Exam

Name_____

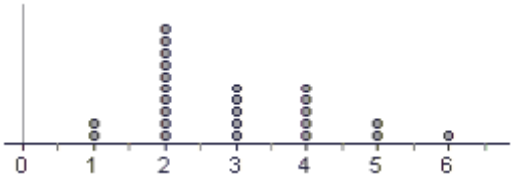
MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

1) Suppose that a Normal model describes the acidity (pH) of rainwater, and that water tested after last week's storm had a *z*–score of 1.8. This means that the acidity of that rain … 1) _____

A) had a pH 1.8 standard deviations higher than that of average rainwater.
B) varied with a standard deviation of 1.8.
C) had a pH 1.8 times that of average rainwater.
D) had a pH 1.8 higher than average rainfall.
E) had a pH of 1.8.

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

2) The distribution below is the number of family members reported by 25 people in the 2010 Census. 2) _____




The best description for the shape of this distribution is

A) skewed right
B) normal
C) bimodal
D) approximately normal
E) skewed left

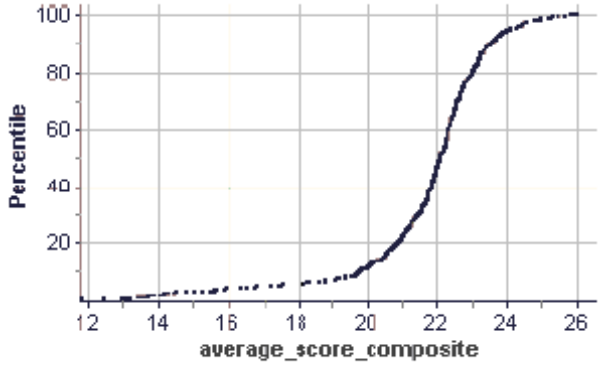
Answer: A
Explanation: A)
B)
C)
D)
E)

3) The advantage of making a stem-and-leaf display instead of a dotplot is that a stem-and-leaf display 3) _____

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

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

4) The average scores for Wisconsin schools on the ACT are summarized in the ogive shown. Estimate the IQR of the average scores. 4) _____




A) 7 B) 50 C) 6 D) 2.8 E) 1.6

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

5) The United States Census collects data on many variables about individuals and households. Which variable is categorical? 5) _____

A) family size
B) hours worked per week
C) annual electricity cost
D) type of residence
E) monthly mortgage

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

6) Which variable about German Shepherds is most likely to be described by a Normal model? 6) _____

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

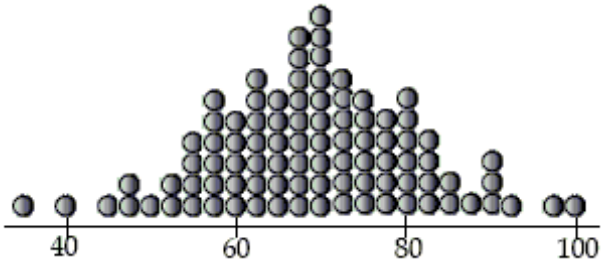
Answer: E
Explanation: A)
B)
C)
D)
E)

7) Suppose that a Normal model described student scores in a history class. Parker has a standardized score (z-score) of +2.5. This means that Parker 7) _____

A) is 2.5 standard deviations above average for the class.
B) has a standard deviation of 2.5.
C) has a score that is 2.5 times the average for the class.
D) is 2.5 points above average for the class.
E) none of these

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

8) The standard deviation of the data displayed in this dotplot is most likely to be… 8) _____




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

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

9) Which type of plot would be most appropriate to display the responses to *type of residence*? 9) _____

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

Answer: E

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

10) The five-number summary of credit hours for 24 students in a statistics class is: 10) _____

| Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|
| 13.0 | 15.0 | 16.5 | 18.0 | 22.0 |

Which statement is true?

A) There is at least one high outlier in the data.
B) There are no outliers in the data.
C) There is at least one low outlier in the data.
D) None of these.
E) There are both low and high outliers in the data.

Answer: B

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

11) Which of these variables is most likely to be bimodal? 11) _____

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

Answer: D

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

12) Suppose that a Normal model describes fuel economy (miles per gallon) for automobiles and that a Saturn has a standardized score (z-score) of +2.2. This means that Saturns . . . 12) _____

A) achieve fuel economy that is 2.2 standard deviations better than the average car.
B) get 2.2 mpg more than the average car.
C) have a standard deviation of 2.2 mpg.
D) get 2.2 miles per gallon.
E) get 2.2 times the gas mileage of the average car.

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

13) Which of these variables is most likely to follow a Normal model for U.S. adults? 13) _____

A) ACT scores
B) commuting time
C) monthly mortgage
D) eye color
E) income

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

14) To see if your school's ACT scores are improving since implementing a new program, the ACT scores for the past several years should be displayed in a 14) _____

A) contingency table
B) timeplot
C) histogram
D) bar graph
E) boxplot

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

15) Which of the following variables would most likely follow a Normal model? 15) _____

A) all of these
B) weights of adult male elephants
C) heights of singers in a co-ed choir
D) scores on an easy test
E) family income

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

16) Which type of plot would be least likely to reveal that a distribution is bimodal? 16) _____

A) dotplot B) histogram C) boxplot D) ogive E) stemplot

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

17) 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? 17) _____

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

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

18) We collect these data from 50 male students. Which variable is categorical? 18) _____

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

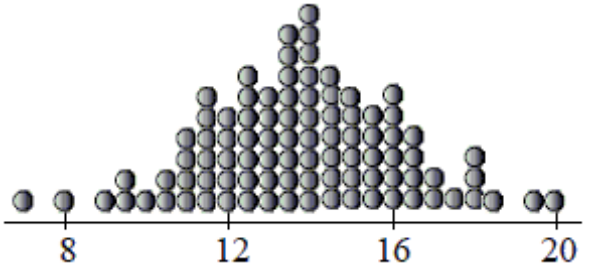
Answer: D
Explanation: A)
B)
C)
D)
E)

19) The SPCA collects the following data about the dogs they house. Which is categorical? 19) _____

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

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

20) The standard deviation of the data displayed in this dotplot is closest to… 20) _____




A) 4 B) 2 C) 0.5 D) 1 E) 3.5

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

21) 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? 21) _____

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

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

22) School administrators collect data on students attending the school. Which of the following variables is quantitative? 22) _____

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

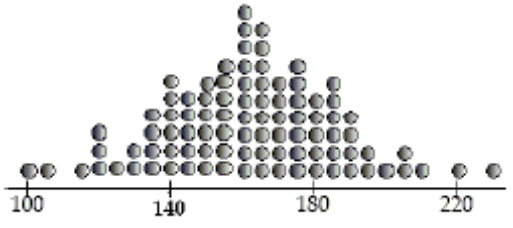
Answer: B
Explanation: A)
B)
C)
D)
E)

23) Which of these variables is most likely to follow a Normal model? 23) _____

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

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

24) The best estimate of the standard deviation of the mens' weights displayed in this dotplot is 24) _____




A) 35 B) 40 C) 15 D) 10 E) 25

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

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

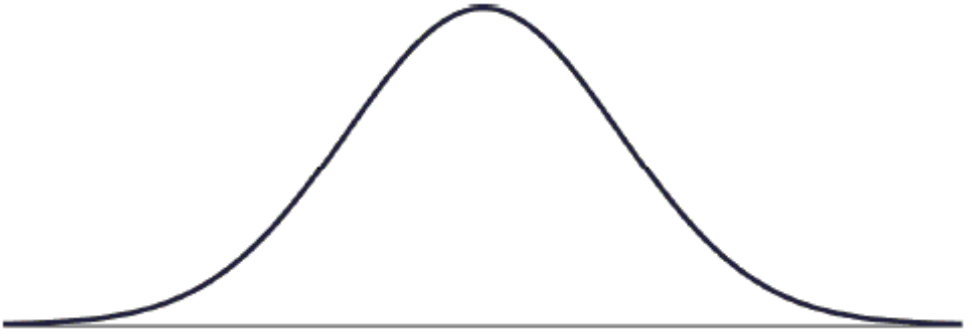
In November 2003 *Discover* published an article on the colonies of ants. They reported some basic information about many species of ants and the results of some discoveries found by myrmecologist Walter Tschinkel of the University of Florida. Information included the scientific name of the ant species, 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.

25) List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units. 25) _____

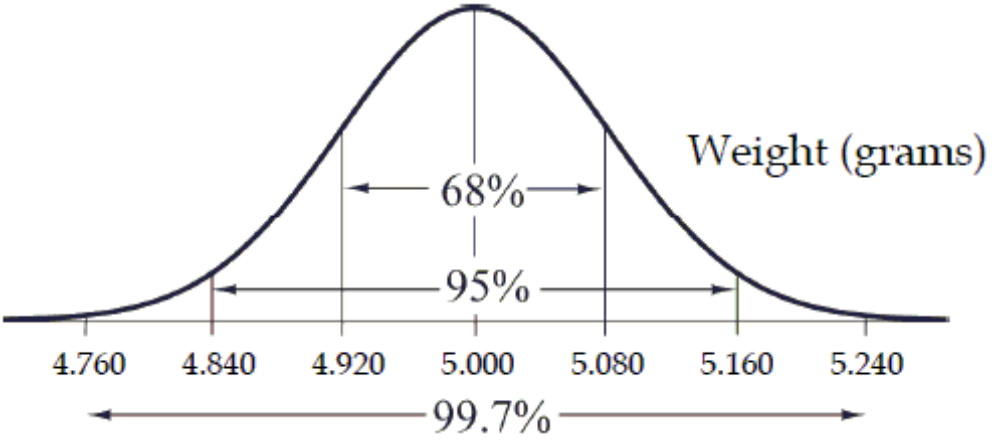
Answer: Categorical: species, geographic location, how new ant colonies begin, and nest design.
Quantitative: nest depth (feet), number of chambers (units), and colony size (units).

Explanation:

26) Nickels minted in the United States are supposed to weigh 5.000 grams. Of course there is some variation in that. The actual weights are pretty well represented by a normal model with a mean of 5.000 g and a standard deviation of about 0.08 g. Draw and clearly label this model. 26) _____



Answer:




Explanation:

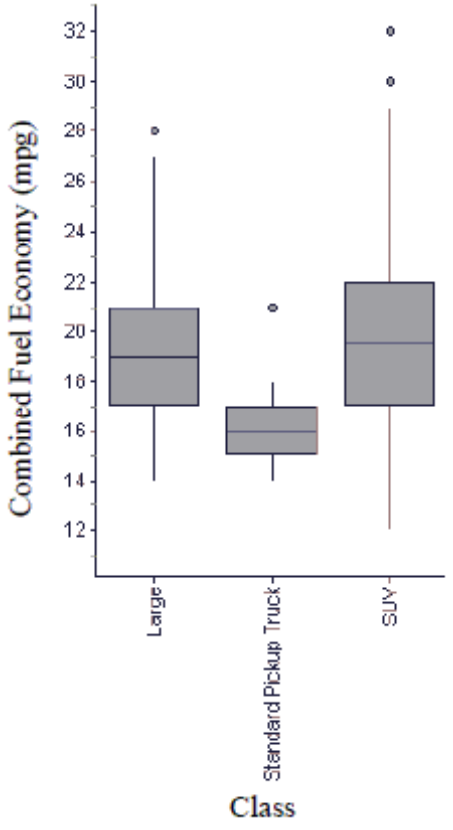
In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses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

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

27) What is the marginal distribution of gender? 27) _____

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

28) The boxplots show fuel economy of 2011 model cars for the classes shown. 28) _____

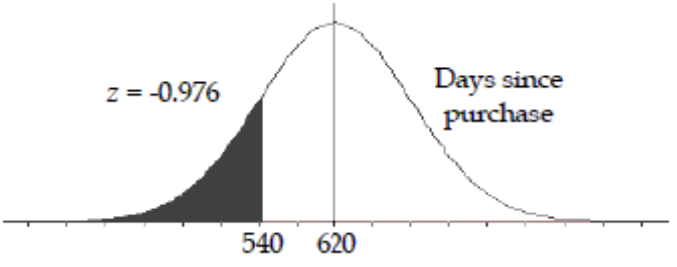



a. Which class offers the car with the best gas mileage, and what is the mileage of that car?
b. Which class has the highest median gas mileage, and how much is it?
c. Which class of car has the smallest range of fuel economy, and what is it?
d. Which class of car has the smallest IQR, and what is it?
e. Which class of car generally gets the best mileage? Explain.

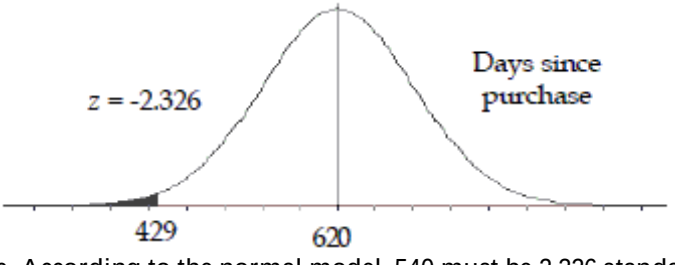
Answer: a. SUV: 32 mpg
b. SUV: 19.5 mpg
c. Pickup: 7 mpg
d. Pickup: 2 mpg
e. SUV; 3/4 of SUVs get better mileage than 3/4 of the Pickup Trucks. SUVs have the highest median, third quartile, and maximum. The first quartile is the same as that of Large cars, and the minimum is the lowest of all three, but overall SUVs appear to have the best mileage.
Explanation:

29) The lifespans of a particular brand of graphing calculator are approximately normally distributed with a mean of 620 days from the purchase date and a standard deviation of 82 days. They will provide a warranty that guarantees a replacement if the calculator stops working within the specified time frame, and are trying to decide what time frame to use. 29) _____
a. If the company sets the warranty at a year and a half (say 540 days), what proportion of calculators will they have to replace?
b. The company does not want to have to replace more than 1% of the calculators they sell. What length of time should they set for the warranty?
c. The company would like to set the warranty for 540 days, and still replace no more than 1% of the calculators sold. Increasing the average life of the calculators is too expensive, but they think they reduce the standard deviation of the lifespans. What standard deviation of lifespans would be needed to make this happen?
d. Explain what achieving a smaller standard deviation means in this context.

