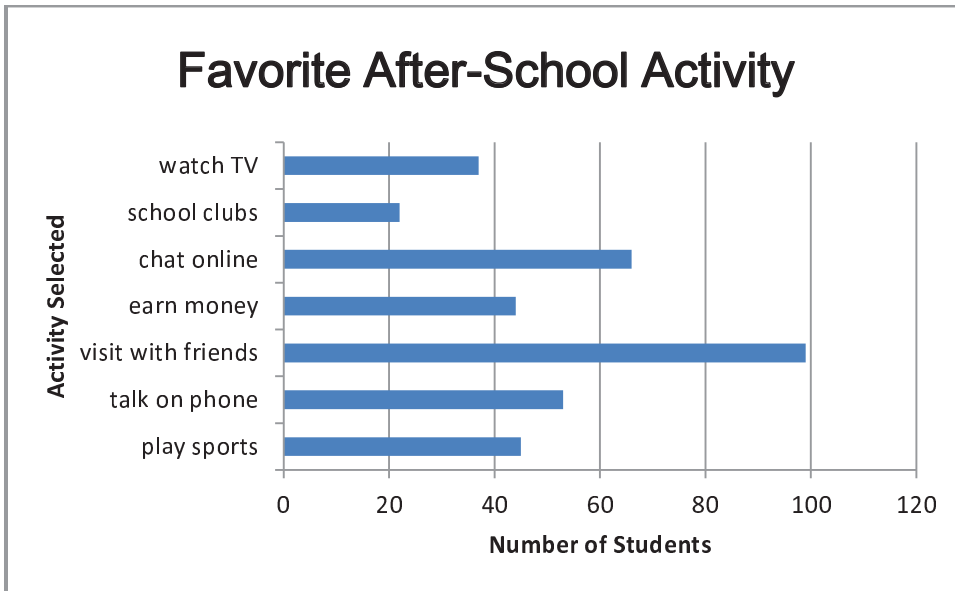


CHAPTER 5 SOLUTIONS

Section 5.1

- 1) Favorite After-School Activity?
 - a) Bar graph



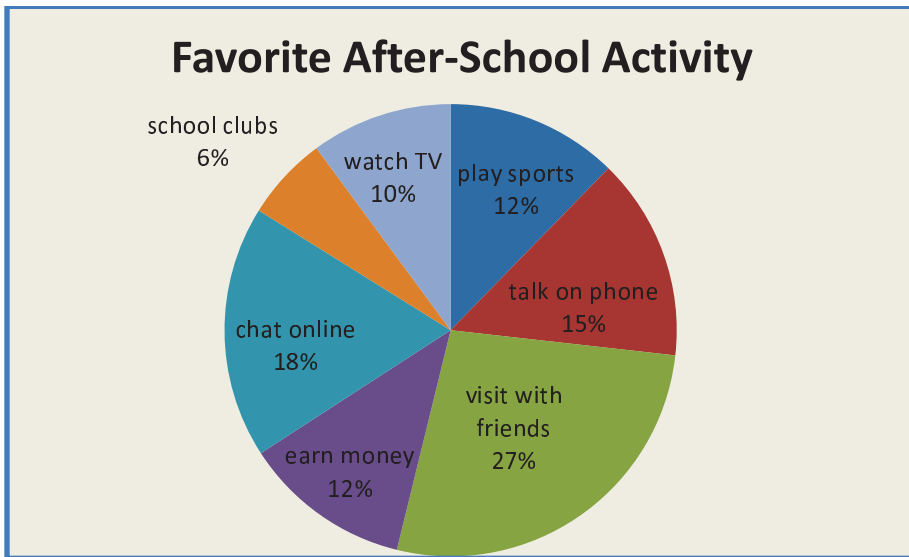
- b) Because the categories add up to the "total" and each student chose exactly one activity. So it makes sense to show this data in a pie chart as "pieces of the total pie" that are students' favorite after-school activity.

c)

Activity	Number of Students	Percent of total	Central Angle
play sports	45	12.30%	44.3°
talk on phone	53	14.50%	52.1°
visit with friends	99	27.00%	97.4°
earn money	44	12.00%	43.3°
chat online	66	18.00%	64.9°
school clubs	22	6.00%	21.6°
watch TV	37	10.00%	36.4°
	366	99.80%	360°

**note: The percents do not add to exactly 100% because of round-off error.*

d) Pie chart



2) Answers will vary. Something like:

During 2007, the majority of bear in North America were Black Bear. According to this graph, there were 800,000 Black Bear in North America. The Brown Bear population was only around 60,000 and there were fewer than 20,000 Polar Bear.

3) Type of Pet?

a-c) Answers will be different for every class based on the survey results.

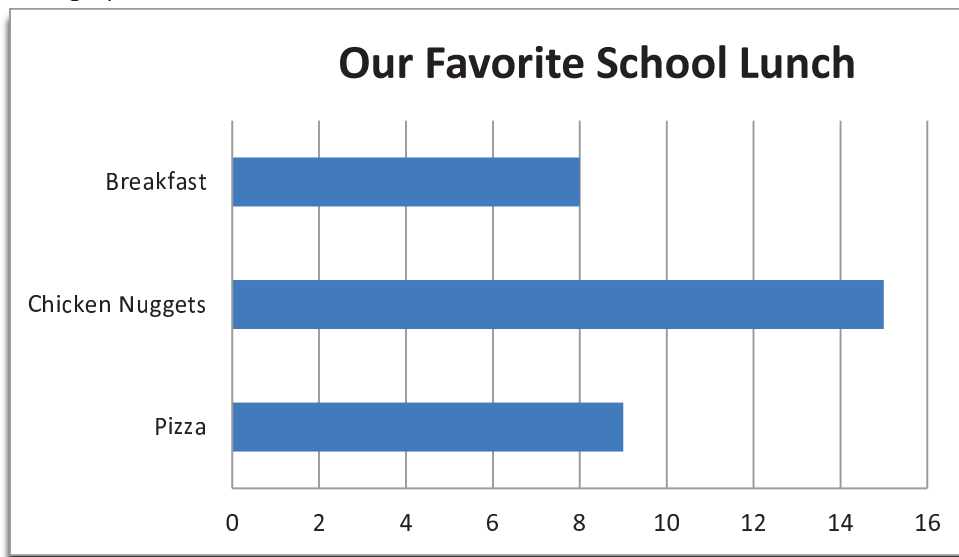
4) Favorite Season?

a-c) Answers will be different for every class based on the survey results.

5) School Lunch graph

a) This graph is misleading because the pictures are different sizes and because they are not lined up. The chicken nuggets are on two lines, so it isn't very clear at first glance that they were a much more popular choice.

b) A bar graph would be better.



- 6) Favorite Foods?
a-c) Answers will be different for every class based on the survey results.
- 7) Minnesota Wild
a) Each player
b) # = jersey number; POS = position; GP = games played; G = goals; A = assists; P = points; +/- = plus/minus; PIM = penalty minutes; PP = power play goals; SH = shorthanded goals; GW = game winning goals; S = shots; S% = shot percentage
c) # and POS are the only categorical variables. The rest are all numerical.
- 8) $2 * 2 * 2 * 2 * 2 * 4 * 4 * 4 * 4 * 4 = 32,768$ possible ways
- 9) $P(\text{all correct}) = \frac{1}{32768} = 0.0000305 = 0.00305\%$

Section 5.2

1)

	Mean	Median	Mode	range
a)	34.33	29	None	50
b)	44.57	48	22	54
c)	62.8	62	None	100

2) 171.6 pounds

3) Heights

- a) 61.2 feet
- b) 38.4 feet
- c) 4.98 feet

4) Mean = 31, median = 32, mode = none, range = 27

Either the mean or median would be an appropriate measure of central tendency. Something like:
The median number of advertisements received by each family during the given month was 32.

5) Mica's grade

- a) Mean = 63.38, med = 70.5, mode = none, range = 72. Mica will receive a D based on his mean grade.
- b) Mean = 70.88, median = 70.5, mode = 70, range = 24. Mica will receive a C- based on his mean grade. His mean, mode and range all changed. The median did not change.
- c) His mean will be 70.3, so he will have a C-.
- d) His mean will be 7.2, so he will have a C.

