Chapter 2

Describing, Exploring, and Comparing Data

2-2 Frequency Distributions

1. Subtracting the first two consecutive lower class limits indicates that the class width is 100 - 90 = 10. Since there is a gap of 1.0 between the upper class limit of one class and the lower class limit of the next, class boundaries are determined by increasing or decreasing the appropriate class limits by (1.0)/2 = 0.5. The class boundaries and class midpoints are given in the table below.

pressure	class boundaries	class midpoint	frequency
90 - 99	89.5 - 99.5	94.5	1
100 - 109	99.5 - 109.5	104.5	4
110 - 119	109.5 - 119.5	114.5	17
120 - 129	119.5 - 129.5	124.5	12
130 - 139	129.5 - 139.5	134.5	5
140 - 149	139.5 - 149.5	144.5	0
150 - 159	149.5 - 159.5	154.5	1
			40

NOTE: Although they often contain extra decimal points and may involve consideration of how the data were obtained, class boundaries are the key to tabular and pictorial data summaries. Once the class boundaries are obtained, everything else falls into place. Here the first class width is readily seen to be 99.5 - 89.5 = 10.0 and the first midpoint is (89.5 + 99.5)/2 = 94.5. In this manual, class boundaries will typically be calculated first and then used to determine other values. In addition, the sum of the frequencies is an informative number used in many subsequent calculations and will be shown as an integral part of each table.

3. Since the gap between classes as presented is 1.0 the appropriate class limits are increased or decreased by (1.0)/2 = .05 to obtain the class boundaries and the following table.

cholesterol	class boundaries	class midpoint	frequency
0 - 199	-0.5 - 199.5	99.5	13
200 - 399	199.5 - 399.5	299.5	11
400 - 599	399.5 - 599.5	499.5	5
600 - 799	599.5 - 799.5	699.5	8
800 - 999	799.5 - 999.5	899.5	2
1000 - 1199	999.5 - 1199.5	1099.5	0
1200 - 1399	1199.5 - 1399.5	1299.5	1
			40

The class width is 199.5 - (-0.5) = 200; the first midpoint is (-0.5 + 199.5)/2

5. The relative frequency for each class is found by dividing its frequency by 40, the sum of the frequencies. NOTE: As before, the sum is included as an integral part of the table. For relative frequencies, this should always be 1.000 (i.e., 100%) and serves as a check for the calculations.

pressure	relative frequency
90 - 99	.025
100 - 109	.100
110 - 119	.425
120 - 129	.300
130 - 139	.125
140 - 149	.000
150 - 159	.025
	1.000

7. The relative frequency for each class is found by dividing its frequency by 40, the sum of the frequencies. NOTE: In #5, the relative frequencies were expressed as decimals; here they are expressed as percents. The choice is arbitrary.

cholesterol	relative frequency
0 - 199	.325
200 - 399	.275
400 - 599	.125
600 - 799	.200
800 - 999	.050
1000 -1199	.000
1200 -1399	.025
	1.000

9. The cumulative frequencies are determined by repeated addition of successive frequencies to obtain the combined number in each class and all previous classes. NOTE: Consistent with the emphasis that has been placed on class boundaries, we choose to use upper class boundaries in the "less than" column. Conceptually, pressures occur on a continuum and the integer values reported are assumed to be the nearest whole number representation of the precise measure. An exact pressure of 99.7, for example, would be reported as 100 and fall in the second class. The values in the first class, therefore, are better described as being "less than 99.5" (using the upper class boundary) than as being "less than 100." This distinction becomes crucial in the construction of pictorial representations in the next section. In addition, the fact that the final cumulative frequency must equal the total number (i.e, the sum of the frequency column) serves as a check for calculations. The sum of cumulative frequencies, however, has absolutely no meaning and is not included.

pressure	cumulative frequency
less than 99.5	1
less than 109.5	5
less than 119.9	22
less than 129.5	34
less than 139.5	39
less than 149.5	39
less than 159.5	40

11. The cumulative frequencies are determined by repeated addition of successive frequencies to obtain the combined number in each class and all previous classes. NOTE: Consistent with the emphasis that has been placed on class boundaries, we choose to use upper class boundaries in the "less than" column.

cholesterol	cumulative frequency
less than 199.5	13
less than 399.5	24
less than 599.5	29
less than 799.5	37
less than 999.5	39
less than 1199.5	39
less than 1399.5	40

13. The relative frequencies are determined by dividing the given frequencies by 200, the sum of the given frequencies. The fact that the sum of the relative frequencies is 1.000 provides a check to the arithmetic. For a fair die, we expect each relative frequency to be close to 1/6 = .167. As these relative frequencies all fall between .135 and .210, they do not appear to differ significantly from the values expected for a fair die.

outcome	relative frequency
1	.135
2	.155
3	.210
4	.200
5	.140
6	.160
	1.000

15. For a lower class limit of 0 for the first class and a class width of 50, the frequency distribution is given at the right.
NOTE: The class limits for the first class are 0-49 and not 0-50.

weight (lbs)	frequency
0 - 49	6
50 - 99	10
100 - 149	10
150 - 199	7
200 - 249	8
250 - 299	2
300 - 349	4
350 - 399	3
400 - 449	3
450 - 499	0
500 - 549	1
	54

17. The separate frequency distributions are given below.

male head circumference	frequency
34.0 - 35.9	2
36.0 - 37.9	0
38.0 - 39.9	5
40.0 - 41.9	29
42.0 - 43.9	14
	50

female head circumference	frequency
34.0 - 35.9	1
36.0 - 37.9	3
38.0 - 39.9	14
40.0 - 41.9	27
42.0 - 43.9	5
	50

It appears that the head circumferences tend to be larger for baby boys than for baby girls.

