DECUS
PROGRAM LIBRARY

DECUS NO.	8-202
TITLE	PLOT
AUTHOR	J. J. Spruit and L. R. Davila
COMPANY	Fels Research Institute Yellow Springs, Ohio
DATE	October 1, 1969
SOURCE LANGUAGE	PAL III

PLOT

Calculation of Linear Regression Line, Plotting of Data Points
and Plotting of Linear Regression Line

DECUS Program Library Write-up DECUS No. 8-202

SCALE DETERMINATION

The computer will calculate a scale for the x axis and the y axis from the minimum and maximum x and y values, respectively. Therefore, one must enter as the first two "data points" value for x and y that will determine the x and y scales. These two points will be plotted on the graph paper, but will not be included in the calculation of the linear regression line.

Example One

If you want to plot on a half page (8 1/2 x 11) x values which are between 1.0 and 8.0, and y values between 1.0 and 5.0, then it would be convenient to add the two points 0,0 and 9,6. The x scale would then be typed out as 0.1 and the y scale 0.1, indicating that each little square on the graph paper corresponds to 0.1 unit. The x and y axis values for the heavy graph lines would be 1.00, 2.00, 3.00, 4.00, 5.00, etc.

Example Two

Say the x values on the half page are between 3.2 and 4.5. Good beginning and end points would be 3 and 4.8 respectively. The ten lines on the x axis should be marked: 3.00, 3.20, 3.40, 3.60, 3.80, 4.00, 4.20, 4.40, 4.60, 4.80.

General rule: find two end points, between which the actual data points will fall, and so that the difference between the two end points can be divided by the number of large squares (listed below) on the corresponding axis giving a convenient scale on that axis.

Half page x axis contains 9 large squares
Half page y axis contains 6 large squares
Full page x axis contains 14 large squares
Full page y axis contains 9 large squares

The actual markings of the axes can be done with the Alphanumeric Plotter Program, ALPHA, DECUS No. 8-203, or automatically when an 8K memory is available.

LOCATIONS

The program uses the following locations:

5, 7, 40-62 for the floating point package
2, 3, 15, 16, 17 for all sections of the program
20-177 for all sections of the program
200-2453, 4650-4751 for all sections of the program
2460-3553 for storage of x values, 3 locations per data point
3554-4647 for storage of y values, 3 locations per data point

A total of 190 x-y combinations can be stored.

OPERATING INSTRUCTIONS

1. Load binary tapes for this program in the usual manner. Also, the Floating Point Package, No. 3 for EAE Type 182, Digital 8-25-U BIN (Fels Tape #7) should be loaded. When extended memory is available, the Alphanumeric Plotter Program should be loaded in the basic memory (Field 0) and PLOT and the Floating Point Package both in Field 1. After the points and linear regression line are plotted, the computer will then go to the Alphanumeric Program automatically.

2. Set RSW to 0200 in Field 1 (or Field 0, if no extended memory is available). "LOAD ADD" and "START".

3. For a half page, put bit 10 up. For a full page, put bit 10 down. These switch settings are the same as used in the Alphanumeric Plotter Program.

4. When a print-out of the entered data is required, put bit 1 up. When no print-out is needed put bit 1 down. When entering the data by hand from the keyboard, choose the print-out option. When entering the data from a paper tape, not typing out will speed up the operation.

5. When the user does not want the regression line to be drawn on the graph, put bit 9 up. The computer will then calculate the regression line, type out the values, type the x scale and y scale, and plot the points. Then it will go to a next regression line calculation. Put bit Ø up when automatic scaling is required.

6. Before entering the data, set the plotter pen manually exactly on point A, indicated in Figure 1 or Figure 2, depending on whether the user wants a half page or a full page format. Location 1320 contains 5331 which causes the computer to bypass a section of the program which automatically sets the pen on point A. We found that the pen might be off-center and so does not go exactly to point A. Therefore, we inserted the JMP. The user is invited to try the subroutine by changing 1320 to 7300. It is possible that the counters M1000 and N18 need adjusting.

7. To enter the data from the keyboard or to punch a paper tape:

"RETURN" or letter y (see explanation at the end of this paragraph).

x_o SPACE y_o SPACE These values must be the additional points used

x_1 SPACE y_1 SPACE to calculate the scales, as explained above.

x_2 SPACE y_2 SPACE This is the first actual data point.

x_n SPACE y_n SPACE This last y-value must be followed by RETURN (EOL).

