DEPARTMENT OF POLITICAL SCIENCE
AND
INTERNATIONAL RELATIONS
Research Methods Posc 302

MANAGING AND ANALYZING DATA

I.   TODAY'S SESSION:
  A.   Research project designs
  B.   Alternative presentations
  C.   Writing tips:
    1.   Testing for sentence fragments:
      i.   Does the sentence have a **verb?**
      ii.   Does it have a **subject**?
      iii. Is the sentence consist **only of a clause**?
    2.   Check for comma splices:
      i.   *The table shows how age affects voting,* ***t****he young participate less frequently than older people.*
        1)   This should be: *The table shows how age affects voting****.*** ***T****he young participate less frequently than older people.*

II. RESEARCH DESIGNS:
  A.   General observation:
    1.   Assignment 5 suggests that many people seem to be picking variables at random.
      i.   If you are comparing groups, as in the "gap" studies, **think** about how or in what ways the groups would differ.
      ii.   Example: on what issues would blacks and whites agree?
    2.   Try to pick a **consistent** list of variables such as attitudes toward crime **or** attitudes toward government activity **or** attitudes toward social controversies.
      i.   But **don't** pick one or two from each list.
  B.   "Independent" variable:
    1.   I have in effect suggested an individual variable for each topic.
  C.   Political culture
    1.   Since the concept of culture is the most sophisticated of those assigned, I have made the analysis much easier by simply stating that region of residence should be used as an indicator.
    2.   You should be able to find a variable that "measures" region. Use it or some combination of categories as an indicator or operational definition of region.
   i.   The values will thus be "East," West," "South," and so on or

perhaps "Deep South," "West." and "North."
    ii.   Assign each person in the survey to a region.
    iii.   Make sure that you tell the reader exactly how the variable has been defined.
   3.   Suppose you find a difference between southerners and northerners' attitudes toward capital punishment. Is this difference due to race? to income? to education?
    i.   Control for variables that might explain differences.
    ii.   But of course if there are no differences there is no reason to control.

D.   Has public opinion become more conservative?
   1.   In this project the independent variable is a bit unusual for it boils down to **year** or "year of interview."
    i.   Time series data: observations (e.g., interviews) are collected at different time periods.
      1)   The General Social Survey provides a example.
      2)   Polls or surveys are conducted every year.
      3)   **Respondents differ from year to year.**
        a)   Hence, it is not possible to see how particular people change over time.
        b)   One can only investigate the public as a whole.
    ii.   By contrast a **panel** study interviews or studies the **same** people at different times.
      1)   The General Social Survey is **not** a panel study. Don't make the mistake of thinking that you are tracking the same group of individuals over time.
    iii.   A person's score or value on the variable "year" is the particular year in which he or she was interviewed.
      1)   Interpret scores on other variables as usual.
   2.   Generic analysis:

Belief/Year	1980	1984	1988	1992
Worse	23%	30%	35	40
Same	44	29	25	30
Better	33	41	40	30
Totals (Number)	100 (988)	100 (897)	100 (967)	100 (1007)

III.   ALTERNATIVE PRESENTATIONS:

     A.   The cross-tabulation table that we have used extensively provides a nice way to summarize information obtained in these research projects including the time series study.

          1.   Consider the previous table. One could display the "percent better" with a line plot.

               i.   A line plot graphs each time period, in this case each year, against the corresponding values.

               ii.   Example: 1980 vs 23; 1984 vs 30; 1988 vs 35; etc.

          2.   More generally a line plot consists of two **series** plotted on an X-Y coordinate system.

               i.   Example
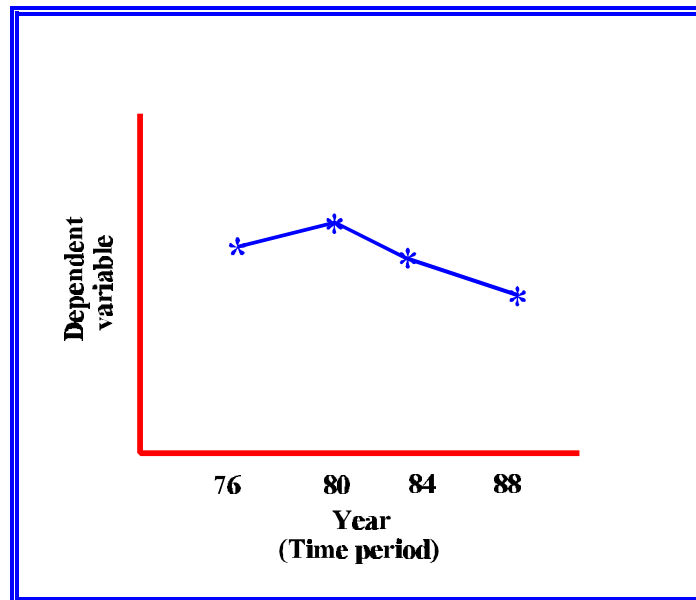
                    1)   Note: time period (year) is the independent variable series.