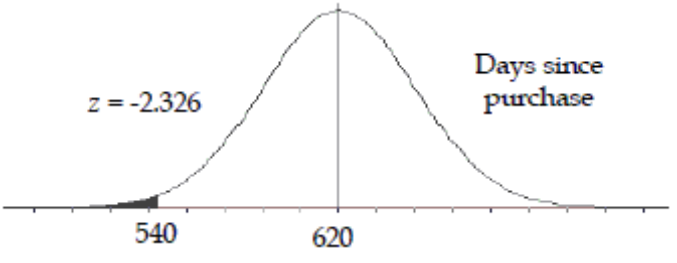
Answer: a. According to the normal model, we expect about 16.75% of the calculators to break down before 540 days.According to the normal model, we expect about 16.75% of the calculators to break down before 540 days.




b. According to the normal model, they would need to set the warranty at 429 days.




c. According to the normal model, 540 must be 2.326 standard deviations below the mean, so the standard deviation is about 34.4 days.




d. It means they would have to make the calculators more consistent in their

Answer: lifespan.
Explanation:

30) 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] 30) _____

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 $-1.88 = \frac{23 - 26}{\sigma}$ ; then $\sigma = 1.6$ inches

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

Explanation:

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.

31) Describe the W's, if the information is given: 31) _____

· Who:
· What:
· When:
· Where:
· How:
· Why:

Answer: · Who: SUV's currently on the market. We don't know how many models.
· What: brand of vehicle, price, type of transmission, fuel economy, acceleration, braking distance, and 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:

32) 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. 32) _____

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:

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** |

33) Find each percent. 33) _____

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:

A research company frequently monitors trends in the use of social media by American Adults. The results of one survey of 1846 randomly selected adults looked at social media use versus age group. The table summarizes the survey results.

| | | Age Group | | | | |
|---|---|---|---|---|---|---|
| | | 18-29 | 30-49 | 50-64 | 65+ | **Total** |
| Uses Social Media | Yes | 328 | 417 | 288 | 114 | **1147** |
| | No | 67 | 125 | 265 | 242 | **699** |
| | **Total** | **395** | **542** | **553** | **356** | **1846** |

34) Find each percent. 34) _____

a) What percent of adults surveyed are social media users aged 30-49?

b) What percent of the social media users are aged 30-49?

c) What percent of adults aged 30-49 are social media users?

Answer: a) 22.6%
b) 36.4%
c) 76.9%

Explanation:

35) Dimes minted in the United States average 2.286 g with a standard deviation of 0.06 g. A couple chemistry students were trying out their teacher's new scale by weighing a bunch of coins. The found a nickel that weighed 5.19 g and a dime that weighed 2.45 g. Which coin was more exceptionally heavy? Explain. 35) _____

Answer: The dime is more unusual. The nickel weighed a bit more than 2 standard deviations above the average weight ($z = 2.375$, while the dime was closer to 3 standard deviations more than the average for dimes ($z = 2.733$).

Explanation:

36) The five-number summary for the fuel economy (in miles per gallon) of year 2011 midsize cars is: 36) _____

| Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|
| 13 | 21 | 23 | 26 | 50 |

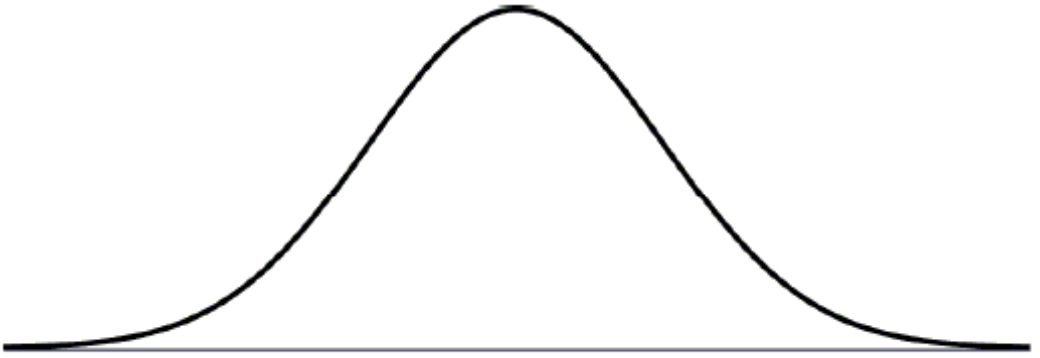
a. Would you expect the mean gas mileage of all midsize cars to be higher or lower than the median? Explain.

b. One model of Volkswagon gets 34 mpg, one model of Toyota gets 28 mpg, and one model of Bentley gets 13 mpg. Are any of these cars outliers? Explain.

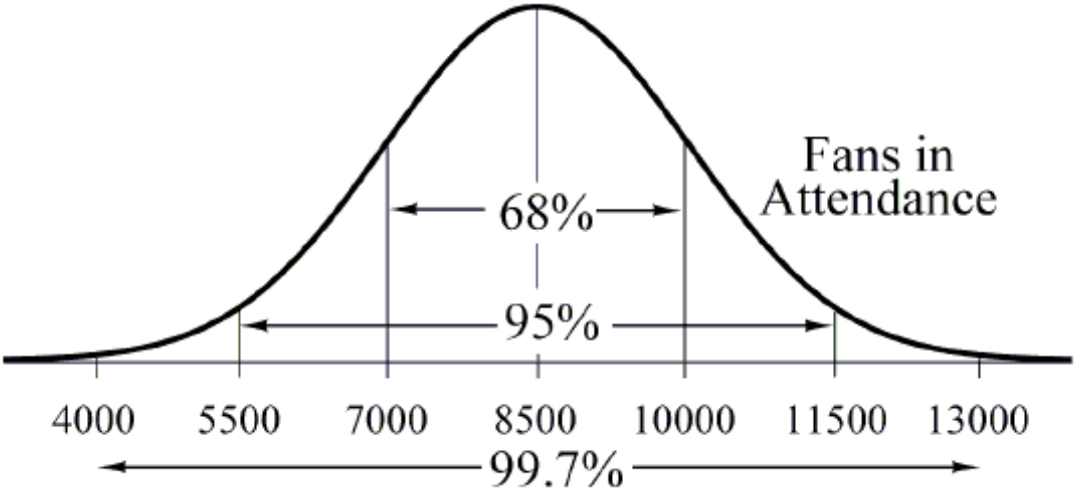
Answer: a. Probably higher. The data appear to be skewed to the right.
b. IQR = 26 – 21 = 5. Since 1.5(IQR) = 7.5, the fences are 21 – 7.5 = 13.5 and 26 + 7.5 = 33.5. The Volkswagon with 34 mpg is more than 1.5 IQRs above Q3 and the Bentley with 13 mpg is more than 1.5 IQR below Q1, so they could both be considered outliers.

Explanation:

37) Owners of a minor league baseball team believe that a Normal model is useful in 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. 37) _____



Answer:




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) Which of the W's are unknown for these data? 38) _____

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

Explanation:

39) Consider the following part of a data set: 39) _____

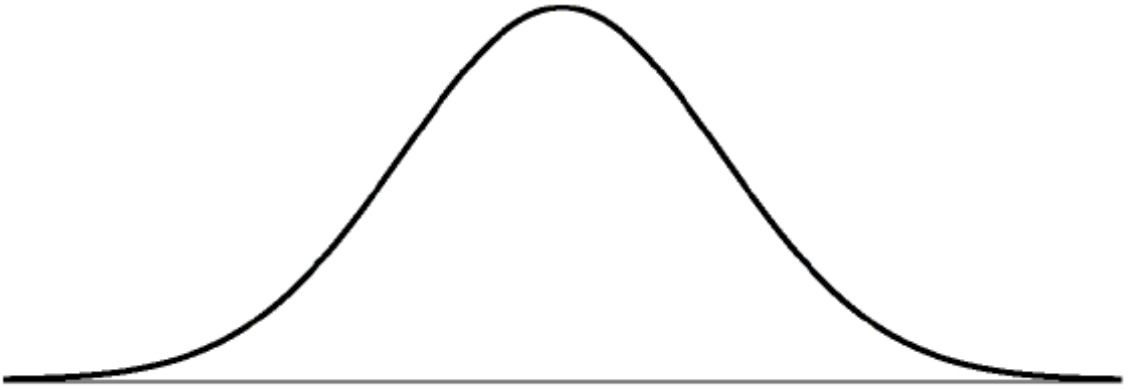
| Age (years) | Sex | Only child? | Height (inches) | Weight (pounds) | Credit Hours | GPA | Major |
|---|---|---|---|---|---|---|---|
| 21 | Female | Yes | 67.00 | 140.0 | 16 | 3.60 | animal science |
| 20 | Female | No | 62.00 | 130.0 | 18 | 3.86 | biology |
| 28 | Female | No | 64.00 | 188.0 | 21 | 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.

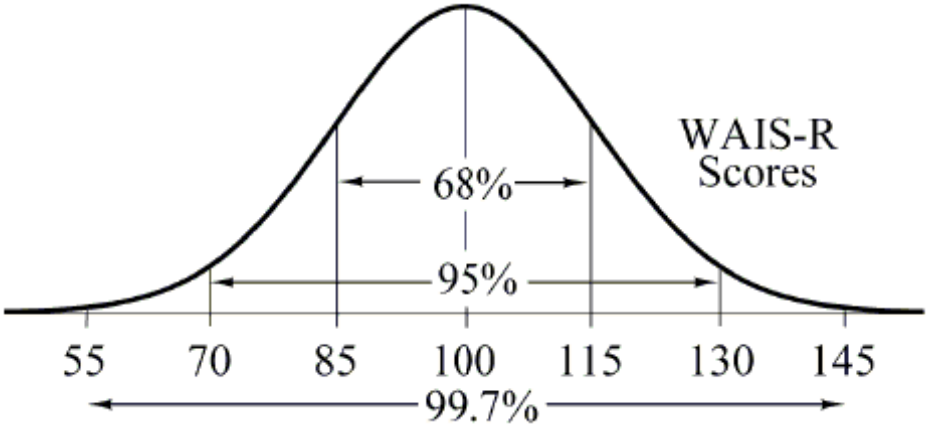
Answer: Categorical: sex, only child?, major
Quantitative: age (years), height (inches), weight (pounds), credit hours, GPA

Explanation:

40) The Wechsler Adult Intelligence Scale - Revised (WAIS-R) follow a Normal model with mean 100 and standard deviation 15. Draw and clearly label this model. 40) _____



Answer:




Explanation:

41) 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. 41) _____

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. $z = \frac{97 - 98.7}{0.7} = -2.43$, 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: $-0.67 = \frac{y - 98.7}{0.7}$ and $0.67 = \frac{y - 98.7}{0.7}$.

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, $\sigma = \frac{0.5}{1.34} = 0.37$. Our new standard deviation is 0.37°F.

Explanation:

42) Repair bills An automobile service shop reported the summary statistics shown for repair bills (in $) for their customers last month. 42) _____

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

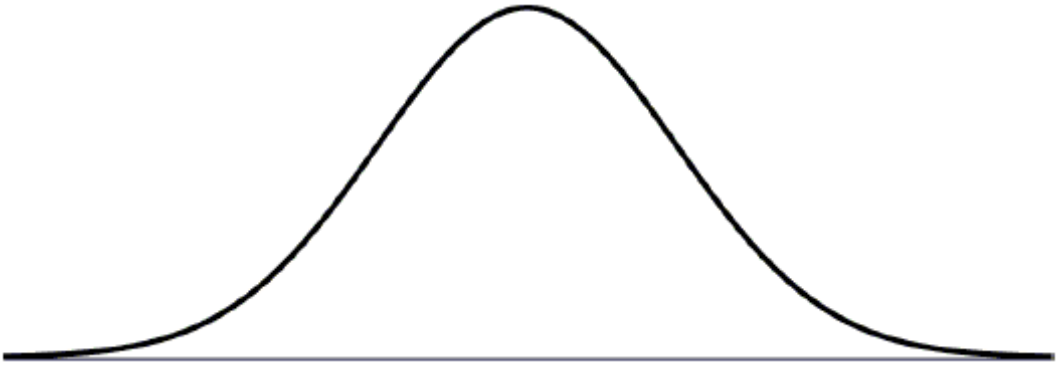
a. Were any of the bills outliers? Show how you made your decision.
b. After checking out a problem with your car the service manager gives you an estimate of "*only* $90." Is he right to imply that your bill will be unusually low? Explain briefly.

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:

43) 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. 43) _____

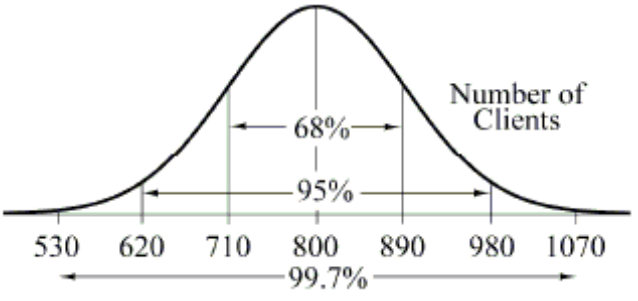
a. Draw and clearly label this model.



