Stats Modeling the World 3rd Edition Bock Test Bank

Exam

Name_____

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

1) Which of those variables is most likely to follow a Normal model?   1) _____
   A) number of cigarettes smoked daily
   B) number of TV sets at home
   C) head circumference
   D) eye color
   E) hours of homework last week

   Answer: C
   Explanation: A)
                   B)
                   C)
                   D)
                   E)

2) Suppose that a Normal model described student scores in a history class. Parker has a standardized   2) _____
   score (z-score) of +2.5. This means that Parker
   A) has a score that is 2.5 times the average for the class.
   B) has a standard deviation of 2.5.
   C) none of these
   D) is 2.5 points above average for the class.
   E) is 2.5 standard deviations above average for the class.

   Answer: E
   Explanation: A)
                   B)
                   C)
                   D)
                   E)

3) The five-number summary of credit hours for 24 students in a statistics class is: 3) _____

Min	Q1	Median	Q3	Max
13.0	15.0	16.5	18.0	22.0

   From this information we know that
   A) there are both low and high outliers in the data.
   B) none of these
   C) there are no outliers in the data.
   D) there is at least one low outlier in the data.
   E) there is at least one high outlier in the data.

   Answer: C
   Explanation: A)
                   B)
                   C)
                   D)
                   E)

1

4) Which of those variables is most likely to be bimodal?　　　　　　　　　　　　　　　　4) _____
   A) eye color
   B) number of TV sets at home
   C) number of cigarettes smoked daily
   D) hours of homework last week
   E) head circumference

Answer: C
Explanation: A)
                B)
                C)
                D)
                E)

5) Which variable about German Shepherds is most likely to be described by a Normal model?　　5) _____
   A) veterinary costs
   B) weight
   C) number of days housed
   D) age
   E) breed

Answer: B
Explanation: A)
                B)
                C)
                D)
                E)

6) The advantage of making a stem-and-leaf display instead of a dotplot is that a stem-and-leaf　　6) _____
      display
   A) preserves the individual data values.
   B) satisfies the area principle.
   C) none of these
   D) shows the shape of the distribution better than a dotplot.
   E) A stem-and-leaf display is for quantitative data, while a dotplot shows categorical data.

Answer: A
Explanation: A)
                B)
                C)
                D)
                E)

7) Which of the following variables would most likely follow a Normal model?　　　　　　　　7) _____
   A) scores on an easy test
   B) heights of singers in a co-ed choir
   C) all of these
   D) family income
   E) weights of adult male elephants

Answer: E
Explanation: A)
                B)
                C)
                D)
                E)

8) If we want to discuss any gaps and clusters in a data set, which of the following should not be chosen to display the data set?

8) _____

    A) stem-and-leaf plot
    B) any of these would work
    C) dotplot
    D) histogram
    E) boxplot

Answer: E
Explanation: A)
                    B)
                    C)
                    D)
                    E)

9) The SPCA collects the following data about the dogs they house. Which data is categorical?

9) _____

    A) age
    B) weight
    C) veterinary costs
    D) breed
    E) number of days housed

Answer: D
Explanation: A)
                    B)
                    C)
                    D)
                    E)

10) We collect these data from 50 male students. Which variable is categorical?

10) _____

    A) eye color
    B) number of cigarettes smoked daily
    C) number of TV sets at home
    D) head circumference
    E) hours of homework last week

Answer: A
Explanation: A)
                    B)
                    C)
                    D)
                    E)

11) The best estimate of the standard deviation of the mens' weights displayed in this dotplot is   11) _____


```
100        140        180        220
```
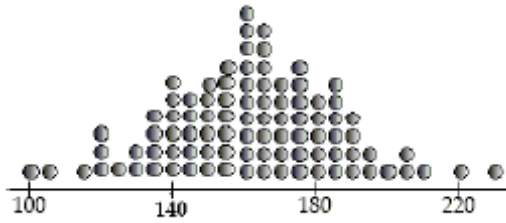
   A) 25   B) 35   C) 10   D) 40   E) 15
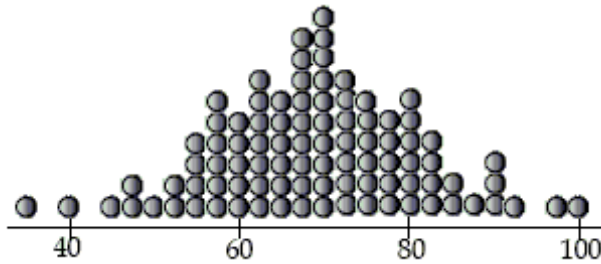
Answer: A
Explanation: A)
                B)
                C)
                D)
                E)

12) The standard deviation of the data displayed in this dotplot is most likely to be…   12) _____


```
     40         60         80        100
```

   A) 12   B) 8   C) 20   D) 5   E) 18

Answer: A
Explanation: A)
                B)
                C)
                D)
                E)

13) Suppose that a Normal model describes fuel economy (miles per gallon) for automobiles and that a   13) _____
Saturn has a standardized score (z-score) of +2.2. This means that Saturns . . .
   A) get 2.2 times the gas mileage of the average car.
   B) achieve fuel economy that is 2.2 standard deviations better than the average car.
   C) get 2.2 miles per gallon.
   D) get 2.2 mpg more than the average car.
   E) have a standard deviation of 2.2 mpg.

Answer: B
Explanation: A)
                B)
                C)
                D)
                E)

4

14) School administrators collect data on students attending the school. Which of the following variables is quantitative?

14) _____

   A) whether the student is in AP* classes
   B) class (freshman, soph., junior, senior)
   C) none of these
   D) grade point average
   E) whether the student has taken the SAT

Answer: D
Explanation: A)
                B)
                C)
                D)
                E)

15) A professor has kept records on grades that students have earned in his class. If he wants to examine the percentage of students earning the grades A, B, C, D, and F during the most recent term, which kind of plot could he make?

15) _____

   A) pie chart   B) dotplot   C) histogram   D) timeplot   E) boxplot

Answer: A
Explanation: A)
                B)
                C)
                D)
                E)

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses from boys and girls.

	Male	Female	Total
Bus	30	34	64
Ride	37	45	82
Personal	19	23	42
Total	86	102	188

16) Do you think mode of transportation is independent of gender? Give statistical evidence to support your conclusion.

16) _____

Answer: The way students get to school does seem to be independent of gender. Overall, 34% of students ride the bus, compared to 35% of the boys and 33% of the girls. 44% of all students caught rides with someone and 22% used personal transportation, almost the same as the percentages for boys (43% and 22%) or girls (44% and 23%) separately. These data provide little indication of a difference in mode of transportation between boys and girls at this school.

Explanation:

17) One of the reasons that the Monitoring the Future (MTF) project was started was "to study changes in the beliefs, attitudes, and behavior of young people in the United States." Data are collected from 8th, 10th, and 12th graders each year. To get a representative nationwide sample, surveys are given to a randomly selected group of students. In Spring 2004, students were asked about alcohol, illegal drug, and cigarette use. Describe the W's, if the information is given. If the information is not given, state that it is not specified.

17) _____

• Who:

• What:

• When:

• Where:

• How:

• Why:

Answer: Who: 8th, 10th, and 12th graders
What: alcohol, illegal drug, and cigarette use
When: Spring 2004
Where: United States
How: survey
Why: "to study changes in the beliefs, attitudes, and behavior of young people in the United States"
Explanation:

18) The five-number summary for the weights (in pounds) of fish caught in a bass tournament is:

18) _____

Min	Q1	Median	Q3	Max
2.3	2.8	3.0	3.3	4.5

a. Would you expect the mean weight of all fish caught to be higher or lower than the median? Explain.
b. You caught 3 bass weighing 2.3 pounds, 3.9 pounds, and 4.2 pounds. Were any of your fish outliers? Explain.

Answer: a. Probably higher. The data appear to be skewed to the right.
b. IQR = 3.3 - 2.8 = 0.5. Since 1.5(IQR) = 0.75, the fences are 2.8 - 0.75 = 2.05 and 3.3 + 0.75 = 4.05. The fish weighing 4.2 pounds is more than 1.5 IQRs outside the quartiles, so it could be considered an outlier.
Explanation:

19) At www.census.gov you can create a "population pyramid" for any country. These pyramids are back-to-back histograms. This pyramid shows Mexico's 2000 female population and the census bureau's projection for 2050. Write a few sentences summarizing the changes that are forecast.

19) _____

Mexico

Source: U.S. Census Bureau, International Data Base.
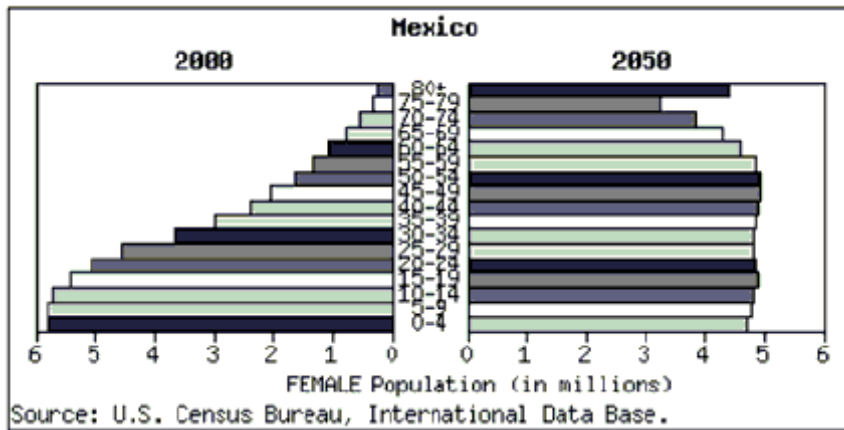
Answer: The Census Bureau projects dramatic changes in the female population of Mexico over the next 50 years. The current distribution of ages is strongly skewed to the right with most of the women under 30 and far fewer 50 and above. By 2050 the population will become more uniform across age groups from 0 to 60, and we anticipate an unusually large number of women over 80.

Explanation:

20) Cats and dogs The table shows whether students in an introductory statistics class like dogs and/or cats.

20) _____

		Like Dogs		
		Yes	No	Total
Like Cats	Yes	194	21	215
	No	110	10	120
	Total	304	31	335

a. What is the marginal distribution (in %) of "liking dogs"?
b. What is the conditional distribution (in %) of "liking dogs" for students who like cats?
c. What kind of display(s) would you use to examine the association between "liking dogs" and "liking cats"? (Just name a graph(s).)
d. Do "liking dogs" and "liking cats" appear to be independent? Give statistical evidence to support your conclusion.

Answer: a. Yes: 90.7%   No: 9.3%
b. Yes: 90.2%   No: 9.8%
c. segmented bar graph or pie charts
d. Perhaps. There is little difference between the percents of those who like dogs, depending on whether they like cats. Of those who like cats, only 90.2% like dogs, compared to 90.7% overall.

Explanation:

21) All students in a physical education class completed a basketball free-throw shooting event and the highest number of shots made was 32. The next day a student who had just transferred into the school completed the event, making 35 shots. Indicate whether adding the new student's score to the rest of the data made each of these summary statistics increase, decrease, or stay about the same.

    a. mean
    b. median
    c. range
    d. IQR
    e. standard deviation

21) _____

Answer: a. mean: increase
        b. median: stay about the same
        c. range: increase
        d. IQR: stay about the same
        e. standard deviation: increase

Explanation:

22) The body temperature of students is taken each time a student goes to the nurse's office. The five-number summary for the temperatures (in degrees Fahrenheit) of students on a particular day is:

22) _____

Min	Q1	Median	Q3	Max
96.6°	97.85°	98.25°	98.6°	101.8°

a. Would you expect the mean temperature of all students who visited the nurse's office to be higher or lower than the median? Explain.
b. After the data were picked up in the afternoon, three more students visited the nurse's office with temperatures of 96.7°, 98.4°, and 99.2°. Were any of these students outliers? Explain.

Answer: a. The mean temperature of all students would probably be higher than the median. Using the five-number summary, it appears the data are skewed to the right.
        b. IQR = 98.6° - 97.85° = 0.75°. Since 1.5(IQR) = 1.125°, the fences are 97.85° - 1.125° = 96.725° and 98.6° + 1.125° = 99.725°. The lowest temperature (96.7°) being added to the data set is smaller than the lower fence (96.725°) so it is an outlier on the low end. The highest temperature (99.2°) being added to the data set is not above the upper fence (99.725°) so it is not an outlier on the high end.

Explanation:

In November 2003 *Discover* published an article on the colonies of ants. They reported some basic information about many species of ants and the results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida. Information included the scientific name of the ant species, the geographic location, the depth of the nest (in feet), the number of chambers in the nest, and the number of ants in the colony. The article documented how new ant colonies begin, the ant-nest design, and how nests differ in shape, number, size of chambers, and how they are connected, depending on the species. It reported that nest designs include vertical, horizontal, or inclined tunnels for movement and transport of food and ants.

23) List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units.

23) _____

Answer: Categorical: species, geographic location, how new ant colonies begin, and nest design.
        Quantitative: nest depth (feet), number of chambers (units), and colony size (units).

Explanation:

24) A company's manufacturing process uses 500 gallons of water at a time. A "scrubbing"   24) _____
machine then removes most of a chemical pollutant before pumping the water into a
nearby lake. Legally the treated water should contain no more than 80 parts per million of
the chemical, but the machine isn't perfect and it is costly to operate. Since there's a fine if
the discharged water exceeds the legal maximum, the company sets the machine to attain
an average of 75 ppm for the batches of water treated. They believe the machine's output
can be described by a Normal model with standard deviation 4.2 ppm. (SHOW WORK.)

a. What percent of the batches of water discharged exceed the 80ppm standard?
b. The company's lawyers insist that they have not more than 2% of the water over the
limit. To what mean value should the company set the scrubbing machine? Assume the
standard deviation does not change.
c. Because achieving a mean that low would raise the costs too much, they decide to leave
the mean set at 75 ppm and try to reduce the standard deviation to achieve the "only 2%
over" goal. Find the new standard deviation needed.
d. Explain what achieving a smaller standard deviation means in this context.

Answer: a. $z = \dfrac{80 - 75}{4.2} = 1.19$

$P(z \geq 1.19) \approx 0.117$
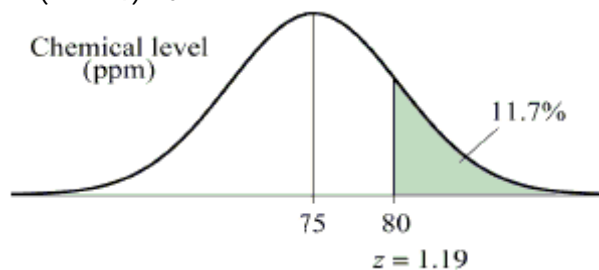
According to the normal model, we expect about 11.7% of the batches to exceed the
80ppm standard.

b. $z = 2.054$

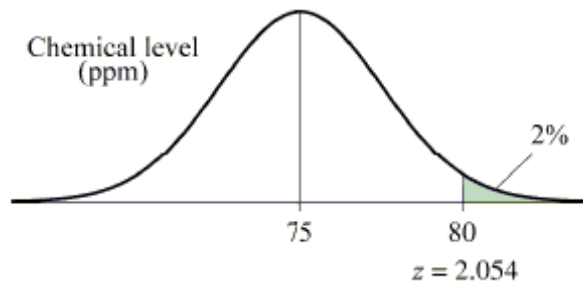$\dfrac{80 - \mu}{4.2} = 2.054$

$\mu = 80 - 2.054(4.2) = 71.37$
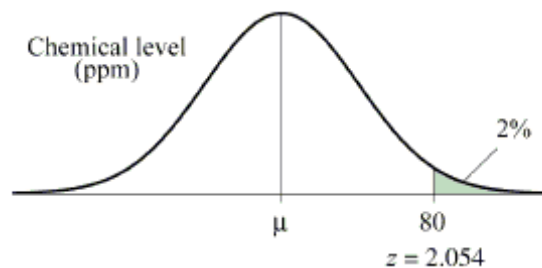
According to the Normal model, a mean of about 71.37ppm would need to be
achieved.

c. $\dfrac{80 - 75}{\sigma} = 2.054$

$\sigma \approx 2.434$

Answer:

According to the Normal model, the new standard deviation would need to be at most 2.43ppm.

d. The scrubber must be more consistent in its performance from batch to batch.

