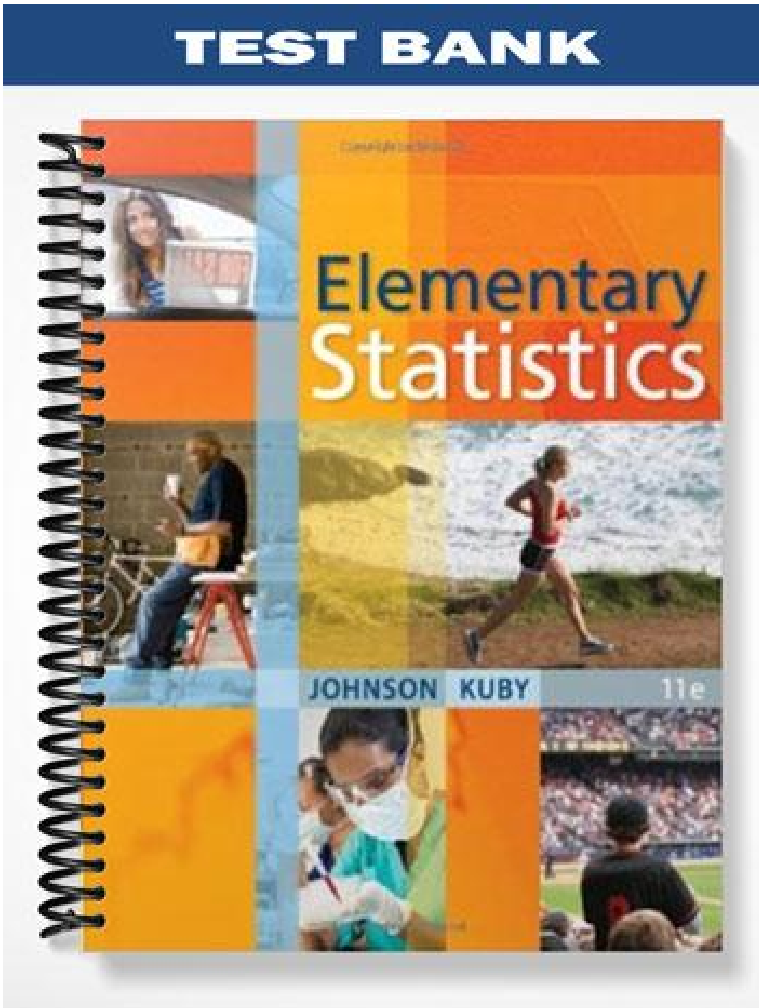

TEST BANK
Elementary
Statistics
JOHNSON KUBY
11e

SHORT ANSWER

1. Circle graphs and bar graphs are graphs that are used to summarize qualitative, or attribute, or categorical data.

   ANS:
   T

   PTS: 1

2. Circle graphs (pie diagrams) show the amount of data that belong to each category as a proportional part of a circle.

   ANS:
   T

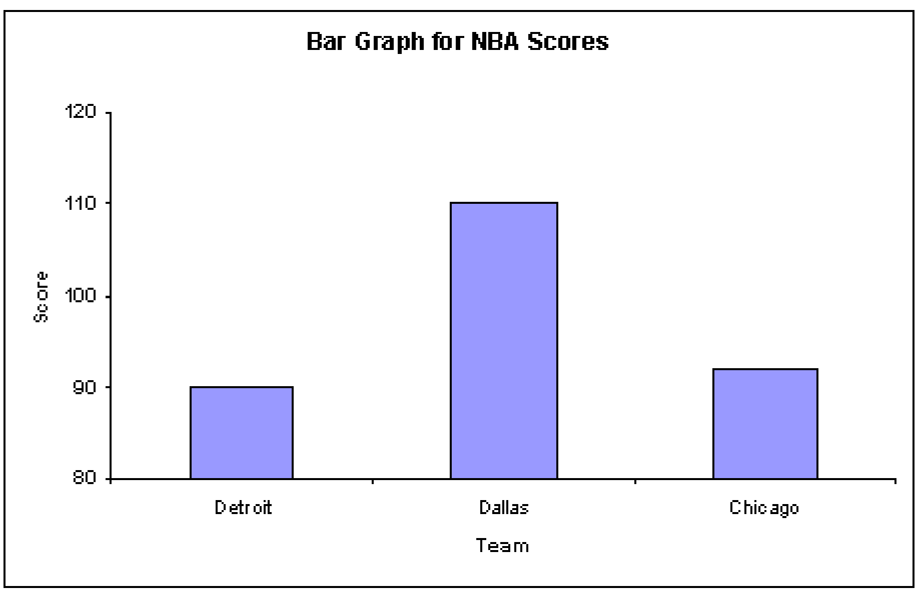
   PTS: 1

The points scored by the winning teams on opening night of a recent NBA season are shown in the table below:

Team	Detroit	Dallas	Chicago
Score	90	110	92

3. Draw a bar graph of these scores using a vertical scale ranging from 80 to 120.

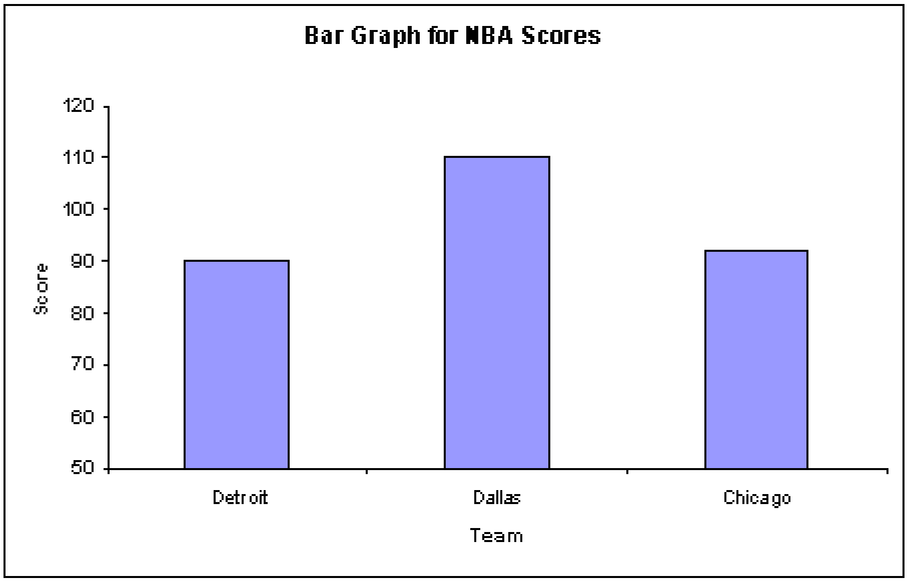
   ANS:



PTS: 1

4. Draw a bar graph of the scores using a vertical scale ranging from 50 to 120.

ANS:



PTS: 1

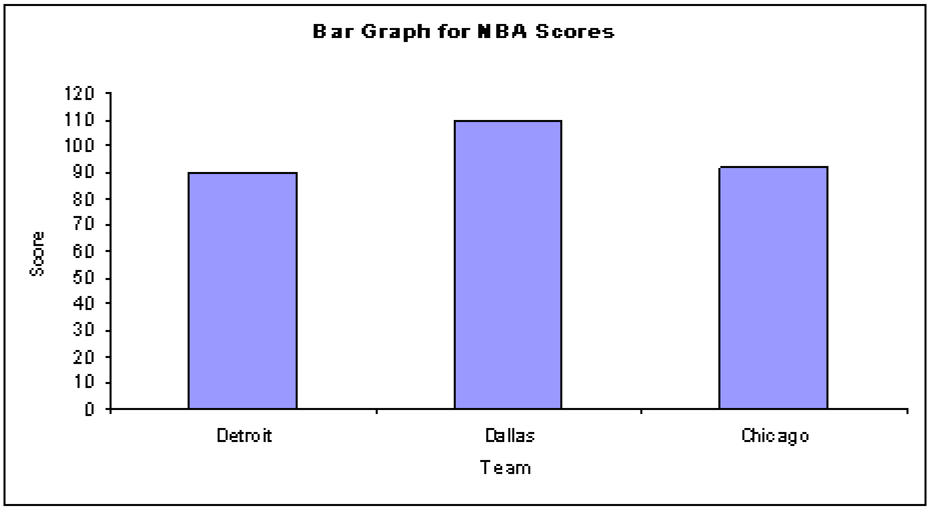
5. In which bar graph does it appear that the NBA scores vary more? Why?

ANS:
Bar graph in question 27 emphasizes the variation in the scores as it focuses only on the variation and not the relative size of the scores.

PTS: 1

6. How could you create an accurate representation of the relative size and variation between these scores? Draw this new bar graph.

ANS:
An accurate representation of both the size and variation of the values would be best served by starting the vertical scale at zero.



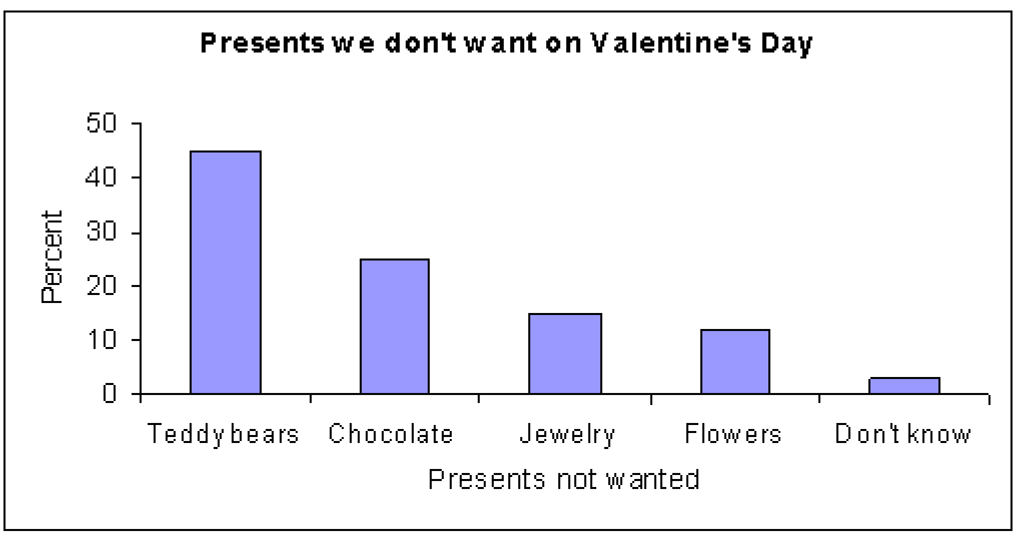
PTS: 1

What not to get them on Valentines Day! A recent study among adults in USA shows that adults prefer not to receive certain items as gifts on Valentine's Day as shown below:

Teddy bears: 45%; Chocolate: 25%; Jewelry: 15%; Flowers: 12%; Don't Know: 3%.

