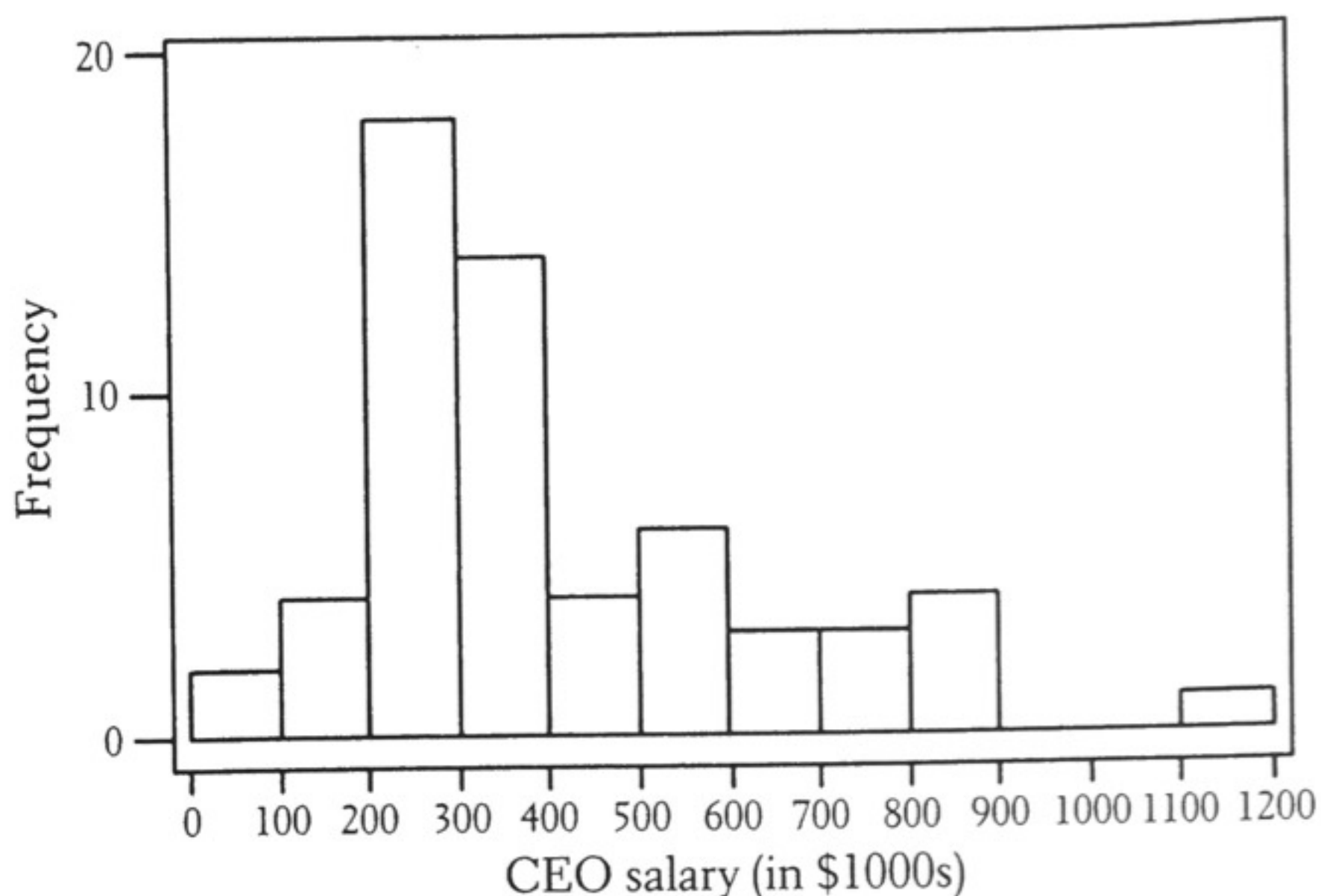


Exploring Data

1.14



The distribution is skewed to the right with a peak in the 200s class. The spread is approximately 1100 (\$21,000 to \$1,103,000) and the center is located at 350 (\$350,000). There is one outlier in the 1100s class.

1.15 (b) The distribution is symmetric with a peak at class (chest size) 40. The center is also located at 40. The spread is 15 (33 to 48). Assuming that the sample is representative of all members of the population, the distribution would provide a useful guide to those making clothing for the militiamen. From the frequency table, it is easy to estimate the percentage of all militiamen who have a certain chest size. The production of uniforms can reflect this distribution.

