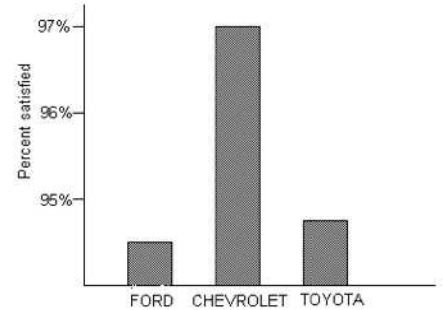


1. The following bar graph gives the percent of owners of three brands of trucks who are satisfied with their truck. From this graph, we may conclude that
- (a) owners of other brands of trucks are less satisfied than the owners of these three brands.
 - (b) Chevrolet owners are substantially more satisfied than Ford or Toyota owners.
 - (c) there is very little difference in the satisfaction of owners for the three brands.
 - (d) Chevrolet probably sells more trucks than Ford or Toyota.
 - (e) a pie chart would have been a better choice for displaying these data.



2. Here are the IQ test scores of 10 randomly chosen fifth-grade students:
 145 139 126 122 125 130 96 110 118 118
- To make a stemplot of these scores, you would use as stems
- (a) 0 and 1.
 - (b) 09, 10, 11, 12, 13, and 14.
 - (c) 96, 110, 118, 122, 125, 126, 130, 139, and 145.
 - (d) 0, 2, 3, 5, 6, 8, 9.
 - (e) None of the above is a correct answer.

3. If a distribution is skewed to the right,
- (a) the mean must be greater than the median.
 - (b) the mean and median must be equal.
 - (c) the mean must be less than the median.
 - (d) the mean is either equal to or less than the median.
 - (e) It's impossible to tell without seeing the data.

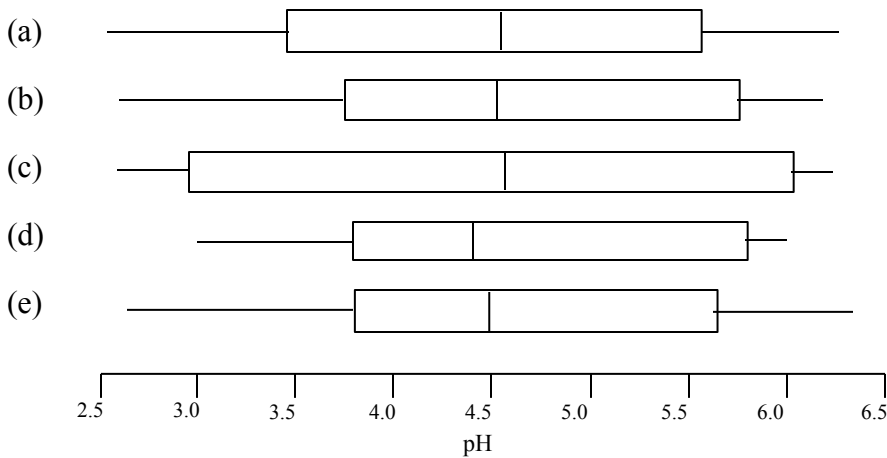
4. Rainwater was collected in water collectors at 30 different sites near an industrial complex and the amount of acidity (pH level) was measured. The data ranged from pH 2.6 to pH 6.3. The following stemplot of the data was constructed.

```

2|679
3|237789
4|1222446899
5|0556788
6|0233
    
```

Key: 3|7 = pH 3.7

Which of the following boxplots is a correct representation of the same distribution?



5. A sample of 252 high school students were asked, “If you had \$1000 to contribute to one kind of charitable organization, which type of organization would you choose? Below is a two-way table of responses to this question and gender.

	Organization				
Gender	Education	Environment	Health	International Aid	Other
Female	19	33	50	28	10
Male	23	29	28	17	13

Which of the following conclusions seems to be supported by the data?

