Transformation, Plot, and Regression: Liberty Ship Production

(Information) The purpose of this lab is to transform data, plot it, and to do a regression analysis. These skills are used in the physical and life sciences lab courses. The historical data for this exercise is production data for Liberty ships during World War II.

Table 1. Liberty Ship Production Data at Shipyard A[1]

Order	5	10	25	30	75	100	125	150	175	200	225	250	285
Manhours	1094	894	647	659	529	424	376	395	395	388	376	353	280

Order: Order number of Liberty Ships built in Yard A in World War II

Manhours: Thousands of man-hours per ship

(Procedure:)

1. Start a new Excel file.
2. Save the file under the file name "Transformation Plot Regression"
    a. File | Save As:
    b. Select the directory or folder to save the file in.
    c. Enter "Transformation Plot Regression" in the "File name:" edit box.
    d. Click the "Save" button in the "Save As" dialog box.
3. Change the sheet name to "Original Data".
    a. Double-click on the active (white) tab at the bottom of the Excel document window.
    b. Enter the name "Original Data".
    

4. Enter column headings between the "Order", "LN Order", and "Manhours".
    a. Click on cell A1. Enter **Order** as the column label.
    b. Click on cell B1. Enter **LN Order** as the column label.
    c. Click on cell C1. Enter **Manhours** as the column label.
5. In column A, beginning with row 2, enter the data for number of ships on order from the "Order" row of Table 1.

[1] A.C. Rosander, Applications of Quality Control to the Service Industries, American Society for Quality Control (ASQC), Quality Press (1985), p. 114, downloaded from Carnegie Mellon University Data Sets on 22 May 2003.

6. In column C, beginning with row 2, enter the data for the number of manhours per ship actually used in building a ship from the "Manhours" row of Table 1.
7. For each row, compute in column B the natural logarithm of the corresponding value in column A. Strategy: Insert a formula in row 2, and copy that formula to other rows.
    a. Insert a formula in Cell B2 to compute the natural logarithm of the number in Cell A2.
        i. Click on cell B2.
        ii. Click on the equal (=) sign next to the formula edit box. If Excel 2002 (Excel XP) is being used, click in the formula edit window to the right of the insert formula icon  , and type an equal sign (=).
        iii. Click the down arrow in the selection box to the left, and select "More functions…". Do this even if "LN" already is visible in the selection box. The learning goal is to find functions if they are not immediately visible.
        iv. In the "Paste Function" dialog box,
            1. under "Function category:" selection list, select "Math & Trig".
            2. In the "Function name:" selection list, select "LN".
            3. Click "OK" in the "Paste Function" dialog box.
        v. Drag the "LN" dialog box to a location that will not obscure cell A2.
        vi. Click on cell A2.
        vii. Click the "OK" button in the "LN" dialog box.
    b. Copy the formula of Cell B2 to Cells B3 through B14.
        i. Place the mouse cursor over the copy handle (small solid square box) in the lower right corner of cell B2.
        ii. Left drag the copy handle to cell B14. This causes the formula to be copied to other rows, computing the natural log of the number of ships on order.
        iii. Click in any other cell in order to deselect Cells B2:B14.
8. Document the spreadsheet in the header and footer.
    a. File | Page Setup
    b. On the "Sheet" tab, in the "Print" section, checkmark "Gridlines".
    c. On the "Page" tab, select "Portrait".
    d. On the "Header/Footer" tab, select "Custom Header".

i. Enter your name in the header “Left Section”.

ii. Enter the title “Transformation and Plot” in the header “Center Section”.

iii. Enter the Date and Time in the header “Right Section”.

1. Click on the Date icon. 
2. Insert a space.
3. Click on the Time icon. 

iv. Click the “OK” button in the “Header” dialog box.

v. Click the “OK” button in the “Page Setup” dialog box.

9. Save the file. Click on the “File Save” icon in the Standard Toolbar. 

10. Preview and Print the Formula sheet.

a. Use **Ctrl** + (**Left Single Quote**) to select the **formula** sheet.