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 $z = -0.674$,

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

$-60.66 = x - 800$

$x = 739.34$

So the first quartile is at 740 clients.

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

1085 ($60^{th}$ percentile) has $z = +0.253$

$1085 - 450 = (0.253 + 1.645)\sigma$

$\sigma \approx 334.5$

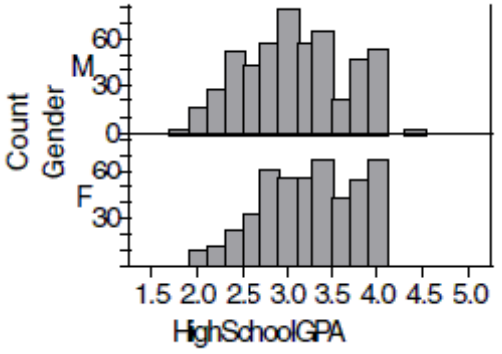
$\mu + 0.253(334.5) = 1085$

$\mu \approx 1000.4$

N(1000.4, 334.5)

Explanation:

44) One thousand students from a local university were sampled to gather information such as 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. 44) _____




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.

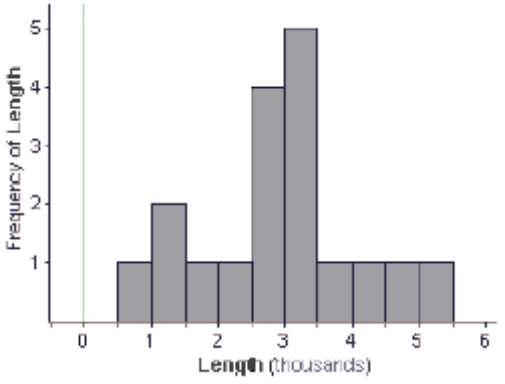
Explanation:

45) There are 18 roller coasters in Virginia for which the lengths (in feet) were reported. Those lengths are listed in the table at the right. 45) _____

| | | | | | |
|---|---|---|---|---|---|
| 1385 | 2757 | 2835 | 2231 | 3828 | 1312 |
| 3369 | 3240 | 5100 | 3157 | 2700 | 600 |

a. Sketch a histogram for these data.
b. Find the mean and standard deviation of the roller coaster lengths.
c. Is it appropriate to use the mean and standard deviation to summarize these data? Explain.
d. Describe the distribution of roller coaster lengths.

Answer: a.




b. $\bar{x}$ = 2918.3ft, $s$ = 1172.92 ft
c. Yes, the data are roughly unimodal and symmetric with no outliers.
d. The mean roller coaster length is 2918 ft, ranging from 600 ft to 5100 ft. The distribution is roughly symmetric, with typical lengths clustered between 2000 ft and 4000 ft.

Explanation:

46) Book sales A publishing company pays its sales staff $600 a week plus a commission of $0.50 per book sold. For example, a salesman who sold 440 books earned 600 + 0.50(440) = $820. 46) _____

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:

A research company frequently monitors trends in the use of social media by American Adults. The results of one survey of 1846 randomly selected adults looked at social media use versus age group. The table summarizes the survey results.

| | | Age Group | | | | |
|---|---|---|---|---|---|---|
| | | 18-29 | 30-49 | 50-64 | 65+ | **Total** |
| Uses Social Media | Yes | 328 | 417 | 288 | 114 | **1147** |
| | No | 67 | 125 | 265 | 242 | **699** |
| | **Total** | **395** | **542** | **553** | **356** | **1846** |

47) What is the marginal distribution of age groups? 47) _____

Answer: There were 395 adults aged 18-29, 542 aged 30-49, 553 aged 50-64, and 356 that were 65 or older.

Explanation:

48) Commuting to work The table shows how a company's employees commute to work. 48) _____

| | 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 a graph.)

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:

49) City planners wanted to know how many people live in a typical housing unit so they compiled data from hundreds of forms that had been submitted in various city offices. Summary statistics are shown in the table. 49) _____

| | |
|---|---|
| $\bar{x}$ | 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

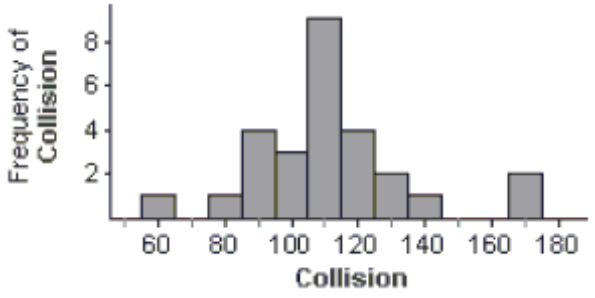
d. Mean = 40 + 5(2.53) = $52.65; SD = 5(1.50) = $7.50

Explanation:

50) Auto insurance The Insurance Institute for Highway Safety publishes ratings for all models of vehicles to compare the relative risk of payouts. 100 is the mean rating for all vehicles. A rating of 122 means the vehicle is 22% worse than average. The table shows the summary statistics for the collision ratings of 27 midsize cars. 50) _____

| | |
|---|---|
| Min | 57 |
| Q1 | 99 |
| Median | 109 |
| Q3 | 122 |
| Max | 173 |
| Mean | 110.9 |
| SD | 23.99 |

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




Answer: a. Yes. IQR = 122 – 99 = 23. The upper fence for outliers is one and a half IQR's above the third quartile, or 122 + 1.5(23) = 156.5, and the lower fence is 99 – 1.5(23) = 64.5. The maximum of 173 and the minimum of 57 are both outside these fences, so there is at least one outlier on each end.
b. The distribution is symmetric, but the existence of outliers means the median and IQR are a better choice.

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** |

51) Identify the variables and tell whether each is categorical or quantitative. 51) _____

Answer: Gender and Breed; both categorical.
Explanation:

52) 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: 52) _____

| 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:

53) Soft drinks A restaurant owner wanted to improve the efficiency of his employees. One way he tried to do this was to buy a machine that will automatically dispense 16 oz. of soda into a glass rather than have the employee hold the button on the dispenser. The actual amount dispensed by the machine can be represented by the model $N(16.2, 0.3)$ 53) _____

a. Draw and clearly label the model.

b. The sales representative who sold him the machine said, "95% of the glasses you fill with soda will fall between _____ and _____." Fill in the blanks based on the normal model, then comment on this claim.

c. What is the 3rd quartile of amounts dispensed?

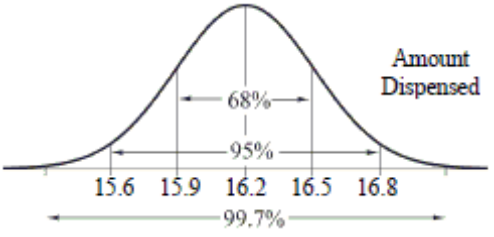
d. If a glass will actually hold 16.7 oz. of soda, what percent of the time would you expect the glass to overflow?

e. The manufacturer wants to reduce the overflow rate to only 1%. Assuming the mean amount dispensed 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 98% of their glasses be safe from overflowing, but 70% will have more than 16 oz., reducing customer complaints. What Normal model parameters is that manufacturer claiming? Show your work. N( ___, ___ )

Answer: a.




b. 15.6, 16.8. The claim is a bit too strong. This model should provide a useful estimate of what might happen, but is not certain to predict what actually will happen.

c. 16.4 oz.

d. 14.7%

e. 0.214 (should include sketches of labeled curves.)

f. A smaller standard deviation means that the machine would be more consistent with the amount it dispenses.

g. For 16 oz. $z = -0.253$ and for 16.7 oz. $z = 0.524$. Thus the difference of 0.7 oz. is $0.524 - -0.253 = 0.777$ standard deviations. The model is $N(16.23, 0.901)$

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.

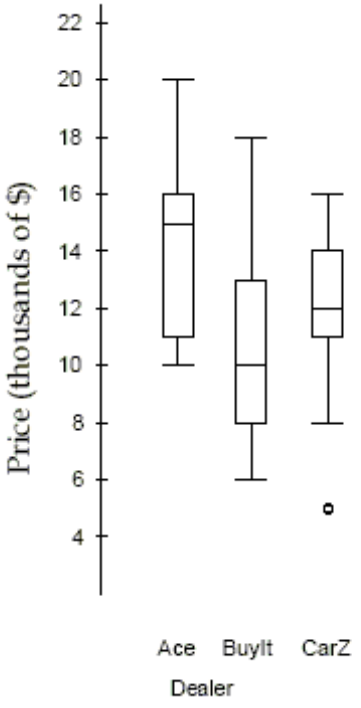
| | Male | Female | Total |
|---|---|---|---|
| **Bus** | 30 | 34 | **64** |
| **Ride** | 37 | 45 | **82** |
| **Personal** | 19 | 23 | **42** |
| **Total** | **86** | **102** | **188** |

54) Do you think mode of transportation is independent of gender? Give statistical evidence to support your conclusion. 54) _____

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:

55) The boxplots show prices of used cars (in thousands of dollars) advertised for sale at three different car dealers. 55) _____




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:

In July 2013, the Federal Drug Administration approved a new version of a drug used to treat opium dependence. The old version of the drug had received complaints about a bitter taste, an aftertaste, and that it took a long time to dissolve. The goal of the new version was to get more patients to take the drug as prescribed by addressing these issues. In addition to these improvements, experimenters monitored the existence and types of side effects of the drug.

56) Describe the W's, if the information is given: 56) _____

· Who:
· What:
· When:
· Where:
· How:
· Why:

Answer: · Who: Patients being treated for opium addiction
· What: Taste, aftertaste, time to dissolve, patient compliance, types of side effects
· When: Prior to July 2013
· Where: United States
· How: Clinical trials
· Why: To determine whether patients would comply with treatment better with the new version of the drug.

Explanation:

57) Cellphones *ConsumerReports.org* evaluated the price and performance of 99 models of cellphones. Computer output gives these summaries for the prices: 57) _____

| Min | Q1 | Median | Q3 | Max | MidRange | Mean | TrMean | SD |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 50 | 200 | 400 | 200 | 96.36 | 90.21 | 107.23 |

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

Answer: a. $IQR = Q_3 - Q_1 = 200 - 0 = 200$
$1.5(IQR) = 1.5(200) = 300$
$Q_3 + 1.5(IQR) = 200 + 300 = 500$; $Max(400) < Q_3 + 1.5(IQR)$, so no high outliers.
$Q_1 - 1.5(IQR) = 0 - 300 = -300$; $Min(0) > Q_1 - 1.5(IQR)$, so no low outliers.
b. At least 25% of the phones had price of $0. $31.95 is well above that, so no, the advertised price would not be a very low price. (Or: The advertised price is much less than one standard deviation below the mean. This is not an unusually low price.)

Explanation:

58) At a large business, employees must report to work at 7:30 A.M. The arrival times of employees can be described by a Normal model with mean of 7:22 A.M. and a standard deviation of four minutes. 58) _____

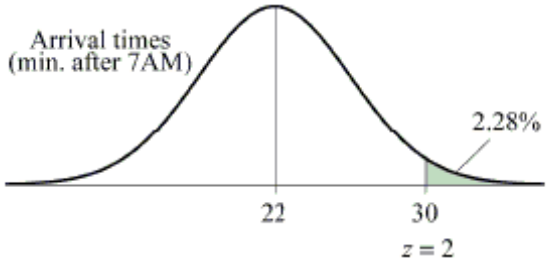
a. What percent of employees are late on a typical work day?
b. A psychological study determined that the typical worker needs five minutes to adjust to their surroundings before beginning their duties. What percent of this business' employees arrive early enough to make this adjustment?
c. Because late 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.

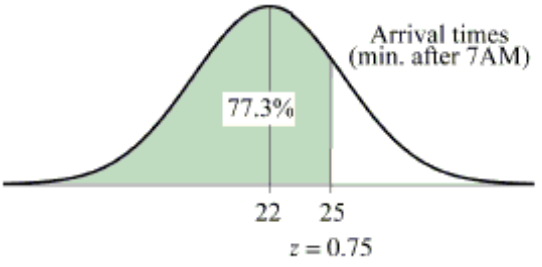
$$P(\text{time} > 7{:}30) = P(z > \frac{30 - 22}{4} = 2) = 0.0228$$

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




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

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

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

Explanation:

In July 2013, the Federal Drug Administration approved a new version of a drug used to treat opium dependence. The old version of the drug had received complaints about a bitter taste, an aftertaste, and that it took a long time to dissolve. The goal of the new version was to get more patients to take the drug as prescribed by addressing these issues. In addition to these improvements, experimenters monitored the existence and types of side effects of the drug.

