Contents

Part V From the Data at Hand to the World at Large

Part VI Learning About the World

Part VII Inference When Variables are Related

Chapter 1 – Stats Starts Here

1. **Voters.** The response, party affiliation, is a categorical variable.

2. **Job growth.** The response, change in workforce size, is a categorical variable.

3. **Medicine.** The company is measuring minutes, a quantitative variable.

4. **Stress.** The researcher is measuring heart rate, a quantitative variable.

5. **The news.** Answers will vary.

6. **The Internet.** Answers will vary.

7. **Bicycle and pedestrian safety.** *Who* – pedestrians and bicyclists killed or severely injured in New York City between 2010 and 2014. *What* – proportion of pedestrians and bicyclists killed or injured by left-turning and right-turning vehicles. *Population of interest* – Answers may vary. Perhaps: All pedestrians and bicyclists killed or severely injured in New York City between 2010 and 2014 *or* all pedestrians and bicyclists killed or severely injured in New York City in any year *or* all pedestrians and bicyclists killed or severely injured in major U.S. cities. Discuss with students why these and other solutions may be problematic.

8. **Investments.** *Who* – 30 companies with similar retirement plans. *What* – 401(k) employee participation rates. *Population of interest* – All companies with similar retirement plans.

9. **Fake news.** *Who* – Middle school, high school, and college students in 12 states. *What* – Ability to evaluate the quality of information found in different online resources. *Population of interest* – All U.S middle school, high school, and college students.

10. **Biological instinct.** *Who* – 40 undergraduate women. *What* – Whether or not the women could identify the sexual orientation of men based on a picture. *Population of interest* – All women.

11. **Blindness.** *Who* – 24 patients. *What* – Whether or not stem cell therapy was effective in treating Stargardt's disease and/or dry age-related macular degeneration. *Population of interest* – All people with these eye conditions.

12. **Molten iron.** *Who* – 10 castings at Cleveland Casting. *What* – The pouring temperature (in degrees Fahrenheit) of molten iron. *Population of interest* – All castings at Cleveland Casting.

13. Weighing bears. *Who* – 54 bears. *What* – Weight, neck size, length (no specified units), and sex. *When* – Not specified. *Where* – Not specified. *Why* - Since bears are difficult to weigh, the researchers hope to use the relationships between weight, neck size, length, and sex of bears to estimate the weight of bears, given the other, more observable features of the bear.
How – Researchers collected data on 54 bears they were able to catch. *Variables* – There are 4 variables: weight, neck size, and length are quantitative variables, and sex is a categorical variable. No units are specified for the quantitative variables. *Concerns* – The researchers are (obviously!) only able to collect data from bears they were able to catch. This method is a good one, as long as the researchers believe the bears caught are representative of all bears, in regard to the relationships between weight, neck size, length, and sex.

14. Schools. *Who* – Students. *What* – Age (probably in years, though perhaps in years and months), race or ethnicity, number of absences, grade level, reading score, math score, and disabilities/special needs. *When* – This information must be kept current. *Where* – Not specified. *Why* – Keeping this information is a state requirement. *How* – The information is collected and stored as part of school records. *Variables* – There are seven variables. Race or ethnicity, grade level, and disabilities/special needs are categorical variables. Number of absences, age, reading test score, and math test score are quantitative variables. *Concerns* – What tests are used to measure reading and math ability, and what are the units of measure for the tests?

15. Arby's menu. *Who* – Arby's sandwiches. *What* – type of meat, number of calories (in calories), and serving size (in ounces). *When* – Not specified. *Where* – Arby's restaurants. *Why* – These data might be used to assess the nutritional value of the different sandwiches. *How* – Information was gathered from each of the sandwiches on the menu at Arby's, resulting in a census. *Variables* – There are three variables. Number of calories and serving size are quantitative variables, and type of meat is a categorical variable.

16. Party and the environment. *Who* – American voters. *What* – Gender, age (in years), race, party affiliation, education, whether or not the person was "worried a great deal" about climate change, air pollution, and pollution of waterways. *When* – 2017. *Where* – United States. *Why* – The information was gathered for presentation in a Gallup public opinion poll. *How* – Poll. *Variables* – There are eight variables. Gender, race, party affiliation, education, whether or not the person was "worried a great deal" about climate change, air pollution, and pollution of waterways are categorical variables; age is a quantitative variable.

17. Babies. *Who* – 882 births. *What* – Mother's age (in years), length of pregnancy (in weeks), type of birth (caesarean, induced, or natural), level of prenatal care (none, minimal, or adequate), birth weight of baby (unit of measurement not specified, gender of baby (male or female), and baby's health problems (none, minor, major). *When* – 1998-2000. *Where* – Large city hospital. *Why* – Researchers were investigating the impact of prenatal care on newborn health. *How* – It is not specifically stated. *Variables* – There are seven variables. Type of birth, level of prenatal care, gender of baby, and baby's health problems are categorical variables; mother's age, length of pregnancy, and birth weight of baby are quantitative variables.

18. Flowers. *Who* – 385 species of flowers for 47 years. 385(47) = 18,095 cases. *What* – Date of first flowering (in days). *When* – An unspecified 47 year period. *Where* – Southern England. *Why* – The researchers believe that this indicates a warming of the overall climate. *How* – Not specified. *Variables* – Date of first flowering is a quantitative variable. The number of years, 47, is also a variable.

19. Herbal medicine. *Who* – Patients. *What* – Herbal cold remedy or sugar solution, and cold severity on a scale of 0-5. *When* – Not specified. *Where* – Major pharmaceutical firm. *Why* – Scientists were testing the efficacy of an herbal compound on the severity of the common cold. *How* – The scientists conducted an experiment. *Variables* – There are two variables. Type of treatment (herbal or sugar solution) is a categorical variable, and severity rating is a quantitative variable. The subjectivity of "cold severity" is a concern that should be raised about this study.

20. Vineyards. *Who* – Vineyards. *What* – Size of vineyard (in acres), number of years in existence, state, varieties of grapes grown, average case price (in dollars), gross sales (probably in dollars), and percent profit. *When* – Not specified. *Where* – United States. *Why* – Business analysts hoped to provide information that would be helpful to producers of American grapes. *How* – Not specified. *Variables* – There are seven variables. State and variety of grapes grown are categorical variables; size of vineyard, number of years in existence, average case price, gross sales, and percent profit are quantitative variables.

21. Streams. *Who* – Streams. *What* – A number of variables including: name of stream, substrate of the stream (limestone, shale, or mixed), acidity of the water (measured in pH), temperature (in degrees Celsius), and BCI (unknown units). *When* – Not specified. *Where* – Upstate New York. *Why* – Research is conducted for an ecology class. *How* – Not specified. *Variables* – There are five variables. Name and substrate of the stream are categorical variables; acidity, temperature, and BCI are quantitative variables.

22. Fuel economy. *Who* – Every model of automobile. *What* – Vehicle manufacturer, vehicle type, weight (probably in pounds), horsepower (in horsepower), and gas mileage (in miles per gallon) for city and highway driving. *When* – This information is collected currently. *Where* – United States. *Why* – The Environmental Protection Agency uses the information to track fuel economy of vehicles. *How* – The data is collected from the manufacturer of each model. *Variables* – There are six variables. Manufacturer and type of car are categorical variables; weight, horsepower, city mileage, and highway mileage are quantitative variables.

23. Refrigerators. *Who* – 148 models of French door style refrigerators. *What* – Brand, price (probably in dollars), temperature performance, temperature uniformity, energy efficiency, noise, ease of use (the five previous variables are measured as Excellent, Very Good, Good, Fair, or Poor), number of doors, capacity (cu. ft.), exterior height (in.), exterior width (in.), and exterior depth (in.). *When* – 2017. *Where* – Not stated. *Why* – The information was compiled to provide information to the readers of *Consumer Reports. How* – Not specified. *Variables* – There are 11 variables. Brand, temperature performance, temperature uniformity, energy efficiency, noise, and ease of use are categorical variables; price, number of doors, capacity, exterior height, exterior width, and exterior depth are quantitative variables.

