Review 1

Name

1. Scientists estimate that the distribution of the life span of the Galápagos Islands giant tortoise is approximately normal with mean **100** years and standard deviation **15** years. Based on the estimate, which of the following is closest to the age of a Galápagos Islands giant tortoise at the **90th** percentile of the distribution?

(A) **80** years

(B) **115** years

(C) **120** years

(D) **125** years

(E) **130** years

2.

	Job	No Job	Total
Juniors	13	5	18
Seniors	13	26	39
Total	26	31	57

A survey of 57 students was conducted to determine whether or not they held jobs outside of school. The two-way table above shows the number of students by employment status (job, no job), and class (juniors, seniors). Which of the following best describes the relationship between employment status and class?

Review 1

(A) There appears to be no association, since the same number of juniors and seniors have jobs.

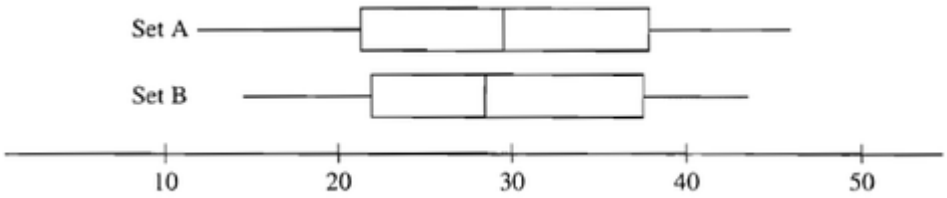
(B) There appears to be no association, since close to half of the students have jobs.

(C) There appears to be an association, since there are more seniors than juniors in the survey.

(D) There appears to be an association, since the proportion of juniors having jobs is much larger than the proportion of seniors having jobs.

(E) A measure of association cannot be determined from these data.

3.



The boxplots above summarize two data sets, A and B. Which of the following must be true?

1. Set A contains more data than Set B.
2. The box of Set A contains more data than the box of Set B.
3. The data in Set A have a larger range than the data in Set B.

(A) I only

(B) III only

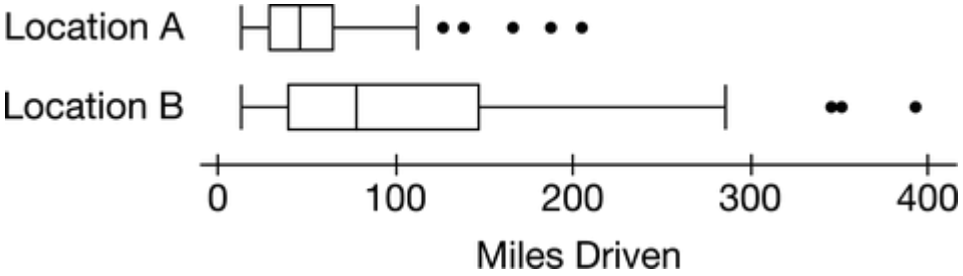
(C) I and II only

(D) II and III only

(E) I, II, and III

Review 1

4. A car rental agency has two locations in a city. The boxplots below summarize the miles driven for one day of single-day car rentals at each location.



Based on the boxplots, which statement provides the best comparison of the two locations?

(A) The number of single-day rentals is greater for location A than for location B.

(B) The number of single-day rentals is less for location A than for location B.

(C) Compared with location A, the miles driven for location B display more variability, and the median is greater.

(D) Compared with location A, the miles driven for location B display less variability, and the median is greater.

(E) Compared with location A, the miles driven for location B display less variability, and the median is about the same.

5. At a college the scores on the chemistry final exam are approximately normally distributed, with a mean of 75 and a standard deviation of 12. The scores on the calculus final are also approximately normally distributed, with a mean of 80 and a standard deviation of 8. A student scored 81 on the chemistry final and 84 on the calculus final. Relative to the students in each respective class, in which subject did this student do better?

Review 1

(A) The student did better in chemistry.

(B) The student did better in calculus.

(C) The student did equally well in each course.

(D) There is no basis for comparison, since the subjects are different from each other and are in different departments.

(E) There is not enough information for comparison, because the number of students in each class is not known.

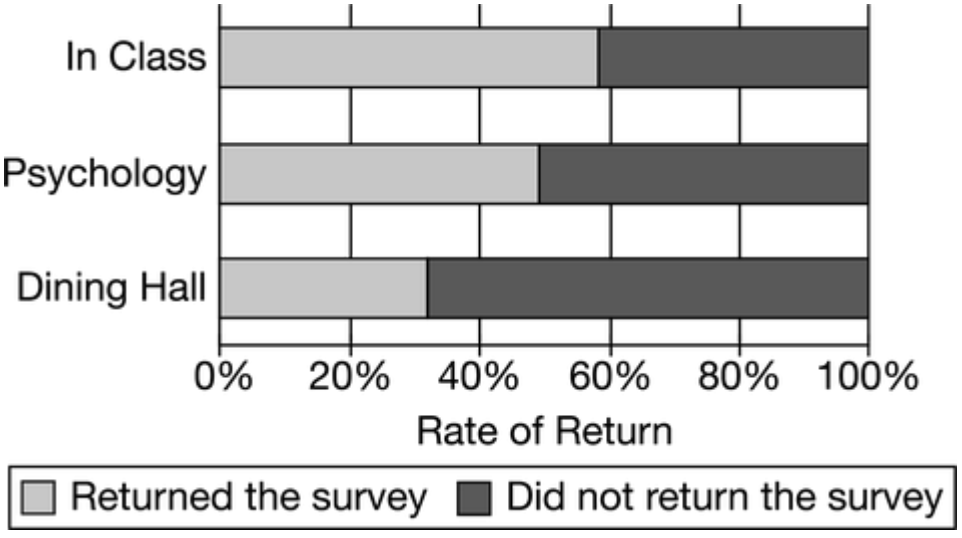
6. College researchers wanted to know under what conditions people are more likely to complete and return a survey. As part of a study, the researchers prepared three sets of identical surveys and used three methods of delivering and returning the surveys. The methods are described as follows.