Explanation:

25) Human body temperatures taken through the ear are typically 0.5°F higher than body temperatures taken orally. Making this adjustment and using the 1992 Journal of the American Medical Association article that reports average oral body temperature as 98.2°F, we will assume that a Normal model with an average of 98.7°F and a standard deviation of 0.7°F is appropriate for body temperatures taken through the ear.

25) _____

a. An ear temperature of 97°F may indicate hypothermia (low body temperature). What percent of people have ear temperatures that may indicate hypothermia?

b. Find the interquartile range for ear temperatures.

c. A new thermometer for the ear reports that it is more accurate than the ear thermometers currently on the market. If the average ear temperature reading remains the same and the company reports an IQR of 0.5°F, find the standard deviation for this new ear thermometer.

Answer: a. $z = \dfrac{97 - 98.7}{0.7} = -2.43$, so $P(z < -2.43) = 0.0075$

About 0.75% of people have ear temperatures that may indicate hypothermia.

b. The z-scores associated with the IQR are $z = -0.67$ and $z = 0.67$. So, we need to solve for y in each of the following equations: $-0.67 = \dfrac{y - 98.7}{0.7}$ and $0.67 = \dfrac{y - 98.7}{0.7}$.

We get $y = 98.7 - 0.67(0.7) = 98.2$ and $y = 98.7 + 0.67(0.7) = 99.2$. The interquartile range is IQR = 99.2°F - 98.2°F = 1.0°F.

c. The new IQR is 0.5°F, while the old IQR was 1.0°F. So, we want

IQR $= [98.7 + 0.67\sigma] - [98.7 - 0.67\sigma] = 0.5$, or $1.34\sigma = 0.5$. Thus, $\sigma = \dfrac{0.5}{1.34} = 0.37$. Our

new standard deviation is 0.37°F.

Explanation:

26) The side-by-side boxplots show the cumulative college GPAs for sophomores, juniors, and
seniors taking an intro stats course in Autumn 2003.
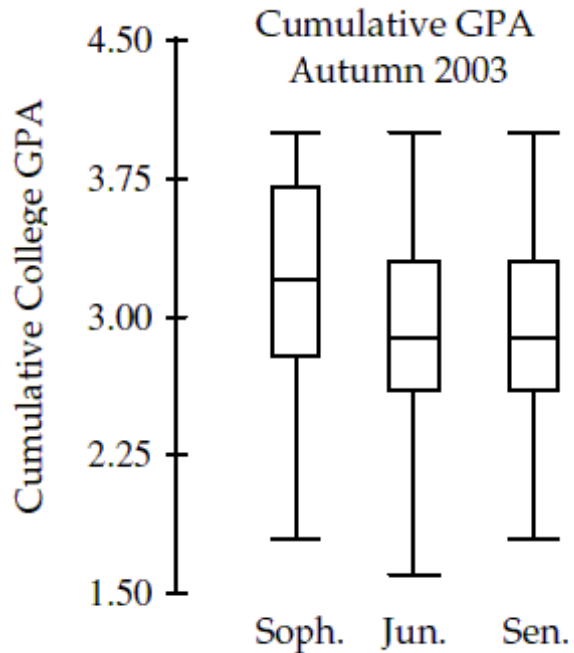
a. Which class (sophomore, junior, or senior) had the lowest cumulative college GPA?
What is the approximate value of that GPA?
b. Which class had the highest median GPA, and what is that GPA?
c. Which class had the largest range for GPA, and what is it?
d. Which class had the most symmetric set of GPAs? The most skewed set of GPAs?

Answer: a. The junior class had the lowest cumulative GPA, about 1.6.
          b. The sophomore class had the highest median cumulative GPA, about 3.2.
          c. The junior class had the largest range for GPA, about 2.4.
          d. The junior class had the most symmetric set of GPAs. The sophomore class had
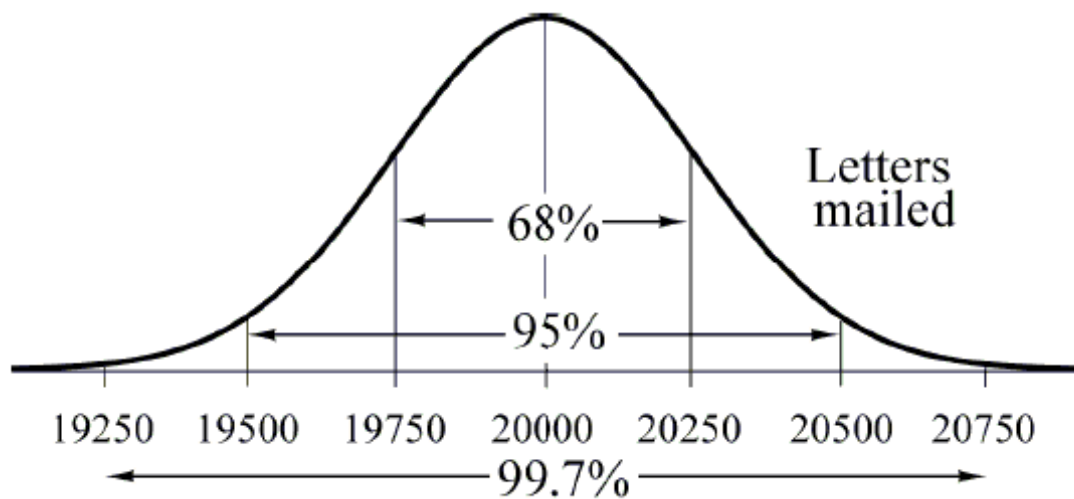   the most skewed set of GPAs, skewed to the left.
Explanation:

27) The Postmaster of a city's Post Office believes that a Normal model is useful in projecting the number of letters which will be mailed during the day. They use a mean of 20,000 letters and a standard deviation of 250 letters. Draw and clearly label this model.

27) _____
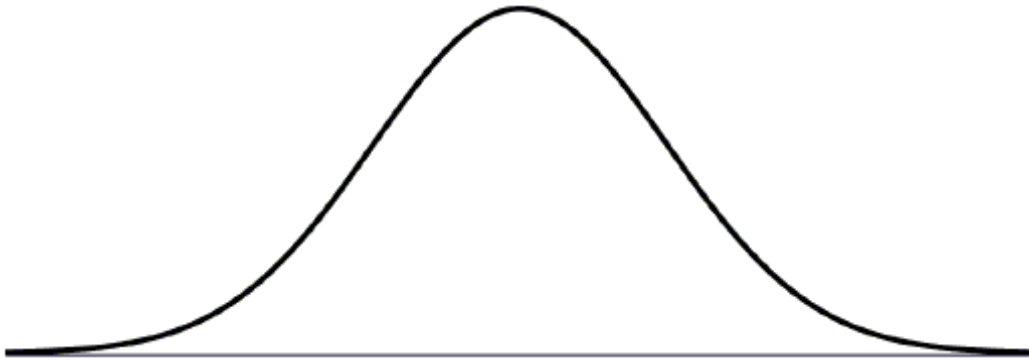
Answer:

Explanation:

28) Adult female Dalmatians weigh an average of 50 pounds with a standard deviation of 3.3 pounds. Adult female Boxers weigh an average of 57.5 pounds with a standard deviation of 1.7 pounds. One statistics teacher owns an underweight Dalmatian and an underweight Boxer. The Dalmatian weighs 45 pounds, and the Boxer weighs 52 pounds. Which dog is more underweight? Explain.

28) _____

Answer: Dalmation: $z_D = \dfrac{45 - 50}{3.3} = -1.52$

Boxer: $z_B = \dfrac{52 - 57.5}{1.7} = -3.24$

The Dalmatian is 1.52 standard deviations underweight, while the Boxer is 3.24 standard deviations underweight. So, the Boxer is more underweight.

Explanation:

29) On Monday, a class of students took a big test, and the highest score was 92. The next day, a   29) _____
student who had been absent made up the test, scoring 100. Indicate whether adding that
student's score to the rest of the data made each of these summary statistics increase,
decrease, or stay about the same:

    a. mean
    b. median
    c. range
    d. IQR
    e. standard deviation

Answer: a. mean: increase
          b. median: stay about the same
          c. range: increase
          d. IQR: stay about the same
          e. standard deviation: increase

Explanation:

30) Suppose that the student taking 22 credit hours in the data set in the previous question was   30) _____
actually taking 28 credit hours instead of 22 (so we would replace the 22 in the data set
with 28). Indicate whether changing the number of credit hours for that student would
make each of the following summary statistics increase, decrease, or stay about the same:

    a. mean
    b. median
    c. range
    d. IQR
    e. standard deviation

Answer: a. mean: increase
          b. median: stay about the same
          c. range: increase
          d. IQR: stay about the same
          e. standard deviation: increase

Explanation:

31) The boxplots show prices of used cars (in thousands of dollars) advertised for sale at three 31) _____
different car dealers.
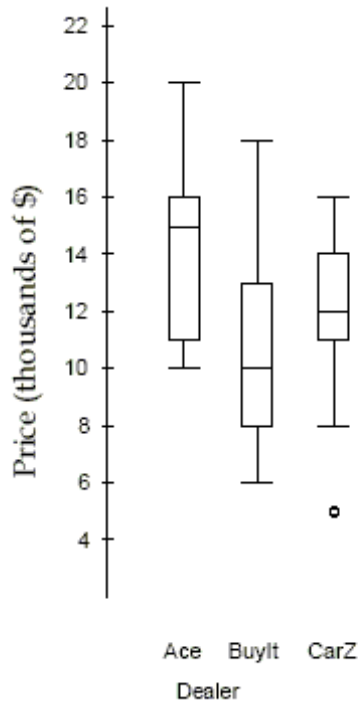
Dealer

a. Which dealer offers the cheapest car offered, and at what price?
b. Which dealer has the lowest median price, and how much is it?
c. Which dealer has the smallest price range, and what is it?
d. Which dealer's prices have the smallest IQR, and what is it?
e. Which dealer generally sells cars cheapest? Explain.

Answer: a. Car Z: $5000
b. BuyIt: $10,000
c. Ace: $10,000
d. CarZ: $3000
e. BuyIt; half of their cars are cheaper than any of the cars at Ace, and 25% of their cars are cheaper than all but one car at CarZ. The third quartile of their prices is well below the third quartile at CarZ, and below even the median price at Ace.

Explanation:

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	132
Dachshund	49	47	96
Golden Retriever	58	33	91
Labrador	37	41	78
Dalmatian	45	28	73
Other breeds	86	67	153
Total	348	275	623

32) Identify the variables and tell whether each is categorical or quantitative.   32) _____

   Answer: Gender and Breed; both categorical.
   Explanation:

33) Although most of us buy milk by the quart or gallon, farmers measure daily production in   33) _____
   pounds. Guernsey cows average 39 pounds of milk a day with a standard deviation of 8
   pounds. For Jerseys the mean daily production is 43 pounds with a standard deviation of 5
   pounds. When being shown at a state fair a champion Guernsey and a champion Jersey
   each gave 54 pounds of milk. Which cow's milk production was more remarkable?
   Explain.

   Answer: The Jersey's milk production was comparatively higher. That cow gave slightly
          more than 2 standard deviations above the average amount of milk ($z = 2.2$), while
          the Guernsey gave less than 2 standard deviations more than the average for
          Guernseys ($z = 1.875$).
   Explanation:

34) Repair bills An automobile service shop reported the summary statistics shown for repair   34) _____
   bills (in $) for their customers last month.

Min	27
Q1	88
Median	132
Q3	308
Max	1442
Mean	284
SD	140

   a. Were any of the bills outliers? Show how you made your decision.
   b. After checking out a problem with your car the service manager gives you an estimate of
   "*only* $90." Is he right to imply that your bill will be unusually low? Explain briefly.

   Answer: a. Yes. IQR = 308 - 88 = 220. The upper fence for outliers is one and a half IQR's
          above the third quartile, or 308 + 1.5(220) = 638. The maximum repair bill was $1442,
          well above $638, so it is certainly an outlier.
          b. No. $90 is higher than over 25% of the bills, so it is not unusual.
   Explanation:

35) A brake and muffler shop reported the repair bills, in dollars, for their customers   35) _____
yesterday.

88	283	312	290	172
154	400	381	346	181
203	118	143	252	227
56	192	292	213	422

a. Sketch a histogram for these data.
b. Find the mean and standard deviation of the repair costs.
c. Is it appropriate to use the mean and standard deviation to summarize these data?
Explain.
d. Describe the association of repair costs.

Answer: a.
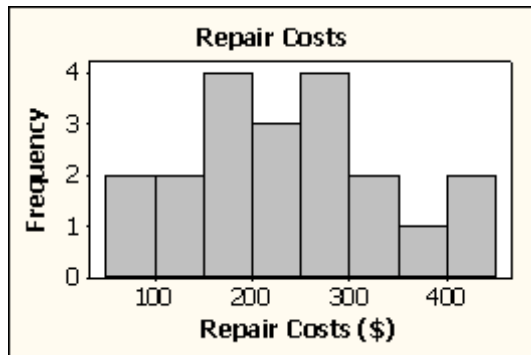
b. $\bar{x}$ = $236.25; s = $103.43
c. Yes, the data are roughly unimodal and symmetric with no outliers.
d. The repair costs averaged $236.25, ranging from $56 to $422 with a standard
deviation of $103.43. The distribution was approximately symmetric, with typical
repair costs clustered between $150 and $300.

Explanation:

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	132
Dachshund	49	47	96
Golden Retriever	58	33	91
Labrador	37	41	78
Dalmatian	45	28	73
Other breeds	86	67	153
Total	348	275	623

36) Write a sentence or two about the conditional relative frequency distribution of the breeds   36) _____
among female respondents.

   Answer: Among females, 20.9% chose Yorkshire Terriers, 14.2% Dachshunds, 16.7% Golden
            Retrievers, 10.6% Labs, and 12.9% Dalmatians. The remaining 24.7% of females
            preferred other breeds.
   Explanation:

37) Commuting to work The table shows how a company's employees commute to work.   37) _____

	Transportation			
Job Class	Car	Bus	Train	Total
Management	26	20	44	90
Labor	56	106	168	330
Total	82	126	212	420

   a. What is the marginal distribution (in %) of mode of transportation?
      Car _____ Bus _____ Train _____
   b. What is the conditional distribution (in %) of mode of transportation for management?
      Car _____ Bus _____ Train _____
   c. What kind of display would you use to show the association between job class and mode
   of transportation? (Just name the graph(s).)
   d. Do job classification and mode of transportation appear to be independent? Give
   statistical evidence to
      support your conclusion.

