

Factorial Repeated Measures ANOVA

Summary	1
Data Input.....	3
Analysis Summary	6
Analysis Options	7
Profile Plot	8
Sphericity Tests and Adjustments.....	9
Table of Means	11
Means Plot	13
Multiple Range Tests	15
Interaction Plot.....	18
Interaction Analysis	19
Trend Analysis	21
Trend Plot.....	22
Reports	23
Observed versus Predicted.....	24
Residual Plots.....	25
Unusual Residuals.....	28

Summary

The **Factorial Repeated Measures ANOVA** procedure analyzes experimental designs in which multiple measurements have been made on the same subjects, and in which one or two factors differentiate the subjects. In many cases, the measurements are taken sequentially in order to examine how the effect of the between-subject factors varies over time. The need for a special procedure stems from potential dependence between observations made on a given subject.

Included in the procedure is Mauchley's test for *sphericity*, which is a condition requiring that the variance between any two estimated treatment means be the same. Corrections to the standard F-test are included for cases when sphericity is not present.

The procedure assumes that:

1. There are one or more between-subject factors and a single within-subject factor, all with fixed effects and arranged in a factorial design.
2. An equal number of subjects are assigned to each combination of the between-subject factors.
3. Each subject is measured at each level of the within-subject factor.

Sample StatFolio: *repeat2.sgp*

Sample Data

The file *drugstudy.sgd* contains the results of an experiment reported by Milliken and Johnson (1984) in which 24 patients were each given one of 3 drugs (one of which was a control) in order to assess the effect of those drugs on their heart rate. Each patient's heart rate was measured 4 times at 5-minute increments after the drug was administered. A section of the data is shown below:

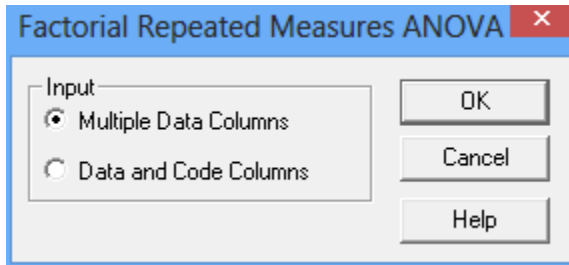
<i>Subject</i>	<i>Drug</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>
1	AX23	72	86	81	77
2	BWW9	85	86	83	80
3	CONTROL	69	73	72	74
4	AX23	78	83	88	81
5	BWW9	82	86	80	84
6	CONTROL	66	62	67	73
...		
8	CONTROL	71	70	65	63

Primary interest centers on comparing the mean heart rates at different points in time and in comparing the drugs.

Note: The observation for *Subject #8, Drug = Control* in Table 26.3 of Milliken and Johnson is incorrectly listed as 65. To obtain the results in subsequent tables of that chapter, it should be equal to 63.

Data Input

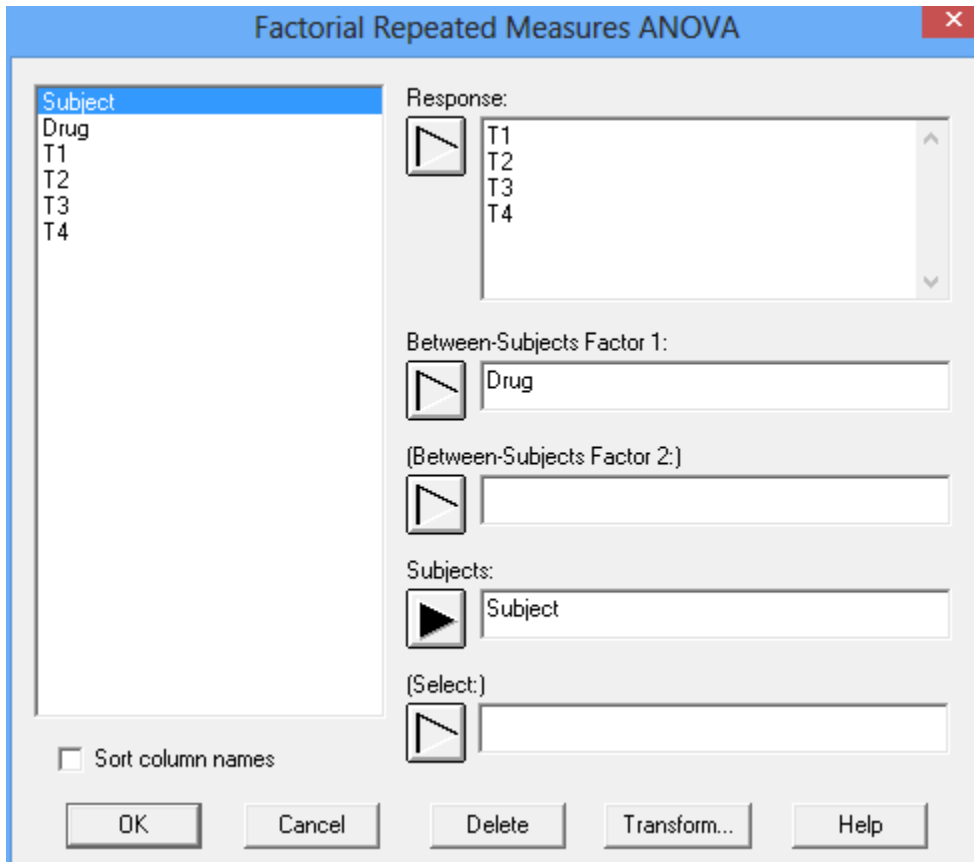
When the *Factorial Repeated Measures ANOVA* procedure is selected from the main menu, an initial dialog box is displayed on which to indicate how the data have been entered:



- **Multiple Data Columns:** indicates that each measurement taken from the same subject has been placed into a separate column.
- **Data and Code Columns:** indicates that all measurements have been placed into a single column, with a second column indicating which subject each observation comes from.

Multiple Data Columns

If the data have been placed in separate columns for each subject, the column names must be entered on the second dialog box:



- **Response:** two or more numeric columns containing the measurements, one row for each subject.
- **Between-Subject Factor 1:** numeric or non-numeric column containing the levels of a factor that varies between subjects.
- **Between-Subject Factor 2:** optional numeric or non-numeric column containing the levels of a second factor that varies between subjects.
- **Subjects:** numeric or non-numeric column identifying the subjects.
- **Select:** subset selection.

After pressing *OK*, the *Analysis Options* dialog box is displayed on which names may be entered for the response variable and the within-subject factor:

Factorial Repeated Measures ANOVA Options

Response: Heart rate

Within-Subject Factor: Time

Factor: Drug

Model

Factor levels (drag to change order):

OK

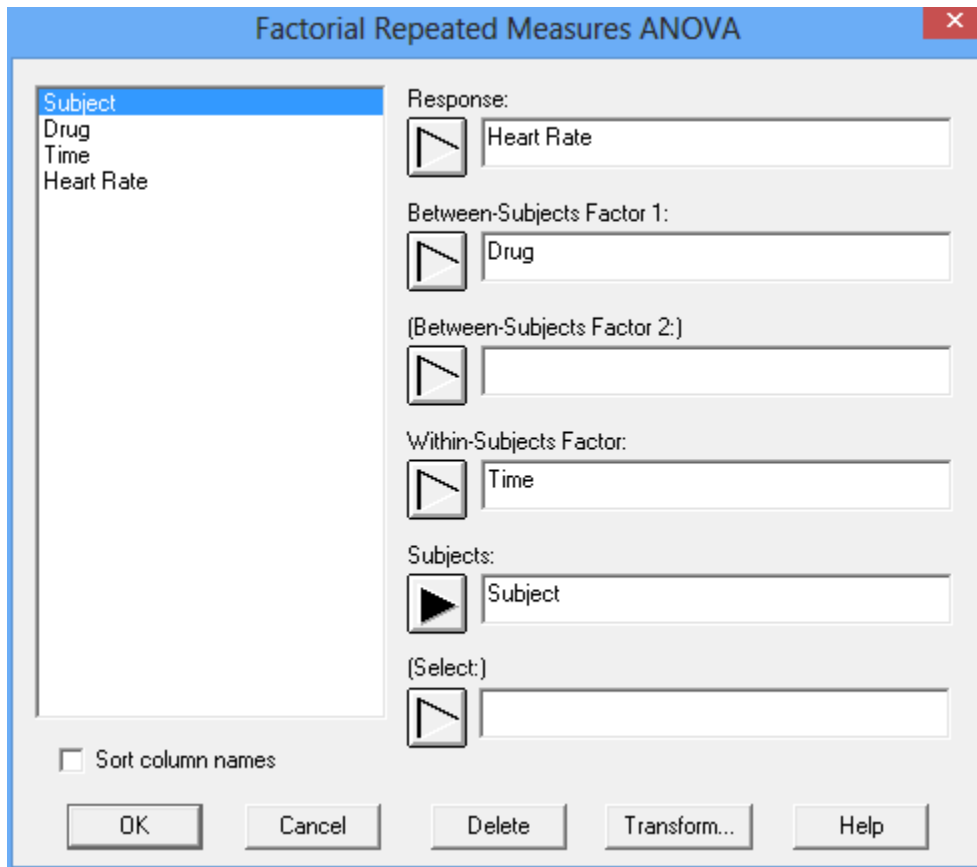
Cancel

Help

These labels will be used in the output tables and graphs. Note: the *factor level* fields will be populated once the data are read.