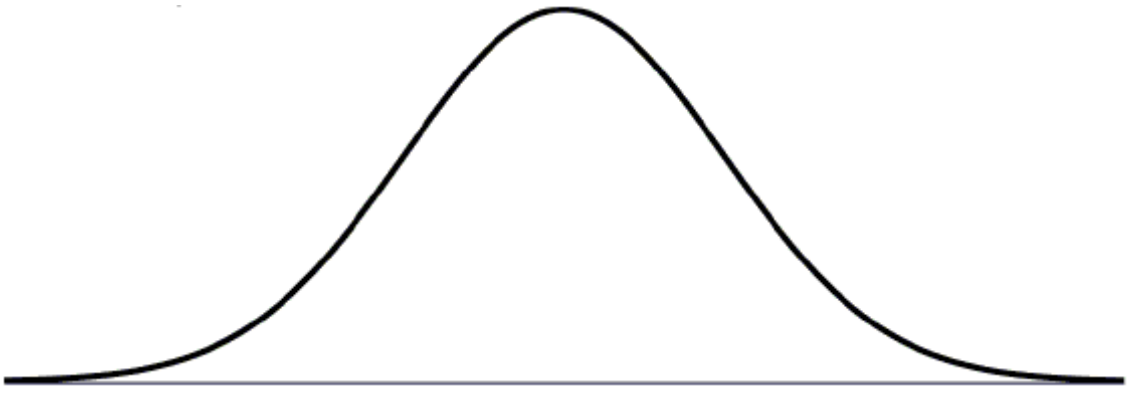
59) List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units. 59) _____

Answer: Categorical: taste, aftertaste, compliance, side effects
Quantitative: time to dissolve (minutes?)

Explanation:

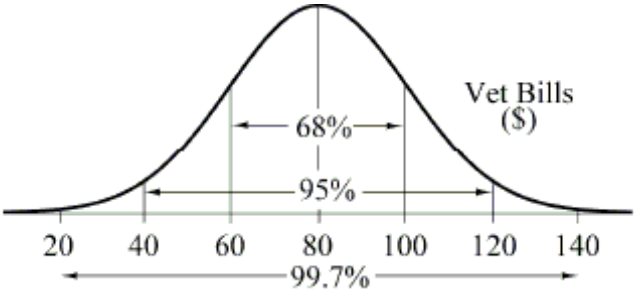
60) Veterinary costs Costs for standard veterinary services at a local animal hospital follow a Normal model with a mean of $80 and a standard deviation of $20. 60) _____

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. $z = \frac{125 - 80}{20} = 2.25$, 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

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

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

Explanation:

61) Embryonic stem cells A Pew Research survey asked Americans their feelings on medical use of embryonic stem cells. Say they surveyed 340 people and got the results summarized in the table. 61) _____

| | Morally Wrong | Not A Moral Issue | Morally Acceptable | Total |
|---|---|---|---|---|
| Conservative | 36 | 33 | 30 | 99 |
| Moderate | 29 | 62 | 58 | 149 |
| Liberal | 14 | 42 | 36 | 92 |
| Total | 79 | 137 | 124 | 340 |

a. What percent of the moderates said it is morally acceptable?
b. What is the conditional relative frequency distribution of belief for for conservatives?
c. If you wanted to show the association between political affiliation and feelings toward medical use of embryonic stem cells, what kind of graph would you make? (Just name it.)
d. Is there evidence of an association between political affiliation and feelings toward medical use of embryonic stem cells? Explain briefly.

Answer: a. 58/149 = 39%%
b. 36% morally wrong, 33% not a moral issue, 30% morally acceptable
c. segmented bar graphs, or pie charts
d. Yes. Liberals are more likely than conservatives to say morally acceptable (42% to 30%) and not a moral issue (45% to 33%). Conservatives are more likely than liberals to say morally wrong (36% to 15%). But moderates were more likely than either conservatives or liberals to say not a moral issue.

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** |

62) Identify the variables and tell whether each is categorical or quantitative. 62) _____

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

63) A company's manufacturing process uses 500 gallons of water at a time. A "scrubbing" 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.) 63) _____

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

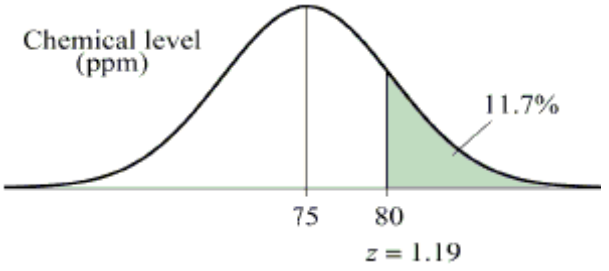
limit. To what mean value should the company set the scrubbing machine? Assume the standard deviation does not change.

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. $z = \frac{80 - 75}{4.2} = 1.19$

$P(z \geq 1.19) \approx 0.117$

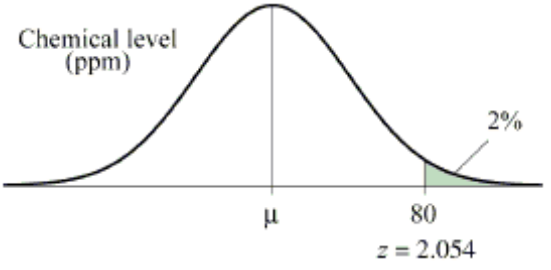



According to the normal model, we expect about 11.7% of the batches to exceed the 80ppm standard.

b. $z = 2.054$

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

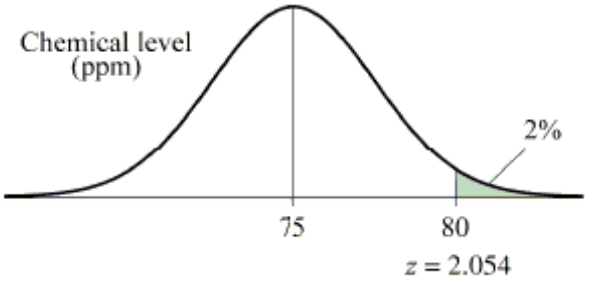
$\mu = 80 - 2.054(4.2) = 71.37$




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

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

$\sigma \approx 2.434$



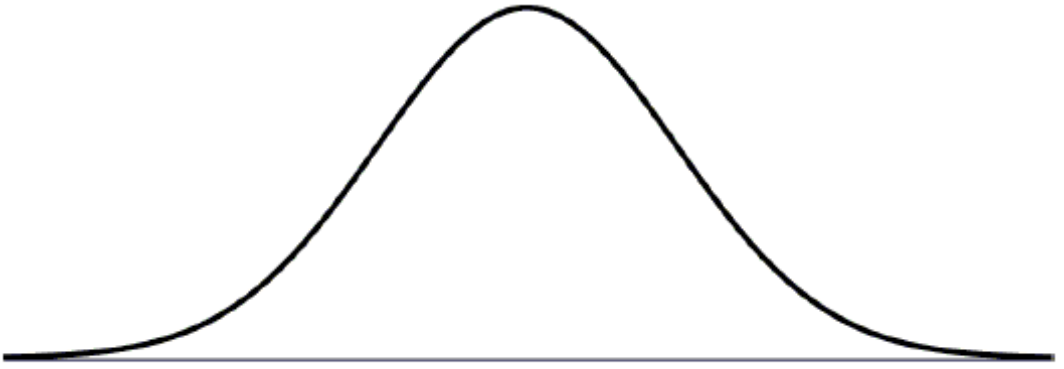

According to the Normal model, the new standard deviation would need to be at most 2.43ppm.

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

Explanation:

64) Copy machines A manufacturer claims that lifespans for their copy machines (in months) can be described by a Normal model N(42, 7). Show your work. 64) _____

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?

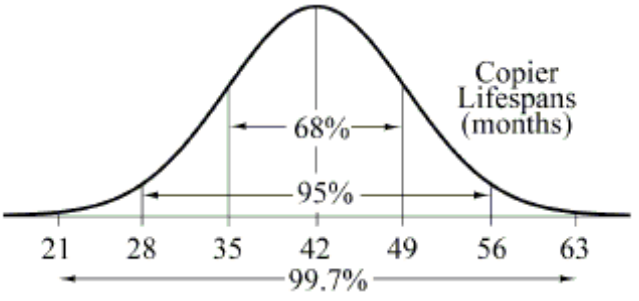
d. What percent of the copiers are expected to fail before 36 months?

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).

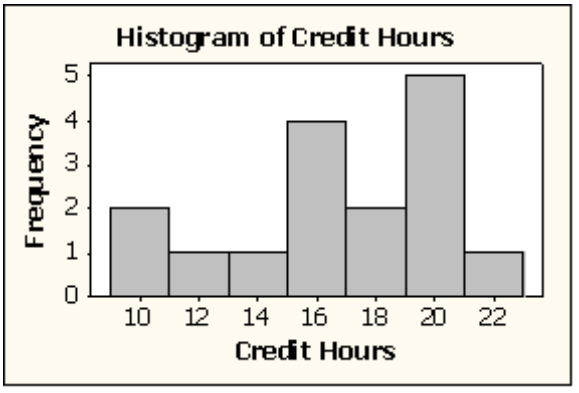
Explanation:

65) A survey conducted in a college intro stats class asked students about the number of credit hours they were taking that quarter. The number of credit hours for a random sample of 16 students is given in the table. 65) _____

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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:

A research company frequently monitors trends in the use of social media by American Adults. The results of one survey of 1846 randomly selected adults looked at social media use versus age group. The table summarizes the survey results.

| | | Age Group | | | | |
|---|---|---|---|---|---|---|
| | | 18-29 | 30-49 | 50-64 | 65+ | **Total** |
| Uses Social Media | Yes | 328 | 417 | 288 | 114 | **1147** |
| | No | 67 | 125 | 265 | 242 | **699** |
| | **Total** | **395** | **542** | **553** | **356** | **1846** |

66) Write a sentence or two about the conditional relative frequency distribution of ages of social media users. 66) _____

Answer: More social media users in the survey (36.4%) were aged 30-49 than any other age group. Next was the 18-29 age group at 28.5%, then the 50-64 group at 25.1%, and the smallest group of social media users (9.6%) was the 65 and older group.

Explanation:

67) 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. 67) _____

· 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:

68) 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. 68) _____

| | |
|---|---|
| $\overline{x}$ | 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).
$z = \frac{20 - 33.39}{5.66} = -2.366$ 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

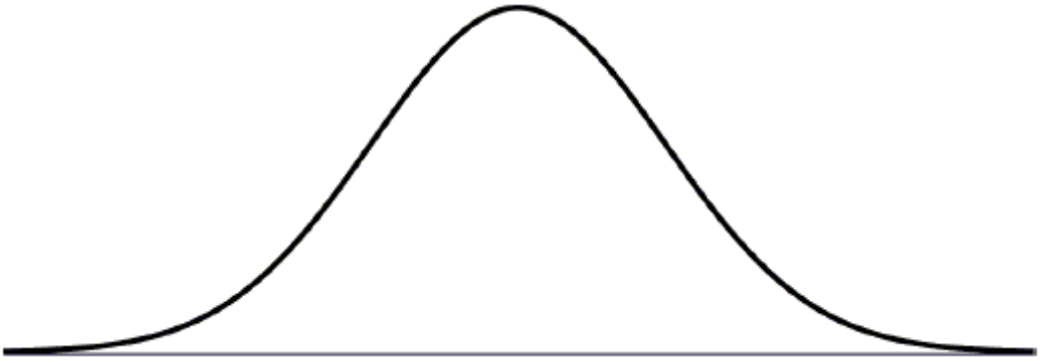
Explanation:

69) While the scales histograms are the same, there is something that could be improved so that we could compare these two distributions better. Identify this improvement and explain why it would be better. 69) _____

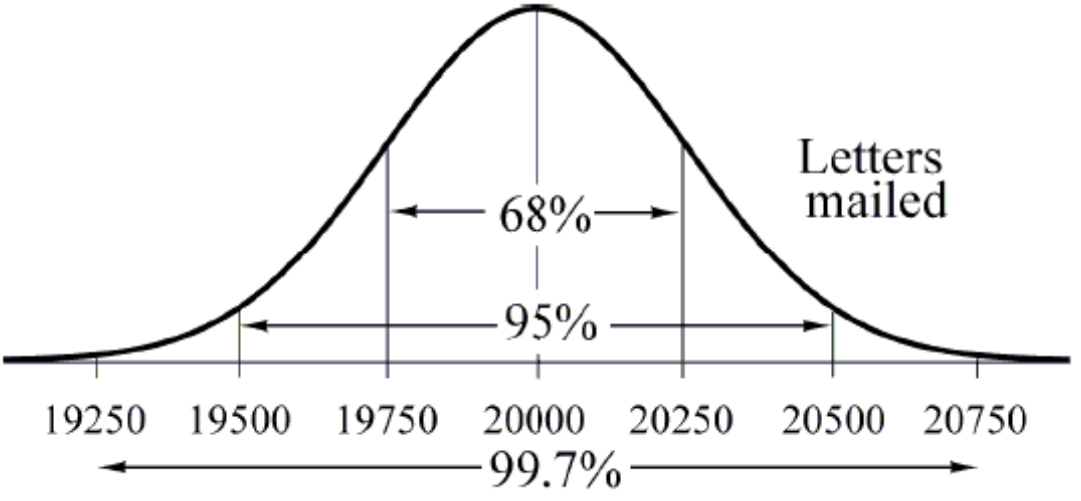
Answer: The frequency scales are not the same for the two histograms. If we converted each of the frequency histograms to a relative frequency histogram, we would be better able to compare the frequencies for each distribution at the heights.