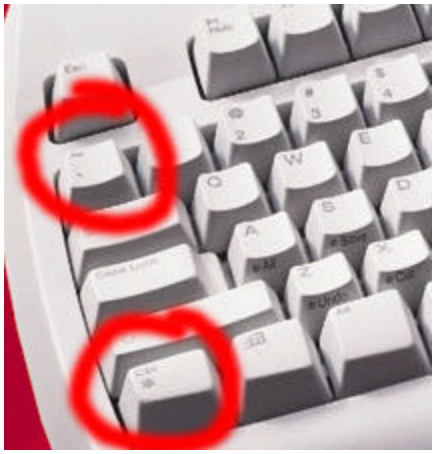


Figure 1. Control + Left Single Quote[2]

- The Left Single Quote is to the left of the “1” key at the top of the typing keys.

b. Click the “Print Preview” icon on the Standard Toolbar.

c. Click the “Print” button in the “Print Preview” window.

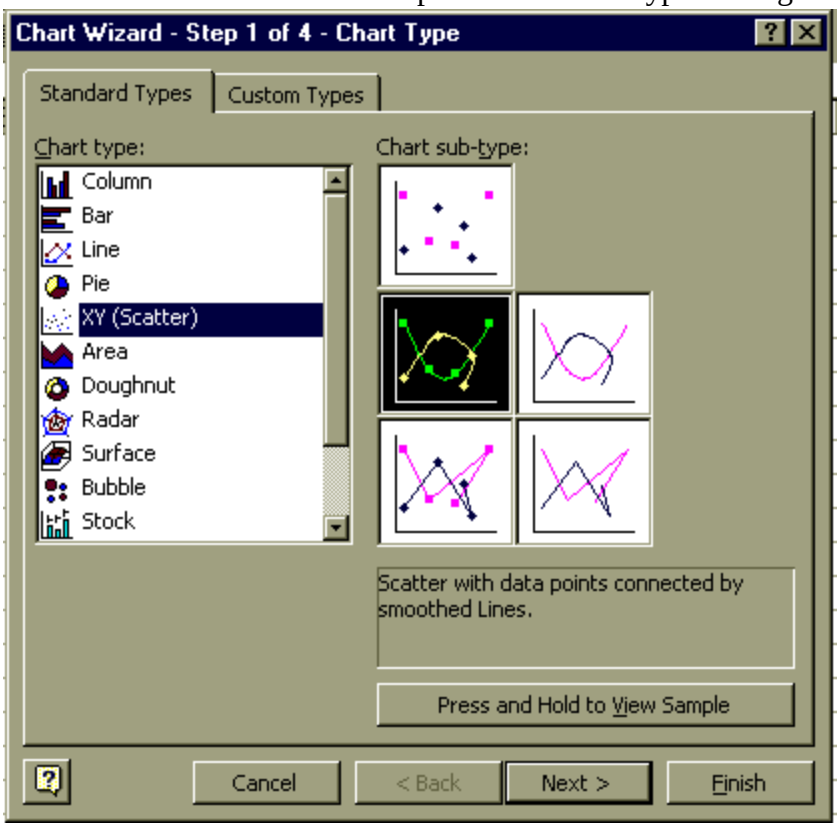
d. In the “Print Range” section of the “Print” dialog box, select “Page(s) From: **1** To: **1**”.

e. Click the “OK” button in the “Print” dialog box.

[2] Illustration by Joshua B. L. Caldwell.