Data and Code Columns

If the data from all subjects have been placed into a single column, enter the name of that column and the column identifying the factor levels:



- **Response:** numeric column containing the measurements to be analyzed.
- **Between-Subject Factors:** numeric or non-numeric columns containing the levels of one or more factors that vary between subjects.
- **Within-Subject Factor:** numeric or non-numeric column identifying the levels that define the repeated factor.
- **Subjects:** numeric or non-numeric column identifying the subjects.
- **Select:** subset selection.

Analysis Summary

The *Analysis Summary* shows information about the fitted model. The top section of the output is shown below:

<u>Factorial Repeated Measures ANOVA</u>					
Dependent variable: Heart rate					
Between-subject factors:					
Drug					
Within-subject factor: Time					
Subjects: Subject					
Number of observations: 96					
Analysis of Variance for Heart rate					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	4487.94	32	140.248	18.83	0.0000
Residual	469.219	63	7.44792		
Total (Corr.)	4957.16	95			
Type III Sums of Squares					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Drug	1333.0	2	666.5	5.99	0.0088
ERROR(Subject)	2337.91	21	111.329	14.95	0.0000
Time	289.615	3	96.5382	12.96	0.0000
Time*Drug	527.417	6	87.9028	11.80	0.0000
ERROR(Time)	469.219	63	7.44792		
Total (corrected)	4957.16	95			

Included in the output are:

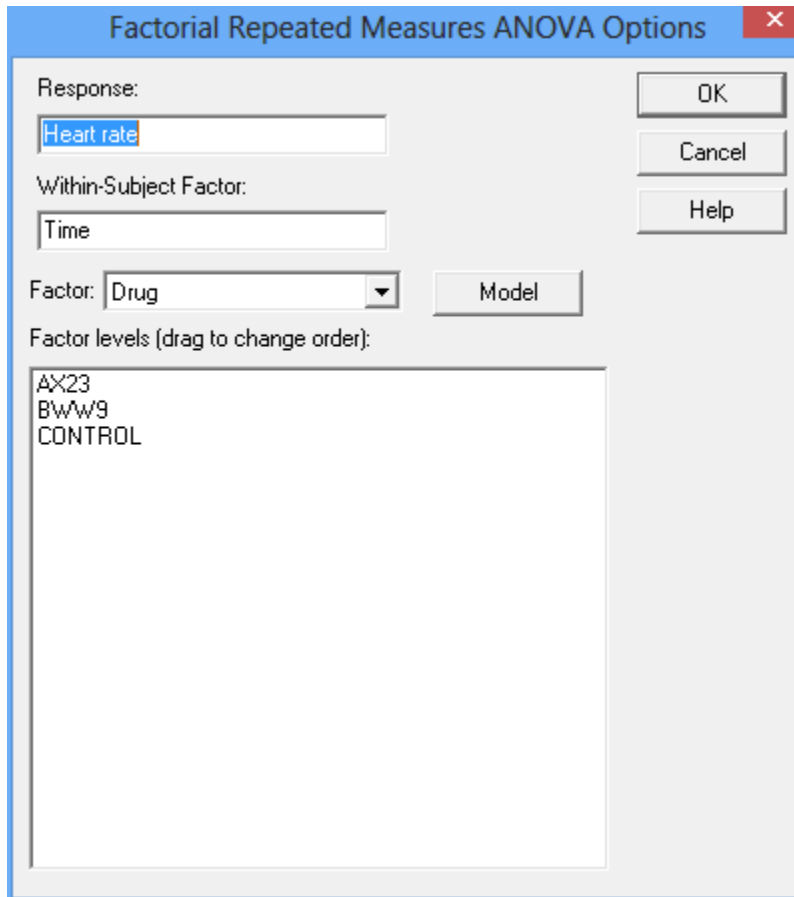
- **Analysis of Variance:** a decomposition of the sum of squares for *Y* into components for the model and for the residuals. The F-test tests the statistical significance of the model as a whole. A small P-value (less than 0.05 if operating at the 5% significance level) indicates that at least one factor in the model is significantly related to the dependent variable. In the current example, the model is highly significant.
- **Type III Sums of Squares:** decomposition of the model sum of squares into components for each factor. Small P-values indicate significant effects. In this example, all four effects are highly significant.

The rest of the *Analysis Summary* output is identical to that of the *General Linear Models* procedure and is described in the PDF file for that procedure.

It should be noted that the P-Values for any effect involving the within-subject factor (*Time* and *Drug*Time*) shown in this table are only appropriate if the data exhibit a property known as *sphericity*. Sphericity is present if the variance of the difference between any pair of estimated level means is the same. A later section shows how to test the assumption of sphericity and how to correct for its absence if necessary.

Analysis Options

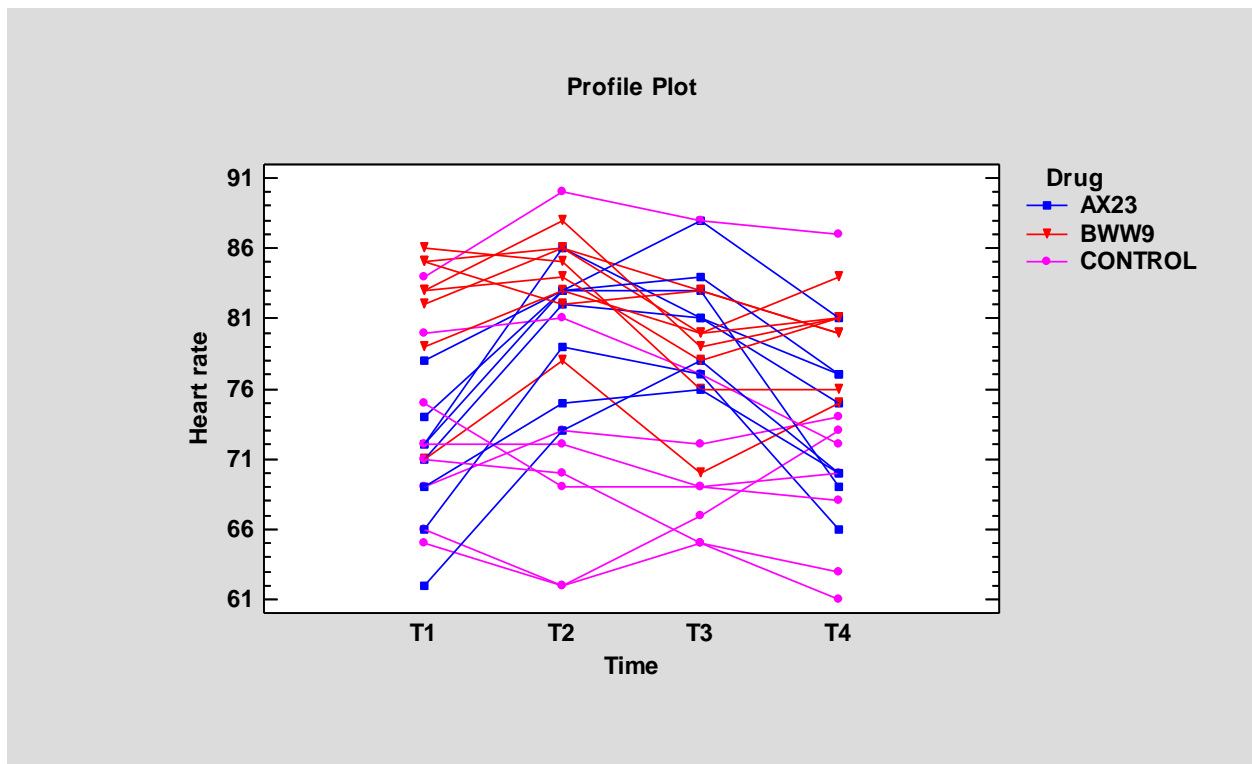
After the initial summary has been displayed, the *Analysis Options* dialog box may be used to reorder the factor levels:



To change the order of the factor levels, drag them to new positions with your mouse.