• In Class: The surveys were given to students in a class, and students were asked to return completed surveys to their instructor.

• Psychology: The surveys were given to students participating in a psychology experiment, and students were asked to return completed surveys to a collection box in the hallway of the psychology building.

• Dining Hall: The surveys were given to students in the dining hall, and students were asked to return completed surveys to a collection box outside the dining hall.

The graph shows the percent of surveys returned and not returned for each delivery method.



Which statement about delivery method and rate of survey return is supported by the graph?

Review 1

(A) There is a positive association between delivery method and rate of return.

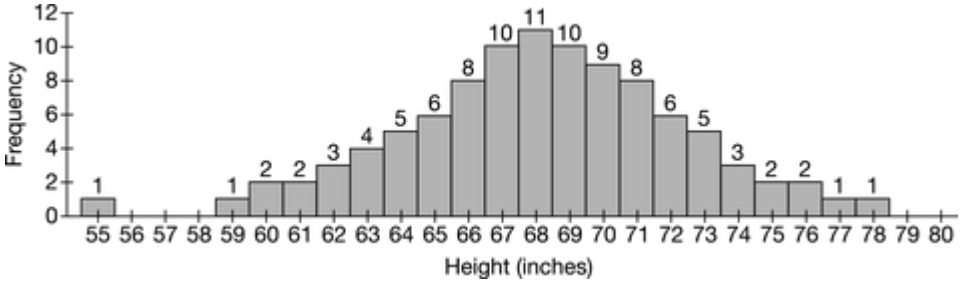
(B) There is a negative association between delivery method and rate of return.

(C) The number of surveys given using the Dining Hall delivery method was less than the number given using either of the other delivery methods.

(D) The Psychology delivery method displays the most symmetric results; the other delivery methods display skewed results.

(E) The In Class delivery method had the greatest rate of return, and the Dining Hall delivery method had the least rate of return.

7. The histogram shows the distribution of heights, in inches, of 100 adult men.



Based on the histogram, which of the following is closest to the interquartile range, in inches, of the distribution?

(A) 2

(B) 5

(C) 9

(D) 12

(E) 15

Review 1

The following question(s) refer to the following information.

Every Thursday, Matt and Dave's Video Venture has "roll-the-dice" day. A customer may choose to roll two fair dice and rent a second movie for an amount (in cents) equal to the numbers uppermost on the dice, with the larger number first. For example, if the customer rolls a two and a four, a second movie may be rented for \$0.42. If a two and two are rolled, a second movie may be rented for \$0.22. Let X represent the amount paid for a second movie on roll-the-dice day. The expected value of X is \$0.47 and the standard deviation of X is \$0.15.

8. If a customer rolls the dice and rents a second movie every Thursday for 30 consecutive weeks, what is the approximate probability that the total amount paid for these second movies will exceed \$15.00?

(A) 0

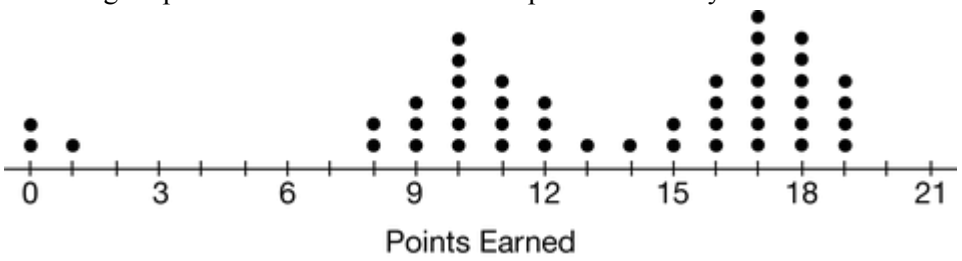
(B) 0.09

(C) 0.14

(D) 0.86

(E) 0.91

9. A group of students played a game in which they earned points for answering questions correctly. The following dotplot shows the total number of points earned by each student.



Which of the following is the best description of the distribution of points earned?