24. Walking in circles. *Who* – 32 people. *What* – Sex, height, handedness, the number of yards walked before going out of bounds, and the side of the field on which the person walked out of bounds. *When* – Not specified. *Where* – Not specified. *Why* – The researcher was interested in whether people naturally walk in circles when lost. *How* – Data were collected by observing the people on the field, as well as by measuring and asking the participants. *Variables* – There are 5 variables. Sex, handedness, and side of the field are categorical variables; height and number of yards walked are quantitative variables.

25. Kentucky Derby 2016. *Who* – Kentucky Derby races. *What* – Year, winner, jockey, trainer, owner, and time (in minutes, seconds, and hundredths of a second. *When* – 1875 – 2016. *Where* – Churchill Downs, Louisville, Kentucky. *Why* –Not specified. *How* – Official statistics are kept for the race each year. *Variables* – There are 6 variables. Winner, jockey, trainer and owner are categorical variables; year and time are quantitative variables.

26. Indy 2016. *Who* – Indy 500 races. *What* – Year, winner, time (in minutes, seconds, and hundredths of a second), and average speed (in miles per hour). *When* – 1911 – 2016. *Where* – Indianapolis Motor Speedway. *Why* – Not specified. *How* – Official statistics are kept for the race every year. *Variables* – There are 4 variables. Winner is a categorical variable, while year, time, and average speed are quantitative variables.

Chapter 2 – Displaying and Describing Categorical Data

1. **Graphs in the news.** Answers will vary.

2. **Graphs in the news II.** Answers will vary.

3. **Tables in the news.** Answers will vary.

4. **Tables in the news II.** Answers will vary.

5. **Movie genres.**

   a) A pie chart seems appropriate from the movie genre data. Each movie has only one genre, and the 891 movies constitute a "whole".

   b) "Other" is the least common genre. It has the smallest region in the chart.

6. **Movie ratings.**

   a) A pie chart seems appropriate for the movie rating data. Each movie has only one rating, and the 891 movies constitute a "whole".

   b) The most common rating is R. It has the largest region on the chart.

7. **Movie genres again.**

   a) Thriller/Suspense films were more common than Adventure films. The bar for Thriller/Suspense is taller than the bar for Adventure.

   b) This is easier to see on the bar chart. The percentages are so close that the difference is nearly indistinguishable in the pie chart. Also, the bar chart is organized by height while the pie chart is not, making it difficult to compare genres with areas similar in proportion.

8. **Movie ratings again.**

   a) The least common rating was NC-17. It has the shortest bar.

   b) While it is easy in both the pie chart and the bar chart, it may be easier in the pie chart. In the pie chart, ratings are ordered clockwise by increasing area while in the bar chart, G and NC-17 are inconsistent with an increasing order.

9. **Movie ratings.**

   i. C (This chart has 4 ratings and the proportion of G ratings is smallest.)

   ii. A (This chart has 4 ratings and a slightly higher proportion of G ratings.)

   iii. D (This chart has 3 ratings with PG more common than R.)

   iv. B (This chart has 3 ratings with PG and R roughly equally as common.)

10. **Marriage in decline.**

    i. D (This bar and pie chart are the only ones for which Bad Thing and No Difference are not roughly equally common)

ii. A (Bad Thing and No Difference are roughly equal, while Don't Know/No Response is noticeably the least frequent response)

iii. B or C (These charts are indistinguishable.)

iv. B or C (These charts are indistinguishable.)

11. Magnet schools.

There were 1,755 qualified applicants for the Houston Independent School District's magnet schools program. Approximately 53% were accepted, 17% were wait-listed, and the other 30% were turned away for lack of space.

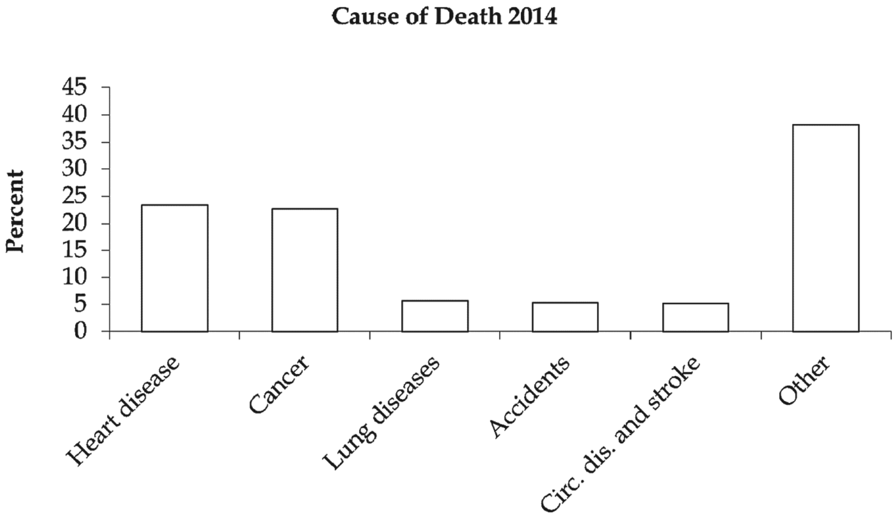
12. Magnet schools again.

There were 1,755 qualified applicants for the Houston Independent School District's magnet schools program. Approximately 29.5% were Black or Hispanic, 16.6% were Asian, and 53.9% were white.

13. Causes of death 2014.

a) Yes, it is reasonable to assume that heart or lung diseases caused approximately 29% of U.S. deaths in 2014, since there is no possibility for overlap. Each person could only have one cause of death.

b) Since the percentages listed add up to 61.9%, other causes must account for 38.1% of US deaths.



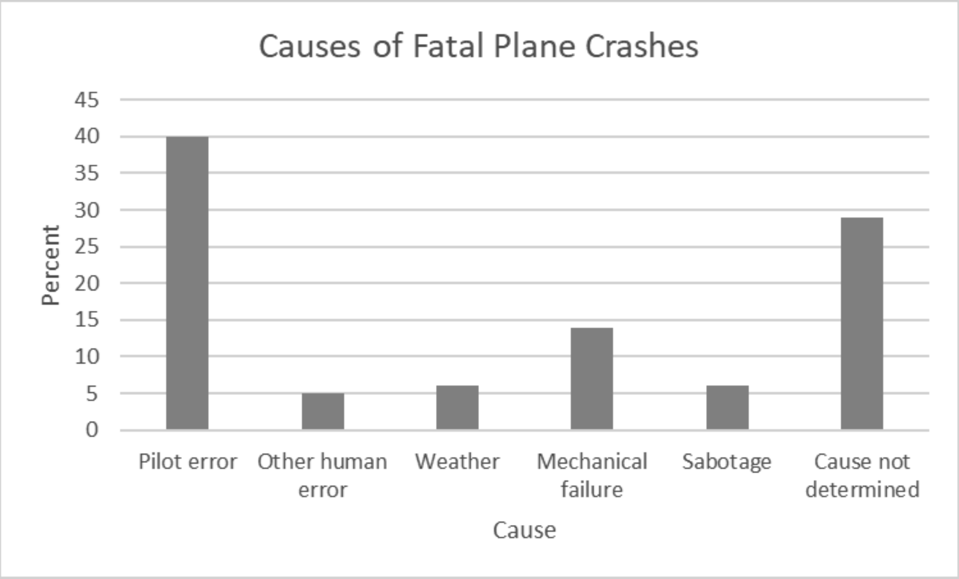
c) A bar chart is a good choice (with the inclusion of the "Other" category). Since causes of US deaths represent parts of a whole, a pie chart would also be a good display.

14. Plane crashes.

a) As long as each plane crash had only one cause, it would be reasonable to assume that weather or mechanical failures were the causes of about 20% of recent plane crashes.

b) Since the percentages listed add up to 71%, other causes (not determined) must account for 29% of recent plane crashes.

c) A relative frequency bar chart is a good choice. A pie chart would also be a good display, as long as each plane crash has only one cause.