- (a) The second choice for most females in the sample is environment.
 (b) There is no association between gender and choice of organization.
 (c) The proportion of males who said they would contribute to an environmental organization was higher than the proportional of females who said they would contribute such an organization.
 (d) None of the students surveyed said they would contribute to religious organizations.
 (e) The marginal totals of Organization is 140, 110.
6. A small company that prints custom t-shirts has 6 employees, one of whom is the owner and manager. Suppose the owner makes \$120,000 per year and the other employees make between \$40,000 and \$50,000 per year. One day, the owner decides to give himself a \$30,000 raise. Which of the following describes how the standard deviation and interquartile range of salaries would change because of the raise?
 (a) The standard deviation and interquartile range would both increase.
 (b) Neither the standard deviation nor the interquartile range would change.
 (c) The standard deviation would increase and the interquartile range would not change.
 (d) The standard deviation would decrease and the interquartile range would not change.
 (e) The interquartile range would increase and the standard deviation would not change.
7. The mean speed of vehicles in the “cars only” lanes of the New Jersey turnpike is 68 miles per hour. The mean speed of vehicles in the “any vehicle” lanes is 64 miles per hour. What must be true about the mean speed of all vehicles on the turnpike, assuming these are the only types of lanes?
 (a) It could be any number between 64 and 68 miles per hour.
 (b) It must be larger than the median speed.
 (c) It must be larger than 66 miles per hour.
 (d) It must be 66 miles per hours.
 (e) We don’t have enough information to draw any conclusion about the mean speed of all vehicles.
8. The mean birth weight of infants born at a certain hospital in the month of April was 128 oz. with a standard deviation of 10.2 oz. Which of the following is a correct interpretation of standard deviation?
 (a) All the infants born in April weighed between 117.8 oz. and 138.2 oz.
 (b) About half the infants born in April weighed between 117.8 oz. and 138.2 oz.
 (c) The difference between the mean weight and the median weight of infants born in April was 10.2 oz.
 (d) The distance between the weight of each infant born in April and the mean weight was, on average, about 10.2 oz.
 (e) The mean weight of infants born in subsequent months is likely to be within 10.2 oz. of the mean weight in April.
9. At the beginning of the school year, a high-school teacher asks every student in her classes to fill out a survey that asks for their age, gender, the number of year they have lived at their current address, their favorite school subject, and whether they plan to go to college after high school. Which of the following best describes the variables that are being measured?
 (a) four quantitative variables
 (b) five quantitative variables
 (c) two categorical variables and two quantitative variables
 (d) two categorical variables and three quantitative variables
 (e) three categorical variables and two quantitative variables

10. A medical researcher collects health data on many women in each of several countries. One of the variables measured for each woman in the study is her weight in pounds. The following list gives the five-number summary for the weights of adult women in one of the countries.

Country A: 92, 110, 120, 160, 240

About what percent of Country A women weigh between 110 and 240 pounds?

- (a) 50%
 - (b) 65%
 - (c) 75%
 - (d) 85%
 - (e) 95%
11. A small company estimating its photocopying expenses finds that the mean number of copies made per day for the past 12 months is 258 copies per day with a standard deviation of 24 copies per day. Which of the following is a correct interpretation of standard deviation?
- (a) The number of copies made per day was always between 234 and 282.
 - (b) About 95% of the time, the number of copies made per day was always between 234 and 282.
 - (c) The difference between the mean number of copies made per day and the median number of copies made per day was between 234 and 282.
 - (d) On average, the number of copies made each day was about 24 copies per day away from the mean, 258.
 - (e) 1.5 times the interquartile range of copies made per day is 24.
12. When testing water for chemical impurities, results are often reported as bdl, that is, below detection limit. The following are the measurements of the amount of lead in a series of water samples taken from inner-city households (in parts per million):

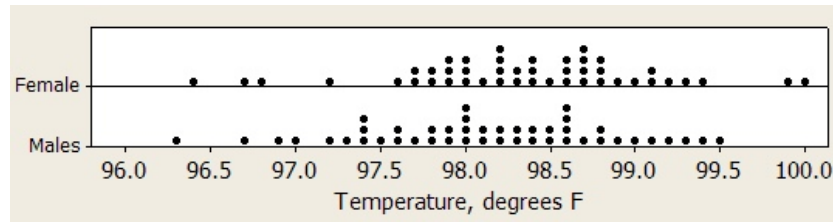
5, 7, 12, bdl, 10, 8, bdl, 20, 6

Which of the following statements can we be sure is true?