6) Deals on Wheels

Car Model	Retail Price	Dealer's Cost	Amt. of Mark-Up	% of Mark-Up
Nissan Sentra	\$24,500	\$18,750	5750	30.7%
Ford Fusion	26450	21300	5150	24.2%
Hyundai Elantra	22660	16600	2760	13.9%
Chevrolet Malibu	25200	22100	3100	14.0%
Pontiac Sunfire	16725	14225	2500	17.6%
Mazda 5	27600	22150	5450	24.6%
Toyota Corolla	14280	13000	1280	9.8%
Honda Accord	28500	25370	3130	12.3%
Volkswagen Jetta	29700	27350	2350	8.6%
Subaru Outback	32450	28775	3675	12.8%

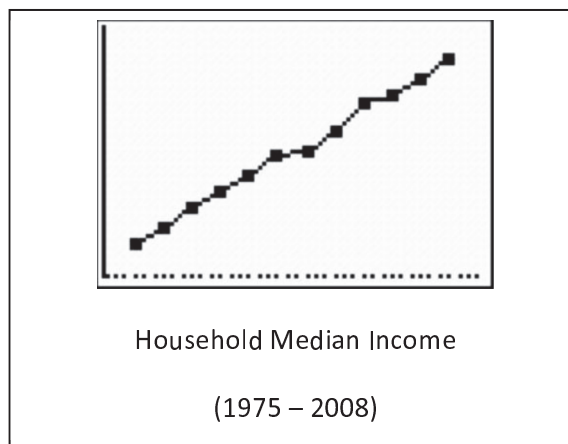
- a) See table
- b) See table
- c) Mean = 16.8% ; median = 13.9%; mode = none; range = 21.7
- d) The first three positions stayed the same, but the rest were in a different order. For example, The Subaru Outback was in fourth place based on amount of mark-up and in seventh place based on percent of mark-up.

7) Answers will vary. Something like:

The price for platinum has generally been increasing from 1960 to 2005. The price for an ounce of platinum in 2005 is approximately nine times as much as the price was in 1960. It remained fairly constant during the 60's and the 90's. There were several spikes in price, most notably around 1980. It then dropped significantly around 1982.

8) Screenshot from a TI-84+

a)



b) Answers will vary. Something like:

The graph shows that the median household income in the U.S. has been increasing at a fairly constant rate from 1975 to 2008. The median income in 2008 is more than four times that of the median household income in 1975. There are no significant spikes or drops in median household income during these 33 years.

9) Permutation because of the “specific horse” assignments.

$${}_{10}P_5 = 30,240 \text{ ways to assign the first shift}$$

10) Combination because they all going to the same activity.

$${}_{10}C_4 = 210 \text{ ways to choose a group of 4}$$

11) Combination because the order you say the toppings doesn't matter – you get the same 3 topping pizza.

$${}_{12}C_3 = 220 \text{ pizzas}$$

12) FCP because of the different categories.

$$3 * 7 * 8 = 168 \text{ outfits}$$

13) Permutation because they are getting different prizes.

$${}_{11}P_3 = 990 \text{ ways to award prizes}$$

Section 5.3

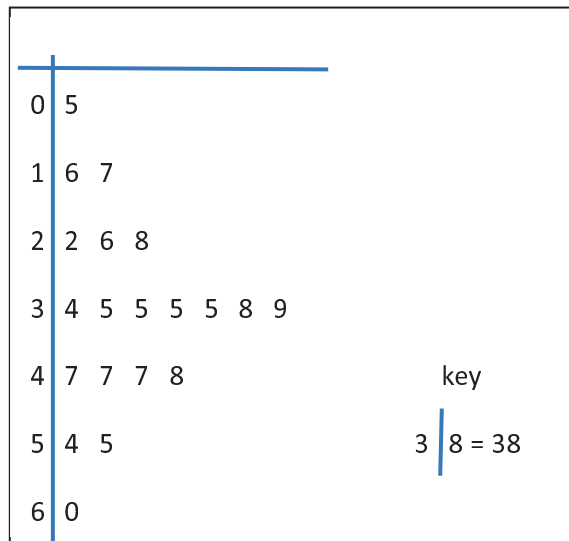
1)

a) $\bar{x} = 65$; Med = 70; mode = 70; range = 64

b) *The percentage of the paper packaging used in a country that is recycled ranges from 34% in Estonia to 98% in Japan. The distribution is basically symmetrical, but is very spread out. There appear to be a few low outliers and Japan seems to be a high outlier at 98%. The majority of countries included in this data recycle between 56% and 84% of the paper packaging. Seventy percent is both the mode and median for the percent of paper packaging that is recycled.*

2)

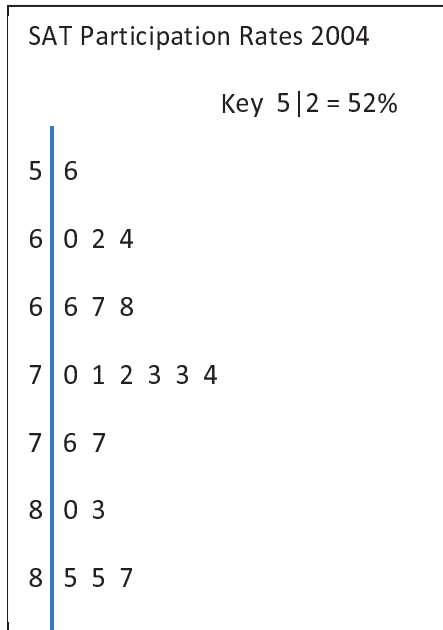
a)



b) *The number of animals treated by the local veterinarian school, over the given 20 day period, ranged from 5 to 60 animals per day. The distribution is quite symmetrical and does not include any outliers. The median number of animals treated was 35 per day.*

3) SAT

a)



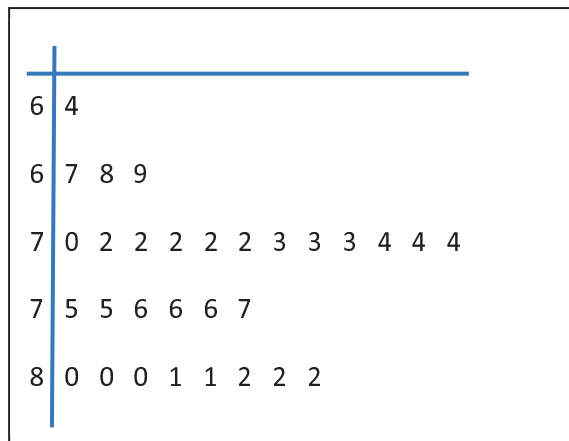
b) The median is 72.5% of students taking the SAT in 2004 for these states.

c) These are the 20 states (19 states and D.C.) with the highest rates of students participating in the SAT. Therefore, if we included the SAT participation rates from the other 31 states, our statistics will all be lower.

d) Among the 20 states with the highest rate of SAT participation, the percent of students who took the SAT during 2004 ranged from 56% in Oregon to 87% in New York. The median among these 20 states was 72.5% taking the SAT. The distribution is roughly symmetrical with no outliers. The nine locations with the highest SAT participation are all in the northeast part of the United States.

4) A

a) Split-stem



b) The data appears to be skewed to the left. There is a high concentration of values between 70 and 74.

c) Answers will vary. Something like: *This might be the ages of the members of the Lion's club.*

5) Wow So Fit

a) Mean = 75.62; Med = 77; Mode = 92

b) *The mode would not be appropriate by itself because it is not typical of the majority of ratings. Either the mean or the median would be an appropriate answer. However, the median makes it clear that half of the game critics gave a score above 77 and half gave a score below 77. So, this might give the best impression of what the average rating for this game was.*

6) Dot plots

a)

i) Skewed left, possible low outliers, quite spread out

ii) One low outlier, causing it to be skewed left. Without the outlier it would be fairly symmetrical.

iii) Very spread out, symmetrical and bimodal, no real outliers.

iv) Strongly skewed to the right, no outliers.

b) Mean or median

i) Median will be greater. Mean is pulled toward low outliers.

ii) Median will be greater. Mean is pulled toward low outlier.