19. The separate relative frequency distributions are given below, to the right of the figure giving the actual frequencies. The relative frequencies were obtained by dividing the actual frequencies for each gender by the total frequencies for that gender.

M	age	F
11	19-28	8
43	29-38	18
31	39-48	4
22	49-58	7
4	59-68	2
111		39

male ages	relative frequency
19-28	.099
29-38	.387
39-48	.279
49-58	.198
59-68	.036
	1.000

female ages	relative frequency
19-28	.205
29-38	.462
39-48	.103
49-58	.179
59-68	.051
	1.000

Both genders have more runners in the 29-38 age group than in any other ten year spread. But the second most populous category is the one above that for the males, and the one below that for the females. It appears that the male runners tend to be slightly older than the female runners.

21. Assuming that "start the first class at 200 lb" refers to the first lower class limit produces the frequency table given at the right.

weight (lbs)	frequency
200 - 219	6
220 - 239	5
240 - 259	12
260 - 279	36
280 - 299	87
300 - 319	28
320 - 339	0
340 - 359	0
360 - 379	0
380 - 399	0
400 - 419	0
420 - 439	0
440 - 459	0
460 - 479	0
480 - 499	0
500 - 519	1
	175

In general, an outlier can add several rows to a frequency table. Even though most of the added rows have frequency zero, the table tends to suggest that these are possible valid values – thus distorting the reader's mental image of the distribution.

2-3 Visualizing Data

1. The answer depends upon what is meant by "the center." Since the ages seem to range from about 10 to about 70, 40 could be called the center value - in the sense that it is half way between the lowest and highest points on the horizontal axis that represent observed data. Since the ages are concentrated at the lower end of the scale, an age near 24 could be called the center value - in the sense that about ½ of the 131 ages appear to be below 24 and about ½ of the ages appear to be above 24. Or an age near 26 could be called the center value - in the sense that a fulcrum placed under 26 on the horizontal would appears to nearly "balance" the histogram.

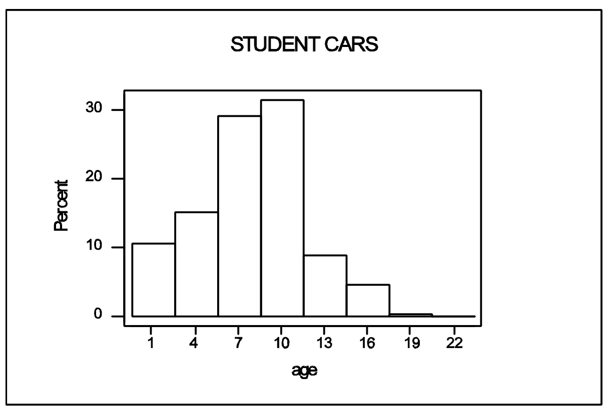
3. The percentage younger than 30 is about (1+38+38+16)/131 = 93/131 = .710 = 71.0%.

5. Since, the shading representing Group A covers about 40% of the "pie," the approximate percentage of people with Group A blood is 40%. If the chart is based on a sample of 500 people, approximately (.40)(500) = 200 people had Group A blood.

7. Obtain the relative frequencies by dividing each frequency by the total frequency for each sample. The two relative frequency histograms are given, using the same scale, below at the right. Each sample distribution spreads out in both directions from a "typical" (or most frequently occurring value), but the faculty data appears to be shifted to the left by about one interval. Since each interval represents 3 years, the faculty cars are about 3 years newer than the student cars. In addition, the majority of the student car ages fall below the class with the highest frequency, while faculty car ages tend to occur above their most populous class.

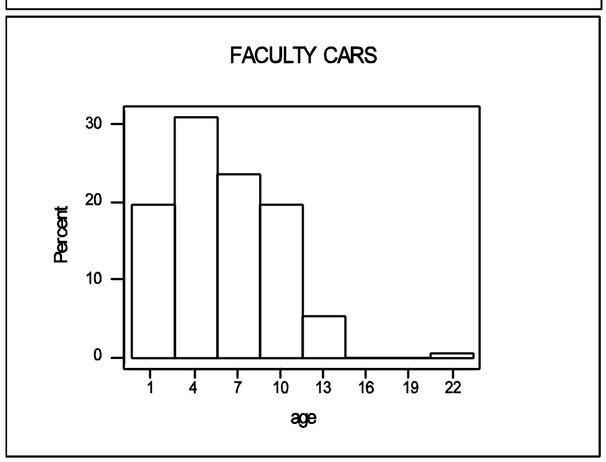
STUDENTS

age	frequency	relative frequency
0- 2	23	.106
3- 5	33	.152
6- 8	63	.290
9-11	68	.313
12-14	19	.088
15-17	10	.046
18-20	1	.005
21-23	0	.000
	217	1.000

