1. Screen Example: some data on age at death

2. Chapter 3. A Picture is Worth a 1000 words

- a. Make a Picture
 - i. How? -1^{st} count and sort get frequencies (example)
 - ii. Relative Frequency Table (uses percentages in addition to counts)
- b. Why?
 - i. To describe the DISTRIBUTION of data
 - ii. Better organization
- c. Area Principle
 - i. Basically, be careful when you graph numbers (p. 16). Properly, if a value or a range of values makes up 20% of the total, it should also make up 20% of a graph (to be fair).

3. Contingency Tables (pp. 18-19)

Sometimes, we want to examine two or more categorical variables simultaneously. We construct what is called a "two-way table" or a "crosstabulation" or a "contingency table". An example:

Q. Do you think things in this country are generally going in the right direction or are they seriously off on the wrong track?

	Men	Women	Total
Right Direction	331	214	545
Off on the wrong track	285	421	706
Don't Know	46	55	101
Total	662	690	1352

The margins of the table are totals. The "marginals" give a frequency distribution (much like the relative frequency table above). The cells of the table give the counts for combinations of answers to the question and gender.

Often, tables like this are presented with percentages. While percentages give us a clearer picture, it can also complicate things. Example. These are called cell percentages:

		¥ ¥	
	Men	Women	Total
Right Direction	24.5%	15.8%	40.3%
Off on the wrong track	21.1%	31.1%	52.2%
Don't Know	3.4%	4.1%	7.5%
Total	49.0%	51.0%	100.0%

Sometimes folks like row percentages:

	Men	Women	Total
Right Direction	50.0%	31.0%	40.3%
Off on the wrong track	43.1%	61.0%	52.2%
Don't Know	6.9%	8.0%	7.5%
Total	100.0%	100.0%	100.0%

And yes, there are column percentages too:

	Men	Women	Total
Right Direction	60.7%	39.3%	100.0%
Off on the wrong track	40.4%	59.6%	100.0%
Don't Know	45.5%	54.5%	100.0%
Total	49.0%	51.0%	100.0%

Your choice of which percentages to use depends on the question being asked. We'll look at contingency tables again just before the first exam in this class.

4. Chapter 3 after page 19 is optional, pp. 23-24 are good for your general education

5. Chapter 4. Examining Quantitative Data with the histogram (page 37 of your text.)

- a. First sort the data in increasing order according to their values
- b. Choose class intervals, ideally, equal in size (easiest to work with)
- c. Count the number of observations that fall into each class interval
- d. Construct rectangles with area proportional to the relative frequencies (percentages) with which the observations fall in each interval
- e. All of the observations should be included (area totals 100%)

Back to our first dataset, what follows is a histogram of age.



Here is a histogram for a different set of age data, what does this tell you:



And one more histogram from yet another set of data, and what does this one tell you:



The data sources were: the UCLA area, the West Hollywood area around 1990 and the city of Beverly Hills.

6. Things to be aware of with respect to histograms

- Shape:
 - Unimodal, bimodal, multimodal, uniform
 - Symmetric or skewed? Tails?
 - *Exceptions? Outliers* and Gaps
- Center
- Spread
- *Remark:* the histogram can take all kinds of shapes.
- Usually histograms are interesting to compare and a single histogram is not usually interesting by itself

7. Time Plots (pp. 43-44) go ahead and skip this section

8. What Can Go Wrong? (pp. 45-46)

The whole point of chapters 3 and 4 is to help you convey information in a meaningful and truthful way using graphics and to help you understand that there are tools which can help you compare data from different sources