Review 1

(A) Approximately normal

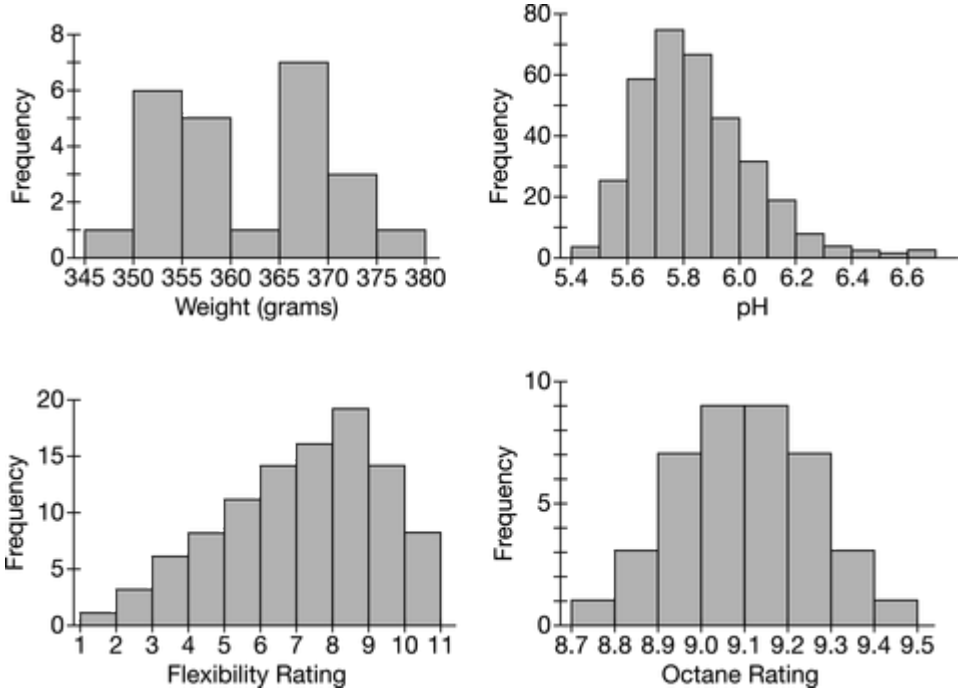
(B) Bimodal without a gap

(C) Bimodal with a gap

(D) Skewed to the right without a gap

(E) Skewed to the right with a gap

10. The distributions of four variables are shown in the following histograms.



Which of the following shapes is NOT represented by one of the four distributions?

Review 1

(A) Uniform

(B) Bimodal

(C) Skewed to the left

(D) Skewed to the right

(E) Symmetric and unimodal

11. Resting heart rates, in beats per minute, were recorded for two samples of people. One sample was from people in the age-group of 20 years to 30 years, and the other sample was from people in the age-group of 40 years to 50 years. The five-number summaries are shown in the table.

Age-Group (years)	Minimum	Q1	Median	Q3	Maximum
20 to 30	60	71	72	75	84
40 to 50	60	70	73	76	85

The values 60, 62, and 84 were common to both samples. The three values are identified as outliers with respect to the age-group 20 years to 30 years because they are either 1.5 times the interquartile range (IQR) greater than the upper quartile or 1.5 times the IQR less than the lower quartile.

Using the same method for identifying outliers, which of the three values are identified as outliers for the age-group 40 years to 50 years?

Review 1

(A) None of the three values is identified as an outlier.

(B) Only 60 is identified as an outlier.

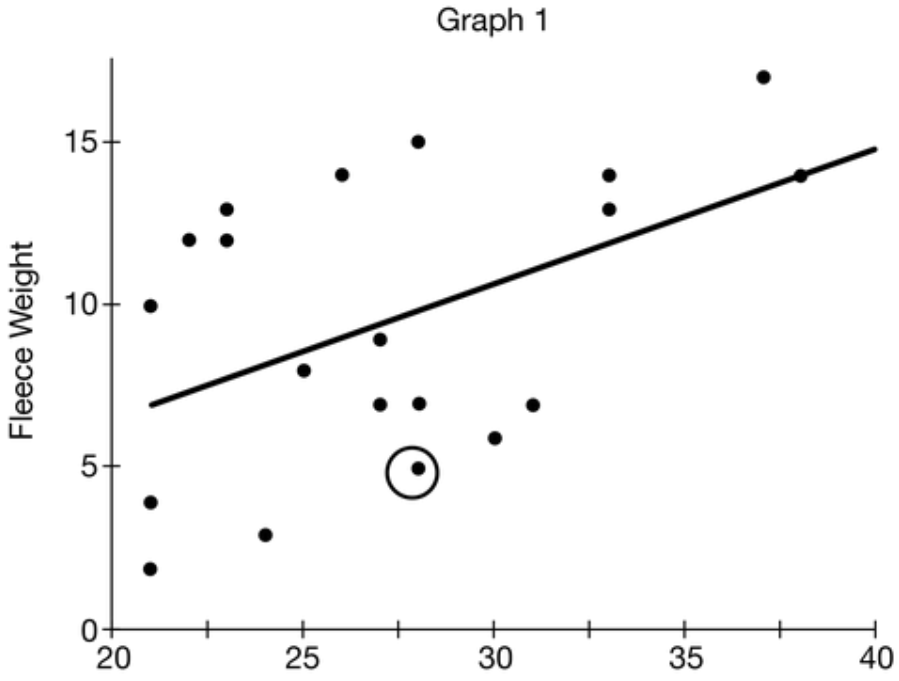
(C) Only 60 and 62 are identified as outliers.