   Answer: a. Car: 19.5%   Bus: 30%   Train: 50.5%
            b. Car: 28.9%   Bus: 22.2%   Train: 48.9%
            c. segmented bar graph, or pie charts
            d. No, there is a difference between the percents in two types of transportation - Car
               and Bus categories, depending on the Job Classification.
                                Car   Bus   Train
               Management   28.9%   22.2%   48.9%
               Labor   17.0%   32.1%   50.9%
            Although about half of each group take the train, management are more likely than
            labor to come by car and less likely to take a bus.
   Explanation:

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	132
Dachshund	49	47	96
Golden Retriever	58	33	91
Labrador	37	41	78
Dalmatian	45	28	73
Other breeds	86	67	153
Total	348	275	623

38) Find each percent.   38) _____

    a. What percent of the responses were from males who favor Labradors?

    b. What percent of the male responses favor Labradors?

    c. What percent of the people who choose Labradors were males?

    Answer: a. 6.6%
             b. 14.9%
             c. 52.6%
    Explanation:

In June 2003 *Consumer Reports* published an article on some sport-utility vehicles they had tested recently. They reported some basic information about each of the vehicles and the results of some tests conducted by their staff. Among other things, the article told the brand of each vehicle, its price, and whether it had a standard or automatic transmission. They reported the vehicle's fuel economy, its acceleration (number of seconds to go from zero to 60 mph), and its braking distance to stop from 60 mph. The article also rated each vehicle's reliability as much better than average, better than average, average, worse, or much worse than average.

39) List the variables. Indicate whether each variable is categorical or quantitative. If the   39) _____
    variable is quantitative, tell the units.

    Answer: Categorical: brand, transmission type, reliability
                Quantitative: price (US$), fuel economy (mpg), acceleration (seconds), braking
                distance (probably feet?)
    Explanation:

40) Consider the following part of a data set: 40) _____

Age (years)	Sex	Only child?	Height (inches)	Weight (pounds)	Credit Hours	GPA	Major
21	Female	Yes	67.00	140.0	16	3.60	animal science
20	Female	No	62.00	130.0	18	3.86	biology
28	Female	No	64.00	188.0	21	3.25	psychology
21	Male	No	65.00	140.0	15	2.95	psychology
24	Female	No	67.00	130.0	20	3.00	anthropology
22	Male	Yes	68.00	135.0	15	2.94	journalism

List the variables in the data set. Indicate whether each variable is treated as categorical or quantitative in this data set. If the variable is quantitative, state the units.

Answer: Categorical: sex, only child?, major
        Quantitative: age (years), height (inches), weight (pounds), credit hours, GPA
Explanation:

41) At a large business, employees must report to work at 7:30 A.M. The arrival times of   41) _____
employees can be described by a Normal model with mean of 7:22 A.M. and a standard
deviation of four minutes.

a. What percent of employees are late on a typical work day?
b. A psychological study determined that the typical worker needs five minutes to adjust to
their surroundings before beginning their duties. What percent of this business' employees
arrive early enough to make this adjustment?
c. Because late employees are a distraction and cost companies money, all employees need
to be on time to work. If the mean arrival time of employees does not change, what
standard deviation would the arrival times need to ensure virtually all employees are on
time to work?
d. Explain what achieving a smaller standard deviation means in the context of this
problem.

Answer: a. Employees are late if they arrive after 7:30 A.M.

$$P(\text{time} > 7:30) = P\left(z > \frac{30 - 22}{4} = 2\right) = 0.0228$$
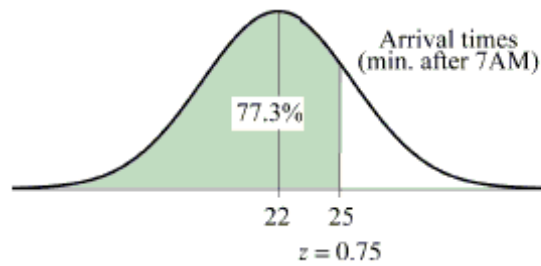
According to the Normal model, about 2.28% of employees are expected to arrive
after 7:30 AM.

b. $P(x \le 25) = P\left(z < \frac{25 - 22}{4} = 0.75\right) = 0.773$

According to the Normal model, about 77.3% of employees arrive at work before
7:25 AM.

19

Answer:

Arrival times
(min. after 7AM)

77.3%

22   25

$z = 0.75$
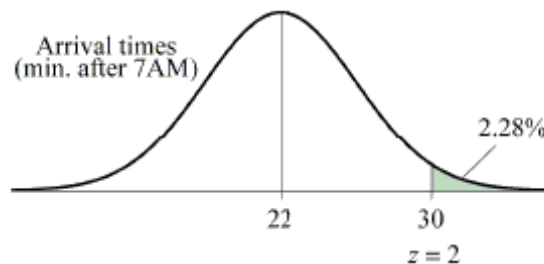
c. Virtually all times lie within 3 standard deviations of the mean. (Accept other reasonable z-scores greater than 3). If $z = 3 \leq \dfrac{30 - 22}{\sigma}$, then $3\sigma \leq 8$, so $\sigma \leq 2.67$ minutes.

d. A smaller standard deviation would mean greater consistency in arrival times.

Explanation:

42) House calls A local plumber makes house calls. She charges $30 to come out to the house and $40 per hour for her services. For example, a 4-hour service call costs $30 + $40(4) = $190.

42) _____

a. The table shows summary statistics for the past month. Fill in the table to find out the cost of the service calls.

Statistic	Hours of Service Call	Cost of Service Call
Mean	4.5	
Median	3.5	
SD	1.2	
IQR	2.0	
Minimum	0.5	

b. This past month, the time the plumber spent on one service call corresponded to a z-score of -1.50. What was the z-score for the cost of that service call?

Answer: a.

| Statistic	Hours of Service Call	Cost of Service Call
Mean	4.5	$210
Median	3.5	$170
SD	1.2	$48
IQR	2.0	$80
Minimum	0.5	$50

b. -1.50

Explanation:

43) A manufacturer claims that lifespans for their copy machines (in months) can be described   43) _____
by a Normal model N(42, 7). Show your work.
a. Draw and clearly label the model.
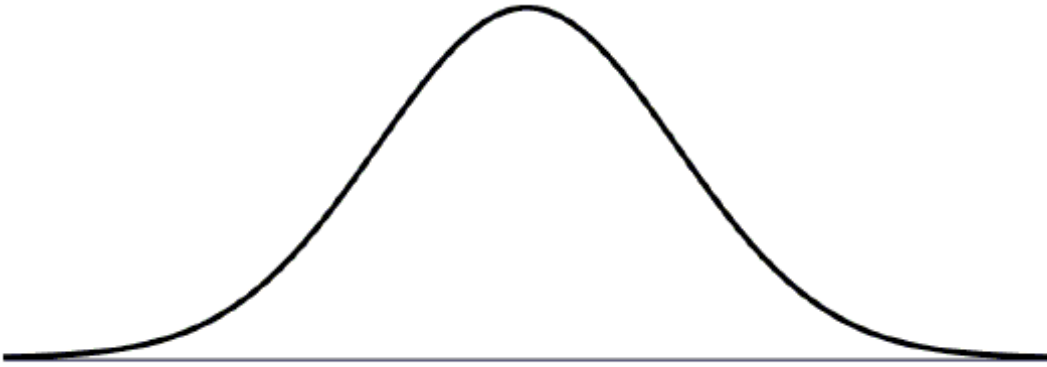
b. A company with several large office buildings buys 200 of these copiers. The salesman tells the boss "190 (95%) of your new copiers will last between _____ and _____ months." Comment on this claim.

c. What is the 3rd quartile of copier lifespans?
d. What percent of the copiers are expected to fail before 36 months?
e. The manufacturer wants to reduce the 36-month failure rate to only 10%. Assuming the mean lifespan will stay the same, what standard deviation must they achieve?
f. Briefly explain what that change in standard deviation means in this context.
g. A competing manufacturer says that not only will 90% of their copiers last at least 36 months, 65% will last at least 42 months. What Normal model parameters is that manufacturer claiming? Show your work. N(____ , ____)

Answer: a.
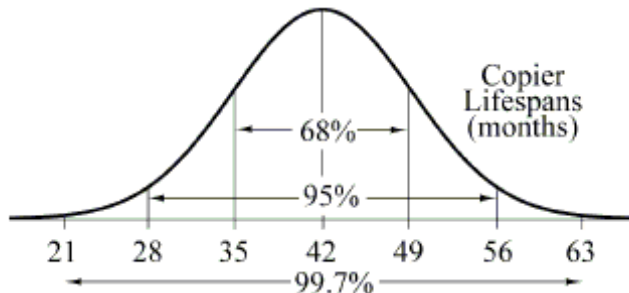
b. 28, 56. The claim is probably false. This model should provide a useful estimate of what might happen, but is not certain to predict what actually will happen.
c. 46.7 months
d. 19.6%
e. 4.7 months (should all include sketches of labeled curves)
f. A smaller standard deviation means that the copiers would be more consistent in their lifespans.
g. For 36 months $z = -1.28$ and for 42 months $z = -0.385$. Thus the difference of 6 months is $1.28 - 0.385 = 0.895$ standard deviations. The model is N(44.6, 6.7).

Explanation:

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from "Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes" (Shanthy A. Bowman, *Journal of the American Dietetic Association*, 102(9), pp. 1234–1239):

<table>
<tr><td></td><td></td><td colspan="4" align="center">Nationwide Food Survey Years</td></tr>
<tr><td></td><td></td><td>1987-1988</td><td>1989-1991</td><td>1994-1996</td><td>Total</td></tr>
<tr><td rowspan="3">Drinks Fluid Milk</td><td>Yes</td><td>354</td><td>502</td><td>366</td><td>1222</td></tr>
<tr><td>No</td><td>226</td><td>335</td><td>366</td><td>927</td></tr>
<tr><td>Total</td><td>580</td><td>837</td><td>732</td><td>2149</td></tr>
</table>

44) Do you think that milk consumption by young girls is independent of the nationwide survey year? Use statistics to justify your reasoning.

44) _____

Answer: No. 56.9% of all young girls surveyed reported drinking milk, but 60% of the young girls reported drinking milk in the 1989-1991 survey. Since these percentages differ, milk consumption and year are not independent.

Explanation:

45) Health Insurance *The World Almanac and Book of Facts 2004* reported the percent of people not covered by health insurance in the 50 states and Washington, D.C., for the year 2002. Computer output gives these summaries for the percent of people not covered by health insurance:

45) _____

Min	Q1	Median	Q3	Max	Mean	SD
7.9	10.8	13.4	16.7	25.8	13.9	3.6

a. Were any of the states outliers? Explain how you made your decision.
b. A histogram of the data is as follows:
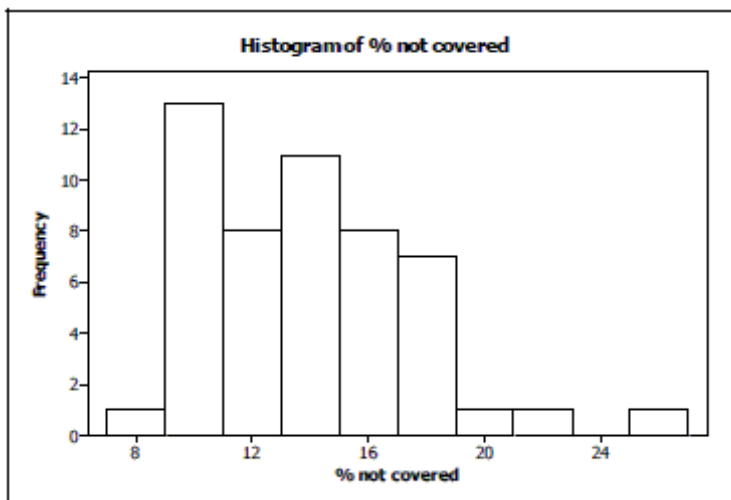
Is it more appropriate to use the mean and standard deviation or the median and IQR to describe these data? Explain.

Answer: a. IQR = Q3 - Q1 = 16.7 - 10.8 = 5.9
      1.5(IQR) = 1.5(5.9) = 8.85
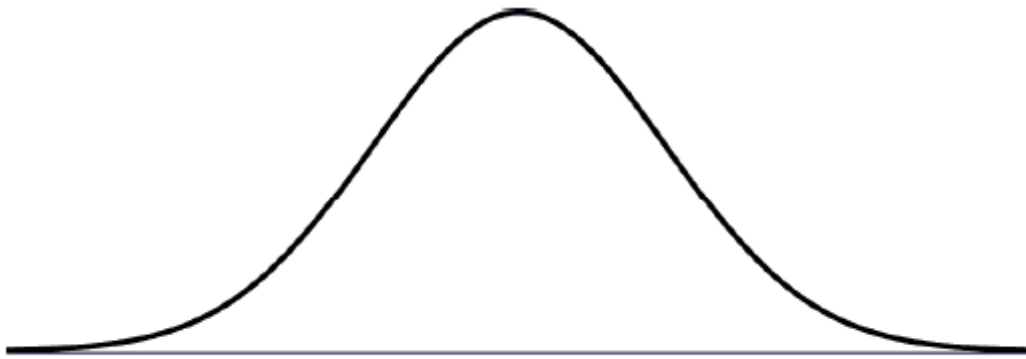   Q3 + 1.5(IQR) = 16.7 + 8.85 = 25.55 < Max, so there is at least one high outlier
      Q1 - 1.5(IQR) = 10.8 - 8.85 = 1.95 < Min, so there are no low outliers
    b. It is more appropriate to use the median and IQR to describe these data, since the distribution is skewed right.
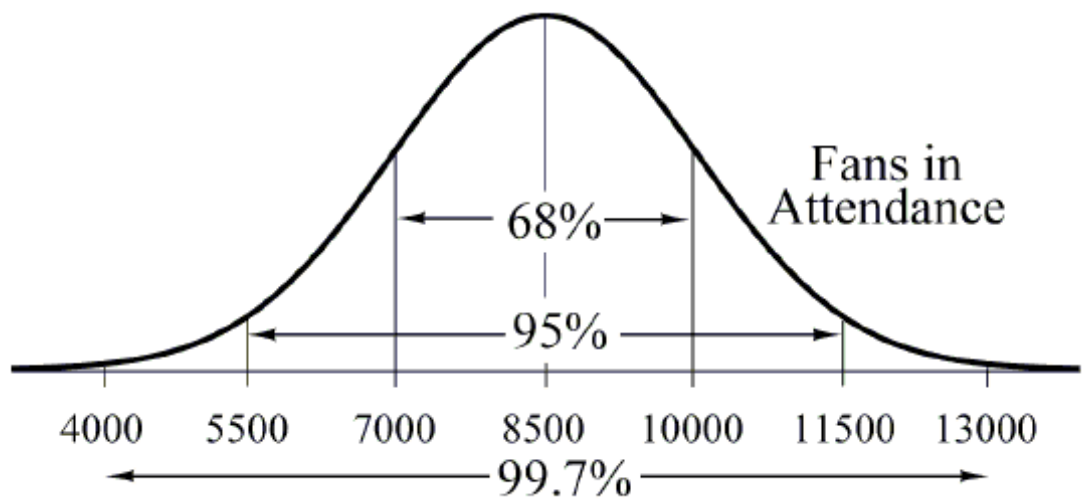
Explanation:

46) Owners of a minor league baseball team believe that a Normal model is useful in   46) _____
projecting the number of fans who will attend home games. They use a mean of 8500 fans
and a standard deviation of 1500 fans. Draw and clearly label this model.

