

**... STATISTIC PLOT****PURPOSE**

Generates a statistic versus index plot for a given statistic.

**DESCRIPTION**

A statistic plot consists of subsample statistic versus subsample index. The subsample statistic is the value of some statistic for the data in the subsample. The statistic plot is used to answer the question: "Does the subsample statistic change over different subsamples?"

The plot consists of:

Vertical axis = subsample statistic;  
Horizontal axis = subsample index.

In addition, a horizontal line is drawn representing the full sample statistic. The appearance of the 2 traces is controlled by the first 2 settings of the LINES, CHARACTERS, SPIKES, BARS, and similar attributes.

**SYNTAX 1**

<stat> STATISTIC PLOT <y> <x> <SUBSET/EXCEPT/FOR qualification>

where <stat> is one of the following statistics:

MEAN, MIDMEAN, MEDIAN, TRIMMED MEAN, WINDSORIZED MEAN,  
SUM, PRODUCT, SIZE (or NUMBER or SIZE),  
STANDARD DEVIATION, STANDARD DEVIATION OF MEAN,  
VARIANCE, VARIANCE OF THE MEAN,  
RELATIVE STANDARD DEVIATION, RELATIVE VARIANCE,  
RANGE, MIDRANGE, MAXIMUM, MINIMUM, EXTREME,  
LOWER HINGE, UPPER HINGE, LOWER QUARTILE, UPPER QUARTILE,  
<FIRST/SECOND/THIRD/FOURTH/FIFTH/SIXTH/SEVENTH/EIGHTH/ NINTH/TENTH> DECILE,  
SKEWNESS, KURTOSIS, NORMAL PPCC,  
AUTOCORRELATION, AUTOCOVARIANCE,  
SINE FREQUENCY, SINE AMPLITUDE,  
CP, CPK, EXPECTED LOSS, PERCENT DEFECTIVE,  
TAGUCHI SN0 (or SN), TAGUCHI SN+ (or SNL),  
TAGUCHI SN- (or SNS), TAGUCHI SN00 (or SN2);

<y> is the response (= dependent) variable;

<x> is the subsample identifier variable (this variable appears on the horizontal axis);

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

This syntax is used for statistics that require a single variable to compute.

**SYNTAX 2**

<stat> STATISTIC PLOT <y1> <y2> <x> <SUBSET/EXCEPT/FOR qualification>

where <stat> is one of the following statistics:

LINEAR INTERCEPT, LINEAR SLOPE, LINEAR RESSD, LINEAR CORRELATION;

<y1> is the first response (= dependent) variable;

<y2> is the second response (= dependent) variable;

<x> is the subsample identifier variable (this variable appears on the horizontal axis);

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

This syntax is used for variables that require two statistics to compute. If a linear fit is performed, the first variable is the dependent variable while the second variable is the independent variable.

**EXAMPLES**

MEAN PLOT Y X  
STANDARD DEVIATION PLOT Y X1

**NOTE 1**

The subcommands (e.g., MEAN PLOT) are documented individually.

## NOTE 2

Although DATAPLOT supports this command for a large number of statistics, there may be cases where you want it for an unsupported statistic. The following example shows how to compute the rank correlation (assume Y1 and Y2 are the response variables and TAG is the group identifier).

```
LET TAGDIST = DISTINCT TAG
LET NGROUP = SIZE TAGDIST
LOOP FOR K = 1 1 NGROUP
    LET IGROUP TAGDIST(K)
    LET A = RANK CORRELATION Y1 Y2 SUBSET TAG = IGROUP
    LET YNEW(K) = A
    LET XNEW(K) = K
END OF LOOP
LET A = RANK CORRELATION Y1 Y2
LET YNEW2 = DATA A A
LET XNEW2 = DATA 1 NGROUP
PLOT YNEW XNEW AND
PLOT YNEW2 XNEW2
```

This basic idea can be easily adapted to other statistics (even ones that are not built-in to DATAPLOT). It can also be adapted to statistics requiring any arbitrary number of variables to compute.

## DEFAULT

None

## SYNONYMS

On most of the commands, the word STATISTIC is optional and is usually omitted (e.g., the mean plot is documented under MEAN PLOT rather than MEAN STATISTIC PLOT). The one exception is for the AUTOCORRELATION STATISTIC PLOT where the word STATISTIC is required.