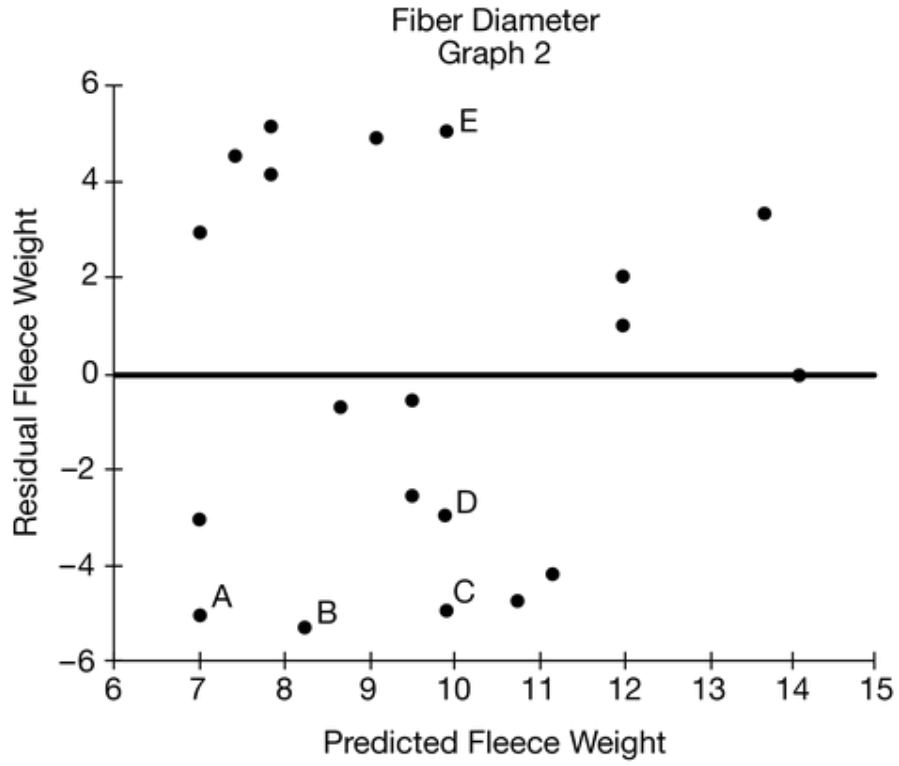
(D) Only 60 and 84 are identified as outliers.

(E) The three values are all identified as outliers.

Review 1

12. Data were collected on the fiber diameter and the fleece weight of wool taken from a sample of **20** sheep. The data are shown in the following graphs. Graph **1** is a scatterplot of fleece weight versus fiber diameter with the respective least-squares regression line shown. Graph **2** is the associated plot of the residuals versus the predicted values.





One point is circled on graph **1**. Five points labeled A, B, C, D, and E are identified on graph **2**. Which point on graph **2** represents the residual for the circled point on graph **1** ?

(A) A

(B) B

(C) C

(D) D

(E) E

13. The height and age of each child in a random sample of children was recorded. The value of the correlation coefficient between height and age for the children in the sample was 0.8. Based on the least-squares regression line created from the data to predict the height of a child based on age, which of the following is a correct statement?

(A) On average, the height of a child is 80% of the age of the child.

(B) The least-squares regression line of height versus age will have a slope of 0.8.

(C) The proportion of the variation in height that is explained by a regression on age is 0.64.

(D) The least-squares regression line will correctly predict height based on age 80% of the time.

(E) The least-squares regression line will correctly predict height based on age 64% of the time.

14. There is a linear relationship between the number of chirps made by the striped ground cricket and the air temperature. A least squares fit of some data collected by a biologist gives the model

$\hat{y} = 25.2 + 3.3x \qquad 9 < x < 25,$

where x is the number of chirps per minute and $\hat{y}$ is the estimated temperature in degrees Fahrenheit. What is the estimated increase in temperature that corresponds to an increase of 5 chirps per minute?

Review 1

(A) 3.3 ° F

(B) 16.5 ° F

(C) 25.2 ° F

(D) 28.5 ° F

(E) 41.7 ° F

15. The back-to-back stem-and-leaf plot below gives the percentage of students who dropped out of school at each of the 49 high schools in a large metropolitan school district.

School Year 1989-1990		School Year 1992-1993
	0	4
9 9 9 9 8 8 7	0	5 6 6 6 7 7 7 8 8 8 9 9
4 4 4 4 4 3 3 2 2 2 2 1 1 1 1 0	1	0 0 0 0 1 1 1 1 2 2 2 3 3 4 4 4 4
9 9 9 7 7 6 6 6 6 5	1	5 5 5 6 6 6 6 7 7 7 7 8
4 2 2 2 1 0 0	2	1 3
8 8 8 7 6	2	
2	3	0 1 1 2
7 6 6	3	5
	4	

For 1992-1993, 1|2 represents 12%.

Which of the following statements is NOT justified by these data?

Review 1

(A) The drop-out rate decreased in each of the 49 schools between the 1989-90 and 1992-1993 school years.

(B) For the school years shown, most students in the 49 schools did not drop out of high school.

(C) In general drop-out rates decreased between the 1989-90 and 1992-1993 school years.

(D) The median drop-out rate of the 49 high schools decreased between the 1989-90 and 1992-1993 school years.

(E) The spread between the schools with the lowest drop-out rates and those with the highest drop-out rates did not change much between the 1989-90 and 1992-1993 school years.

16. Gina's doctor told her that the standardized score (z- score) for her systolic blood pressure, as compared to the blood pressure of other women her age, is 1.50. Which of the following is the best interpretation of this standardized score?

(A) Gina's systolic blood pressure is 150.

(B) Gina's systolic blood pressure is 1.50 standard deviations above the average systolic blood pressure of women her age.

(C) Gina's systolic blood pressure is 1.50 above the average systolic blood pressure of women her age.