Explanation:

70) 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. 70) _____

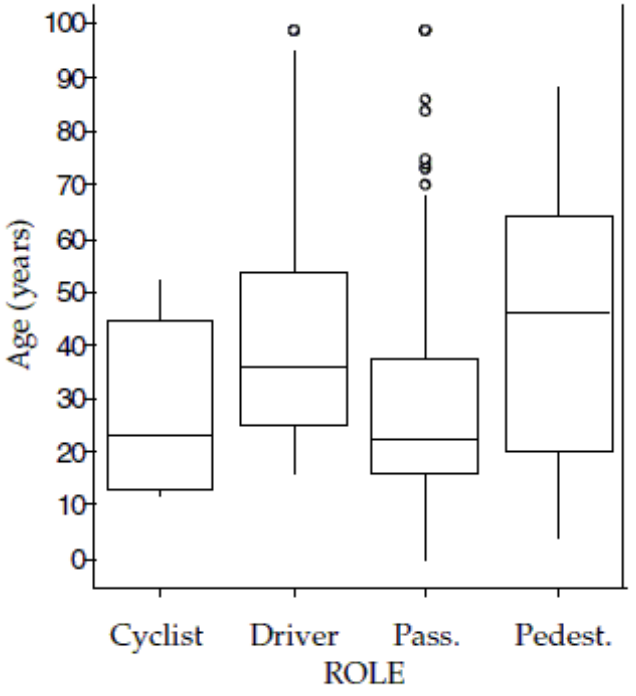


Answer:




Explanation:

71) The boxplots show the age of people involved in accidents according to their role in the accident. 71) _____




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:

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** |

72) Do you think the breed selection is independent of gender? Give statistical evidence to support your conclusion. 72) _____

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:

73) There is a proposal to replace the shortest roller coaster above with one that has a length of 1,200 ft. Indicate whether changing that roller coaster's length would make each of these summary statistics increase, decrease, or stay about the same. 73) _____

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

Answer: a. mean: Increase
b. median: Same
c. range: Increase
d. IQR: Same
e. standard deviation: Decrease

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.

74) Describe the W's, if the information is given: 74) _____

· 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:

75) Although most of us buy milk by the quart or gallon, farmers measure daily production in 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. 75) _____

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:

76) Cats and dogs The table shows whether students in an introductory statistics class like dogs and/or cats. 76) _____

| | | 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.)
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:

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.

77) List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units. 77) _____

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

Explanation:

A research company frequently monitors trends in the use of social media by American Adults. The results of one survey of 1846 randomly selected adults looked at social media use versus age group. The table summarizes the survey results.

| | | Age Group | | | | |
|---|---|---|---|---|---|---|
| | | 18-29 | 30-49 | 50-64 | 65+ | **Total** |
| Uses Social Media | Yes | 328 | 417 | 288 | 114 | **1147** |
| | No | 67 | 125 | 265 | 242 | **699** |
| | **Total** | **395** | **542** | **553** | **356** | **1846** |

78) Do you think social media use is independent of age? Give statistical evidence to support your conclusion. 78) _____

Answer: Social media use does not appear to be independent of age. Overall, 21.4% of adults surveyed were 18-29 years old, but 28.6% of social media users are in that age group. And 19.2% of adults surveyed were 65 and older, but only 9.9% of social media users were. In general, older groups seem to be underrepresented among social media users.

Explanation:

79) Which of the W's are unknown for these data? 79) _____

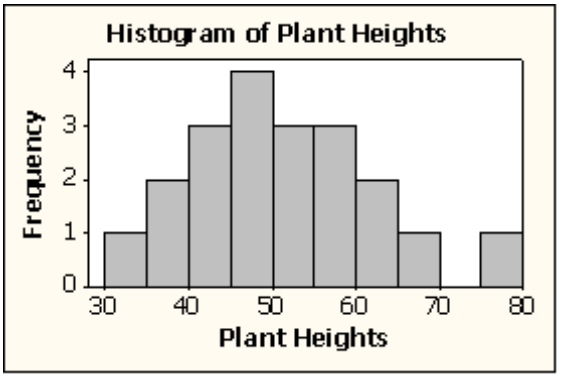
Answer: We don''t know when the adults were surveyed.
Explanation:

80) The students in a biology class kept a record of the height (in centimeters) of plants for a class experiment. 80) _____

| | | | | |
|---|---|---|---|---|
| 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 distribution 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:

81) Soda cans A machine that fills cans with soda fills according to a Normal model with mean 12.1 ounces and standard deviation 0.05 ounces. 81) _____

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. $z = \frac{12 - 12.1}{0.05} = -2.0$, 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. $-2.33 = \frac{12. - 12.1}{\sigma} \Rightarrow -2.33\sigma = -0.1 \Rightarrow \sigma = \frac{-0.1}{-2.33} = 0.043$. The standard deviation would need
to be 0.043 ounces.

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

Explanation:

82) Hardwood mulch Hardwood mulch is sold by the cubic yard. (But they just call it 'yards' to be confusing.) One lawn and garden store has a truck that they say can carry up to 7 yards for delivery. (It can actually hold a bit more.) Of course, when they fill the truck they don't get exactly 7 yards of mulch. They weigh it to determine the actual amount for billing purposes. They charge $28 per yard for the mulch, plus $25 for delivery. 82) _____
a. Shown are some summary statistics describing the distribution of the actual amounts of mulch in full loads. Fill in the table to include those statistics for the cost including delivery.

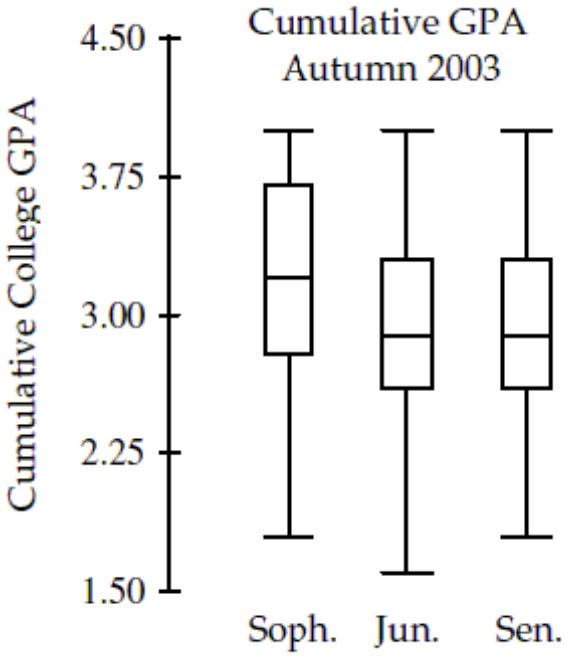
| Statistic | Yards Of Mulch | Cost incl. Delivery |
|---|---|---|
| Mean | 7.1 | |
| Standard deviation | 0.18 | |
| Q1 | 6.8 | |
| Median | 7.2 | |
| IQR | 0.5 | |

b. Your delivery is among this set of data, and it has a *z*-score of –0.84 for the distribution of yards of mulch. What is your *z*-score for the cost including delivery?

Answer: a. $223.80, $5.04, $215.40, $226.60, $14
b. $z = -0.84$

Explanation:

83) The side-by-side boxplots show the cumulative college GPAs for sophomores, juniors, and seniors taking an intro stats course in Autumn 2003. 83) _____




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

Explanation:

84) The five-number summary for the weights (in pounds) of fish caught in a bass tournament is: 84) _____

| 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.

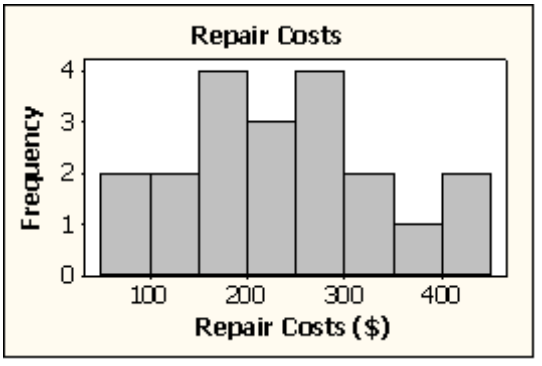
Explanation:

85) A brake and muffler shop reported the repair bills, in dollars, for their customers yesterday. 85) _____

| | | | | |
|---|---|---|---|---|
| 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 distribution 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:

86) A statistics teacher gave her class a 15 point quiz. The summary statistics for the students' scores are shown in the table at the right. 86) _____

| | |
|---|---|
| $\bar{x}$ | 10.95 points |
| $s$ | 2.481 points |
| min | 4 |
| Q1 | 9.5 |
| median | 12 |
| Q3 | 12 |
| max | 15 |

a. Notice that the median score and the third quartile are the same. Explain how this can be.

b. One student's parent heaped praise on him for scoring 13, saying it was an amazing score. Comment on whether that praise is deserved using the summary statistics as support.

c. To convert these raw scores to a score out of 100, the teacher multiplies each score by six, then adds 10. (We can debate the wisdom of such a strategy later!). What is the median converted score? And the IQR?

d. What are the mean and standard deviation of the converted test scores?

Answer: a. At least 25% of the students units had a score of 12.

b. The parent might be saying that the student did well compared to his usual performance, but as the class scores go it's not exceptionally strong. While we can say it was better than 75% of the students, we can't say anything stronger than that.

c. Median = 6(12) + 10 = 82 points; IQR = 6(12 – 9.5) = 15 points

d. Mean = 6(10.95) + 10 = 75.7 points; SD = 6(2.481) = 14.866 points

Explanation:

87) Students taking an intro stats class reported the number of credit hours that they were taking that quarter. 87) _____

Summary statistics are shown in the table.

| | |
|---|---|
| $\overline{x}$ | 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:

88) Suppose that the student taking 22 credit hours in the data set in the previous question was 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: 88) _____

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:

A research company frequently monitors trends in the use of social media by American Adults. The results of one survey of 1846 randomly selected adults looked at social media use versus age group. The table summarizes the survey results.

| | | Age Group | | | | |
|---|---|---|---|---|---|---|
| | | 18-29 | 30-49 | 50-64 | 65+ | **Total** |
| Uses Social Media | Yes | 328 | 417 | 288 | 114 | **1147** |
| | No | 67 | 125 | 265 | 242 | **699** |
| | **Total** | **395** | **542** | **553** | **356** | **1846** |

89) Identify the variables and tell whether each is categorical or quantitative. 89) _____

Answer: Age is numerical, but the grouping treats it as categorical, and social media use is categorical.

Explanation:

90) 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. 90) _____

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:

91) The five-number summary for midterm scores (number of points; the maximum possible score was 50 points) from an intro stats class is: 91) _____

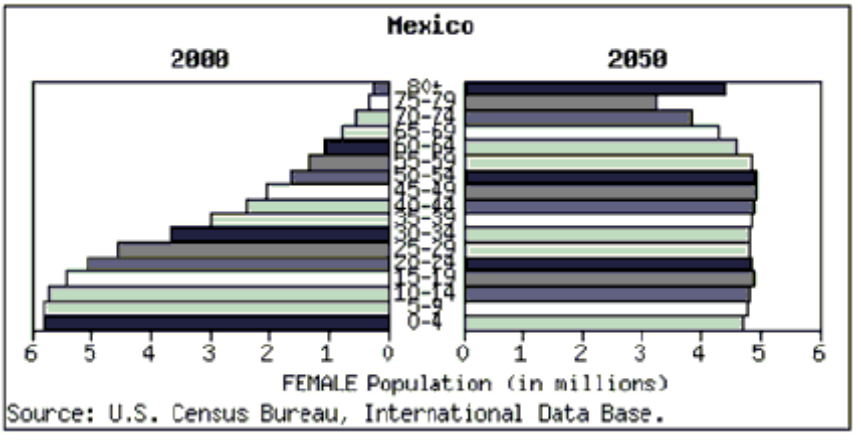
| Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|
| 16.5 | 32 | 39 | 43.5 | 48.5 |

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:

92) 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. 92) _____




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:

93) Light bulbs are measured in lumens (light output), watts (energy used), and hours (life). A 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. 93) _____

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

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:

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** |

94) What is the marginal distribution of milk consumption? 94) _____

Answer: Yes: 1,222; No: 927

Explanation:

95) Paying for purchases One day a store tracked the way shoppers paid for their purchases. Their data are summarized in the table. 95) _____

| | 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:

96) 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. 96) _____

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:

97) 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. 97) _____

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

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

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

Explanation:

98) On Monday, a class of students took a big test, and the highest score was 92. The next day, a 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: 98) _____

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:

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** |

99) Find each percent. 99) _____

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:

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** |

100) What is the marginal distribution of breeds? 100) _____

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

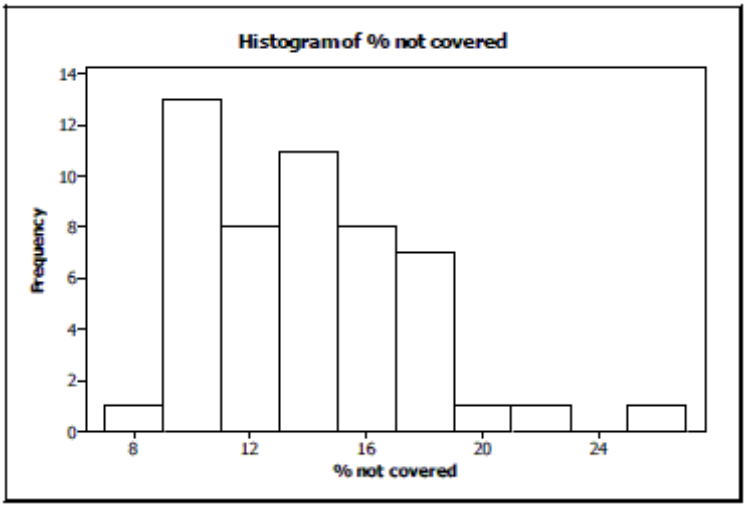
Explanation:

101) 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: 101) _____

| 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:

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.