15. Oil spills as of 2016.

a) Grounding, accounting for 150 spills, is the most frequent cause of oil spillage for these 460 spills. A substantial number of spills, 136, were caused by collision. Less prevalent causes of oil spillage in descending order of frequency were hull failures, other/unknown causes, fire/explosions, and equipment failure.

b) If being able to differentiate between these close counts is required, use the bar chart. Since each spill only has one cause, the pie chart is also acceptable as a display, but it's difficult to tell whether, for example, there is a greater percentage of spills caused by fire/explosions or hull failure. If you want to showcase the causes of oil spills as a fraction of all 460 spills, use the pie chart.

16. Winter Olympics 2016.

a) There are too many categories to construct an appropriate display. In a bar chart, there are too many bars. In a pie chart, there are too many slices. In each case, we run into difficulty trying to display those countries that didn't win many medals.

b) Perhaps we are primarily interested in countries that won many medals. We might choose to combine all countries that won fewer than 6 medals into a single category. This will make our chart easier to read. We are probably interested in number of medals won, rather than percentage of total medals won, so we'll use a bar chart. A bar chart is also better for comparisons.

17. Global warming.

Perhaps the most obvious error is that the percentages in the pie chart add up to 141%, when they should, of course, add up to 100%. This means that each individual region and any resulting sums will occupy less area of the display than their percentages imply. Furthermore, the three-dimensional perspective view distorts the regions in the graph, violating the area principle. The regions corresponding to "Global warming isn't happening" and "Can't reduce global warming even if it is happening" should be the same size, at 25% of respondents. However, the "Global warming isn't happening" region looks bigger. Always use simple, two-dimensional graphs.

18. Modalities.

a) The bars have false depth, which can be misleading. This is a bar chart, so the bars should have space between them.

b) Since each trainer was asked to list 3 modalities, the expected sum should be 300% rather than 100%.

19. Teen smokers.

According to the Monitoring the Future study, teen smoking brand preferences differ somewhat by region. Although Marlboro is the most popular brand in each region, with about 58% of teen smokers preferring this brand in each region, teen smokers from the South prefer Newports at a higher percentage than teen smokers from the West, 22.5% to approximately 10%, respectively. Camels are more popular in the West, with 9.5% of teen smokers preferring this brand, compared to only 3.3% in the South. Teen smokers in the West are also more likely to have to particular brand than teen smokers in the South. 12.9% of teen smokers in the West have no particular brand, compared to only 6.7% in the South. Both regions have 9% of teen smokers that prefer one of over 20 other brands.

20. Handguns.

76% of handguns involved in Milwaukee buyback programs are small caliber, while only 20.3% of homicides are committed with small caliber handguns. Along the same lines, only 19.3% of buyback handguns are of medium caliber, while 54.7% of homicides involve medium caliber handguns. A similar disparity is seen in large caliber handguns. Only 2.1% of buyback handguns are large caliber, but this caliber is used in 10.8% of homicides. Finally, 2.2% of buyback handguns are of other calibers, while 14.2% of homicides are committed with handguns of other calibers. Generally, the handguns that are involved in buyback programs are not the same caliber as handguns used in homicides in Milwaukee.

21. Movie genres and ratings.

a) 452 of these films were rated R. 452/1,529 ≈ 29.56%

b) 124 of these films were R-rated comedies. 124/1,529 ≈ 8.1%

c) 124 of the 452 R-rated films were comedies. 124/452 ≈ 27.43%

d) 124 of the 312 comedies were R-rated. 124/312 ≈ 39.74%

22. Not the labor force.

a) 2207 of the unemployed population were available to work now. 2207/12,872 ≈ 17.1%

b) 1,048 of the unemployed population were available to work now and aged 25 to 54 years. 1,048/12,872 ≈ 8.14%

c) 208 of 4,158 unemployed 16-24 year olds were in school or training. 208/4,158 ≈ 5%

d) 4,158 of the unemployed population were aged 16-24 years. 4,158/12,872 ≈ 32.3%

23. Seniors.

A table with marginal totals is to the right.

Plans	White	Minority	TOTAL
4-year college	198	44	242
2-year college	36	6	42
Military	4	1	5
Employment	14	3	17
Other	16	3	19
TOTAL	268	57	325

a) 268 seniors were white. 268/325 ≈ 82.5%

b) 42 seniors are planning to attend a 2-year college. 42/325 ≈ 13%

c) 36 seniors are white and planning to attend 2-year colleges. 36/325 ≈ 11.1%

d) 36 of the 268 white seniors are planning to attend 2-year colleges. 36/268 ≈ 13.4%

e) There are 42 graduates planning to attend 2-year colleges. 36 are white. 36/42 ≈ 85.7%

24. Politics.

a) There are 192 students taking Intro Stats. Of those, 115 are male. 115/192 ≈ 59.9%.

b) 27 students in the course consider themselves to be "Conservative". 27/192 ≈ 14%.

c) There are 115 males taking Intro Stats. Of those, 21 consider themselves to be "Conservative". 21/115 ≈ 18.26%.

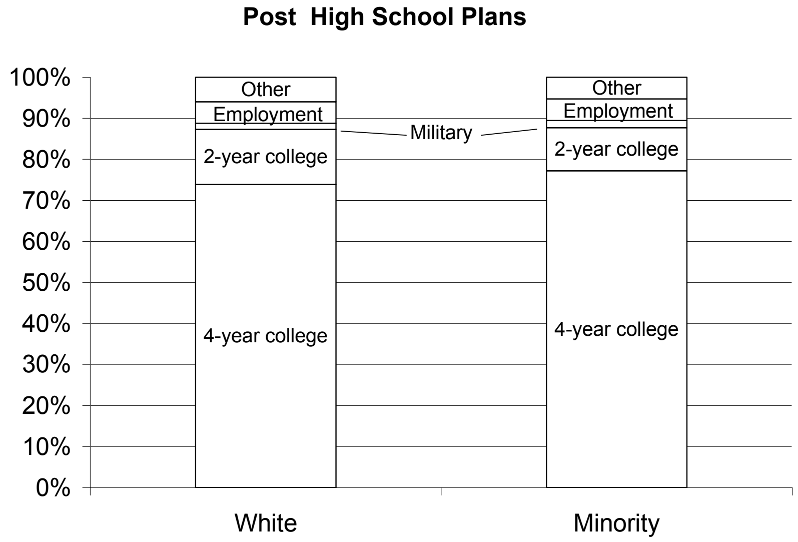
d) 21 of the students in the course are males who consider themselves to be "Conservative". 21/192 ≈ 10.94%

25. More about seniors.

a) For white students, 73.9% plan to attend a 4-year college, 13.4% plan to attend a 2-year college, 1.5% plan on the military, 5.2% plan to be employed, and 6.0% have other plans.

b) For minority students, 77.2% plan to attend a 4-year college, 10.5% plan to attend a 2-year college, 1.8% plan on the military, 5.3% plan to be employed, and 5.3% have other plans.

c) A segmented bar chart is a good display of these data:



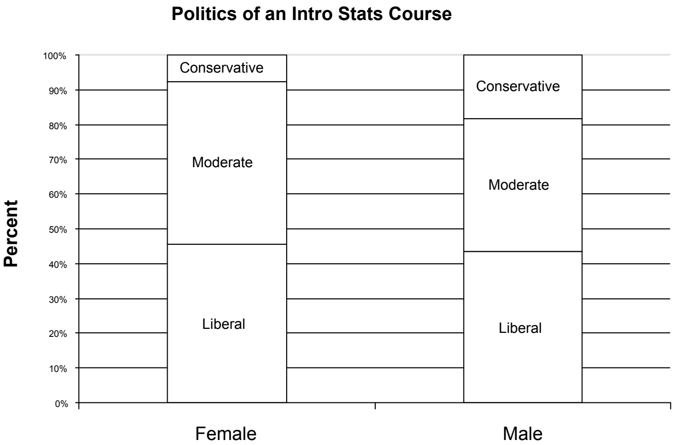
d) The conditional distributions of plans for Whites and Minorities are similar:
White – 74% 4-year college, 13% 2-year college, 2% military, 5% employment, 6% other.
Minority – 77% 4-year college, 11% 2-year college, 2% military, 5% employment, 5% other.
Caution should be used with the percentages for Minority graduates, because the total is so small. Each graduate is almost 2%. Still, the conditional distributions of plans are essentially the same for the two groups. There is little evidence of an association between race and plans for after graduation.