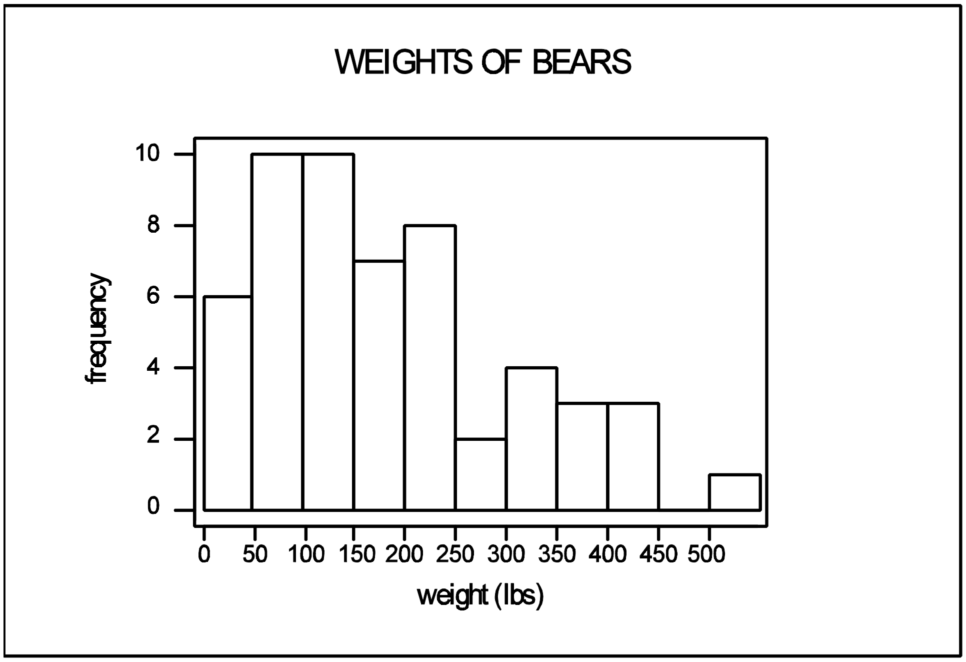


FACULTY

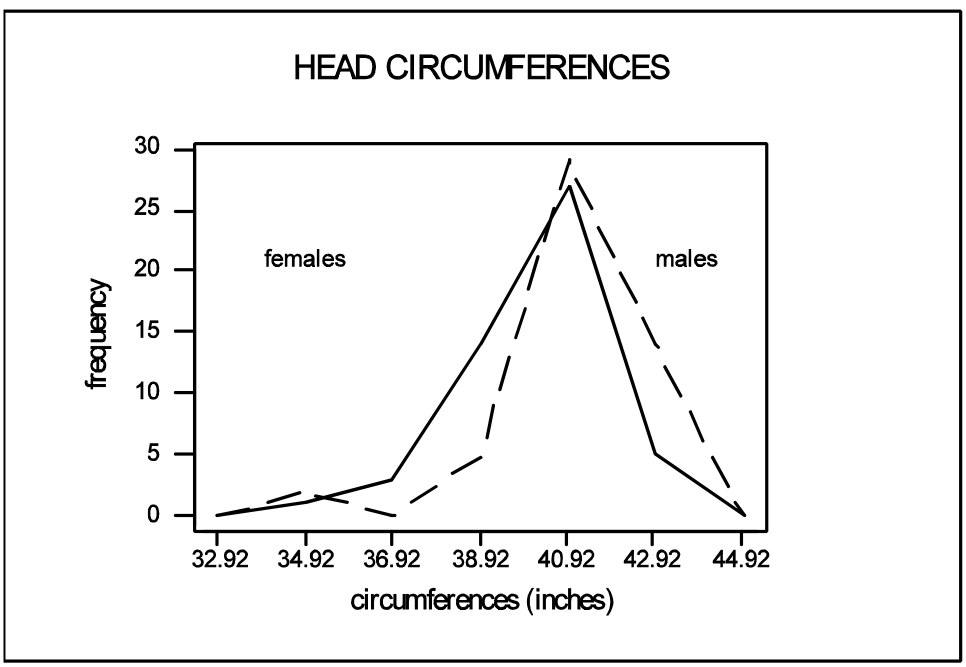
age	frequency	relative frequency
0- 2	30	.197
3- 5	47	.309
6- 8	36	.237
9-11	30	.197
12-14	8	.053
15-17	0	.000
18-20	0	.000
21-23	1	.007
	152	1.000



9. See the figure below. The class boundaries are -0.5,49.5,99.5,149.5,…549.5. The bars extend from class boundary to class boundary. For visual simplicity, the horizontal axis has been labeled 0,50,150,etc. The "center" of the weights (i.e., the "balance point" of the histogram) appears to be about 190.

