Section 2 – One Factor Tutorial

Introduction

In this tutorial you will build a one factor design using Design-Ease® software. This type of design can be very useful for simple comparisons such as who's the best supplier, or which type of raw material works best. The data for this example comes from the Stat-Ease bowling league. Three bowlers must compete for the last position on the team. They each bowl six games (see data below).

Game	Pat	Mark	Shari
1	160	165	166
2	150	180	158
3	140	170	145
4	167	185	161
5	157	195	151
6	148	175	156
Mean	153.7	178.3	156.2

Bowling Scores

The captain knows better than just to simply pick the bowler with the highest score. Maybe it's a fluke that Mark scored highest and Pat's score is low. He wants to know if the scores are significantly different, given the variability in individual scores.

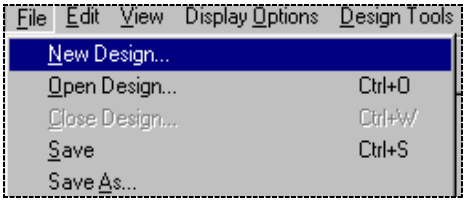
This one factor case study provides a good introduction to the power of simple comparative design of experiments (DOE). It will exercise a number of handy features provided by Design-Ease software. We won't explain all features displayed – some will be covered in follow-up tutorials. Many other features and outputs will be covered only in the help system, which can be accessed by clicking on Help on the main menu, or in most places via a right click or by pressing the F1 key (context sensitive).

Design the Experiment

We will assume that you are familiar with the graphical user interface on your computer and the use of a mouse. Start the program by finding and double clicking the Design-Ease icon. You will then see the main menu and tool bar.

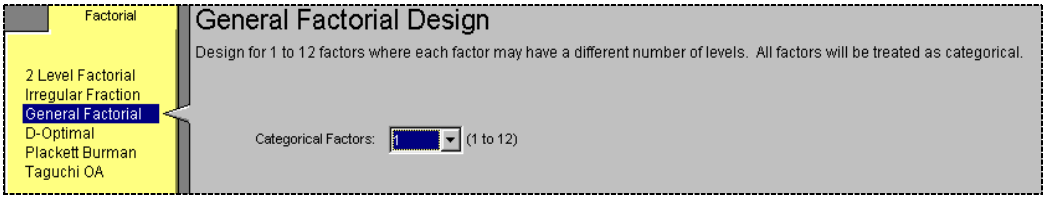
Main Menu and Tool Bar

Click on **File** in the menu bar. (Unavailable items are shown in a secondary color.) (If you prefer using the keyboard, press the Alt key and underlined letter, in this case Alt F, simultaneously.)



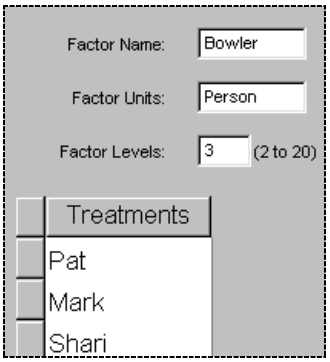
File Menu

Select the **New Design** item with your mouse. (The blank-sheet icon 🗋 on the left of the toolbar is a quicker route to this screen. If you'd like to check this out, press Cancel to re-activate the tool bar.) The **Factorial** tab comes up by default. Select **General Factorial** for this design because the factor is categorical. (If your factor is numerical, such as time or temperature, consider doing a one-factor response surface design, which can be found in the more advanced Design-Expert® software published by Stat-Ease, Inc.) Leave the number of factors at its default level of **1** and then click on **Continue**.



General Factorial Design

Type **Bowler** as the name of the factor. **Tab** to the **Units** field and enter **Person**. Then **Tab** to the **Levels** field and enter **3**. Click on the **Treatments** field and enter **Pat**, **Mark**, and **Shari**.

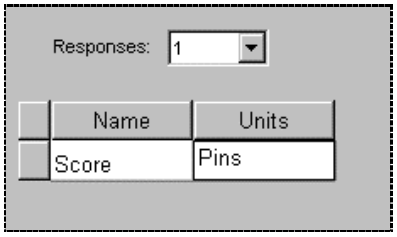


General Factorial Design Builder Dialog Box - Completed

Press **Continue** to specify the remaining design options. Click on the **Replicates** field and type **6** (each bowler bowls 6 games). **Tab** to the **Blocks** field but leave it blank. The number of experiments should now be 18.

Design Options Entered

Press **Continue**. Leave the number of **Responses** at the default of **1**. Then click on the **Name** box and enter **Score**. **Tab** to the **Units** field and enter **Pins**.