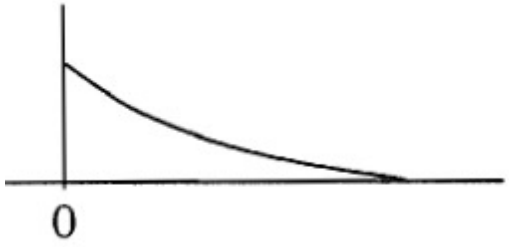
(D) Gina's systolic blood pressure is 1.50 times the average systolic blood pressure for women her age.

(E) Only 1.5% of women Gina's age have a higher systolic blood pressure than she does.

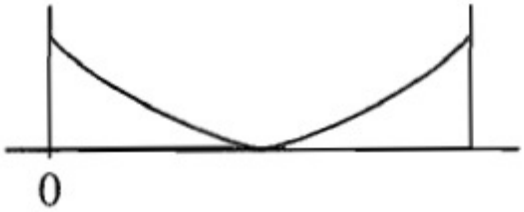
17. For which of the following distributions is the mean greater than the median?

Review 1

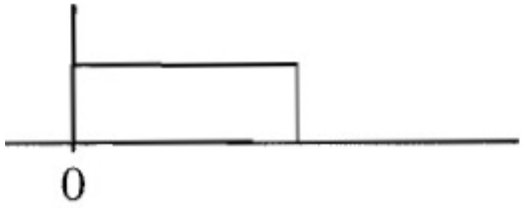
A



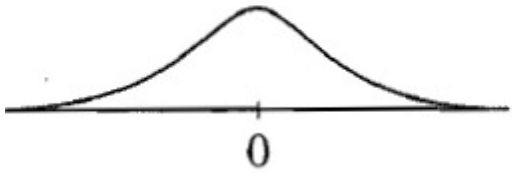
B



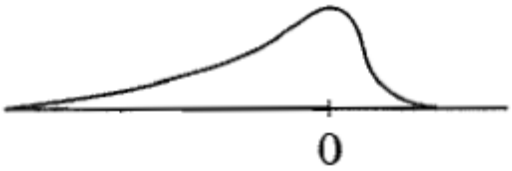
C



D



E



18. The distribution of number of hours worked by volunteers last year at a large hospital is approximately normal with mean 80 and standard deviation 7. Volunteers in the top 20 percent of hours worked will receive a certificate of merit. If a volunteer from last year is selected at random, which of the following is closest to the probability that the volunteer selected will receive a certificate of merit given that the number of hours the volunteer worked is less than 90 ?

A 0.077

B 0.123

C 0.134

D 0.618

E 0.923

19. The distribution of the weights of loaves of bread from a certain bakery follows approximately a normal distribution. Based on a very large sample, it was found that 10 percent of the loaves weighed less than 15.34 ounces, and 20 percent of the loaves weighed more than 16.31 ounces. What are the mean and standard deviation of the distribution of the weights of the loaves of bread?

A $\mu = 15.82, \sigma = 0.48$

B $\mu = 15.82, \sigma = 0.69$

C $\mu = 15.87, \sigma = 0.50$

D $\mu = 15.93, \sigma = 0.46$

E $\mu = 16.00, \sigma = 0.50$

20. The heights of adult women are approximately normally distributed about a mean of 65 inches with a standard deviation of 2 inches. If Rachael is at the 99th percentile in height for adult women, then her height, in inches, is closest to

(A) 60

(B) 62

(C) 68

(D) 70

(E) 74

21. Sean and Evan are college roommates who have part-time jobs as servers in restaurants. The distribution of Sean's weekly income is approximately normal with mean $225 and standard deviation $25. The distribution of Evan's weekly income is approximately normal with mean $240 and standard deviation $15. Assuming their weekly incomes are independent of each other, which of the following is closest to the probability that Sean will have a greater income than Evan in a randomly selected week?

(A) 0.067

(B) 0.159

(C) 0.227

(D) 0.303

(E) 0.354

Review 1

22. A random sample of 1,092 people were asked whether color was a consideration in buying a new car. They were also asked to identify one additional feature that was important. The responses are shown in the table.

	Color Consideration			
	Yes	No	Maybe	Total
Comfort	40	96	12	148
Cost	108	68	8	184
Performance	62	62	12	136
Reliability	128	116	4	248
Safety	152	192	32	376
Total	490	534	68	1,092

Which of the following is closest to the proportion of people who responded no to color consideration and who identified safety as the additional feature that was important?

(A) 0.18

(B) 0.34

(C) 0.36

(D) 0.49

(E) 0.51

23. A company wanted to determine the health care costs of its employees. A sample of 25 employees were interviewed and their medical expenses for the previous year were determined. Later the company discovered that the highest medical expense in the sample was mistakenly recorded as 10 times the actual amount. However, after correcting the error, the corrected amount was still greater than or equal to any other medical expense in the sample. Which of the following sample statistics must have remained the same after the correction was made?