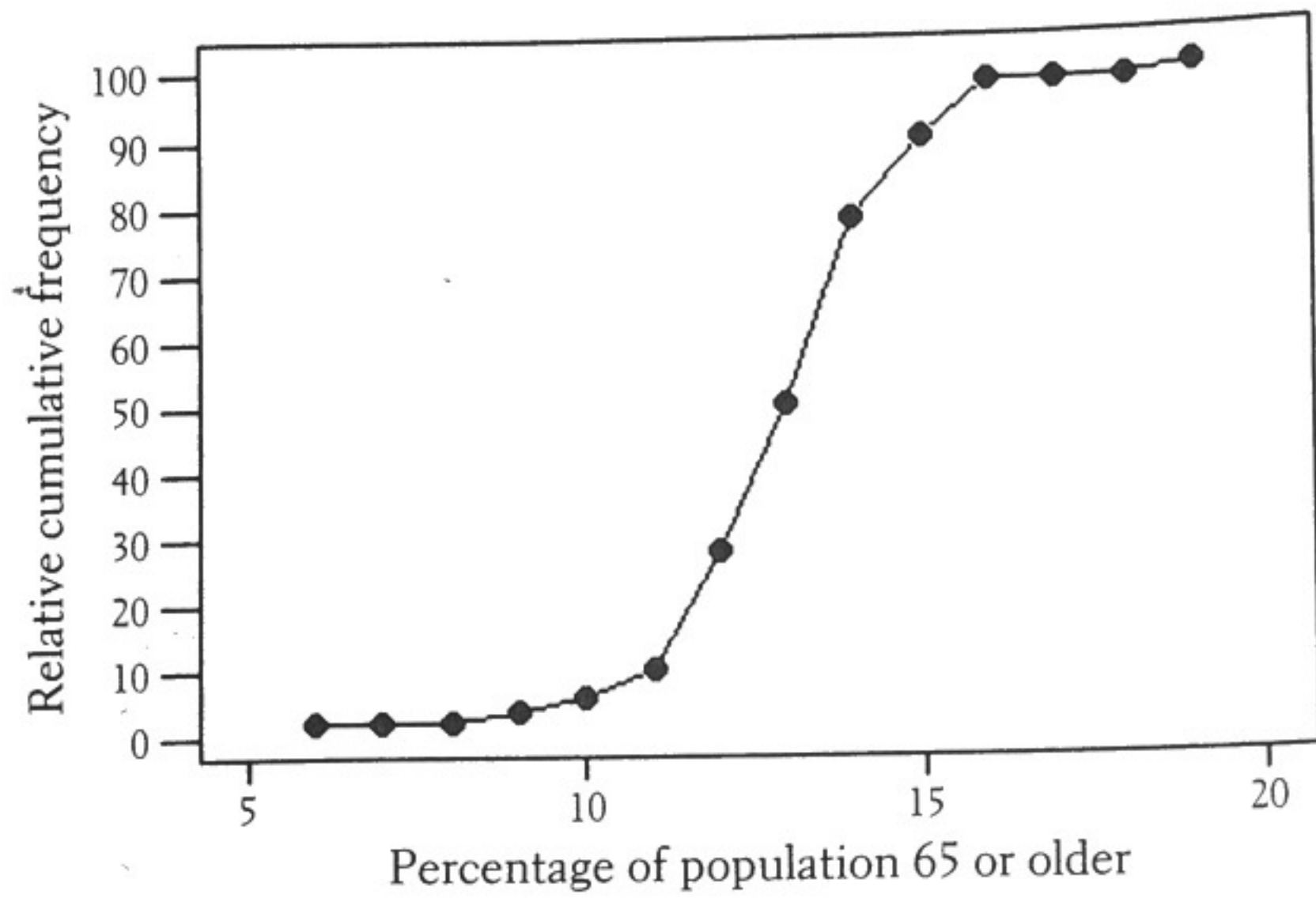
1.16 (a) Roughly symmetric, though it might be viewed as SLIGHTLY skewed to the right. (b) About 15%. (39% of the stocks had a total return less than 10%, while 60% had a return less than 20%. This places the center of the distribution somewhere between 10% and 20%.) (c) The smallest return was between -70% and -60%, while the largest was between 100% and 110%. (d) 23% (1 + 1 + 1 + 1 + 3 + 5 + 11).

1.17 (a) Skewed to the right; center at about 3 (31 less than 3, 11 equal to 3, 23 more than 3); spread: 0 to 10. No outliers. (b) About 23% (15 out of 65 years).

1.18 Lightning histogram: centered at noon (or more accurately, somewhere from 11:30 to 12:30). Spread is from 7 to 17 (or more accurately, 6:30 AM to 17:30, i.e., 5:30 PM). Shakespeare histogram: centered at 4, spread from 1 to 12.

1.19 (a)

