

PROFILE PLOT

PURPOSE

Generates a profile plot.

DESCRIPTION

A profile plot is a graphical data analysis technique for examining the relative behavior of all variables in a multivariate data set. The profile plot consists of a sequence of equi-spaced vertical spikes with each spike representing a different variable in the multivariate data set. An individual profile plot examines the behavior of all such variables but only for a specified subset of the data (e.g., looking at all the attributes of car performance, but only for a particular car, such as Chevrolet). The total length of a given spike is uniformly set to unity for sake of reference. The “data length” of a given spike is proportional to the magnitude of the variable for the subset relative to the maximum magnitude of the variable across all subsets. (Thus we are looking at the ratio of the “local” value of the variable to the “global” maximum of the variable.) An interconnecting line cutting across each spike at the “data length” gives the profile plot its unique appearance and name.

SYNTAX

```
PROFILE PLOT <y1> <y2> ... <yk>                <SUBSET/EXCEPT/FOR qualification>
```

where <y1> ... <yk> are the response variables;

and where the <SUBSET/EXCEPT/FOR qualification> must be given. It is not optional for this command as it is for most other DATAPLOT commands.

EXAMPLES

```
PROFILE PLOT Y1 Y2 Y3 Y4 Y5 SUBSET AUTO 4
```

```
PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 SUBSET STATE 25
```

NOTE 1

A few variations of the profile plot exist, all of which DATAPLOT can easily generate by judicious use of the components in the LINE, CHARACTER, SPIKE, and BAR commands (and their attribute setting commands). For example, suppose there are k variables in the profile plot (and so k spikes), then

1. element 1 of LINES controls the appearance of the interconnecting line (this is usually SOLID);
2. element 1 of SPIKE and SPIKE LINE controls the existence and appearance of the spikes from the base out to the interconnecting line (some analysts prefer this to be SOLID, others prefer this to be DOTTED).

When using the SPIKE and SPIKE LINES commands in this context, note that both must be used if you want spikes to appear. The SPIKE (ON/OFF) command sets whether spikes will appear or not while the SPIKE LINES command sets the desired line type (SOLID, DOTTED, DASHED, etc.). Thus SPIKE ON and SPIKE LINES SOLID would for a profile plot turn the spikes on and set them to solid, respectively. Bars can be generated as an alternative to spikes. Some of the alternatives are demonstrated in the first program example below.

NOTE 2

The generation of multiple profile plots per page is typical (one profile plot for each subset of interest). This is easily done in DATAPLOT by using the PROFILE PLOT command in conjunction with the MULTIPLOT and LOOP commands. The following example generates 50 profile plots on the same page with each profile plot consisting of 6 variables:

```
MULTIPLOT 5 10
LOOP FOR K = 1 1 50
    PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 SUBSET STATE K
END OF LOOP
```

DEFAULT

None

SYNONYMS

None

RELATED COMMANDS

STAR PLOT	=	Generates a (multivariate) star plot.
ANDREWS PLOT	=	Generates an Andrews (multivariate) plot.
LINES	=	Sets the type for plot lines.

SPIKES	=	Sets the type for plot spikes.
MULTILOT	=	Allows multiple plots per page.
LOOP	=	Starts a loop (iteration).
^	=	Allows string and value substitution.

REFERENCE

“Graphical Methods of Data Analysis,” Chambers, Cleveland, Kleiner, and Tukey, Wadsworth, 1983 (pp. 162-163).