Answer:

Fans in Attendance

68%

95%

4000   5500   7000   8500   10000   11500   13000

99.7%

Explanation:

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	132
Dachshund	49	47	96
Golden Retriever	58	33	91
Labrador	37	41	78
Dalmatian	45	28	73
Other breeds	86	67	153
Total	348	275	623

47) What is the marginal distribution of breeds?   47) _____

    Answer: There were 132 Yorkshire terrier responses, 96 Dachshund responses, 91 Golden
           Retriever responses, 78 Labrador responses, 73 Dalmation responses, and 153 Other
           responses.
    Explanation:

48) Which of the W's are unknown for these data?   48) _____

    Answer: We do not know how or when the people were surveyed, or where the local dog
           show was located.
    Explanation:

49) Testbank Starts at Ch. 2   49) _____

    Answer: Testbank Starts at Ch. 2
    Explanation:

In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses from boys and girls.

	Male	Female	Total
Bus	30	34	64
Ride	37	45	82
Personal	19	23	42
Total	86	102	188

50) What is the marginal distribution of gender?   50) _____

    Answer: There are 86 males and 102 females.
    Explanation:

51) Which of the W's are unknown for these data?   51) _____

   Answer: We don't know how or when the students were surveyed, nor where the school is.
   Explanation:

52) Book sales A publishing company pays its sales staff $600 a week plus a commission of   52) _____
    $0.50 per book sold. For example, a salesman who sold 440 books earned 600 + 0.50(440) =
    $820.

    a. The table shows summary statistics for the number of books the large sales staff sold last
    week.   Fill in the table to show the statistics for the pay these people earned.

Statistic	Books Sold	$ Earned
Mean	640	
Standard deviation	360	
IQR	450	
Maximum	1420	

    b. The newest employee had a pretty good week. Among all the salespeople her pay
    corresponded to a z-score of +1.80. What was the z-score of the number of books she sold?

    Answer: a.

Statistic	Books Sold	$ Earned
Mean	640	$920
Standard deviation	360	$180
IQR	450	$225
Maximum	1420	$1310

          b. +1.80
    Explanation:

53) While the scales histograms are the same, there is something that could be improved so   53) _____
    that we could compare these two distributions better. Identify this improvement and
    explain why it would be better.

    Answer: The frequency scales are not the same for the two histograms. If we converted each
             of the frequency histograms to a relative frequency histogram, we would be better
             able to compare the frequencies for each distribution at the heights.
    Explanation:

54) Salary conversions You learn that your company is sending you and several other employees to staff a new office in China. While there everyone will earn the equivalent of their current salary, converted to Chinese currency at the rate of 8 yuans per dollar. In addition, everyone will earn a weekly foreign living allowance of 200 yuans. For example, since you are earning $1000 per week, your weekly salary in China will be $1000 \times 8 + 200 = 8200$ yuans.

a. Shown are some summary statistics describing the current salaries of this group being sent overseas. Fill in the table to show what these statistics will be for the salaries you all will earn while in China.

Statistic	In the US	In China
Minimum salary	$400	
Standard deviation	$250	
Median	$750	
IQR	$300	

b. Among this group of employees going to China, your US salary has a z-score of +1.20. What will your new z-score be, based on everyone's China salary?

Answer: a. 3400 yuans, 2000, 6200, 2400
          b. $z = +1.20$

Explanation:

In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses from boys and girls.

	Male	Female	Total
Bus	30	34	64
Ride	37	45	82
Personal	19	23	42
Total	86	102	188

55) Find each percent.

a. What percent of the students are girls who ride the bus?

b. What percent of the girls ride the bus?

c. What percent of the bus riders are girls?

Answer: a. 18.1%
          b. 33.3%
          c. 53.1%

Explanation:

26

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	**132**
Dachshund	49	47	**96**
Golden Retriever	58	33	**91**
Labrador	37	41	**78**
Dalmatian	45	28	**73**
Other breeds	86	67	**153**
Total	**348**	**275**	**623**

56) Do you think the breed selection is independent of gender? Give statistical evidence to support your conclusion.

56) _____

Answer: The breed selection does not appear to be independent of gender. Overall, 56% of the respondents were females, but females were over-represented among those who favored Golden Retrievers (64%) and Dalmatians (62%), yet a much lower percentage (47%) among those who chose Labradors.

Explanation:

57) Exercising Owners of an exercise gym believe that a Normal model is useful in projecting the number of clients who will exercise in their gym each week. They use a mean of 800 clients and a standard deviation of 90 clients.

57) _____

a. Draw and clearly label this model.
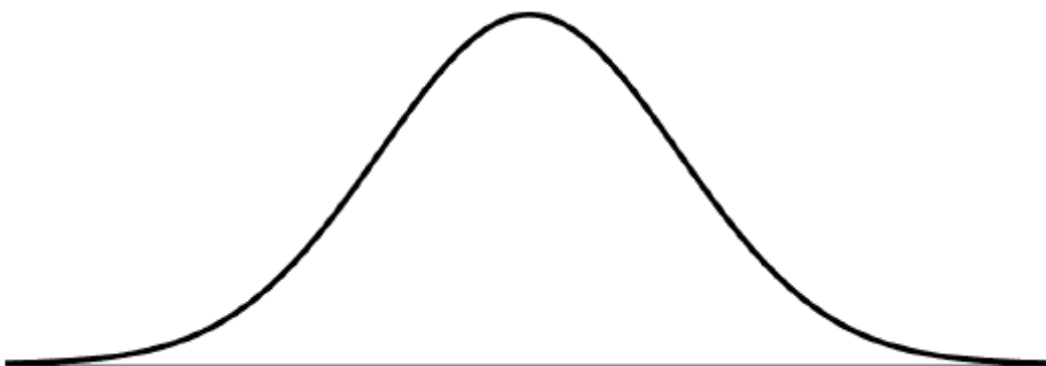
b. What is the first quartile of the weekly number of clients? [Show work]

c. An owner of another gym reports that 5% of the time their gym has fewer than 450 clients, and 40% of the time the gym has more than 1085 clients. What parameters should that owner use for his Normal model?   N(___ , ___)

Answer: a.
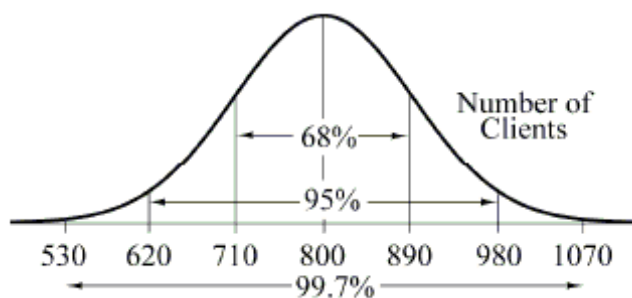
b. $Q_1 \Rightarrow P = 0.25$ and $z = -0.674$,

$$-0.674 = \frac{x - 800}{90} \qquad -60.66 = x - 800$$

x = 739.34
So the first quartile is at 740 clients.

c. 450 (5th percentile) has $z = -1.645$

1085 (60th percentile) has $z = +0.253$
1085 - 450 = (0.253 + 1.645)σ
$\sigma \approx 334.5$
$\mu + 0.253(334.5) = 1085$
$\mu \approx 1000.4$
N(1000.4, 334.5)

Explanation:

58) One thousand students from a local university were sampled to gather information such as 58) _____
gender, high school GPA, college GPA, and total SAT scores. The results were used to
create histograms displaying high school grade point averages (GPA's) for both males and
females. Compare the grade distribution of males and females.
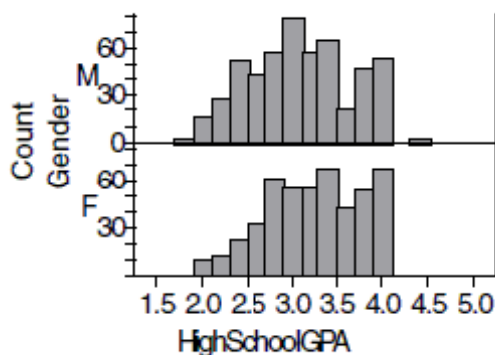
Answer: The distributions of high school GPA for both males and females are skewed to the
left, and both distributions appear to be centered at a GPA of about 3.0. The
distribution of male GPA appears slightly more spread out than the distribution of
female GPA.

Explanation:

59) The boxplots show the age of people involved in accidents according to their role in the accident.
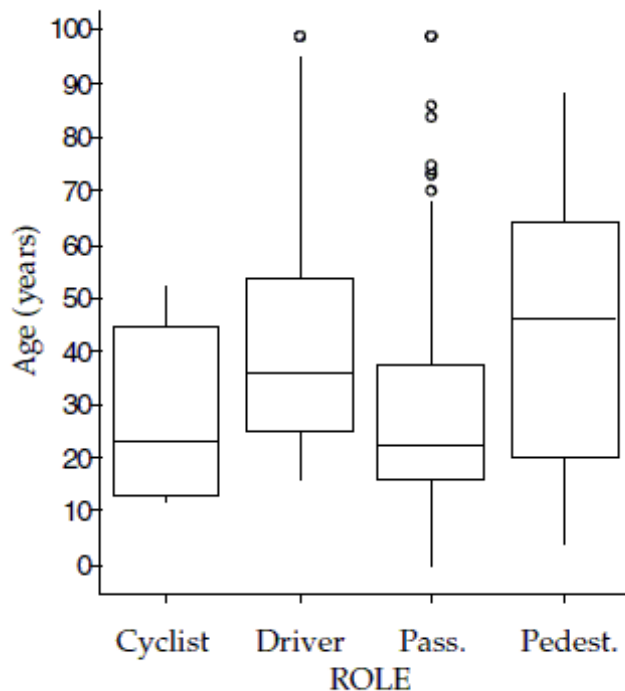
a. Which role involved the youngest person, and what is the age?

b. Which role had the lowest median age, and what is the age?

c. Which role had smallest range of ages, and what is it?

d. Which role had the largest IQR of ages, and what is it?

e. Which role generally involved the oldest people? Explain.

Answer: a. Passenger, less than 1 year
   b. Passenger, 21 years
   c. Cyclist, 40 years
   d. Pedestrian, 44 years
   e. Pedestrian. While the oldest person involved in an accident is not a pedestrian, the median age for pedestrians is almost 45 years, while the median ages in the other groups are between 22 and 35 years old. The oldest 50% of the Pedestrian group, from 45 to 87 years, is generally older than the youngest 75% of two groups - Cyclist and Passenger, and only the Driver group has any of its middle 50% as old. The Driver and Passenger groups have a few people older than the Pedestrian group.

Explanation:

60) City planners wanted to know how many people live in a typical housing unit so they   60) _____
        compiled data
    from hundreds of forms that had been submitted in various city offices. Summary statistics
        are shown in
    the table.

$\bar{x}$	2.53 people
s	1.40 people
min	1
Q1	1
median	2
Q3	3
max	10

a. Notice that the minimum occupancy and the first quartile are the same. Explain how this
can be.
b. The city classifies residences housing 4 or more people as "high occupancy". Would you
consider 4 occupants to be unusually high? Explain.
c. The city bases their garbage disposal fee on the occupancy level of the home or
apartment. The annual fee is $40 plus $5 per person, so a single occupant pays $45 and the
homes with 10 people pay $40 +$5(10) = $90 a year. What is the median fee paid? And the
IQR?
d. What are the mean and standard deviation of the garbage disposal fees?

Answer: a. At least 25% of the housing units have only one resident.

        b: 4 residents is just above the 3rd quartile, and only about one standard deviation
        above the mean. It is not an unusually high number of residents.
        c. Median = 40 + 5(2) = $50; IQR = 5(3 - 1) = $10
        d. Mean = 40 + 5(2.53) = $52.65; SD = 5(1.50) = $7.50
Explanation:

61) Concrete thickness A roadway construction process uses a machine that pours concrete onto the roadway and measures the thickness of the concrete so the roadway will measure up to the required depth in inches. The concrete thickness needs to be consistent across the road, but the machine isn't perfect and it is costly to operate. Since there's a safety hazard if the roadway is thinner than the minimum 23 inch thickness, the company sets the machine to average 26 inches for the batches of concrete. They believe the thickness level of the machine's concrete output can be described by a Normal model with standard deviation 1.75 inches. [Show work]

a. What percent of the concrete roadway is under the minimum depth?

b. The company's lawyers insist that no more than 3% of the output be under the limit. Because of the expense of operating the machine, they cannot afford to reset the mean to a higher value. Instead they will try to reduce the standard deviation to achieve the "only 3% under" goal. What standard deviation must they attain?

c. Explain what achieving a smaller standard deviation means in this context.

Answer: a. The concrete roadway is under minimum depth when less then 23 inches in thickness.

$$z = \frac{23 - 26}{1.75} = -1.7 \rightarrow P = 0.043,$$ so the model suggests about 4.3% is under the

minimum depth

b. $P = 0.03 \rightarrow z = -1.88$, so $-1.88 = \dfrac{23 - 26}{\sigma}$; then $\sigma = 1.6$ inches

c. A smaller standard deviation means that the thickness of the concrete will be more consistent.

Explanation:

62) The students in a biology class kept a record of the height (in centimeters) of plants for a   62) _____
class experiment.

49	67	38	55	62
54	36	41	56	43
48	75	44	60	48
52	48	53	59	32

a. Sketch a histogram for these data.
b. Find the mean and standard deviation of the plant heights.
c. Is it appropriate to use the mean and standard deviation to summarize these data? Explain.
d. Describe the association of plant heights.

Answer: a.
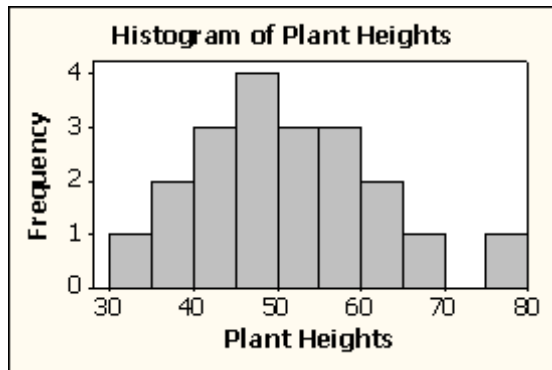
b. $\bar{x}$ = 51.0 cm; s = 10.6 cm
c. Yes, the data are roughly unimodal and symmetric with no outliers.
d. The data are roughly symmetric with no outliers; however there is a small gap from 70 to 75 cm. The average plant height is 51.0 centimeters, with a standard deviation of 10.6 centimeters. The range of plant heights is 43 centimeters. The distribution of plant heights has a mode between 45 and 49 centimeters.
Explanation:

In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses from boys and girls.

	Male	Female	Total
Bus	30	34	64
Ride	37	45	82
Personal	19	23	42
Total	86	102	188

63) Write a sentence or two about the conditional relative frequency distribution of modes of   63) _____
transportation for the boys.