26. Politics revisited.

a) The females in this course were 45.5% Liberal, 46.8% Moderate, and 7.8% Conservative.

b) The males in this course were 43.5% Liberal, 38.3% Moderate, and 18.3% Conservative.

c) A segmented bar chart comparing the distributions is at the right.



d) Politics and sex do not appear to be independent in this course. Although the percentage of liberals was roughly the same for each sex, females had a greater percentage of moderates and a lower percentage of conservatives than males.

27. Magnet schools revisited.

a) There were 1,755 qualified applicants to the Houston Independent School District's magnet schools program. Of those, 292 were Asian. 292/1,755 ≈ 16.6%

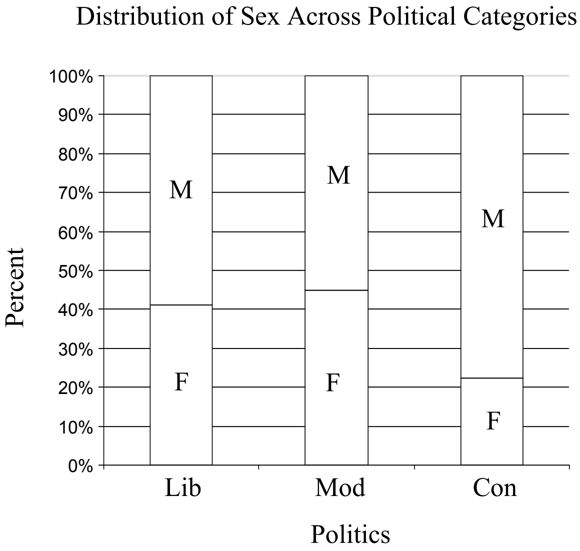
b) There were 931 students accepted to the magnet schools program. Of those, 110, were Asian. 110/931 ≈ 11.8%.

c) There were 292 Asian applicants. Of those, 110 were accepted. 110/292 ≈ 37.7%.

d) There were 1,755 total applicants. Of those, 931 were accepted. 931/1,755≈ 53%.

28. More politics.

a)



b) The percentage of males and females varies across political categories. The percentage of self-identified Liberals and Moderates who are female is about twice the percentage of Conservatives who are female. This suggests that *sex* and *politics* are not independent.

29. Back to school.

There were 1,755 qualified applicants for admission to the magnet schools program. 53% were accepted, 17% were wait-listed, and the other 30% were turned away. While the overall acceptance rate was 53%, 93.8% of Blacks and Hispanics were accepted, compared to only 37.7% of Asians, and 35.5% of whites. Overall, 29.5% of applicants were Black or Hispanics, but only 6% of those turned away were Black or Hispanic. Asians accounted for 16.6% of applicants, but 25.3% of those turned away. It appears that the admissions decisions were not independent of the applicant's ethnicity.

30. Cars.

a) In order to get percentages, first we need totals. Here is the same table, with row and column totals. Foreign cars are defined as non-American. There are 45+102=147 non-American cars or 147/359 ≈ 40.95%.

	Driver		
Origin	Student	Staff	Total
American	107	105	212
European	33	12	45
Asian	55	47	102
Total	195	164	359

b) There are 212 American cars of which 107 or 107/212 ≈ 50.47% were owned by students.

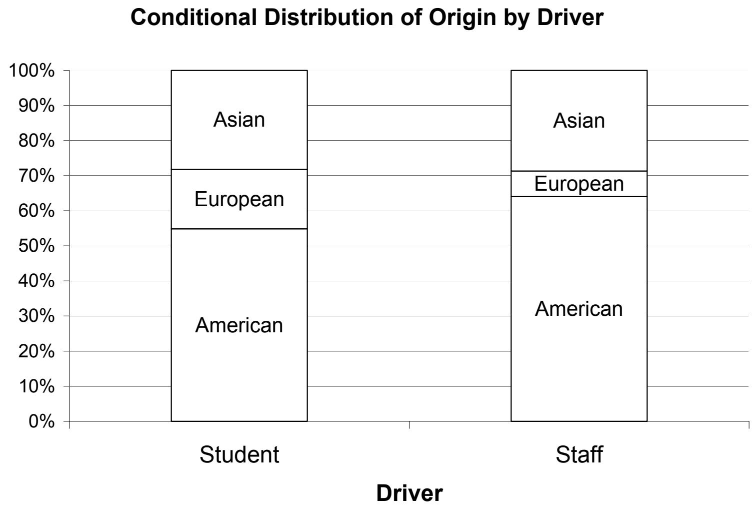
c) There are 195 students of whom 107 or 107/195 ≈ 54.87% owned American cars.

d) The marginal distribution of Origin is displayed in the third column of the table at the right: 59% American, 13% European, and 28% Asian.

Origin	Totals
American	212 (59%)
European	45 (13%)
Asian	102 (28%)
Total	359

e) The conditional distribution of Origin for Students is: 54.8% (107 of 195) American, 17% (33 of 195) European, and 28% (55 of 195) Asian.
The conditional distribution of Origin for Staff is:
64% (105 of 164) American, 7% (12 of 164) European, and 29% (47 of 164) Asian.

f) The percentages in the conditional distributions of Origin by Driver (students and staff) seem slightly different. Let's look at a segmented bar chart of Origin by Driver, to compare the conditional distributions graphically.



The conditional distributions of Origin by Driver have similarities and differences. Although students appear to own a higher percentage of European cars and a smaller percentage of American cars than the staff, the two groups own nearly the same percentage of Asian cars. However, because of the differences, there is evidence of an association between Driver and Origin of the car.

31. Super students.

a) Some possible observations: The most common choice of super power among students was the ability to Fly at just over 30%. The least common choice was Super Strength at less than 10%. Freeze Time and Telepathy were both chosen by just over 20% of students, the second most frequent choices.

b) It is likely that the ability to Fly is characteristic of the whole population because it was chosen both most frequently and just over 30% in each of the samples. Super Strength is likely to be the least popular choice in the population as it was the least selected characteristic in two of the three other samples. In the third sample, it was only slightly more popular than Invisibility but still by only 10% of students. It is also likely that Freeze Time and Telepathy are equally desirable and just above 20%. Telepathy was consistently chosen by slightly more than 20%. Freeze Time was chosen by about 26% of students in one sample, but was consistently at 20% in the others.

32. Super students II.

a) Some possible observations: Males were about twice as likely to select Fly for their Super Power compared to females. Males were also more likely to choose Freeze Time or Super Strength than females. Females, on the other hand, were far more likely to select Telepathy than males, about 40% to 8% respectively.

b) It appears that observation that males were twice as likely to choose Fly compared to females was a quirk of the first sample. Subsequent samples suggest the females are equally as likely to choose Fly as males. The three new samples support the claim that males are more likely to choose Freeze Time and Super Strength compared to females. While females in the subsequent surveys were more likely to choose Telepathy compared to males, consistent with the first sample, the contrast may not be as severe as in two of the samples females were only twice as likely to choose Telepathy.