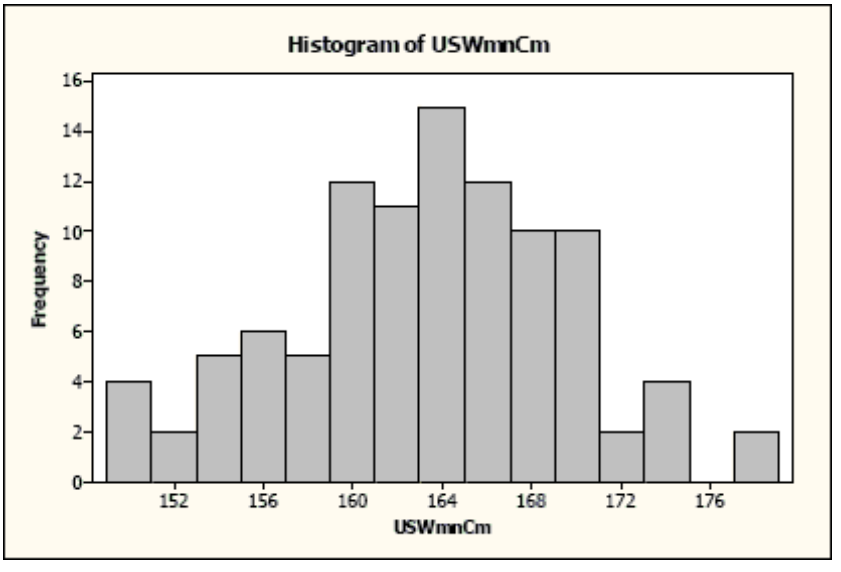
| | Female | Male | Total |
|---|---|---|---|
| **Yorkshire Terrier** | 73 | 59 | **132** |
| **Dachshund** | 49 | 47 | **96** |
| **Golden Retriever** | 58 | 33 | **91** |
| **Labrador** | 37 | 41 | **78** |
| **Dalmatian** | 45 | 28 | **73** |
| **Other breeds** | 86 | 67 | **153** |
| **Total** |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

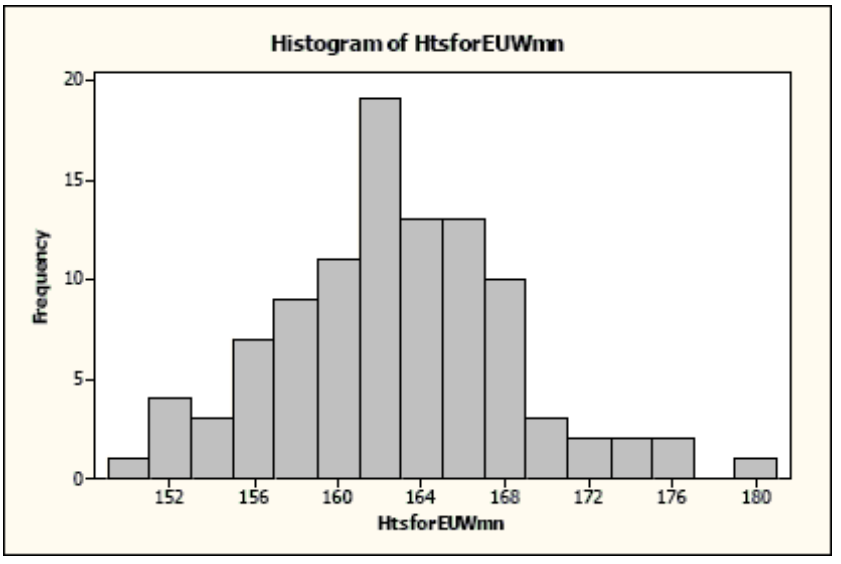
102) Write a sentence or two about the conditional relative frequency distribution of the breeds among female respondents. 102) _____

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:

103) The following are histograms for the heights of 100 US women and the heights of 100 European women: 103) _____





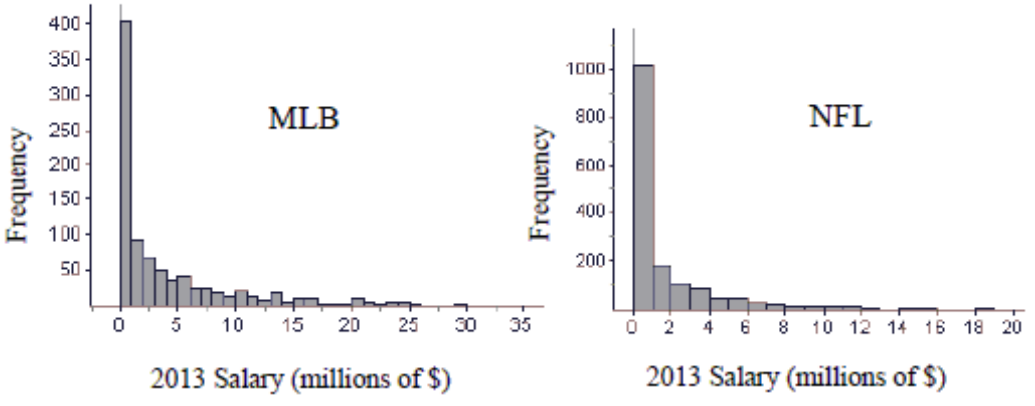

Compare the two distributions of the women's heights. Be sure to talk about shape, center, and spread.

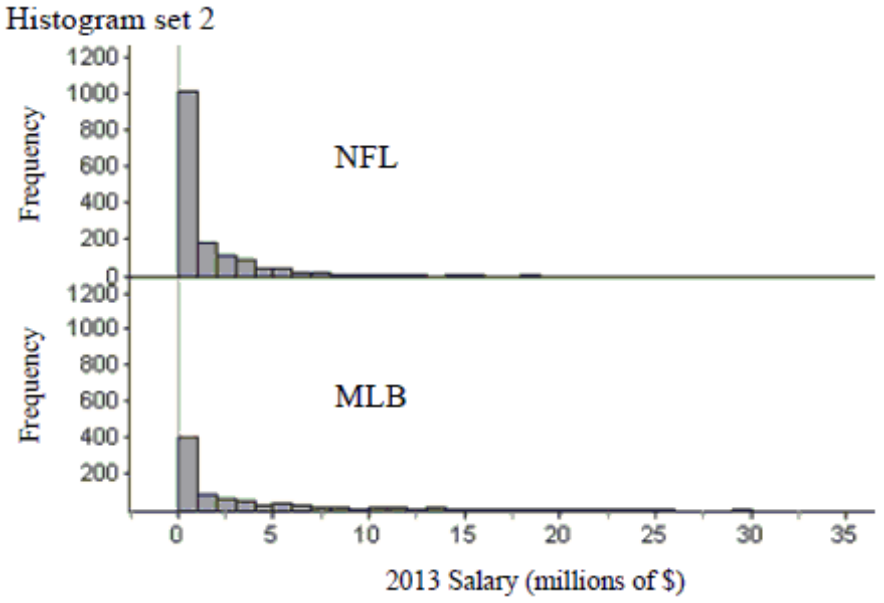
Answer: Both distributions are unimodal and roughly symmetric. Each distribution appears to be centered around 164cm. The heights for the US women appear to be more spread out than those for the European women.

Explanation:

104) How do sports salaries compare? Two sets of histograms below show the distributions of salaries for Major League Baseball and the National Football League. Give two reasons why the second set of histograms makes it easier to compare the distributions. Then write a few sentences comparing the salary distributions for the two sports. 104) _____

Histogram set 1






Answer: The second set of histograms is stacked vertically, and they use the same scales on the axes. Both distributions are strongly skewed to the right. MLB salaries are higher (the median is higher, and there are fewer low salaries and the highest salaries are much higher.) The MLB salaries also have much more variability (greater spread).

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** |

105) Find the following: 105) _____

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:

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** |

106) Which of the W's are unknown for these data? 106) _____

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

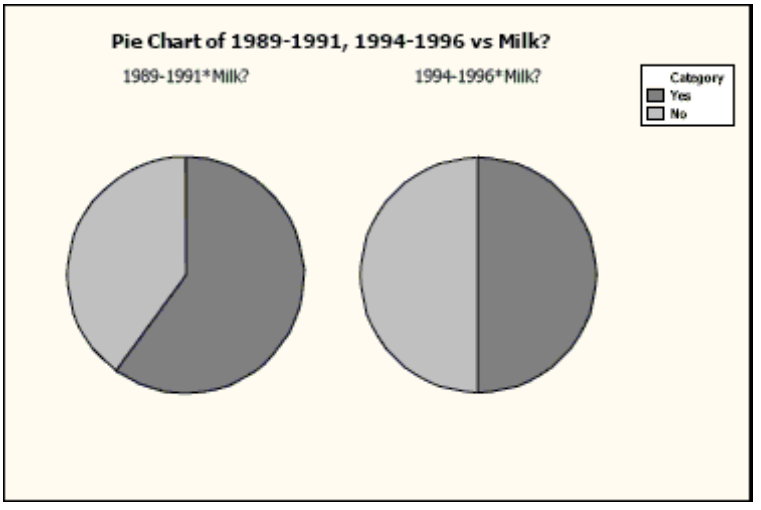
107) Write a sentence or two about the conditional relative frequency distribution of modes of transportation for the boys. 107) _____

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:

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** |

108) Consider the following pie charts of a subset of the data above: 108) _____




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:

109) Do you think that milk consumption by young girls is independent of the nationwide survey year? Use statistics to justify your reasoning. 109) _____

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:

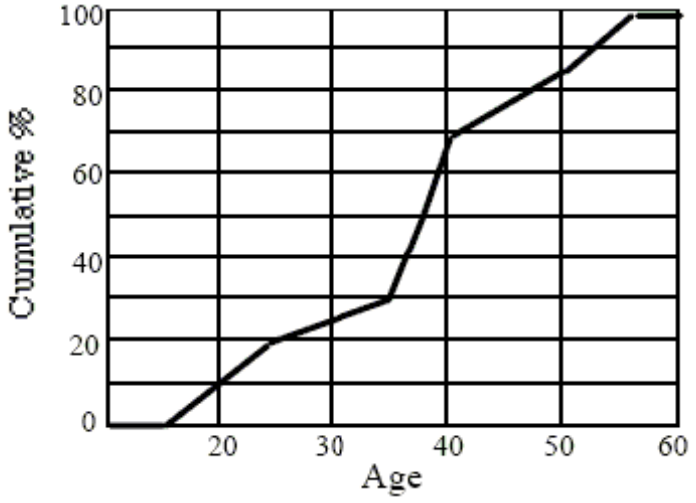
MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

110) Last weekend police ticketed 18 men whose mean speed was 72 miles per hour, and 30 women going an average of 64 mph. Overall, what was the mean speed of all the people ticketed? 110) _____

A) none of these
B) 69 mph
C) 67 mph
D) 68 mph
E) It cannot be determined.

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

111) The ages of people attending the opening show of a new movie are summarized in the ogive shown. Estimate the IQR of the ages. 111) _____




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

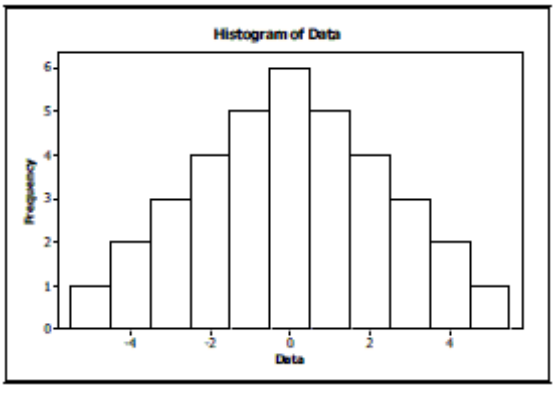
Answer: C
Explanation: A)
B)
C)
D)
E)

112) 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? 112) _____

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

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

113) Which is true of the data shown in the histogram? 113) _____




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) III only
C) I, II, and III
D) I and III
E) I only

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

114) Suppose that a Normal model describes the acidity (pH) of rainwater, and that water tested after last week's storm had a $z$-score of 1.8. This means that the acidity of that rain… 114) _____

A) had a pH 1.8 times that of average rainwater.
B) varied with a standard deviation of 1.8
C) had a pH 1.8 higher than average rainfall.
D) had a pH 1.8 standard deviations higher than that of average rainwater.
E) had a pH of 1.8.

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

115) 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? 115) _____

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

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

116) Suppose a Normal model describes the number of pages printer ink cartridges last. If we keep track of printed pages for the 47 printers at a company's office, which must be true? 116) _____

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) I, II, and III
B) I only
C) II only
D) none
E) II and III

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

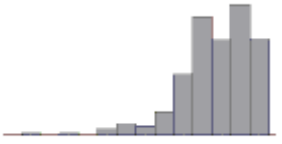
117) We might choose to display data with a stemplot rather than a boxplot because a stemplot 117) _____

I. reveals the shape of the distribution.
II. is better for large data sets.
III. displays the actual data.

A) I only
B) I, II, and III
C) III only
D) I and III
E) II only

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

118) Which is true of the data whose distribution is shown? 118) _____

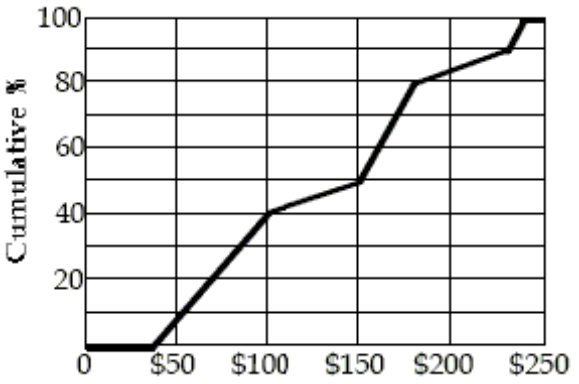


I. The distribution is skewed to the right.
II. The mean is probably smaller than the median.
III. We should summarize with median and IQR.

