

1.14



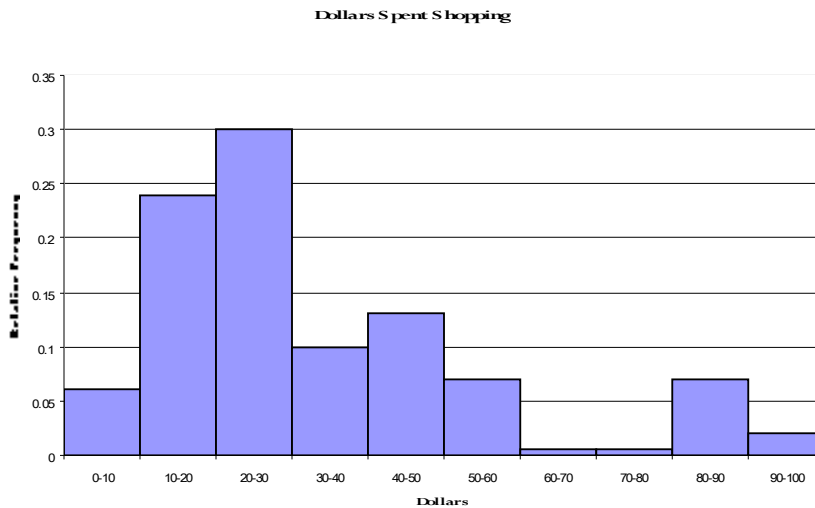
The distribution of CEO salaries is slightly skewed to the right. It is centered at around 360 to 400 thousand dollars. The salaries range from \$21,000 to \$1.103 million. The \$1.103 million is a possible outlier.

1.16

- a) The distribution of total returns is very symmetric.
- b) The distribution is centered around 15% total return.
- c) The smallest percent total return was about -65% and the largest was about 105%.
- d) About $11 + 5 + 3 + 4(1) \% = 23\%$ of the stocks lost money.

1.20

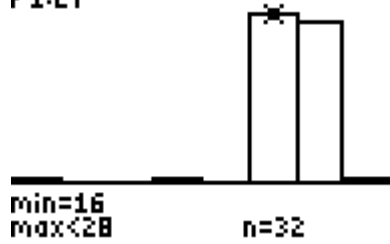
- a) The center is at about \$30. \$30 is the 50th percentile, which means half the customers spent more than this amount, while half spent less.
- b) A shopper who spent \$17 would be in about the 20th percentile.
- c)



1.30

(b) Here is a histogram of the data.

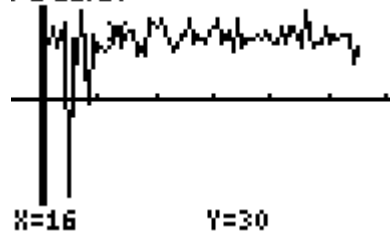
F1:L1



(c) The distribution of the deviations from 24800 ns is skewed left because of possible outliers at -2 and -44 ns. The center is around 27-28 ns.

(d) Below is a timeplot of the data, where X = the number of measurement (1 to 66) and Y = the deviation in ns from 24800.

F1:L2:L1



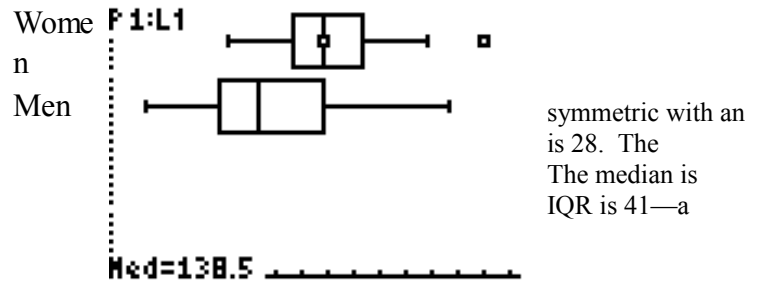
(e) The timeplot shows that as time progressed, Newcomb's measurements became more and more consistent. That is, there was less variation in his measurements.

1.36

(a) Here are the side-by-side boxplots.

(c) The distribution for the women is pretty outlier at 200. The median is 138.5 and the IQR distribution for the men is a bit skewed right. 113.5, notably lower than the women's, and the much larger spread than the women's scores.