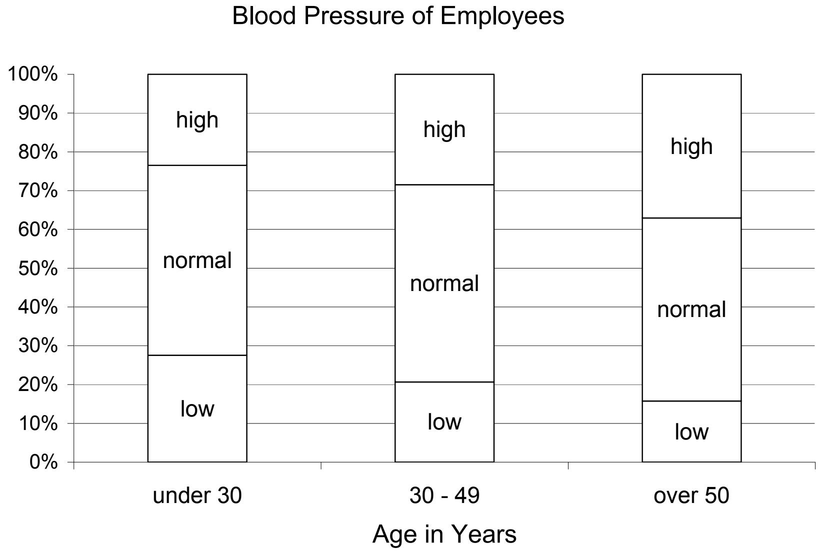
33. Blood pressure.

a) The marginal distribution of blood pressure for the employees of the company is the total column of the table, converted to percentages. 20% low, 49% normal and 31% high blood pressure.

Blood pressure	under 30	30 - 49	over 50	Total
low	27	37	31	95
normal	48	91	93	232
high	23	51	73	147
Total	98	179	197	474

b) The conditional distribution of blood pressure within each age category is:
Under 30 : 28% low, 49% normal, 23% high
30 – 49 : 21% low, 51% normal, 28% high
Over 50 : 16% low, 47% normal, 37% high

c) A segmented bar chart of the conditional distributions of blood pressure by age category is below.

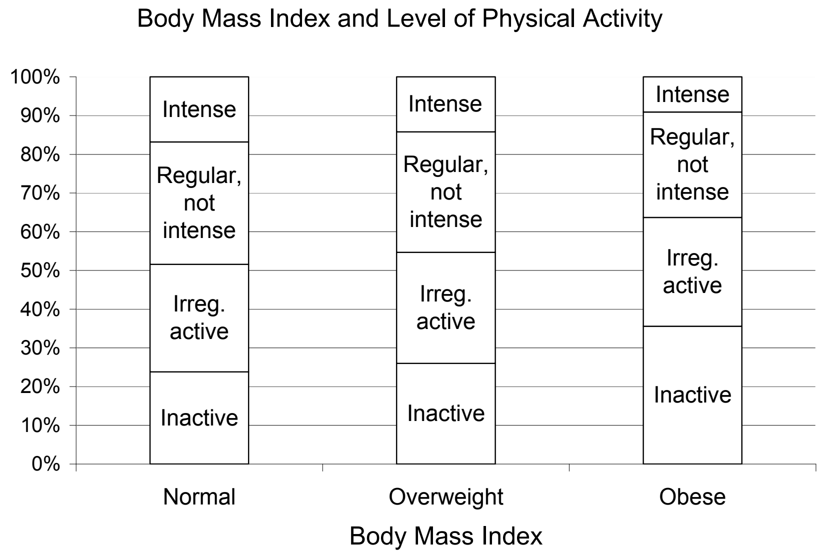


d) In this company, as age increases, the percentage of employees with low blood pressure decreases, and the percentage of employees with high blood pressure increases.

e) No, this does not prove that people’s blood pressure increases as they age. Generally, an association between two variables does not imply a cause-and-effect relationship. Specifically, these data come from only one company and cannot be applied to all people. Furthermore, there may be some other variable that is linked to both age and blood pressure. Only a controlled experiment can isolate the relationship between age and blood pressure.

34. Obesity and exercise.

a) Participants were categorized as Normal, Overweight or Obese, according to their Body Mass Index. Within each classification of BMI (column), participants self reported exercise levels. Therefore, these are column percentages. The percentages sum to 100% in each column, *not* across each row.



b) A segmented bar chart of the conditional distributions of level of physical activity by Body Mass Index category is at the right.

c) No, even though the graphical displays provide strong evidence that lack of exercise and BMI are not independent. All three BMI categories have nearly the same percentage of subjects who report “Regular, not intense” or “Irregularly active”, but as we move from Normal to Overweight to Obese we see a decrease in the percentage of subjects who report “Regular, intense” physical activity (16.8% to 14.2% to 9.1%), while the percentage of subjects who report themselves as “Inactive” increases. While it may seem logical that lack of exercise causes obesity, association between variables does not imply a cause-and-effect relationship. A lurking variable (for example, overall health) might influence both BMI and level of physical activity, or perhaps lack of exercise is *caused by* obesity. Only a controlled experiment could isolate the relationship between BMI and level of physically activity.

35. Anorexia.

These data provide no evidence that Prozac might be helpful in treating anorexia. About 71% of the patients who took Prozac were diagnosed as "Healthy", while about 73% of the patients who took a placebo were diagnosed as "Healthy". Even though the percentage was higher for the placebo patients, this does not mean that Prozac is hurting patients. The difference between 71% and 73% is not likely to be statistically significant.

36. Antidepressants and bone fractures.

These data provide evidence that taking a certain class of antidepressants (SSRI) might be associated with a greater risk of bone fractures. Approximately 10% of the patients taking this class of antidepressants experience bone fractures. This is compared to only approximately 5% in the group that were not taking the antidepressants.

37. Driver's licenses 2014.

a) There are 8.5 million drivers under 20 and a total of 214.1 million drivers in the U.S. 8.5/214.1 ≈ 3.97%

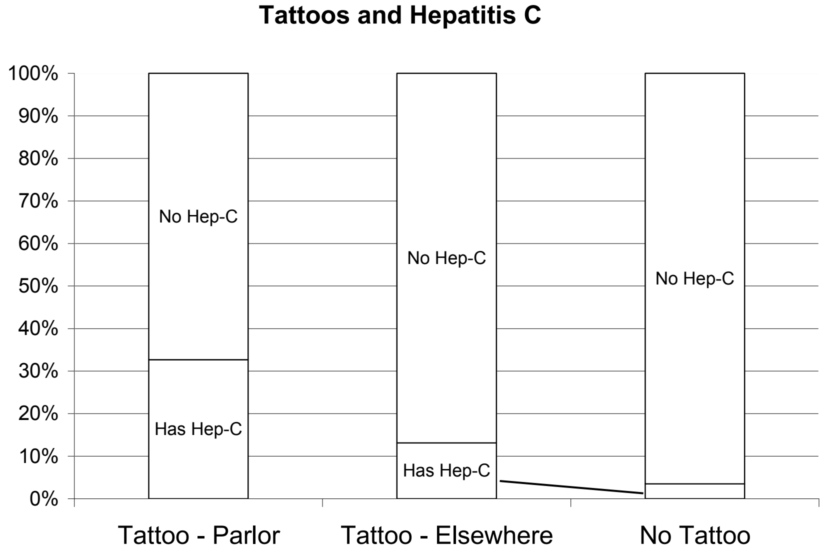
b) There are 105.9 million males out of 214.1 million U.S. drivers. 105.9/214.1 ≈ 49.46%

c) Each age category appears to have about 50% male and 50% female drivers. At the youngest ages, males form the slight majority of drivers. This percentage shrinks until the percentages are 50% male and 50% female for middle aged drivers. The percentage of male drivers continues to shrink until, at around age 45, female drivers hold a slight majority. This continues into the 85 and over category.

d) There is a slight association between age and gender of U.S. drivers. Younger drivers are slightly more likely to be male, and older drivers are slightly more likely to be female.

38. Tattoos.

The study by the University of Texas Southwestern Medical Center provides evidence of an association between having a tattoo and contracting hepatitis C. Around 33% of the subjects who were tattooed in a commercial parlor had hepatitis C, compared with 13% of those tattooed elsewhere, and only 3.5% of those with no tattoo. If having a tattoo and having hepatitis C were independent, we would have expected these percentages to be roughly the same.



