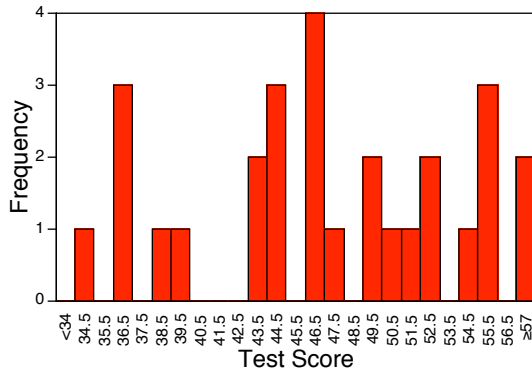


Chapter 3-Displaying Data

3.1 Katz et al (1990) No Passage Group:



There is too little data to say very much about the shape of this distribution, but it certainly isn't looking normally distributed.

3.3 I would use stems of 3*, 3., 4*, 4. 5*, and 5. for this display.

3.5 Compared to those who read the passages:

- a) Almost everyone who read the passages did better than the best person who did not read them. Certainly knowing what you are talking about is a good thing (though not always practiced).

b)

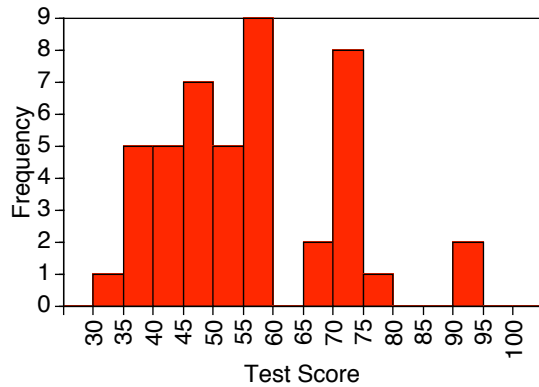
	4	3*	
	68966	3.	
	44343	4*	
	6669697	4.	
	42102	5*	
	57557	5.	5669
		6*	
		6.	66
		7*	21232231
		7.	5
		HI	91 93

Notice that I have entered the data in the order in which I encountered them, rather than in increasing order. It makes it easier.

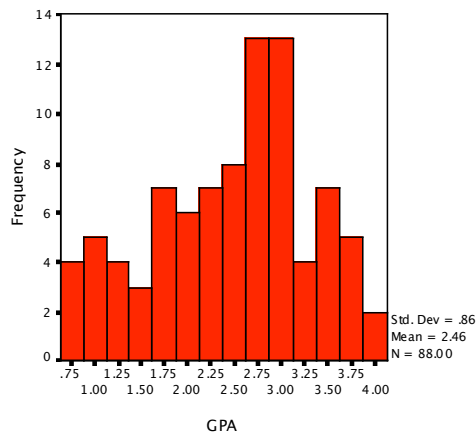
c) It is obvious that the two groups are very different in their performance. We would be worried if they weren't.

d) This is an Internet exercise with no fixed answer. That source is far more advanced than the students would be at this time, but I think that they should be able to read it if they just skip over what they don't understand.

3.7 A bimodal set of data can be created by combining the two groups from the Katz et al. study. The NoPassage group's scores will be well below the Passage group's scores.



3.9 Histogram of ADDSC scores



3.11 (1) Mexico has very many young people and very few old people, while Spain has a more even distribution. (2) The difference between males and females is more pronounced at most ages in Spain than it is in Mexico. (3) You can see the high infant mortality rate in Mexico.

3.13 The distribution of those whose attendance is poor is far more spread out than the distribution of normal attendees. This would be expected because a few very good students can score well on tests even when they don't attend, but most of the poor attenders are generally poor students who would score badly no matter what. The difference between the average grades of these two groups is obvious.

