For Practice

1. Josiah and Allen have the same *ABILITY* to run the mile. However, Josiah is a very consistent runner, while Allen isn't. Sketch dotplots for both runners (on the same scale) that show the results of 10 *PERFORMANCES* for each runner.

2. Rachel and Brittany have the same *ABILITY* to throw the discus. However, Rachel is much more consistent than Brittany. Sketch boxplots for both throwers (on the same scale) that show the results of 10 *PERFORMANCES* for each thrower.

3. Adam Dunn is known as one of the most consistent home run hitters in baseball. Here are his home run totals for the years 2004–2009: 46, 40, 40, 40, 40, 38. Calculate *and* interpret the mean absolute deviation of these *PERFORMANCES*.

4. Albert Pujols of the St. Louis Cardinals is also considered a very consistent hitter. Here are his home run totals for the years 2001–2009: 37, 34, 43, 46, 41, 49, 32, 37, 47. Calculate *and* interpret the mean absolute deviation of these *PERFORMANCES*.

5. In the 2009 LPGA Championship, winner Anna Nordqvist had scores of 66, 70, 69, and 68 while runner-up Lindsey Wright posted scores of 70, 68, 69, and 70.
   (a) Use dotplots to display these data and discuss any differences in consistency.
   (b) Calculate *and* interpret the mean absolute deviations for both golfers. Does this confirm what you saw in the graphs? Explain.

6. Of the top 2 running backs in 2008, who was more consistent, Adrian Peterson or Michael Turner? The following data show the number of yards gained for each running back during the 2008 regular season:

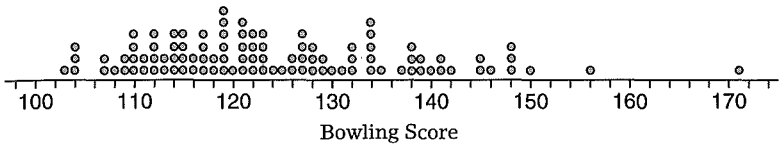
Rushing yards Michael Turner	220	42	104	56	121	54	58	139	96	81	117	120	61	152	70	208
Rushing yards Adrian Peterson	103	160	77	80	32	111	121	139	192	85	80	131	105	165	76	103

   (a) Use dotplots to display these data and discuss any differences in consistency.
   (b) Calculate *and* interpret the mean absolute deviations for both running backs. Does this confirm what you saw in the graphs? Explain.

7. Using points scored as the variable, create two distributions of 5 basketball *PERFORMANCES* that have different means but the same MAD.

8. Using points scored as the variable, create a distribution of 5 basketball *PERFORMANCES* that have a MAD = 0.

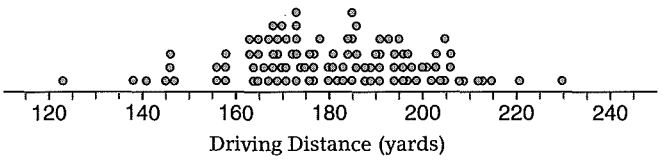
9. Here is Nick's distribution of bowling scores during the last month. The standard deviation of this distribution is 13.



(a) Interpret Nick's standard deviation.

(b) Andre has the same *ABILITY* to bowl as Nick, but he is less consistent. In fact, Andre's standard deviation is twice as big. Draw a dotplot of what you think Andre's distribution of bowling scores looks like.

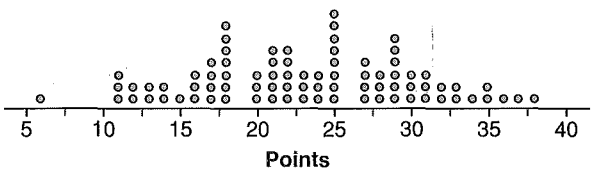
10. Here is Chelsea's distribution of driving distances (in yards) from her last trip to the driving range. The standard deviation of this distribution is 19.2 yards.



