in the sports section. Most students in the regular section scored between 300–380 points, with a center of approximately 340 points. In contrast, many students in the sports section scored less than 300 points, and the center was approximately 310–320 points.
- e. No, some sports students scored more points than some regular students. The statistical tendency means that a typical student in the regular section scored more points than a typical student in the sports section.
- f. The proportions are found by dividing the counts by 29 for the regular section and by 28 for the sports section. These proportions are

Regular section:	.552 good	.379 fair	.069 poor
Sports section:	.250 good	.536 fair	.214 poor

- g. The bar graphs follow:



- h. The bar graphs reveal similar results to the dotplots: Students in the regular section tended to score higher than those in the sports section. More than half of the regular students were in the good category, compared to only one-fourth of the students in the sports section. At the other extreme, only 6.9% of the regular students did poor work, compared to 21.4% of sports students.
- i. You cannot draw a cause-and-effect conclusion between the type of section and student performance. You will study these issues again in the next topic, but one key is that students self-selected which section to take. Perhaps those who chose to take the sports section had lower academic aptitude than those who selected the