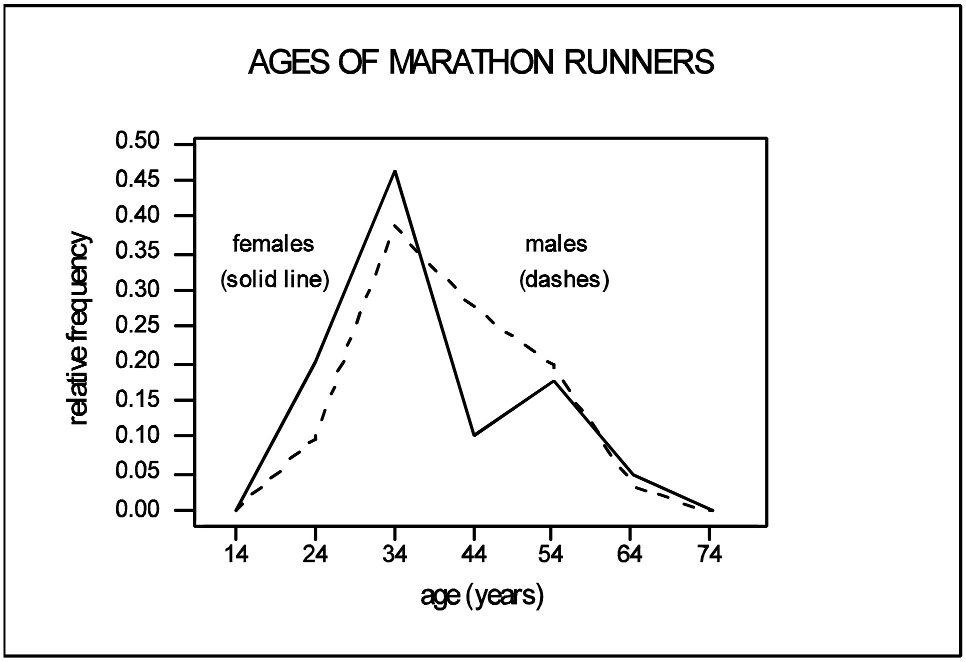


11. See the figure below. The frequencies are plotted above the class midpoints, and "extra midpoints are added so that both polygons begin and end with a frequency of zero. The females are represented by the solid line, and the males by the dashes.



Both polygons have approximately the same shape, but the male circumferences appear to be slightly larger.

13. See the figure below. The frequencies are plotted above the class midpoints, and "extra midpoints are added so that both polygons begin and end with a frequency of zero. The females are represented by the solid line, and the males by the dashes.

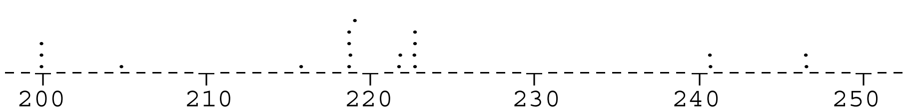


The two polygons are almost identical, except for the "dip" at age 44 for the females. It could be that physical or family conditions that typically occur about that age for females are not conducive to running marathons. Because of that dip, it appears that the male runners are slightly older than their female counterparts.
NOTE: The class limits for the first class were given as 19-28. Since age is reported as of last birthday and not to the nearest year, the upper class boundary extends all the way to (but not including) 29. This means that the class midpoint is 24 and not 23.5.

15. The original numbers are listed by the row in which they appear in the stem-and-leaf plot.

stem	leaves	original numbers
20	0005	200, 200, 200, 205
21	69999	216, 219, 219, 219, 219
22	2233333	222, 222, 223, 223, 223, 223, 223
23		
24	1177	241, 241, 247, 247

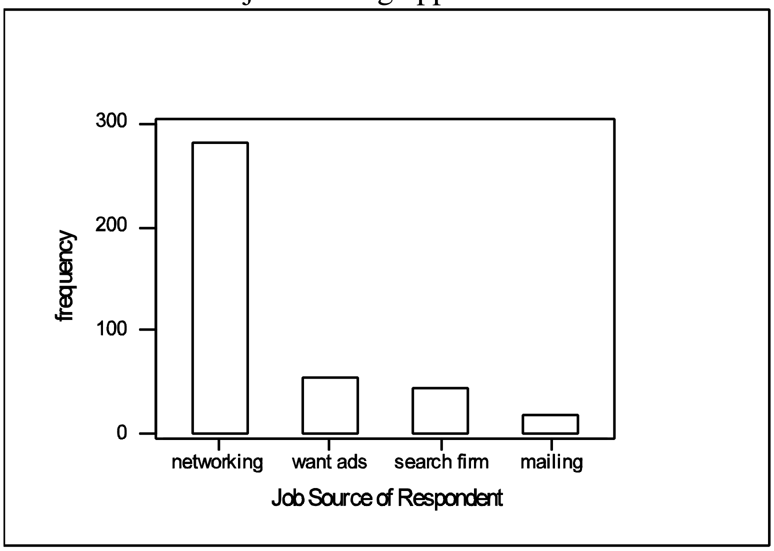
17. The dotplot is constructed using the original scores as follows. Each space represents 1 unit.



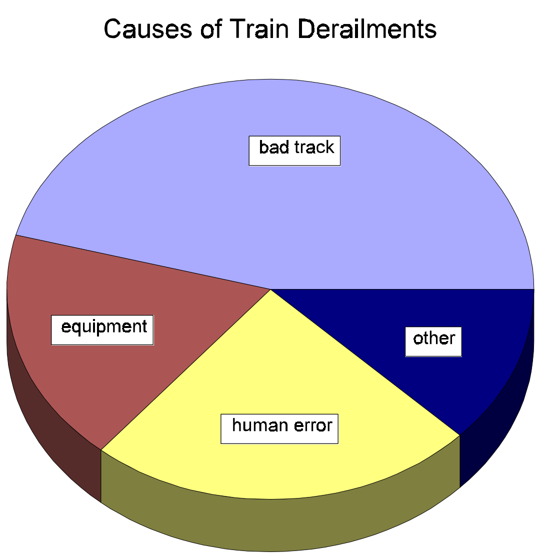
19. The expanded stem-and-leaf plot on the next page is one possibility. NOTE: The text claims that stem-and-leaf plots enable us to "see the distribution of data and yet keep all theinformation in the original list." Following the suggestion to round the nearest inch not only loses information but also uses subjectivity to round values exactly half way between. Since always rounding such values "up" creates a slight bias, many texts suggest rounding toward the even digit – so that 33.5 becomes 34, but 36.5 becomes 36. The technique below of using superscripts to indicate the occasional decimals is both mathematically clear and visually uncluttered.