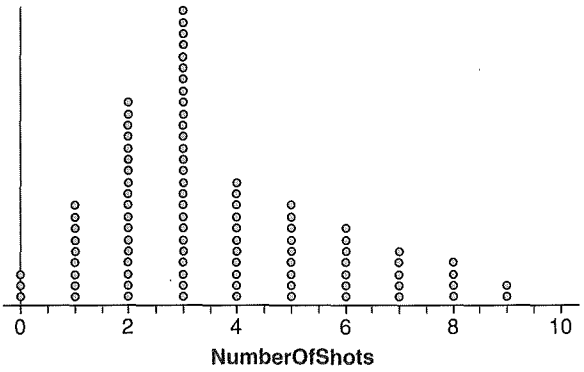
(a) Interpret Chelsea's standard deviation.

(b) Shauntelle has the same *ABILITY* to drive a golf ball as Chelsea, but she is more consistent. In fact, Shauntelle's standard deviation is only half as big as Chelsea's. Draw a dotplot of what you think Shauntelle's distribution of driving distances looks like.

11. The dotplot below shows the number of points scored by NBA player Tim Duncan in his 81 regular season games during the 2002–2003 season. Is the standard deviation of this distribution closest to 2, 7, or 15? Explain.



12. The dotplot below shows the number of shots taken by NHL player Sidney Crosby in his 81 regular season games during the 2009–2010 season. Is the standard deviation of this distribution closest to 2, 5, or 10? Explain.



13. In the 2008–2009 regular season, Kobe Bryant of the Los Angeles Lakers averaged 26.8 points per game with a standard deviation of 8.6 points.

(a) Interpret Kobe's standard deviation.

(b) If he played one additional game and scored 27 points, what effect would that have on the mean and standard deviation? Explain.

(c) If he played one additional game and scored 60 points, what effect would that have on the mean and standard deviation? Explain.

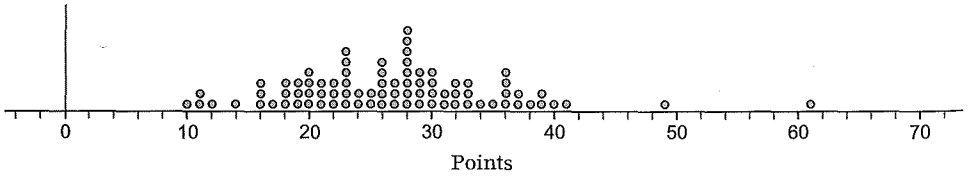
14. In the 2009 regular season, pitcher Tim Lincecum of the San Francisco Giants averaged 8.2 strikeouts per game with a standard deviation of 2.8 strikeouts.

(a) Interpret Lincecum's standard deviation.

(b) If he pitched in one additional game and had only 2 strikeouts, what effect would that have on the mean and standard deviation? Explain.

(c) If he pitched in one additional game and had 8 strikeouts, what effect would that have on the mean and standard deviation? Explain.

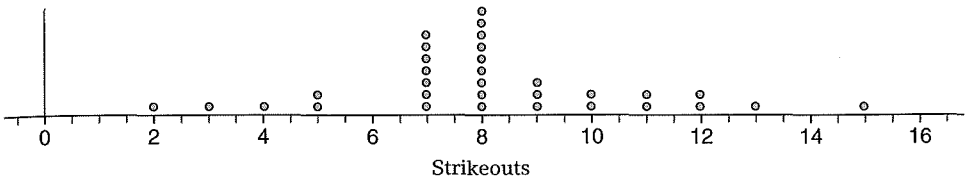
15. Here is a dotplot of Kobe Bryant's point totals for each of the 82 games during the 2008–2009 regular season. The mean of this distribution is 26.8 points, with a standard deviation of 8.6 points.



(a) In what percentage of games did he score within one standard deviation of his mean?

(b) In what percentage of games did he score within two standard deviations of his mean?

16. Here is a dotplot of Tim Lincecum's strikeout totals for each of the 32 games he pitched in during the 2009 regular season. The mean of this distribution is 8.2, with a standard deviation of 2.8.



(a) In what percentage of games was his number of strikeouts within one standard deviation of the mean?

(b) In what percentage of games was his number of strikeouts within two standard deviations of the mean?

17. Tom Brady, quarterback of the New England Patriots, is considered by many people to be one of the best players in the league. Here are his touchdown pass totals for the years 2001–2006: 18, 28, 23, 28, 26, 24.

