

## PERCENT POINT PLOT

### PURPOSE

Generates a percent point plot (commonly referred to as quantile plot in the statistical graphics literature).

### DESCRIPTION

A percent point plot is a graphical data analysis technique for summarizing the distributional information of a variable. The X percent point of a data set is the point where X% of the data is below that point and (100-X)% of the data is above that point. It consists of:

Vertical axis = percent point;

Horizontal axis = percent (0 to 100).

If the value of 50 is chosen on the horizontal axis, then the corresponding value on the vertical axis is the estimated 50% point (that is, the median) from the data. The attributes of the plot can be set by the first setting of the LINE, CHARACTER, SPIKE, and BAR commands (and their corresponding attribute setting commands). This is demonstrated in the sample program below.

### SYNTAX 1

PERCENT POINT PLOT <y> <SUBSET/EXCEPT/FOR qualification>

where <y> is the variable of raw data;

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is used when you have raw data only.

### SYNTAX 2

PERCENT POINT PLOT <y> <x> <SUBSET/EXCEPT/FOR qualification>

where <y> is a variable of pre-computed frequencies;

<x> is a variable of distinct values;

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is used when you have pre-computed frequencies at each data value.

### EXAMPLE

```
PERCENT POINT PLOT Y
PERCENT POINT Y X
PERCENT POINT Y X SUBSET X > 2
```

### NOTE 1

DATAPLOT divides the original data variable into classes in the same manner as it does for a histogram or frequency polygon. The percent points are calculated at the mid-points of these histogram classes. The defaults are the same as for histograms (the class width is 0.3\*standard deviation, 6 classes above and 6 classes below the mean). This tends to leave a wide gap in the middle of the distribution and a large number of 0 and 100 percent point values. This can be improved by using the CLASS LOWER, CLASS UPPER, and CLASS WIDTH commands. A simple method for doing this is demonstrated in the sample program below.

### NOTE 2

The residual-fitted spread plot (or r-f spread plot) recommended by Bill Cleveland (see REFERENCE section) is based on percent point plots. It plots the percent point plot of the fitted values minus their mean and the residuals from a fit side by side. A common scale is used on the vertical scale. This plot gives a visual representation of how influential the fit is. If the spread of the fitted-value distribution is large compared to the spread of the residual distribution, then the fitted variables are influential. The second program example demonstrates this plot. In addition to linear fits, it can be applied to various other types of fits (e.g., nonlinear fits, lowess fits, spline fits).

### DEFAULT

None

### SYNONYMS

None

### RELATED COMMANDS

QUANTILE-QUANTILE PLOT	=	Generates a quantile-quantile plot.
HISTOGRAM	=	Generates a histogram.
FREQUENCY PLOT	=	Generates a frequency plot.
PROBABILITY PLOT	=	Generates a probability plot.

- PIE CHART = Generates a pie chart.
- PLOT = Generates a data or function plot.
- CLASS LOWER = Sets the lower class minimum for histograms, frequency plots, and pie charts.
- CLASS UPPER = Sets the upper class maximum for histograms, frequency plots, and pie charts.
- CLASS WIDTH = Sets the class width for histograms, frequency plots, and pie charts.

APPLICATIONS

Distributional Analysis

REFERENCE

“Graphical Methods for Data Analysis,” Chambers, Cleveland, Kleiner, and Tukey, Wadsworth, 1983.