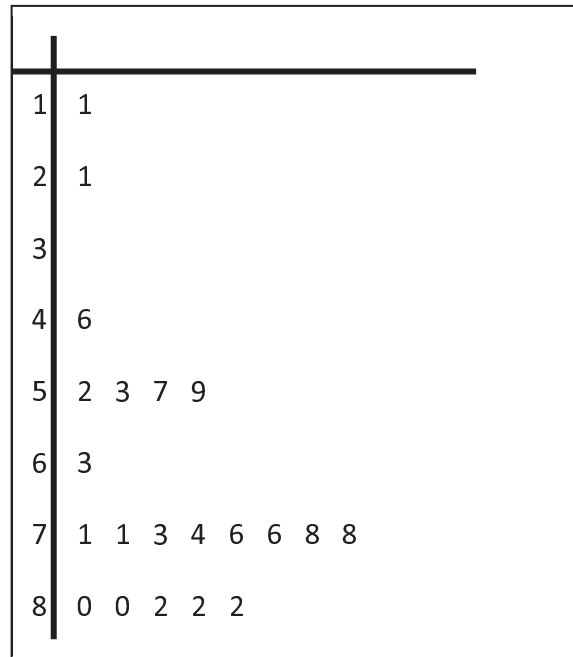
iii) Mean and median would be similar. They would both be where very little actual data values are.

iv) Mean will be greater. Mean is pulled toward tail at high end.

c) Answers will vary. Could be ages, costs, test scores, # miles on a car, # songs on an Ipod, etc.

7) Minnesota Wild

a) Stem plot



b) Mean = 65; med = 73; mode = 82; range = 71

c) The number of games played by these Minnesota Wild players during the 2010-11 regular season ranges from 11 to 82 games. There are two possible low outliers of 11 and 21 games being played during the season. The median number of games played was 73 and the majority of the players played at least 46 games. The distribution is strongly skewed to the left.

8) Minnesota Wild

a) +/- statistic. From the website: http://proicehockey.about.com/od/scoresandstat1/f/plus_minus.htm

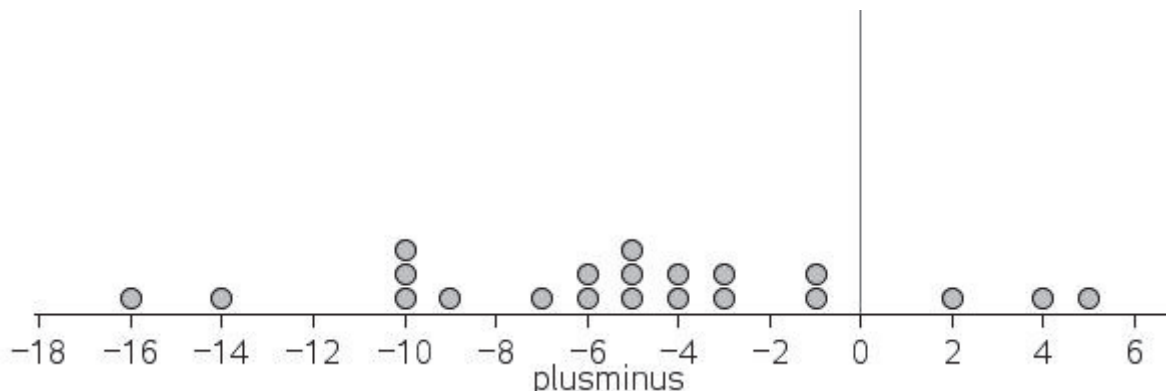
Question: What is the "plus-minus" statistic and how is it calculated?

Answer: When an even-strength or shorthanded goal is scored, every player on the ice for the team scoring the goal is credited with a "plus." Every player on the ice for the team scored against gets a "minus."

A player's overall total is calculated by subtracting the minuses from the pluses. A high plus total is taken to mean that a guy is a good defensive player. But it's a very broad measurement and there has always been disagreement over how useful a statistic it is.

- Power play goals are not used in calculating plus-minus.
- Shorthanded goals are used in calculating plus-minus.
- Penalty shot goals are not used in calculating plus-minus.

b) Dot plot



c) The distribution of the plus-minus statistics for these Minnesota Wild players is bimodal and somewhat symmetrical. The median is -5 points and there are two modes of -5 and -10 points. The plus-minus statistic ranges from -16 for Eric Nystrom to +5 for Clayton Stoner, but there are no extreme outliers.

9) The Simpsons

- a) **Population:** All Springfield residents
- b) **Parameter:** What percent enjoy watching The Simpsons?
- c) **Sample:** 1245 Springfield residents surveyed
- d) **Statistic:** $\hat{p} = \frac{1002}{1245} = 0.805 = 80.5\%$
- e) **Margin of Error:** $m. e. = \pm \frac{1}{\sqrt{1245}} = \pm 0.028 = \pm 2.8\%$
- f) **Estimated 95% CI:** $0.805 - 0.028 = 0.777$ and $0.805 + 0.028 = 833$
(0.777 to 0.833) or (77.7% to 83.3%)
- g) **Confidence Statement:** *We are 95% confident that the true percent of Springfield residents who enjoy watching The Simpsons is between 77.7% and 83.3%.*

Section 5.4

1) Salaries

- a) Approximately 50 employees
- b) Bin width is \$11,000
- c) Answers will vary-something like: *A typical employee of this district makes approximately \$50,000.*

2) Jessica's weight

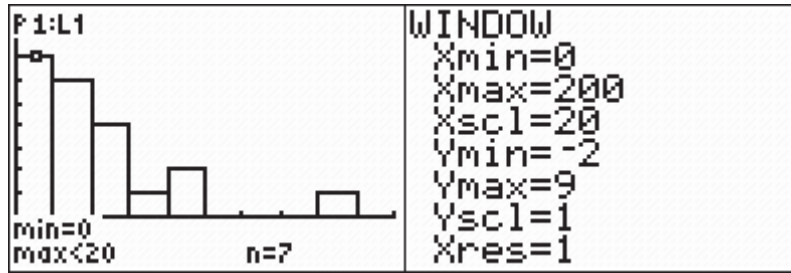
- a) *Jessica's weight has ranged from 131 pounds to 149 pounds during this time period. Her median weight was approximately 140 pounds. The graph shows a symmetrical distribution with no outliers.*
- b) $137 \leq w < 139$

- c) 76.7% of the time ($23/30 = 0.76667$)

3) Men's exercise

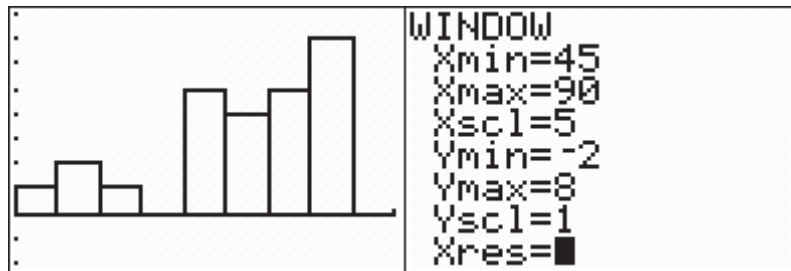
- a) The "over 75" bin is not the same width as the other bins. All others cover a range of ten years.
- b) *According to the graph, the percentage of men who spend at least one hour per week exercising decreases with age. The histogram is strongly skewed to the right and does not appear to have any outliers. Over 60% of the men between 16 and 24 years of age exercise at least one hour per week, but fewer than 5% of the men over 75 do. The biggest drop in time spent exercising happened between the 16 – 24 age group and the 25 – 3 age group.*

- 4) Wild
 a) Histogram



- b) *The number of penalty minutes served by these Minnesota Wild players is strongly skewed to the right. Brad Staubitz is an outlier with 173 penalty minutes. The PIM statistic ranged from 2 minutes to 173 minutes for these Wild players. The median is 34 minutes and the man is 44.14 minutes – the mean is higher due to the skewed shape and the extreme outlier.*

- 5) Life expectancy 2010
 a) Histogram: screenshot from a TI-84 calculator



- b) The median will be higher than the mean because of the left skew of this distribution.
- c) Range = 35 years; Mean = 71.52 years; Med = 74 years; Mode = 82 years
- d) In this case, because the distribution is skewed left, the median is the most appropriate measure of central tendency. The mean is pulled toward the low end by the four countries with very low life expectancies. And the mode is a higher number than most countries have, so it is misleading as well. The median tells us that half of these countries have life expectancies above 74 years and half are below 74 years.