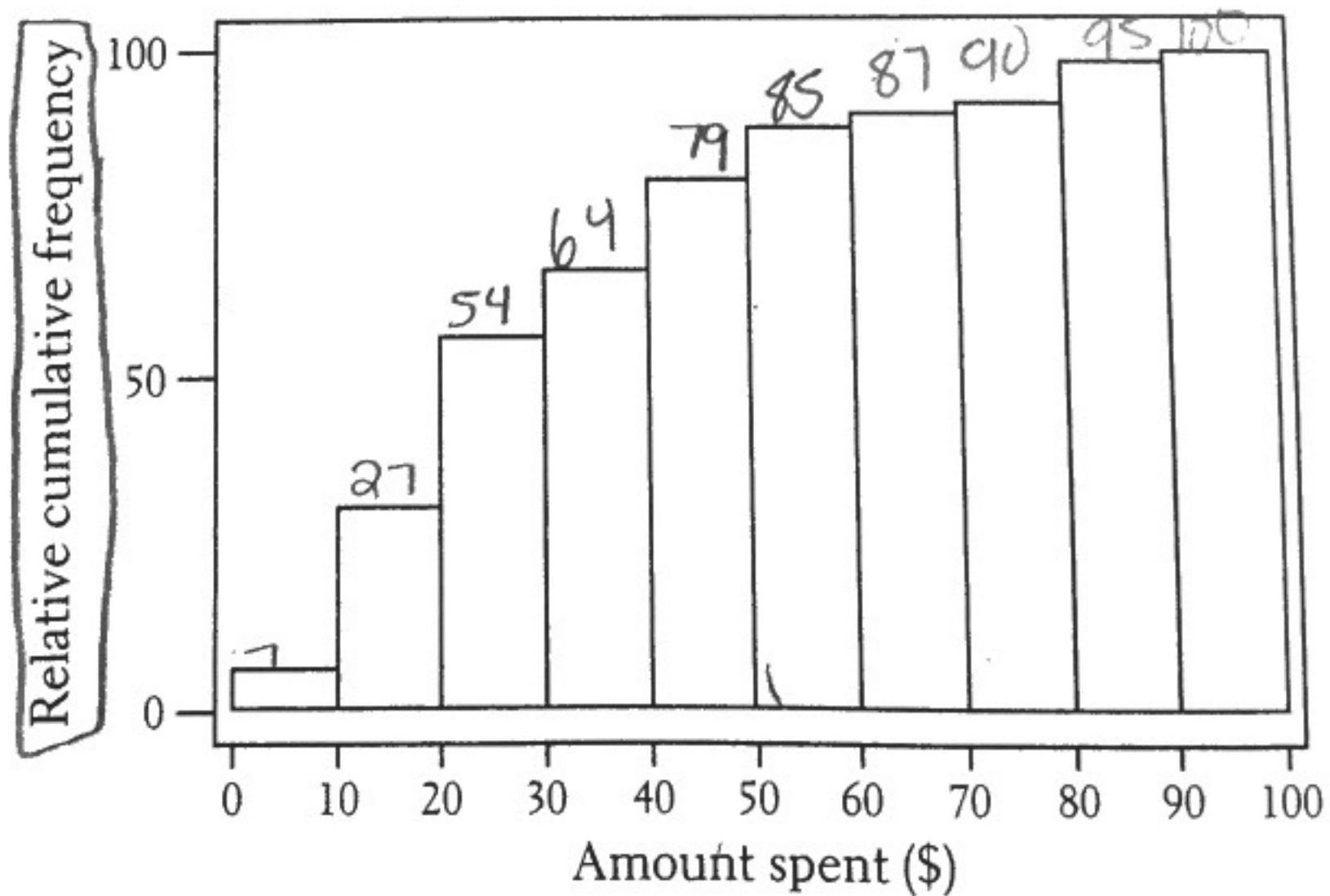
Percent	Cumulative frequency	Relative cumulative frequency	Percent	Cumulative frequency	Relative cumulative frequency
5.0-5.9	1	2%	12.0-12.9	25	50%
6.0-6.9	1	2%	13.0-13.9	39	78%
7.0-7.9	1	2%	14.0-14.9	45	90%
8.0-8.9	2	4%	15.0-15.9	49	98%
9.0-9.9	3	6%	16.0-16.9	49	98%
10.0-10.9	5	10%	17.0-17.9	49	98%
11.0-11.9	14	28%	18.0-18.9	50	100%



- (b) • Percentage of states in which percentage of "65 and older" is less than 15% = 90%, since the point (15, 90) lies on the ogive.
- 40th percentile of distribution $\approx 12.4\%$, since the horizontal line drawn from 40% on the vertical axis intersects the ogive at a point whose horizontal coordinate is approximately 12.4%. Less than 40% of states have 12.4% or less of their population aged 65 or older.
- Answers vary.

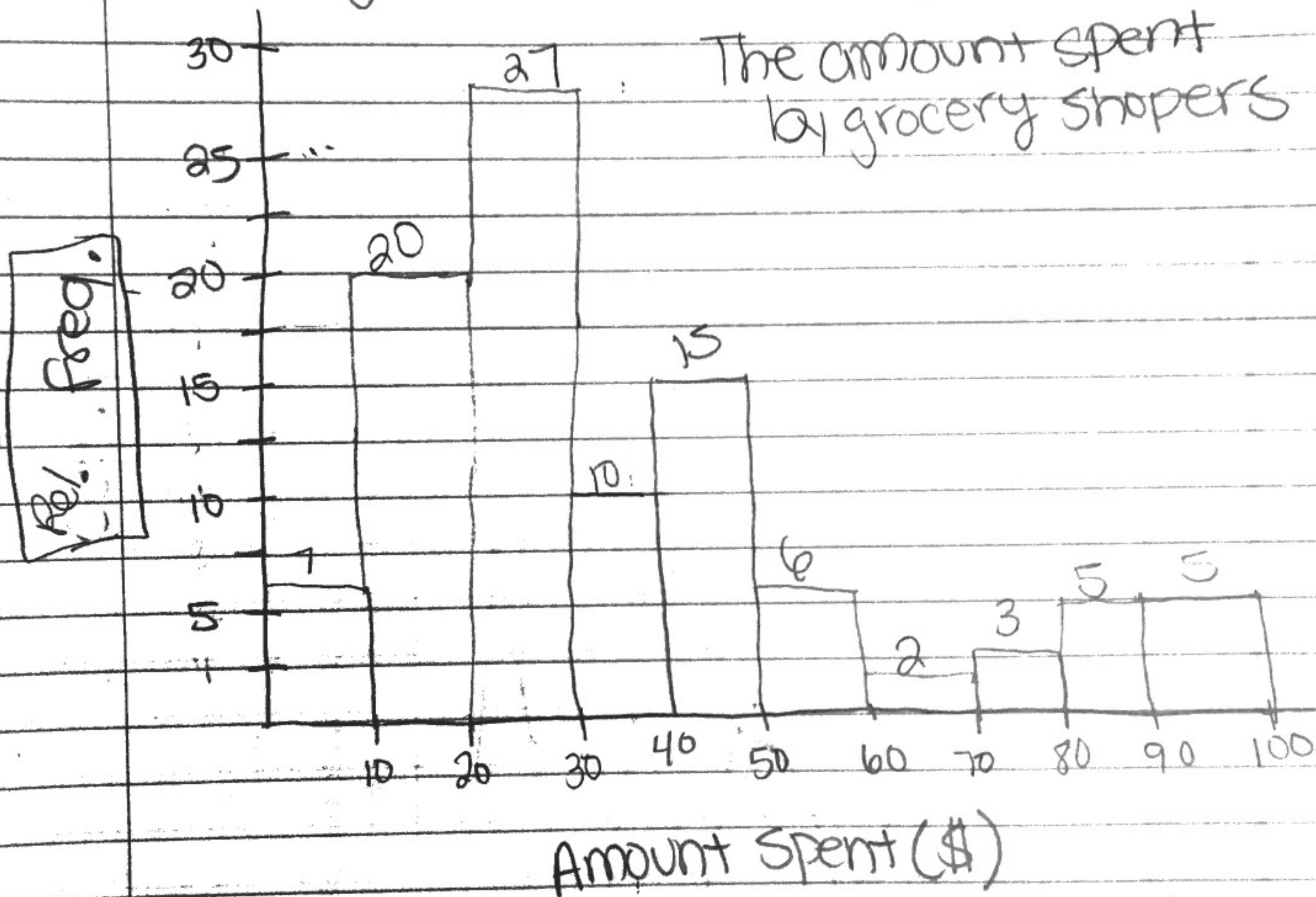
- 1.20 (a) The center corresponds to the 50th percentile. Draw a horizontal line from the value 50 on the vertical axis and determine the point on the ogive where the line intersects the ogive. Then draw a vertical line from this point to the horizontal axis. The line intersects the axis at approximately \$28. Thus, \$28 is the estimate of the center.
- (b) The 20th percentile.