39. Hospitals.

a) The marginal totals have been added to the table:

		Discharge delayed		
		Large Hospital	**Small Hospital**	**Total**
Procedure	**Major surgery**	120 of 800	10 of 50	130 of 850
	Minor surgery	10 of 200	20 of 250	30 of 450
	Total	130 of 1000	30 of 300	160 of 1300

160 of 1300, or about 12.3% of the patients had a delayed discharge.

b) Yes. Major surgery patients were delayed 130 of 850 times, or about 15.3% of the time.
Minor Surgery patients were delayed 30 of 450 times, or about 6.7% of the time.

c) Large Hospital had a delay rate of 130 of 1000, or 13%.
Small Hospital had a delay rate of 30 of 300, or 10%.
The small hospital has the lower overall rate of delayed discharge.

d) Large Hospital: Major Surgery 15% delayed and Minor Surgery 5% delayed. Small Hospital: Major Surgery 20% delayed and Minor Surgery 8% delayed. Even though small hospital had the lower overall rate of delayed discharge, the large hospital had a lower rate of delayed discharge for each type of surgery.

e) No. While the overall rate of delayed discharge is lower for the small hospital, the large hospital did better with *both* major surgery and minor surgery.

f) The small hospital performs a higher percentage of minor surgeries than major surgeries. 250 of 300 surgeries at the small hospital were minor (83%). Only 200 of the large hospital's 1000 surgeries were minor (20%). Minor surgery had a lower delay rate than major surgery (6.7% to 15.3%), so the small hospital's overall rate was artificially inflated. Simply put, it is a mistake to look at the overall percentages. The real truth is found by looking at the rates after the information is broken down by type of surgery, since the delay rates for each type of surgery are so different. The larger hospital is the better hospital when comparing discharge delay rates.

40. Delivery service.

a) Pack Rats has delivered a total of 28 late packages (12 Regular + 16 Overnight), out of a total of 500 deliveries (400 Regular + 100 Overnight). 28/500 = 5.6% of the packages are late. Boxes R Us has delivered a total of 30 late packages (2 Regular + 28 Overnight) out of a total of 500 deliveries (100 Regular + 400 Overnight). 30/500 = 6% of the packages are late.

b) The company should have hired Boxes R Us instead of Pack Rats. Boxes R Us only delivers 2% (2 out of 100) of its Regular packages late, compared to Pack Rats, who deliver 3% (12 out of 400) of its Regular packages late. Additionally, Boxes R Us only delivers 7% (28 out of 400) of its Overnight packages late, compared to Pack Rats, who delivers 16% of its Overnight packages late. Boxes R Us is better at delivering Regular and Overnight packages.

c) This is an instance of Simpson's Paradox, because the overall late delivery rates are unfair averages. Boxes R Us delivers a greater percentage of its packages Overnight, where it is comparatively harder to deliver on time. Pack Rats delivers many Regular packages, where it is easier to make an on-time delivery.

41. Graduate admissions.

a) 1284 applicants were admitted out of a total of 3014 applicants. 1284/3014 = 42.6%

Program	Males Accepted (of applicants)	Females Accepted (of applicants)	Total
1	511 of 825	89 of 108	600 of 933
2	352 of 560	17 of 25	369 of 585
3	137 of 407	132 of 375	269 of 782
4	22 of 373	24 of 341	46 of 714
Total	**1022 of 2165**	**262 of 849**	**1284 of 3014**

b) 1022 of 2165 (47.2%) of males were admitted. 262 of 849 (30.9%) of females were admitted.

c) Since there are four comparisons to make, the table at the right organizes the percentages of males and females accepted in each program. Females are accepted at a higher rate in every program.

Program	Males	Females
1	61.9%	82.4%
2	62.9%	68.0%
3	33.7%	35.2%
4	5.9%	7%

d) The comparison of acceptance rate within each program is most valid. The overall percentage is an unfair average. It fails to take the different numbers of applicants and different acceptance rates of each program. Women tended to apply to the programs in which gaining acceptance was difficult for everyone. This is an example of Simpson's Paradox.

42. Be a Simpson!

Answers will vary. The three-way table below shows one possibility. The number of local hires out of new hires is shown in each cell.

	Company A	**Company B**
Full-time New Employees	40 of 100 = 40%	90 of 200 = 45%
Part-time New Employees	170 of 200 = 85%	90 of 100 = 90%
Total	210 of 300 = 70%	180 of 300 = 60%

Chapter 3 – Displaying and Summarizing Quantitative Data

1. **Histogram.** Answers will vary.
2. **Not a histogram.** Answers will vary.
3. **In the news.** Answers will vary.
4. **In the news II.** Answers will vary.
5. **Thinking about shape.**

   a) The distribution of the number of speeding tickets each student in the senior class of a college has ever had is likely to be unimodal and skewed to the right. Most students will have very few speeding tickets (maybe 0 or 1), but a small percentage of students will likely have comparatively many (3 or more?) tickets.

   b) The distribution of player's scores at the U.S. Open Golf Tournament would most likely be unimodal and slightly skewed to the right. The best golf players in the game will likely have around the same average score, but some golfers might be off their game and score 15 strokes above the mean. (Remember that high scores are undesirable in the game of golf!)

   c) The weights of female babies in a particular hospital over the course of a year will likely have a distribution that is unimodal and symmetric. Most newborns have about the same weight, with some babies weighing more and less than this average. There may be slight skew to the left, since there seems to be a greater likelihood of premature birth (and low birth weight) than post-term birth (and high birth weight).

   d) The distribution of the length of the average hair on the heads of students in a large class would likely be bimodal and skewed to the right. The average hair length of the males would be at one mode, and the average hair length of the females would be at the other mode, since women typically have longer hair than men. The distribution would be skewed to the right, since it is not possible to have hair length less than zero, but it is possible to have a variety of lengths of longer hair.

6. **More shapes.**

   a) The distribution of the ages of people at a Little League game would likely be bimodal and skewed to the right. The average age of the players would be at one mode and the average age of the spectators (probably mostly parents) would be at the other mode. The distribution would be skewed to the right, since it is possible to have a greater variety of ages among the older people, while there is a natural left endpoint to the distribution at zero years of age.

b) The distribution of the number of siblings of people in your class is likely to be unimodal and skewed to the right. Most people would have 0, 1, or 2 siblings, with some people having more siblings.

c) The distribution of pulse rate of college-age males would likely be unimodal and symmetric. Most males' pulse rates would be around the average pulse rate for college-age males, with some males having lower and higher pulse rates.

d) The distribution of the number of times each face of a die shows in 100 tosses would likely be uniform, with around 16 or 17 occurrences of each face (assuming the die had six sides).

7. Cereals.

a) The distribution of the carbohydrate content of breakfast cereals is bimodal, with a cluster of cereals with carbohydrate content around 13 grams of carbs and another cluster of cereals around 22 grams of carbs. The lower cluster shows a bit of skew to the left. Most cereals in the lower cluster have between 10 and 20 grams of carbs. The upper cluster is symmetric, with cereals in the cluster having between 20 and 24 grams of carbs.

b) The cereals with the highest carbohydrate content are Corn Chex, Corn Flakes, Cream of Wheat (Quick), Crispix, Just Right Fruit & Nut, Kix, Nutri-Grain Almond-Raisin, Product 19, Rice Chex, Rice Krispies, Shredded Wheat 'n' Bran, Shredded Wheat Spoon Size, Total Corn Flakes, and Triples.

8. Singers.

a) The distribution of the heights of singers in the chorus is bimodal and symmetric, with a mode at around 65 inches and another mode around 71 inches. No chorus member has height below 60 inches or above 76 inches.

b) The two modes probably represent the mean heights of the male and female members of the chorus.

9. Vineyards.

a) There is information displayed about 36 vineyards and it appears that 28 of the vineyards are smaller than 60 acres. That's around 78% of the vineyards. (75% would be a good estimate!)

b) The distribution of the size of 36 Finger Lakes vineyards is unimodal and skewed to the right. The distribution is tightly centered between 0 and 60 acres. There is an unusually large vineyard over 240 acres in size. The range in size of vineyards is 240 acres including the unusual large vineyard, but only 180 acres if that vineyard is excluded.

