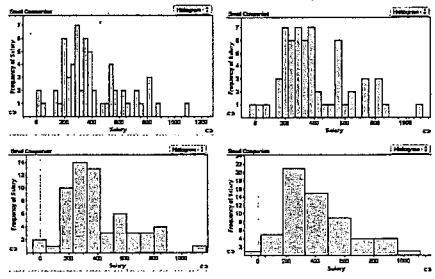
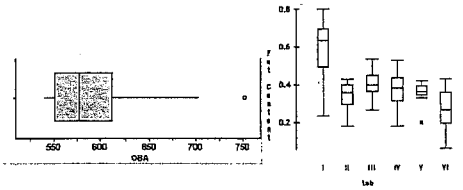
Data graphing facts

Histograms

Salaries of CEOs of small companies (60 data pts)



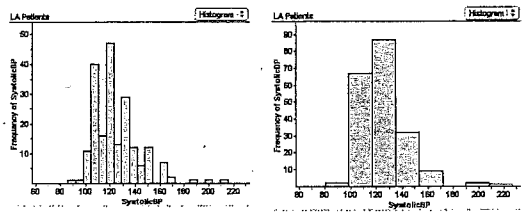
Box Plot (Box and Whisker Plot)



One data set

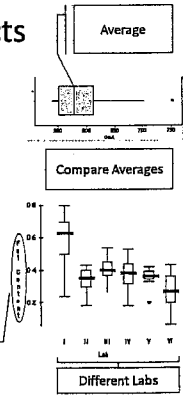
several related data sets

Systolic Blood Pressure: 200 data points

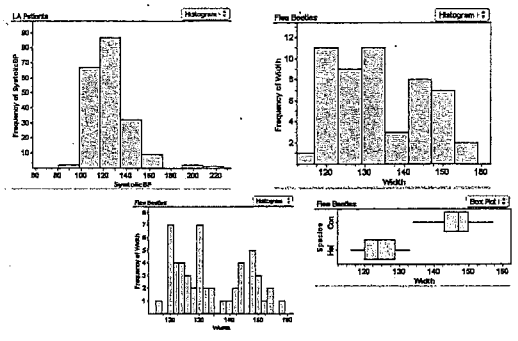


Box Plot Facts

- Your data **must** be numbers of the sort that finding an average is a sensible thing to do
- Box plots are **particularly useful** if you have several related data sets (where taking averages is sensible) that you would like to **compare**
- Usually data sets can be compared if the **same attribute** is being measured for **different populations**

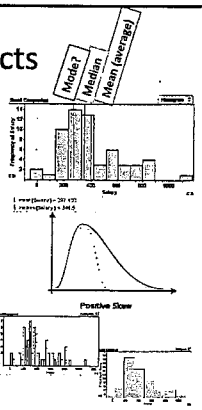


Unimodal and Bimodal Data



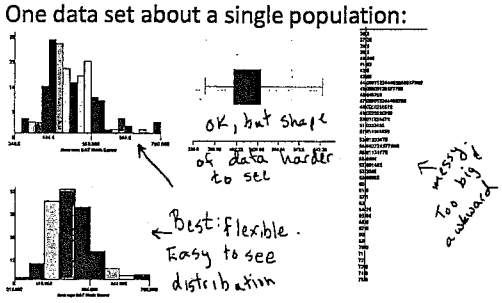
Histogram facts

- Histograms are appropriate for numerical data where finding an average is a sensible thing to do.
- Histograms are good for showing an overall shape, and for seeing if the data is skewed
- Histograms can show more or less detail depending on your choice of bar-width



Comparing numerical data graphs

One data set about a single population:



Stem and Leaf Plots or Stemplots

◀Quiz scores

Stem	Leaf
6	5 9
7	3 8
8	0 2 6 8
9	1 5

Class sizes

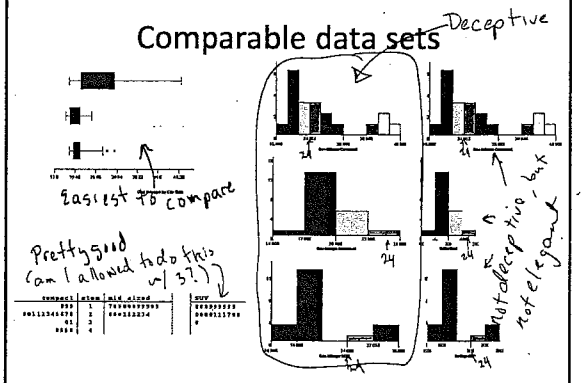
class sizes

Econ 101	stem	Pol 306
9	0	
3 4 5 6 6 7 9	1	4 6 7 8 8
0 1 1 2 5 5 6	2	0 0 4 9

key: "2|0" means "20"

Rounded		Truncated
7 3 1 0	1	0 0 3 6 9
4 0	2	3
8 2	3	2 7

Comparable data sets

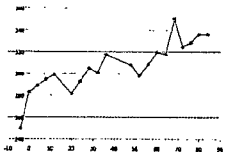


Stem and Leaf Plot Facts

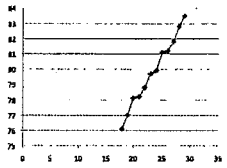
- Can be used only with numerical data for which finding an average makes sense
- All of the data numbers appear in the plot (though often rounded)
- Shows shape of the data distribution
- Easy to make
- **Most useful** if you want to see a **limited number of place values** and if you have a **moderate number of data points** (not too small, not too large)

Line Graphs

Olympic long jumps



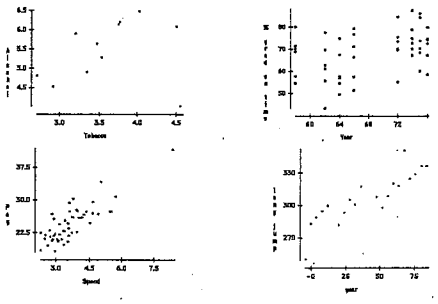
Age Vs Height



Line Graph Facts

- **Best** at showing trends in a numerical data that changes over time
- Always paired data where
  - Independent data is a number or something else ordered (data/time)
  - Only one "y-value" for each "x-value"

Scatter Plot



Scatter Plot Facts

- Scatter plots can only be made for paired numerical data
- Each data pair is data collected taking two different measurements of the same individual
- Scatter plots help show correlation—when the data values are related in some way
- Scatter plots can show change over time, but often do not.