When the x values are to be plotted on the x axis and the y values on the y axis strike or punch "RETURN". When the y values are to be plotted on the x axis and the x values on the y axis, strike or punch Y.

8. When the return is sensed by the computer it will calculate from the data points the linear regression line and type out the values:

B, C, SE	In the formula $y = Bx + C \pm$ standard error	See last page for the formulas used in the calculations.
PCT	Prediction error in percent	
RSQ	r^2	
R	r (correlation coefficient)	
SER	Standard error of r	
N	Number of entries (x-y combinations)	

9. The computer then waits. Now any value of x can be entered followed by a terminator (see below) and the computer will calculate and type out the corresponding y value and the y value ± standard error. Selecting x values close to either end of the graph will allow the user to draw the band around the regression line very easily. For this, two x values will be required; terminate the first x by a space. This will permit the computer to accept another x value. The last x must be terminated by either a RETURN or a decimal point.

RETURN	Will bring the computer back for another regression line calculation.
DECIMAL POINT	Will start the plotting of the points and the regression line plot.
SPACE	Will bring the computer back for another x value.

10. If the x value is an integer, a following decimal point will be interpreted as a real decimal point. In order to terminate the entry, a second decimal point is required. Before striking the terminating decimal point, establish the proper switch setting:

a. Bit 0 down: The plotter pen will stop after the regression line is drawn, ready to start in the Alphanumeric program.

b. Bit 0 up: The plotter pen will go all the way to the right on the x axis after the present plot is completed, and start drawing the x and y axis lines and the axis notations with the Alphanumeric Plotter Program (see 1).

11. When the decimal point is used as terminator, the computer will first calculate the y and y ± standard error, rank the x values in descending order and then calculate the x and y scales and type these out. Then it will start plotting, first two end points and then the actual data points.

12. Figure 1 is a completed plot in half page format. Figure 2 gives the same points in full page format. Figure 3 is a typical type-out (for the half page option of Figure 1).

13. Resume of switch settings:

Bit	Up	Down
0	After plotting the points and regression line, plot the scales automatically.	After plot is finished, prepare for Alphanumeric plot.

Bit	Up	Down
1	Print-out of entered data	No print-out of entered data.
9	Do not plot regression line.	Plot regression line.
10	Half page format.	Full page format.

14. When only the basic memory is available, the two programs should be used one after the other. In this case, location 1721 in this program could be changed to HLT (7402). This will halt the program.

15. In some cases, a slight job will appear at certain intervals in the regression line. This is due to a limitation imposed by the size of the plotter motor step and is more noticeable with the felt pen than with the ball point pen.

Regression and Correlation

$$y = b\chi \div c \pm \text{standard error}$$

$$b = \frac{N\Sigma\chi y - (\Sigma\chi)(\Sigma y)}{N\Sigma\chi^2 - (\Sigma\chi)^2}$$

$$c = \frac{\Sigma y}{N} - \frac{b\Sigma\chi}{N}$$

$$\text{standard error} = \sqrt{\frac{\Sigma(y \text{ calculated} - y \text{ experimental})^2}{N}}$$

$$r^2 = \frac{[N\Sigma\chi y - (\Sigma\chi)(\Sigma y)]^2}{[N\Sigma\chi^2 - (\Sigma\chi)^2][N\Sigma y^2 - (\Sigma y)^2]}$$

$$r = \sqrt{r^2}$$

$$\text{standard error of } r = \frac{1 - r^2}{\sqrt{N}}$$

$$PCT = \sqrt{\frac{\Sigma\left(\frac{y \text{ calculated} - y \text{ experimental}}{y \text{ calculated}}\right)^2}{N-2}}$$