7. Draw a bar graph picturing the percentages of "Presents not wanted".

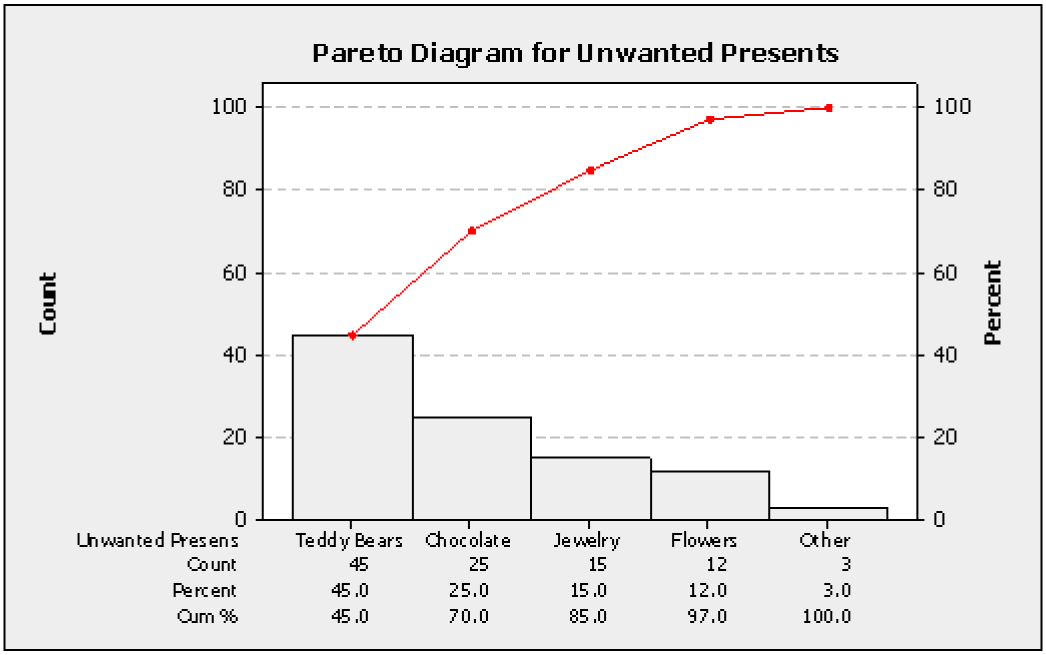
ANS:



PTS: 1

8. Draw a Pareto diagram picturing the "Presents not wanted".

ANS:



PTS: 1

9. If you want to be 80% sure you did not get your valentine something unwanted, what should you avoid buying? How does the Pareto diagram show this?

ANS:
Teddy bears, chocolates, jewelry; these are listed first in the Pareto diagram.

PTS: 1

10. 400 adults are to be surveyed, what frequencies would you expect to occur for each unwanted item listed on the snapshot?

ANS:
The frequencies are 180, 100, 60, 48, and 12 for teddy bears, chocolates, jewelry, flowers, and don't know, respectively.

PTS: 1

The final-inspection defect report for an assembly line is reported on the table and Pareto diagram as shown below:

Defect	Blemish	Scratch	Chip	Bend	Dent	Others
Count	61	50	28	17	13	11

11. What is the total defect count in the report?

ANS:
180 defects

PTS: 1

12. Find the percentage for “chip” defect items.

ANS:
Percent of chip = $(50/180) \cdot 100\% = 15.56\%$

PTS: 1

13. Find the “cum % for bend”, and explain what that value means.

ANS:
$[(61+50+28+17)/180] \cdot 100\% = (156/180) \cdot 100\% = 86.67\%$. The value 86.67% is the sum of the percentages for all defects that occurred more often than Bend, including Bend.

PTS: 1

14. Management has given the production line the goal of reducing their defects by 50%. What two defects would you suggest they give special attention to in working toward this goal? Explain.

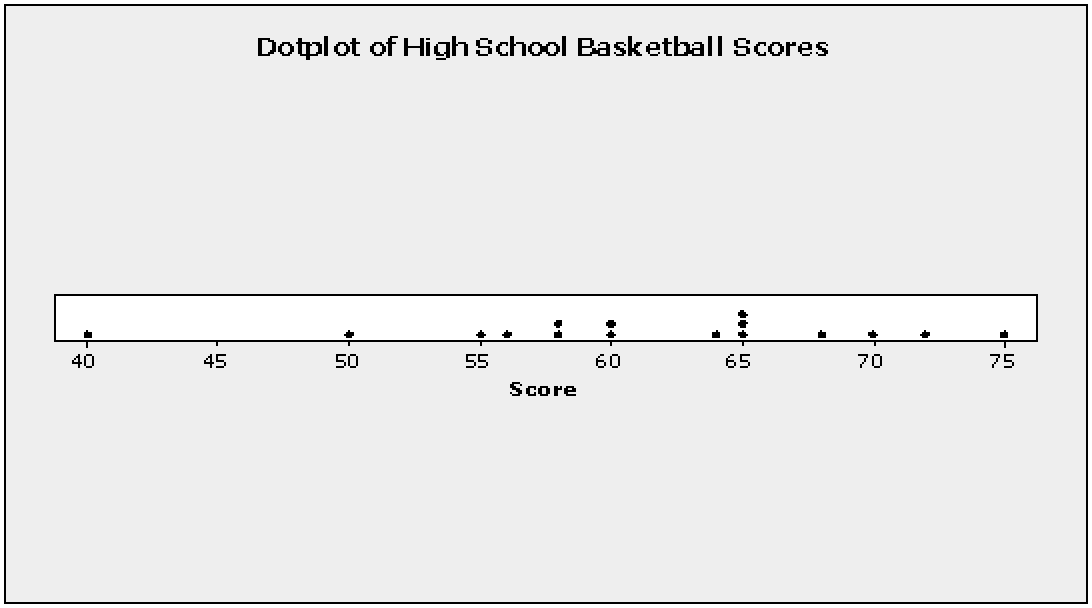
ANS:
The two defects, Blemish and Scratch, total 61.67%. If they can control these two defects, the goal should be within reach.