- 6)
 a) Answers will vary.
 b) Answers will vary
 c) Answers will vary
 d) Answers will vary
 e) Answers will vary

7) Raffle

a) Probability distribution table

Prize Amount X	\$1000	\$250	\$50	\$25	\$0
Probability P(X)	1/500	2/500	5/500	10/500	482/500

b) Expected value before cost: $(1/500)1000 + (2/500)250 + (5/500)50 + (10/500)25 + (482/500)0 = 4$
 Expected value after cost: Expect to win \$4, but it costs \$10 per ticket, so $EV = \$4 - \$10 = -\$6$.
 The expected value per ticket is to lose \$6, on average, per ticket purchased.

c) Fair game: This is not a “fair game” because the players do not “break even in the long run”. The expected winnings per ticket are not equal to the cost to play per ticket.

8) Fish

a) $P(2 \text{ turquoise}) = 7/16 * 7/16 = 49/256 = 0.1914$

b) $P(\text{exactly one gold}) = (\text{either order}) = 4/16 * 12/16 + 12/16 * 4/16 = 2(4/16 * 12/16) = 3/8 = 0.375$

c) $P(\text{pink, then gold}) = 5/16 * 4/16 = 5/64 = 0.078$

9) More Fish

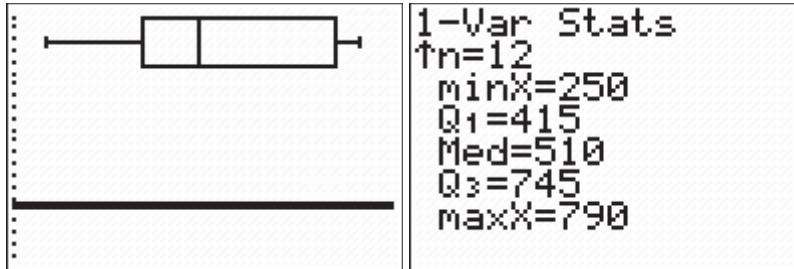
a) $P(2 \text{ pink}) = 5/16 * 4/15 = 1/12 = 0.083$

b) $P(\text{exactly one turquoise}) = (\text{either order}) = 7/16 * 9/15 + 9/16 * 7/15 = ({}_{7}C_1 * {}_{9}C_1) / ({}_{16}C_2) = 21/40 = 0.525$

c) $P(\text{no gold}) = 12/16 * 11/15 = 11/20 = 0.55$

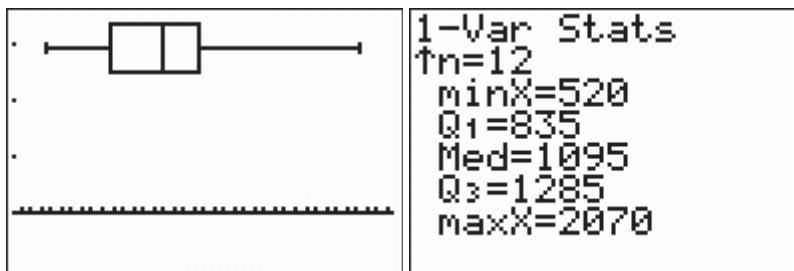
Section 5.5

- 1) McDonald's sandwiches
 - a) Calories from fat: Median = 240 cal; IQR = 160 cal
 - b) Cholesterol: Median = 82.5 mg; IQR = 67.5 mg
- 2) Calories
 - a) 5# sum = { 250, 415, 510, 745, 790} & box plot



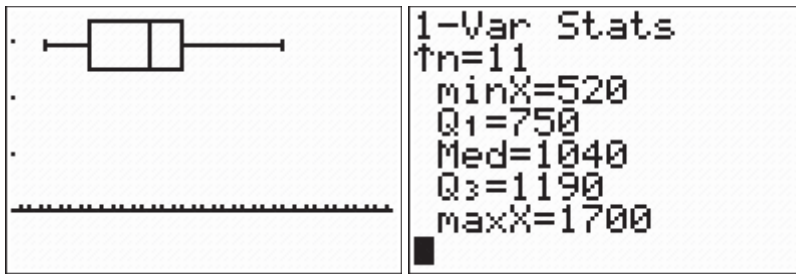
- b) Outlier test
$$415 - 1.5*(330) = -80$$
so, no low outliers
$$745 + 1.5*(330) = 1120$$
so no high outliers
 - c) *The calories in these McDonald's sandwiches range from 250 to 790 calories. The median number of calories is 510 for these sandwiches. There are no outliers and the box plot does not show any skew-ness. However, there is a big gap in the distribution because there are no sandwiches with calories between 540 and 740 calories.*

- 3) Sodium content
 - a) Box plot



- b) Median = 1095 mg; IQR = 450 mg
 - c) Mean = 1130.8333 mg ; and standard deviation = 430.8862 mg

With outlier removed:

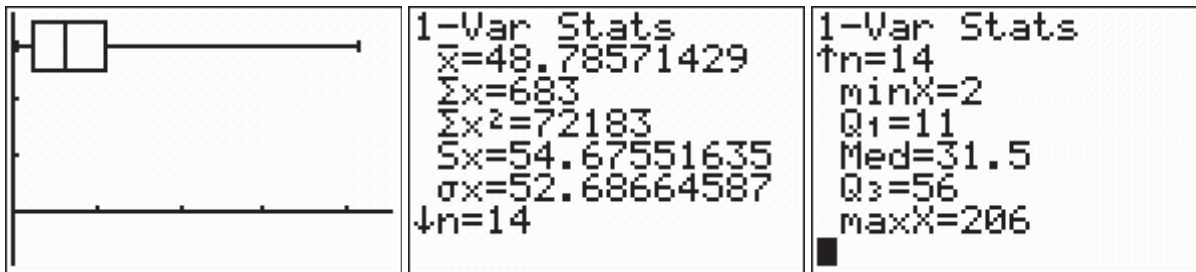


d) Median = 1040 mg ; IQR = 440 mg. They both changed, but the IQR only changed by 10 mg.

e) Mean = 1045.4545 mg ; and standard deviation = 328.6446 mg. Both changed, but the standard deviation changed by more than 100 mg.

4) Energy Saved

a) 5 # summary & box plot



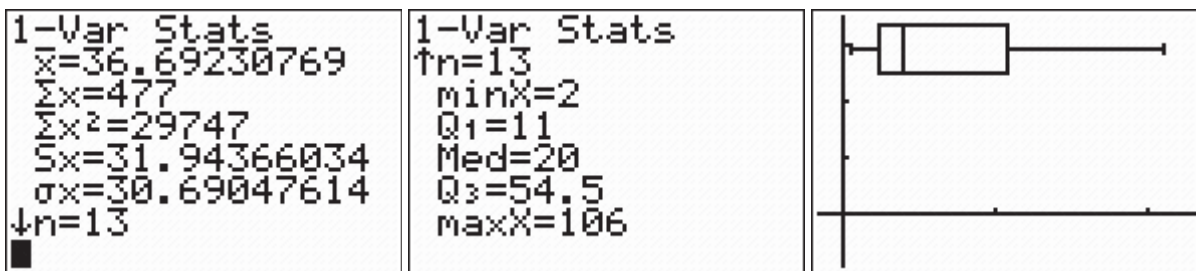
b) Outlier test

$11 - 1.5*(45) = -56.5$ So, there are no low outliers

$56 + 1.5*(45) = 123.5$ So, there is one high outlier, the aluminum cans.