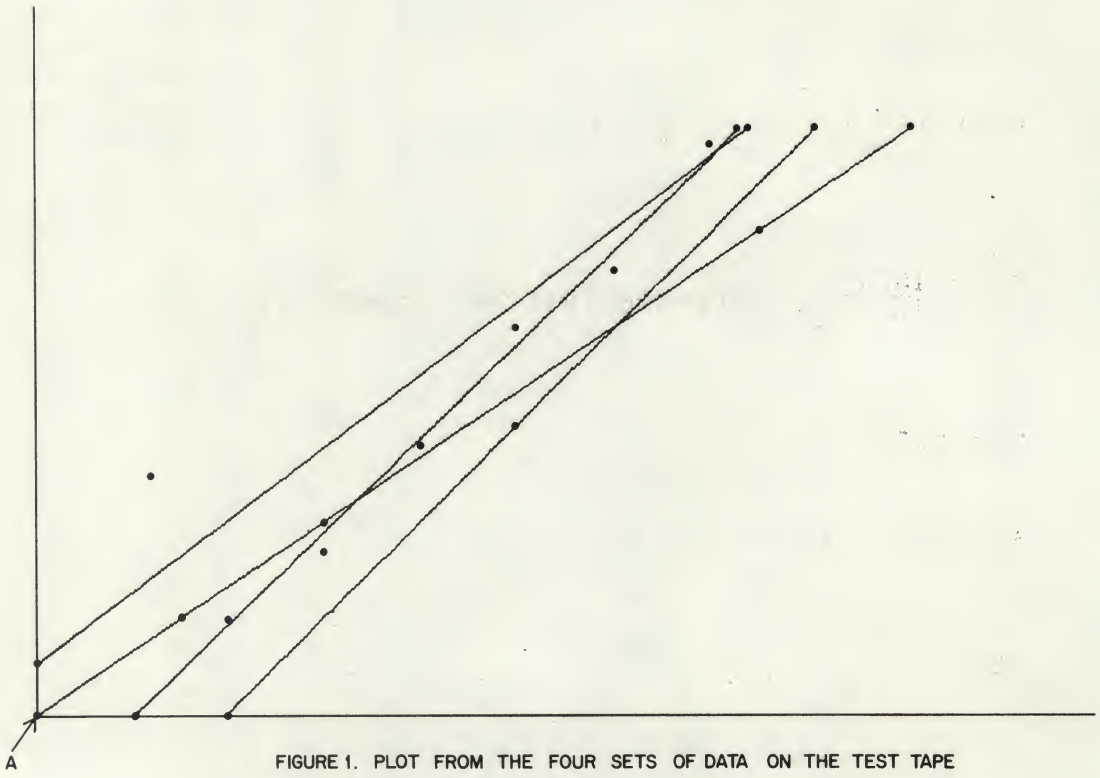

FIGURE 1. PLOT FROM THE FOUR SETS OF DATA ON THE TEST TAPE

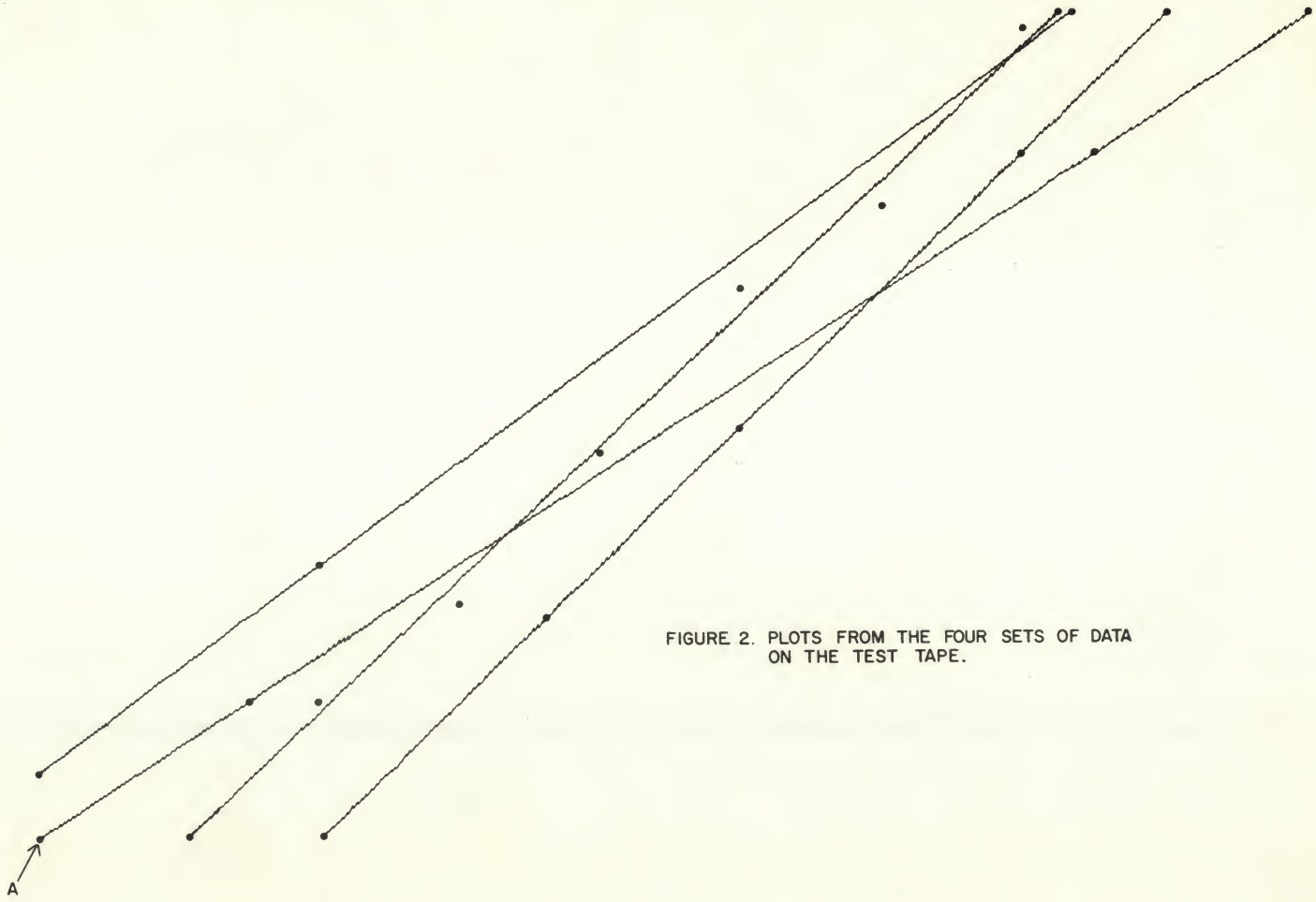

FIGURE 2. PLOTS FROM THE FOUR SETS OF DATA ON THE TEST TAPE.

Figure 3. Type-out for the four sets of data on the test tape for half page format.

```
B    +Ø.6666666E+ØØ
C    +Ø.4768371E-Ø6
SE   +Ø.4396219E-Ø6
PCT  +Ø.1762792E-Ø4
RSQ  +Ø.1ØØØØØØE+Ø1
R    +Ø.1ØØØØØØE+Ø1
SER  +Ø.ØØØØØØØE+ØØ
N    +Ø.5ØØØØØØE+Ø1

+Ø.6666671E+ØØ  +Ø.6666666E+ØØ  +Ø.6666674E+ØØ
X SCALE +Ø.9999999E-Ø1
Y SCALE +Ø.9999999E-Ø1



B    +Ø.9714274E+ØØ
C    -Ø.1Ø28566E+Ø1
SE   +Ø.1385Ø5ØE+ØØ
PCT  +Ø.234Ø5Ø8E+Ø2
RSQ  +Ø.9949423E+ØØ