## RELATED COMMANDS

|               |   |   |
|---------------|---|---|
| CHARACTERS    | = | Sets the type for plot characters.          |
| LINES         | = | Sets the type for plot lines.               |
| BOX PLOT      | = | Generates a box plot.                       |
| CONTROL CHART | = | Generates a control chart.                  |
| PLOT          | = | Generates a data or function plot.          |
| SUMMARY       | = | Computes various statistics for a variable. |

## APPLICATIONS

Exploratory Data Analysis

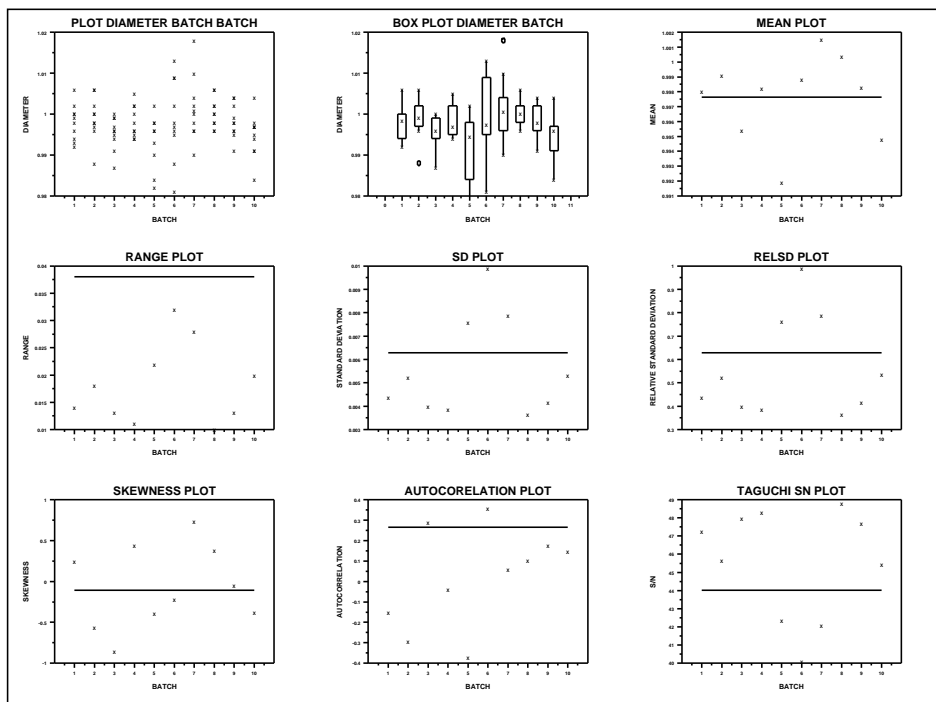
## IMPLEMENTATION DATE

88/2 (a few of the statistics have been added at various times since then, see the date on the individual command)

PROGRAM 1

```

SKIP 25
READ GEAR.DAT DIAMETER BATCH
.
TITLE AUTOMATIC
MULTIPLY 3 3 ; MULTIPLY CORNER COORDINATES 0 0 100 100
CHARACTER X ALL; LINE BLANK ALL
XTIC OFFSET 1 1
XILABEL BATCH; YILABEL DIAMETER
PLOT DIAMETER BATCH BATCH
CHARACTERS BOX PLOT; LINES BOX PLOT; FENCES ON
BOX PLOT DIAMETER BATCH
.
LINE BLANK SOLID
CHARACTER X BLANK
YILABEL MEAN; TITLE MEAN PLOT
MEAN PLOT DIAMETER BATCH
YILABEL RANGE; TITLE RANGE PLOT
RANGE PLOT DIAMETER BATCH
YILABEL STANDARD DEVIATION; TITLE SD PLOT
STANDARD DEVIATION PLOT DIAMETER BATCH
YILABEL RELATIVE STANDARD DEVIATION; TITLE RELSD PLOT
RELSD PLOT DIAMETER BATCH
YILABEL SKEWNESS; TITLE SKEWNESS PLOT
SKEWNESS PLOT DIAMETER BATCH
YILABEL AUTOCORRELATION; TITLE AUTOCORRELATION PLOT
AUTOCORRELATION STATISTICS PLOT DIAMETER BATCH
YILABEL S/N; TITLE TAGUCHI SN PLOT
TAGUCHI SN PLOT DIAMETER BATCH
END OF MULTIPLY
    
```



PROGRAM 2

```

SKIP 25
READ BERGER1.DAT Y X BATCH
.
TITLE AUTOMATIC
XTIC OFFSET 1 1
MULTIPLY 2 2 ; MULTIPLY CORNER COORDINATES 0 0 100 100
LINE BLANK SOLID
CHARACTER X BLANK
YLABEL SLOPE
LINEAR SLOPE PLOT Y X BATCH
YLABEL INTERCEPT
LINEAR INTERCEPT PLOT Y X BATCH
YLABEL CORRELATION
LINEAR CORRELATION PLOT Y X BATCH
YLABEL RESSD
LINEAR RESSD PLOT Y X BATCH
END OF MULTIPLY
    
```

