

... FREQUENCY PLOT**PURPOSE**

Generates a frequency plot.

DESCRIPTION

A frequency plot is a graphical data analysis technique for summarizing the distributional information of a variable. The response variable is divided into equal sized intervals (or bins). The number of occurrences of the response variable is calculated for each bin. The frequency plot then consists of:

- Vertical axis = frequencies or relative frequencies;
- Horizontal axis = response variable (i.e., the mid-point of each interval).

There are 4 types of frequency plots:

1. frequency plot (absolute counts);
2. relative frequency plot (convert counts to proportions);
3. cumulative frequency plot;
4. cumulative relative frequency plot.

The frequency plot and the histogram have the same information except the frequency plot has lines connecting the frequency values whereas the histogram has bars at the frequency values.

SYNTAX 1

```
FREQUENCY PLOT <y>                                <SUBSET/EXCEPT/FOR qualification>
RELATIVE FREQUENCY PLOT <y>                       <SUBSET/EXCEPT/FOR qualification>
CUMULATIVE FREQUENCY PLOT <y>                     <SUBSET/EXCEPT/FOR qualification>
CUMULATIVE RELATIVE FREQUENCY PLOT <y>           <SUBSET/EXCEPT/FOR qualification>>
```

where <y> is the variable of raw data values which will appear on the horizontal axis;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is used when you have raw data only.

SYNTAX 2

```
FREQUENCY PLOT <y> <x>                             <SUBSET/EXCEPT/FOR qualification>
RELATIVE FREQUENCY PLOT <y> <x>                   <SUBSET/EXCEPT/FOR qualification>
CUMULATIVE FREQUENCY PLOT <y> <x>                 <SUBSET/EXCEPT/FOR qualification>
CUMULATIVE RELATIVE FREQUENCY PLOT <y> <x>       <SUBSET/EXCEPT/FOR qualification>
```

where <y> is the variable of pre-computed frequencies to appear on the vertical axis;
<x> is the variable of distinct values to appear on the horizontal axis;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

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

EXAMPLES

```
FREQUENCY PLOT TEMP
RELATIVE FREQUENCY PLOT TEMP
CUMULATIVE FREQUENCY PLOT TEMP
CUMULATIVE RELATIVE FREQUENCY PLOT TEMP
FREQUENCY PLOT COUNTS STATE
RELATIVE FREQUENCY PLOT COUNTS STATE
CUMULATIVE FREQUENCY PLOT COUNTS STATE
CUMULATIVE RELATIVE FREQUENCY PLOT COUNTS STATE
```

NOTE 1

Although DATAPLOT does not have a FREQUENCY TABLE command, one can be generated with the following commands:

```
FREQUENCY PLOT Y
LET YFREQ = YPLOT
LET XVAL = XPLOT
```

Then the variables YFREQ and XVAL essentially contain a frequency table. There is a LET subcommand called FREQUENCY. However, it does not generate a frequency table in the sense that a frequency plot does.

NOTE 2

By default, DATAPLOT uses a class width of 0.3 times the standard deviation of the variable. Use the CLASS WIDTH command to override this default. DATAPLOT also tends to generate a large number of zero frequency classes at the lower and upper tails. This tends to compress the frequency plot on the horizontal axis. Use the XLIMITS command or the CLASS LOWER and CLASS UPPER commands to avoid plotting these zero frequency classes.

NOTE 3

If you want to overlay several frequency plots, specify the axis limits via the XLIMITS and YLIMITS commands before the first FREQUENCY PLOT command. Enter a PRE-ERASE OFF command after the first FREQUENCY PLOT command.

DEFAULT

None

SYNONYMS

A synonym for CUMULATIVE RELATIVE FREQUENCY PLOT is RELATIVE CUMULATIVE FREQUENCY PLOT.

RELATED COMMANDS

HISTOGRAM	=	Generates a histogram.
PIE CHART	=	Generates a pie chart.
PERCENT POINT PLOT	=	Generates a percent point plot.
PROBABILITY PLOT	=	Generates a probability plot.
PPCC PLOT	=	Generates probability plot correlation coefficient 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.
MINIMUM	=	Sets the frame minima for all plots.
MAXIMUM	=	Sets the frame maxima for all plots.
LIMITS	=	Sets the frame limits for all plots.
PLOT	=	Generates a data or function plot.

REFERENCE

Most introductory statistics book discuss frequency polygons and histograms.

“Multivariate Density Estimation,” David Scott, John Wiley, 1992 (chapter 4). This book discusses frequency polygons as “density estimators” and gives optimal criterion for selecting the class width.

APPLICATIONS

Exploratory Data Analysis, Distributional Analysis

IMPLEMENTATION DATE

Pre-1987

PROGRAM

```

SET READ FORMAT F10.1
SKIP 25
READ SUNSPOT.DAT Y
MULTIPLY 2 2
MULTIPLY CORNER COORDINATES 0 0 100 100
TITLE AUTOMATIC
XLIMITS 0 200
XTIC OFFSET 10 40
MAJOR XTIC MARK NUMBER 6
MINOR XTIC MARK NUMBER 3
FREQUENCY PLOT Y
RELATIVE FREQUENCY PLOT Y
CUMULATIVE FREQUENCY PLOT Y
CUMULATIVE RELATIVE FREQUENCY PLOT Y
END OF MULTIPLY
    
```