Profile Plot

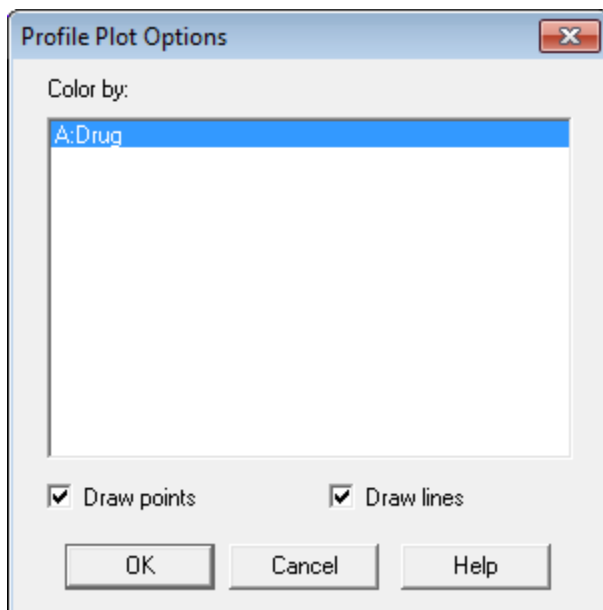
The *Profile Plot* displays the repeated measurements for each subject:



The profile lines are color-coded according to the level of a selected between-subjects factor.

Pane Options

The *Pane Options* dialog box lets you select the between-subjects factor or combination of factors used to color-code the profile lines:



Sphericity Tests and Adjustments

As noted earlier, the F-tests for effects involving the within-subject factor require that a condition known as *sphericity* be present. This pane conducts a test known as Mauchly’s test to determine whether the assumption of sphericity is violated. It generates the following output:

Sphericity Tests and Adjustments						
Mauchly's Sphericity Test						
<i>W</i>	<i>Chi-square</i>	<i>D.f.</i>	<i>P-value</i>			
0.631214	9.07439	5.0	0.106134			
Epsilon						
<i>Huynh-Feldt</i>	<i>Greenhouse-Geisser</i>	<i>Lower-bound</i>				
0.963045	0.776916	0.333333				
Tests of Within-Subject Effects						
<i>Source</i>	<i>Sphericity Correction</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Time	None	289.615	3.0	96.5382	12.96	0.0000
	Huynh-Feldt	289.615	2.88913	100.243	12.96	0.0000
	Greenhouse-Geisser	289.615	2.33075	124.258	12.96	0.0000
	Lower-bound	289.615	1.0	289.615	12.96	0.0017
Time*Drug	None	527.417	6.0	87.9028	11.80	0.0000
	Huynh-Feldt	527.417	5.77827	91.2759	11.80	0.0000
	Greenhouse-Geisser	527.417	4.6615	113.143	11.80	0.0000
	Lower-bound	527.417	2.0	263.708	11.80	0.0004
Residual	None	469.219	63.0	7.44792		
	Huynh-Feldt	469.219	60.6718	7.73372		
	Greenhouse-Geisser	469.219	48.9457	9.58651		
	Lower-bound	469.219	21.0	22.3437		
Multivariate Tests						
<i>Effect</i>	<i>Test</i>	<i>Value</i>	<i>F</i>	<i>Hypothesis D.F.</i>	<i>Error D.F.</i>	<i>P-Value</i>
Time	Wilks' lambda	0.463964	7.31714	3.0	19.0	0.0019
	Pillai trace	0.536036	7.31714	3.0	19.0	0.0019
	Hotelling-Lawley trace	1.15534	7.31714	3.0	19.0	0.0019
	Roy's largest root	1.15534	7.31714	3.0	19.0	0.0019
Time*Drug	Wilks' lambda	0.535332	2.32274	6.0	38.0	0.0523
	Pillai trace	0.531079	2.41029	6.0	40.0	0.0441
	Hotelling-Lawley trace	0.743946	2.23184	6.0	36.0	0.0622
	Roy's largest root	0.491595	3.2773	3.0	20.0	0.0423

The top section shows the results of Mauchly’s sphericity test. Mauchly’s test calculates the test statistic W defined by:

$$W = \frac{\det(A)}{(\text{trace}(A) / d)^d} \tag{1}$$

where $A = M^T S M$, S is the error sum of squares and cross product matrix, and M is an orthonormal matrix based on the within-subject effects being tested. A chi-square statistic is then calculated from

$$\chi^2 = -v \left(1 - \frac{2d^2 + d + 2}{6vd} \right) \log(W) \tag{2}$$

where q is the number of measurements made on each subject, $d = q - 1$, and ν is the degrees of freedom associated with subjects. χ^2 is compared to a chi-square distribution with $d(d+1)/2 - 1$ degrees of freedom. If the P-Value for Mauchley's test is small, then the sphericity assumption is rejected. For the sample data, P is above 0.10, so the sphericity assumption is not rejected at the 10% level of significance.

If sphericity is not present, the degrees of freedom for the F-test may be adjusted by multiplying them by an adjustment factor ε . The output shows the results of F-tests using three possible values for ε (from least conservative to most conservative):

$$\text{Huynh-Feldt: } \varepsilon_{hf} = \min\left(\frac{nd\varepsilon_{gg} - 2}{d\nu - d^2\varepsilon_{gg}}, 1\right) \quad (3)$$

$$\text{Greenhouse-Geisser: } \varepsilon_{gg} = \frac{(\text{trace}(A))^2}{d \times \text{trace}(A^T A)} \quad (4)$$

$$\text{Lower-Bound: } \varepsilon_{lb} = \frac{1}{d} \quad (5)$$

The lower-bound correction is the most severe, but is rarely used. The Greenhouse-Geisser correction is less conservative, but may still overcorrect the degrees of freedom when epsilon is close to 1. The Huynh-Feldt version is less conservative and may be preferable when the true epsilon is 0.75 or higher.

The bottom section shows several tests of the hypothesis that the means at each level of the within-subject factor are all the same, calculated by performing a multivariate analysis of variance (MANOVA) on differences between the level means. The tests do not assume the existence of sphericity and are thus correct regardless of whether that condition exists. If sphericity is present, however, they may not be as powerful as the usual univariate F test. For the sample data, the small P-values for the multivariate tests indicate that there are significant differences among the means at the 4 *times*. In addition, the interaction between *Time* and *Drug* has a P-value close to 0.05, where that interaction would be statistically significant at the 5% level. Further details about the four tests may be found in the documentation for the *General Linear Models* analysis.

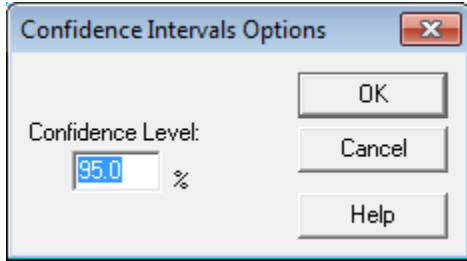
Table of Means

This pane displays the estimated means at each level of the factors, together with estimated standard errors and confidence intervals. A portion of the output is shown below:

Table of Least Squares Means for Heart rate with 95.0 Percent Confidence Intervals					
<i>Level</i>	<i>Count</i>	<i>Mean</i>	<i>Std. Error</i>	<i>Lower Limit</i>	<i>Upper Limit</i>
GRAND MEAN	96	76.4063	0.278536	75.8496	76.9629
Drug					
AX23	32	76.2813	1.86522	72.4023	80.1602
BWW9	32	81.0313	1.86522	77.1523	84.9102
CONTROL	32	71.9063	1.86522	68.0273	75.7852
Subject within Drug					
1 AX23	4	79.0	1.36454	76.2732	81.7268
2 BWW9	4	83.5	1.36454	80.7732	86.2268
3 CONTROL	4	72.0	1.36454	69.2732	74.7268
4 AX23	4	82.5	1.36454	79.7732	85.2268
5 BWW9	4	83.0	1.36454	80.2732	85.7268
6 CONTROL	4	67.0	1.36454	64.2732	69.7268
7 AX23	4	77.25	1.36454	74.5232	79.9768
8 BWW9	4	73.5	1.36454	70.7732	76.2268
9 CONTROL	4	87.25	1.36454	84.5232	89.9768
10 AX23	4	76.75	1.36454	74.0232	79.4768
11 BWW9	4	82.75	1.36454	80.0232	85.4768
12 CONTROL	4	77.5	1.36454	74.7732	80.2268
13 AX23	4	72.0	1.36454	69.2732	74.7268
14 BWW9	4	80.75	1.36454	78.0232	83.4768
15 CONTROL	4	70.75	1.36454	68.0232	73.4768
16 AX23	4	79.5	1.36454	76.7732	82.2268
17 BWW9	4	82.5	1.36454	79.7732	85.2268
18 CONTROL	4	63.25	1.36454	60.5232	65.9768
19 AX23	4	70.75	1.36454	68.0232	73.4768
20 BWW9	4	80.75	1.36454	78.0232	83.4768
21 CONTROL	4	70.25	1.36454	67.5232	72.9768
22 AX23	4	72.5	1.36454	69.7732	75.2268
23 BWW9	4	81.5	1.36454	78.7732	84.2268
24 CONTROL	4	67.25	1.36454	64.5232	69.9768
Time					
T1	24	75.0	0.557073	73.8868	76.1132
T2	24	78.9583	0.557073	77.8451	80.0716
T3	24	77.0417	0.557073	75.9284	78.1549
T4	24	74.625	0.557073	73.5118	75.7382
Time by Drug					
T1 AX23	8	70.5	0.964878	68.5718	72.4282
T1 BWW9	8	81.75	0.964878	79.8218	83.6782
T1 CONTROL	8	72.75	0.964878	70.8218	74.6782
T2 AX23	8	80.5	0.964878	78.5718	82.4282
T2 BWW9	8	84.0	0.964878	82.0718	85.9282
T2 CONTROL	8	72.375	0.964878	70.4468	74.3032
T3 AX23	8	81.0	0.964878	79.0718	82.9282
T3 BWW9	8	78.625	0.964878	76.6968	80.5532
T3 CONTROL	8	71.5	0.964878	69.5718	73.4282
T4 AX23	8	73.125	0.964878	71.1968	75.0532
T4 BWW9	8	79.75	0.964878	77.8218	81.6782
T4 CONTROL	8	71.0	0.964878	69.0718	72.9282