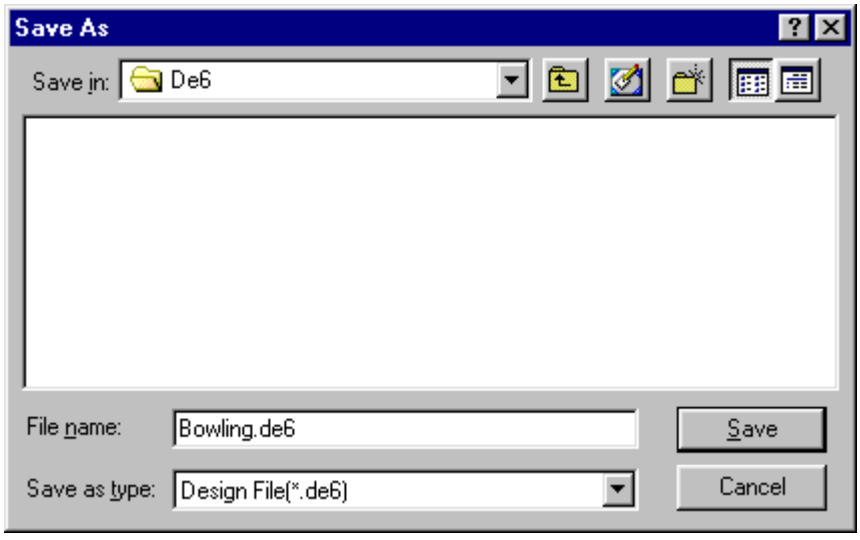


Response Name Dialog Box - Completed

Click on **Continue** to accept the response name. Design-Ease creates the design and takes you to the design layout window.

Save the Design

When you complete the design setup, save it to a file by selecting **File**, **Save As** . Type in the name of your choice (such as **Bowling.de6**) for your data file.

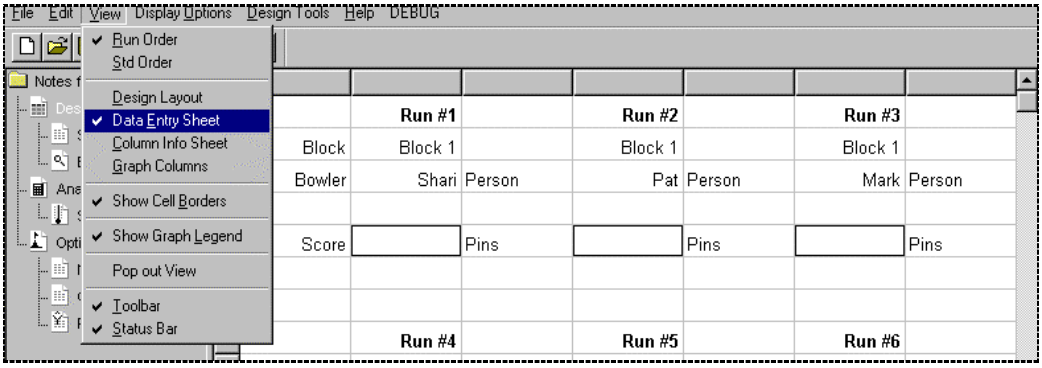


Save As Dialog Box

Then click on **Save**. Now you're protected in case of a system crash.

Create a Data Entry Form

Go to the **View** menu and select **Data Entry Sheet** from the menu to get a recipe sheet of your experiments. It provides space to record the responses.



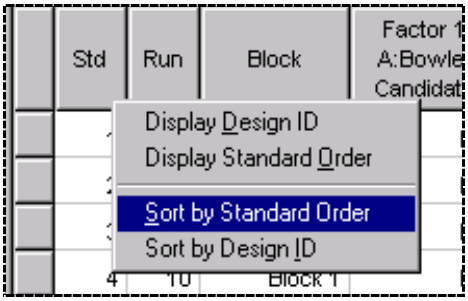
Data Entry Sheet View (Your Run Order May Differ)

It's not necessary for this tutorial, but if you have a printer connected, you can select File, Print and OK to make a hard copy. (You can do the same from the basic design layout if you like that format better.)

Enter the Response Data

When you do your own experiments, you will need to go out and collect the data. Simulate this by doing a **File**, **Exit**. Click on **Yes** if you are prompted to **Save**. Then re-start Design-Ease and use **File**, **Open Design** (or file open icon on the toolbar) to open your data file (**Bowling.de6**).

You should now see the data tabulated in the standard design layout. (If not, go to View and select Design Layout.) For an actual experiment, the runs would have been performed in randomized run order, which is the order Design-Ease defaults to. This run order will be different each time a design is created from scratch. For this example, you must enter the data in the proper order to match up with the correct bowlers, so right click at the top of the **Std** column and choose **Sort by Standard Order**.



Sort Runs by Standard Order

Then enter the responses from the table on page one, or use the following screen as a template.

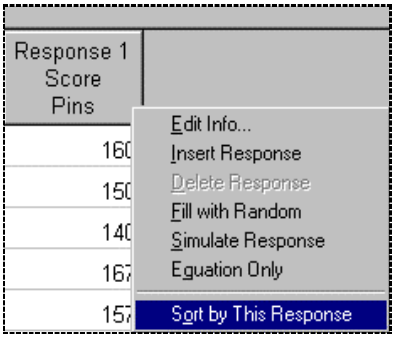
Std	Run	Block	Factor 1 A:Bowler Person	Response 1 Score Pins
1	9	Block 1	Pat	160
2	7	Block 1	Pat	150
3	2	Block 1	Pat	140
4	16	Block 1	Pat	167
5	8	Block 1	Pat	157
6	5	Block 1	Pat	148
7	6	Block 1	Mark	165
8	15	Block 1	Mark	180
9	4	Block 1	Mark	170
10	11	Block 1	Mark	185
11	14	Block 1	Mark	195
12	3	Block 1	Mark	175
13	10	Block 1	Shari	166
14	1	Block 1	Shari	158
15	17	Block 1	Shari	145
16	12	Block 1	Shari	161
17	18	Block 1	Shari	151
18	13	Block 1	Shari	156

Design Layout in Standard Order with Response Data Entered

Your design layout window should now look like that shown above, except for run order. When you do your own experiments, be sure to do the runs and enter the response(s) in randomized order. Standard order should only be used as a convenience for entry of pre-existing designs.

