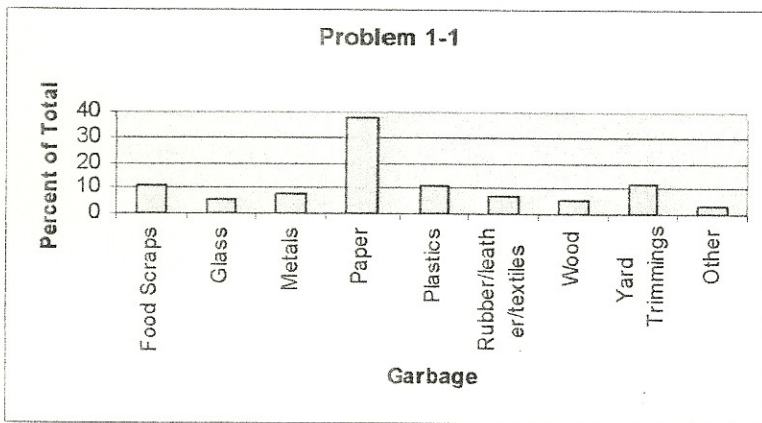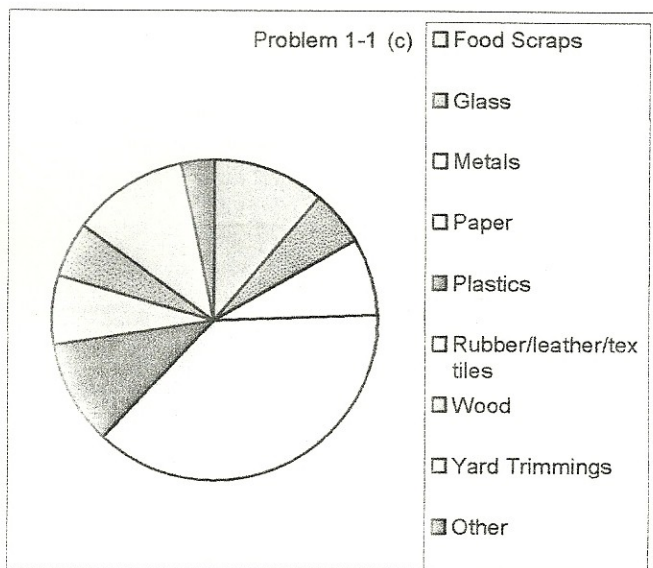## CH1H1: 1, 3, 5, 10, 11

1.   a) The sum of the weights is 231.8 million tons. (This is different from 231.9 because of roundoff error)

b)

c)

3. a) The stemplot gives a better display of the shape of the data...the dotlot is too spread out to notate any shape.

b) The numbers in the left hand column represent the cumulative counts of the observations from the bottom up and the top down. The (3) in the far left column represents the "middle value"

c) Dotplot: the second dot from the left. Stemplot: 0.99 The reason is that these numbers represent the final concentrations "as a multiple of their initial concentrations"...therefore the data point that is closest to the original concentration is the data point closest to 1, which would be 0.99.

5. A stemplot, dotplot or histrogram is an acceptable graph. Here is what the stemplot looks like:

```
1 | 445899
2 | 255667789
3 | 13344555589
4 | 00112345667789
5 | 1224
```

The data is fairly symmetric with a center of about 35. The range is 38…the data ranges from 14 to 54. There are no gaps or outliers.

10. (Shakepeare) The distribution of words in Shakespeare's play is skewed to the right with a center of about 5 or 6 and the range is 11 letters. There are no gaps or outliers.

11. (Presidential Ages)
a) Histogram

b) The distribution is symmetric with a single peak at the center of about 55 years. The range is 27 and there are no outliers.

c) The youngest president was Teddy Roosevelt and the oldest was Ronald Reagan

d) He was among the younger presidents but was not an outlier.

CH1H2: 12, 13, 14, 21, 22, 23, 26

12. a) Stemplot or histogram would be good choices.

```
0 | 000002
1 |
2 |
3 | 89999
4 | 012236678
5 | 12
```

b) The distribution has two clusters of data, one around 0 grams of sugar (the diet drinks) and the other around 40 grams with a gap in the 10's and 20's. The center is in the 30's and the range is 52.