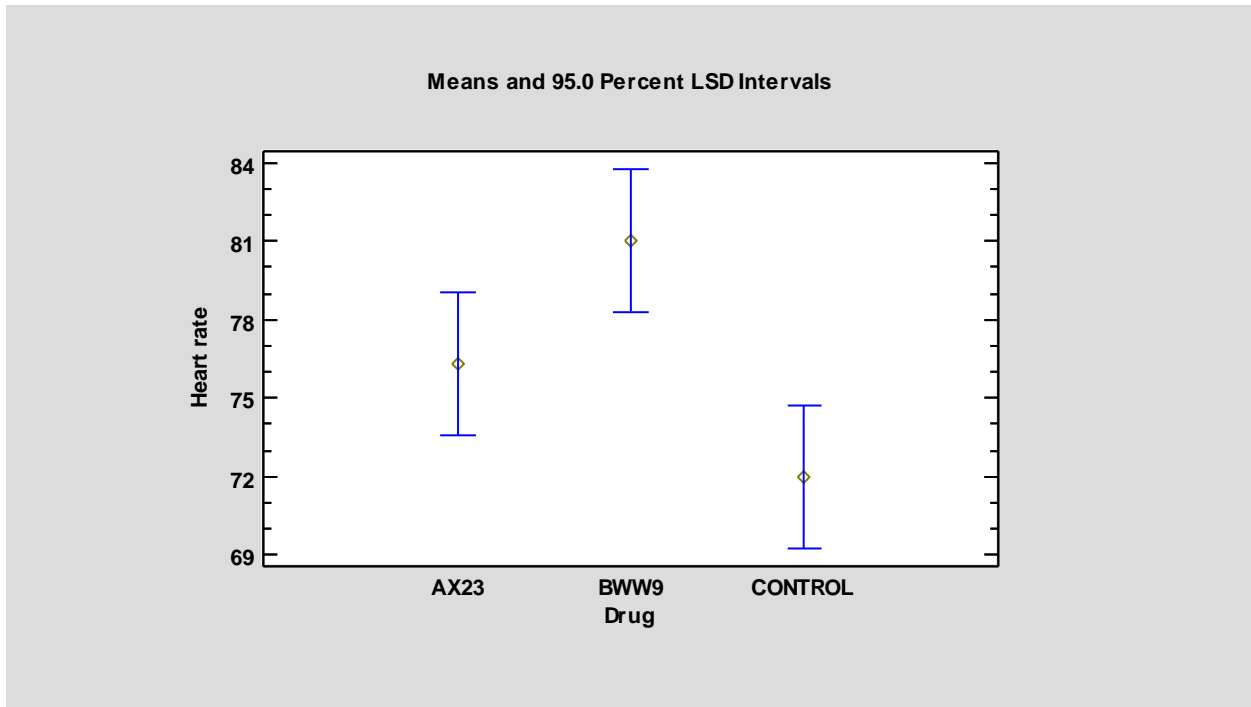
Pane Options

You may use this dialog box to change the confidence level:



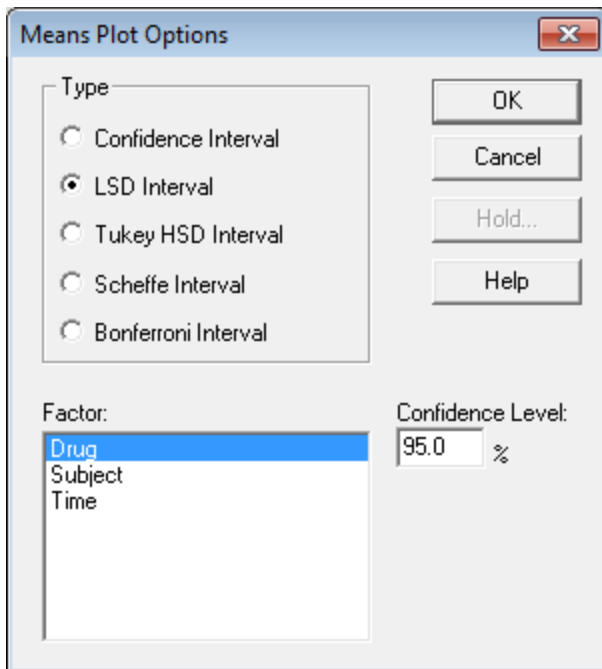
Means Plot

The level means may be plotted together with uncertainty intervals:



The type of intervals and the factor displayed are controlled by the *Pane Options* dialog box.

Pane Options



- **Type:** the method used to construct the intervals.

- **Factor:** the factor to be plotted.
- **Confidence Level:** the level of confidence associated with each interval.

The types of intervals that may be selected are:

Confidence intervals - displays confidence intervals for the level means using the estimated standard errors.

LSD intervals - designed to compare any pair of means with the stated confidence level.

Tukey HSD Intervals - designed for comparing all pairs of means. The stated confidence level applies to the entire family of pairwise comparisons.

Scheffe Intervals - designed for comparing all contrasts. Not usually relevant here.

Bonferroni Intervals - designed for comparing a selected number of contrasts. Tukey's intervals are usually tighter.

Each of the intervals is formed by adding a multiple of the least squares standard error to the estimated mean. The multiple depends upon the method used, as described in the *General Linear Models* documentation.

Multiple Range Tests

For factors that show significant P-Values in the ANOVA, a further analysis can be performed by selecting the *Multiple Range Tests*.

Multiple Comparisons for Heart rate by Drug				
Method: 95.0 percent LSD				
Drug	Count	LS Mean	LS Sigma	Homogeneous Groups
CONTROL	32	71.9063	1.86522	X
AX23	32	76.2813	1.86522	XX
BWW9	32	81.0313	1.86522	X

Contrast	Sig.	Difference	+/- Limits
AX23 - BWW9		-4.75	5.48564
AX23 - CONTROL		4.375	5.48564
BWW9 - CONTROL	*	9.125	5.48564

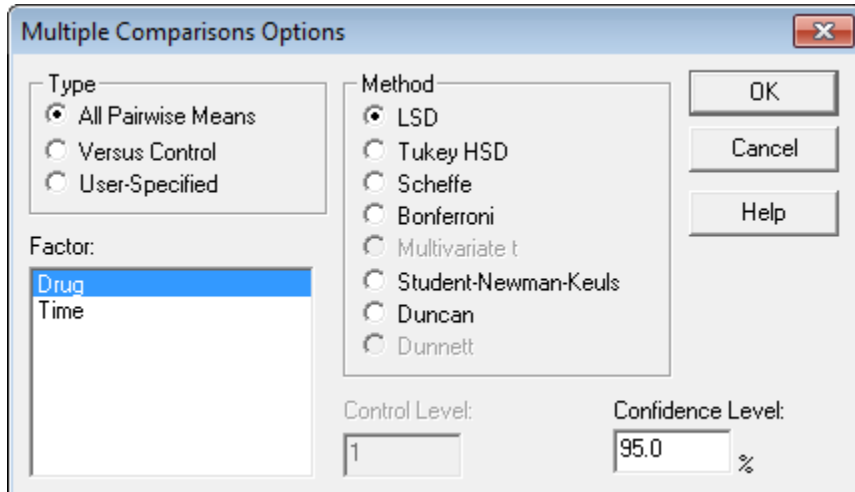
* denotes a statistically significant difference.

The top half of the table displays each of the estimated level means in increasing order of magnitude. It shows:

- **Count** - the number of observations at the specified level of the factor.
- **LS Mean** - the estimated least squares mean. In balanced designs such as these, the least squares mean is equivalent to the average of all observations at the indicated factor level.
- **LS Sigma** – the estimated standard error of the least squares mean.
- **Homogeneous groups** - a graphical illustration of which means are significantly different from which others, based on the contrasts displayed in the second half of the table. Each column of X's indicates a group of means within which there are no statistically significant differences. In the example, there are 2 columns of X's. The first column contains an X for AX23 and CONTROL, indicating they are not significantly different. The second column contains an X for AX23 and BWW9, indicating they are not significantly different. Since no column contains both BWW9 and CONTROL, they are significantly different.