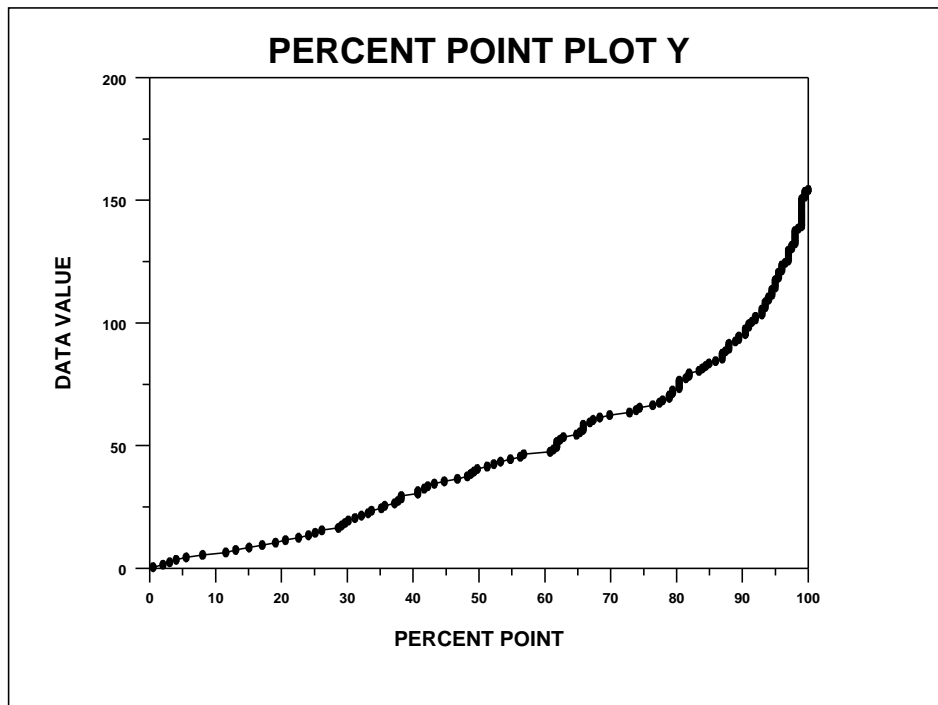
“Visualizing Data,” William Cleveland, Hobart Press, 1993.

IMPLEMENTATION DATE

Pre-1987

PROGRAM 1

```
SKIP 25
READ SUNSPOT2.DAT Y
LET ALOW = MINIMUM Y
LET AHIGH = MAXIMUM Y
CLASS LOWER ALOW
CLASS UPPER AHIGH
CLASS WIDTH 1.0
CHARACTER CIRCLE
CHARACTER FILL ON
CHARACTER SIZE 1.2
XILABEL PERCENT POINT; YILABEL DATA VALUE
TITLE AUTOMATIC
PERCENT POINT PLOT Y
```



## PROGRAM 2 (R-F SPREAD PLOT)

```

SKIP 25
READ BERGER1.DAT Y X BATCH
FIT Y X
LET M = MEAN PRED; LET Y2 = PRED - M
LET ALOW1 = MINIMUM Y2; LET ALOW2 = MINIMUM RES
LET ALOW = MIN(ALOW1,ALOW2)
LET AHIGH1 = MAXIMUM Y2; LET AHIGH2 = MAXIMUM RES
LET AHIGH = MAX(AHIGH1,AHIGH2)
CLASS WIDTH 1.0

.
MULTIPLY 1 2; MULTIPLY CORNER COORDINATES 0 0 100 95
FRAME CORNER COORDINATES 15 20 98 90

.
XLIMITS 0 100; XTIC OFFSET 2 2
MAJOR XTIC MARK NUMBER 6; MINOR XTIC MARK NUMBER 1
YLIMITS -25 30; YTIC OFFSET 2 2
LINE BLANK DASH; CHARACTER O
X1LABEL SIZE 2; X1LABEL PERCENT POINT
TITLE OFFSET 2; TITLE SIZE 3; TITLE FITTED VALUES
CLASS LOWER ALOW1; CLASS UPPER AHIGH1
PERCENT POINT PLOT Y2

.
TITLE RESIDUAL VALUES
CLASS LOWER ALOW2; CLASS UPPER AHIGH2
PERCENT POINT PLOT RES

.
END OF MULTIPLY
HEIGHT 4; JUSTIFICATION CENTER
MOVE 50 96; TEXT R-F SPREAD PLOT

```