13.
    a) The center of the ogive is where the 50% dot is. Find 50% on the y-axis and go horizontally over to the ogive and then go vertically down to the answer: 30.

    b) About 18%

    c) Histogram

14. Rounding to the nearest ten (157 rounds to 160) and doing a regular stemplot yields the following where 1|4 = 140

```
0 | 8
1 | 0001345555677
2 | 067
3 | 6
4 |
```

The book suggests "splitting stems". You can do this, where you "split" the "1"'s into two stems and put half of the "1" numbers on one stem and the rest on another stem. The problem with this is that you lose the visual of the high peak in the 100's.

b) ogive

c) histogram

d) The percentage of glucose levels between 90 and 130 is about 25%. The center is about 148mg/dl and the relative cumulative frequency associated with 130 is 27.78%.
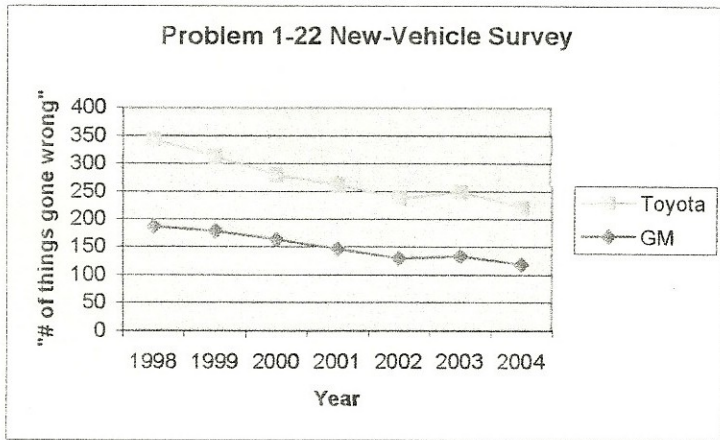
21.

Problem 1-21 College Costs
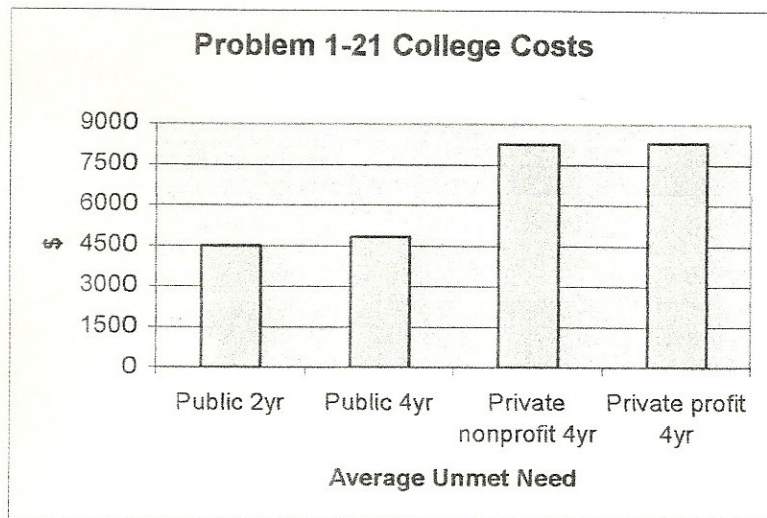
A pie chart would be incorrect because these numbers do not represent parts of a whole.

22.

Problem 1-22 New-Vehicle Survey

Both timeplots suggest that the manufacturers have reduced the "things gone wrong" over time which means they have improved.

23.   a) Alaska: 5.7% and Florida: 17.6%
       b) Without the outliers, the shape is fairly symmetric, with a slight skew to the left and the percentages range from 8.5% to 16.6%

26. Matching: 1c, 2b, 3d, 4a

Ch1H3: 27, 28, 29, 32, 33, 34, 35, 37

27.   a) Mean = $\bar{x} = 85$
       b) Final quiz average = 79.33. The mean is affect by outliers. (Mean is NONresistant)
       c) Bar chart if using categories (A, B, C, D, F) or histogram if using numbers (93-100, etc) (Make histogram)

28.

10	139
11	5
12	669
13	77
14	08
15	244
16	55
17	8
18	
19	
20	0

a) Possibly an outlier at 200. The median is 138.5 and the range is 77 without the 200.
b) The mean is 141.056
c) The mean is larger than the median because the outlier affects the mean...skews it towards the outlier.

29. The annual payroll for the baseball players is 1.2 million x 25 = 30 million. No, you would not be able to answer this question knowing the median.