stem	leaves
3	6 7
4	0 0 1 3 3^5
4	6 6 7 8 8 9
5	0 2 2^5 3 3 4
5	7^3 7^5 8 9 9 9
6	0 0^5 1 1 1^5 2 3 3 3 3^5 4 4 4
6	5 5 6^5 7 7^5 8^5
7	0 0^5 2 2 2 2 3 3^5
7	5 6^5

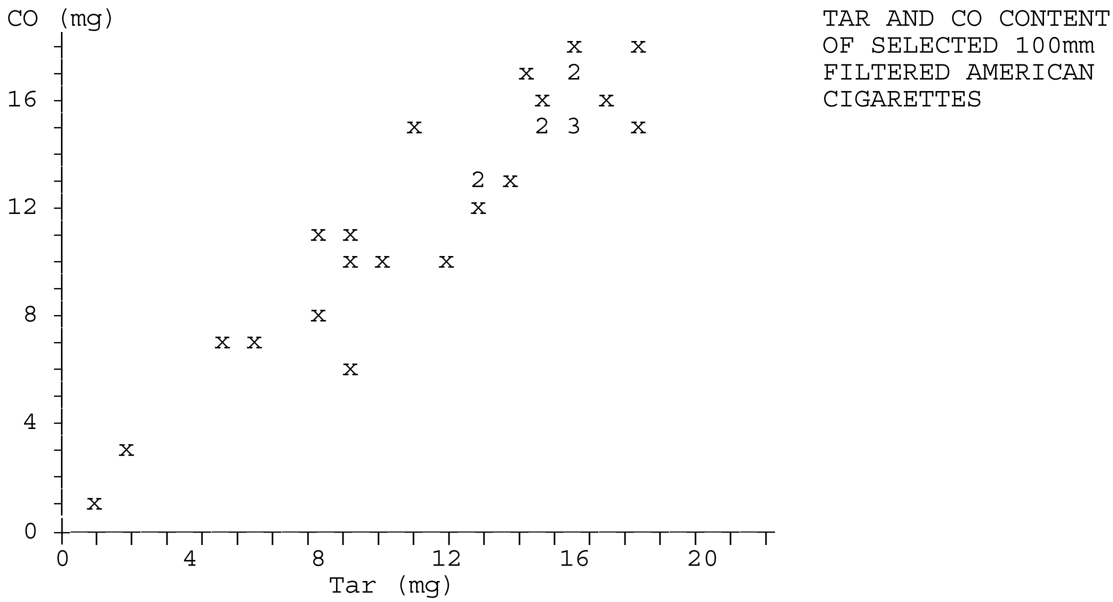
21. See the figure below, with bars arranged in order of magnitude. Networking appears to be the most effective job-seeking approach.



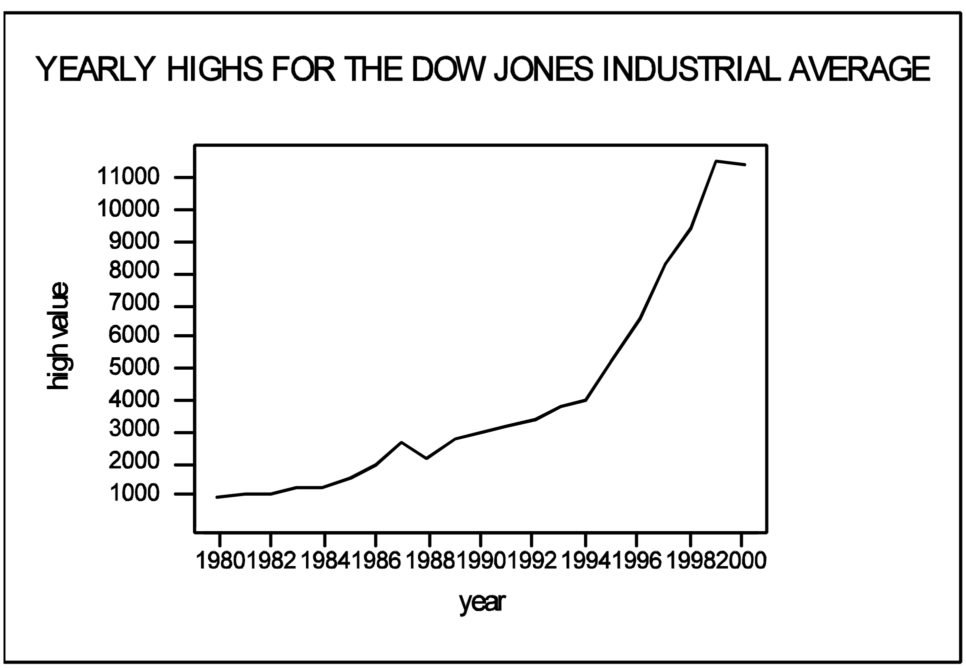
23. See the figure below. The sum of the frequencies is 50; the relative frequencies are 23/50 = 46%, 9/50 = 18%, 12/50 = 24%, and 6/50 = 12%. The corresponding central angles are (.46)360° = 165.6°, (.18)360° = 64.8°, (.24)360° = 86.4°, and (.12)360° = 43.2°. To be complete, the figure needs to be titled with the name of the quantity being measured.



25. The scatter diagram is given below. The figure should have a title, and each axis should be labeled both numerically and with the name of the variable. An "x" marks a single occurrence, while numbers indicate multiple occurrences at a point. Cigarettes high in tar also tend to be high in CO. The points cluster about a straight line from (0,0) to (18,18), indicating that the mg of CO tends to be about equal to the mg of tar.