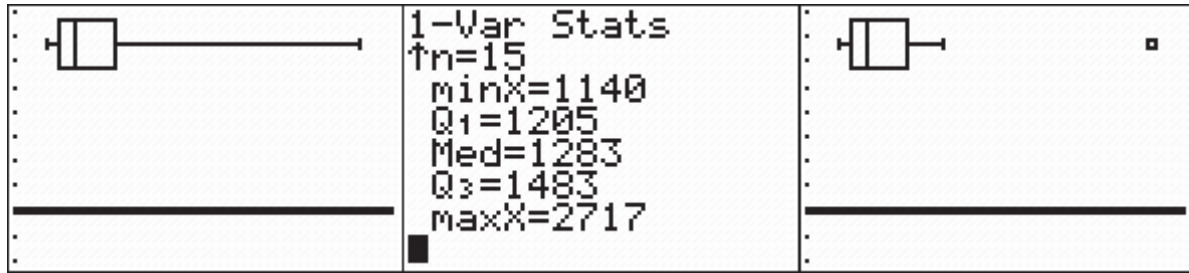
c) Mean = 48.7857 and s.d. = 54.6755. The mean is larger than the median because of the outlier.

d) Delete the aluminum cans data. The median, Q3, and max all changed. Therefore, so did the box plot. The mean and standard deviation both changed also.



5) Tallest buildings in the world

a) 5# summary and box plot & (box plot showing outliers)



b) Outlier test

$1205 - 1.5*(278) = 788$ So, there are no low outliers.

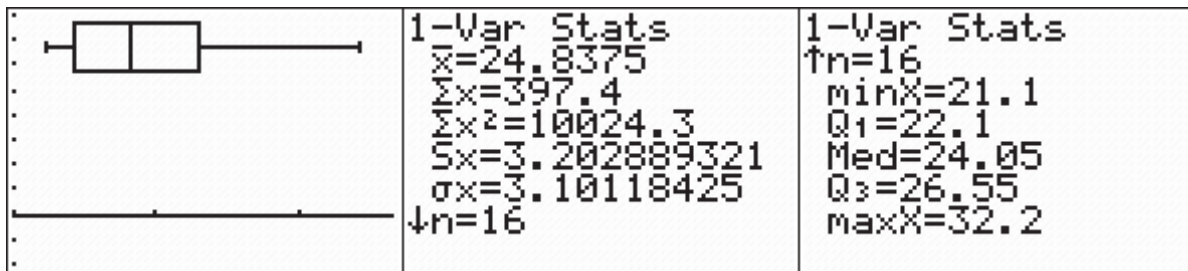
$1483 + 1.5*(278) = 1900$ So, the Burj Dubai building will be an extremely high outlier.

c) *The heights of these tall buildings range from 1140 feet to 2717 feet tall. The distribution is extremely skewed to the right. The Burj Dubai is the tallest building in the world by more than 1,000 feet taller than the next tallest building. Among these buildings it is a high outlier. Without this value, the rest of the heights form a fairly symmetrical distribution. The median height is the CITIC Plaza at 1283 feet tall.*

d) *The middle 50% of these buildings are between 1205 feet tall and 1483 feet tall.*

6) Mean travel time to work

a) Box plot: 5#-summary = {21.1, 22.1, 24.05, 26.55, 32.2 minutes}

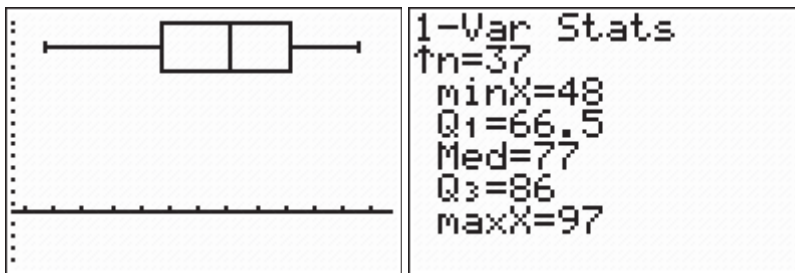


b) *The 'box' shows that the middle 50% of the mean travel times to work are between 22.1 minutes and 26.55 minutes for people in these cities.*

c) *The mean number of minutes it takes for workers in these cities to get to work ranges from 21.1 minutes in Roseville to 32.2 minutes in Albertville. The distribution for travel times is skewed to the right, but does not include any times that are outliers. The median travel time for these workers is 24.05 minutes.*

7) Wow So Fit

- a) 5# sum = { 48, 66.5, 77, 86, 97 }
- b) Box plot



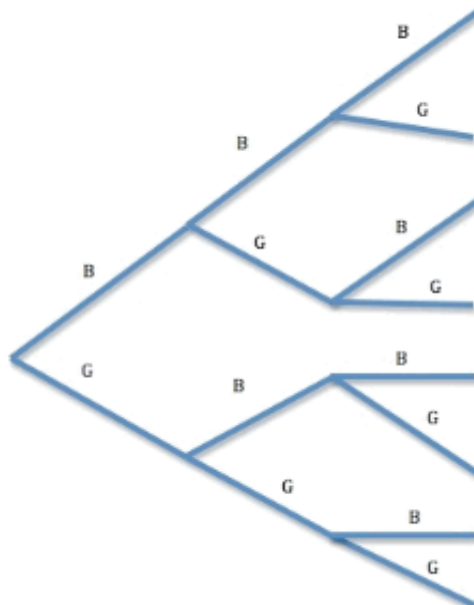
- c) *The median critics' rating for the Wow So Fit game was 77 on a scale of 1 to 100, with 100 being the highest rating. The ratings for this game range from 48 to 97. The distribution is slightly skewed to the left, but includes no outliers.*
- d) The "box" part tells us that the middle 50% of critics gave the Wow So Fit Game a rating between 66.5 and 86, on a scale of 1 to 100.

8) Reliability, validity or bias?

- a) Bias – The person is claiming that the ratings are too high due to the critics receiving a free game.
- b) Validity – The person is claiming that the rating system does not make sense as a way to measure the games because of the lack of guidelines.
- c) Reliability – The person is claiming that the rating system cannot be relied on to give consistent results.

9) Tree diagram

S = {BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG}



10) Probabilities

- a) $P(B,G,B) = 1/8 = 0.125$
- b) $P(\text{exactly two girls}) = 3/8 = 0.375$
- c) $P(\text{at least one boy}) = 7/8 = 0.875$

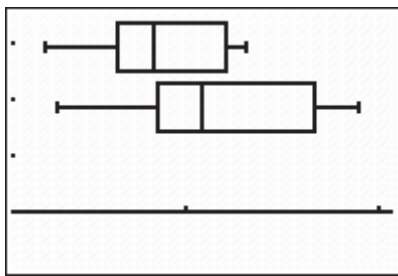
Section 5.6

1) % Daily Value of Total Fat vs. % Daily Value of Saturated Fat

- a) 5 # Summary Total Fat 5 # Summary Saturated Fat

<pre> 1-Var Stats ↑n=12 minX=13 Q1=32 Med=41.5 Q3=60 maxX=65 </pre>	<pre> 1-Var Stats ↑n=12 minX=16 Q1=42 Med=54 Q3=83.5 maxX=95 </pre>
---	---

- b) Parallel box plots

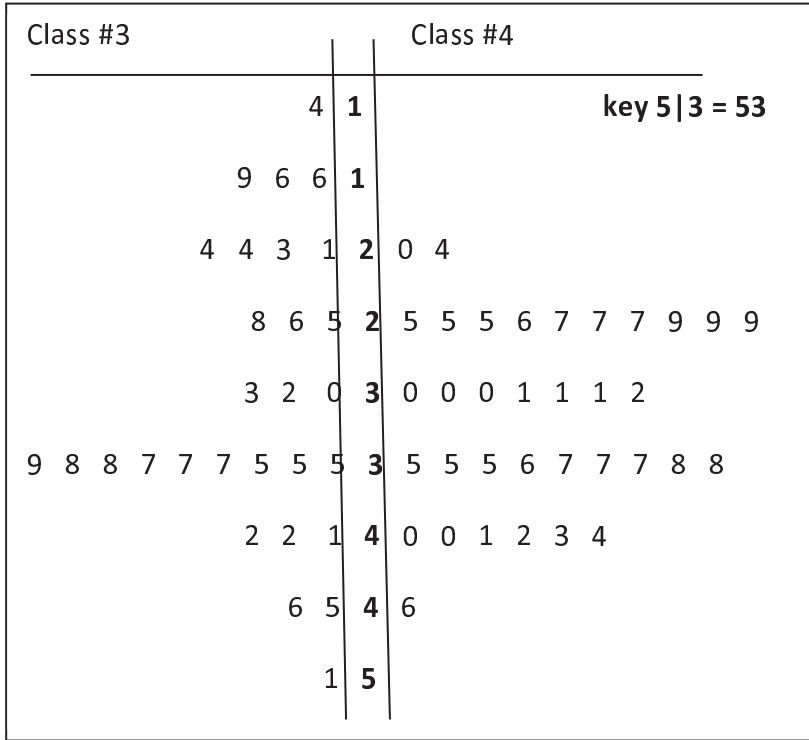


- c) Answers will vary. Possible answer: *The median percent of daily value for saturated fat for these McDonald's sandwiches is 12.5% higher than that of total fat. The first quartile for the saturated fat data is higher than the median for total fat. The saturated fat percentages are more spread out (16% to 95% of daily value) than the total fat percentages (13% to 65% of daily value). Neither distribution has any outliers.*