Answer: More boys (43%) caught rides to school than any other means of transportation. 35% rode the bus while only 22% used personal transportation like biking, skateboarding, or walking.
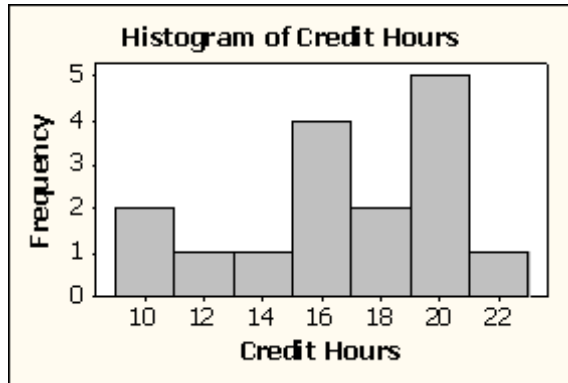Explanation:

64) A survey conducted in a college intro stats class asked students about the number of credit   64) _____
hours they were taking that quarter. The number of credit hours for a random sample of
16 students is given in the table.

10	10	12	14	15	15	15	15
17	17	19	20	20	20	20	22

a. Sketch a histogram of these data.
b. Find the mean and standard deviation for the number of credit hours.
c. Find the median and IQR for the number of credit hours.
d. Is it more appropriate to use the mean and standard deviation or the median and IQR to
    summarize these data? Explain.

Answer: a.

b. $\bar{x}$ = 16.3 credit hours; s = 3.7 credit hours
c. The median is 16.0 credit hours.
   IQR = Q3 - Q1 = 20 - 14.5 = 5.5 credit hours
d. It is more appropriate to use the median and IQR to summarize these data,
because these data are not unimodal and symmetric.

Explanation:

65) The five-number summary for midterm scores (number of points; the maximum possible   65) _____
score was 50 points) from an intro stats class is:

Min	Q1	Median	Q3	Max
16.5	32	39	43.5	48.5

a. Would you expect the mean midterm score of all students who took the midterm to be
higher or lower than the median? Explain.
b. Based on the five-number summary, are any of the midterm scores outliers? Explain.

Answer: a. The mean midterm score of all students would probably be lower than the
median. Using the five-number summary, it appears that the data are skewed to the
left.
b. IQR = 43.5 - 32 =11.5
Q1 - 1.5IQR = 32 - 1.5(11.5) = 14.75
Q3 + 1.5IQR = 43.5 + 1.5(11.5) = 60.75
Since both the maximum and minimum scores fall between these "fences", there are
no outliers in this data set.

Explanation:

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from "Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes" (Shanthy A. Bowman, *Journal of the American Dietetic Association*, *102*(9), pp. 1234–1239):

Nationwide Food Survey Years

Drinks Fluid Milk		1987-1988	1989-1991	1994-1996	Total
	Yes	354	502	366	1222
	No	226	335	366	927
	Total	580	837	732	2149

66) What is the marginal distribution of milk consumption?   66) _____

Answer: Yes: 56.9%; No: 43.1%
Explanation:

67) Soda pop A machine that fills cans with soda fills according to a Normal model with mean   67) _____
12.1 ounces and standard deviation 0.05 ounces.

a. If the cans claim to have 12 ounces of soda each, what percent of cans are under-filled?
b. Management wants to ensure that only 1% of cans are under-filled.
  i. Scenario 1: If the mean fill of the cans remains at 12.1 ounces, what standard deviation does the filling
    machine need to have to achieve this goal?
  ii. Scenario 2: If the standard deviation is to remain at 0.05 ounces, what mean does the filling machine
    need to have to achieve this goal?

Answer: a. $z = \dfrac{12 - 12.1}{0.05} = -2.0$, which suggests that 2.28% of cans are under-filled.

b. A *z*-score of -2.33 has 1% to its left, meaning that 1% of the cans would be under-filled.

  i. $-2.33 = \dfrac{12.- 12.1}{\sigma} \Rightarrow -2.33\sigma = -0.1 \Rightarrow \sigma = \dfrac{-0.1}{-2.33} = 0.043$. The standard deviation would need
    to be 0.043 ounces.

  ii. $-2.33 = \dfrac{12.- \mu}{0.05} \Rightarrow -2.33(0.05) = 12 - \mu \Rightarrow \mu = 12 + 2.33(0.05) = 12.12$. The mean would need to be 12.12 ounces.

Explanation:

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from "Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes" (Shanthy A. Bowman, *Journal of the American Dietetic Association*, *102*(9), pp. 1234-1239):

<table>
<tr><td colspan="5" align="center">Nationwide Food Survey Years</td></tr>
<tr><td></td><td>1987-1988</td><td>1989-1991</td><td>1994-1996</td><td>Total</td></tr>
<tr><td>Yes</td><td>354</td><td>502</td><td>366</td><td>1222</td></tr>
<tr><td>No</td><td>226</td><td>335</td><td>366</td><td>927</td></tr>
<tr><td>Total</td><td>580</td><td>837</td><td>732</td><td>2149</td></tr>
</table>

Drinks Fluid Milk (row labels: Yes, No, Total)

68) Consider the following pie charts of a subset of the data above:　　　　　　68) _____

Pie Chart of 1989-1991, 1994-1996 vs Milk?
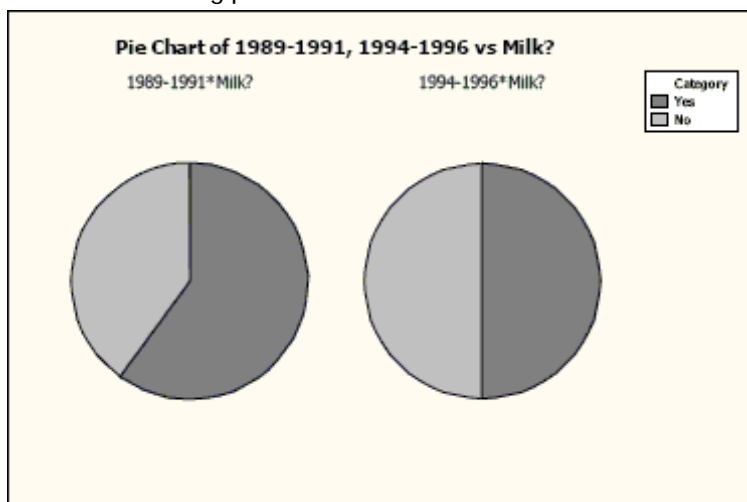
Do the pie charts above indicate that milk consumption by young girls is independent of the nationwide survey year? Explain.

Answer: No. It looks like there is some sort of relationship between milk consumption and nationwide survey year, since the percentage of young girls who reported drinking milk is a larger slice of the pie chart for the 1989-1991 survey than the same response for the 1994-1996 survey.

Explanation:

69) Students taking an intro stats class reported the number of credit hours that they were taking that quarter.

Summary statistics are shown in the table.

69) _____

$\bar{x}$	16.65
s	2.96
min	5
Q1	15
median	16
Q3	19
max	28

a. Suppose that the college charges $73 per credit hour plus a flat student fee of $35 per quarter. For example, a student taking 12 credit hours would pay $35 + $73(12) = $911 for that quarter.

 i. What is the mean fee paid?

 ii. What is the standard deviation for the fees paid?

 iii. What is the median fee paid?

 iv. What is the IQR for the fees paid?

b. Twenty-eight credit hours seems like a lot. Would you consider 28 credit hours to be unusually high? Explain.

Answer: a. i. $35 + $73(16.65) = $1250.45

ii. $73(2.96) = $216.08

iii. $35 + $73(16) = $1203

iv. IQR = $73(19-15) = $292

b. IQR = 19 - 15 = 4 credit hours

High outliers will lie above Q3 + 1.5IQR = 19 + 1.5(4) = 25 credit hours. Since 28 credit hours exceeds 25 credit hours, I would consider 28 credit hours to be unusually high.

Explanation:

70) During a budget meeting, local school board members decided to review class size information to
determine if budgets were correct. Summary statistics are shown in the table.

$\bar{x}$	33.39 students
s	5.66 students
min	17
Q1	29
median	33
Q3	40
max	40

a. Notice that the third quartile and maximum class sizes are the same. Explain how this can be.
b. The school district declares that classes with fewer than 20 students are "too small". Would you consider a class of 20 students to be unusually small? Explain.
c. The school district sets the office supply budgets of their high schools on the enrollment of students. The district budgets each class $12 plus $0.75 per student, so a class with one student receives $12.75 and the classes with 40 students receive 12 + 0.75(40) = $42. What is the median class budget for office supplies? And the IQR?
d. What are the mean and standard deviation of the class office supply budgets?

Answer: a. The top 25 percent of all classes have 40 students enrolled.
  b. Yes, classes with 20 students enrolled seem unusually small. Twenty is well below the first quartile of 29 students, and only slightly above the minimum size (17).
  $z = \dfrac{20 - 33.39}{5.66} = -2.366$ With $z = -2.366$, this size class is over 2 standard deviation
  units below the mean.
  c. Median budget = $12 + $0.75(33) = $36.75
    Q1 budget = $12 + $0.75(29) = $33.75
    Q3 budget = $12 + $0.75(40) = $42.00
    IQR = $42.00 - $33.75 = $8.25
  d. Mean budget = $12 + $0.75(33.39) = $37.04
    Standard deviation = $0.75(5.66) = $4.25
Explanation:

In November 2003 *Discover* published an article on the colonies of ants. They reported some basic information about many species of ants and the results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida. Information included the scientific name of the ant species, the geographic location, the depth of the nest (in feet), the number of chambers in the nest, and the number of ants in the colony. The article documented how new ant colonies begin, the ant-nest design, and how nests differ in shape, number, size of chambers, and how they are connected, depending on the species. It reported that nest designs include vertical, horizontal, or inclined tunnels for movement and transport of food and ants.

71) Describe the W's, if the information is given: 71) _____

   • Who:

   • What:

   • When:

   • Where:

   • How:

   • Why:

Answer: Who: Colonies of ants. "Many species of ants," but no indication of exactly how
         many.
         What: scientific name, geographic location, average nest depth, average number of
         chambers, average colony size, how new ant colonies begin, the ant-nest design,
         and how nests differ in architecture.
         When: November 2003
         Where: not specified
         How: The results of some discoveries found by myrmecologist Walter Tschinkel of
         the University of Florida
         Why: Information of interest to readers of the magazine
Explanation:

72) The Wechsler Adult Intelligence Scale – Revised (WAIS-R) follow a Normal model with   72) _____
mean 100 and
standard deviation 15. Draw and clearly label this model.
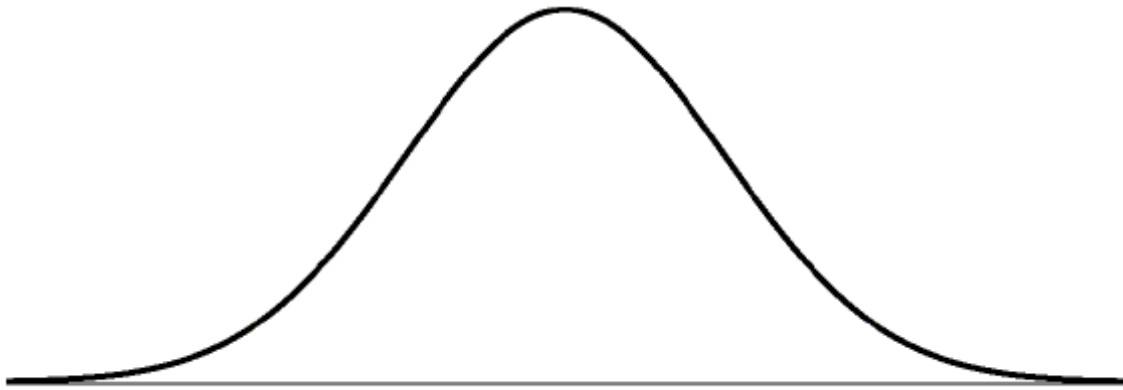
Answer:
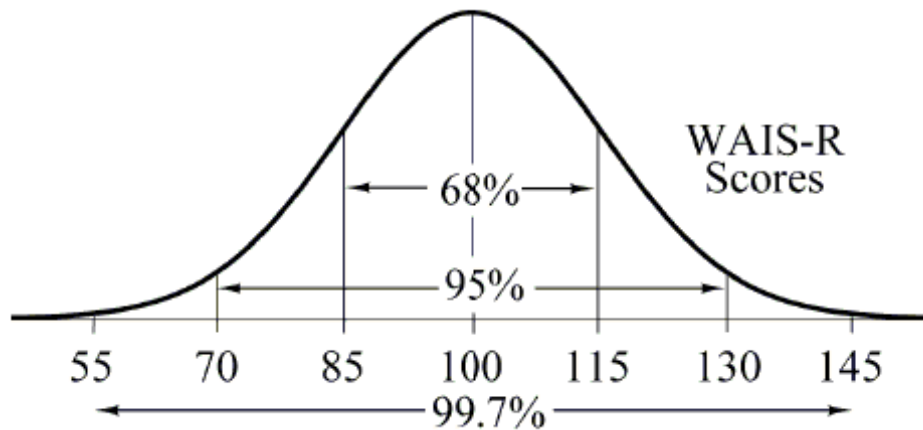
Explanation:

73) Cordless phones In their October 2003 issue, *Consumer Reports* evaluated the price and   73) _____
performance of 23 models of cordless phones. Computer output gives these summaries for
the prices:

Min	Q1	Median	Q3	Max	MidRange	Mean	TrMean	SD
15	30	50	110	200	107.5	71.75	67.63	52.08

a. Were any of the prices outliers? Explain how you made your decision.
b. One of the manufacturers advertises a cordless phone "economy-priced at only $31.95".
Would you consider that to be a very low price? Explain.

Answer: a. IQR = Q_3 - Q_1 = 110 - 30 = 80

1.5(IQR) = 1.5(80) = 120

Q_3 + 1.5(IQR) = 110 + 120 = 230; Max(200) < Q_3 + 1.5(IQR), so no high outliers.

Q_1 - 1.5(IQR) = 30 - 120 = -90; Min(15) > Q_1 - 1.5(IQR), so no low outliers.

b. A price of $31.95 is slightly above the first quartile, so over 25% of phones cost
less. No, the advertised price would not be a very low price. (Or: The advertised
price is only 0.76 standard deviations below the mean. This is not an unusually low
price.)

Explanation:

74) Paying for purchases One day a store tracked the way shoppers paid for their purchases.   74) _____
Their data are summarized in the table.

	Cash	Check	Charge	Total
Male	18	10	12	40
Female	18	12	30	60
Total	36	22	42	100

a. What percent of the men paid cash?
b. What is the conditional relative frequency distribution of payment method for women?
c. If you wanted to show the association between gender and method of payment visually,
what kind of graph would you make? (Just name it.)
d. Is there evidence of an association between gender and method of payment? Explain
briefly.

Answer: a. 45%
         b. 30% cash, 20% check, 50% charge
         c. segmented bar graphs, or pie charts
         d. Yes. Women are more likely to charge their purchases than men (50% to 30%) and
         less likely to
            pay cash (30% to 45%).
Explanation:

In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses from boys and girls.

	Male	Female	Total
Bus	30	34	64
Ride	37	45	82
Personal	19	23	42
Total	86	102	188

75) Identify the variables and tell whether each is categorical or quantitative.   75) _____

Answer: Gender and mode of transportation, both categorical.
Explanation:

40

In June 2003 *Consumer Reports* published an article on some sport-utility vehicles they had tested recently. They reported some basic information about each of the vehicles and the results of some tests conducted by their staff. Among other things, the article told the brand of each vehicle, its price, and whether it had a standard or automatic transmission. They reported the vehicle's fuel economy, its acceleration (number of seconds to go from zero to 60 mph), and its braking distance to stop from 60 mph. The article also rated each vehicle's reliability as much better than average, better than average, average, worse, or much worse than average.