Review 1

(A) Mean

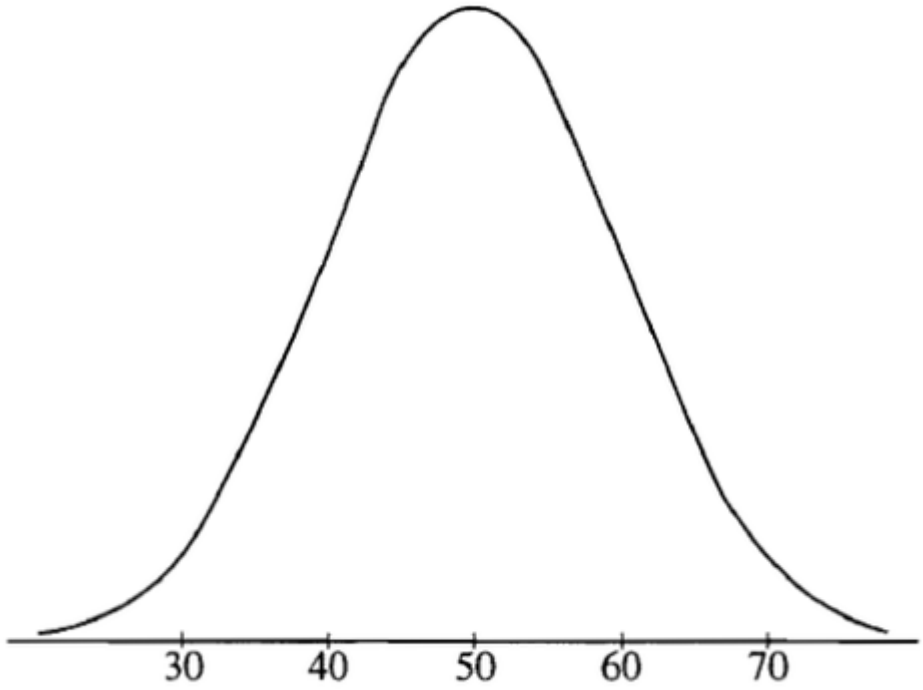
(B) Median

(C) Mode

(D) Range

(E) Variance

24.



Which of the following is the best estimate of the standard deviation of the distribution shown in the figure above?

Review 1

(A) 5

(B) 10

(C) 30

(D) 50

(E) 60

25. A grocery store receives deliveries of corn from two farms, one in Iowa and the other in Ohio. Both farms produce ears of corn with mean weight 1.26 pounds. The standard deviation of the weights of the ears of corn from the farm in Ohio is 0.01 pound greater than that from the farm in Iowa. A randomly selected ear of corn from the farm in Iowa weighed 1.39 pounds, which has a standardized score of 1.645 for the distribution of weights for the Iowa corn. If an ear of corn from the farm in Ohio weighs 1.39 pounds, how many standard deviations from the mean is the weight with respect to the Ohio distribution?

(A) 1.46 standard deviations below the mean

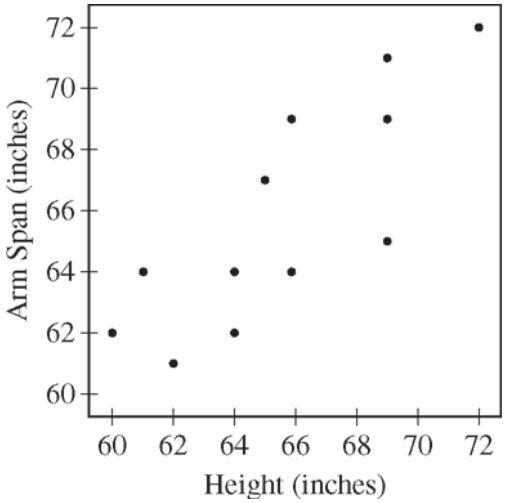
(B) 1.46 standard deviations above the mean

(C) 1.65 standard deviations above the mean

(D) 1.88 standard deviations below the mean

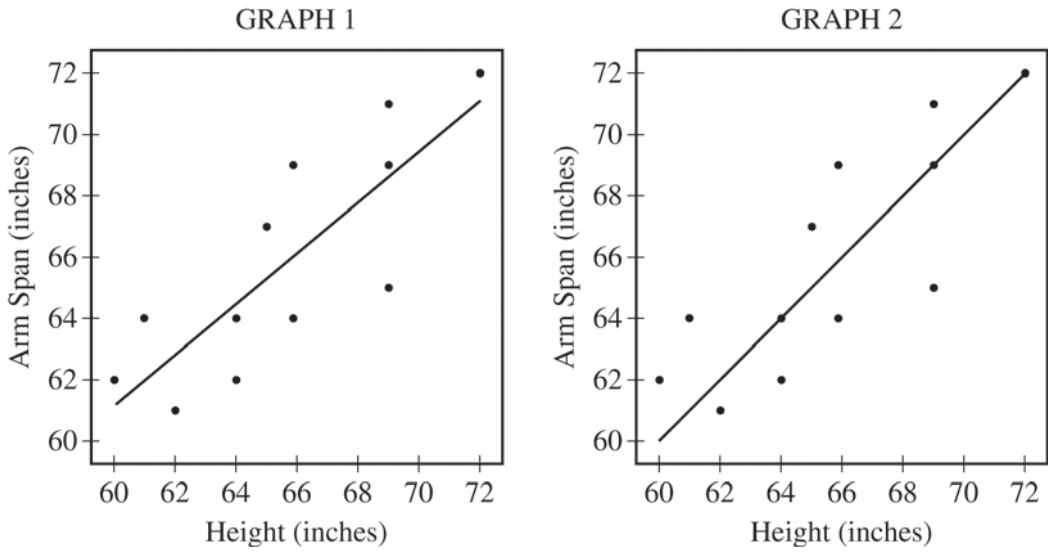
(E) 1.88 standard deviations above the mean

5. A student measured the heights and the arm spans, rounded to the nearest inch, of each person in a random sample of 12 seniors at a high school. A scatterplot of arm span versus height for the 12 seniors is shown.