(a) Calculate *and* interpret the mean absolute deviation (MAD) of these PERFORMANCES.

(b) Calculate *and* interpret the standard deviation of these PERFORMANCES.

(c) In 2007, Tom Brady had a league-leading 50 touchdown passes. Recalculate the MAD and SD including this new value. Briefly summarize the effect of this unusually high value on each measure of variability.

18. In his first four full seasons in the Major Leagues, New York Mets third baseman David Wright hit 27, 26, 30, and 33 home runs.

(a) Calculate *and* interpret the mean absolute deviation (MAD) of these PERFORMANCES.

(b) Calculate *and* interpret the standard deviation of these PERFORMANCES.

(c) In his 5th season, Wright only had 10 home runs. Recalculate the MAD and SD including this new value. Briefly summarize the effect of this unusually low value on each measure of variability.

19. Which quarterback was more consistent over his career, Dan Marino or John Elway? Here are the average number of passing yards per game for both quarterbacks during each season of their careers:

Marino: 201 318 259 297 270 277 250 223 248 257 244 278 262 215 236 219 223

Elway: 151 173 243 218 267 221 203 220 203 187 252 249 248 222 227 216

(a) Draw dotplots to display the distribution of average passing yards for each quarterback. Briefly compare these distributions.

(b) Calculate *and* interpret the standard deviation for both quarterbacks.

(c) Based on your answers to part (b), which quarterback was more consistent? Explain.

20. Have the Pittsburgh Pirates been more consistent in their futility than the New York Yankees have been in their excellence? Here are their number of wins per season for 2000–2009:

Yankees: 87 95 103 101 101 95 97 94 89 103

Pirates: 69 62 72 75 72 67 67 68 67 62

(a) Draw dotplots to display the distribution of wins for each team. Briefly compare these distributions.

(b) Calculate *and* interpret the standard deviation for both teams.

(c) Based on your answers to part (b), which team was more consistent? Explain.

21. A member of a track team was practicing the long jump and recorded the following distances (in centimeters): 570, 575, 578, 579, 583.

(a) Calculate the mean and standard deviation of these distances.

(b) After chatting with a teammate, the jumper realized that he measured his jumps from the back of the board instead of the front of the board. Thus, he had to subtract 20 centimeters from each of his jumps to get the correct measurement for each jump. What effect does the subtraction of 20 centimeters from each jump have on the mean and standard deviation of his distances?

22. Tim Tebow was the quarterback at the University of Florida during the 2006–2009 seasons. Here are the number of touchdowns he completed in each of his four seasons: 5, 32, 30, 21.

(a) Calculate the mean and standard deviation of these totals.

(b) Calculate the deviations from the mean for these totals.

(c) Calculate the mean and standard deviation of the *deviations*. How do these values compare to your answers in part (a)?

23. Is LeBron James a more consistent scorer at home than on the road? During the 2008–2009 regular season at home, he had a mean of 25.35 points, with a standard deviation of 7.94 points in 40 games; while on the road, he had a mean of 31.46 points, with a standard deviation of 8.63 in 41 games.

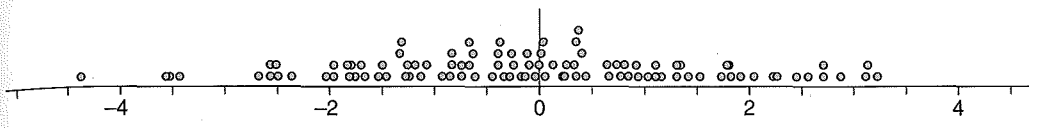
(a) Is 7.94 LeBron's observed standard deviation at home or his true standard deviation at home? Explain.

(b) State the hypotheses you would use to test if LeBron is more consistent at home.

(c) Calculate the difference in standard deviations (road − home) and use this as the test statistic.

(d) Describe how to simulate the distribution of the test statistic, assuming that LeBron is equally consistent at home and on the road.

(e) Here are the results of 100 trials of the simulation. Describe what information is provided by the dotplot.



Simulated Difference in Standard Deviations (Road – Home)

(f) Based on the results of the simulation, estimate and interpret the *p*-value.