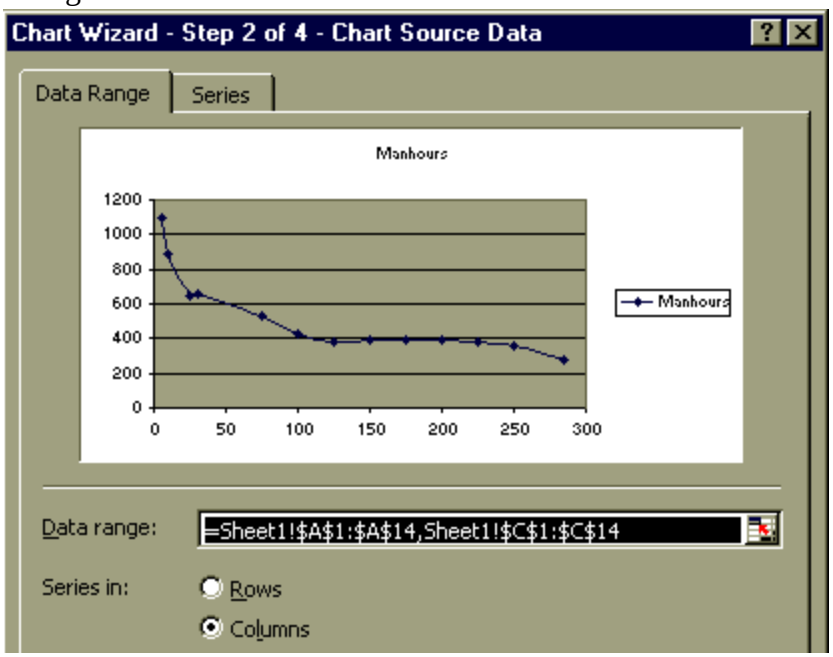
11. Return to the data sheet view by again pressing Use Ctrl + (Left Single Quote).
12. **Plot** "Thousands of Manhours per Ship" (vertical axis, **column C data**) **versus** "Number of Ships on Order" (horizontal axis, **column A data**). Use the Chart icon to select the Scatter chart with solid lines connecting data points.
    a. Select Column C data.
        i. Click on Cell C1.
        ii. Push and hold down the **Shift** key.
        iii. Click on Cell C14.
        iv. Release the Shift key.
    b. Select Column A data.
        i. Push and hold down the **Ctrl** key.
        ii. Click on Cell A1.
        iii. Release the Ctrl key.
        iv. Push and hold down the Shift key.
        v. Click on Cell A14.
    c. Click on the Chart Wizard icon in the Excel Standard Toolbar.

d. Select "XY (Scatter)" in the "Chart type:" section of the "Standard Types" sheet of the "Chart Wizard – Step 1 of 4 – Chart Type" dialog box.



e. Select the "Scatter with data points connected by smoothed Lines" plot by clicking on the plot in column 1, row 2 of the "Chart sub-type:" section of the "Standard Types" sheet of the "Chart Wizard – Step 1 of 4 – Chart Type" dialog box.

f. Click the "Next" button on the "Chart Wizard – Step 1 of 4 – Chart Type" dialog box.

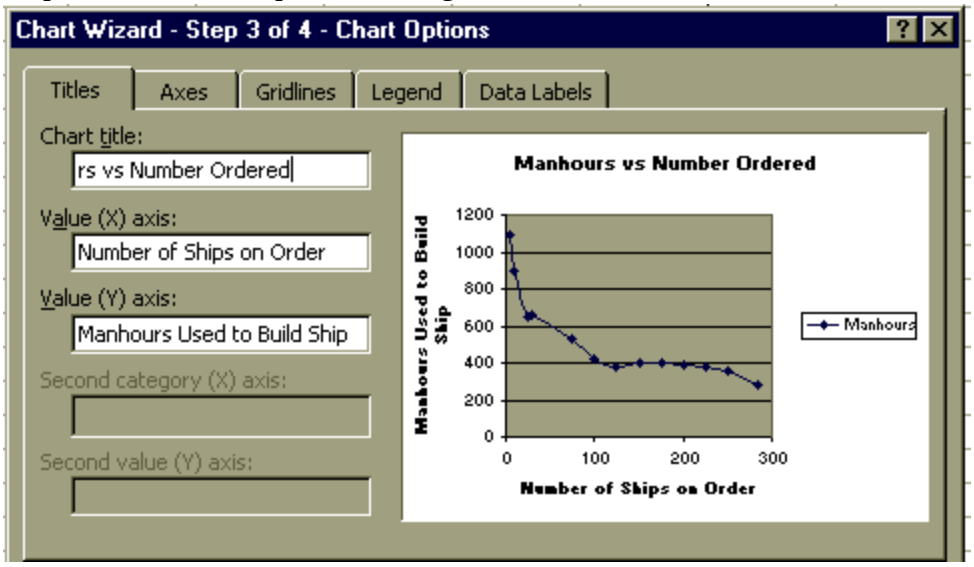


g. Select "Columns" for "Series in:".

h. If "Manhours" does not appear on the displayed chart, the column titles probably were not included. This can be corrected by making sure the number after the first "A" is "1", and the number after the first "C" is "1" in the code in the "Data range:" edit box. This selects the first row cell in each of those columns to be included in the data range.

i. Click the "Next" button on the "Chart Wizard – Step 2 of 4 – Chart Source Data" dialog box.