see next page for another way to do C.



1.20C (another way to think of)

The directions for 1.20C state to construct a histogram from the ogive (as seen in book answers). Lets take this a step further. Does this really give us useful information? Lets construct a histogram from cum. freq. (not rel. cum. freq.)



*Note: Remember, there is a difference in Rel. cum. freq. and cum. freq.

From ogive to Rel. cum. freq.: Keep the same X-axis. Graph the height of each point to the y-axis (as seen in book answers) Your #'s may be slightly different.

To graph ogive to cum freq histogram:

Graph the same x-axis. Then graph the difference between each 2 intervals. (as seen in my answer for 1.20c)

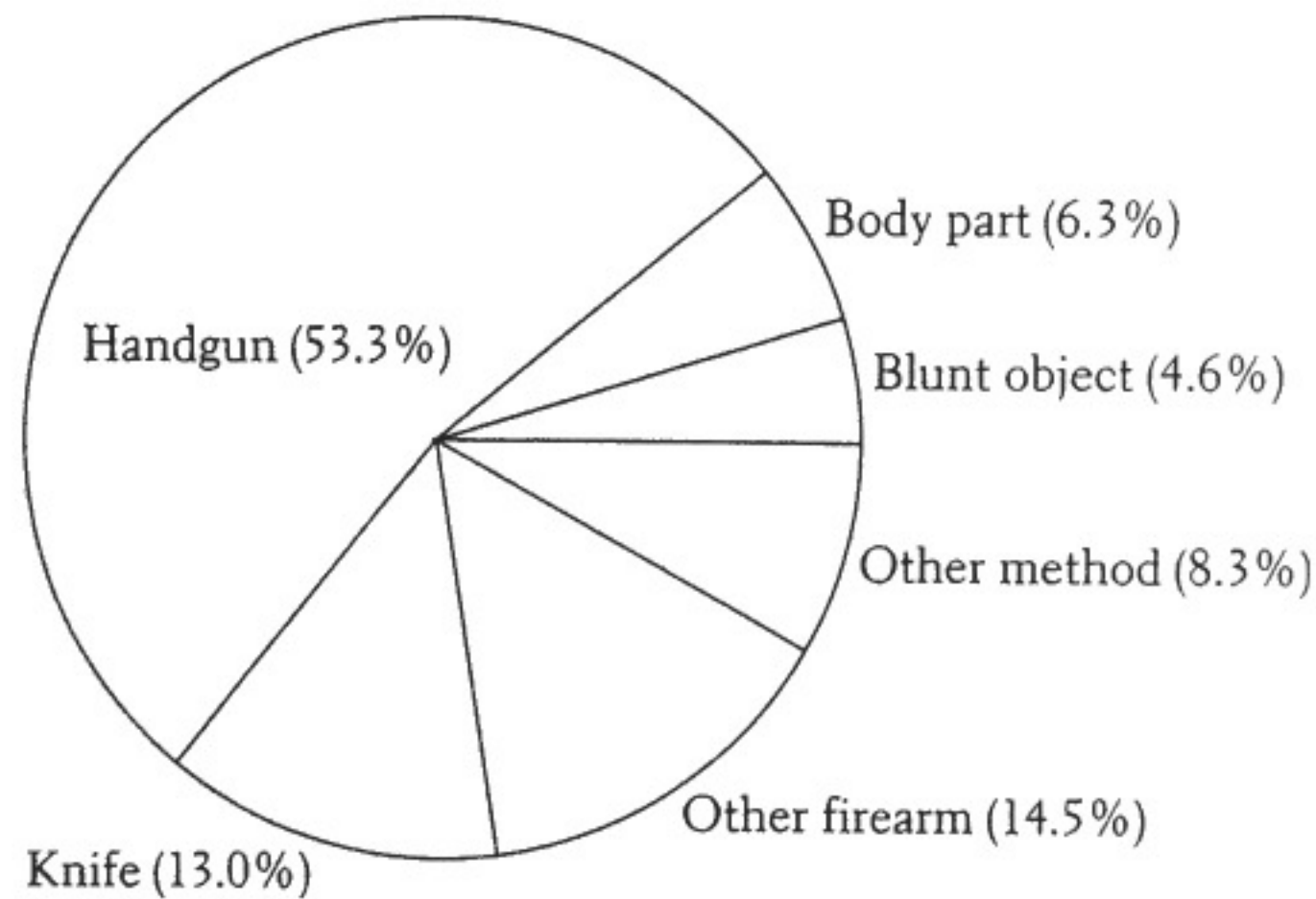
* Sorry for any confusion! Hopefully this will help :)