Save your data by selecting **File**, **Save** from the menu (or the save icon 🖫 on the toolbar).

Before starting the statistical analysis it might be enlightening to simply sort the results. You can do this by right-clicking on the **Response** column and selecting **Sort by This Response**.



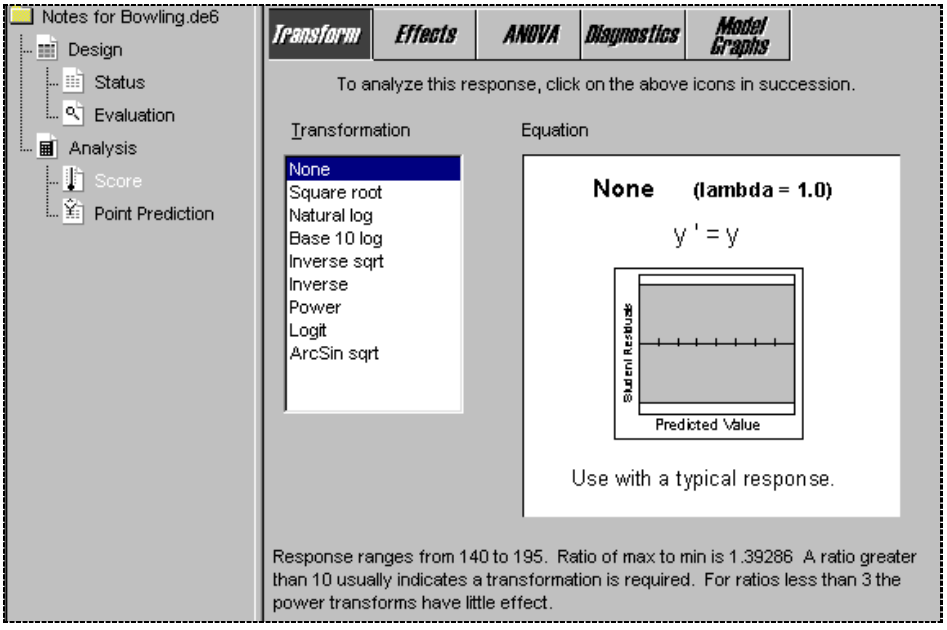
Sorting on a Response Column (also works for factors)

This is a very useful feature. It works on factors as well as responses. In this case you can see that the highest games were bowled by Mark.

Another right-click feature that you can explore if you like, is the **Edit Info**. Among other things, this allows you to change names and levels of factors, or specify decimal places for any of the numeric entries. Click on **OK** when you are ready to move on.

Analyze the Results

Next, begin the data analysis. Click on the **Score** node on the left side of your screen under analysis. The Transform dialog box will be displayed in the main window of Design-Ease on a progressive tool bar. You will click these buttons from left to right and do the complete analysis. It's a very easy process. The first dialog box gives you the option of selecting a transformation for the response, which may improve the statistical properties of the analysis.

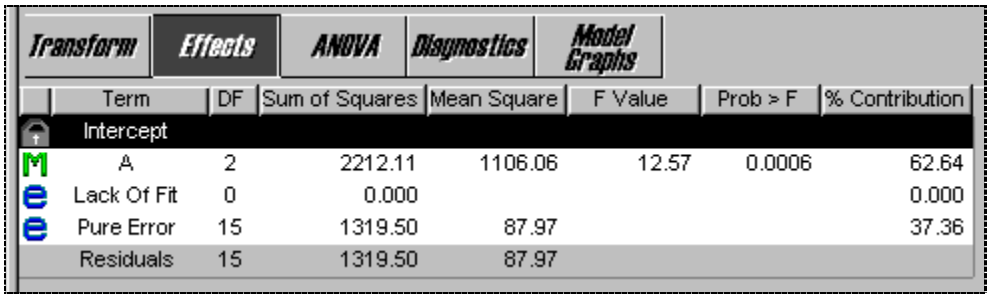


Transformation Button – The Starting Point for the Statistical Analysis

The program provides some data-sensitive advice at the bottom of the screen. In this case, it does not indicate the need for a transformation, so press ahead with the default of **None** by clicking on the ***Effects*** button. (If you want more background on transformations, go to the Help command on the main menu and select Contents. Then Search on "transformations.")

Examine the Analysis

Things get a bit complicated from here on out. If you're not a statistician, you may not understand all the numerical outputs. We advise you attend a basic class on regression, or better yet, a workshop on DOE such as Stat-Ease's Experiment Design Made Easy.