- (a) The mean lead level in the water is about 10 ppm.
 - (b) The mean lead level in the water is about 9 ppm.
 - (c) The median lead level in the water is 7 ppm.
 - (d) The median lead level in the water is 8 ppm.
 - (e) Neither the mean nor the median can be computed because some values are unknown.
13. A sample of 99 distances has a mean of 24 feet and a median of 24.5 feet. Unfortunately, it has just been discovered that the maximum value in the distribution, which was erroneously recorded as 40, actually had a value of 50. If we make this correction to the data, then
- (a) the mean remains the same, but the median is increased.
 - (b) the mean and the median remain the same.
 - (c) the median remains the same, but the mean is increased.
 - (d) the mean and the median are both increased.
 - (e) we do not know how the mean and the median are affected without further calculations, but the variance is increased.
14. A researcher for a consumer products company is field testing a new formula for laundry detergent. He has contracted with 60 families, each with two children, who have agreed to test the product. He randomly assigns 30 families to the group that will use the new formula and 30 to the group that will use the company's current detergent formula. The most important reason for this random assignment is that
- (a) randomization is a good way to create two groups of 30 families that are as similar as possible, so that comparisons can be made between the two groups.
 - (b) randomization eliminates the impact of any confounding variables.
 - (c) randomization reduces the impact of outliers.
 - (d) randomization ensures that the study is double-blind.
 - (e) randomization makes the analysis easier since the data can be collected and entered into the computer in any order.

15. What electrical charges occur in muscles as they get tired? Student subjects are instructed to hold their arms above their shoulders as long as they can. Meanwhile, the electrical activity in their arms is measured. This is
- (a) a randomized comparative experiment.
 - (b) a matched pairs design.
 - (c) an uncontrolled experiment.
 - (d) an observational study.
 - (e) impossible to describe unless more details of the survey are provided.
16. A recent survey by a large-circulation Canadian magazine on the contribution of universities to the economy was circulated to 394 people who the magazine decided “are the most likely to know how important universities are to the Canadian economy.” The main problem with using these results to draw conclusions about the general public’s perception is
- (a) insufficient attention to the placebo effect.
 - (b) response bias.
 - (c) no control group.
 - (d) lack of random selection.
 - (e) lack of random assignment.
17. Which of the following properties is true for all Normal density curves?
- I. They are symmetric.
 - II. The curve reaches its peak at the mean.
 - III. 95% percent of the area under the curve is within one standard deviation of the mean.
- (a) I only
 - (b) II only
 - (c) I and II only
 - (d) I and III only
 - (e) All three statements are correct.
18. The area under the standard Normal curve corresponding to $-0.3 < Z < 1.6$ is
- (a) 0.3273
 - (b) 0.4713
 - (c) 0.5631
 - (d) 0.9542
 - (e) none of the above
19. The distribution of the time it takes for different people to solve a certain crossword puzzle is strongly skewed to the right, with a mean of 30 minutes and a standard deviation of 15 minutes. The distribution of z-scores for those times is:
- (a) Normally distributed, with mean 30 and standard deviation 15.
 - (b) Skewed to the right, with mean 30 and standard deviation 15.
 - (c) Normally distributed, with mean 0 and standard deviation 1.
 - (d) Skewed to the right, with mean 0 and standard deviation 1.
 - (e) Skewed to the right, but the mean and standard deviation cannot be determined without more information.
20. The heights of American men aged 18 to 24 are approximately Normally distributed with mean 68 inches and standard deviation 2.5 inches. Only about 5% of young men have heights outside the range
- (a) 65.5 inches to 70.5 inches
 - (b) 63 inches to 73 inches
 - (c) 60.5 inches to 75.5 inches
 - (d) 58 inches to 78 inches
 - (e) none of the above