27. The time series graph is given below. An investor could use this graph to project where the market might be next year and invest accordingly. The flat portion at the right might be suggesting that the market has reached a point at which it will level off for a while – and that the rapid growth of the last several years may be at end.

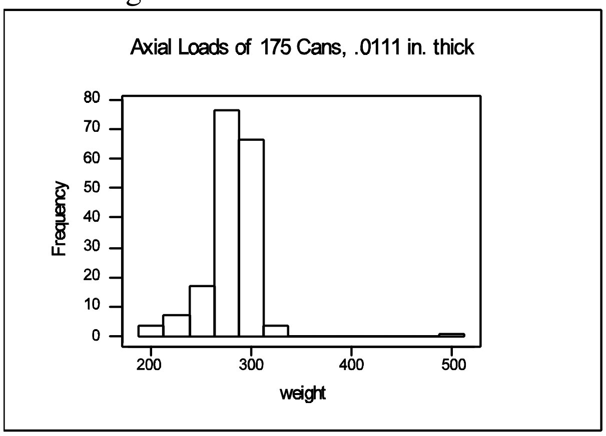


29. According to the figure, 422,000 started and 10,000 returned. 10,000/422,000 = 2.37%

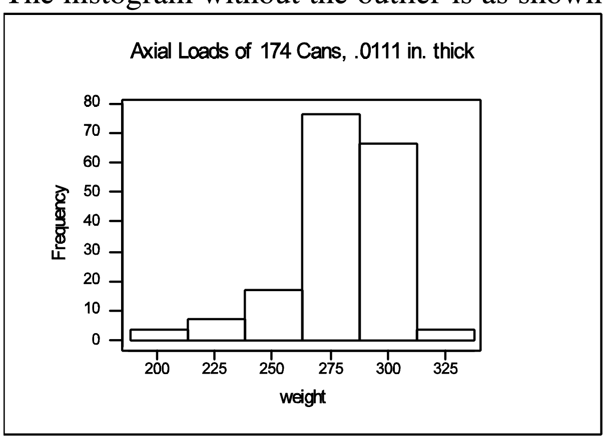
31. The figure indicates the number of men had just dropped to 37,000 on November 9 when the temperature was 16°F (-9°C), and had just dropped to 24,000 on November 14 when the temperature was -6°F (-21°C). The number who died during that time, therefore, was 37,000 - 24,000 = 13,000.

33. NOTE: Exercise #20 of section 2-2 dealt with frequency table representations of this data and specified a class width of 20. Using a class width with an odd number of units of measure allows the class midpoint to have the same number of decimal places as the original data and often produces more appealing visual representations. Here we use a class width of 25 – with class midpoints of 200, 225, etc. – and so the figures will differ from ones made using the previous classes. The histogram bars extend from class boundary to class boundary (i.e., from 187.5 to 212.5 for the first class), but for convenience the labels have been placed at the class midpoints.

a. The histogram with the outlier is as shown below.



b. The histogram without the outlier is as shown below.



c. The basic shape of the histogram does not change, except that a distant piece has been "broken off." NOTE: The redrawn histogram in part (b) should not be an exact copy of the one in part (a) with the distant bar erased. Since removing the distant bar reduces the effective width of the figure significantly, the rule that the height should be approximately 3/4 of the width requires either making the remaining bars wider (and keeping the figure's height) or making them shorter (and keeping the figure's reduced width). To do otherwise produces a figure too tall for its width – and one that tends to visually overstate the differences between classes.

2-4 Measures of Center

NOTE: As it is common in mathematics and statistics to use symbols instead of words to represent quantities that are used often and/or that may appear in equations, this manual employs symbols for the measures of central tendency as follows:

mean = $\bar{x}$ mode = M
median = $\tilde{x}$ midrange = m.r.

Also, this manual generally follows the author's guideline of presenting means, medians and midranges accurate to one more decimal place than found in the original data. The mode, the only measure which must be one of the original pieces of data, is presented with the same accuracy as the original data. This manual will, however, recognize the following two exceptions to these guidelines.
(1) When there are an odd number of data, the median will be one of the original values and will not be reported to one more decimal place. And so the median of 1,2,3,4,5 would be reported as 3.
(2) When the mean falls exactly halfway between two values that meet the guidelines, an extra decimal place ending in 5 will be used. And so the mean of 1,2,3,3 would be reported as 9/4 = 2.25.

1. Arranged in order, the 6 scores are: 0 0 0 176 223 548
   a. $\bar{x} = (\Sigma x)/n = (947)/6 = 157.8$ c. M = 0
   b. $\tilde{x} = (0 + 176)/2 = 88.0$ d. m.r. = (0 + 548)/2 = 274.0
   Whether or not there is a problem with scenes of tobacco use in animated children's films depends on the scene. If a likeable character uses tobacco without any criticism or ill effects, this would suggest to children that such usage is appropriate and acceptable. In most children's films, however, tobacco usage is typically associated with evil and undesirable characters and situations – suggesting that tobacco usage is not appropriate or acceptable for decent persons.

NOTE: The median is the middle score when the scores are arranged in order, and the midrange is halfway between the first and last score when the scores are arranged in order. It is therefore usually helpful to begin by placing the scores in order. This will not affect the mean, and it may also aid in identifying the mode. In addition, no measure of central tendency can have a value lower than the smallest score or higher than the largest score – remembering this helps to protect against gross errors, which most commonly occur when calculating the mean.