32. a) Bonds' mean: 37.0   Bonds' median: 37.0 Since these are so close, we would expect the distribution to be symmetric.

b)

```
1 | 69
2 | 455
3 | 334477
4 | 0255669
5 |
6 |
7 | 3
```

c) His typical number of home runs is in the high 30's. 2001 was a very unusual year.

33. a) Side by side boxplots

b)   Women 5# summary: 101, 123.25, 138.5, 156.75, 200
       Men 5# summary: 70, 95, 114.5, 144.5, 187

c) Women generally score highER than men (all 5# in the 5# summary are largER). The men's data is more spread out than the women's. The shapes are fairly similar, both being slightly skewed to the right.

34. a) Since the shape of the data is roughly symmetric, we would expect the mean to be roughly the same as the median.

b) 5# summary: 42, 51, 54, 59, 69. The median (54) is very close to the mean (54.805)

c) The range of the middle 50% in the IQR which is 59-51 = 8.

d) Boxplot

e) Calc question

35. The IQR is a resistant measure of spread.
EX. Data set A (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) and Data set B (1, 2, 3, 4, 5, 6, 7, 8, 9, 50)
IQR of Data set A: 8 – 3 = 5   IQR of Data set B: 8 – 3 = 5.
The IQR is NOT affected by the outlier (50).

37. a) $Q_1 = 25$ $Q_3 = 45$
b) 1.5(IQR) = 1.5 (45-25) = 1.5(20) = 30. $Q_3 + 30$ = 75. Therefore 73 is NOT an outlier.

39.   a) $\bar{x} = 5.4$ mg/dl

b)

X	$(x-\bar{x})$	$(x-\bar{x})^2$
5.6	0.2	0.04
5.2	-0.2	0.04
4.6	-0.8	0.64
4.9	-0.5	0.25
5.7	0.3	0.09
6.4	1.0	1
		$\sum(x-\bar{x})^2 = 2.06$
		$2.06 \div 5 = 0.415$
		$\sqrt{0.415} = 0.6419$ mg/dl

c) Check on Calculator using Stat/Calc/1-Var Stats

42. An example of a set of numbers with a median of 10 and a mean of 7:
What I think: The middle number has to be 10 and the numbers have to add up to 35 (since 35/5 = 7)
EX. 1, 1, 10, 10, 13

45. If this problem was challenging...make up numbers and try it out!
Before the raise: (30000, 40000, 60000)   After the raise: (31000, 41000, 61000)
a) The mean will rise by $1000 and the median will rise by $1000.
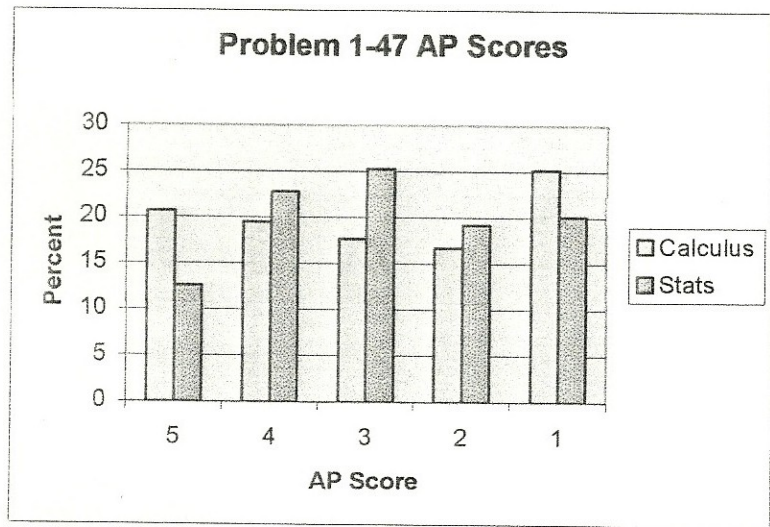
b) No, the IQR will remain the same

c) No, the standard deviation will remain the same.

46. Yes, both measures of spread will change: TRY IT WITH NUMBERS!
Before the raise: (30000, 40000, 60000)   After the raise (31500, 42000, 63000)

47.

Problem 1-47 AP Scores

You could have made two separate bar charts.
The stats scores are more bell shaped than the calculus scores. The center of the stats scores is slightly higher than the calculus scores. The range is the same.