(f) Here are the numerical summaries.



	\bar{x}	s	Min	Q_1	Med	Q_3	Max	IQR
Women	141.06	26.44	101	126	138.5	154	200	28
Men	121.5	32.85	70	98	114.5	143	187	45

1.40

$\bar{x} = 5.4$ and $s_x = 0.642$ mg/dL

Without using 1-Var Stats we could type something like

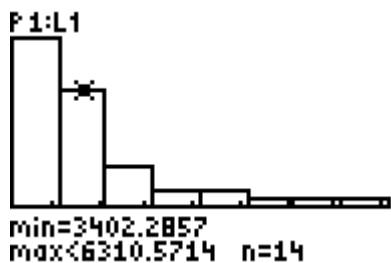
```
√(sum((L1-mean(L1))^2)/5)
.6418722614
```

- 1.43 (a) Here are four numbers with 0 standard deviation: $\{1, 1, 1, 1\}$
- (b) Here are four numbers with the largest standard deviation: $\{0, 0, 10, 10\}$
- (c) There are definitely more choices for (a) but not for (b).

- 1.44 (a) X = the lengths of the cockroaches, so $\bar{x} = 1.5$ inches, and $s_x = 0.4358$ inches.
- (b) $\bar{x} = 3.81$ cm, and $s_x = 1.107$ cm.
- (c) I would estimate that the average length is 3.81 cm. This is not a very good estimate because of the small sample size.

- 1.68 Since the mean is influenced by the presence of outliers and strong skewing in a distribution we can guess that the distribution of jury awards is strongly skewed to the right by a few very high award values. Most of the awards are closer to \$8000, but perhaps a very few of the awards are very high (maybe in the millions). This would cause a large difference in the mean and median values.

- 1.70 Here is a histogram of the population of the states (in 1000's).

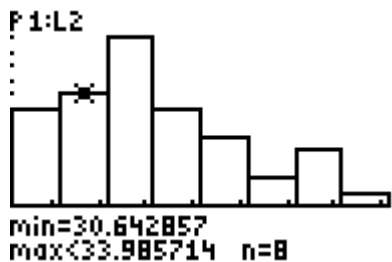


The distribution of the population is strongly skewed right with median 3731.5 thousand and IQR = 5055 thousand. This shape is not surprising since we have a few very large states and many “small” ones. The boxplot shows that the states with 15982, 18976, and 20852 thousand people are outliers. These are Florida, New York, and Texas.

- 1.71 The stemplot of the percent of students taking the SAT is shown below. The stemplot shows that the distribution is multi-peaked, and roughly skewed right. The median is 33% and the mean is 37.8%. The center is marked roughly on the plot below. In fact, none of our data are at the center, so the numerical measures are not useful for estimating averages for this variable.

Stem	Leaf
0	4455578888999
1	1222368
2	15
3	24 ← Center
4	
5	2233
6	1135578
7	2678
8	0
3	2 = 32%

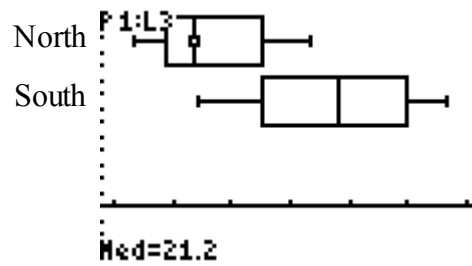
- 1.72 A histogram of average teachers' salaries (in \$1000's) is shown below. This distribution is pretty symmetric with no apparent outliers. The mean is \$5101.5 thousand, and the standard deviation is \$4840.8 thousand.



1.73

Here is the list of percentages of students not completing high school in the two regions with comparative boxplots. Both distributions are roughly symmetric, and numerical summaries are given in the table.

	North East	Southern
	20.8	22.5
	21.2	26.9
	20	25.6
	17.8	29.1
	23.3	35.4
	25.2	21.6
	25.3	35.7
	28	30
	19.2	31.7
		32.9
		24.8
		34
mean	22.31	29.18
st. dev	3.34	4.9



It appears that more students in the southern states do not finish high school than in the northeastern states.