PTS: 1

The points scored during each game by the Big Rapids High School basketball team last season were: 60, 58, 65, 75, 50, 65, 60, 72, 64, 70, 58, 65, 56, 40, 68, and 55.

15. Construct a dotplot of these data.

ANS:



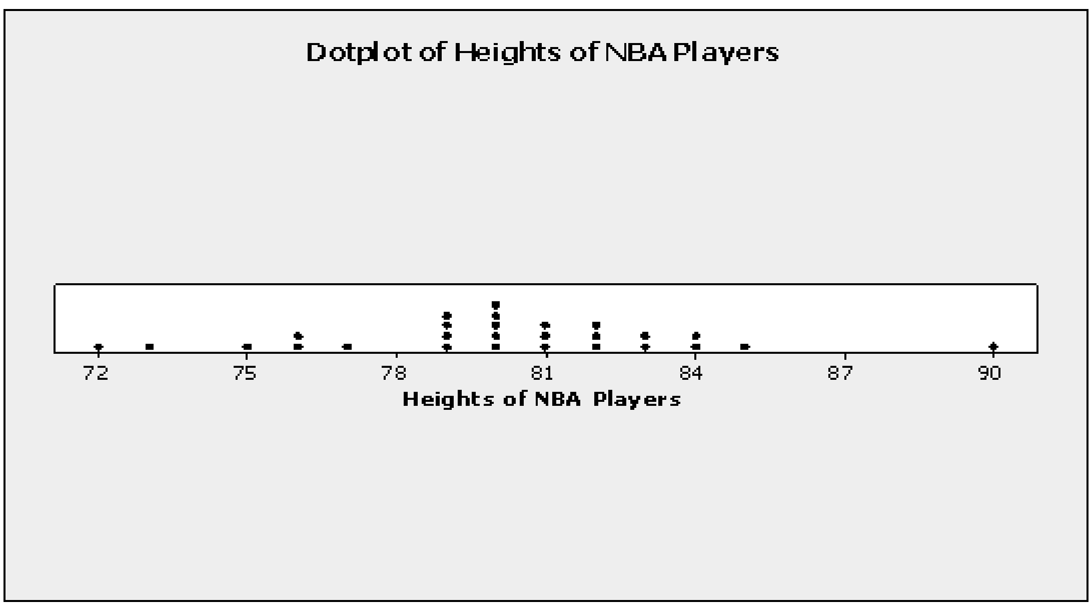
PTS: 1

The data shown below are the heights (in inches) of the basketball players who were the first round picks by the professional NBA teams in a recent year.

83	83	75	80	76	80	81	84	79	80
84	86	72	82	82	79	81	79	80	73
90	82	81	75	77	80	79	76	85	

16. Construct a dotplot of the heights of these players.

ANS:



PTS: 1

17. Use the dotplot in question 37 to uncover the shortest and the tallest players.

ANS:
The shortest player is 72 inches and the tallest player is 90 inches.

PTS: 1

18. Use the dotplot in question 37 to determine the most common height and how many players share that height?

ANS:
The most common height is 80 inches, shared by 5 players.

PTS: 1

19. What feature of the dotplot in question 37 illustrates the most common height?

ANS:
The height of column of dots illustrates the most common height.

PTS: 1

The players on a professional soccer team scored 40 goals during last season.

Player	1	2	3	4	5	6	7	8	9	10	11	12	13
Goals	2	7	3	2	2	5	2	1	6	2	3	2	3

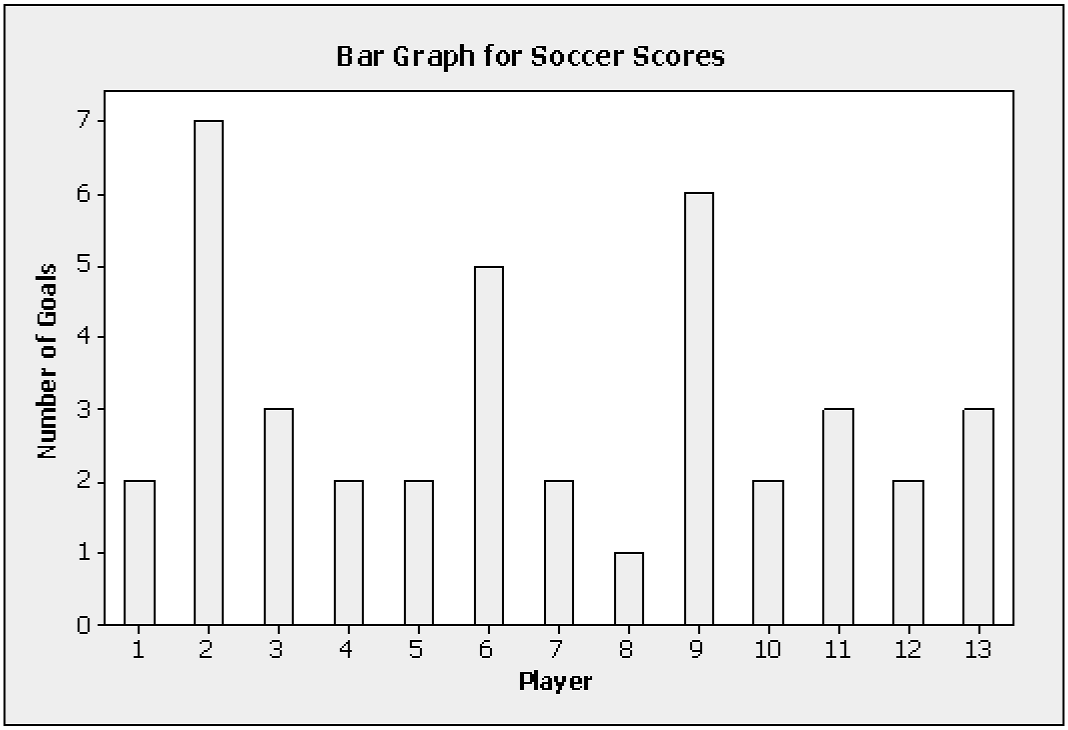
20. If you want to show the number of goals scored by each player, would it be more appropriate to display this information on a bar graph or a histogram? Explain.