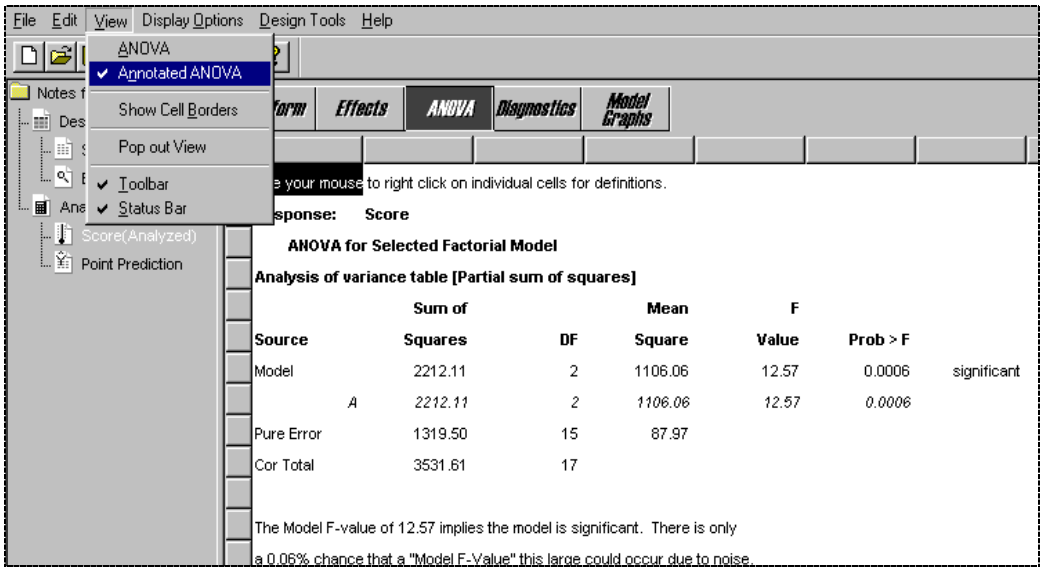


Term	DF	Sum of Squares	Mean Square	F Value	Prob > F	% Contribution
Intercept						
A	2	2212.11	1106.06	12.57	0.0006	62.64
Lack Of Fit	0	0.000				0.000
Pure Error	15	1319.50	87.97			37.36
Residuals	15	1319.50	87.97			

Effects Button Results

The really important outputs on the effects are the F-value and associated probability ("Prob>F"). In this case, there's a very small probability, near 0.06%, that the differences in bowling averages are due to chance variation. In other words, it appears at this stage that the difference between bowlers is significant.

To get more details press the ***ANOVA*** (Analysis of Variance) button. Then select **View**, **Annotated ANOVA**. This tells the same story: there is a significant difference.



Annotated ANOVA Results

Next you see a section of the output that reports various summary statistics.

Std. Dev.	9.38	R-Squared	0.6264
Mean	162.72	Adj R-Squared	0.5766
C.V.	5.76	Pred R-Squared	0.4620
PRESS	1900.08	Adeq Precision	6.442

The "Pred R-Squared" of 0.4620 is in reasonable agreement with the "Adj R-Squared" of 0.5766.

"Adeq Precision" measures the signal to noise ratio. A ratio greater than 4 is desirable. Your ratio of 6.442 indicates an adequate signal. This model can be used to navigate the design space.

Summary Statistics

The annotations tell you what you need to know, but don't be shy about clicking on a number and getting on-line Help via a right-click or the F1 key. In most cases you will then get helpful advice on the particular statistic.

Now click on the scroll down arrow (at the bottom right side of screen) until you get to the section labeled "Treatment Means."

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-Pat	153.67	3.83
2-Mark	178.33	3.83
3-Shari	156.17	3.83

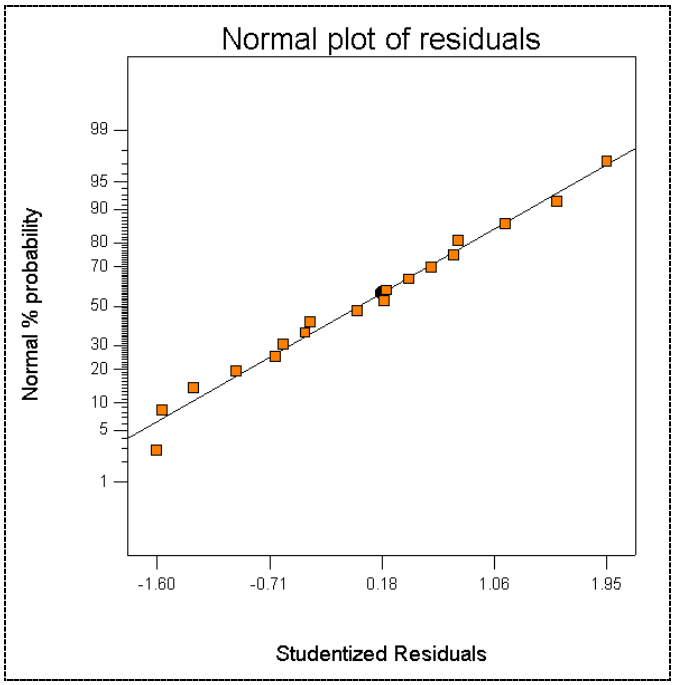
Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-24.67	1	5.41	-4.56	0.0004
1 vs 3	-2.50	1	5.41	-0.46	0.6509
2 vs 3	22.17	1	5.41	4.09	0.0010

Treatment Means

You now can evaluate how the means compare. You can see from the treatment comparisons: Pat is different from Mark (1 vs 2). Pat is not different from Shari (1 vs 3). Mark is different from Shari (2 vs 3).