2) Boys' vs. Girls' Heights

- a) Boys: standard deviation = 3.796 in ; range = 12 in ; IQR = 7 in
 Girls: standard deviation = 3.215 in ; range = 11 in ; IQR = 5 in
- b) *The girls' heights are less spread out than the boys' heights are. All three of these measures of spread are less for girls than they are for boys.*
- c) Boys: mean = 69.133 in ; median = 69 in ; modes = 67 in & 73 in
 Girls: mean = 63.215 in ; median = 63.5 in ; mode = 61 in, 64 in & 66 in
- d) *The median height for girls is 5.5 inches less than it is for boys and the mean height for girls is about 5.9 inches less than it is for boys.*
- e) *The shape of the distributions of heights is roughly symmetrical for both boys and girls. Neither group has any outliers.*

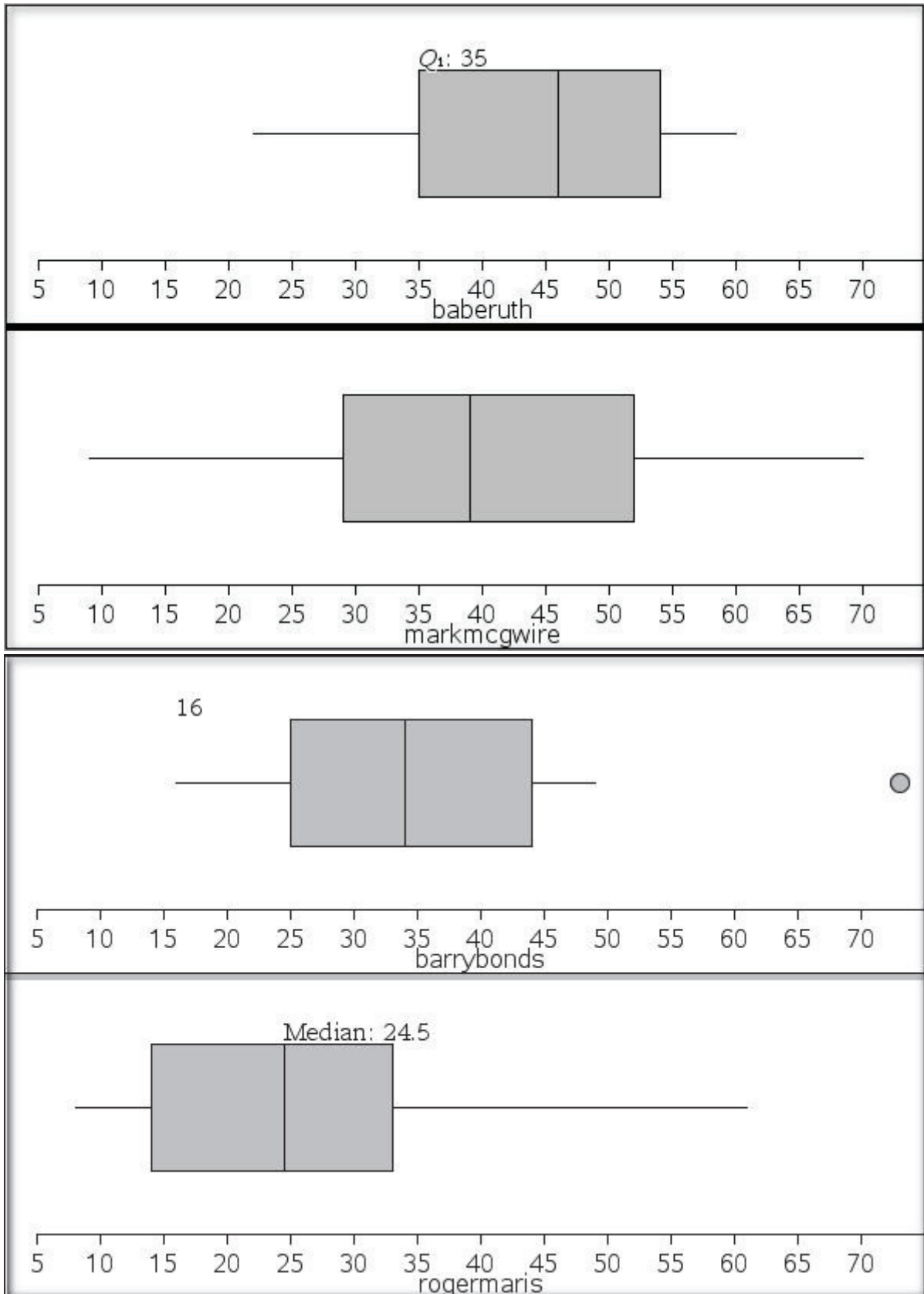
- 3) Common Assessment comparison
 a) Back-to-back (split-stem) plots



- b) Class #3—5number summary = {14, 24, 35, 38.5, 51}
 Class #4 – 5number summary = {20, 27, 31, 38, 46}
- c) Class #3: mean = 32.03 ; standard deviation = 9.77 ; modes = 35 & 37 ; range = 37 ; IQR = 14.5
 Class #4: mean = 32.91 ; standard deviation = 6.51 ; modes = “none” or list all with three occurrences 25, 27, 29, 30, 31, 35, 37 ; range = 26 ; IQR = 11
- d) Answers will vary. Some possible comments are: *The 4th period class had more consistent scores on this common assessment because they only had a range of 26, compared with a range of 37 for the 3rd period class. The 3rd period class had a higher median and maximum score, but it also had a lower minimum and quartile one score. So, the 3rd period class had students who scored very high as well as students who scored very low. Neither class had any scores that are considered to be outliers.*

4) Home-Runs

a) Parallel Box Plots



b) Players' Stats:

Player	Mean	St. Dev.	Range	IQR	Min	Q1	Med	Q3	Max
Babe Ruth	43.93	11.25	38	19	22	35	46	54	60
Mark McGwire	38.67	18.12	61	23	9	29	39	52	70
Barry Bonds	36.06	13.45	57	19	16	25	34	44	73
Roger Maris	26.1	15.61	53	19	8	14	24.5	33	61

- c) Outlier tests: $Q1 - 1.5IQR$ and $Q3 + 1.5IQR$
 Babe Ruth: $35 - 1.5(19) = 6.5$ and $54 + 1.5(19) = 82.5$; no low or high outliers
 Mark McGwire: $29 - 1.5(23) = -5.5$ and $52 + 1.5(23) = 86.5$; no low or high outliers
 Barry Bonds: $25 - 1.5(19) = -3.5$ and $44 + 1.5(19) = 72.5$; no low, but one high outlier of 73
 Roger Maris: $14 - 1.5(19) = -14.5$ and $33 + 1.5(19) = 61.5$; no low or high outliers

d) Answers will vary. Some possible comments could be: *Babe Ruth was much more consistent, he had the smallest range. The Great Bambino's median is higher than all three of the others. Also, his first quartile is higher than the median for Bonds and the third quartile for Maris. Even though Barry Bonds has the highest number, he is not the best hitter among these players because his high number of home runs was an outlier for him. Also, his median is lower than both McGwire and Ruth. Roger Maris and Barry Bonds both have distributions that are skewed to the right. Mark McGwire has a fairly symmetrical graph and Babe Ruth's graph is the only one that is skewed to the left.*

5) Gas Mileage. Answers will vary. Some possible comments are:

The small cars had the biggest range for gas mileage. Vans had the smallest range.

The vans had one low outlier and the small cars had two high outliers for gas mileage.