ANS:
In order to show the number of goals scored by each player, it would be more appropriate to display this information on a bar graph

PTS: 1

21. Construct the appropriate graph for question 211.

ANS:



PTS: 1

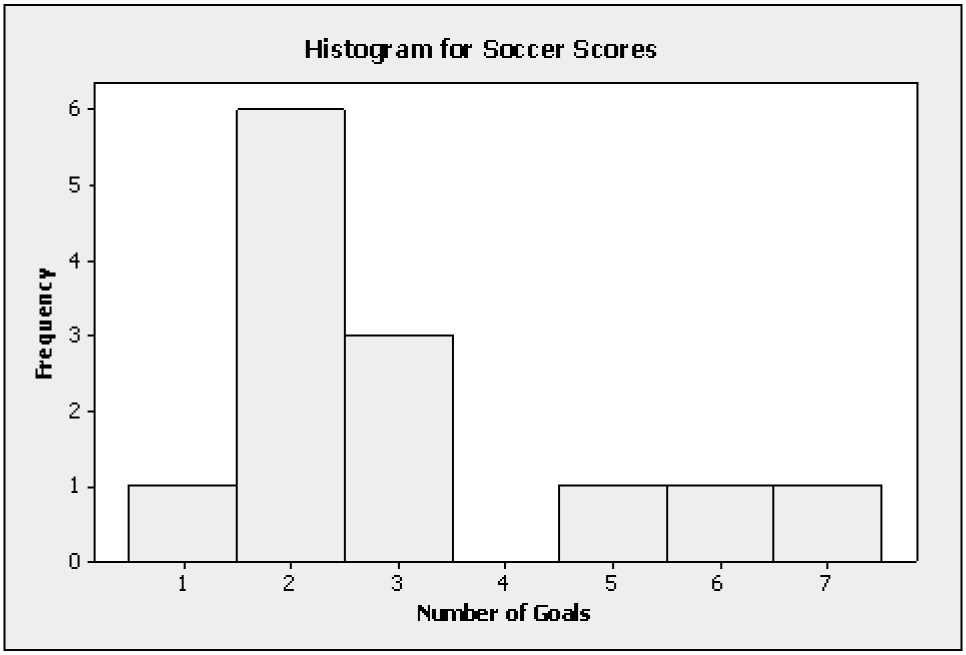
22. If you wanted to show (emphasize) the distribution of scoring by the team, would it be more appropriate to display this information on a bar graph or a histogram? Explain.

ANS:
If we want to emphasize the distribution of scoring by the team, it would be more appropriate to display this information on a histogram.

PTS: 1

23. Construct the appropriate graph for question 23.

ANS:



PTS: 1

The ages of 50 students who are attending a community college in Iowa are shown below:

20	20	19	21	21	22	19	19	21	19
18	21	19	18	22	21	24	20	24	17
21	19	22	19	18	20	23	19	19	20
19	20	21	22	21	20	22	20	21	20
21	19	21	21	19	19	20	19	19	19

24. Prepare an ungrouped frequency distribution of these ages.

ANS:

Age	17	18	19	20	21	23	23	24
Frequency	1	3	16	10	12	5	1	2

PTS: 1

25. Prepare an ungrouped relative frequency distribution of the same data.

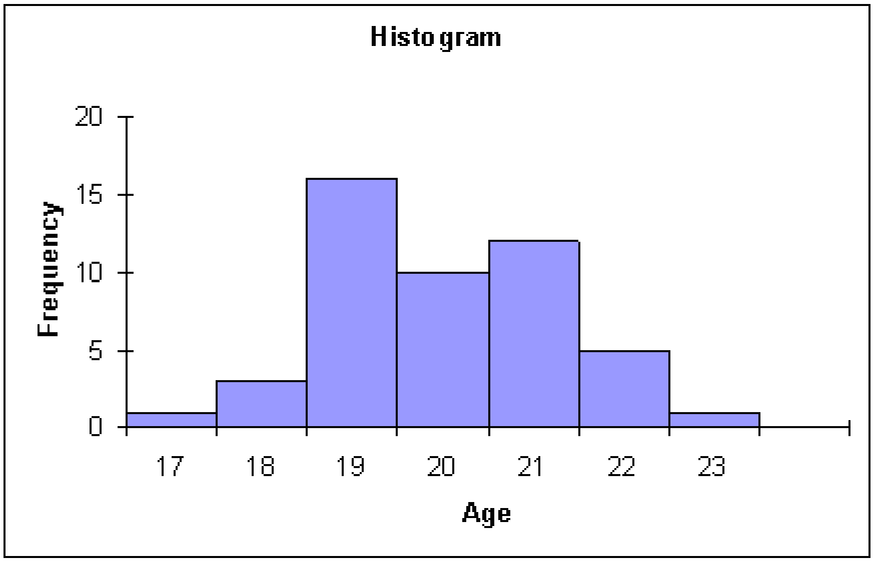
ANS:

Age	17	18	19	20	21	23	23	24
Rel. Freq.	0.02	0.06	0.32	0.20	0.24	0.10	0.02	0.04

PTS: 1

26. Prepare a frequency histogram of these data.

ANS:



PTS: 1

The following frequency distribution provides the number of managers and their annual salaries (in $1000):

Annual Salary ($1000)	15-25	25-35	35-45	45-55	55-65
Number of Managers	24	74	52	38	12

27. Prepare a cumulative frequency distribution for this frequency distribution.

ANS:

Class Boundaries	Cumulative Frequency
$15 \leq x \leq 25$	24
$15 \leq x \leq 25$	98
$15 \leq x \leq 25$	150
$15 \leq x \leq 25$	188
$15 \leq x \leq 25$	200

PTS: 1

28. Prepare a cumulative relative frequency distribution for this frequency distribution.

ANS:

Class Boundaries	Cumulative Frequency
$15 \leq x \leq 25$	0.12
$15 \leq x \leq 25$	0.49
$15 \leq x \leq 25$	0.75

$15 \le x \le 25$	0.94
$15 \le x \le 25$	1.00

PTS: 1

29. Explain why it is possible to find the mean for the data of a quantitative variable, but not for a qualitative variable.

ANS:
Quantitative variable results in numbers for which arithmetic is meaningful; qualitative variable does not

PTS: 1

Starting with a sample of two values 70 and 100, add three data values to your sample to obtain a new sample with certain statistics.

30. What are the three data values such that the new sample has a mean of 100? Justify your answer.

ANS:
Many different answers are possible. The sum of the five numbers needs to be 500; therefore we need any three numbers that total 330, such as 100, 110,120.Thus, the new sample mean $\bar{x} = 500 / 5 = 100$.

PTS: 1

31. What are the three data values such that the new sample has a median of 70? Justify your answer.

ANS:
Many different answers are possible. Need two numbers smaller than 70 and one number larger than 70. For example, we may choose 50, 60, and 80.Thus the five numbers are 50, 60, 70, 80, 100, and the median is 70.

PTS: 1

32. What are the three data values such that the new sample has a mode of 87? Justify your answer.

ANS:
Many different answers are possible. Need multiple 87's. For example, we may choose 87, 87 and 95. Thus, the five numbers are 70, 87, 87, 95, 100, and the mode = 87.

PTS: 1

33. What are the three data values such that the new sample has a midrange of 70? Justify your answer.

ANS:
Many different answers are possible. Need any two numbers that total 140 for the extreme values *L* and *H*, where one is 100 or larger. For example, we may choose the numbers 40, 50, and 60. Thus the five numbers are 40, 50, 60, 70, 100, and midrange = (*L*+*H*)/2 = (40+100)/2 = 70.

PTS: 1

34. What are the three data values such that the new sample has a mean of 100 and a median of 70? Justify your answer.

ANS:
Many different answers are possible. Need two numbers smaller than 70 and one number larger than 70 so that their total is 330. For example, we may choose the numbers 65, 65, and 200. Thus the five numbers are 65, 65, 70, 100, 200. Hence, $\bar{x}$ = 500/5 = 100, and the median is 70.

PTS: 1

35. What are the three data values such that the new sample has a mean of 100 and a mode of 87? Justify your answer.

ANS:
Many different answers are possible. Need two numbers of 87 and a number large enough so that the total of all five numbers is 500. Therefore the three numbers are 87, 87,156. The five numbers are 70, 87, 87, 100, 156. Thus the mode = 87, and $\bar{x}$ = 500 / 5 = 100.

PTS: 1

36. What are the three data values such that the new sample has a mean of 100, a median of 70, and a mode of 87? Justify your answer.

ANS:
Many different answers are possible. There must be two 87's in order to have a mode of 87, and there can only be two data values larger than 70 in order for 70 to be the median, which is impossible since 100 is one of the numbers, and that makes three of the five numbers larger than 70.

PTS: 1

37. Explain the meaning of the following statement “The data value x = 30 has a deviation value of 6.”

ANS:

The value x = 30 is 6 larger than the mean.

PTS: 1

38. Explain the meaning of the following statement "The data value x = 80 has a deviation value of -15."

39. Why the sum of the deviations, $\sum(x-\bar{x})$, is always zero?

ANS:
$\sum(x-\bar{x})$, is always zero because the deviations of x values smaller than the mean (which are negative values) cancel out x values larger than the mean (which are positive).

PTS: 1

Consider the following sample: 26, 49, 9, 42, 60, 11, 43, 26, 30, and 14.

