## CS 5630/6630 Scientific Visualization

**Elementary Plotting Techniques I** 

- Everyone uses plotting
- It is easy to lie or to deceive people with bad plots
- Default plotting tools are terrible
- Most people ignore or are unaware of simple principles



Default Excel Plot



Default Matplotlib/Matlab Plot





- Why are they all different?
- What is good/bad about each?

# Fundamentals of plotting

- Analysis vs. Communication
- Presenting data vs. Presenting correlation
- Vision vs. Understanding



- Principle 1: Make data stand out
  - Avoid superfluidity, clutter, or chartjunk.



Activities of a !Kung woman and her baby

Open Bar and Vertical Lines: Nursing times Closed Bars: Sleeping F: Fretting Slashed Lines: Held by mother Arrows: Picking up and setting down

- Principle 1: Make data stand out
  - Avoid superfluidity, clutter, or chartjunk.



- Principle 1: Make data stand out
  - Avoid superfluidity, clutter, or chartjunk.





- Principle 1: Make data stand out
  - Avoid superfluidity, clutter, or chartjunk.



- Principle 1: Make data stand out
  - Avoid superfluidity, clutter, or chartjunk.



- Principle 1: Make data stand out
  - Avoid superfluidity, clutter, or chartjunk.



- Principle 2: Visual prominence
  - Use visually prominent graphical elements to show the data.



- Principle 2: Visual prominence
  - Use visually prominent graphical elements to show the data.



- Principle 2: Visual prominence
  - Use visually prominent graphical elements to show the data.





- Principle 2: Visual prominence
  - Use visually prominent graphical elements to show the data.



- Principle 3: Scale lines and the data rectangle
  - Use two scale lines (box), add margins for data, tick-marks out, 3-10 tick marks.



- Principle 3: Scale lines and the data rectangle
  - Use two scale lines (box), add margins for data, tick-marks out, 3-10 tick marks.



- Principle 3: Scale lines and the data rectangle
  - Use two scale lines (box), add margins for data, tick-marks out, 3-10 tick marks.





- Principle 3: Scale lines and the data rectangle
  - Use two scale lines (box), add margins for data, tick-marks out, 3-10 tick marks.



- Principle 3: Scale lines and the data rectangle
  - Use two scale lines (box), add margins for data, tick-marks out, 3-10 tick marks.



- Principle 3: Scale lines and the data rectangle
  - Use two scale lines (box), add margins for data, tick-marks out, 3-10 tick marks.



- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.



- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.



- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.



- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.





- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.



- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.



Fig. 1. Xenon-133 activity (picocuries per cubic meter of air) in Albany, New York, for the end of March and early April 1979. The lower trace shows the time-averaged estimates of releases (curies per second) from the Three Mile Island reactor (2). The inset shows detailed values for air samples (gas counting) and concurrent average values for ambient air (Ge diode). Abbreviation: LT, less than.

- Principle 4: Reference lines, labels, notes, and keys
  - Only use when necessary and don't let them obscure data.



- Principle 5: Superposed data sets
  - Symbols should be separable and data sets should be easily visually assembled.





- Principle 5: Superposed data sets
  - Symbols should be separable and data sets should be easily visually assembled.





- Principle 5: Superposed data sets
  - Symbols should be separable and data sets should be easily visually assembled.





- Principle 1: Explanation and conclusions
  - Describe everything, draw attention to major features, describe conclusions



#### Principle 1: Explanation and conclusions

Describe everything, draw attention to major features, describe conclusions

Fig. 2. Tension and the intensity of the 42.9nm layer line during 1-second tetanus at the sarcomere length of 2.2 µm. (a) Tension record averaged over the 40 tetanic contractions required for obtaining the time course of the layer-line intensity. A sartorius muscle was dissected from Rana catesbeiana and tetanized for 1 second at 2-minute intervals. The horizontal line represents the period of stimulation. Tension was recorded with an isometric tension transducer (Shinkoh, type UL). (b) Intensity of the first-order myosin layer line at 42.9 nm. The x-ray source was a rotating-anode generator (Rigaku FR) with a fine focus (1.0 by 0.1 mm) on a copper target. This was operated at 50 kV with a tube current of 70 mA; such a high power was possible with an anode of a large diameter (30 cm) rotating at a high speed (9000 rev/ min). A bent-crystal monochromator was used at a source-to-crystal distance of 25 cm with a viewing angle of 6°. The intensity of the myosin layer line was measured with a scintillation counter combined with a mask; the mask had apertures at the positions of the off-meridional parts of the first-order



layer line. The meridional reflection at 14.3 nm is known to be slightly displaced during contraction, suggesting a minute change in the myosin periodicity (I, 3). It is, therefore, possible that the 42.9-nm layer line is also slightly displaced. However, the possible displacement (14  $\mu$ m at the position of the mask) would be insignificant compared with the width of each aperture (0.8 mm). The intensity measured at the resting state was 1400 count/sec. The intensities during and after tetanus were expressed as percentages of the resting intensity and plotted against time after the first stimulus of each set of stimuli. Each point represents the intensity averaged over a 100-msec period. The first three points represent the measurements made before stimulation.

- Principle 1: Explanation and conclusions
  - Describe everything, draw attention to major features, describe conclusions



Fig. 1. Inhibitory effect of morphine hydrochloride (A) and  $\beta$ -endorphin (B) on wet-shake behavior in rats. Antagonism was by antibody to cerebroside sulfate (AS). A volume of 2  $\mu$ I was delivered into the PAG in 1- $\mu$ I increments with a 1-minute interval in between. The control consisted of saline (Sal). Morphine HCI (M-HCI) and  $\beta$ -endorphin were preceded by saline, AS, or naloxone (Nal) as indicated. The dose of morphine HCI and naloxone refers to the chloride salt. Each point is the datum for one animal. The AS + morphine HCI groups are all significantly different (t-test, P < .005) from the corresponding morphine HCI groups alone (A). The group receiving AS + 0.0375  $\mu$ g of  $\beta$ -endorphin is also significantly different (P < .005) from the group receiving 0.0375  $\mu$ g of  $\beta$ -endorphin (B).

- Principle 2: Use all of the available space
  - Fill the data rectangle, only use zero if you need it



• Principle 2: Use all of the available space

• Fill the data rectangle, only use zero if you need it



- Principle 2: Use all of the available space
  - Fill the data rectangle, only use zero if you need it



#### inciple 3: Juxtaposed data sets Clear Understanding Make sure scales match and graphs are aligned



- Principle 3: Juxtaposed data sets
  - Make sure scales match and graphs are aligned





- Principle 4: Log scales
  - Used to show percentage change, multiplicative factors and skewness



- Principle 4: Banking to 45°
  - Aspect ratio is important for judging rate of change

# CO2 VisTrails Demo

# Summary of Principles

- Clear Vision
  - 1. Make data stand out
  - 2. Visual prominence
  - 3. Scale lines and data rectangle
  - 4. Superposed data sets
- Clear Understanding
  - 1. Explanations and conclusions
  - 2. Use all available space
  - 3. Juxtaposed data sets
  - 4. Log scaling
  - 5. Banking to 45°

# Summary of Principles



- Why are they all different?
- What is good/bad about each?

### Quiz on Principles

• What is wrong with this plot?



Computing in Science & Engineering Sep/Oct 2007 page 8

# Quiz on Principles

• What is wrong with this plot?



Computing in Science & Engineering Sep/Oct 2007 page 14

### Quiz on Principles

• What is wrong with this plot?



Computing in Science & Engineering Sep/Oct 2007 page 94