(g) Based on the *p*-value, make an appropriate conclusion.

(h) If your conclusion is in error, which type of error did you make? Explain.

24. Was the St. Louis Cardinals offense more consistent at home? During the 2009 regular season at home, they had a mean of 4.21 runs scored, with a standard deviation of 2.80 runs scored in 81 games; while on the road, they had a mean of 4.80 runs scored, with a standard deviation of 3.38 runs scored in 81 games.

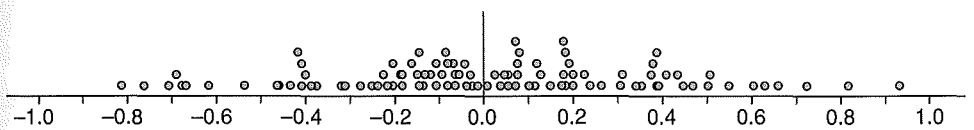
(a) Is 3.38 the Cardinals' observed standard deviation on the road or their true standard deviation on the road? Explain.

(b) State the hypotheses you would use to test if the Cardinals' offense was more consistent at home.

(c) Calculate the difference in standard deviations (road − home) and use this as the test statistic.

(d) Describe how to simulate the distribution of the test statistic, assuming that the Cardinals' offense was equally consistent at home and on the road.

(e) Here are the results of 100 trials of the simulation. Describe what information is provided by the dotplot.



Simulated Difference in Standard Deviations (Home – Road)

(f) Based on the results of the simulation, estimate and interpret the *p*-value.

(g) Based on the *p*-value, make an appropriate conclusion.

(h) If your conclusion is in error, which type of error did you make? Explain.

25. In NASCAR, it is very important for a crew to be able to predict how long a set of tires should last. It is also important that all 4 tires wear out at the same rate! So, when evaluating different brands, consistency is important. To compare two brands of tires, A and B, a crew bought 2 sets of 4 brand A tires and 2 sets of 4 brand B tires and ran each set (in random order) for a total of 200 miles at race speed. At the end, a trained evaluator rated the amount of wear on each tire using a scale from 0 (no wear) to 100 (completely worn out).

(a) Explain why it was important to randomly assign the order in which the tires were used.

(b) How did this experiment use the concept of control? What else should be controlled in this experiment?

(c) Could this be run as a double-blind experiment? Explain.

Here are the results of the experiment (0 = no wear and 100 = totally worn out):

Brand A	83	81	81	84	77	79	82	81
Brand B	63	68	82	87	75	71	68	70

(d) Graph these distributions and briefly compare them. Why might a crew prefer brand A? Why might they prefer brand B?

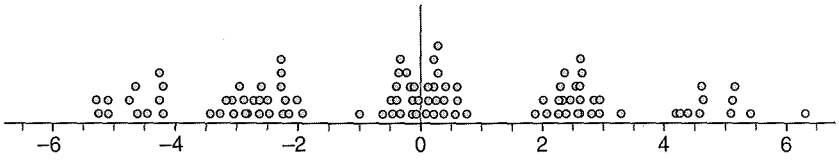
(e) Calculate and interpret the standard deviations of both distributions.

(f) To test whether Brand A tires wear more consistently, what hypotheses should you use?

(g) Calculate the difference in standard deviations (A − B) and use this as the test statistic.

(h) Describe how to simulate the distribution of the test statistic, assuming the brands are equally consistent.

(i) Use the results of the simulation below to estimate the *p*-value and make an appropriate conclusion.



26. Suppose that you wanted to compare the consistency of two different brands of golf balls (one cheap brand and one expensive brand). After all, it's not just about how far you can hit the ball—it's also important to be able to predict how far the ball will travel with as much precision as possible. To investigate, an experienced golfer hit 10 of each brand of ball, in random order, always using the same club.

(a) Explain why it was important to randomly assign the order in which the golf balls were hit.

(b) How did this experiment use the concept of control? What else should be controlled in this experiment?

(c) Could this be run as a blind experiment? Explain.

Here are the results of the experiment (in yards):

Cheap	193	179	207	196	187	201	183	189	188	197
Expensive	203	196	200	197	198	201	208	201	211	205

(d) Graph these distributions and briefly compare them. Give two reasons why you might prefer the expensive brand.