3. Arranged in order, the 16 scores are:
   .03 .07 .09 .13 .13 .17 .24 .30 .39 .43 .43 .44 .45 .47 .47 .48
   a. $\bar{x} = (\Sigma x)/n = (4.72)/16 = .295$ c. M = .13,.43,.47 [tri-modal]
   b. $\tilde{x} = (.30 + .39)/2 = .345$ d. m.r. = (.03 + .48)/2 = .255
   No; this sample means weights each brand of cereal equally and does not take into account which of the cereals have higher (or lower) rates of consumption.

5. Arranged in order, the 15 scores are:
   .12 .13 .14 .16 .16 .16 .17 .17 .17 .18 .21 .24 .24 .27 .29
   a. $\bar{x} = (\Sigma x)/n = (2.81)/15 = .187$ c. M = .16,.17 [bi-modal]
   b. $\tilde{x} = .17$ d. m.r. = (.12 + .29)/2 = .205
   Yes; these values appear to be significantly above the allowed maximum.

7. Arranged in order, the 20 scores are:
   15 17 17 17 17 17 17 18 18 18 18 18 19 19 19 19 20 21 21 21
   a. $\bar{x} = (\Sigma x)/n = (366)/20 = 18.3$ c. M = 17
   b. $\tilde{x} = (18 + 18)/2 = 18.0$ d. m.r. = (15 + 21)/2 = 18.0
   The sample results appear to be very consistent – i.e., there appears to be little variation from person to person. This suggests that there is little variation in the population, and

that the sample mean should be a good estimate of the population mean – in the sense that other samples would likely produce similar results that vary little from this sample.

9. Arranged in order, the scores are as follows.
JV: 6.5 6.6 6.7 6.8 7.1 7.3 7.4 7.7 7.7 7.7
Pr: 4.2 5.4 5.8 6.2 6.7 7.7 7.7 8.5 9.3 10.0

Jefferson Valley	Providence
n = 10	n = 10
$\bar{x} = (\Sigma x)/n = (71.5)/10 = 7.15$	$\bar{x} = (\Sigma x)/n = (71.5)/10 = 7.15$
$\tilde{x} = (7.1 + 7.3)/2 = 7.20$	$\tilde{x} = (6.7 + 7.7)/2 = 7.20$
M = 7.7	M = 7.7
m.r. = (6.5 + 7.7)/2 = 7.10	m.r. = (4.2 + 10.0)/2 = 7.10

Comparing only measures of central tendency, one might suspect the two sets are identical. The Jefferson Valley times, however, are considerably less variable.
NOTE: This is the reason most banks have gone to the single waiting line. While it doesn't make service faster, it makes service times more equitable by eliminating the "luck of the draw" – i.e., ending up by pure chance in a fast or slow line and having unusually short or long waits.

11. Arranged in order, the scores are as follows.
McDonald's: 92 118 128 136 153 176 192 193 240 254 267 287
Jack in the Box: 74 109 109 190 229 255 270 300 328 377 428 481

McDonald's	Jack in the Box
n = 12	n = 12
$\bar{x} = (\Sigma x)/n = (2236)/12 = 186.3$	$\bar{x} = (\Sigma x)/n = (3150)/12 = 262.5$
$\tilde{x} = (176 + 192)/2 = 184.0$	$\tilde{x} = (255 + 270)/2 = 262.5$
M = [none]	M = 109
m.r. = (92 + 287)/2 = 189.5	m.r. = (74 + 481)/2 = 277.5

McDonald's appears to be faster. Yes; the difference appears to be significant.

13. The following values were obtained for the head circumference data, where x_i indicates the i^{th} score from the ordered list.

males	females
n = 50	n = 50
$\bar{x} = (\Sigma x)/n = 2054.9/50 = 41.10$	$\bar{x} = (\Sigma x)/n = 2002.4/50 = 40.05$
$\tilde{x} = (x_{25}+x_{26})/2$	$\tilde{x} = (x_{25}+x_{26})/2)/2$
$= (41.1+41.1)/2 = 41.10$	$= (40.2+40.2 = 40.20$

No; on the basis of the means and medians alone, there does not appear to be a significant difference between the genders.

15. The following values were obtained for Boston rainfall, where x_i indicates the ith score from the ordered list.

Thursday	Sunday
n = 52	n = 52
$\bar{x} = (\Sigma x)/n = 3.57/52 = .069$	$\bar{x} = (\Sigma x)/n = 3.52/52 = .068$
$\tilde{x} = (x_{26}+x_{27})/2 = (.00 + .00)/2 = .000$	$\tilde{x} = (x_{26}+x_{27})/2 = (.00 + .00)/2 = .000$

If "it rains more on weekends" refers to the amount of rain, the data do not support the claim. The amount of rainfall appears to be virtually the same for Thursday and Sunday. If "it rains more on weekends" refers to the frequency of rain (regardless of the amount), then the proportions of days on which there was rain would have to be compared.

17. The x values below are the class midpoints from the given frequency table.

x	f	x·f
44.5	8	356.0
54.5	44	2398.0
64.5	23	1483.5
74.5	6	447.0
84.5	107	9041.5
94.5	11	1039.5
104.5	1	104.5
	200	14870.0

$\bar{x} = (\Sigma x \cdot f)/n$
$= (14870.0)/200$
$= 74.35$

NOTE: The mean time was calculated to be 74.35 minutes. According to the rule given in the text, this value should be rounded to one decimal place. The text describes how many decimal places to present in an answer, but not the actual rounding process. When the figure to be rounded is <u>exactly</u> half-way between two values (i.e., the digit in the position to be discarded is a 5, and there are no further digits because the calculations have "come out even"), there is no universally accepted rounding rule. Some authors say to always round up such a value; others correctly note that always rounding up introduces a consistent bias, and that the value should actually be rounded up half the time and rounded down half the time. And so some authors suggest rounding toward the even value (e.g., .65 becomes .6 and .75 becomes .8), while others simply suggest flipping a coin. In this manual, answers <u>exactly</u> half-way between will be reported without rounding (i.e., stated to one more decimal than usual).