The second half of the table displays a comparison between each pair of level means.

- **Difference** - the difference between the two least squares means.
- **Limits** - an interval estimate of that difference, using the currently selected multiple comparisons procedure.
- **Sig.** - An asterisk is placed next to any difference that is statistically significantly different from 0 at the currently selected significance level, i.e., any interval that does not contain 0.

Pane Options

- **Type:** type of contrasts to be created.
- **Factor:** factor to be analyzed.
- **Method:** the method used to make the multiple comparisons.
- **Control Level:** if *Type* is set to *Versus Control*, the number of the level against which all other levels will be compared.
- **Confidence Level:** the level of confidence used by the selected multiple comparison procedure.

The available methods are:

LSD - forms a confidence interval for each pair of means at the selected confidence level using Student's t distribution. This procedure is due to Fisher and is called the *Least Significant Difference* procedure, since the magnitude of the limits indicates the smallest difference between any two means that can be declared to represent a statistically significant difference. It should only be used when the F-test in the ANOVA table indicates significant differences amongst the level means. The probability of making a Type I error α applies to each pair of means separately. If making more than one comparison, the overall probability of calling at least one pair of means significantly different when they are not may be considerably larger than α .

Tukey HSD - widens the intervals to allow for multiple comparisons amongst all pairs of means using Tukey's T. Tukey called his procedure the *Honestly Significant Difference* procedure since it controls the experiment-wide error rate at α . If all of the means are equal, the probability of declaring *any* of the pairs to be significantly different in the entire experiment equals α . Tukey's procedure is more conservative than Fisher's LSD procedure, since it makes it harder to declare any particular pair of means to be significantly different.

Scheffe - designed to permit the estimation of all possible contrasts amongst the sample means (not just pairwise comparisons).

Bonferroni - designed to permit the estimation of any preselected number of contrasts. These limits are usually wider than Tukey's limits when all pairwise comparisons are being made.

Multivariate t – designed for sets of linearly independent combinations of the means.

Student-Newman-Keuls - Unlike the previous methods, this method does not create intervals for the pairwise differences. Instead, it sorts the means in increasing order and then begins to separate them into groups according to values of the Studentized range distribution. Eventually, the means are separated into homogeneous groups within which there are no significant differences.

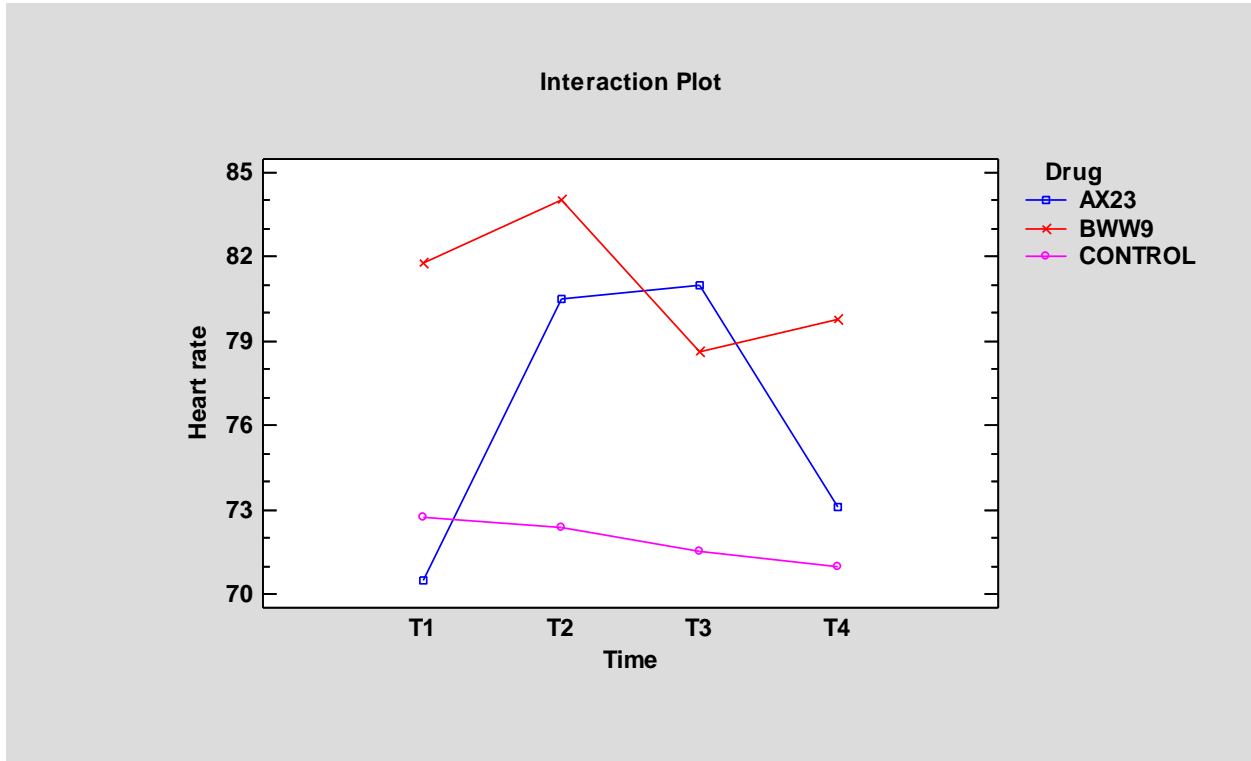
Duncan - similar to the Student-Newman-Keuls procedure, except that it uses a different critical value of the Studentized range distribution when defining the homogeneous groups. A detailed discussion of the Duncan and Student-Newman-Keuls procedures is given by Milliken and Johnson (1992).

Dunnett – designed for pairwise comparisons when one level is a control.

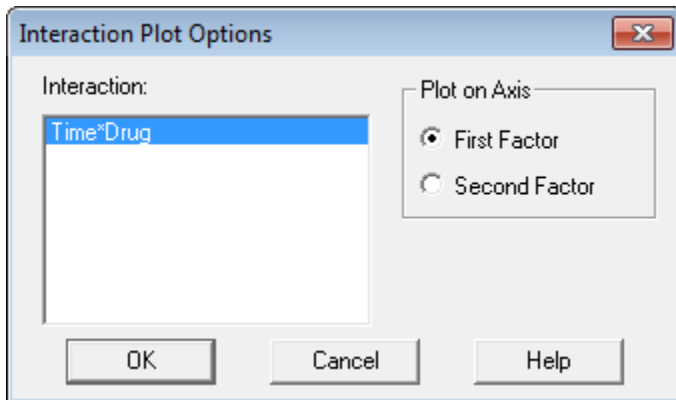
For details on creating a user-specified contrast, refer to the *General Linear Models* procedure documentation.

Interaction Plot

The *Interaction Plot* plots the mean response at all combinations of any 2 between-subject or within-subject factors. For example, the plot below displays means at the 12 combinations of *Drug* and *Time*:



Pane Options



- **Interaction:** the interaction to be plotted.
- **Plot on Axis:** the factor in the interaction that is to be displayed on the X axis. The other factor will be used to define the lines.

Interaction Analysis

The *Interaction Analysis* displays the mean response at all combinations of one between-subject factor and the within-subject factor:

Interaction Analysis for Heart rate by Drug*Time				
Comparison of Time Means at Same Drug				
	AX23	BWW9	CONTROL	
T1	70.5 (a)	81.75 (a,b)	72.75 (a)	
T2	80.5 (b)	84.0 (a)	72.375 (a)	
T3	81.0 (b)	78.625 (c)	71.5 (a)	
T4	73.125 (a)	79.75 (b,c)	71.0 (a)	
Method: 95.0 percent LSD (2.72683)				
Comparison of Drug Means at Same Time				
	T1	T2	T3	T4
AX23	70.5 (a)	80.5 (a)	81.0 (a)	73.125 (a)
BWW9	81.75 (b)	84.0 (a)	78.625 (a)	79.75 (b)
CONTROL	72.75 (a)	72.375 (b)	71.5 (b)	71.0 (a)
Method: 95.0 percent LSD (5.97171)				

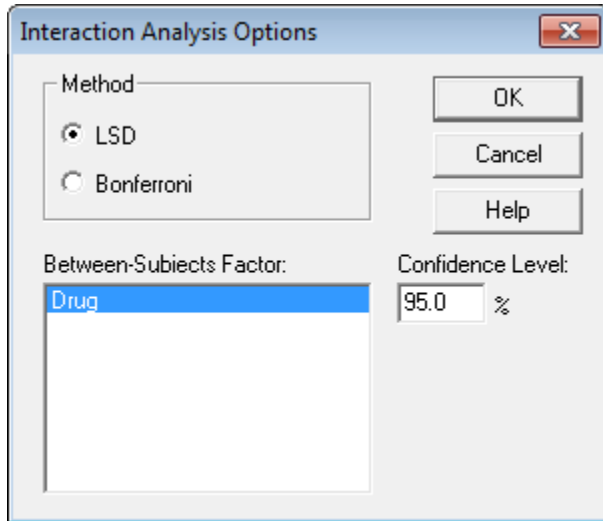
The output also contains comparisons between the means. In the top table, the within-subjects means are compared at each level of the between-subjects factor. In the bottom table, the between-subjects means are compared at each level of the within-subjects factor. Each column is divided into homogeneous groups, which are groups within which the means are not significantly different using the LSD (Least Significant Difference) procedure. Any 2 means within a column that have the same letter are not significantly different at the indicated confidence level.