j. Make the following entries in the "Titles" sheet of the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box:



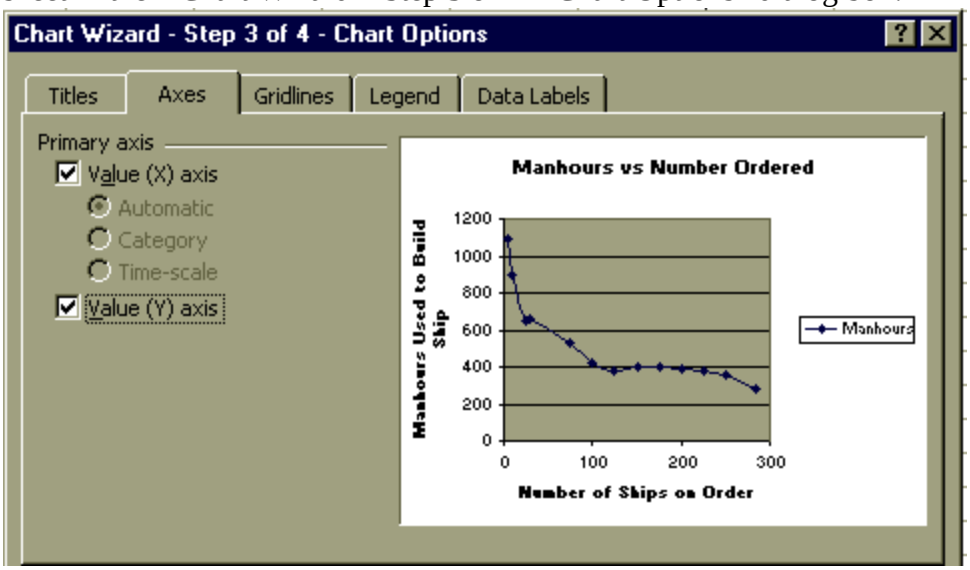
i. Enter “Manhours vs Number Ordered” into the “Chart title:” edit box.

ii. Enter “Number of Ships on Order” into the “Value (X) axis:” edit box.

iii. Enter “Manhours Used to Build Ship” into the “Value (Y) axis:” edit box.

k. Click on the “Axes” tab display the “Axes” sheet of the “Chart Wizard – Step 3 of 4 – Chart Options” dialog box.

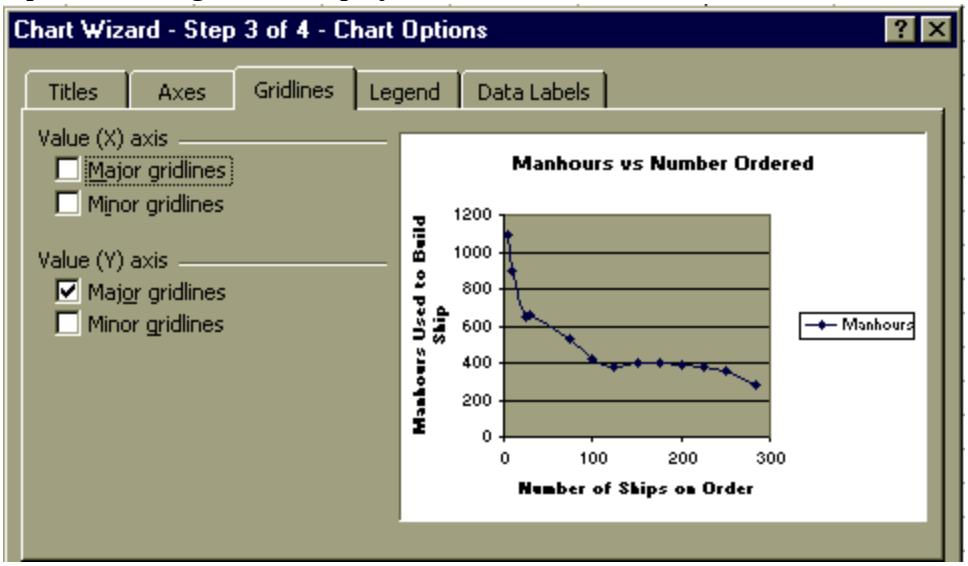
l. Make the following selections in the “Primary axis” section of the “Axes” sheet in the “Chart Wizard – Step 3 of 4 – Chart Options” dialog box:



i. Checkmark “Value (X) axis”.

