Describe – Multiple Response Variables

This procedure calculates rowwise statistics for multiple columns of numeric values. This type of data is typical of samples that are collected periodically over time, perhaps as input for an X-bar control chart.

The data for this analysis consist of *m* samples from a population, each consisting of one or more measurements. Let

$x_{i,j}$ = i-th measurement in sample *j*

n_j = size of sample *j*

Access

Highlight: two or more *Response* columns.

Select: *Describe* from the main menu.

Output Page 1: A scatterplot of the data by row number.

Output Page 2: Box-and-whisker plots for the data in each row.

Output Page 3: A plot of rowwise means with standard errors.

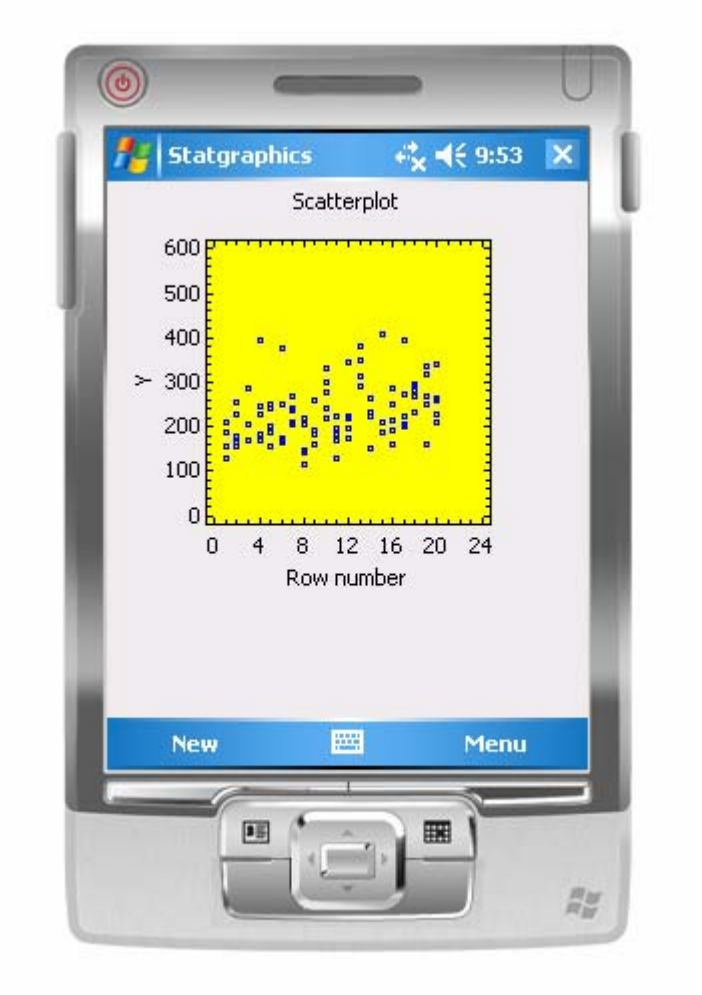
Sample Data

The file *wafers.sgm* contains data on *m* = 20 samples, each containing the measured resistivity of *n* = 5 silicon wafers. The first five samples are shown below:

Sample	Time	Wafer 1	Wafer 2	Wafer 3	Wafer 4	Wafer 5
1	1:00	211.2	128.4	154.9	186.8	156.7
2	2:00	155.4	255.7	165.8	227.9	178.8
3	3:00	171.0	205.6	285.7	284.9	207.7
4	4:00	181.1	169.0	393.1	246.2	228.2
5	5:00	158.0	239.6	200.2	250.8	189.1