76) Describe the W's, if the information is given: 76) _____

• Who:

• What:

• When:

• Where:

• How:

• Why:

Answer: Who: SUV's currently on the market. We don't know how many models.
What: brand, price, transmission type, fuel economy, acceleration, braking distance, reliability
When: prior to June 2003
Where: not specified, probably the United States
How: testing the vehicles by driving each
Why: information for potential consumers

Explanation:

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from "Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes" (Shanthy A. Bowman, *Journal of the American Dietetic Association*, *102*(9), pp. 1234–1239):

Nationwide Food Survey Years

		1987-1988	1989-1991	1994-1996	Total
Drinks Fluid Milk	Yes	354	502	366	1222
	No	226	335	366	927
	Total	580	837	732	2149

77) Find the following: 77) _____

   a. What percent of the young girls reported that they drink milk?

   b. What percent of the young girls were in the 1989-1991 survey?

   c. What percent of the young girls who reported that they drink milk were in the 1989-1991 survey?

   d. What percent of the young girls in 1989-1991 reported that they drink milk?

   Answer: a. 56.9%
                 b. 38.9%
                 c. 41.1%
                 d. 60.0%
   Explanation:

78) Light bulbs are measured in lumens (light output), watts (energy used), and hours (life). A   78) _____
   standard white light bulb has a mean life of 675 hours and a standard deviation of 50
   hours. A soft white light bulb has a mean life of 700 hours and a standard deviation of 35
   hours. In a test at a local science competition, both light bulbs lasted 750 hours. Which light
   bulb's life span was better? Explain.

   Answer: Standard light bulb: $z = \dfrac{750 - 675}{50} = 1.5$ Soft light bulb: $z = \dfrac{750 - 700}{35} = 1.4286$

   The standard light bulb lasted more than 1.5 standard deviations above its mean life,
   compared to the soft light bulb at 1.4286 standard deviations above its mean. The
   standard light bulb's performance was slightly better.
   Explanation:

79) Veterinary costs Costs for standard veterinary services at a local animal hospital follow a 　79) _____
Normal model with a mean of $80 and a standard deviation of $20.

a. Draw and clearly label this model.
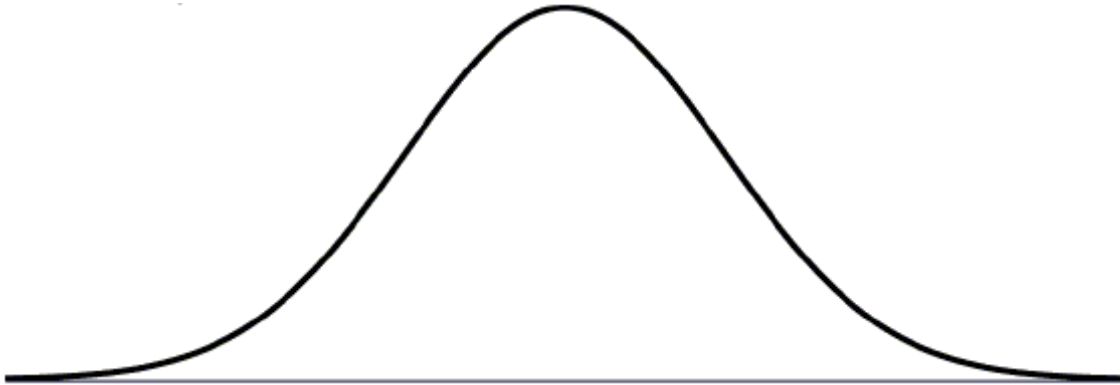
b. Is it unusual to have a veterinary bill for $125? Explain.

c. What is the IQR for the costs of standard veterinary services? Show your work.

Answer: a.
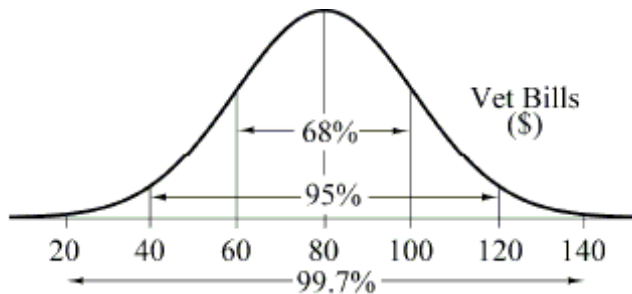
b. $z = \dfrac{125 - 80}{20} = 2.25$, more than 2 standard deviations above the mean bill. A

veterinary bill of $125 is unusual.

c. Q1 has $z = -0.67$ and Q3 has $z = +0.67$, so

$-0.67 = \dfrac{y - 80}{20} \Rightarrow y = 80 - 0.67(20) = 66.6$ and

$+0.67 = \dfrac{y - 80}{20} \Rightarrow y = 80 + 0.67(20) = 93.4$. The IQR = Q3 - Q1 = 93.4 - 66.6 =

$26.80.

Explanation:

80) The following are histograms for the heights of 100 US women and the heights of 100
European women:

Histogram of USWmnCm

Histogram of HtsforEUWmn
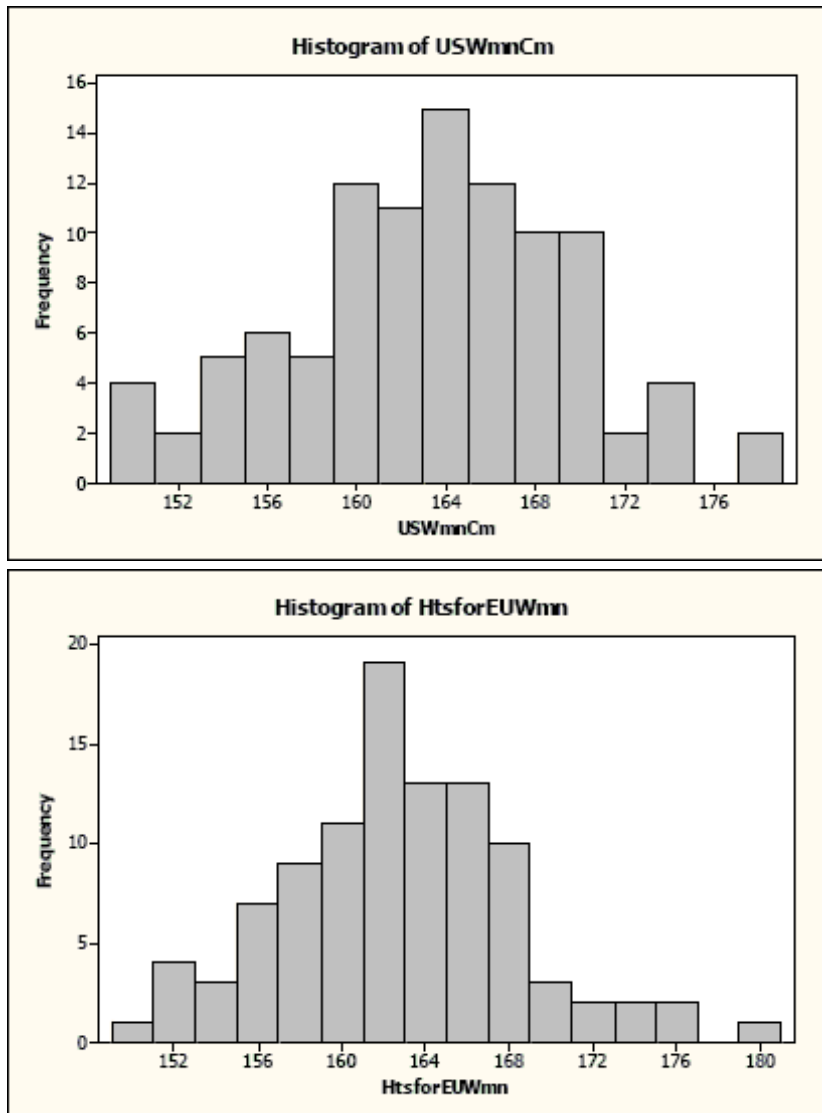
Compare the two distributions of the women's heights. Be sure to talk about shape, center, and spread.

Answer: Both distributions are unimodal and roughly symmetric. Each distribution appears to be centered around 164cm. The heights for the US women appear to be more spread out than those for the European women.

Explanation:

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

81) Which is true of the data shown in the histogram?   81) _____
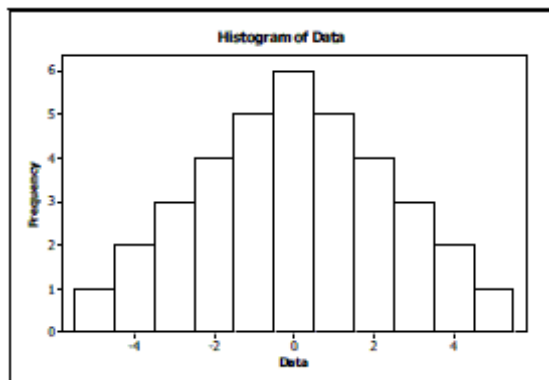
I.   The distribution is approximately symmetric.
II. The mean and median are approximately equal.
III. The median and IQR summarize the data better than the mean and standard deviation.
   A) I and II
   B) I only
   C) I, II, and III
   D) I and III
   E) III only

Answer: A
Explanation: A)
               B)
               C)
               D)
               E)

82) The veterinary bills for the dogs are summarized in the ogive below. Estimate the IQR of these   82) _____
    expenses.
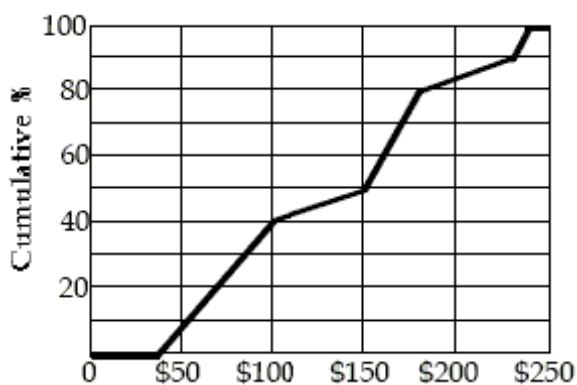
   A) $50   B) $200   C) $75   D) $150   E) $100

Answer: E
Explanation: A)
               B)
               C)
               D)
               E)

45

83) We might choose to display data with a stemplot rather than a boxplot because a stemplot   83) _____
   I. reveals the shape of the distribution.
   II. is better for large data sets.
   III. displays the actual data.
      A) II only
      B) I only
      C) I and III
      D) III only
      E) I, II, and III

   Answer: C
   Explanation: A)
                   B)
                   C)
                   D)
                   E)

84) Which is true of the data shown in the histogram?   84) _____
   I. The distribution is skewed to the right.
   II. The mean is probably smaller than the median.
   III. We should use median and IQR to summarize these data.
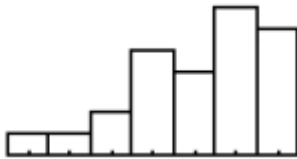
      A) II only
      B) II and III only
      C) I only
      D) III only
      E) I, II, and III

   Answer: B
   Explanation: A)
                   B)
                   C)
                   D)
                   E)

85) The mean number of hours worked for the 30 males was 6, and for the 20 females was 9. The   85) _____
   overall mean number of hours worked…
      A) none of these
      B) cannot be determined
      C) is 7.5
      D) is 7.2
      E) is 6.5

   Answer: D
   Explanation: A)
                   B)
                   C)
                   D)
                   E)

86) Your Stats teacher tells you your test score was the 3rd quartile for the class. Which is true?   86) _____
   I. You got 75% on the test.
   II. You can't really tell what this means without knowing the standard deviation.
   III. You can't really tell what this means unless the class distribution is nearly Normal.
    A) none of these
    B) I only
    C) II and III
    D) II only
    E) III only

Answer: A
Explanation: A)
               B)
               C)
               D)
               E)

87) Which is true of the data whose distribution is shown?   87) _____
   I. The distribution is skewed to the right.
   II. The mean is probably smaller than the median.
   III. We should summarize with mean and standard deviation.
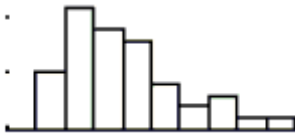
    A) II only
    B) I, II, and III
    C) I only
    D) II and III
    E) I and II

Answer: C
Explanation: A)
               B)
               C)
               D)
               E)

88) Suppose that a Normal model describes the acidity (pH) of rainwater, and that water tested after   88) _____
last week's storm had a z-score of 1.8. This means that the acidity of that rain…
    A) varied with a standard deviation of 1.8
    B) had a pH 1.8 higher than average rainfall.
    C) had a pH 1.8 times that of average rainwater.
    D) had a pH of 1.8.
    E) had a pH 1.8 standard deviations higher than that of average rainwater.

Answer: E
Explanation: A)
               B)
               C)
               D)
               E)

89) Suppose a Normal model describes the number of pages printer ink cartridges last. If we keep track   89) _____
of printed pages for the 47 printers at a company's office, which must be true?
I.   The page counts for those ink cartridges will be normally distributed.
II. The histogram for those page counts will be symmetric.
III. 95% of those page counts will be within 2 standard deviations of the mean.
   A) none
   B) I, II, and III
   C) II only
   D) II and III
   E) I only

Answer: A
Explanation: A)
                 B)
                 C)
                 D)
                 E)

90) Environmental researchers have collected rain acidity data for several decades. They want to see if   90) _____
there is any evidence that attempts to reduce industrial pollution have produced a trend toward
less acidic rainfall. They should display their data in a(n)…
   A) bar graph
   B) contingency table
   C) boxplot
   D) timeplot
   E) histogram

Answer: D
Explanation: A)
                 B)
                 C)
                 D)
                 E)

91) The ages of people attending the opening show of a new movie are summarized in the ogive shown. Estimate the IQR of the ages.

91) _____
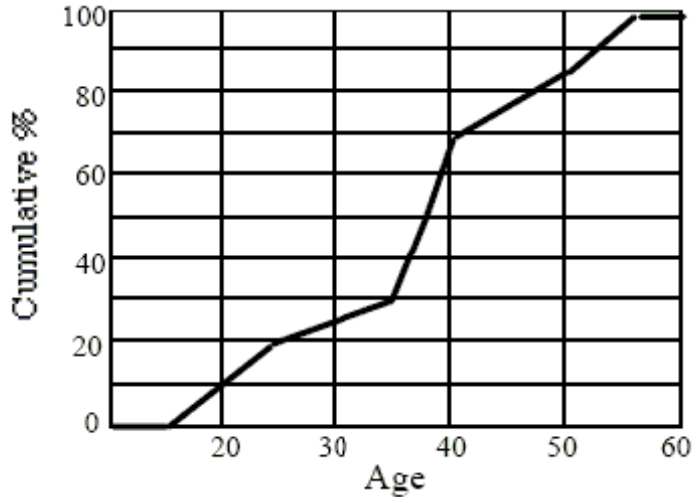
A) 21   B) 5   C) 13   D) 37   E) 30

Answer: C
Explanation: A)
                  B)
                  C)
                  D)
                  E)

92) The SPCA has kept these data records for the past 20 years. If they want to show the trend in the number of dogs they have housed, what kind of plot should they make?