3.15 The stem-and-leaf displays in Exercise 3.14 support the sequential processing hypothesis.

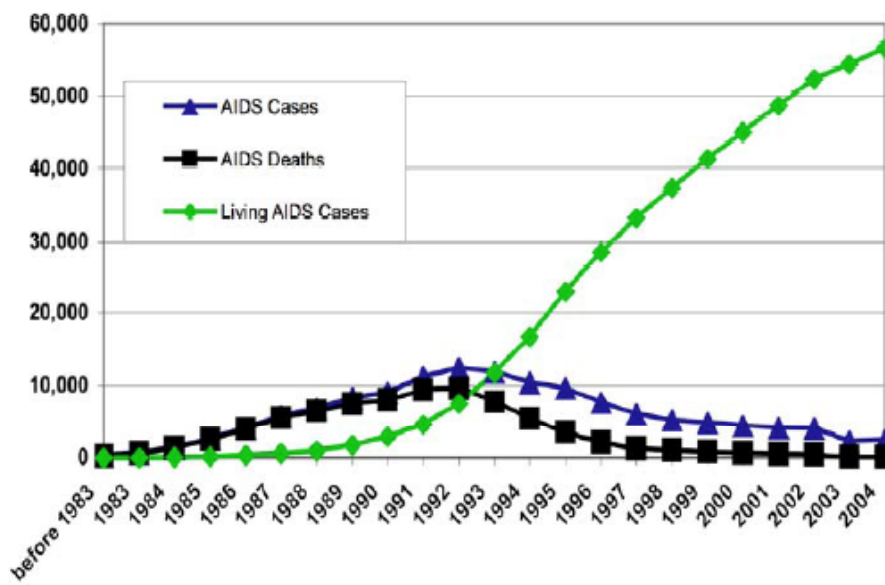
3.17 The data points are probably not independent in that data set. As time went on, there would be changes in the subject's performance. At first he might get better with practice,

but then fatigue would start to set in. Since the data are given in the order in which they were collected, at least within each condition, data nearer in time should be more similar than data farther apart in time.

3.19 The amount of shock that a subject delivers to a white participant does not vary as a function of whether or not that subject has been insulted by the experimenter. However, the black participants do suffer more shocks when the subject has been insulted.

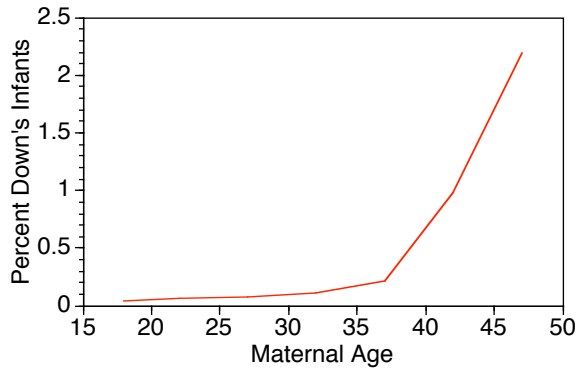
3.21 These data graph the response variable against time.

AIDS Cases, Deaths and Living AIDS Cases in California through 2004



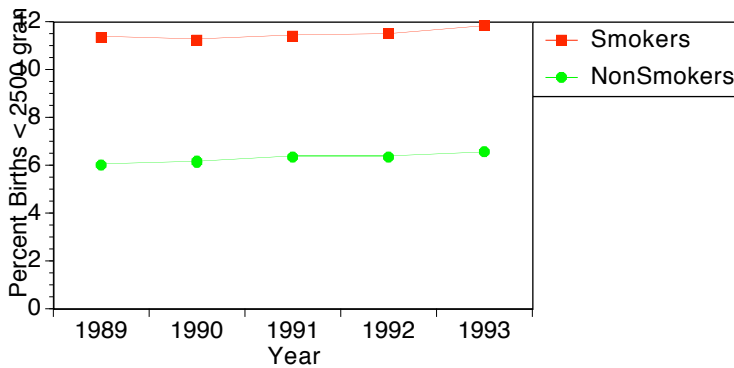
It is clear that the number of AIDS cases in California began to drop about 1992, with an even steeper drop in deaths due to AIDS. (There were very few deaths in 2004.) But notice the dramatic increase in the number of people who continue to live with AIDS, primarily due to new drug therapies.

3.23 There is a tremendous increase in Down's Syndrome in children born to older mothers. This increase doesn't really take off until mothers are in their 40s, but with parents delaying having children, this is a potential problem.



3.25 Smoking and low birthweight:

The data are given as the percentage of births for each group that were less than 2500 grams.



The relationship is unlikely to be a fluke because it is so consistent year after year. You can see that within each group there is very little variability.

3.27 White females have a longer life expectancy than black females, but the difference has shrunk considerably since 1920, though recent changes have been modest.