                    2)   The scores plot for each time period are the values of the dependent variable.

                        a)   Pick one series per plot.

                        b)   That is, in the previous table the values of the dependent variable are "Better," "Same," and "Worse." **Plot only the percentages for one of these values.**

     B.   Spreadsheet programs.

          1.   Programs such as Microsoft Excel, Corel Quattro Pro, SPSS, Minitab, and countless provide an easy way to manipulate data and create various types

of charts.
2.   First enter or transfer the data into the spreadsheet. Then use it to create a line plot.

IV.   EXAMPLE:
    A.   Microsoft Excel is perhaps the most commonly used spreadsheet program on campus and so we can illustrate its use in data analysis and presentation.
        1.   Spreadsheet.
           i.   The key idea is the spreadsheet that consists of a matrix of rows and columns.
           ii.   Each row stores a particular individual or year or in our language a unit of analysis.
           iii.   Columns hold the values of variables for the particular individual.
           iv.   Example:
                1)   Suppose the value of percent "Worse" for 1980 is 23.
                2)   Suppose furthermore 1976 is the first case.
                3)   Then the spreadsheet's first row will represent 1980 and 23 will be stored in a column named, say, "Worse."
                4)   If the value for the year 1984 is 30, then we'll find a 30 in the "Worse" column.
           v.   Example: here's what an Excel spreadsheet containing the data in the previous table looks like.

Figure 2:Excel Spreadsheet

           vi.   Here's a Quattro Pro spreadsheet. It works more or less the same way, right?
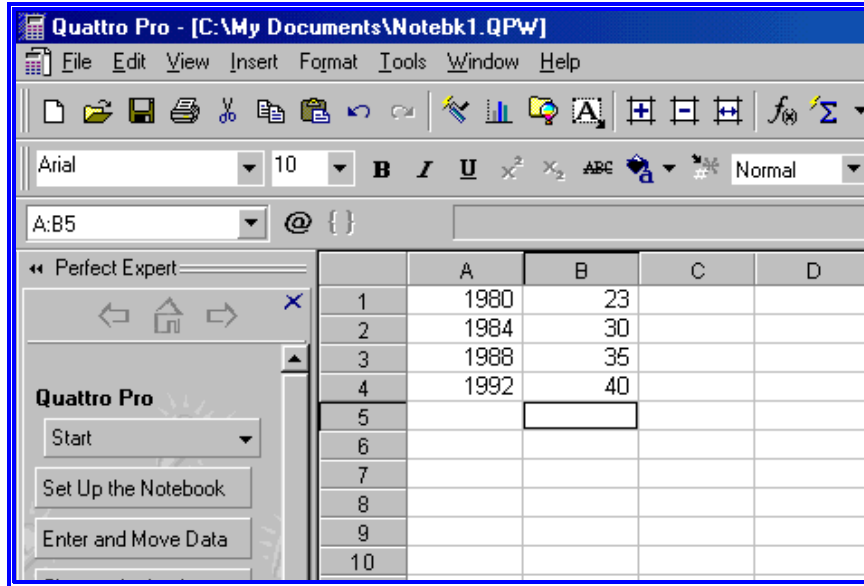
Figure 3: Quattro Pro Spreadsheet

B.   Making a plot with Excel:
      1.   Highlight the column containing the data to be graphed.
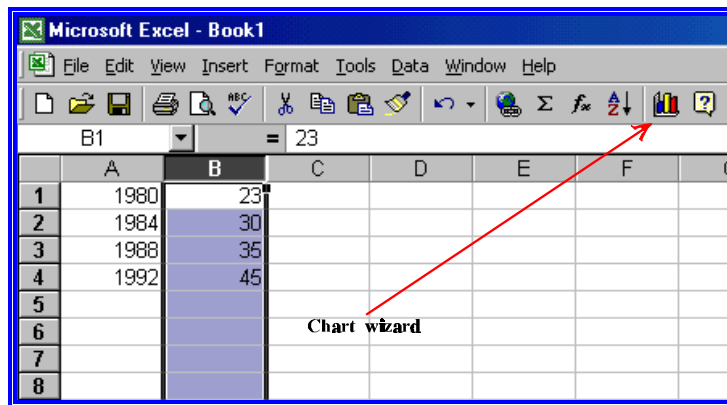   i.   Example