(a) Based on the scatterplot, describe the relationship between arm span and height for the sample of 12 seniors.

Let x represent height, in inches, and let y represent arm span, in inches. Two scatterplots of the same data are shown below. Graph 1 shows the data with the least squares regression line $\hat{y} = 11.74 + 0.8247x$, and graph 2 shows the data with the line $y = x$.



(b) The criteria described in the table below can be used to classify people into one of three body shape categories: square, tall rectangle, or short rectangle.

Square	Tall Rectangle	Short Rectangle
Arm span is equal to height.	Arm span is less than height.	Arm span is greater than height.

(i) For which graph, 1 or 2, is the line helpful in classifying a student's body shape as square, tall rectangle, or short rectangle? Explain.

(ii) Complete the table of classifications for the 12 seniors.

Classification	Square	Tall Rectangle	Short Rectangle
Frequency			

(c) Using the best model for prediction, calculate the predicted arm span for a senior with height 61 inches.

1. An administrator at a large university is interested in determining whether the residential status of a student is associated with level of participation in extracurricular activities. Residential status is categorized as on campus for students living in university housing and off campus otherwise. A simple random sample of 100 students in the university was taken, and each student was asked the following two questions.

   - Are you an on campus student or an off campus student?
   - In how many extracurricular activities do you participate?

   The responses of the 100 students are summarized in the frequency table shown.

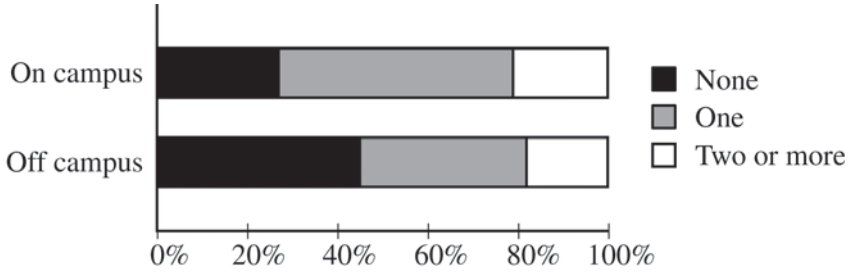
	Residential Status		
Level of Participation in Extracurricular Activities	On campus	Off campus	Total
No activities	9	30	39
One activity	17	25	42
Two or more activities	7	12	19
Total	33	67	100

   (a) Calculate the proportion of on campus students in the sample who participate in <u>at least one</u> extracurricular activity and the proportion of off campus students in the sample who participate in <u>at least one</u> extracurricular activity.

   On campus proportion:

   Off campus proportion:

   The responses of the 100 students are summarized in the segmented bar graph shown.



   (b) Write a few sentences summarizing what the graph reveals about the association between residential status and level of participation in extracurricular activities among the 100 students in the sample.

GO ON TO THE NEXT PAGE.

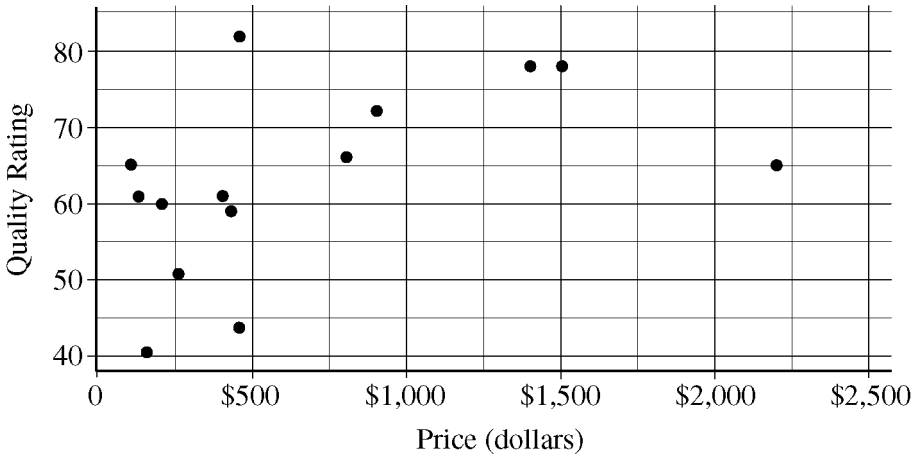
(c) After verifying that the conditions for inference were satisfied, the administrator performed a chi-square test of the following hypotheses.

H_0 : There is no association between residential status and level of participation in extracurricular activities among the students at the university.

H_a : There is an association between residential status and level of participation in extracurricular activities among the students at the university.

The test resulted in a *p*-value of 0.23. Based on the *p*-value, what conclusion should the administrator make?