92) _____

A) bar graph   B) histogram   C) boxplot   D) pie chart   E) timeplot

Answer: E
Explanation: A)
                  B)
                  C)
                  D)
                  E)

93) Two sections of a class took the same quiz. Section A had 15 students who had a mean score of 80, and Section B had 20 students who had a mean score of 90. Overall, what was the approximate mean score for all of the students on the quiz?

93) _____

A) It cannot be determined.
B) 85.7
C) none of these
D) 84.3
E) 85.0

Answer: B
Explanation: A)
                  B)
                  C)
                  D)
                  E)

94) Which of the following data summaries are changed by adding a constant to each data value? 94) _____
    I. the mean
    II. the median
    III. the standard deviation
      A) I only
      B) I, II, and III
      C) I and II
      D) I and III
      E) III only

Answer: C
Explanation: A)
                B)
                C)
                D)
                E)

95) Last weekend police ticketed 18 men whose mean speed was 72 miles per hour, and 30 women 95) _____
going an average of 64 mph. Overall, what was the mean speed of all the people ticketed?
      A) 67 mph
      B) It cannot be determined.
      C) none of these
      D) 68 mph
      E) 69 mph

Answer: A
Explanation: A)
                B)
                C)
                D)
                E)

1) C
2) E
3) C
4) C
5) B
6) A
7) E
8) E
9) D
10) A
11) A
12) A
13) B
14) D
15) A
16) The way students get to school does seem to be independent of gender. Overall, 34% of students ride the bus, compared to 35% of the boys and 33% of the girls. 44% of all students caught rides with someone and 22% used personal transportation, almost the same as the percentages for boys (43% and 22%) or girls (44% and 23%) separately. These data provide little indication of a difference in mode of transportation between boys and girls at this school.
17) Who: 8th, 10th, and 12th graders
What: alcohol, illegal drug, and cigarette use
When: Spring 2004
Where: United States
How: survey
Why: "to study changes in the beliefs, attitudes, and behavior of young people in the United States"
18) a. Probably higher. The data appear to be skewed to the right.
b. IQR = 3.3 - 2.8 = 0.5. Since 1.5(IQR) = 0.75, the fences are 2.8 - 0.75 = 2.05 and 3.3 + 0.75 = 4.05. The fish weighing 4.2 pounds is more than 1.5 IQRs outside the quartiles, so it could be considered an outlier.
19) The Census Bureau projects dramatic changes in the female population of Mexico over the next 50 years. The current distribution of ages is strongly skewed to the right with most of the women under 30 and far fewer 50 and above. By 2050 the population will become more uniform across age groups from 0 to 60, and we anticipate an unusually large number of women over 80.
20) a. Yes: 90.7%   No: 9.3%
b. Yes: 90.2%   No: 9.8%
c. segmented bar graph or pie charts
d. Perhaps. There is little difference between the percents of those who like dogs, depending on whether they like cats. Of those who like cats, only 90.2% like dogs, compared to 90.7% overall.
21) a. mean: increase
b. median: stay about the same
c. range: increase
d. IQR: stay about the same
e. standard deviation: increase
22) a. The mean temperature of all students would probably be higher than the median. Using the five-number summary, it appears the data are skewed to the right.
b. IQR = 98.6° - 97.85° = 0.75°. Since 1.5(IQR) = 1.125°, the fences are 97.85° - 1.125° = 96.725° and 98.6° + 1.125° = 99.725°. The lowest temperature (96.7°) being added to the data set is smaller than the lower fence (96.725°) so it is an outlier on the low end. The highest temperature (99.2°) being added to the data set is not above the upper fence (99.725°) so it is not an outlier on the high end.
23) Categorical: species, geographic location, how new ant colonies begin, and nest design.
Quantitative: nest depth (feet), number of chambers (units), and colony size (units).

24) a. $z = \dfrac{80 - 75}{4.2} = 1.19$

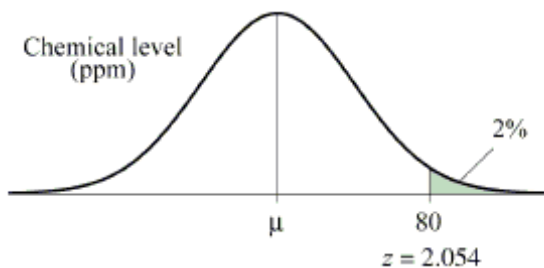$P(z \geq 1.19) \approx 0.117$
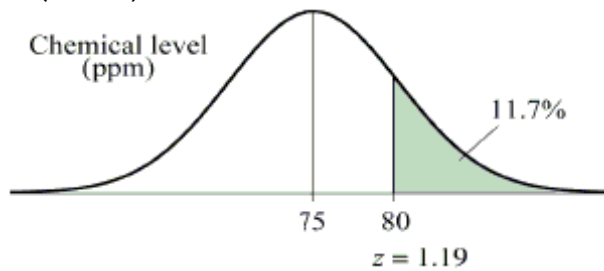
According to the normal model, we expect about 11.7% of the batches to exceed the 80ppm standard.

b. $z = 2.054$

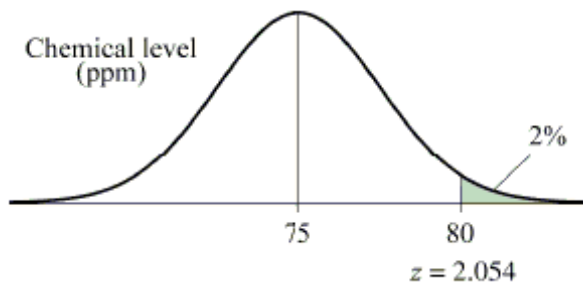$\dfrac{80 - \mu}{4.2} = 2.054$

$\mu = 80 - 2.054(4.2) = 71.37$

According to the Normal model, a mean of about 71.37ppm would need to be achieved.

c. $\dfrac{80 - 75}{\sigma} = 2.054$

$\sigma \approx 2.434$

According to the Normal model, the new standard deviation would need to be at most 2.43ppm.

d. The scrubber must be more consistent in its performance from batch to batch.

25) a. $z = \dfrac{97 - 98.7}{0.7} = -2.43$, so $P(z < -2.43) = 0.0075$

About 0.75% of people have ear temperatures that may indicate hypothermia.

b. The z-scores associated with the IQR are $z = -0.67$ and $z = 0.67$. So, we need to solve for y in each of the following

equations: $-0.67 = \dfrac{y - 98.7}{0.7}$ and $0.67 = \dfrac{y - 98.7}{0.7}$. We get $y = 98.7 - 0.67(0.7) = 98.2$ and $y = 98.7 + 0.67(0.7) = 99.2$. The

interquartile range is IQR = 99.2°F - 98.2°F = 1.0°F.

c. The new IQR is 0.5°F, while the old IQR was 1.0°F. So, we want

IQR = $[98.7 + 0.67\sigma] - [98.7 - 0.67\sigma] = 0.5$, or $1.34\sigma = 0.5$. Thus, $\sigma = \dfrac{0.5}{1.34} = 0.37$. Our new standard deviation is 0.37°F.

26) a. The junior class had the lowest cumulative GPA, about 1.6.

b. The sophomore class had the highest median cumulative GPA, about 3.2.

c. The junior class had the largest range for GPA, about 2.4.

d. The junior class had the most symmetric set of GPAs. The sophomore class had the most skewed set of GPAs, skewed to the left.

27)
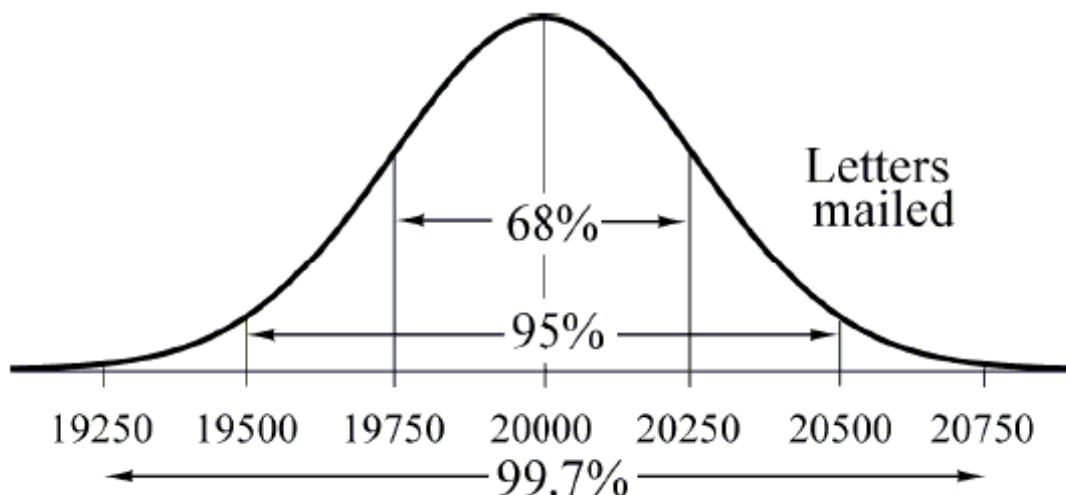
28) Dalmation: $z_D = \dfrac{45 - 50}{3.3} = -1.52$

Boxer: $z_B = \dfrac{52 - 57.5}{1.7} = -3.24$

The Dalmatian is 1.52 standard deviations underweight, while the Boxer is 3.24 standard deviations underweight. So, the Boxer is more underweight.

29) a. mean: increase

b. median: stay about the same

c. range: increase

d. IQR: stay about the same

e. standard deviation: increase

30) a. mean: increase

b. median: stay about the same

c. range: increase

d. IQR: stay about the same

e. standard deviation: increase

31) a. Car Z: $5000
   b. BuyIt: $10,000
   c. Ace: $10,000
   d. CarZ: $3000
   e. BuyIt; half of their cars are cheaper than any of the cars at Ace, and 25% of their cars are cheaper than all but one car at CarZ. The third quartile of their prices is well below the third quartile at CarZ, and below even the median price at Ace.
32) Gender and Breed; both categorical.
33) The Jersey's milk production was comparatively higher. That cow gave slightly more than 2 standard deviations above the average amount of milk ($z = 2.2$), while the Guernsey gave less than 2 standard deviations more than the average for Guernseys ($z = 1.875$).
34) a. Yes. IQR = 308 - 88 = 220. The upper fence for outliers is one and a half IQR's above the third quartile, or 308 + 1.5(220) = 638. The maximum repair bill was $1442, well above $638, so it is certainly an outlier.
   b. No. $90 is higher than over 25% of the bills, so it is not unusual.
35) a.
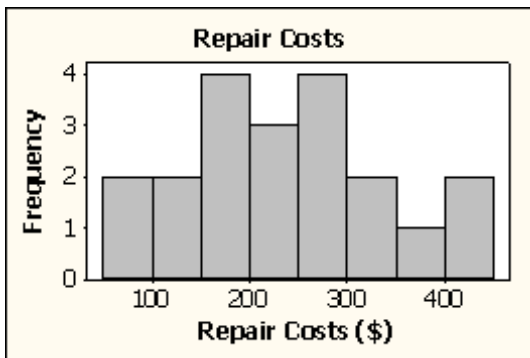
   b. $\bar{x}$ = $236.25; s = $103.43
   c. Yes, the data are roughly unimodal and symmetric with no outliers.
   d. The repair costs averaged $236.25, ranging from $56 to $422 with a standard deviation of $103.43. The distribution was approximately symmetric, with typical repair costs clustered between $150 and $300.
36) Among females, 20.9% chose Yorkshire Terriers, 14.2% Dachshunds, 16.7% Golden Retrievers, 10.6% Labs, and 12.9% Dalmatians. The remaining 24.7% of females preferred other breeds.
37) a. Car: 19.5%   Bus: 30%   Train: 50.5%
   b. Car: 28.9%   Bus: 22.2%   Train: 48.9%
   c. segmented bar graph, or pie charts
   d. No, there is a difference between the percents in two types of transportation - Car and Bus categories, depending on the Job Classification.

	Car	Bus	Train
Management	28.9%	22.2%	48.9%
Labor	17.0%	32.1%	50.9%

   Although about half of each group take the train, management are more likely than labor to come by car and less likely to take a bus.
38) a. 6.6%
   b. 14.9%
   c. 52.6%
39) Categorical: brand, transmission type, reliability
   Quantitative: price (US$), fuel economy (mpg), acceleration (seconds), braking distance (probably feet?)
40) Categorical: sex, only child?, major
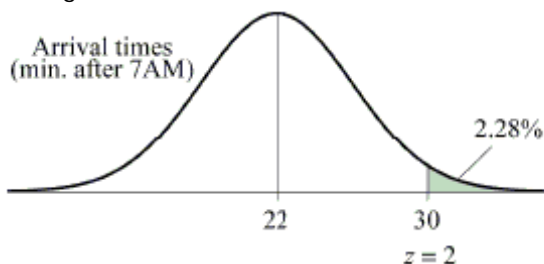   Quantitative: age (years), height (inches), weight (pounds), credit hours, GPA
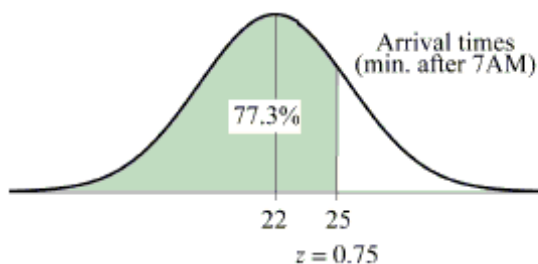
41) a. Employees are late if they arrive after 7:30 A.M.

$P(\text{time} > 7:30) = P(z > \dfrac{30 - 22}{4} = 2) = 0.0228$

According to the Normal model, about 2.28% of employees are expected to arrive after 7:30 AM.

b. $P(x \leq 25) = P(z < \dfrac{25 - 22}{4} = 0.75) = 0.773$

According to the Normal model, about 77.3% of employees arrive at work before 7:25 AM.

c. Virtually all times lie within 3 standard deviations of the mean. (Accept other reasonable z-scores greater than 3). If

$z = 3 \leq \dfrac{30 - 22}{\sigma}$, then $3\sigma \leq 8$, so $\sigma \leq 2.67$ minutes.

d. A smaller standard deviation would mean greater consistency in arrival times.

42) a.

Statistic	Hours of Service Call	Cost of Service Call
Mean	4.5	$210
Median	3.5	$170
SD	1.2	$48
IQR	2.0	$80
Minimum	0.5	$50

b. -1.50

43) a.

b. 28, 56. The claim is probably false. This model should provide a useful estimate of what might happen, but is not certain to predict what actually will happen.

c. 46.7 months

d. 19.6%

e. 4.7 months (should all include sketches of labeled curves)

f. A smaller standard deviation means that the copiers would be more consistent in their lifespans.

g. For 36 months $z = -1.28$ and for 42 months $z = -0.385$. Thus the difference of 6 months is $1.28 - 0.385 = 0.895$ standard deviations. The model is $N(44.6, 6.7)$.

44) No. 56.9% of all young girls surveyed reported drinking milk, but 60% of the young girls reported drinking milk in the 1989-1991 survey. Since these percentages differ, milk consumption and year are not independent.

45) a. IQR = Q3 - Q1 = 16.7 - 10.8 = 5.9
   1.5(IQR) = 1.5(5.9) = 8.85
   Q3 + 1.5(IQR) = 16.7 + 8.85 = 25.55 < Max, so there is at least one high outlier
   Q1 - 1.5(IQR) = 10.8 - 8.85 = 1.95 < Min, so there are no low outliers

b. It is more appropriate to use the median and IQR to describe these data, since the distribution is skewed right.