For example, the first column in the top table shows the observed means at each *Time* level for the AX23 drug. Since times T1 and T4 both have an “a”, they are not significantly different. Similarly, times T2 and T3 both have a “b” and are also not significantly different. On the other hand, times T1 and T2 do not share a common letter and thus are significantly different.

It is possible for a single mean to have more than one letter, as for drug *BWW9* in the top table. That column indicates 3 homogeneous groups:

- (a) T1 and T2
- (b) T1 and T4
- (c) T3 and T4

For that drug, the data are not sufficient to divide the times into distinct groups.

Pane Options

- **Method:** the method used to compare the means. See the discussion under *Multiple Range Tests* above.
- **Between-Subjects Factor:** the between-subjects factor to be included in the table.
- **Confidence level:** the confidence level for the LSD or Bonferroni procedure.

Trend Analysis

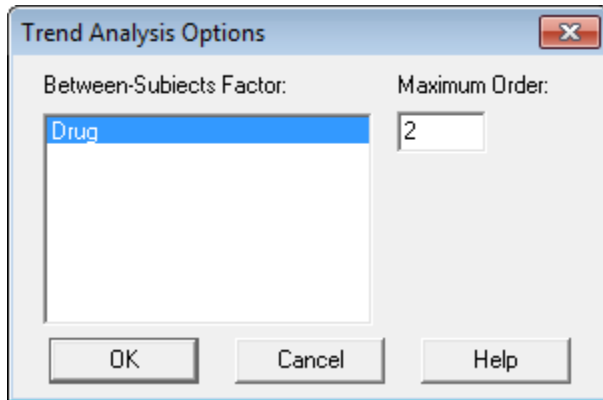
If the order of the conditions has meaning, as it does for the sample data, the pattern of the means exhibited in the *Profile Plot* can be examined using a trend analysis. A trend analysis uses orthogonal polynomials to decompose the pattern of the means into linear, quadratic, cubic and higher order components. It produces a table similar to that shown below:

Drug	Trend	Contrast	Std. error	t	P-value
AX23	linear	1.87271	0.964878	1.94	0.0568
AX23	quadratic	-8.9375	0.964878	-9.26	0.0000
BWW9	linear	-2.54353	0.964878	-2.64	0.0105
BWW9	quadratic	-0.5625	0.964878	-0.58	0.5620
CONTROL	linear	-1.36959	0.964878	-1.42	0.1607
CONTROL	quadratic	-0.0625	0.964878	-0.06	0.9486

Small P-Values suggest the need for a trend component of the order indicated.

For the sample data, both drugs AX23 and BWW9 have at least one P-value that is less than 0.05. This indicates that the trend over time evident in the *Profile Plot* is statistically significant for those drugs. The CONTROL, on the other hand, has no significant trend.

Pane Options

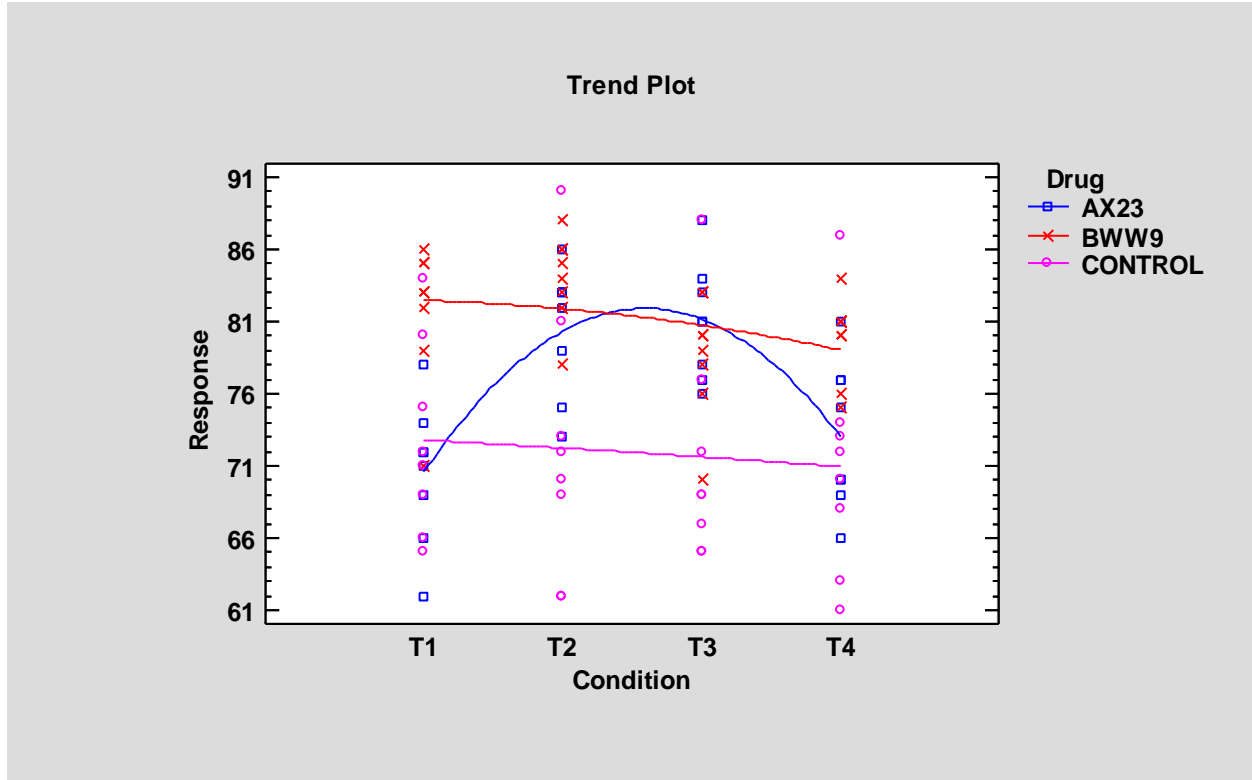


Between-Subjects Factor: the between-subjects factor to be included in the table.

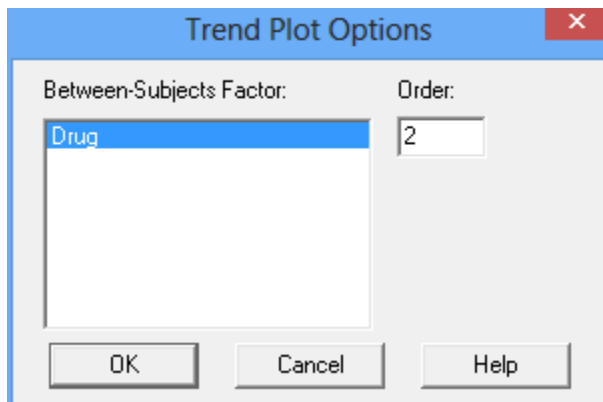
Maximum Order: maximum order component to be included in table (cannot exceed the number of levels of the repeated measures factor minus 1).

Trend Plot

This pane plots polynomials fitted to the means at each level of the repeated measures factor. The polynomials may be plotted for any between-subject factor. The plot below shows quadratic polynomials for each of the 2 drugs and the control:



Pane Options



Between-Subjects Factor: the between-subjects factor to be included in the table.

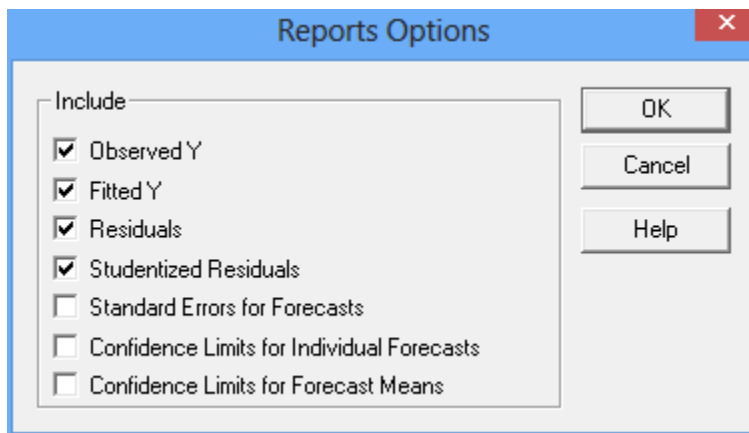
Order: order of the fitted polynomial (cannot exceed the number of levels of the repeated measures factor minus 1).