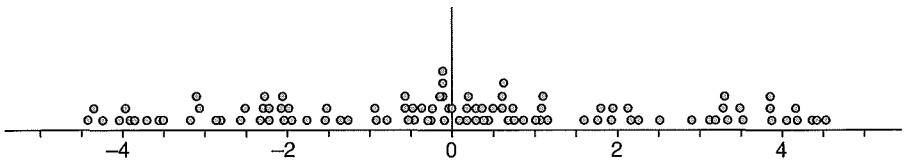
(e) Calculate and interpret the standard deviations of both distributions.

(f) To test whether the golfer is more consistent with the expensive balls, what hypotheses should you use?

(g) Calculate the difference in standard deviations (cheap − expensive) and use this as the test statistic.

(h) Describe how to simulate the distribution of the test statistic, assuming that the golfer is equally consistent with both types of balls.

(i) Use the results of the simulation below to estimate the *p*-value and make an appropriate conclusion.



Simulated Difference in Standard Deviations (Cheap − Expensive)

27. In the National Hockey League, the Detroit Red Wings and the Pittsburgh Penguins faced off for the 2008−2009 Stanley Cup Championship. Clearly, these were two of the best teams in the NHL during that particular season. How do the two teams compare over the last 10 years? Was one team more consistent? Here are the total points for each team during the previous 9 seasons. For a hockey team, points = 2 (number of wins) + (number of ties) + (number of overtime/shootout losses).

Season	99−00	00−01	01−02	02−03	03−04	05−06	06−07	07−08	08−09
Detroit	108	111	116	110	109	124	113	115	112
Pittsburgh	88	96	69	65	58	58	105	102	99

(a) Graph these distributions and briefly compare them. Which team seems more consistent?

(b) To test whether the Detroit Red Wings are more consistent, what hypotheses should you use? What test statistic should you use? What is the value of the test statistic?

(c) Describe how to simulate the distribution of the test statistic, assuming that the two teams are equally consistent.

(d) Conduct at least 20 trials of your simulation. Use the results of the simulation to estimate the *p*-value and make an appropriate conclusion.

(e) The point total for Detroit in the 2005–2006 season (124) was an outlier. If you were to remove it from Detroit's distribution, what effect will this have on the standard deviation? Will removing the outlier make it easier or harder to show that Detroit was the more consistent team? Explain.

28. In the 1990s and 2000s, the Atlanta Braves were a model of consistency, winning 14 straight division titles and 1 World Series championship. However, during that same time period, another team in the same division, the Florida Marlins, won 2 World Series championships but suffered some terrible years as well. Were the Braves significantly more consistent than the Marlins? Here are their winning percentages for the years 1993 through 2005:

Atlanta: 64.2 59.6 62.5 59.3 62.3 65.4 63.6 58.6 54.3 63.1 62.3 59.3 55.6

Florida: 39.5 44.3 46.9 49.4 56.8 33.3 39.5 49.1 46.9 48.8 56.2 51.2 51.2

(a) Graph these distributions and briefly compare them. Which team seems more consistent?

(b) To test whether the Atlanta Braves are more consistent, what hypotheses should you use? What test statistic should you use? What is the value of the test statistic?

(c) Describe how to simulate the distribution of the test statistic, assuming that the two teams are equally consistent.

(d) Conduct at least 20 trials of your simulation. Use the results of the simulation to estimate the *p*-value and make an appropriate conclusion.

(e) If your conclusion is in error, which type of error did you make? Explain.

29. For a quarterback, being able to throw passes a specific distance is more important than how far he can throw the ball. In other words, it is more important that a quarterback's passes be consistent so he doesn't under- or overthrow his receivers. In hot weather, when the quarterback's hands might be slippery with sweat, will wearing gloves make him a more consistent passer? Design an experiment to answer this question.

(a) What are the explanatory and response variables in your experiment?

(b) How will you incorporate randomization in your experiment?

(c) What variables are important to control during this experiment?

(d) Can the quarterback be blind in this experiment? Explain why this matters.

(e) State the hypotheses you are interested in testing and the test statistic you will use.