Figure 4: Highlight Series to Be Plotted

      2.   To make a graph click on the "chart wizard."
            i.   Choose "Line," then choose the simplest plot as shown next.
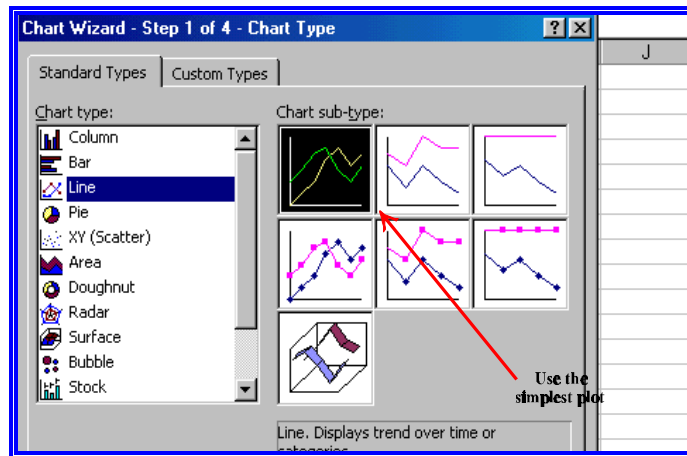
Figure 5: Choosing Chart Type

ii.   Click Next.
iii.   Click on Series tab in the next dialogue box.
        1)   Look for "Category (X) axis labels." Click the icon on the
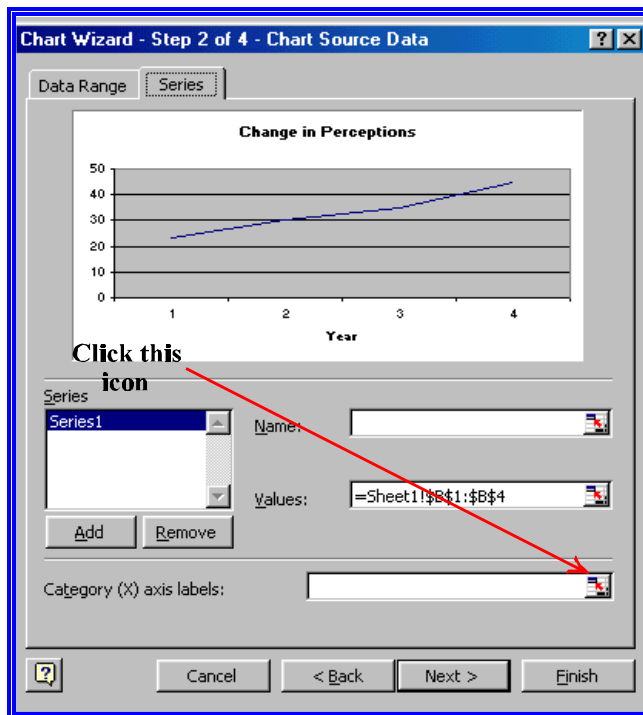   right to minimize the box. See next figure.

Figure 6: Click on Category (X) Icon

2)   Doing so minimizes the dialog box so that you can now
highlight the "year," column which contains the labels for
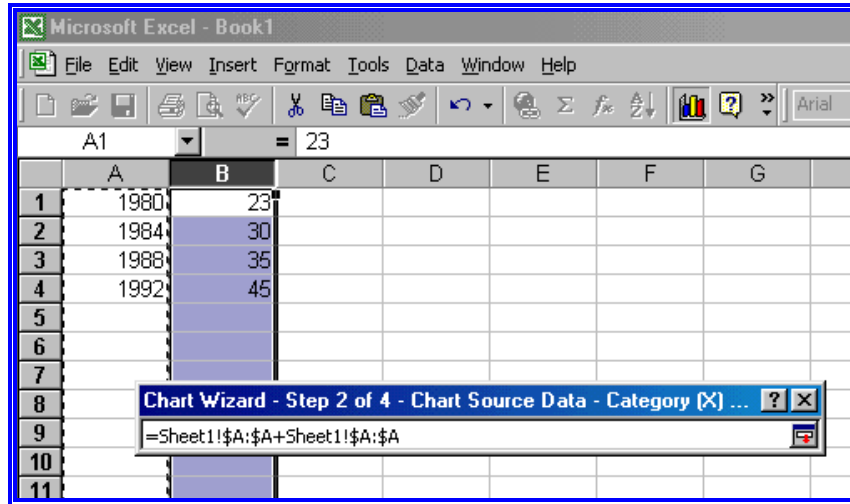the X axis. See below.