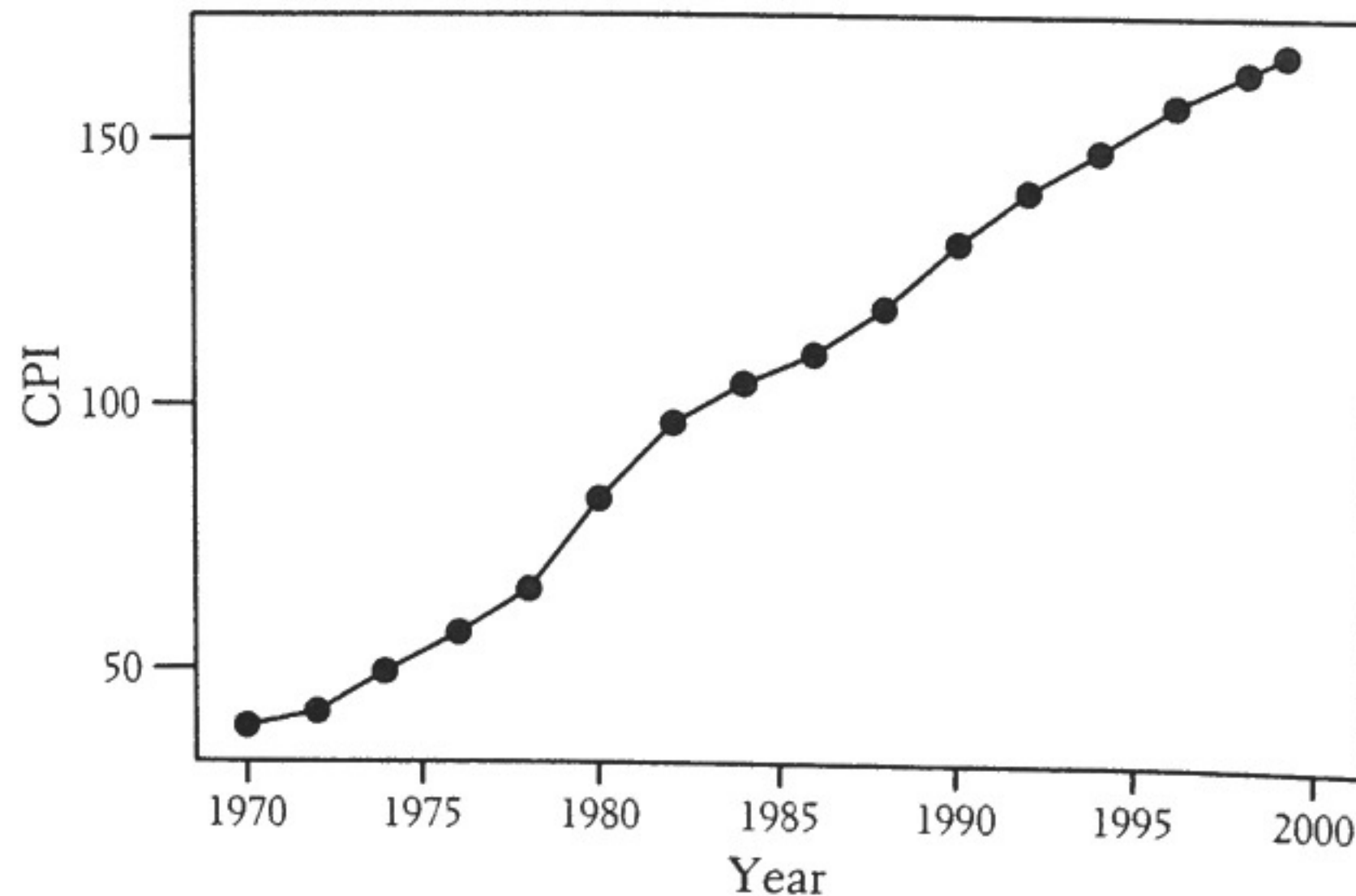
1.23 Gender, party voted for: Categorical
 Age, income: Quantitative

1.24 (a) Car makes: a bar chart or pie chart. Car age: a histogram or stemplot. (b) Study time: a histogram or stemplot. Change in study hours: a time plot (average hours studied vs. time). (c) A bar chart or pie chart.