1. We all “know” that the body temperature of a healthy person is 98.6°F. In reality, the actual body temperature of individuals varies. Here are dot plots and summary statistics of the body temperatures for 90 healthy individuals (45 males and 45 females).

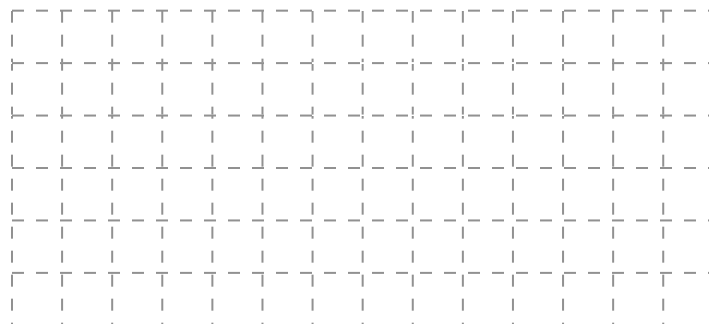


Variable	Count	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Female	45	98.344	0.111	0.745	96.400	97.900	98.400	98.800	100.000
Males	45	98.122	0.110	0.735	96.300	97.600	98.100	98.600	99.500

- (a) Determine if there are any outliers in the *female* distribution. Show your work.

- (b) Write a few sentences comparing the distribution of body temperatures for adult males and females.

- (c) Make a histogram for the *male* distribution.

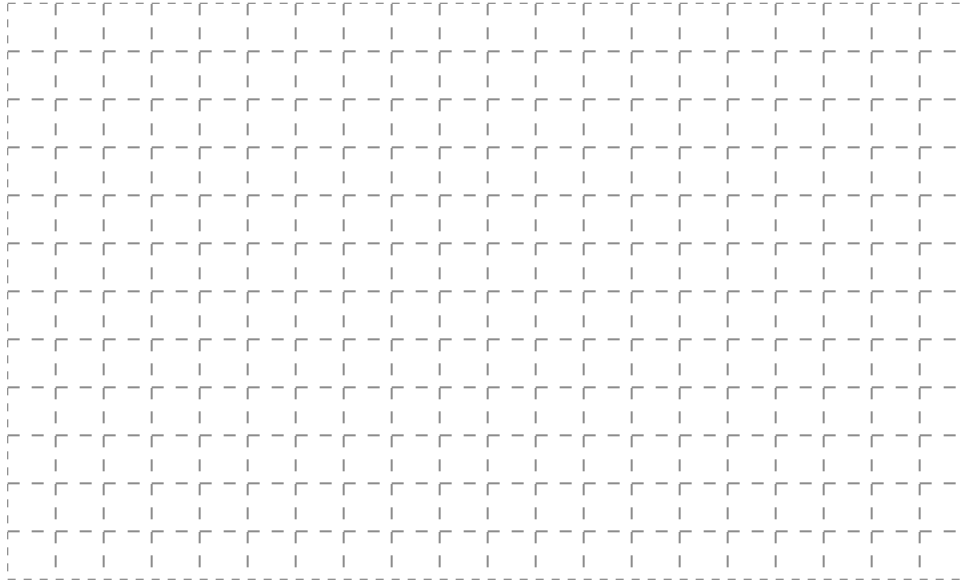


- (d) Suppose that we measured the temperature of an additional male and his temperature was 98.2°F. What effect will including this observation have on the mean and standard deviation of the *male* distribution? Explain.

2. A research study asked children which of four different emotions they associated with the color red. The response and gender of each child are given in the following table.