A) II only
B) I, II, and III
C) I only
D) II and III
E) I and II

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

119) The veterinary bills for the dogs are summarized in the ogive below. Estimate the IQR of these expenses. 119) _____




A) $200 B) $100 C) $150 D) $50 E) $75

Answer: B

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

120) Environmental researchers have collected rain acidity data for several decades. They want to see if 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)... 120) _____

A) contingency table
B) histogram
C) bar graph
D) timeplot
E) boxplot

Answer: D

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

121) Your Stats teacher tells you your test score was the 3rd quartile for the class. Which is true? 121) _____

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.

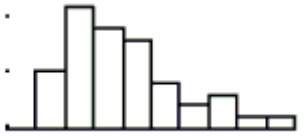
A) III only
B) I only
C) II only
D) none of these
E) II and III

Answer: D

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

122) Which is true of the data whose distribution is shown? 122) _____

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) II and III
D) I only
E) I and II

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

123) The mean number of hours worked for the 30 males was 6, and for the 20 females was 9. The overall mean number of hours worked… 123) _____

A) is 6.5
B) cannot be determined
C) is 7.2
D) none of these
E) is 7.5

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

124) Which of the following summaries are changed by adding a constant to each data value? 124) _____

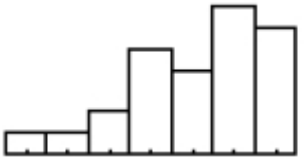
I. the mean
II. the median
III. the standard deviation

A) I and III
B) I, II, and III
C) I only
D) I and II
E) III only

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

125) Which is true of the data shown in the histogram? 125) _____

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.

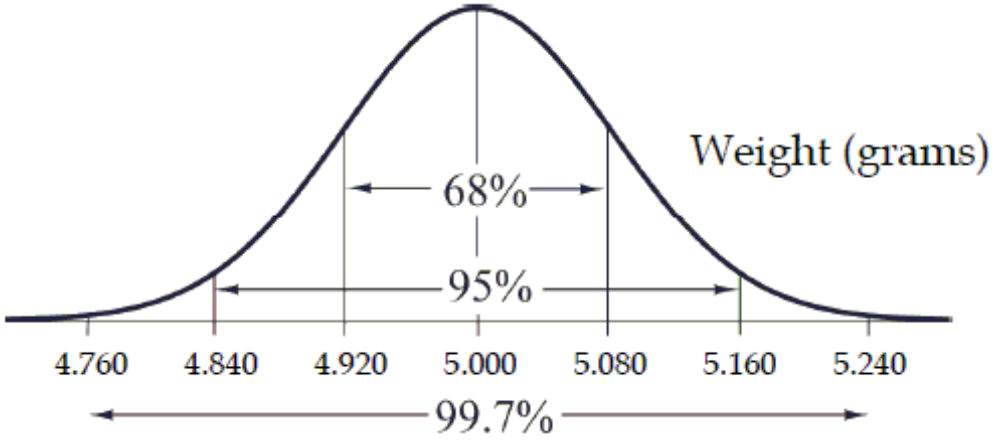


A) II and III only
B) I, II, and III
C) I only
D) II only
E) III only

Answer: A

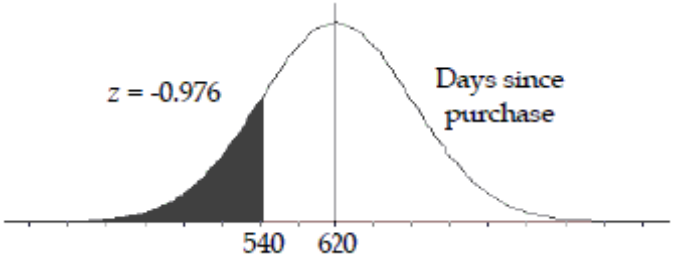
Explanation: A)
B)
C)
D)
E)

1) A
2) A
3) D
4) E
5) D
6) E
7) A
8) C
9) E
10) B
11) D
12) A
13) A
14) B
15) B
16) C
17) A
18) D
19) E
20) B
21) A
22) B
23) E
24) E
25) Categorical: species, geographic location, how new ant colonies begin, and nest design.
Quantitative: nest depth (feet), number of chambers (units), and colony size (units).
26)




27) There are 86 males and 102 females.
28) a. SUV: 32 mpg
b. SUV: 19.5 mpg
c. Pickup: 7 mpg
d. Pickup: 2 mpg
e. SUV; 3/4 of SUVs get better mileage than 3/4 of the Pickup Trucks. SUVs have the highest median, third quartile, and maximum. The first quartile is the same as that of Large cars, and the minimum is the lowest of all three, but overall SUVs appear to have the best mileage.

29) a. According to the normal model, we expect about 16.75% of the calculators to break down before 540 days.According to the normal model, we expect about 16.75% of the calculators to break down before 540 days.




b. According to the normal model, they would need to set the warranty at 429 days.


c. According to the normal model, 540 must be 2.326 standard deviations below the mean, so the standard deviation is about 34.4 days.


d. It means they would have to make the calculators more consistent in their lifespan.

30) 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 $-1.88 = \frac{23 - 26}{\sigma}$ ; then $\sigma = 1.6$ inches

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

31) · Who: SUV's currently on the market. We don't know how many models.
· What: brand of vehicle, price, type of transmission, fuel economy, acceleration, braking distance, and 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

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

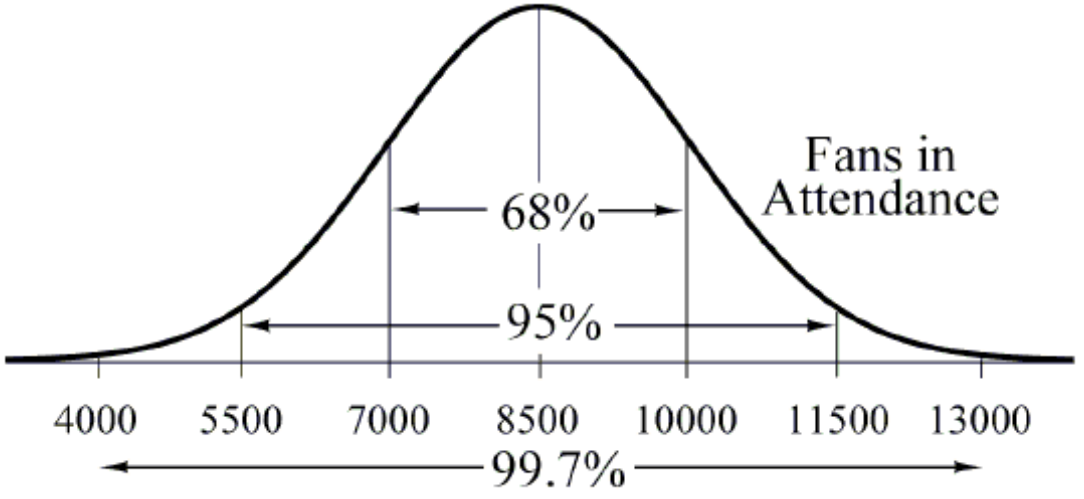
33) a. 6.6%
b. 14.9%
c. 52.6%

34) a) 22.6%
b) 36.4%
c) 76.9%

35) The dime is more unusual. The nickel weighed a bit more than 2 standard deviations above the average weight ($z$ = 2.375, while the dime was closer to 3 standard deviations more than the average for dimes ($z$ = 2.733).

36) a. Probably higher. The data appear to be skewed to the right.
b. IQR = 26 – 21 = 5. Since 1.5(IQR) = 7.5, the fences are 21 – 7.5 = 13.5 and 26 + 7.5 = 33.5. The Volkswagon with 34 mpg is more than 1.5 IQRs above Q3 and the Bentley with 13 mpg is more than 1.5 IQR below Q1, so they could both be considered outliers.

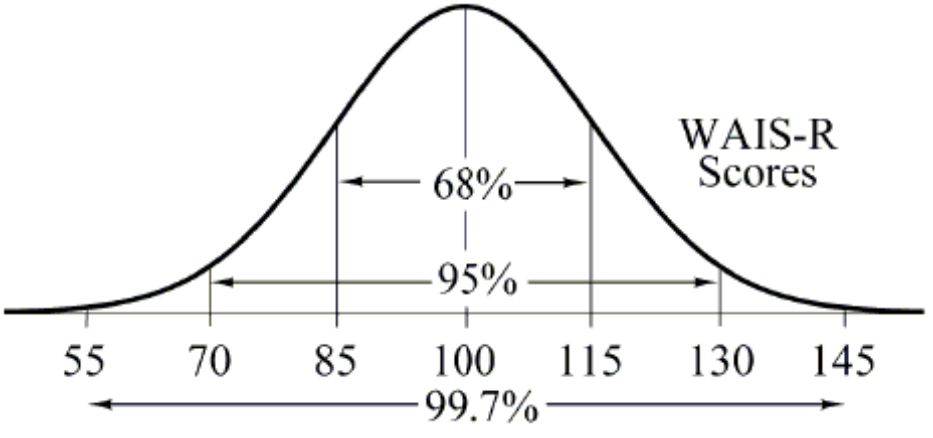
37)




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

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

40)



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

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

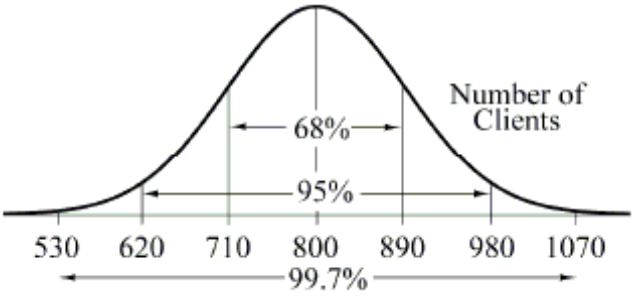
b. The $z$-scores associated with the IQR are $z = -0.67$ and $z = 0.67$. So, we need to solve for $y$ in each of the following equations: $-0.67 = \frac{y - 98.7}{0.7}$ and $0.67 = \frac{y - 98.7}{0.7}$. We get $y = 98.7 - 0.67(0.7) = 98.2$ and $y = 98.7 + 0.67(0.7) = 99.2$. The interquartile range 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, $\sigma = \frac{0.5}{1.34} = 0.37$. Our new standard deviation is 0.37°F.

42) 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.

43) a.




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

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

$-60.66 = x - 800$

$x = 739.34$

So the first quartile is at 740 clients.

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

1085 ($60^{th}$ percentile) has $z = +0.253$

$1085 - 450 = (0.253 + 1.645)\sigma$

$\sigma \approx 334.5$

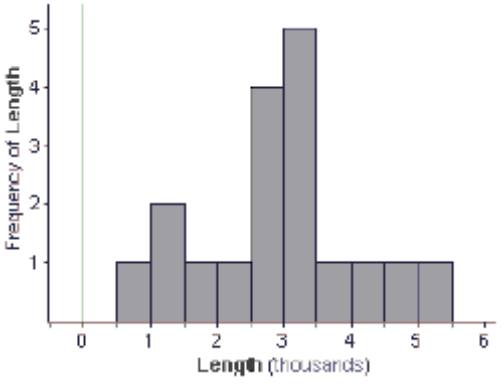
$\mu + 0.253(334.5) = 1085$

$\mu \approx 1000.4$

N(1000.4, 334.5)

44) 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.

45) a.




b. $\bar{x}$ = 2918.3ft, $s$ = 1172.92 ft

c. Yes, the data are roughly unimodal and symmetric with no outliers.

d. The mean roller coaster length is 2918 ft, ranging from 600 ft to 5100 ft. The distribution is roughly symmetric, with typical lengths clustered between 2000 ft and 4000 ft.

46) a.

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

b. +1.80

47) There were 395 adults aged 18-29, 542 aged 30-49, 553 aged 50-64, and 356 that were 65 or older.

48) 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.

49) 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

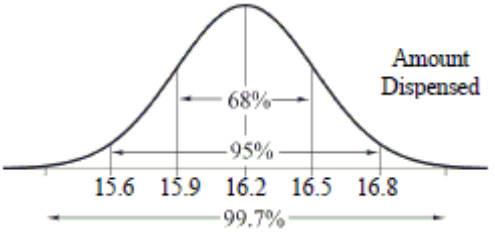
50) a. Yes. IQR = 122 – 99 = 23. The upper fence for outliers is one and a half IQR's above the third quartile, or 122 + 1.5(23) = 156.5, and the lower fence is 99 – 1.5(23) = 64.5. The maximum of 173 and the minimum of 57 are both outside these fences, so there is at least one outlier on each end.

b. The distribution is symmetric, but the existence of outliers means the median and IQR are a better choice.

51) Gender and Breed; both categorical.

52) 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.

53) a.




b. 15.6, 16.8. The claim is a bit too strong. This model should provide a useful estimate of what might happen, but is not certain to predict what actually will happen.
c. 16.4 oz.
d. 14.7%
e. 0.214 (should include sketches of labeled curves.)
f. A smaller standard deviation means that the machine would be more consistent with the amount it dispenses.
g. For 16 oz. $z = -0.253$ and for 16.7 oz. $z = 0.524$. Thus the difference of 0.7 oz. is 0.524 – -0.253 = 0.777 standard deviations. The model is $N(16.23, 0.901)$

54) 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.

55) 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.