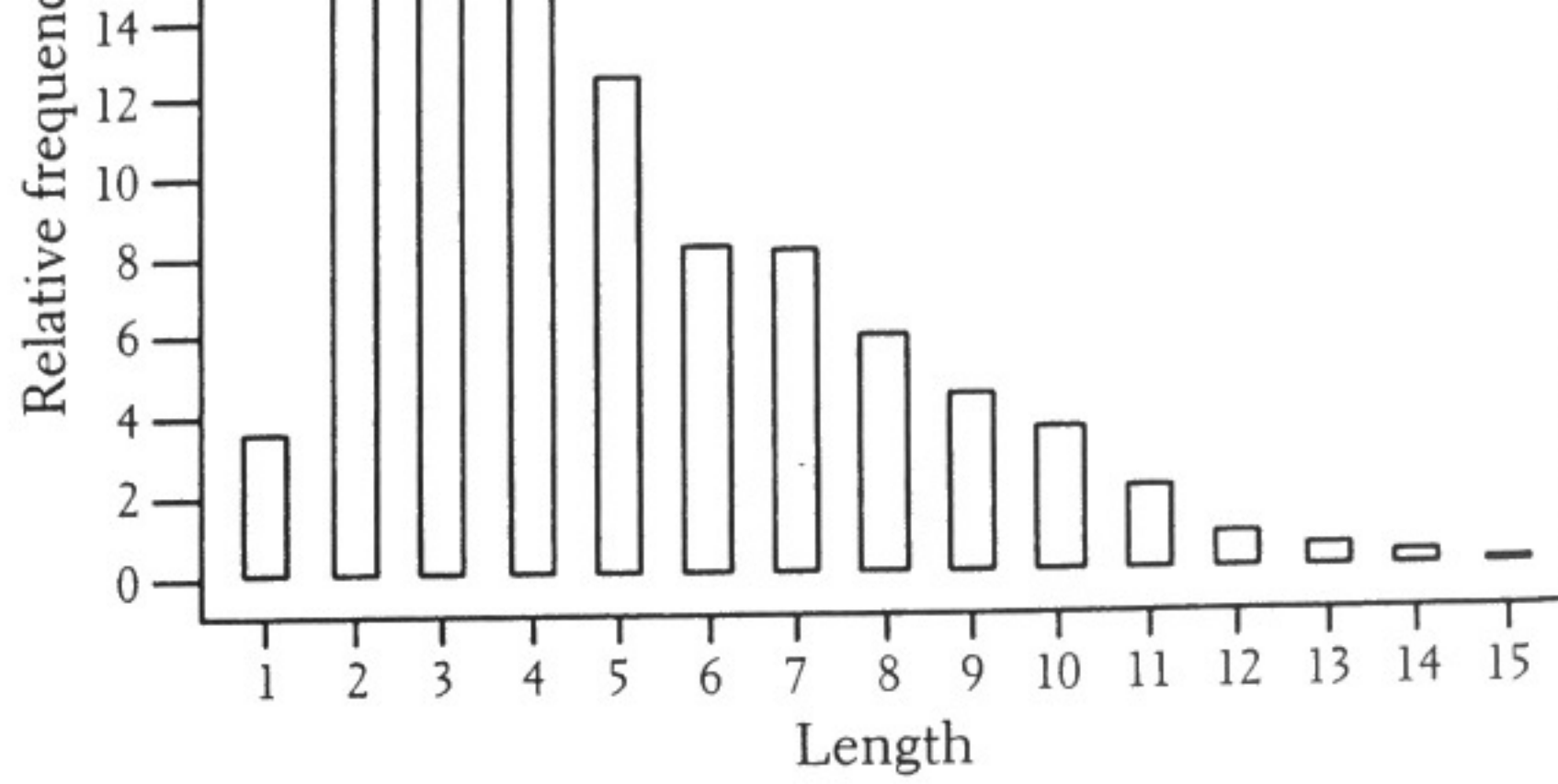
1.25 An "Other Methods" plot is needed because the sum of the percentages for the other categories is less than 100%.



1.26 (a)



(c) Prices rose steadily during this period. There was no reversal of this trend in any of the periods under study.



The distribution is skewed to the right with a single peak. There are no gaps or outliers.

(b) Shakespeare was somewhat more likely to use short words and somewhat less likely to use extremely long words than *Popular Science*. However, the distributions have strongly similar shapes.