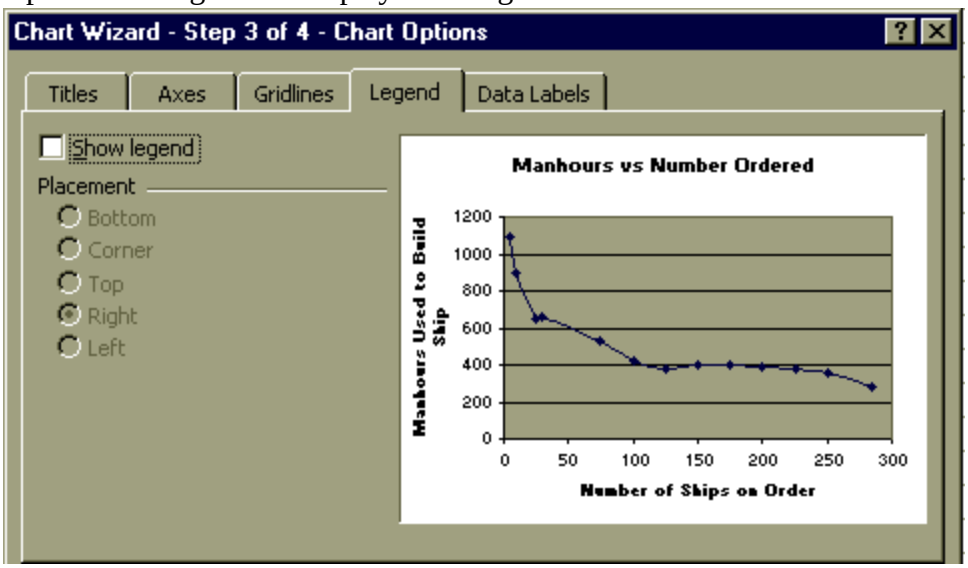
ii. Checkmark “Value (Y) axis”.

m. Click on the "Gridlines" tab of the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box to display the "Gridlines" sheet.



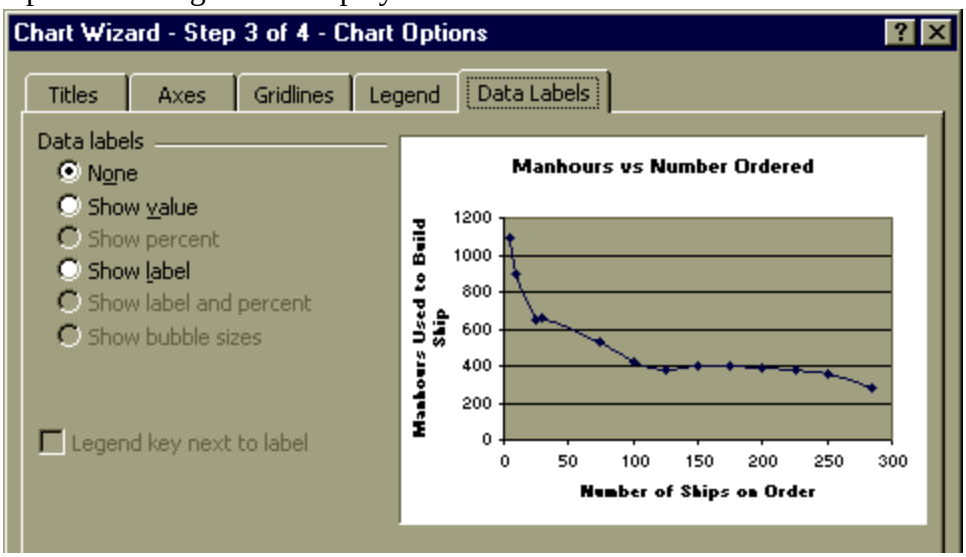
n. Ensure that the only the "Value (Y) axis Major gridlines" box is checked on the "Gridlines" sheet of the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box.

o. Click on the "Legend" tab tab of the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box to display the "Legend" sheet.