10. Run times.

The distribution of runtimes is unimodal and skewed to the right. The shortest runtime was around 28.5 minutes and the longest runtime was around 35.5

minutes. A typical run time was between 30 and 31 minutes, and the majority of runtimes were between 29 and 32 minutes. It is easier to run slightly slower than usual and end up with a longer runtime than it is to run slightly faster than usual and end up with a shorter runtime. This could account for the skew to the right seen in the distribution.

11. Heart attack stays.

a) The distribution of length of stays is skewed to the right, so the mean is larger than the median.

b) The distribution of the length of hospital stays of female heart attack patients is bimodal and skewed to the right, with stays ranging from 1 day to 36 days. For one of the modes, there is a population of patients whose stay was only one day, possibly because the patient died. The second mode of distribution is centered around 8 days, with the majority of the hospital stays lasting between 2 and 15 days. There are a relatively few hospital stays longer than 27 days. The full range of the data appears to be around 35 days.

c) It is important to think about the data as two populations. For the first population, a single number will suffice: 1 day. For the second population, the median and IQR would be used to summarize the distribution of hospital stays since the distribution is strongly skewed.

12. Emails.

a) The distribution of the number of emails sent is skewed to the right, so the mean is larger than the median.

b) The distribution of the number of emails received from each student by a professor in a large introductory statistics class during an entire term is unimodal and skewed to the right, with the number of emails varying from 1 to 21 emails for a range of 20 emails. The distribution is centered at about 1 to 2 emails, with many students only sending 1 email. There is one highly unusual number of emails in the distribution, a student who sent 21 emails. The next highest number of emails sent was only 8. Ignoring the unusually high number, the range is only 7 emails.

c) The median and IQR would be used to summarize the distribution of the number of emails received, since the distribution is strongly skewed.

13. Super Bowl points 2017.

a) The median number of points scored in the first 51 Super Bowl games is 46 points.

b) The first quartile of the number of points scored in the first 51 Super Bowl games is 37 points. The third quartile is 56 points.

c) IQR = Q3 – Q1 = 56 – 37 = 19.
Using the Outlier Rule (1.5 IQRs beyond quartiles):

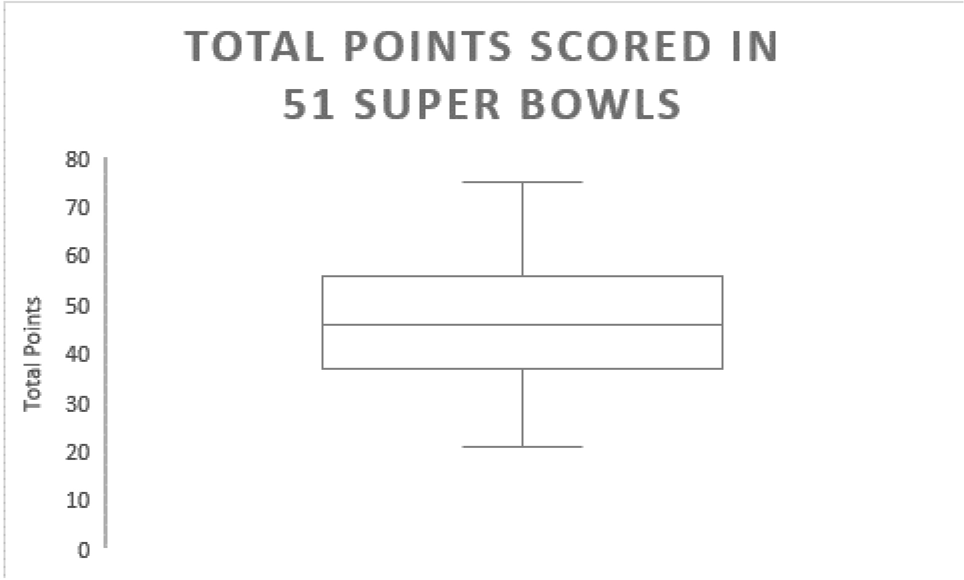
Upper Fence: Q3 + 1.5(IQR) = 56 + 1.5(19) = 84.5

The highest number of points scored in a Super Bowl is 75, which is lower than the upper fence of 84.5. There are no high outliers.

Lower Fence: Q1 – 1.5(IQR) = 37 – 1.5(19) = 8.5

The lowest number of points scored in a Super Bowl is 21, which is higher than the lower fence of 8.5. There are no low outliers.

d) A boxplot of the number of points scored in the first 51 Super Bowl games is below



e) The distribution of total scores from the first 51 Super Bowls appears to be symmetric, centered at 46 points, and with a range of 54 points (75 – 21).

14. Super Bowl wins 2017.

a) The median winning margin in the first 51 Super Bowl games is 12 points.

b) The first quartile of the winning margin in the first 51 Super Bowl games is 4 points. The third quartile is 19 points.

c) IQR = Q3 – Q1 = 19 – 4 = 15.
Using the Outlier Rule (1.5 IQRs beyond quartiles):

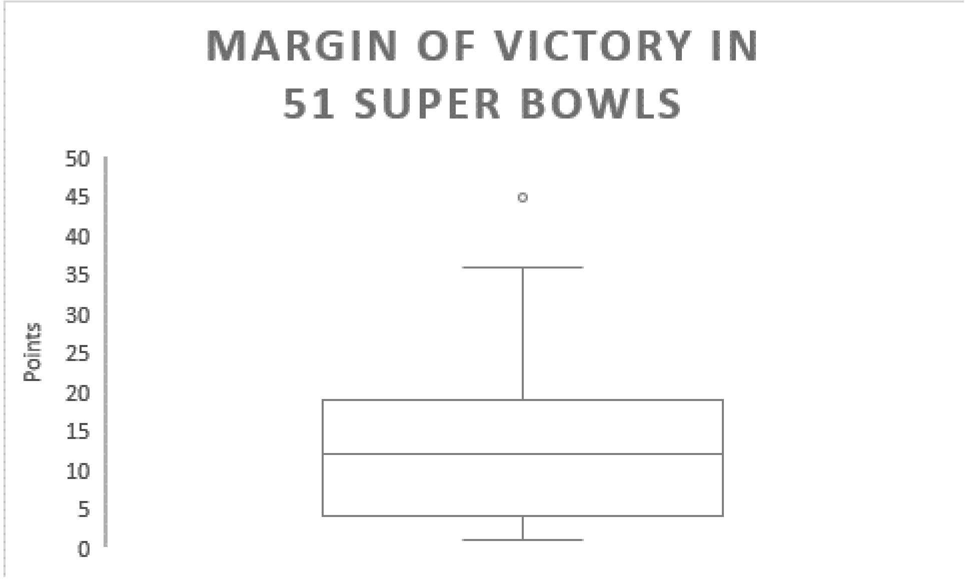
Upper Fence: Q3 + 1.5(IQR) = 19 + 1.5(15) = 41.5

The highest winning mardgin in a Super Bowl is 45 points, which is the only margin higher than the upper fence of 40. This winning margin is an outlier.

Lower Fence: Q1 1 1.5(IQR) = 4 - 1.5(15) = -18.5

The lowest winning margin in a Super Bowl is 1, which is higher than the lower fence of -18.5. There are no low outliers.

d) A boxplot of the number of points scored in the first 51 Super Bowl games is below.



e) The distribution of winning margins from the first 51 Super Bowls is skewed towards the higher numbers, centered at 12 points, and with a range of 44 points (45 – 1) including the unusually high margin. Without that margin, the range would be 35 points (36 – 1).

15 Details.

a) The boxplots do not show clusters and gaps, nor locations of multiple modes.

b) Boxplots can give only general ideas about overall shape, and should not be used when more detail is needed.

16. Opposites.

a) The tallest bars are in the narrow regions of the boxplot.

b) The histogram shows taller bars clustered on the left and smaller bars more spread out to the right. The boxplot shows narrower regions on the left, and a longer right side of the box and a long whisker stretching to the right, as well as some potential outliers on the right.

c) The histogram shows more precisely the locations of peaks and gaps.

d) The boxplot shows the locations of points flagged as outliers by the 1.5* IQR outlier rule.