Analyze Residuals

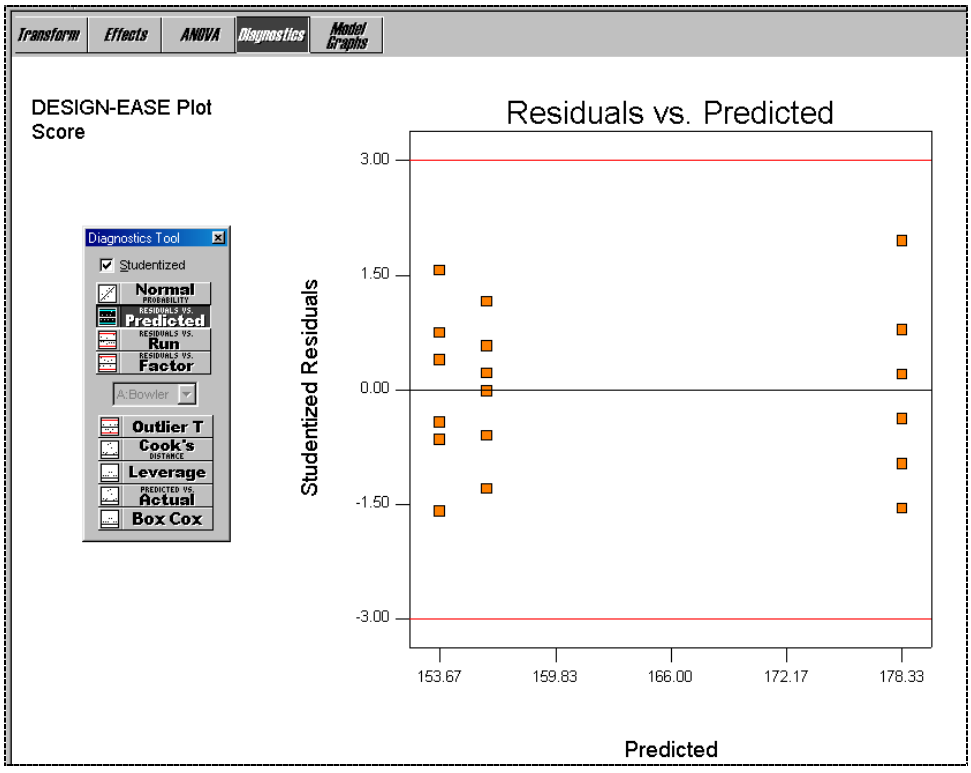
Click on the ***Diagnostics*** button. Ideally this will be a straight line, indicating no abnormalities. In this case the plot looks OK.



Normal Probability of Studentized Residuals

You may reposition the line by dragging the line (place the mouse pointer on the line, hold down the left button, and move the mouse) or its "pivot point."

Select **Residuals vs. Predicted** from the list shown on the **Diagnostics Tool**.

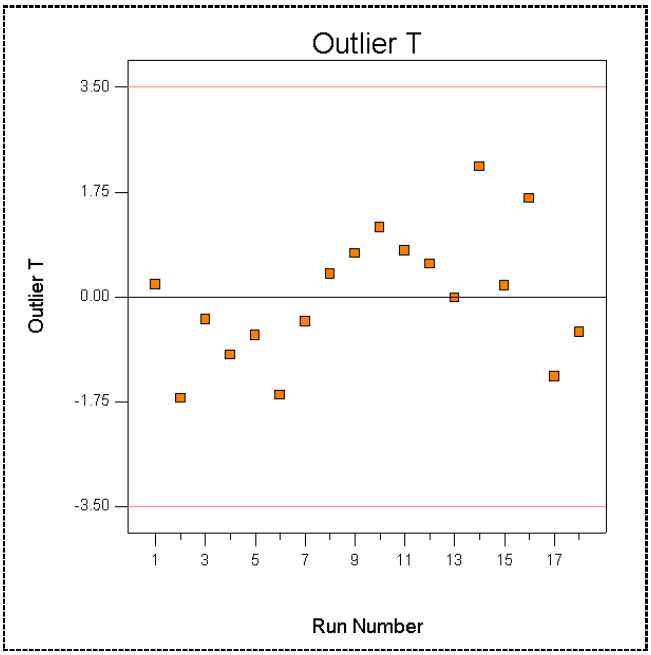


Studentized Residuals versus Predicted Values

Notice that the label on the Y-axis says “Studentized” residuals. This is a rescaling that reflects the number of standard deviations between the actual and predicted response values. More details on this may be found by searching Help. The actual residuals can be displayed by unchecking the “Studentized” default mode on the Diagnostics Tool. However, when some runs have greater leverage (another statistical term to look up in Help!), only the Studentized form of residuals will produce valid diagnostic graphs. Therefore we advise that you not turn this feature off.

The size of the studentized residual should be independent of its predicted value. In other words the spread of the studentized residuals should be approximately the same for each bowler. In this case the plot looks OK.

Select **Outlier T** from the **Diagnostics Tool** to see if any points stand out.