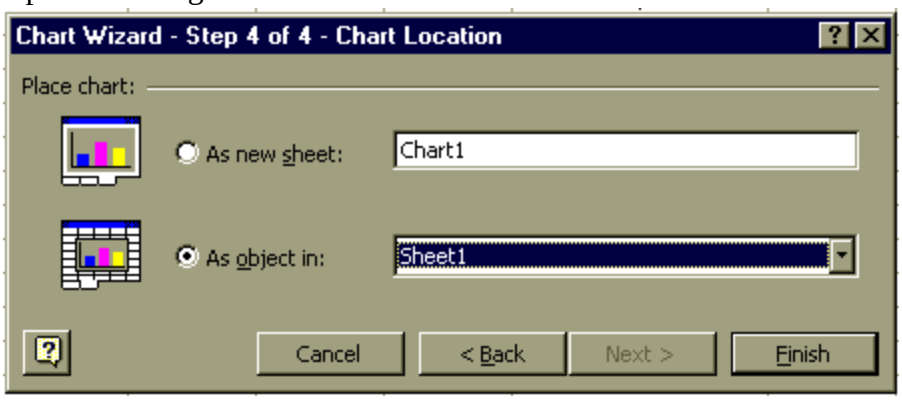
p. Ensure the "Show legend" box is not checked. If it is checked, click on the checkmark to uncheck it. This will cause the legend key to be removed.

q. Click on the "Data Labels" tab of the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box to display the "Data Labels" sheet.



r. Ensure that "None" is selected in the "Data labels" section of the "Data Labels" sheet in the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box.

s. Click on the "Next" button on the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box.



t. Select "As object in:" in the "Place chart:" section of the "Chart Wizard – Step 4 of 4 – Chart Location" dialog box. This causes the plot to be inserted on the data sheet rather on a new sheet.

u. Click the "Finish" button in the "Chart Wizard – Step 4 of 4 – Chart Location" dialog box.

13. Drag the plot so that the upper left corner of the plot is in the upper left corner of cell A16.

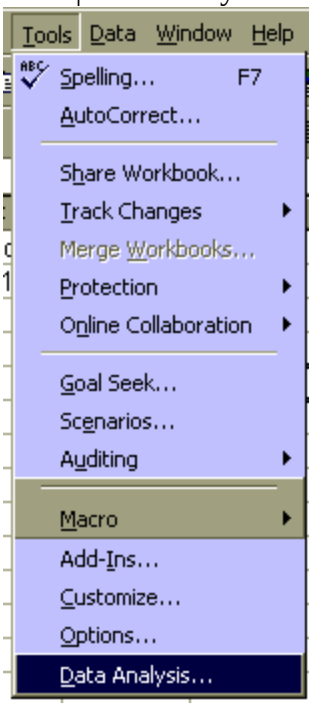
14. Click on the "Save" icon in the Excel Standard Toolbar.

15. **Plot** "Thousands of Manhours per Ship" (vertical axis, **column C data**) **versus** "LN of Number of Ships on Order" (**column B data**). Use the Chart icon to select the Scatter chart with solid lines connecting data points.

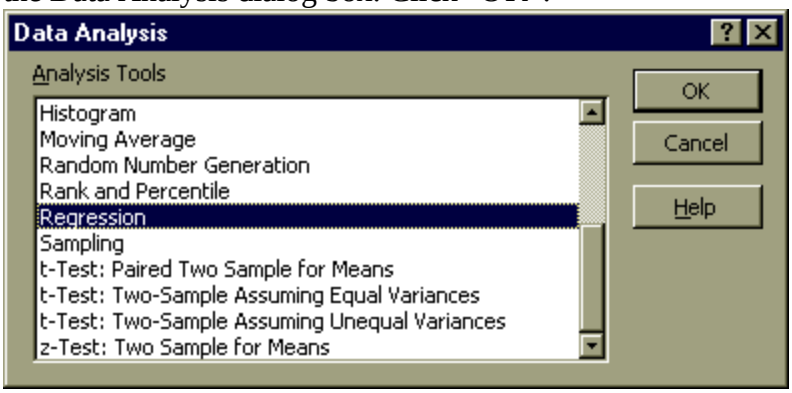
a. Select columns of data to be plotted.
    i. Select cell B1.
    ii. Push and hold the Shift key.
    iii. Select cell C14.
b. Repeat steps 12.c through 12.g.
c. If "Manhours" does not appear on the displayed chart, the column titles probably were not included. This can be corrected by making sure the number after the first "B" is "1", and the number after the first "C" is "1" in the code in the "Data range:" edit box. This selects the first row cell in each of those columns to be included in the data range.
d. Click the "Next" button on the "Chart Wizard – Step 2 of 4 – Chart Source Data" dialog box.
e. Make the following entries in the "Titles" sheet of the "Chart Wizard – Step 3 of 4 – Chart Options" dialog box:"
    i. Enter "Manhours vs Log of Number Ordered" into the "Chart title:" edit box.
    ii. Enter "Log of Number of Ships on Order" into the "Value (X) axis:" edit box.
    iii. Enter "Manhours Used to Build Ship" into the "Value (Y) axis:" edit box.
f. Do steps 12.k through 12.u.
g. In the "Center Section" of the Header, enter "LN of Number of Ships on Order versus Thousands of Manhours per Ship" in the "Center Section".