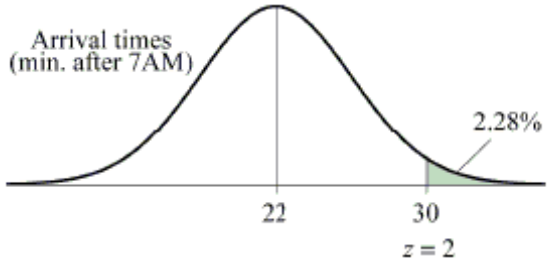
56) · Who: Patients being treated for opium addiction
· What: Taste, aftertaste, time to dissolve, patient compliance, types of side effects
· When: Prior to July 2013
· Where: United States
· How: Clinical trials
· Why: To determine whether patients would comply with treatment better with the new version of the drug.

57) a. $IQR = Q_3 - Q_1 = 200 - 0 = 200$
$1.5(IQR) = 1.5(200) = 300$
$Q_3 + 1.5(IQR) = 200 + 300 = 500$; Max(400) < $Q_3 + 1.5(IQR)$, so no high outliers.
$Q_1 - 1.5(IQR) = 0 - 300 = -300$; Min(0) > $Q_1 - 1.5(IQR)$, so no low outliers.
b. At least 25% of the phones had price of $0. $31.95 is well above that, so no, the advertised price would not be a very low price. (Or: The advertised price is much less than one standard deviation below the mean. This is not an unusually low price.)

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

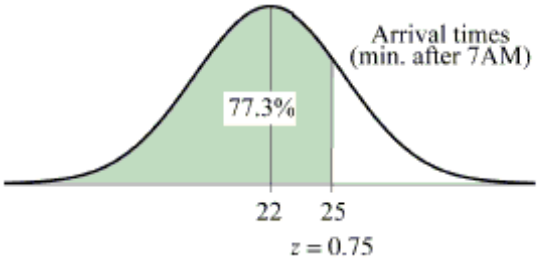
P(time > 7:30) = P($z > \frac{30 - 22}{4} = 2$) = 0.0228

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




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

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

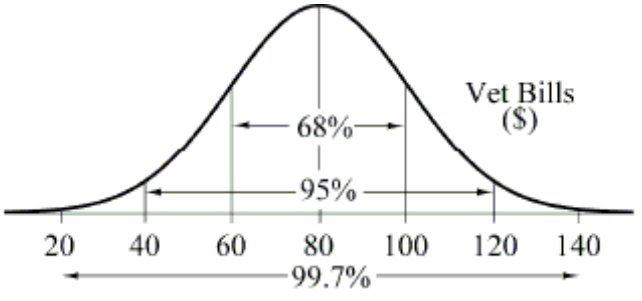



c. Virtually all times lie within 3 standard deviations of the mean. (Accept other reasonable z-scores greater than 3). If $z = 3 \le \frac{30 - 22}{\sigma}$, then $3\sigma \le 8$, so $\sigma \le 2.67$ minutes.

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

59) Categorical: taste, aftertaste, compliance, side effects
Quantitative: time to dissolve (minutes?)

60) a.




b. $z = \frac{125 - 80}{20} = 2.25$, 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

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

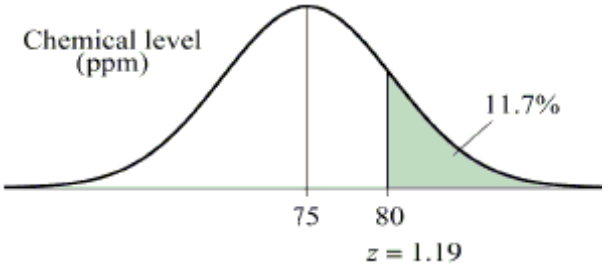
$+0.67 = \frac{y - 80}{20} \Rightarrow y = 80 + 0.67(20) = 93.4$. The IQR = Q3 - Q1 = 93.4 - 66.6 = $26.80.

61) a. 58/149 = 39%%
b. 36% morally wrong, 33% not a moral issue, 30% morally acceptable
c. segmented bar graphs, or pie charts
d. Yes. Liberals are more likely than conservatives to say morally acceptable (42% to 30%) and not a moral issue (45% to 33%). Conservatives are more likely than liberals to say morally wrong (36% to 15%). But moderates were more likely than either conservatives or liberals to say not a moral issue.

62) Gender and mode of transportation, both categorical.

63) a. $z = \frac{80 - 75}{4.2} = 1.19$

$P(z \geq 1.19) \approx 0.117$

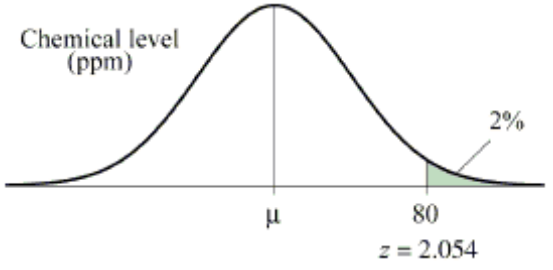



According to the normal model, we expect about 11.7% of the batches to exceed the 80ppm standard.

b. $z = 2.054$

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

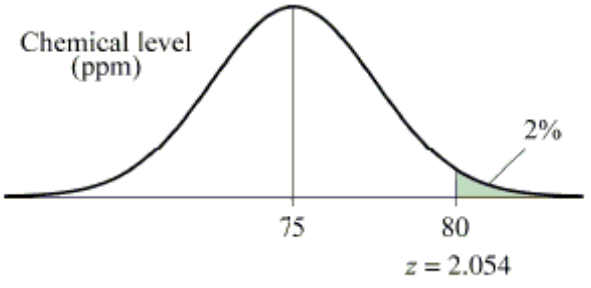
$\mu = 80 - 2.054(4.2) = 71.37$




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

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

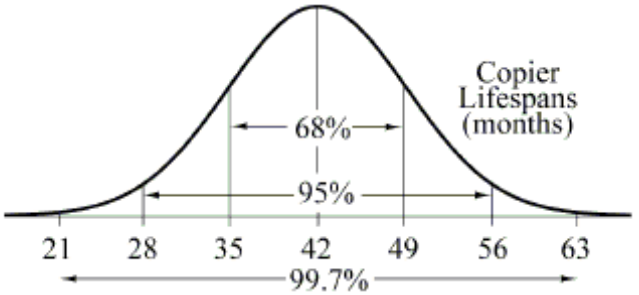
$\sigma \approx 2.434$




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.

64) 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

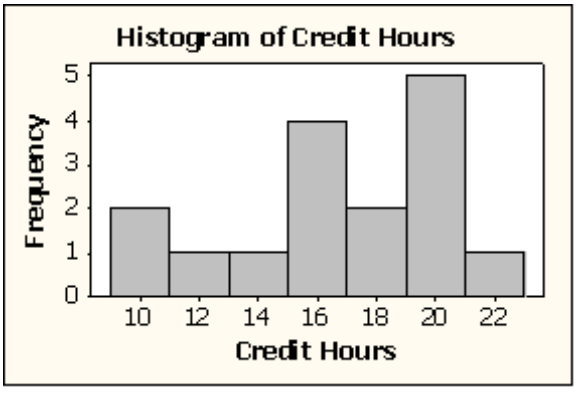
d. 19.6%

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).

65) 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.

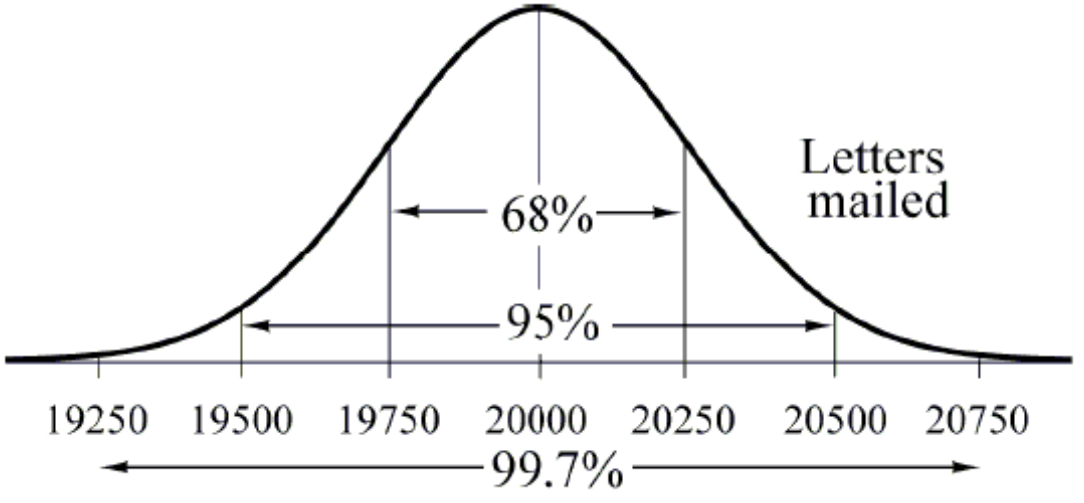
66) More social media users in the survey (36.4%) were aged 30-49 than any other age group. Next was the 18-29 age group at 28.5%, then the 50-64 group at 25.1%, and the smallest group of social media users (9.6%) was the 65 and older group.

67) · 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"

68) 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).
$z = \frac{20 - 33.39}{5.66} = -2.366$ 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

69) 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.

70)




71) 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.

72) 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.

73) a. mean: Increase
b. median: Same
c. range: Increase
d. IQR: Same
e. standard deviation: Decrease

74) · 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

75) 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 ($z = 1.875$).

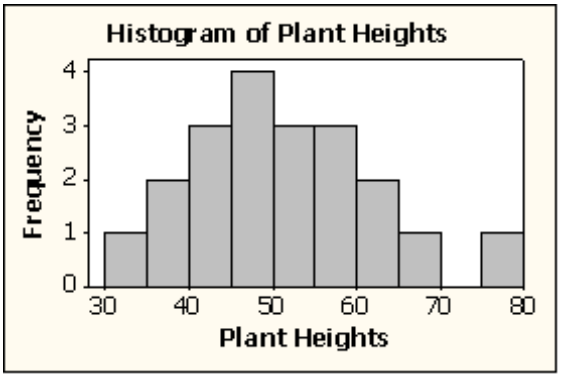
76) a. Yes: 90.7% No: 9.3%
b. Yes: 90.2% No: 9.8%
c. segmented bar graph or pie charts
d. Perhaps. There is little difference between the percents of those who like dogs, depending on whether they like cats. Of those who like cats, only 90.2% like dogs, compared to 90.7% overall.

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

78) Social media use does not appear to be independent of age. Overall, 21.4% of adults surveyed were 18-29 years old, but 28.6% of social media users are in that age group. And 19.2% of adults surveyed were 65 and older, but only 9.9% of social media users were. In general, older groups seem to be underrepresented among social media users.

79) We don''t know when the adults were surveyed.

80) 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.

81) a. $z = \frac{12 - 12.1}{0.05} = -2.0$, 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. $-2.33 = \frac{12. - 12.1}{\sigma} \Rightarrow -2.33\sigma = -0.1 \Rightarrow \sigma = \frac{-0.1}{-2.33} = 0.043$. The standard deviation would need to be 0.043 ounces.

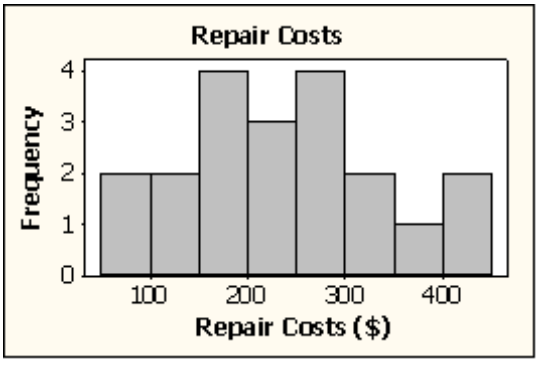
ii. $-2.33 = \frac{12. - \mu}{0.05} \Rightarrow -2.33(0.05) = 12 - \mu \Rightarrow \mu = 12 + 2.33(0.05) = 12.12$. The mean would need to be 12.12 ounces.

82) a. $223.80, $5.04, $215.40, $226.60, $14
b. $z = -0.84$

83) 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 senior class had the most symmetric set of GPAs. The sophomore class had the most skewed set of GPAs, skewed to the left.

84) 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.

85) 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.

86) a. At least 25% of the students units had a score of 12.
b. The parent might be saying that the student did well compared to his usual performance, but as the class scores go it 's not exceptionally strong. While we can say it was better than 75% of the students, we can't say anything stronger than that.
c. Median = 6(12) + 10 = 82 points; IQR = 6(12 – 9.5) = 15 points
d. Mean = 6(10.95) + 10 = 75.7 points; SD = 6(2.481) = 14.866 points

87) 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.

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

89) Age is numerical, but the grouping treats it as categorical, and social media use is categorical.

90) 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

91) 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.

92) 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.

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

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.

94) Yes: 1,222; No: 927

95) 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%).

96) a. 3400 yuans, 2000, 6200, 2400
b. $z = +1.20$

97) Dalmation: $z_D = \frac{45 - 50}{3.3} = -1.52$

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

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

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

99) a. 18.1%
b. 33.3%
c. 53.1%

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

101) 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.
102) 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.
103) 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.
104) The second set of histograms is stacked vertically, and they use the same scales on the axes. Both distributions are strongly skewed to the right. MLB salaries are higher (the median is higher, and there are fewer low salaries and the highest salaries are much higher.) The MLB salaries also have much more variability (greater spread).
105) a. 56.9%
b. 38.9%
c. 41.1%
d. 60.0%
106) We don't know how or when the students were surveyed, nor where the school is.
107) 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.
108) 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.
109) 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.
110) C
111) C
112) A
113) A
114) D
115) E
116) D
117) D
118) D
119) B
120) D
121) D
122) D
123) C
124) D
125) A