46)
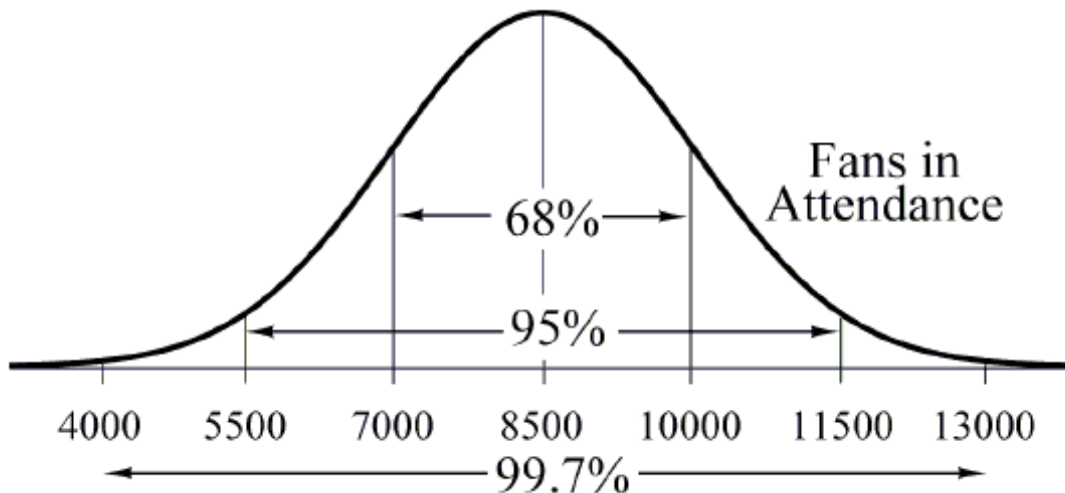
47) There were 132 Yorkshire terrier responses, 96 Dachshund responses, 91 Golden Retriever responses, 78 Labrador responses, 73 Dalmation responses, and 153 Other responses.

48) We do not know how or when the people were surveyed, or where the local dog show was located.

49) Testbank Starts at Ch. 2

50) There are 86 males and 102 females.

51) We don't know how or when the students were surveyed, nor where the school is.

52) a.

Statistic	Books Sold	\$ Earned
Mean	640	\$920
Standard deviation	360	\$180
IQR	450	\$225
Maximum	1420	\$1310

   b. +1.80
53) The frequency scales are not the same for the two histograms. If we converted each of the frequency histograms to a relative frequency histogram, we would be better able to compare the frequencies for each distribution at the heights.
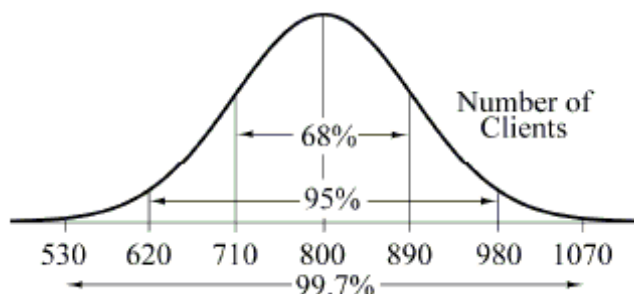54)   a. 3400 yuans, 2000, 6200, 2400
   b. $z = +1.20$
55) a. 18.1%
   b. 33.3%
   c. 53.1%
56) The breed selection does not appear to be independent of gender. Overall, 56% of the respondents were females, but females were over-represented among those who favored Golden Retrievers (64%) and Dalmatians (62%), yet a much lower percentage (47%) among those who chose Labradors.
57) a.

   b. $Q_1 \Rightarrow P = 0.25$ and $z = -0.674$,

$$-0.674 = \frac{x - 800}{90} \qquad -60.66 = x - 800$$

   $x = 739.34$
   So the first quartile is at 740 clients.

   c. 450 (5th percentile) has $z = -1.645$

   1085 (60th percentile) has $z = +0.253$
   $1085 - 450 = (0.253 + 1.645)\sigma$
   $\sigma \approx 334.5$
   $\mu + 0.253(334.5) = 1085$
 $\mu \approx 1000.4$
   N(1000.4, 334.5)

58) The distributions of high school GPA for both males and females are skewed to the left, and both distributions appear to be centered at a GPA of about 3.0. The distribution of male GPA appears slightly more spread out than the distribution of female GPA.

59) a. Passenger, less than 1 year
  b. Passenger, 21 years
  c. Cyclist, 40 years
  d. Pedestrian, 44 years
  e. Pedestrian. While the oldest person involved in an accident is not a pedestrian, the median age for pedestrians is almost 45 years, while the median ages in the other groups are between 22 and 35 years old. The oldest 50% of the Pedestrian group, from 45 to 87 years, is generally older than the youngest 75% of two groups - Cyclist and Passenger, and only the Driver group has any of its middle 50% as old. The Driver and Passenger groups have a few people older than the Pedestrian group.

60) a. At least 25% of the housing units have only one resident.

  b: 4 residents is just above the 3rd quartile, and only about one standard deviation above the mean. It is not an unusually high number of residents.
  c. Median = 40 + 5(2) = $50; IQR = 5(3 - 1) = $10
  d. Mean = 40 + 5(2.53) = $52.65; SD = 5(1.50) = $7.50

61) a. The concrete roadway is under minimum depth when less then 23 inches in thickness.

  $z = \dfrac{23 - 26}{1.75} = -1.7 \rightarrow P = 0.043$, so the model suggests about 4.3% is under the minimum depth

  b. $P = 0.03 \rightarrow z = -1.88$, so $-1.88 = \dfrac{23 - 26}{\sigma}$; then $\sigma = 1.6$ inches

  c. A smaller standard deviation means that the thickness of the concrete will be more consistent.

62) a.
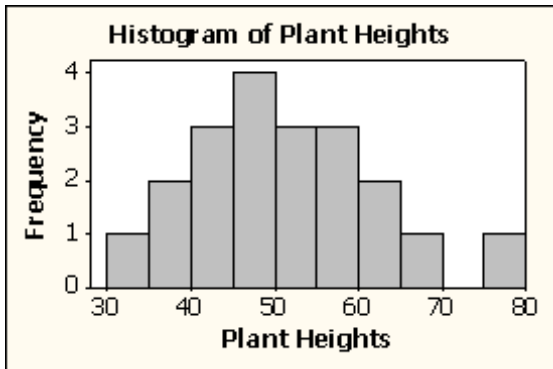
  b. $\bar{x}$ = 51.0 cm; s = 10.6 cm
  c. Yes, the data are roughly unimodal and symmetric with no outliers.
  d. The data are roughly symmetric with no outliers; however there is a small gap from 70 to 75 cm. The average plant height is 51.0 centimeters, with a standard deviation of 10.6 centimeters. The range of plant heights is 43 centimeters. The distribution of plant heights has a mode between 45 and 49 centimeters.

63) More boys (43%) caught rides to school than any other means of transportation. 35% rode the bus while only 22% used personal transportation like biking, skateboarding, or walking.

64) a.
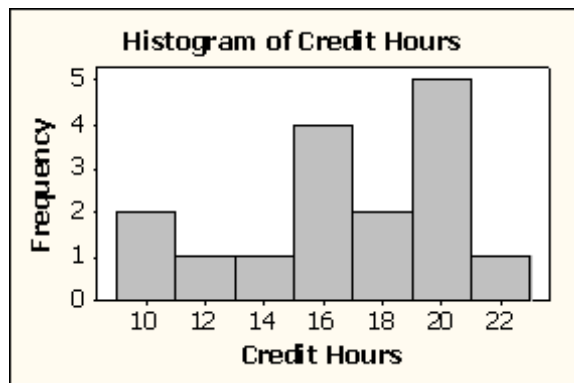
b. $\bar{x}$ = 16.3 credit hours; s = 3.7 credit hours

c. The median is 16.0 credit hours.

   IQR = Q3 - Q1 = 20 - 14.5 = 5.5 credit hours

d. It is more appropriate to use the median and IQR to summarize these data, because these data are not unimodal and symmetric.

65) a. The mean midterm score of all students would probably be lower than the median. Using the five-number summary, it appears that the data are skewed to the left.

   b. IQR = 43.5 - 32 = 11.5

   Q1 - 1.5IQR = 32 - 1.5(11.5) = 14.75

   Q3 + 1.5IQR = 43.5 + 1.5(11.5) = 60.75

   Since both the maximum and minimum scores fall between these "fences", there are no outliers in this data set.

66) Yes: 56.9%; No: 43.1%

67) a. $z = \dfrac{12 - 12.1}{0.05}$ = -2.0, which suggests that 2.28% of cans are under-filled.

   b. A z-score of -2.33 has 1% to its left, meaning that 1% of the cans would be under-filled.

   i. $-2.33 = \dfrac{12. - 12.1}{\sigma} \Rightarrow -2.33\sigma = -0.1 \Rightarrow \sigma = \dfrac{-0.1}{-2.33}$ = 0.043. The standard deviation would need

   to be 0.043 ounces.

   ii. $-2.33 = \dfrac{12. - \mu}{0.05} \Rightarrow -2.33(0.05) = 12 - \mu \Rightarrow \mu = 12 + 2.33(0.05)$ = 12.12. The mean would need to be 12.12 ounces.

68) No. It looks like there is some sort of relationship between milk consumption and nationwide survey year, since the percentage of young girls who reported drinking milk is a larger slice of the pie chart for the 1989-1991 survey than the same response for the 1994-1996 survey.

69) a. i. $35 + $73(16.65) = $1250.45

   ii. $73(2.96) = $216.08

   iii. $35 + $73(16) = $1203

   iv. IQR = $73(19-15) = $292

   b. IQR = 19 - 15 = 4 credit hours

   High outliers will lie above Q3 + 1.5IQR = 19 + 1.5(4) = 25 credit hours. Since 28 credit hours exceeds 25 credit hours, I would consider 28 credit hours to be unusually high.

70) a. The top 25 percent of all classes have 40 students enrolled.

   b. Yes, classes with 20 students enrolled seem unusually small. Twenty is well below the first quartile of 29 students, and only slightly above the minimum size (17).

   $z = \dfrac{20 - 33.39}{5.66} = -2.366$ With $z = -2.366$, this size class is over 2 standard deviation units below the mean.

   c. Median budget = $12 + $0.75(33) = $36.75
   Q1 budget = $12 + $0.75(29) = $33.75
   Q3 budget = $12 + $0.75(40) = $42.00
   IQR = $42.00 - $33.75 = $8.25

   d. Mean budget = $12 + $0.75(33.39) = $37.04
   Standard deviation = $0.75(5.66) = $4.25

71) Who: Colonies of ants. "Many species of ants," but no indication of exactly how many.
   What: scientific name, geographic location, average nest depth, average number of chambers, average colony size, how new ant colonies begin, the ant-nest design, and how nests differ in architecture.
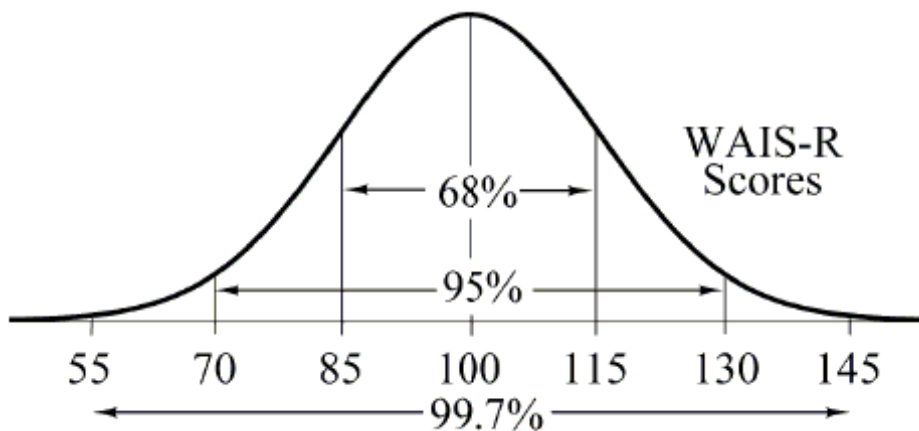   When: November 2003
   Where: not specified
   How: The results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida
   Why: Information of interest to readers of the magazine

72)

73) a. IQR = Q_3 - Q_1 = 110 - 30 = 80

   1.5(IQR) = 1.5(80) = 120

   Q_3 + 1.5(IQR) = 110 + 120 = 230; Max(200) < Q_3 + 1.5(IQR), so no high outliers.

   Q_1 - 1.5(IQR) = 30 - 120 = -90; Min(15) > Q_1 - 1.5(IQR), so no low outliers.

   b. A price of $31.95 is slightly above the first quartile, so over 25% of phones cost less. No, the advertised price would not be a very low price. (Or: The advertised price is only 0.76 standard deviations below the mean. This is not an unusually low price.)

74) a. 45%
   b. 30% cash, 20% check, 50% charge
   c. segmented bar graphs, or pie charts
   d. Yes. Women are more likely to charge their purchases than men (50% to 30%) and less likely to pay cash (30% to 45%).

75) Gender and mode of transportation, both categorical.

Answer Key
Testname: PART1

76) Who: SUV's currently on the market. We don't know how many models.
   What: brand, price, transmission type, fuel economy, acceleration, braking distance, reliability
   When: prior to June 2003
   Where: not specified, probably the United States
   How: testing the vehicles by driving each
   Why: information for potential consumers

77) a. 56.9%
    b. 38.9%
    c. 41.1%
    d. 60.0%

78) Standard light bulb: $z = \dfrac{750 - 675}{50} = 1.5$ Soft light bulb: $z = \dfrac{750 - 700}{35} = 1.4286$

   The standard light bulb lasted more than 1.5 standard deviations above its mean life, compared to the soft light bulb at 1.4286 standard deviations above its mean. The standard light bulb's performance was slightly better.

79) a.
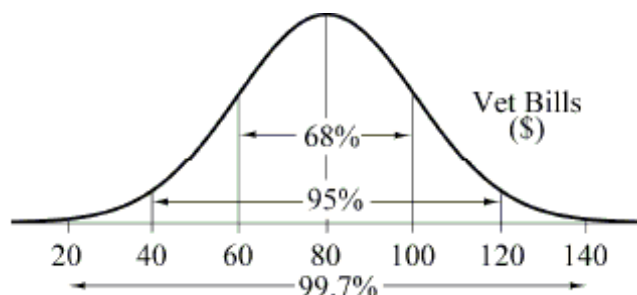
   b. $z = \dfrac{125 - 80}{20} = 2.25$, more than 2 standard deviations above the mean bill. A veterinary bill of \$125 is unusual.

   c. Q1 has $z = -0.67$ and Q3 has $z = +0.67$, so

   $-0.67 = \dfrac{y - 80}{20} \Rightarrow y = 80 - 0.67(20) = 66.6$ and

   $+0.67 = \dfrac{y - 80}{20} \Rightarrow y = 80 + 0.67(20) = 93.4$. The IQR = Q3 – Q1 = 93.4 – 66.6 = \$26.80.

80) Both distributions are unimodal and roughly symmetric. Each distribution appears to be centered around 164cm. The heights for the US women appear to be more spread out than those for the European women.

81) A
82) E
83) C
84) B
85) D
86) A
87) C
88) E
89) A
90) D
91) C
92) E
93) B
94) C
95) A