1.28

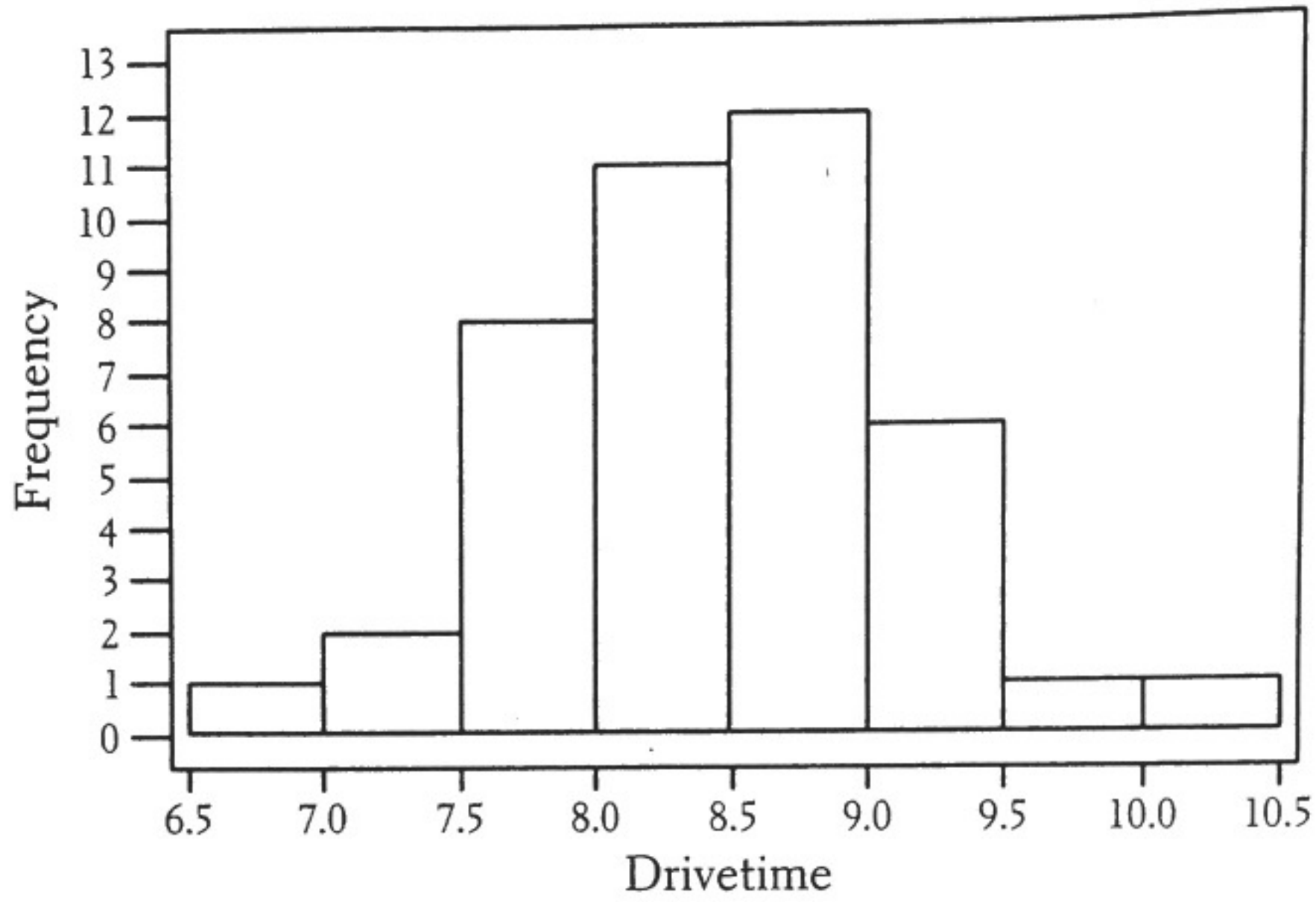
48	8
49	
50	7
51	0
52	6799
53	04469
54	2467
55	03578
56	12358
57	59
58	5

Stem = first two digits Leaf = last digit.

The distribution is roughly symmetric with one value (4.88) that is somewhat low. The center of the distribution is between 5.4 and 5.5.

Based on the plot, we would estimate the Earth's density to be about halfway between 5.4 and 5.5.

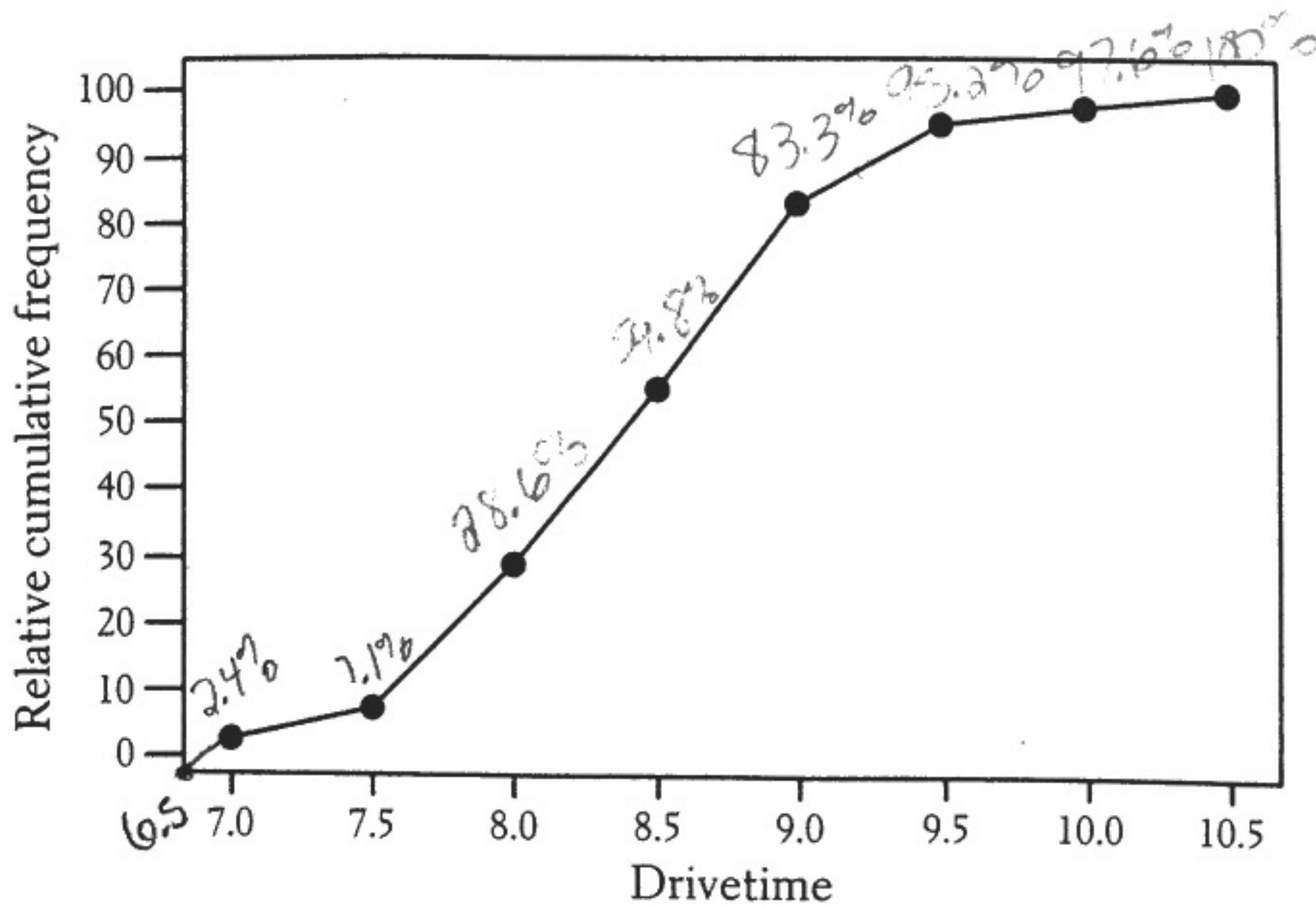
1.29 (a)