Reports

The *Reports* pane displays information about each of the observations in the data. A section of the report is shown below:

Model Results for Heart rate						
Row	Subject	Time	Observed Value	Fitted Value	Residual	Studentized Residual
1	1	T1	72.0	73.2188	-1.21875	-0.548199
2	1	T2	86.0	83.2187	2.78125	1.26397
3	1	T3	81.0	83.7188	-2.71875	-1.23486
4	1	T4	77.0	75.8438	1.15625	0.51996
5	2	T1	85.0	84.2188	0.78125	0.350909
6	2	T2	86.0	86.4688	-0.46875	-0.210412
7	2	T3	83.0	81.0938	1.90625	0.860461
8	2	T4	80.0	82.2188	-2.21875	-1.00365
9	3	T1	69.0	72.8438	-3.84375	-1.76769
10	3	T2	73.0	72.4688	0.53125	0.238491
11	3	T3	72.0	71.5938	0.40625	0.182341
12	3	T4	74.0	71.0938	2.90625	1.32235
13	4	T1	78.0	76.7188	1.28125	0.576459
14	4	T2	83.0	86.7187	-3.71875	-1.70745
15	4	T3	88.0	87.2188	0.78125	0.350909
16	4	T4	81.0	79.3438	1.65625	0.746522

Pane Options

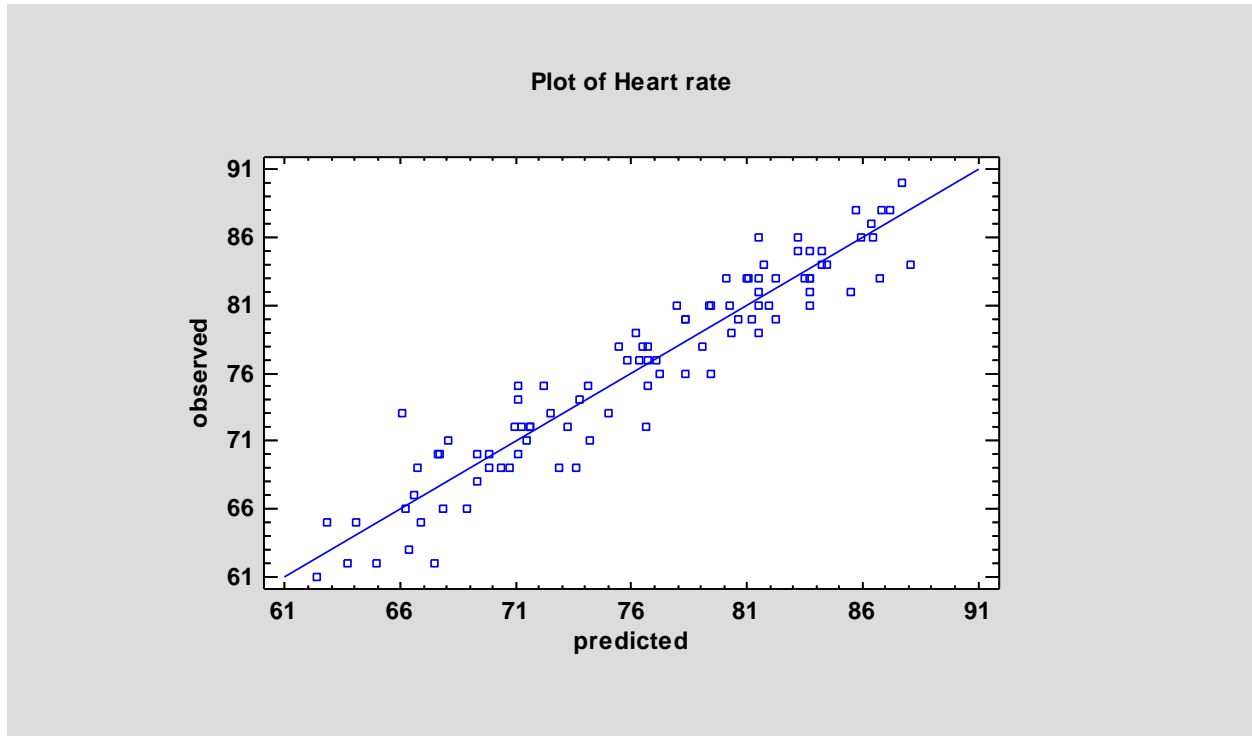


You may include:

- *Observed Y* – the observed values of the response variable.
- *Fitted Y* – the predicted values from the fitted model.
- *Residuals* – the ordinary residuals (observed minus predicted).
- *Studentized Residuals* – the Studentized deleted residuals.
- *Standard Errors for Forecasts* – the standard errors for new observations at each combination of the factors.
- *Confidence Limits for Individual Forecasts* – confidence intervals for new observations.
- *Confidence Limits for Forecast Means* – confidence intervals for the mean value of Y at each combination of the factors.

Observed versus Predicted

The *Observed versus Predicted* plot shows the observed values of Y on the vertical axis and the predicted values \hat{Y} on the horizontal axis.



If the model fits well, the points should be randomly scattered around the diagonal line. Any change in variability from low values of Y to high values of Y might indicate the need to transform the response variable before fitting a model to the data.

Residual Plots

As with all statistical models, it is good practice to examine the residuals. When fitting a statistical model, the residuals are defined by

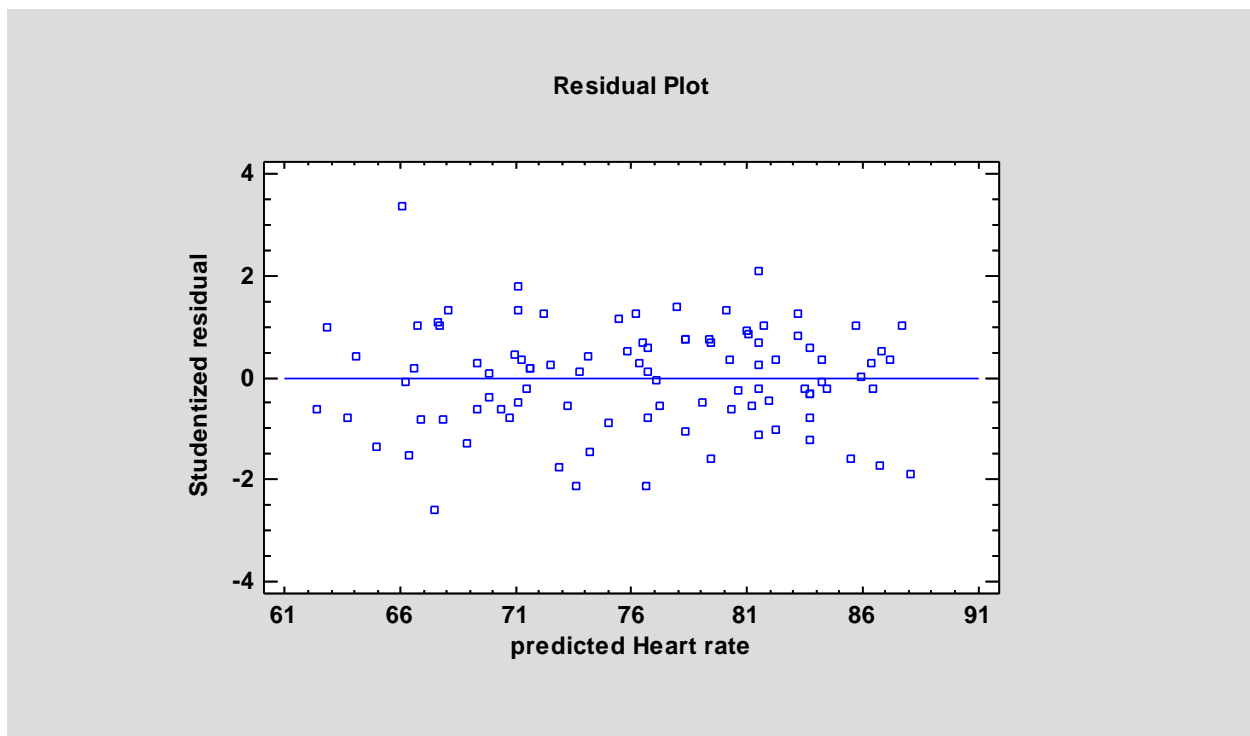
$$e_i = y_i - \hat{y}_i \quad (6)$$

i.e., the residuals are the differences between the observed data values and the fitted model.