APPLICATIONS

Multivariate Analysis

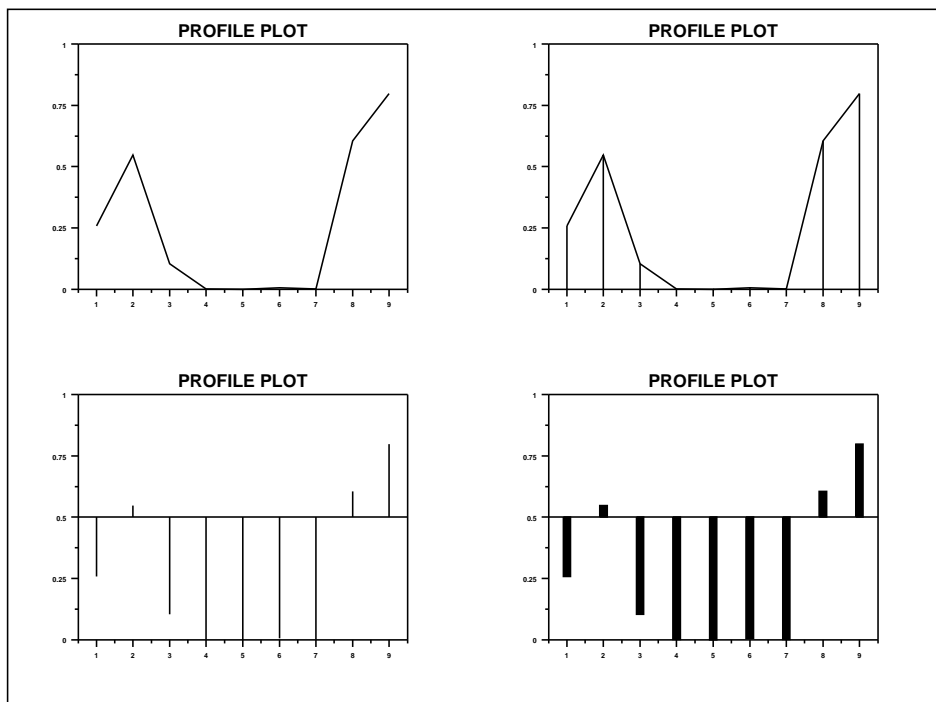
IMPLEMENTATION DATE

88/3

PROGRAM 1

```

DIMENSION 100 COLUMNS
SKIP 25
COLUMN LIMITS 20 132
READ AUTO79.DAT Y1 TO Y9
LET N = SIZE Y1
LET CAR = SEQUENCE 1 1 N
.
MULTIPLY 2 2; MULTIPLY CORNER COORDINATES 0 0 100 100
TITLE PROFILE PLOT
XLIMITS 1 9; XTIC OFFSET 0.5 0.5
MAJOR XTIC MARK NUMBER 9
YLIMITS 0 1
.
PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 SUBSET CAR 1
.
SPIKE ON
PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 SUBSET CAR 1
.
LINE BLANK; SPIKE BASE 0.5
PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 SUBSET CAR 1
LINE SOLID; MOVEDATA 0.5 0.5; DRAWDATA 9.5 0.5
.
LINE BLANK; SPIKE OFF
BAR ON; BAR BASE 0.5
BAR FILL ON
BAR WIDTH 0.2
PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 SUBSET CAR 1
LINE SOLID; MOVEDATA 0.5 0.5; DRAWDATA 9.5 0.5
END OF MULTIPLY
    
```



PROGRAM 2

```

DIMENSION 100 COLUMNS
SKIP 25
COLUMN LIMITS 20 132
READ AUTO79.DAT Y1 TO Y9
LET N = SIZE Y1; LET CAR = SEQUENCE 1 1 N
.
COLUMN LIMITS 1 19; SKIP 0
LOOP FOR K = 1 1 25
    LET K1 = 25+K
    ROW LIMITS K1 K1
    READ STRING AUTO79.DAT S^K
END OF LOOP
.
XLIMITS 1 9; XTIC OFFSET 0.5 0.5; MAJOR XTIC MARK NUMBER 9
YLIMITS 0 1; MAJOR YTIC MARK NUMBER 6; MINOR TIC MARK NUMBER 1
TIC LABEL SIZE 4
.
MULTIPLY 5 5; MULTIPLY CORNER COORDINATES 0 0 100 100
XFRAME OFF; Y2FRAME OFF; FRAME CORNER COORDINATES 15 10 95 95
.
LINE BLANK; BAR ON; BAR BASE 0.5; BAR FILL ON; BAR WIDTH 0.5
LEGEND 1 COORDINATES 55 5; LEGEND JUSTIFICATION CENTER; LEGEND SIZE 4
LOOP FOR K = 1 1 25
    LEGEND 1 ^S^K
    PROFILE PLOT Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 SUBSET CAR K
END OF LOOP
END OF MULTIPLY
    
```