17. Adoptions.

a) The histogram and boxplot of the distribution of adoptions both show that the vast number of states had 1000 or fewer adoptions. They both also show that a couple of states had an unusually large number of adoptions, causing the distribution to be skewed towards higher numbers of adoptions. Because 4 states are outliers, the gaps in the distribution are also evident in both distributions. Both displays show the range of approximately 5,500 adoptions.

b) This histogram shows that the distribution is unimodal and that there are no gaps in the number of adoptions for the 47 states without an unusual number of adoptions, given the bin size of 400.

c) Median is the better measure of center, since the distribution of adoptions has outliers in one direction. Median is more resistant to outliers than the mean.

d) IQR is a better measure of spread, since the distribution of adoptions has outliers. IQR is more resistant to outliers than the standard deviation.

18. Adoptions again.

a) Both displays show a high concentration in the number of adoptions per 100,000 around 13-14. Both displays also show that the distribution has some unusually high adoptions rates resulting in an overall skewed distribution towards the higher rates. Both displays also show a range of about 45 adoptions per 100,000.

b) The histogram shows that the distribution is unimodal and has two slightly unusually high rates around 29 and 35.

c) Since the distribution of adoption rates is skewed with high outliers, the median is the better summary of center.

d) IQR is a better measure of spread, since the distribution of adoption rates is skewed with high outliers. IQR is more resistant to outliers than the standard deviation.

19. Camp sites.

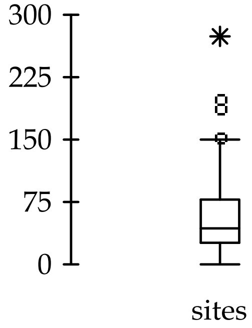
a) The distribution of the number of campsites in public parks in Vermont is skewed to the right, so median and IQR are appropriate measures of center and spread.

b) IQR = Q3 – Q1 = 78 – 28 = 50.
Using the Outlier Rule (1.5 IQRs beyond quartiles):

Upper Fence: $$\begin{aligned} Q3 + 1.5(IQR) &= 78 + 1.5(50) \\ &= 78 + 75 \\ &= 153 \end{aligned}$$

Lower Fence: Well below 0 campsites.

There are 3 parks with greater than 180 campsites. These are definitely outliers. There are 2 parks with between 150 and 160 campsites each. These may be outliers as well.



c) A boxplot of the distribution of number of campsites is at the right.

d) The distribution of the number of campsites at public parks in Vermont is unimodal and skewed to the right. The center of the distribution is approximately 44 campsites. The distribution of campsites is quite spread out, with several high outliers. These parks have in excess of 150 campsites each.

20. Outliers.

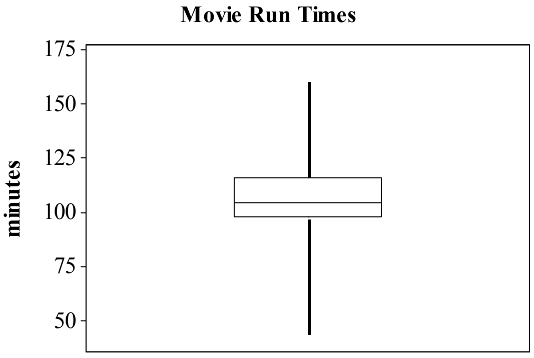
a) IQR = Q3 – Q1 = 116 – 98 = 18.
Using the Outlier Rule (1.5 IQRs beyond quartiles):

Upper Fence: Q3 + 1.5(IQR) = 116 + 1.5(18) = 143

The longest run time is 160 minutes, which is higher than the upper fence of 143. This run time is an outlier. There may be others.

Lower Fence: Q1 – 1.5(IQR) = 98 – 1.5(18) = 71
The shortest run time is 43 minutes, which is lower than the lower fence of 71. This run time is an outlier. There may be others.

b) We do not have enough information to know if there are other outliers, or where the last non-outlier point is on either side, so the boxplot can be drawn without outliers as shown. We can't say much about the shape of the distribution, since many different shapes could underlie the boxplot. The most we could say is that the distribution of movie run times is skewed slightly to the right.



21. Standard deviation I.

a) Set 2 has the greater standard deviation. Both sets have the same mean, 6, but set two has values that are generally farther away from the mean.
SD(Set 1) = 2.24 SD(Set 2) = 3.16

b) Set 2 has the greater standard deviation. Both sets have the same mean (15), maximum (20), and minimum (10), but 11 and 19 are farther from the mean than 14 and 16.
SD(Set 1) = 3.61 SD(Set 2) = 4.53

c) The standard deviations are the same. Set 2 is simply Set 1 + 80. Although the measures of center and position change, the spread is exactly the same.
SD(Set 1) = 4.24 SD(Set 2) = 4.24

22. Standard deviation II.

a) Set 2 has the greater standard deviation. Both sets have the same mean (7), maximum (10), and minimum (4), but 6 and 8 are farther from the mean than 7.
SD(Set 1) = 2.12 SD(Set 2) = 2.24

b) The standard deviations are the same. Set 1 is simply Set 2 + 90. Although the measures of center and position are different, the spread is exactly the same.
SD(Set 1) = 36.06 SD(Set 2) = 36.06

c) Set 2 has the greater standard deviation. The central 4 values of Set 2 are simply the central 4 values of Set 1 +40, but the maximum and minimum of Set 2 are farther away from the mean than the maximum and minimum of Set 1.
Range(Set 1) = 18 and Range(Set 2) = 22. Since the Range of Set 2 is greater than the Range of Set 1, the standard deviation is also larger.
SD(Set 1) = 6.03 SD(Set 2) = 7.24

23. Pizza prices.

The mean and standard deviation would be used to summarize the distribution of pizza prices, since the distribution is unimodal and symmetric.

24. Neck size.

The mean and standard deviation would be used to summarize the distribution of neck sizes, since the distribution is unimodal and symmetric.

25. Pizza prices again.

a) The mean pizza price is closest to $2.60. That's the balancing point of the histogram.

b) The standard deviation in pizza prices is closest to $0.15, since that is the typical distance to the mean. There are no pizza prices as far as $0.50 of $1.00.

26. Neck sizes again.

a) The mean neck size is closest to 15 inches. That's the balancing point of the histogram.

b) The standard deviation in neck sizes is closest to 1 inch, because a typical value lies about 1 inch from the mean. There are a few points as far away as 3 inches from the mean, and none as far away as 5 inches. Those are too large to be the standard deviation.

27. Movie lengths.

a) A typical movie would be around 105 minutes long. This is near the center of the unimodal and roughly symmetric histogram.

b) You would be surprised to find that your movie ran for 150 minutes. Only 2 movies ran that long.

c) The mean run time would probably be roughly the same as the median run time, since the distribution is reasonably symmetric.

28. Golf drives 2015.

a) The distribution of golf drives is roughly unimodal and symmetric, with a typical drive of around 290 yards. Professional golfers on the men's PGA tour had drives that were as short as about 270 yards, and as long as about 320 yards for a range of roughly 50 yards. There is a slightly unusual high and low average length, 320 and 260 yards, respectively.

b) Approximately 25% of professional male golfers drive less than 280 yards.

c) The mean is 288.69 yards and the median is 288.7 yards. Any estimate between 285 and 295 yards is reasonable.

d) The distribution of golf drives is approximately symmetric, so the mean and the median should be relatively close.

29. Movie lengths II.

a) i) The distribution of movie running times is fairly consistent, with the middle 50% of running times between 97 and 115 minutes. The interquartile range is 18 minutes.

ii) The standard deviation of the distribution of movie running times is 17.3 minutes, which indicates that movies typically have running times fairly close to the mean running time. On average, movie running times differed from the mean running time by only 17.3 minutes.

b) Since the distribution of movie running times is reasonably symmetric, either measure of spread would be appropriate.