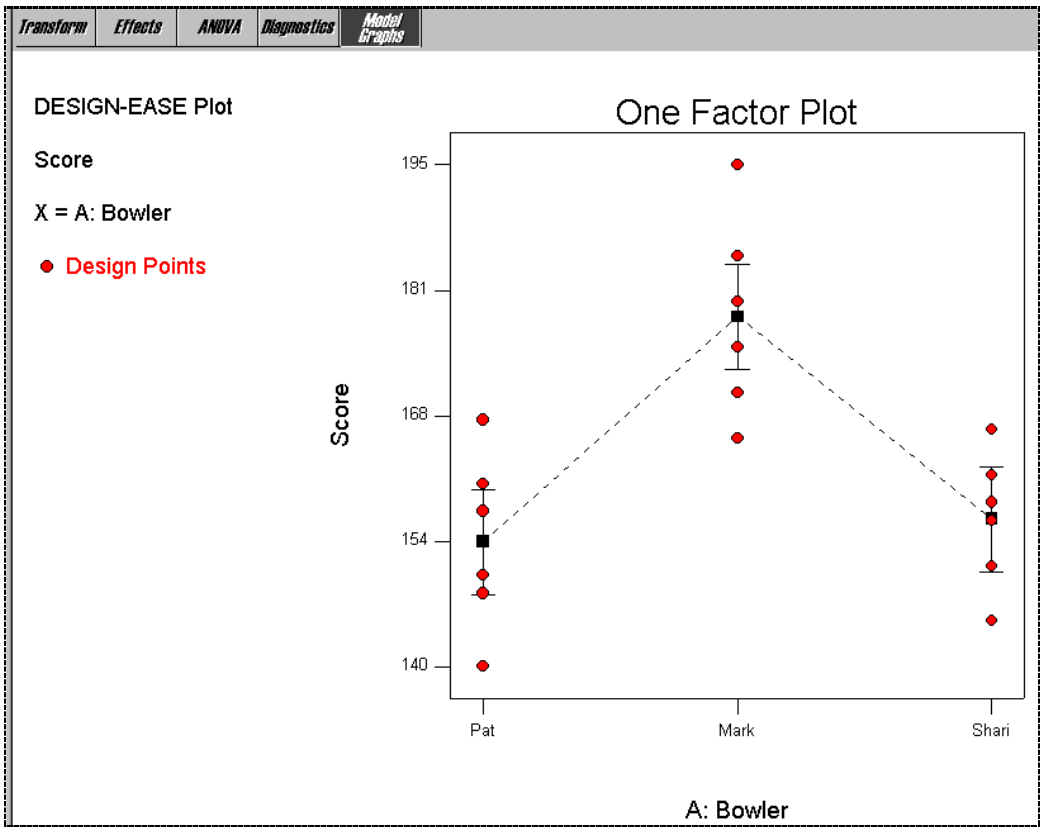
Outlier T Plot (pattern may differ on your display due to random run order)

On the outlier t plot we are looking for points outside the plus and minus 3.5 standard deviation limits, not patterns. (Due to the randomization of run order, your pattern may differ from that shown.) All points in this case fall within the limits.

You can check other residual plots as well, but the three reproduced above are the most important. Since there’s no indication of abnormality, it’s OK to move on to the model graph. This will tell the story about the effect of changing bowlers.

View the Means and Data Plot

Select the ***Model Graphs*** button from the progressive tool bar to display a plot containing all of the response data and the average value at each level of the treatment (factor). This plot gives an excellent overview of the data and the effect of the factor levels on the mean and spread of the response.



One Factor Graph

The squares represent the mean responses for each factor level (bowler). The vertical bars represent the 95% least significant difference (LSD) intervals for each treatment. Mark's LSD bars do not overlap with Pat's or Shari's, so we can say with at least 95% confidence that Mark's mean is significantly higher than the means of the other two bowlers. You can get a numerical value for the length of the LSD bar by clicking on one of the squares that represent the predicted score. It appears to the left of the graph. Give this a try. (You can also click on any round points to get the actual score. Check it out!) Pat and Shari's LSD bars do overlap, so we cannot say which of them is the better bowler.

Save the results by going to **File**, **Save**. You can now **Exit** Design-Ease if you like, or keep it open and go on to your next design of experiments project.

Modifying the Design Layout (Advanced Topic)

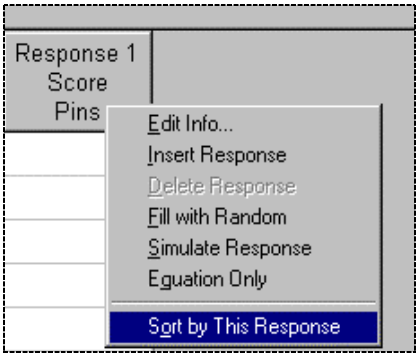
Design-Ease software offers a great deal of flexibility to modify data in its design layout. Let's see what can be done in the bowling case. (Warning: we now depart from the simple and straightforward design laid out earlier. If your brain is already full with new things, skip this section or come back to it later.)

The outcome of the bowling match appears to be definitive, especially from Mark's point of view. However, Pat and Shari demand one more chance to prove themselves worthy of the team. Mark objects and an argument ensues. To achieve compromise, the captain decides to toss out the highest and lowest games for each of the three bowlers, and replace them with two new scores each. The results can be seen below.