	Joy	Happiness	Love	Anger
Male	28	20	40	18
Female	61	25	80	60

- (a) Make a graph to display the relationship between gender and response.



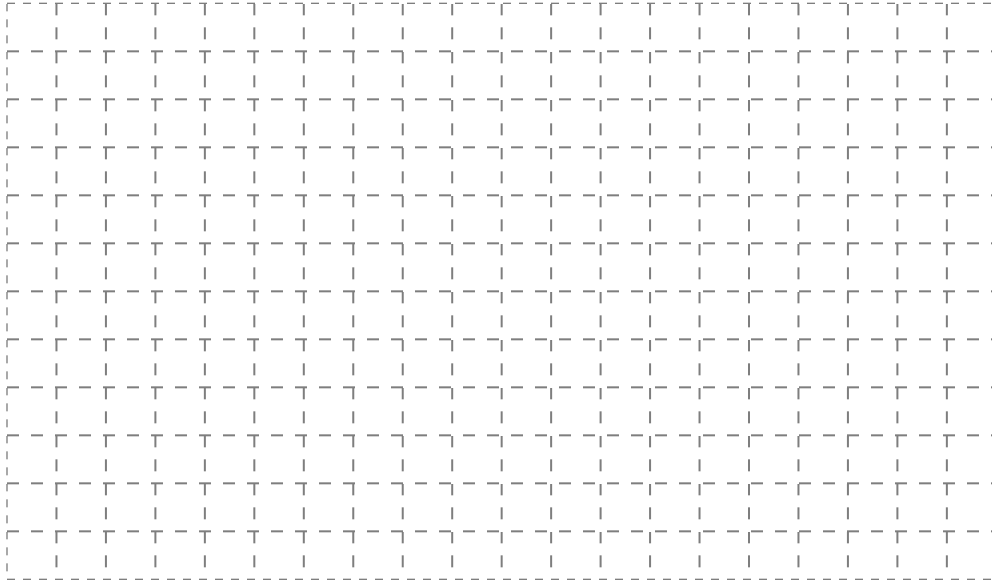
- (b) Based on your graph, briefly discuss the association between gender and response.

3. Two additional measures of center are the midrange (average of the minimum and maximum) and the midquartile (average of Q_1 and Q_3). Which of these, if either, are resistant measures of center? Explain.
4. As a researcher for a pharmaceutical company, you are designing a study to test the effectiveness of a new treatment for migraine headaches. You have been given a list of 126 people willing to participate in the trial. The first 70 people are female; the remaining 56 are male. Describe a design for this experiment. Be sure to include a description of how you assign individuals to the treatment groups.

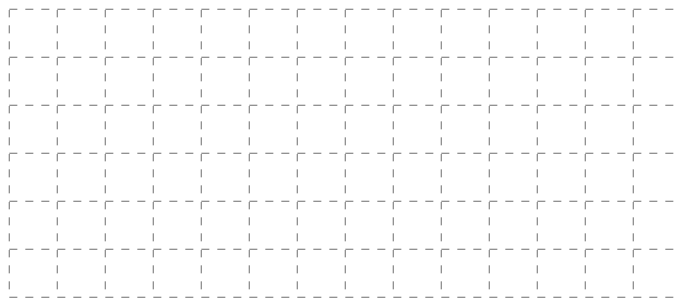
5. The data below is the number of unprovoked attacks by alligators on people in Florida each year for a 33-year period.

6	12	2	4	17	4	6	10	3	9	13
9	15	14	6	18	1	9	6	6	11	24
14	14	5	17	17	5	13	22	20	3	5

(a) Construct a histogram for this distribution. Choose an appropriate bin width, and be sure to provide a label and scale for each axis.



(b) Draw a boxplot for this distribution. Be sure to label the plots and provide a scale.



(c) What numerical measures of center and spread would be best to use for this distribution? Explain your choice.