16. Drag the plot so that the upper left corner of the plot is in the upper left corner of cell A31.
17. Click on an empty cell to deselect the plot.
18. Click on the "Save" icon in the Excel Standard Toolbar.
19. **Preview and print the data and plots on one sheet.**
20. Do a **linear regression** on "Thousands of Manhours per Ship" (**column C data**) **versus** "Number of Ships on Order" (**column A data**).
    a. Select the Regression program of Excel.

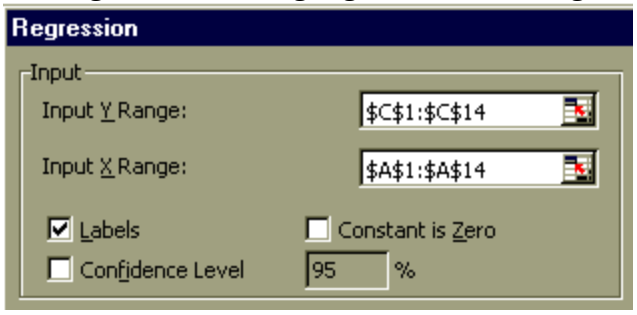
i. Tools | Data Analysis.



ii. Choose "Regression" from the "Analysis Tools" selection list of the Data Analysis dialog box. Click "OK".



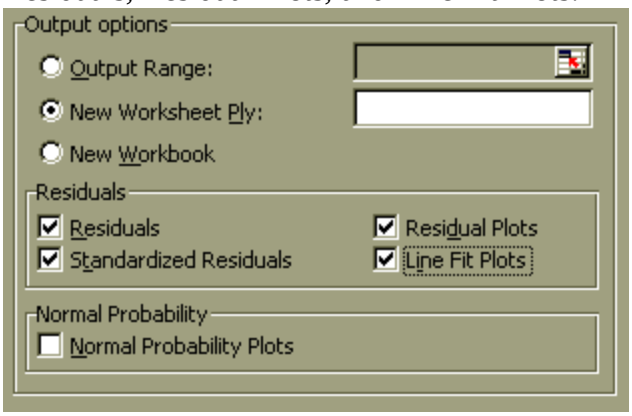
b. In the "Input" section of the "Regression" dialog box, click in the "Print Y-Range" edit box. Highlight cells C1 through C14.



c. In the "Input" section of the "Regression" dialog box, click in the "Print X-Range" edit box. Highlight cells A1 through A14.

d. In the "Input" section of the "Regression" dialog box, checkmark the "Labels" check box.

e. In the "Output options" section of the "Regression" dialog box, in the "Residuals" subsection, check all boxes: Residuals, Standardized Residuals, Residual Plots, and Line Fit Plots.



f. Click the "OK" button in the "Regression" dialog box.

g. Change the sheet name to "Linear Regression".

   i. Double-click on the active (white) tab at the bottom of the Excel document window.

   ii. Enter the name "Linear Regression".