R    +Ø.9974679E+ØØ
SER  +Ø.1911628E-Ø2
N    +Ø.7ØØØØØØE+Ø1

+Ø.3342856E+Ø1  +Ø.32Ø435ØE+Ø1  +Ø.348136ØE+Ø1
X SCALE +Ø.9999999E-Ø1
Y SCALE +Ø.9999999E-Ø1



B    +Ø.1ØØØØØØE+Ø1
C    -Ø.1999999E+Ø1
SE   +Ø.3693562E-Ø6
PCT  +Ø.288675ØE+Ø2
RSQ  +Ø.ØØØØØØØE+ØØ
R    +Ø.1ØØØØØØE+Ø1
SER  +Ø.ØØØØØØØE+ØØ
N    +Ø.5ØØØØØØE+Ø1

+Ø.25ØØØØØE+Ø1  +Ø.25ØØØØØE+Ø1  +Ø.25ØØØØØE+Ø1
X SCALE +Ø.9999999E-Ø1
Y SCALE +Ø.9999999E-Ø1


B    +Ø.75ØØØØØE+ØØ
C    +Ø.5ØØØØØØE+ØØ
SE   +Ø.ØØØØØØØE+ØØ
PCT  +Ø.ØØØØØØØE+ØØ
RSQ  +Ø.1ØØØØØØE+Ø1
R    +Ø.1ØØØØØØE+Ø1
SER  +Ø.ØØØØØØØE+ØØ
N    +Ø.4ØØØØØØE+Ø1

+Ø.425ØØØØE+Ø1 +Ø.425ØØØØE+Ø1  +Ø.425ØØØØE+Ø1
X SCALE +Ø.9999999E-Ø1
Y SCALE +Ø.9999999E-Ø1
```