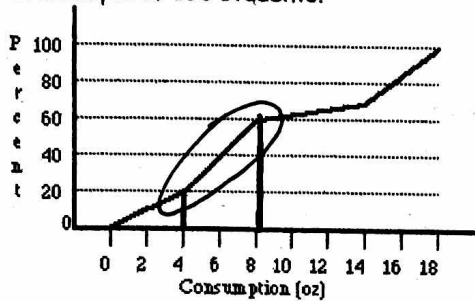


1. You measure the age, marital status, and earned income of an SRS of 1463 women. The number and type of variables you have measured is

- (a) 1463; all quantitative.
- (b) four; two categorical and two quantitative.
- (c) four; one categorical and three quantitative.
- (d) three; two categorical and one quantitative.
- (e) three; one categorical and two quantitative.**

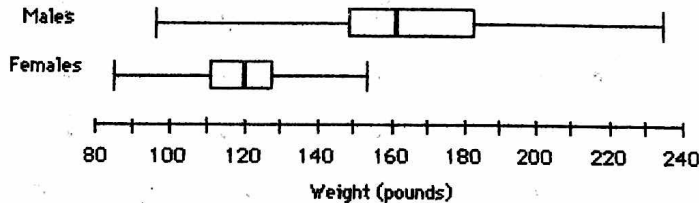
2. The following is an ogive on the number of ounces of alcohol (one ounce is about 30 mL) consumed per week in a sample of 150 students.



A study wished to classify the students as "light", "moderate", "heavy" and "problem" drinkers by the amount consumed per week. About what percentage of students are moderate drinkers, i.e., consume between 4 and 8 ounces per week?

- (a) 60%
- (b) 20%
- (c) 40%**
- (d) 80%
- (e) 50%

3. The weights of the male and female students in a class are summarized in the following boxplots:



Which of the following is NOT correct?

- (a) About 50% of the male students have weights between 150 and 185 lbs.
- (b) About 25% of female students have weights more than 130 lbs.
- (c) The median weight of male students is about 162 lbs.
- (d) The mean weight of female students is about 120 because of symmetry.
- (e) The male students have less variability than the female students.**

4. The test grades for a certain class were entered into a Minitab worksheet, and then "Descriptive Statistics" were requested. The results were:

MTB > Describe 'Grades'. *mean < median*

	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN	MIN	MAX	Q1	Q3	IQR
Grades	28	74.71	76.00	75.50	12.61	2.38	35.00	94.00	68.00	84.00	16

You happened to see, on a scrap of paper, that the lowest grades were 35, 57, 59, 60, ... but you don't know what the other individual grades are. Nevertheless, a knowledgeable user of statistics can tell a lot about the dataset simply by studying the set of descriptive statistics above.

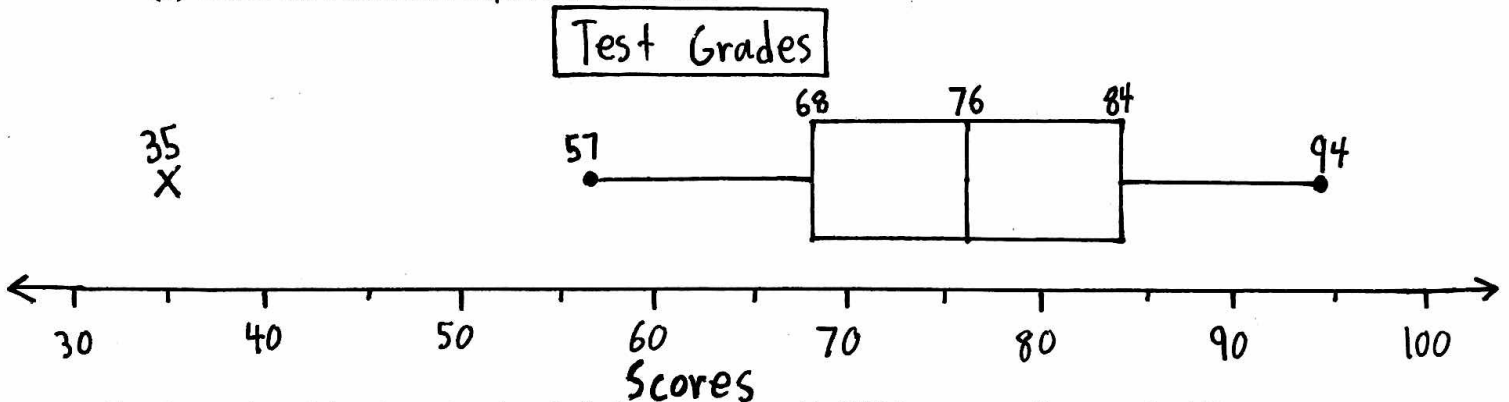
(a) Write a brief description of what the results tell you about the distribution of grades. Be sure to address:

- the general shape of the distribution
- unusual features, including possible outliers
- the middle 50% of the data
- any significance in the difference between the mean and the median

• The distribution of the grades is *roughly* ~~approximately~~ symmetric since the mean and the median are very close. The center is the mean (74.71) and the spread is the standard deviation (12.61). There is one outlier at 35. (could say skewed left)

- Low outliers: $Q_1 - 1.5(IQR) \rightarrow 68 - 1.5(16) = 44$
- Higher outliers: $Q_3 + 1.5(IQR) \rightarrow 84 + 1.5(16) = 108$

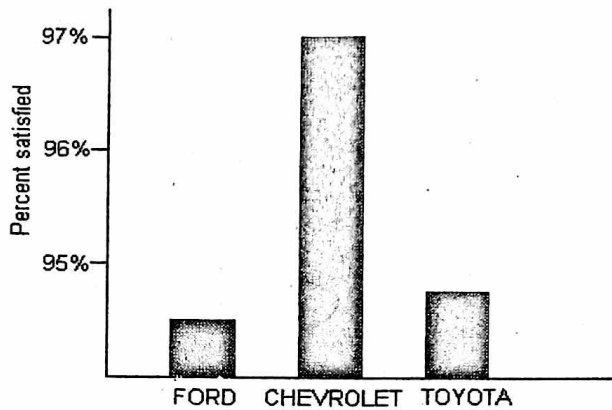
(b) Construct a modified boxplot for these data.



5. A reporter wishes to portray baseball players as overpaid. Which measure of center should he report as the average salary of major league players?

- (a) The mean.
- (b) The median.
- (c) The mode.
- (d) Either the mean or median. It doesn't matter since they will be equal.
- (e) Neither the mean nor median. Both will be much lower than the actual average salary.

6. The following bar graph gives the percent of owners of three brands of trucks who are satisfied with their truck.



From this graph we may legitimately conclude

- (a) Owners of other brands of trucks are less satisfied than the owners of these three brands.
 - (b) Chevrolet owners are substantially more satisfied than Ford or Toyota owners.
 - (c) There is very little difference in the satisfaction of owners for the three brands.
 - (d) Chevrolet probably sells more trucks than Ford or Toyota.
 - (e) A pie chart would have been a better choice for displaying this data.
7. The mean salary of all female workers is \$35,000. The mean salary of all male workers is \$41,000. What must be true about the mean salary of all workers?
- (a) It must be \$38,000.
 - (b) It must be larger than the median salary.
 - (c) It could be any number between \$35,000 and \$41,000.
 - (d) It must be larger than \$38,000.
 - (e) It cannot be larger than \$40,000.
8. During the early part of the 1994 baseball season, many sports fans and baseball players noticed that the number of home runs being hit seemed to be unusually large. Here are the data on the number of home runs hit by American and National League teams.

American League 35, 40, 43, 49, 51, 54, 57, 58, 58, 64, 68, 68, 75, 77

National League 29, 31, 42, 46, 47, 48, 48, 53, 55, 55, 63, 63, 67

Title: Homeruns in 1994

- (a) Construct an appropriate graph for comparing the number of home runs hit in the two leagues.

American Leaf	Stem	National Leaf
	2	9
5	3	1
9 3 0	4	2 6 7 8 8
8 8 7 4 1	5	3 5 5 5
8 8 4	6	3 3 7
7 5	7	

key:
 2/9 = 29 hrs (National)
 5/3 = 35 hrs (American)

- (b) Calculate numerical summaries of the number of home runs hit in the two leagues. Which of these numbers would be most appropriate for comparing the two leagues? Explain.

American: $\bar{x} = 56.93 \checkmark$	National: $\bar{x} = 50.14 \checkmark$
med = 57.5	med = 50.5
$S_x = 12.69 \checkmark$	$S_x = 11.13 \checkmark$

- (c) Write a few sentences comparing the distributions of home runs in the two leagues.

On average, the American League teams hit many more homeruns in 1994. There was also greater variability in the American League data.