39. Find the variance.

ANS:
294.89

PTS: 1

40. Find the standard deviation

ANS:
17.17

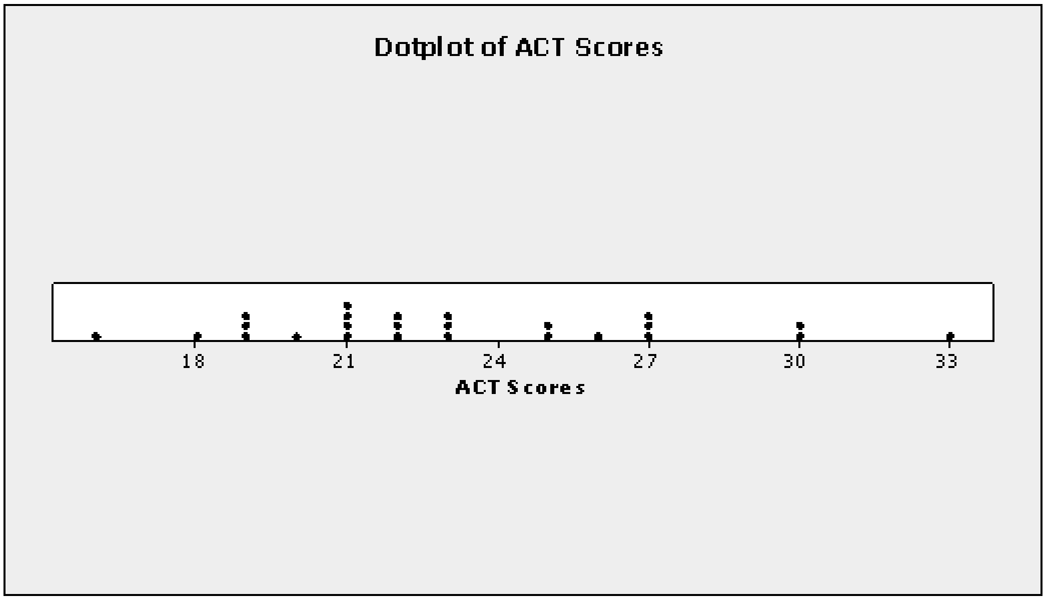
PTS: 1

Below are the ACT scores attained by the 25 members of a local high school graduating class.

23	26	25	19	33	21	21	22	21	27
19	25	18	23	22	30	27	27	23	16
21	19	20	30	22					

41. Draw a dotplot of the ACT scores.

ANS:



PTS: 1

42. Using the concept of depth, describe the position of 26 in the set of 25 ACT scores in two different ways.

ANS:
The data values in ascending are:

16	18	19	19	19	20	21	21	21	21
22	22	22	23	23	23	25	25	26	27
27	27	30	30	33					

Therefore, the value 26 is in the 19^{th} position from L = 16, and in the 7^{th} position from H = 33.

PTS: 1

43. Find P_5 for the ACT scores.

ANS:
nk / 100 =(25)(5) / 100 = 1.25. Hence, $d(P_5)$ = 2, and P_5 = 18

PTS: 1

44. Find P_{10} for the ACT scores.

ANS:
nk / 100 =(25)(10) / 100 = 2.5. Hence $d(P_{10})$ = 3, and P_{10} = 19

PTS: 1

45. Find P_{20} for the ACT scores.

ANS:
nk / 100 =(25)(20) / 100 = 5. Hence $d(P_{20})$ = 5.5, and P_{20} =(19+20)/2 = 19.5

PTS: 1

46. Find P_{99} for the ACT scores.

ANS:
Since *k* = 99 > 50, subtract 99 from 100 and use 100 – *k* in place of *k* to determine the depth, which is then counted from the largest-valued data *H*. Therefore, *n*(100 – *k*) / 100 = 25 (1) / 100 = 0.25; then *d*(P_{99}) = 1, and P_{99} = 33

PTS: 1

47. Find P_{90} for the ACT scores.

ANS:
Since *k* = 90 > 50, subtract 90 from 100 and use 100 – *k* in place of *k* to determine the depth, which is then counted from the largest-valued data *H*. Therefore, P_{90}: *n*(100 – *k*) / 100 = 25 (10) / 100 = 2.5; then *d*(P_{90}) = 3, and P_{90} = 30

PTS: 1

48. Find P_{80} for the ACT scores.

ANS:
Since *k* = 80 > 50, subtract 80 from 100 and use 100 – *k* in place of *k* to determine the depth, which is then counted from the largest-valued data *H*. Therefore, *n*(100 – *k*) / 100 = 25 (20) / 100 = 5; then *d*(P_{80}) = 5.5, and P_{80} = (27+27) / 2 = 27

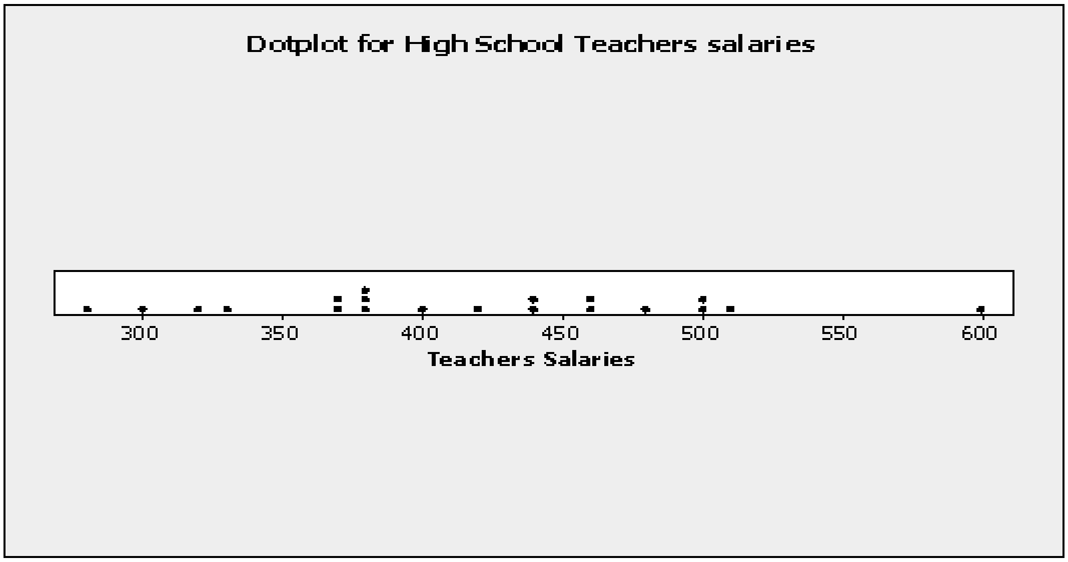
PTS: 1

The annual salaries (in $100) of high school teachers employed at one of the high schools in Kent County, Michigan are listed below:

600	440	461	419	397	477	464	275	507	497
332	373	440	373	501	382	377	301	323	383

49. Draw a dotplot of the salaries.

ANS:



PTS: 1

50. Using the concept of depth, describe the position of 332 in the set of 20 salaries in two different ways.

ANS:
The data values in ascending are:

275	301	323	332	373	373	377	382	383	397
419	440	440	461	464	477	497	501	507	600

Therefore, the value 332 is in the 4th position from L = 270, and in the 17th position from H = 33.

PTS: 1

51. Find the first quartile for these salaries, and interpret the result.

ANS:
nk / 100 =(20)(25) / 100 =5.0. Hence $d(Q_1)$ =5.5, and Q_1 = (373+373)/2 = 373 or $37,300. This means that at most 25% of high school teachers' salaries are lower than $37,300 and at most 75% are higher.

PTS: 1

52. Find the third quartile for these salaries, and interpret the result.

ANS:
Since k = 75 > 50, subtract 75 from 100 and use 100 – k in place of k to determine the depth, which is then counted from the largest-valued data H. Therefore,
n(100 – k) / 100 = 20 (25) / 100 = 5.0; then d (Q_3) = 5.5, and Q_3 = (464+477)/2 = 470.5 or $47,050.
This means that at most 75% of high school teachers' salaries are lower than $47,50 and at most 25% are higher.

PTS: 1

53. In general, the median, the midrange, and the midquartile are not necessarily the same
value. Each is the middle value, but by different definitions of “middle”. What property does the distribution need for these three measures to all be the same value?

ANS:
The distribution of the data needs to be symmetric for these three measures to all be the
same value.

PTS: 1

54. For a particular sample, the mean is 4.74, and the standard deviation is 3.10. What score in the sample has a *z*-score equal to –0.40?

ANS:

3.5

PTS: 1

55. What does it mean to say that *x* = 163 has a standard score of +1.60?

ANS:
It means that *x* = 163 is 1.60 standard deviations above the mean.

PTS: 1

56. What does it mean to say that a particular value of *x* has a *z* score of -1.94?

ANS:
It means that value of *x* is 1.94 standard deviations below the mean.

PTS: 1

57. In general, the standard score is a measure of what?

ANS:
The standard score is a measure of the number of standard deviations from the mean.

PTS: 1

58. In which of these situations (A, B, or C) is the *x*-value lowest in relation to the sample from which it comes? These samples come from three different populations.

Situation A: $x = 6$, $\bar{x} = 20.0$, $s = 9.0$
Situation B: $x = 350$, $\bar{x} = 400.0$, $s = 20.0$
Situation C: $x = 1.6$, $\bar{x} = 2.00$, $s = 0.30$

ANS:

In situations A, B, and C; z = -1.56, -2.50, and –1.33, respectively. In situation B we see the lowest z-score of –2.50. Therefore, the x-value in B is lowest in relation to the sample from which it comes.

PTS: 1

The mean lifetime of a certai**5** n tire is 50,000 miles and the standard deviation is 2,500 miles.

59. If we assume the mileages are normally distributed, approximately what percentage of all such tires will last between 42,500 and 57,500 miles?

ANS:
According to the Empirical Rule, approximately 99.7% of all such tires will last between 42,500 and 57,500 miles (i.e., within three standard deviations of the mean).

PTS: 1

60. If we assume nothing about the shape of distribution, approximately what percentage of all such tires will last between 42,500 and 57,500 miles?

ANS:
According to Chebyshev's Theorem, at least 89% of all such tires will last between 42,500 and 57,500 miles (i.e., within three standard deviations of the mean).

PTS: 1

The average clean-up time for a crew of a medium-size firm is 80.0 hours and the standard deviation is 6.5 hours. Assuming that the Empirical Rule is appropriate.

61. What proportion of the time will it take the clean-up crew 93.0 or more hours to clean the plant?

ANS:
z = (93 -80) / 6.5 = 2. Therefore, 93.0 is 2 standard deviations above the mean. Hence, 2.5% of the time more than 93.0 hours will be required.

PTS: 1

62. The total clean-up time will fall within what interval 95% of the time?

ANS:
95% of the time, the total clean-up time will fall within 2 standard deviations of the mean; that is 80.0 $\pm$ 2 (6.5) or from 67 to 93 hours.

PTS: 1

63. According to Chebyshev's Theorem, what percent of a set of data will be more than three standard deviations from the mean?

ANS:

About 11%

PTS: 1

Chebyshev’s Theorem can be stated in an equivalent form to that given in your book. For example, to say “at least 75% of the data fall within two standard deviations of the mean” is equivalent to stating that “at most, 25% will be more than two standard deviations away from the mean”.

64. At most, what percentage of a distribution will be three or more standard deviations from the mean?

ANS:
At most 11%

PTS: 1

65. At most, what percentage of a distribution will be four or more standard deviations from the mean?

ANS:
At most 6.25%

PTS: 1