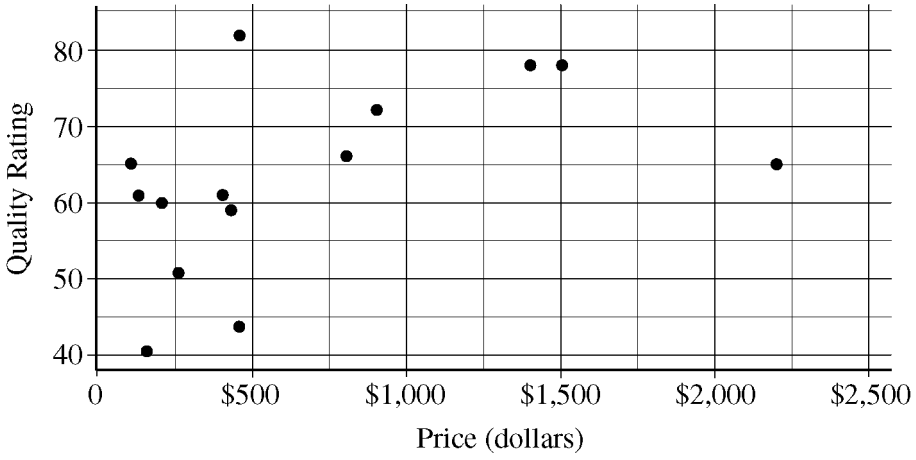
1. The scatterplot below displays the price in dollars and quality rating for 14 different sewing machines.



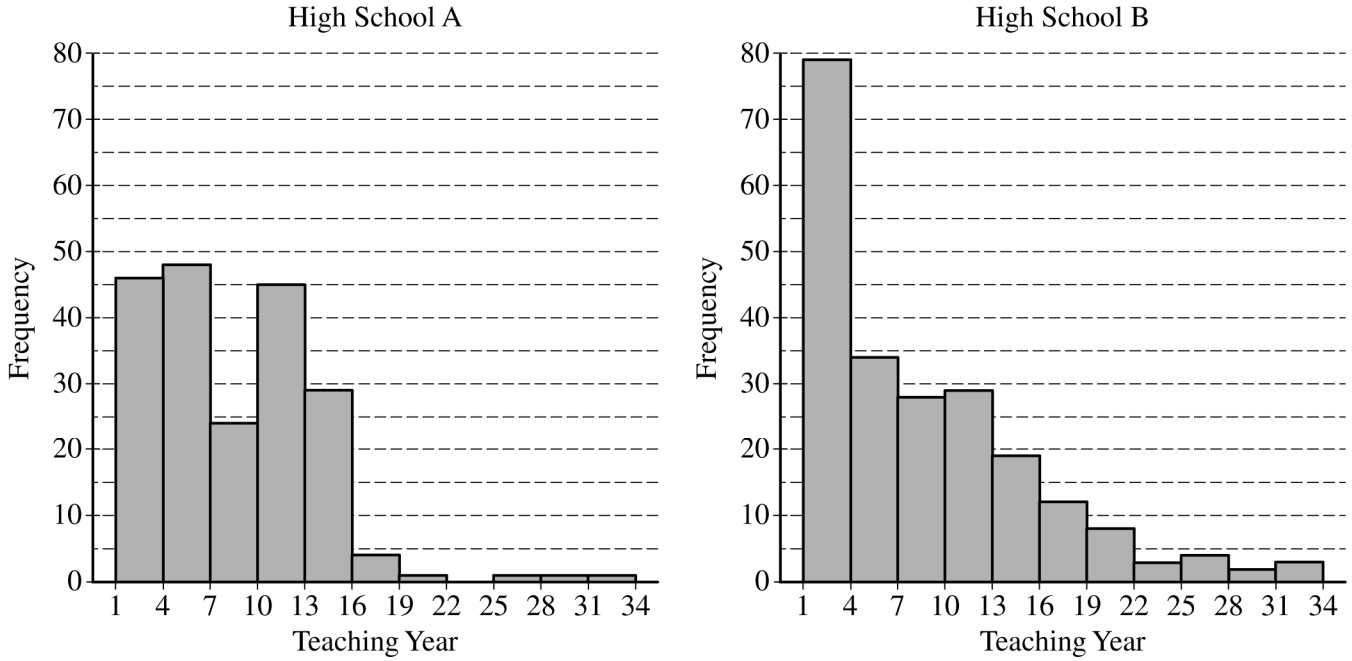
(a) Describe the nature of the association between price and quality rating for the sewing machines.

(b) One of the 14 sewing machines substantially affects the appropriateness of using a linear regression model to predict quality rating based on price. Report the approximate price and quality rating of that machine and explain your choice.

(c) Chris is interested in buying one of the 14 sewing machines. He will consider buying only those machines for which there is no other machine that has both higher quality and lower price. On the scatterplot reproduced below, circle all data points corresponding to machines that Chris will consider buying.



5. The following histograms summarize the teaching year for the teachers at two high schools, A and B.



Teaching year is recorded as an integer, with first-year teachers recorded as 1, second-year teachers recorded as 2, and so on. Both sets of data have a mean teaching year of 8.2, with data recorded from 200 teachers at High School A and 221 teachers at High School B. On the histograms, each interval represents possible integer values from the left endpoint up to but not including the right endpoint.

(a) The median teaching year for one high school is 6, and the median teaching year for the other high school is 7. Identify which high school has each median and justify your answer.

(b) An additional 18 teachers were not included with the data recorded from the 200 teachers at High School A. The mean teaching year of the 18 teachers is 2.5. What is the mean teaching year for all 218 teachers at High School A?

(c) The standard deviation of the teaching year for the 221 teachers at High School B is 7.2. If one teacher is selected at random from High School B, what is the probability that the teaching year for the selected teacher will be within 1 standard deviation of the mean of 8.2 ? Justify your answer.