Figure 7: Select Column Containing X Axis Labels.

3)   The boarder around the data flashes indicating that its
values have been selected as labels for the X axis. Then
click the icon on the Chart Wizard again to enlarge it.

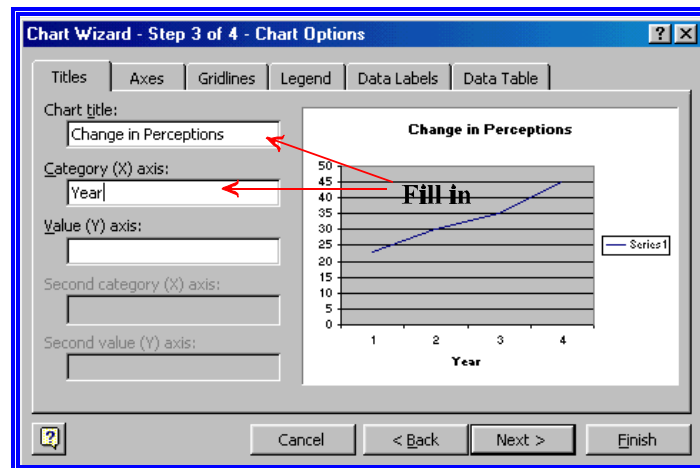iv.   Fill in the boxes in the "Titles" dialog box, as shown next.

Figure 8: Use Titles and Labels

v.   Now open Gridlines tab and close all grid lines.

vi.   Open Legend tab and click off "Show legend."

        vii.   Click Next and choose "As new sheet."
        viii.   Click Finish.
3.   You can then copy the graph and paste it in your Word document.
        i.   Note: I actually coped the Excel chart and copied it into WordPerfect. It may not work for you, however.
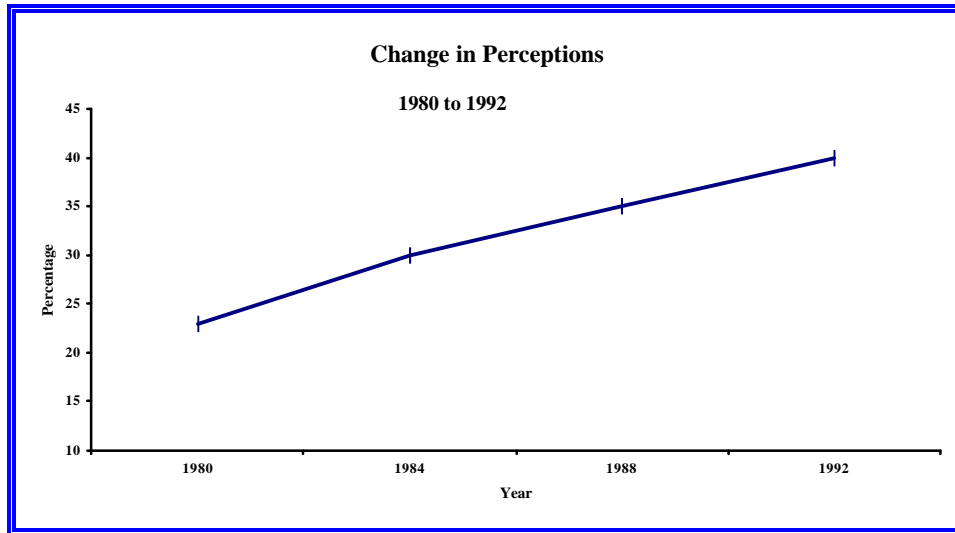   ii.   Here's what I got.

Figure 9: Finished Time Series Plot

4.   Editing the graph.
        i.   We'll explore ways to edit and fine tune the plot or graph.

V.   NEXT TIME:
    A.   Measures of association and correlation
    B.   Reading:
        1.   Look carefully at Johnson and Joslyn, *Research Methods* Chapter 14 for a discussion of how to organize and write a research report.