30. Dale is trying out for his school's bowling team and knows that the coach values consistency just as much as a high average score. Design an experiment to determine whether using a lighter bowling ball will make Dale a more consistent bowler.

(a) What are the explanatory and response variables in your experiment?

(b) How will you incorporate randomization in your experiment?

(c) What variables are important to control during this experiment?

(d) Can Dale be blind in this experiment? Explain why this matters.

(e) State the hypotheses you are interested in testing and the test statistic you will use.

CHAPTER REVIEW EXERCISES

31. A track coach is trying to decide which of two runners to include on the 4×400 relay team for an upcoming meet. Here are the most recent 400-meter times (in seconds) for each of the runners:

Amy	61.8	61.3	64.6	60.6	63.2	64.5	63.0	59.8
Janelle	62.2	62.5	63.0	63.2	62.8	61.7	62.5	62.9

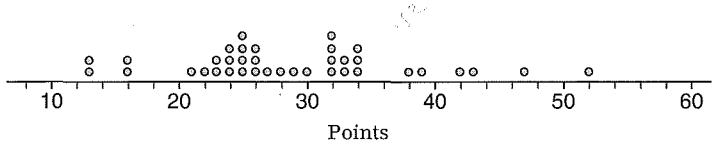
(a) Using dotplots, graph these two distributions and briefly compare them.

(b) Calculate *and* interpret the mean absolute deviation (MAD) for both athletes.

(c) Explain why the coach might choose Amy to be in the relay instead of Janelle.

(d) Explain why the coach might choose Janelle to be in the relay instead of Amy.

32. In the 2010–2011 college basketball season, Jimmer Fredette of Brigham Young University was named the National Player of the Year by a variety of organizations. The dotplot below shows his distribution of points scored during the season.[6] The mean of this distribution is 28.9 points, with a standard deviation of 8.7 points.



(a) Interpret the standard deviation of points scored.

(b) In what percentage of his games did Jimmer score within one standard deviation of the mean? In what percentage of his games did he score within two standard deviations of the mean?

(c) If Jimmer played one additional game and scored 60 points, what would happen to the mean and standard deviation of the distribution? Explain.

(d) If Jimmer played one additional game and scored 25 points, what would happen to the mean and standard deviation of the distribution? Explain.

(e) Is 8.7 points Jimmer's observed standard deviation or true standard deviation? Explain.

33. For games played at home, the Colorado Rockies store their baseballs in a humidor to keep the balls from getting too hard in the dry mountain air. What effect does this have on the baseballs? Does it make them more consistent since they are stored in a regulated climate? To investigate, 10 baseballs were stored in a regulated, humid environment and 10 baseballs were stored in an unregulated, less humid environment. In random order, each ball was dropped from a fixed distance onto a hard surface and the height the ball reached after its first bounce was recorded.

(a) Explain why it is important to randomly assign which balls are stored in which environment and to randomly assign the order in which the balls were dropped.

(b) How did this experiment use the concept of control? What else should be controlled in this experiment?

(c) Could blinding be used in this experiment? Explain.

Here are the heights of each bounce (in feet):

Stored in humid environment	23	22	23	24	24	23	22	23	24	22
Not stored in humid environment	22	23	23	24	24	26	27	22	24	25

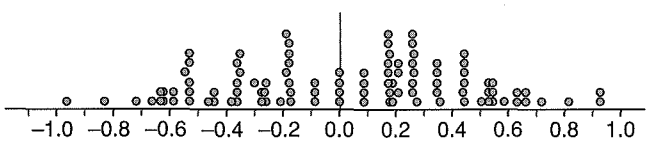
(d) Graph these distributions and briefly compare them.

(e) To test whether balls stored in a humid environment are more consistent, what hypotheses should you use?

(f) Calculate the difference in standard deviations (humid − not) and use this as the test statistic.

(g) Describe how to simulate the distribution of the test statistic, assuming the balls stored in each environment are equally consistent.

(h) The results of 100 trials of the simulation are shown below. Use the results of the simulation to estimate the *p*-value and make an appropriate conclusion.



Simulated Difference in Standard Deviations (Humid – Not)

(i) If your conclusion is in error, which type of error did you make? Explain.