The distribution is roughly symmetric with no clear outliers.

(b)

Drivetime	Cum. freq.	Rel. cum. freq.
6.5 - 7.0	1	2.4%
7.0 - 7.5	3	7.1%
7.5 - 8.0	12	28.6%
8.0 - 8.5	23	54.8%
8.5 - 9.0	35	83.3%
9.0 - 9.5	40	95.2%
9.5 - 10.0	41	97.6%
10.0 - 10.5	42	100%



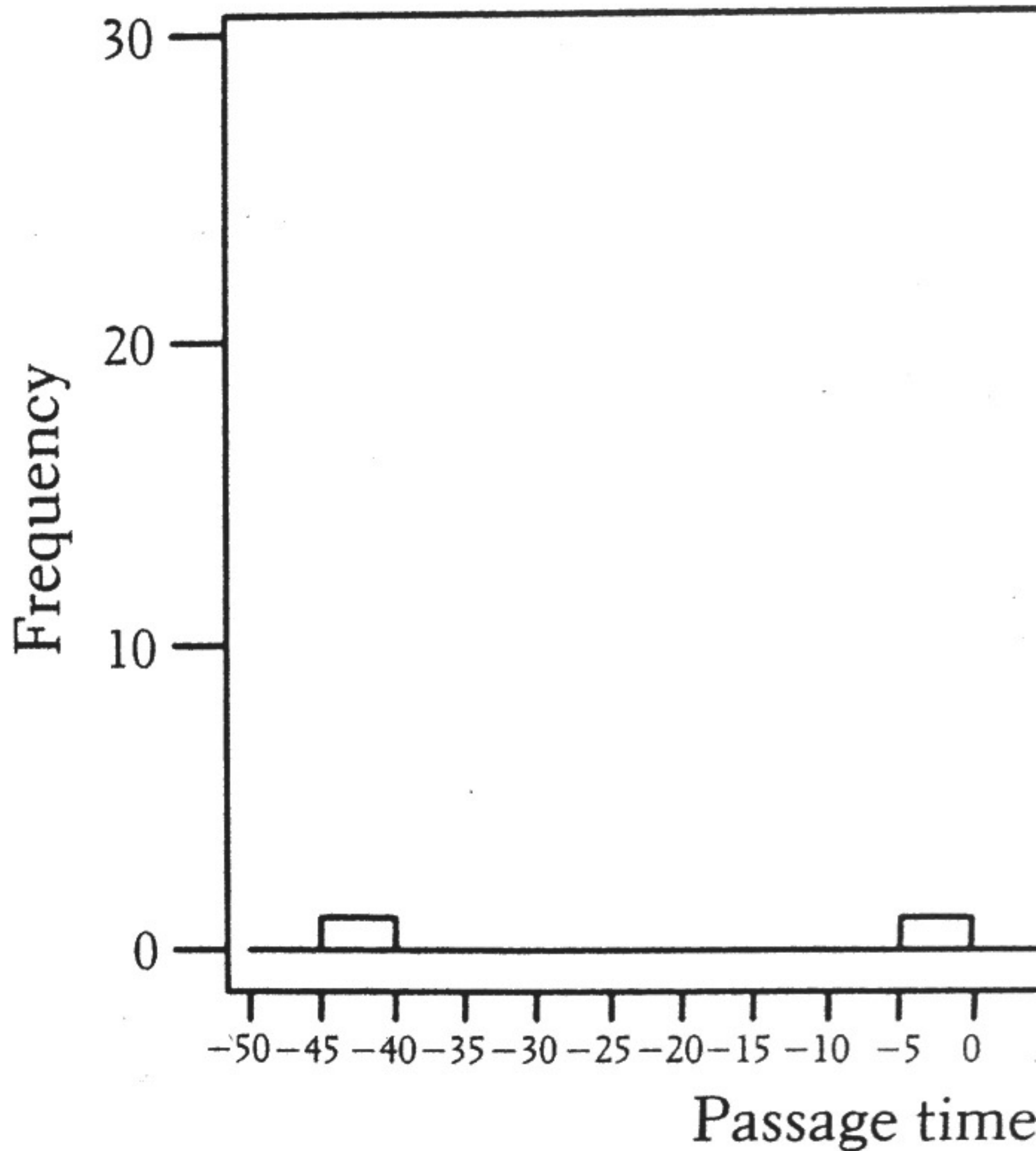
Exploring Data

1.29

(c) Center ≈ 8.5 , 90th percentile ≈ 9.4

(d) $8.0 \approx 28$ th percentile

1.30 (a)



A stemplot would have much the same appearance
what less practical because of the large number of