Block	Game	Pat	Mark	Shari
1	1	160	~~165~~	~~166~~
1	2	150	180	158
1	3	~~140~~	170	~~145~~
1	4	~~167~~	185	161
1	5	157	~~195~~	151
1	6	148	175	156
2	1	162	175	163
2	2	153	180	166

Bowling Scores with High and Low Games Replaced by Two New Games

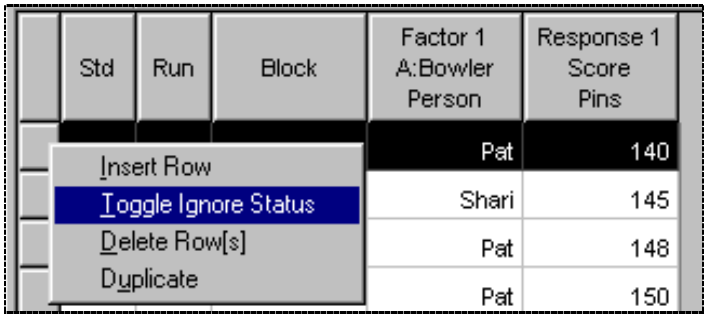
To enter this data, you must modify the original design layout. If you exited the program, re-start Design-Ease and use **File, Open Design** to open your data file (**Bowling.de6**). Otherwise, click on the **Design** node at the upper left of your screen. You now should see the bowling data from before. Right click at the top of the **Response** column and choose **Sort by This Response**.



Sorting Runs by Response

The next task will be to exclude the low and high games for each of the three bowlers. Right-click on the square button to the left of the first row (Pat's low game) and select

Toggle Ignore Status as shown below. (The data can be restored by repeating the Toggle Ignore Status procedure. Give it a try!)



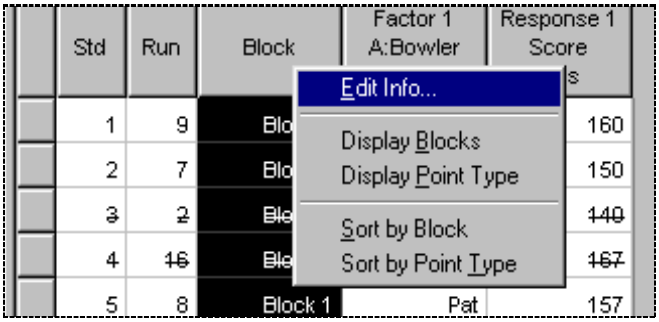
Ignoring Pat's Low Game

Select **Toggle Ignore Status** for Shari's low game, which should be next in line. Go down a number of rows and do the same for Mark's low game. Then for each person's high games, which appear towards the bottom of the design layout, also select **Toggle Ignore Status**. When you finish, your screen should look like that below.

Std	Run	Block	Factor 1 A:Bowler Person	Response 1 Score Pins
3	~~2~~	~~Block 1~~	~~Pat~~	~~140~~
~~15~~	~~17~~	~~Block 1~~	~~Shari~~	~~145~~
6	5	Block 1	Pat	148
2	7	Block 1	Pat	150
17	18	Block 1	Shari	151
18	13	Block 1	Shari	156
5	8	Block 1	Pat	157
14	1	Block 1	Shari	158
1	9	Block 1	Pat	160
16	12	Block 1	Shari	161
~~7~~	~~6~~	~~Block 1~~	~~Mark~~	~~165~~
~~13~~	~~10~~	~~Block 1~~	~~Shari~~	~~166~~
~~4~~	~~16~~	~~Block 1~~	~~Pat~~	~~167~~
9	4	Block 1	Mark	170
12	3	Block 1	Mark	175
8	15	Block 1	Mark	180
10	11	Block 1	Mark	185
~~11~~	~~14~~	~~Block 1~~	~~Mark~~	~~195~~

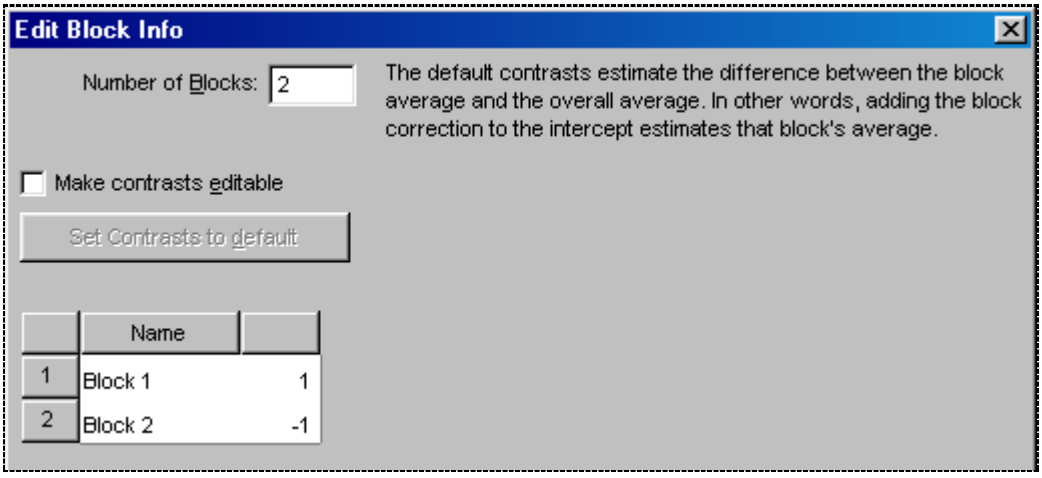
Design Layout After Ignoring Low and High Games for Each of the Bowlers

Return the layout to the original order by right-clicking on the **Std** column heading and choosing **Sort by Standard Order**. Now, create a new block by right-clicking at the top of the **Block** column and choosing **Edit Info**.