OTHER APPLICATIONS

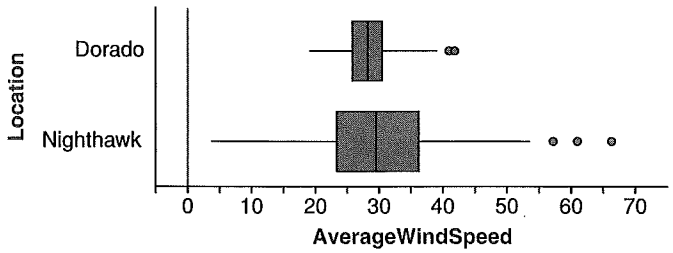
34. Here are the average high temperatures in each month for two different cities, Los Angeles, CA, and St. Louis, MO:[7]

MONTH	LOS ANGELES	ST. LOUIS
January	68	38
February	70	45
March	70	55
April	73	66
May	75	77
June	80	86
July	84	91
August	85	88
September	83	81
October	79	69
November	73	54
December	69	42

(a) Using dotplots, graph these two distributions and briefly compare them.

(b) Calculate and interpret the mean absolute deviation (MAD) for both cities. Which city has the most consistent high temperatures?

35. The use of alternative energy sources has grown tremendously in the last few years. To capture the energy provided by air currents, "wind farmers" have built arrays of windmills that turn wind into electricity. Because no electricity can be generated when there is no wind, potential wind farmers are scouting locations that have a consistent supply of wind. The graphs below show the average daily wind speeds (in miles per hour) during the last year at two different locations, Dorado Canyon and Nighthawk Valley.



(a) The standard deviation of wind speeds in Dorado Canyon is 3.99 mph. Interpret this value.

(b) There are two outliers in the distribution of wind speeds in Dorado Canyon. If these outliers were removed from the distribution, what would happen to the mean and standard deviation? Explain.

(c) Is the standard deviation of wind speeds in Nighthawk Valley smaller or larger than 3.99 mph? Explain how you know this.

(d) Explain why the farmer might choose to settle in Dorado Canyon.

(e) Explain why the farmer might choose to settle in Nighthawk Valley.

36. Nurses at a hospital want to know whether a newer, more expensive type of thermometer measures temperatures more precisely. That is, they want to know whether the new type provides more consistent temperature measurements than the cheaper alternative that they currently use. To answer this question, the nurses use 5 new thermometers and 5 old thermometers to conduct an experiment. In the experiment, they will use each of the 10 thermometers, in random order, to measure the temperature of a patient.

(a) Explain why it is important to randomly assign the order in which the thermometers are used.

(b) How did this experiment use the concept of control? What else should be controlled in this experiment?

(c) Could blinding be used in this experiment? Explain.

Here are the temperatures using each thermometer (in degrees Fahrenheit):

New thermometer	98.2	98.2	98.1	97.9	98.3
Old thermometer	98.2	98.3	98.0	98.5	97.8

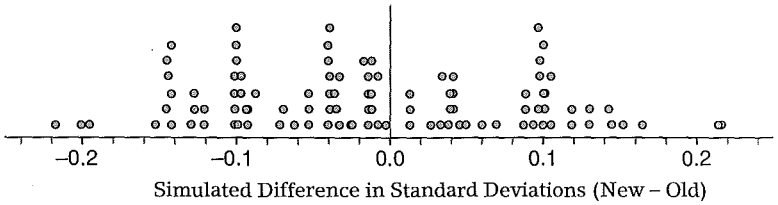
(d) Graph these distributions and briefly compare them.

(e) To test whether the new thermometers are more consistent, what hypotheses should you use?

(f) Calculate the difference in standard deviations (new − old) and use this as the test statistic.

(g) Describe how to simulate the distribution of the test statistic, assuming the two types of thermometers are equally consistent.

(h) The results of 100 trials of the simulation are shown below. Use the results to estimate the *p*-value and make an appropriate conclusion.



Simulated Difference in Standard Deviations (New – Old)

(i) Describe a Type I and a Type II error in the context of this question and briefly discuss the consequences of each type of error. Which error is it possible that you made in your conclusion?