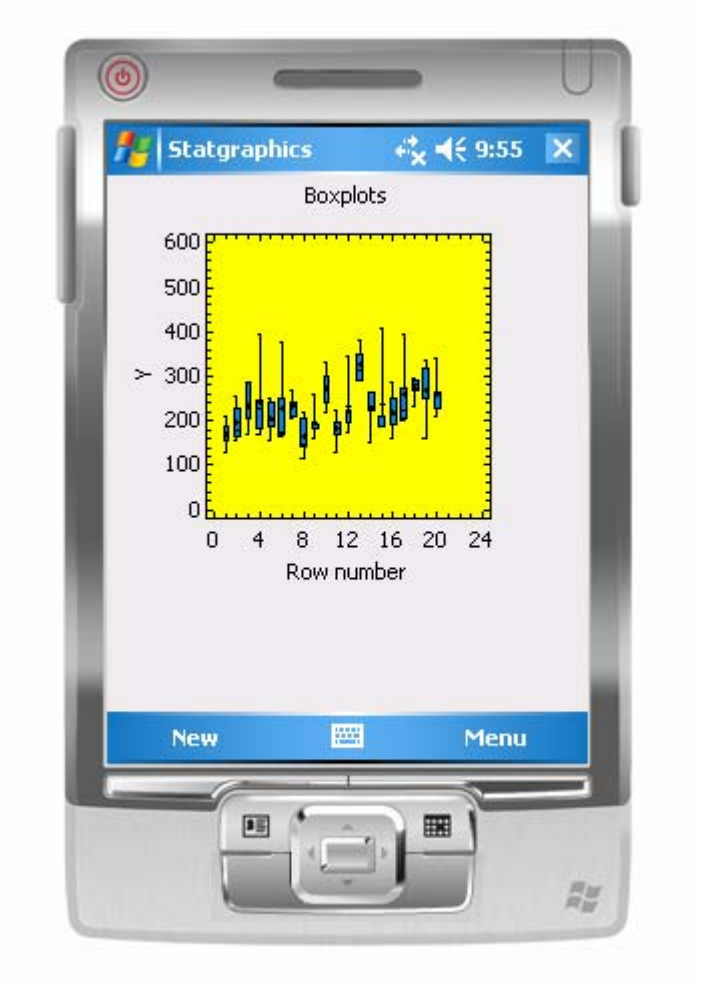
Scatterplot

The *Scatterplot* displays the data by row number.



Boxplots

The *Boxplots* page creates a box-and-whisker plot for the data in each row.

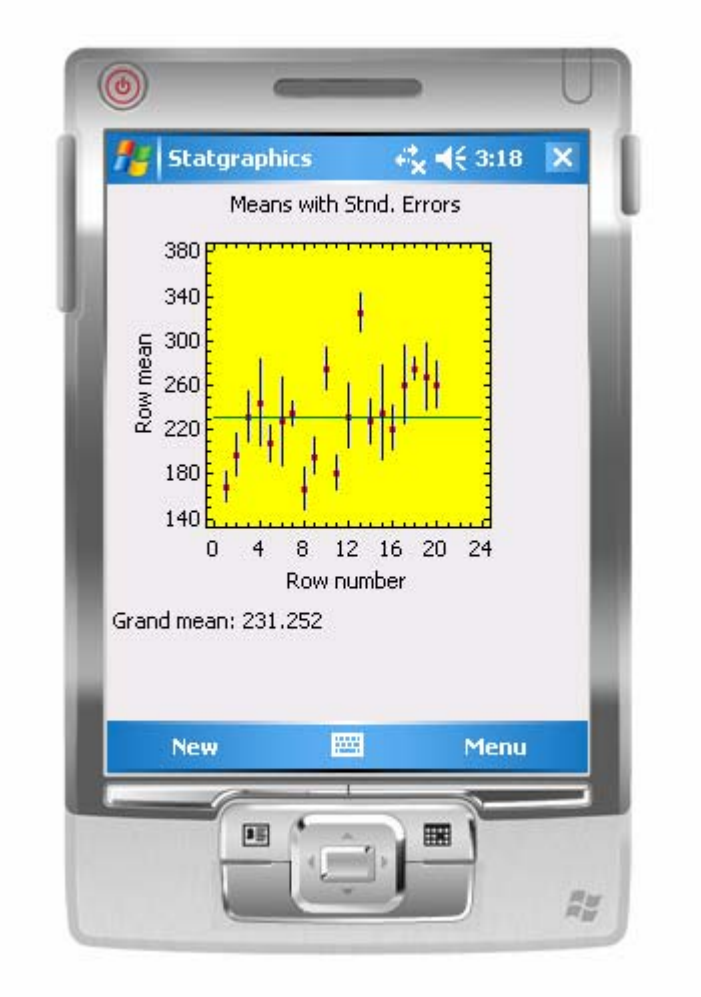


The plot is constructed in the following manner:

- A box is drawn extending from the *lower quartile* of each sample to the *upper quartile*. This is the interval covered by the middle 50% of the data values when sorted from smallest to largest.
- A horizontal line is drawn at the *median* (the middle value).
- A plus sign is placed at the location of the sample mean.
- Whiskers are drawn from the edges of the box to the largest and smallest data values.

Means Plot

The *Means Plot* page displays the mean of each row as a point symbol, with standard error bars above and below each mean.



The point symbols are located at the sample means

$$\bar{x}_j = \frac{\sum_{i=1}^{n_j} x_{i,j}}{n_j} \tag{1}$$

The error bars extend one standard error above and below the sample means, where the standard error of the *j*-th mean is calculated from

$$se_j = \sqrt{\frac{\sum_{i=1}^{n_j} \left(x_{i,j} - \bar{x}_j\right)^2 / \left(n_j - 1\right)}{n_j}} \tag{2}$$