19. The x values below are the class midpoints from the given frequency table.

x	f	x·f
43.5	25	1087.5
47.5	14	665.0
51.5	7	360.5
55.5	3	166.5
59.5	1	59.5
	50	2339.0

$\bar{x} = (\Sigma x \cdot f)/n$
$= (2339.0)/50$
$= 46.8$

The mean speed of 46.8 mi/hr of those ticketed by the police is more than 1.5 times the posted speed limit of 30 mi/hr. NOTE: This indicates nothing about the mean speed of *all* drivers, a figure which may or may not be higher than the posted limit.

21. a. Arranged in order, the original 54 scores are:

```
 26  29  34  40  46  48  60  62  64  65  76  79  80  86  90
 94 105 114 116 120 125 132 140 140 144 148 150 150 154 166
166 180 182 202 202 204 204 212 220 220 236 262 270 316 332
344 348 356 360 365 416 436 446 514
```

$\bar{x} = (\Sigma x)/n = (9876)/54 = 182.9$

b. Trimming the highest and lowest 10% (or 5.4 = 5 scores), the remaining 44 scores are:

```
 48  60  62  64  65  76  79  80  86  90  94 105 114 116 120
125 132 140 140 144 148 150 150 154 166 166 180 182 202 202
204 204 212 220 220 236 262 270 316 332 344 348 356 360
```

$\bar{x} = (\Sigma x)/n = (7524)/44 = 171.0$

c. Trimming the highest and lowest 20% (or 10.8 = 11 scores), the remaining 32 scores are:

```
 79  80  86  90  94 105 114 116 120 125 132 140 140 144 148
150 150 154 166 166 180 182 202 202 204 204 212 220 220 236
262 270
```

$\bar{x} = (\Sigma x)/n = (5093)/32 = 159.2$

In this case, the mean gets smaller as more scores are trimmed. In general, means can increase, decrease, or stay the same as more scores are trimmed. The mean decreased here because the higher scores were farther from the original mean than were the lower scores.

23. If n=10 values have $\bar{x} = (\Sigma x)/n = 75.0$, then $\Sigma x = 750$.
    a. Since the sum of the 9 given values is 698, the 10th value must be 750-698 = 52.
    b. Generalizing the result in part (a), n-1 values can be freely assigned. The nth value must be whatever is needed (whether positive or negative) to reach the sum $(\Sigma x) = n \cdot \bar{x}$ determined by the given mean.

25. The w values below are the weights – which can sum to any value, since the weighted mean is found by dividing by Σw.

x	w	x·w
65	.15	9.75
83	.15	12.45
80	.23	12.00
90	.15	13.50
92	.40	36.80
	1.00	84.50

$\bar{x} = (\Sigma x \cdot w)/(\Sigma w)$
$= (84.50)/(1.00)$
$= 84.50$

27. Let $\bar{x}_h$ stand for the harmonic mean: $\bar{x}_h = n/[\Sigma(1/x)] = 2/[1/40 + 1/60] = 2/[.0417]$
$= 48.0$

29. R.M.S. $= \sqrt{\Sigma x^2/n} = \sqrt{[(110)^2 + (0)^2 + (-60)^2 + (12)^2]/4} = \sqrt{15844/4} = \sqrt{3961} = 62.9$

2-5 Measures of Variation

NOTE: Although not given in the text, the symbol R will be used for the range throughout this manual. Remember that the range is the difference between the highest and the lowest scores, and not necessarily the difference between the last and the first values as they are listed. Since calculating the range involves only the subtraction of 2 original pieces of data, that measure of variation will be reported with the same accuracy as the original data.

1.

x	$x-\bar{x}$	$(x-\bar{x})^2$	x^2
0	-157.83	24910.3089	0
0	-157.83	24910.3089	0
0	-157.83	24910.3089	0
176	18.17	330.1489	30976
223	65.17	4247.1289	49729
548	390.17	152232.6289	300304
947	0.02	231540.8834	381009

$\bar{x} = (\Sigma x)/n = 947/6 = 157.83$

R = 5480 = 548

by formula 2-4,
$s^2 = \Sigma(x-\bar{x})^2/(n-1)$
$= 231540.8834/5$
$= 46308.17$
$= 46308.2$

by formula 2-5,
$s^2 = [n(\Sigma x^2) - (\Sigma x)^2]/[n(n-1)]$
$= [6(381009) - (947)^2]/[6(5)]$
$= [13888888889245]/[30]$
$= 46308.17$
$= 46308.2$

$s = \sqrt{46308.17} = 215.2$

The times appear to vary widely.

NOTE: When finding the square root of the variance to obtain the standard deviation, use all the decimal places of the variance, and not the rounded value reported as the answer. The best way to do this is either to keep the value on the calculator display or to place it in the memory. Do not copy down all the decimal places and then re-enter them to find the square root, as that could introduce round-off and/or copying errors.

When using formula 2-4, constructing a table having the first three columns shown above helps to organize the calculations and makes errors less likely. In addition, verify that $\Sigma(x-\bar{x}) = 0$ [except for a possible small discrepancy at the last decimal, due to using a rounded value for the mean] before proceeding – if such is not the case, there is an error and further calculation is