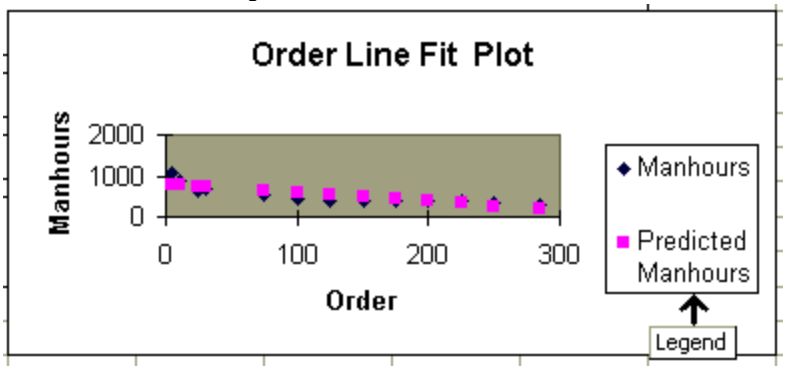
h. AutoFit the size of columns A, B, C, and D to display labels completely.

   i. In the grey column headings at the top of the Excel spreadsheet, select columns A, B, C, and D.

   ii. Place the mouse cursor at the right edge of column D in the top grey area.

   iii. Double-click. The columns will automatically adjust width to the widest entry in each selected column.

i. Click on the legend in the "Order Line Fit Plot" and use the Delete key to remove it from the plot.



j. Drag the "Order Line Fit Plot" so the upper left corner is at the upper left corner of cell E1. Size the plot so the lower right corner is at the lower right corner of cell J10.

k. Drag the "Order Residual Plot" so the upper left corner is at the upper left corner of cell E23. Size the plot so the lower right corner is at the lower right corner of cell J32.

21. Setup to print linear regression analysis and plots.

a. In File | Page Setup, on the "Page" sheet, select the "Landscape" orientation.

b. On the "Header/Footer" tab, select "Custom header".

i. Enter your name in the header "Left Section".

ii. Enter the title "Linear Regression" in the header "Center Section".

iii. Enter the Date and Time in the header "Right Section".

1. Click on the Date icon. 

2. Insert a space.

3. Click on the Time icon. 

iv. Click the "OK" button in the "Header" dialog box.

c. On the "Header/Footer" tab, select "Custom footer".

i. Click in the "Left section".

ii. Insert the code for displaying the file name by clicking on the File icon. 

iii. Click in the "Center section".

iv. Insert the code for displaying the sheet name by clicking on the "Tab" icon. 

v. Click the "OK: button in the "Footer" dialog box.

d. Click the "OK" button in the "Page Setup" dialog box.

e. Click the "Print Preview" icon in the Standard Toolbar. Notice that the chart almost fits onto one page.

f. Click "Setup…" in the "Print Preview" toolbar at the top of the Print Preview window.

g. Click "Page" in the "Page Setup" dialog box.

h. In the "Scaling" section, select "Fit to: **1** page(s) wide by **1** tall".

i. Click the "OK" button in the "Page Setup" dialog box.

j. Click the "Close" button in the "Print Preview" toolbar.

22. Click on the "Save" icon in the Standard Toolbar.

23. Print the linear regression analysis.

24. Do a **logarithmic regression on** "Thousands of Manhours per Ship" (**column C data**) **versus** "LN Number of Ships on Order" (**column B data**). Repeat the above procedure, with the following changes:
    a. Click on the "Original Data" tab at the bottom of the Excel document window.
    b. Do steps 20.a through 20.b.
    c. In the "Input" section of the "Regression" dialog box, click in the "Print X-Range" edit box and delete any code in the edit box. Highlight cells B1 through B14.
    d. Do steps 20.d through 20.f.
    e. Change the sheet name to "Logarithmic Regression".
        i. Double-click on the active (white) tab at the bottom of the Excel document window.
    f. Enter the name "Logarithmic Regression".
    g. Do step 20.h.
    h. Click on the legend in the "LN Order Line Fit Plot" and use the Delete key to remove it from the plot.
    i. Drag the "LN Order Line Fit Plot" so the upper left corner is at the upper left corner of cell E1. Size the plot so the lower right corner is at the lower right corner of cell J10.
    j. Drag the "LN Order Residual Plot" so the upper left corner is at the upper left corner of cell E23. Size the plot so the lower right corner is at the lower right corner of cell J32.
25. Setup to print linear regression analysis and plots.
    a. In File | Page Setup, on the "Page" sheet, select the "Landscape" orientation.
    b. On the "Header/Footer" tab, select "Custom header".
        i. Enter your name in the header "Left Section".
    c. Enter the title "Logarithmic Regression" in the header "Center Section".
    d. Do steps 21.b.iii through 20.h.

26. Print the logarithmic regression analysis.