This procedure creates various types of residual plots, depending on the settings in *Pane Options*.

Scatterplot versus Predicted Values

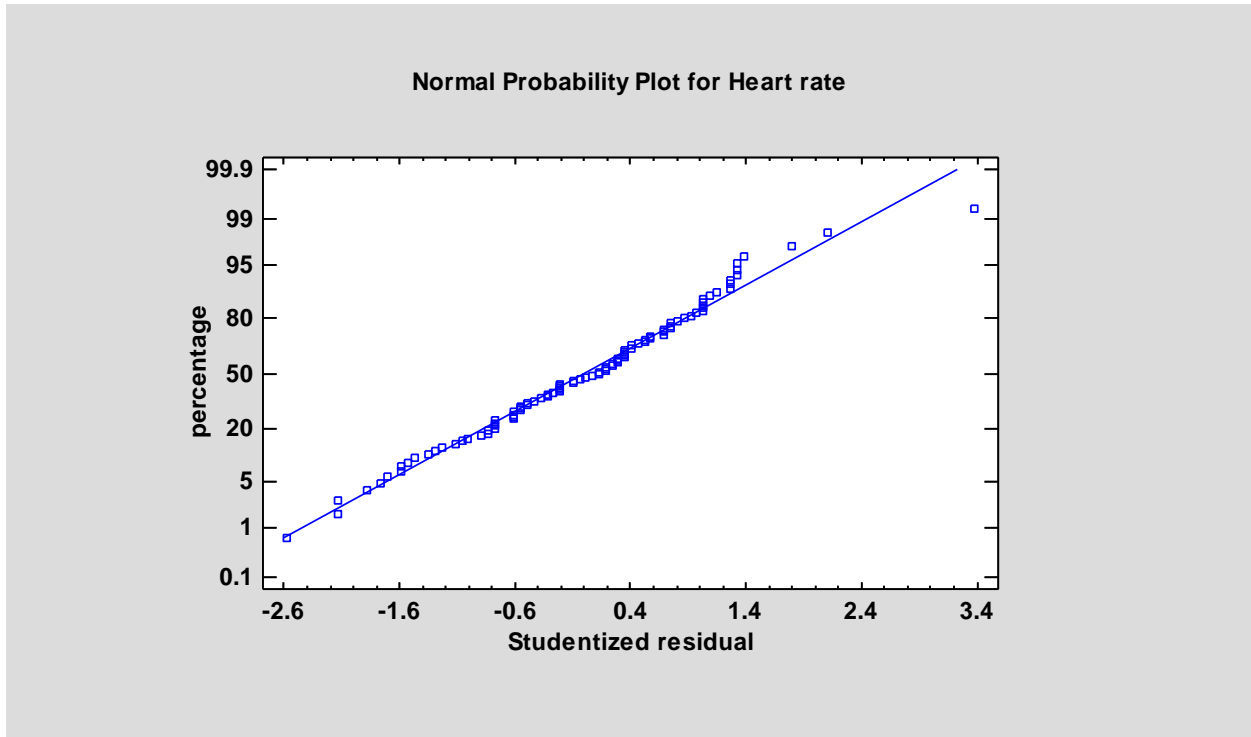
This plot is helpful in visualizing any possible dependence of the residual variance on the mean, which might necessitate a transformation of the response variable.



The above plot shows a fairly constant variance, except for one possible outlier.

Normal Probability Plot

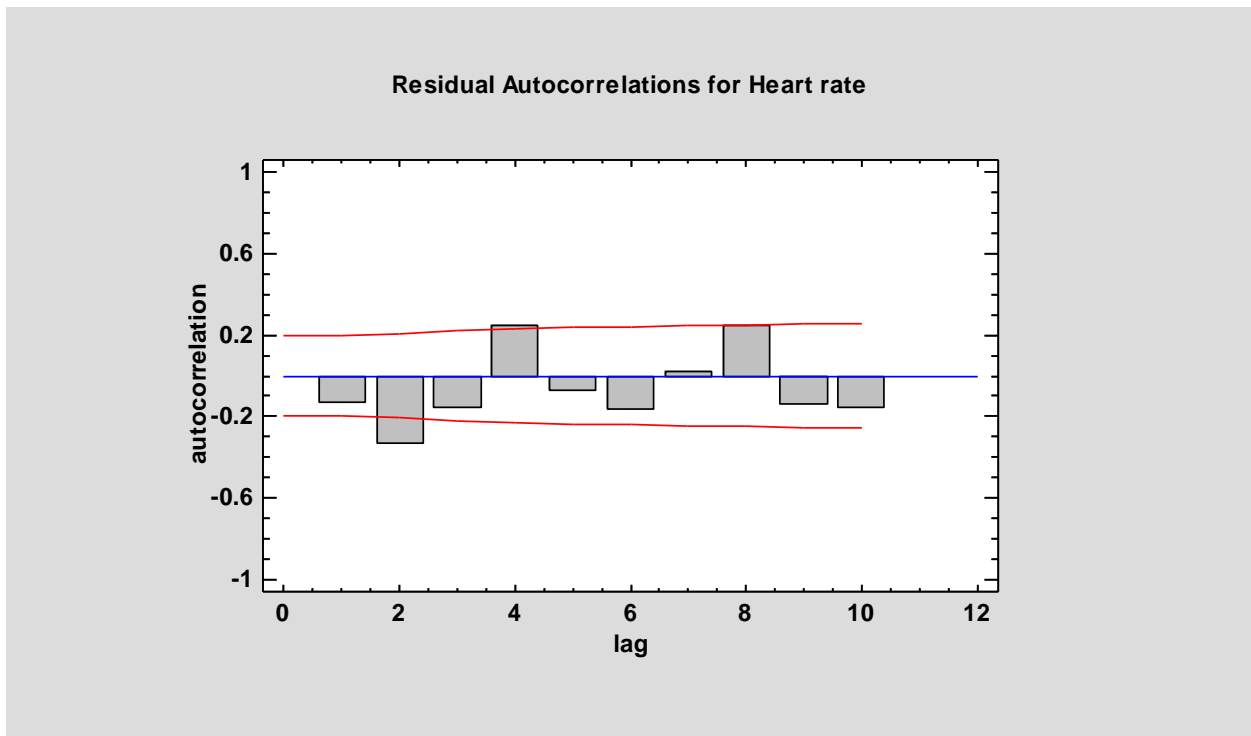
This plot can be used to determine whether or not the deviations around the line follow a normal distribution, which is the assumption used to form the prediction intervals.



If the deviations follow a normal distribution, they should fall approximately along a straight line.

Residual Autocorrelations

This plot calculates the autocorrelation between residuals as a function of the number of rows between them in the datasheet.



Any bars extending beyond the probability limits would indicate significant dependence between residuals separated by the indicated “lag”, which would violate the assumption of independence made when fitting the statistical model.

Pane Options

- **Plot:** the type of residuals to plot:
 1. *Residuals* – the residuals from the least squares fit.
 2. *Studentized residuals* – the difference between the observed values y_i and the predicted values \hat{y}_i when the model is fit using all observations except the i -th, divided by their estimated standard error. These residuals are sometimes called *externally deleted residuals*, since they measure how far each value is from the fitted model when that model is fit using all of the data except the point being considered. This is important, since a large outlier might otherwise affect the model so much that it would not appear to be unusually far away from the line.
- **Type:** the type of plot to be created. A *Scatterplot* is used to test for curvature. A *Normal Probability Plot* is used to determine whether the model residuals come from a normal distribution. An *Autocorrelation Function* is used to test for dependence between consecutive residuals.
- **Plot Versus:** for a *Scatterplot*, the quantity to plot on the horizontal axis.
- **Number of Lags:** for an *Autocorrelation Function*, the maximum number of lags. For small data sets, the number of lags plotted may be less than this value.
- **Confidence Level:** for an *Autocorrelation Function*, the level used to create the probability limits.

Unusual Residuals

Once the model has been fit, it is useful to study the residuals to determine whether any outliers exist that should be removed from the data. The *Unusual Residuals* pane lists all observations that have Studentized residuals of 2.0 or greater in absolute value.

Unusual Residuals for Heart rate						
			<i>Predicted</i>		<i>Studentized</i>	
<i>Row</i>	<i>Subject</i>	<i>Time</i>	<i>Y</i>	<i>Y</i>	<i>Residual</i>	<i>Residual</i>
22	6	T2	62.0	67.4688	-5.46875	-2.58
24	6	T4	73.0	66.0938	6.90625	3.37
40	10	T4	69.0	73.5938	-4.59375	-2.14
48	12	T4	72.0	76.5938	-4.59375	-2.14
53	14	T1	86.0	81.4688	4.53125	2.10

Studentized residuals greater than 3.0 in absolute value correspond to points more than 3 standard deviations from the fitted model, which is a rare event for a normal distribution. For the sample data, there is one Studentized residual greater than 3.0.