The Q1 for small cars is higher than the Q3 for all other types of cars. This means that 75% of small cars get better gas mileage than at least 75% of each of the other types of cars.

For the vans, the Q3 and maximum value are the same.

The compact and midsize graphs are fairly symmetrical.

The small and sporty graphs are both skewed to the right (higher numbers).

The large and vans are both skewed to the left (lower numbers).

6) Graphs

a) Graph I: symmetrical and bell shaped

Graph II: symmetrical and bell shaped

Graph III: symmetrical and bimodal

Graph IV: uniform (symmetrical and constant)

b) The mean and median of all of these would be around 52

c) The mean, median, and range would be the same for all of these.

d) Graph I would have the smallest standard deviation, because it is the least spread out and has the most observations near the mean. Graph III would have the largest standard deviation, because it is the most spread out and has the most observations far from the mean.

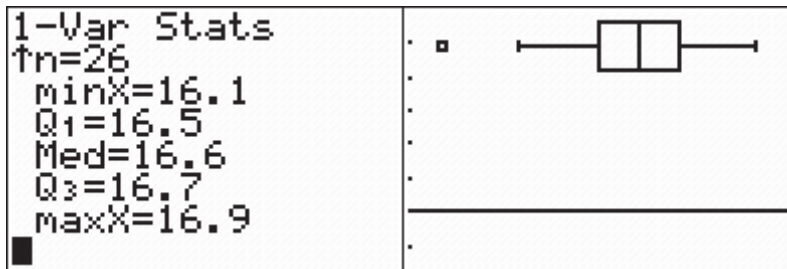
e) The mean and standard deviation would be appropriate Graphs I and II because of the symmetrical, bell shapes. The five-number summary would be more appropriate for Graphs III and IV because they are not bell shaped and have many observations far from the mean.

Section 5.7 Review

1) C

2) Weight

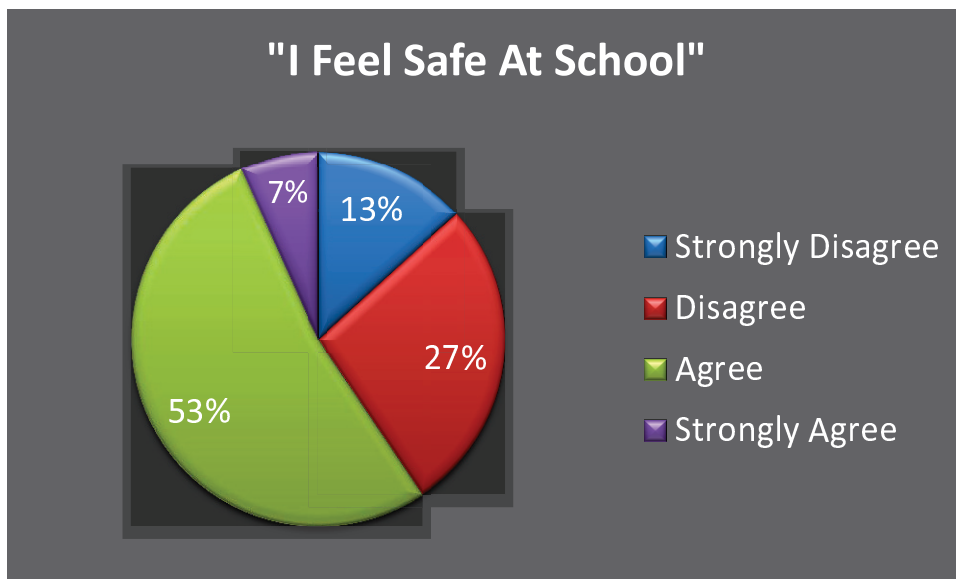
- a) Mean = 16.592 ounces ; standard deviation = 0.165 ounces ; mode = 16.6 ounces ; range = 0.8 ounces
- b) Five-number summary = {16.1, 16.5, 16.6, 16.7, 16.9} and box plot



- c) The dot plot is more informative because you can see every actual value. The mode and the shape are both very clear in the dot plot.
- d) 11.54% weigh more than 16.75 ounces. So, he will save money 88.46% of the time.

3)

- a) Pie chart



- b) Only 7% of the students surveyed said that they "strongly agree" that they feel safe at school. The most common reaction to "I feel safe at school" among these students was "agree", with 53% of the students responding this way. However, 2 out of every 5 of the students said that the "disagree" or "strongly disagree" with this statement.
- c) The committee should probably be concerned because 40% of the students do not feel safe at school. This is nearly half of the student body who have concerns for their safety.

- 4) B) the pennies
- 5) D) the doughnut holes
- 6) A) the heights of all 12th graders
- 7) C) the favorite movies
- 8) A
- 9) B
- 10) A
- 11) D
- 12) E
- 13) B
- 14) C
- 15) A
- 16) E
- 17) B

18) Cellphones & Landlines

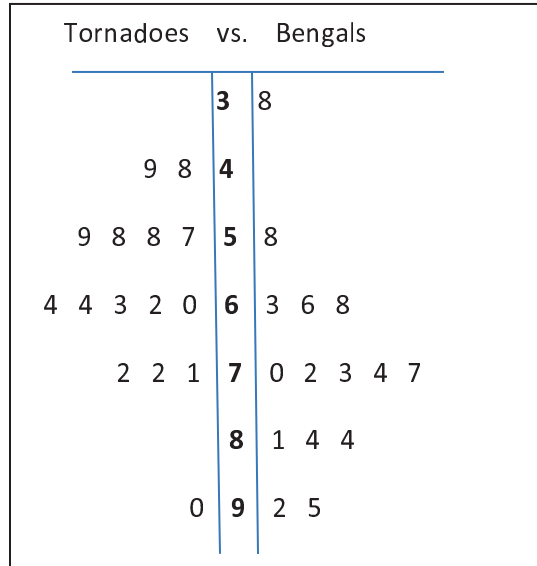
- a) Time plots
- b) The cellphone graph shows a steady rise in the number of cell phones per capita starting after 1985. It is increasing at a fast pace, but appears to be leveling off somewhat.

The landlines graph shows a rise in the number of landlines from 1985 to around 2000. After 2000 the number of landlines per capita has been dropping off at a fairly steep rate.

- c) Both were rising from 1985 to 2000. After 2000, the cellphones increased at a faster rate while the number of landlines declined. This shows that as cellphones have become more prevalent, many people no longer have landlines. The highest that the number of landlines reached was 0.7 per capita, but the number of cell phones has surpassed 0.85 and the trend shows that it is still increasing.
- d) Approximate per capita numbers: 0.2 in 1997; 0.75 in 2005; 1.15 in 2018
- e) Approximate per capita numbers: 0.7 at the peak; 2000; 0.25 in 2015

19) Tornadoes vs. Bengals

a) Back-to-back stem plot



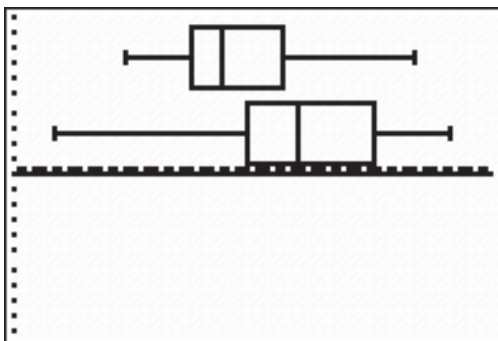
b) Tornadoes: 5# summary = {48,58.62,71,90}, mean = 63.133 pts. & standard deviation = 10.315 pts.

<pre> 1-Var Stats x̄=63.13333333 Σx=947 Σx²=61277 Sx=10.3154992 σx=9.965719018 ↓n=15 </pre>	<pre> 1-Var Stats ↑n=15 minX=48 Q1=58 Med=62 Q3=71 maxX=90 </pre>
---	---

Bengals: 5# summary = {38,66,73,84,95}, mean = 73 pts. & standard deviation = 14.147 pts.

<pre> 1-Var Stats x̄=73 Σx=1095 Σx²=82737 Sx=14.14718548 σx=13.66747965 ↓n=15 </pre>	<pre> 1-Var Stats ↑n=15 minX=38 Q1=66 Med=73 Q3=84 maxX=95 </pre>
--	---

c) Parallel box plots

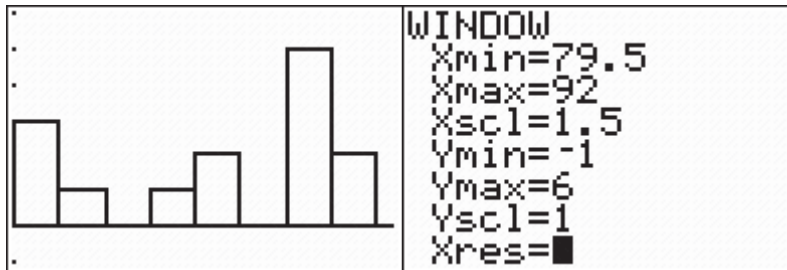


- d) Answers will vary. *The Bengals generally scored more points than the Tornadoes did in these 15 basketball games. The median number of points for the Bengals is higher than the third quartile for the Tornadoes. Also, the mean number of points is 10 points higher for the Bengals. The Bengals have one low outlier and their graph is skewed left, while the Tornadoes have one high outlier and their graph is skewed right. Even though the spreads are similar, the Bengals only had one game with less than 58 points, but the Tornadoes only had one game with more than 72 points. Based on only these points, the Bengals appear to be the better basketball team over this time period.*
- e) These are only the points for our team. Did they win or lose the games? How many points did the opposing teams score? Were any players out due to illness or injury?

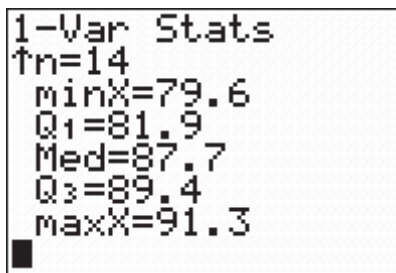
20) Percent of High School Graduates for several states

- a) Histogram

window used

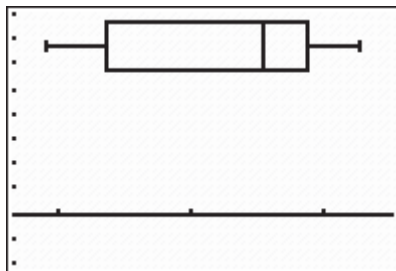


- b) 5# summary = {79.6%, 81.9%, 87.7%, 89.4%, 91.3%}



- c) Outlier test. $IQR = 89.4 - 81.9 = 7.5$
 $81.9 - 1.5(7.5) = 70.65$ There are no low outliers.
 $89.4 + 1.5(7.5) = 100.65$ There are no high outliers:

- d) Box plot

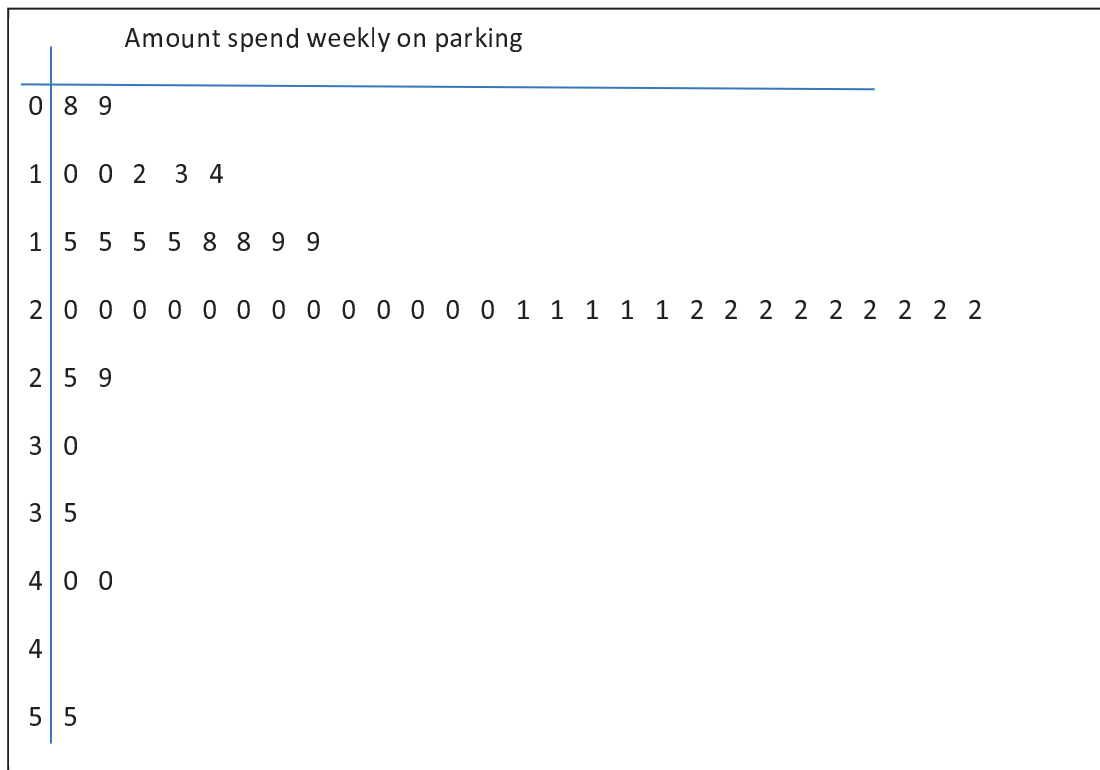


- e) Range = 11.7% ; IQR = 7.5%; modes = 73% and 85 %
- f) Mean = 86.27% ; standard deviation = 4.23%

- g) Mean < median. (1.43% difference)
- h) The five number summary would be more appropriate, because the graph is not symmetrical and bell shaped. There are many values far from the center of the distribution which has a strong effect on the standard deviation.
- i) Answers will vary. *The percent of people, 25 years and older, who are high school graduates for these states ranges from 79.6 % in Mississippi to 91.3% in Minnesota. The distribution for this 2010 data is not symmetrical, but doesn't contain any outliers. The histogram shows that these states form three different groupings. The median rate of high school graduates for these states is 86.27% for 2010 according to the U.S. Census website.*
- j) *According to this Census data, Minnesota has 91.3% of its residents over 25 years of age who are high school graduates. Minnesota is the highest among these states.*

21) Parking Costs

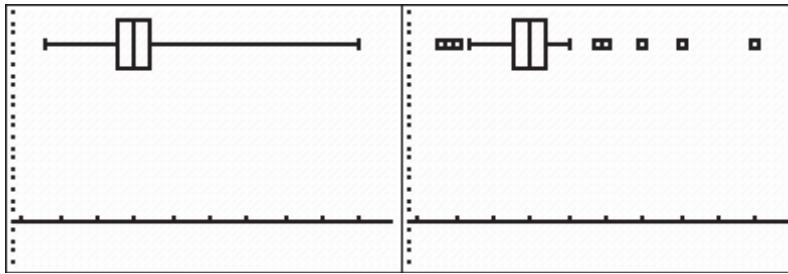
- a) Split-stem plot



- b) 5# summary = {8, 18, 20, 22, 50}

- c) Outlier test IQR = 22 - 18 = 4
 $18 - 1.5(4) = 12$ Four low outliers of \$8, \$9, \$10, and \$10 dollars per week.
 $22 + 1.5(4) = 28$ Six high outliers of \$29, \$30, \$35, \$40, \$40, and \$50 dollars per week.

d) Box plot and Modified box plot



e) Range = \$42 ; IQR = \$ 4 ; mode = \$20

f) Mean = \$20.90 ; standard deviation = \$7.65

g) Mean > median (slightly higher because it skewed to the right)

h) In this case, the five-number summary would be more appropriate, because the distribution has so many outliers and is so spread out.

i) *The amount of money spent weekly on parking by this sample of 50 employees has a large spread. It ranges from \$8 to \$50 per week. . However, the IQR is quite small. The middle 50% of these people spend between \$18 and \$22 dollars weekly on parking. There are ten outliers, four on the low end (\$8, \$9, \$10 & \$10) and six on the high end (\$29, \$30, \$35, \$40, \$40, & \$50). The graph is skewed to the right. The median and mode for the amount spent on parking is \$20 per week. The majority of employees spend less than \$22 per week.*