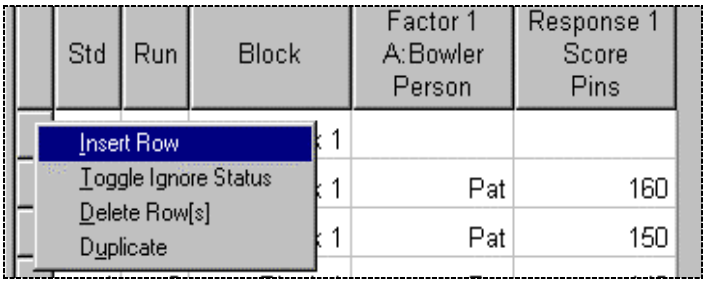
Creating a New Block

You now see a form, which allows users to give whatever name they want to the block(s). Don't bother doing this now. As shown below, change the **Number of Blocks** to 2. Press the **Tab** key to see this change take effect. Then press **OK**.



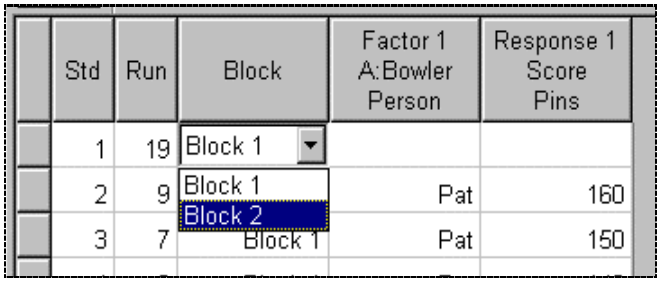
Add Block Form

Now you are ready to begin adding and/or duplicating rows. This can be accomplished in different ways, depending on your ingenuity. We will follow a procedure that exercises as many of the editing features as possible, so it may not be the most elegant approach. Right click on the square button at the left of the first row to bring up the editing menu. Move the mouse over the first selection **Insert Row** and click.



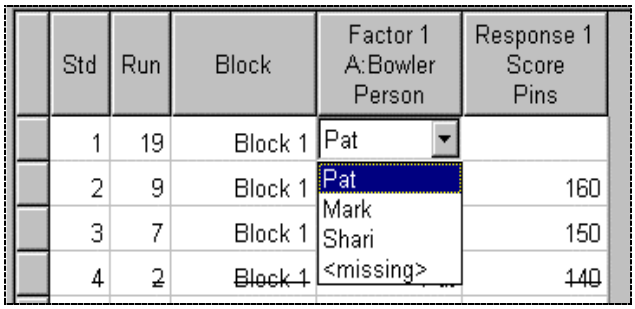
Inserting a Row

You now will see a new row with blanks for the bowler and the score. Click on the block field and then on the list arrow. Select **Block 2** as shown below.



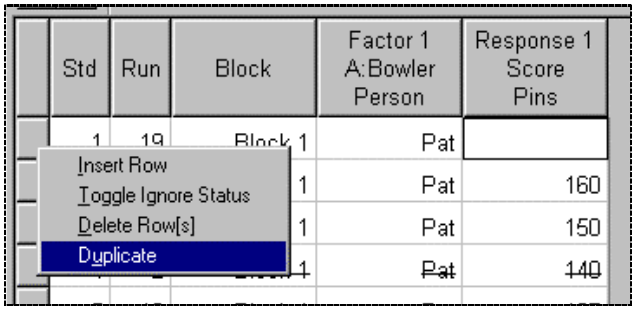
Changing Block Number

Next, click on the blank field for bowler and then on the list arrow. Select **Pat** (the default value). (If this were a numerical field, you would simply enter the value.)



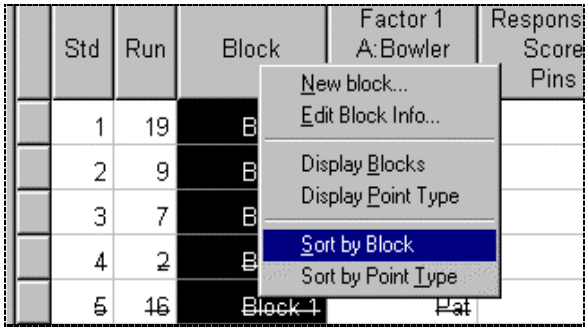
Entering a Categorical Value for Factor

Again right-click on the open square at the left of the first row to bring up the editing menu. Move the mouse over the selection to **Duplicate** and click.



Duplicating a Row

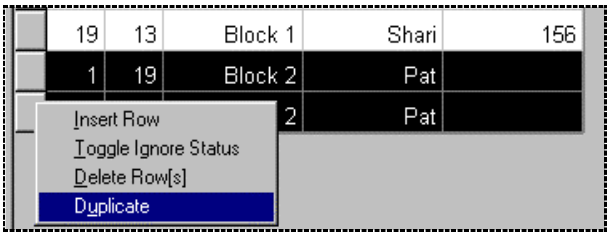
Right-click on the **Block** column heading and **Sort by Block**.



Sorting by Block

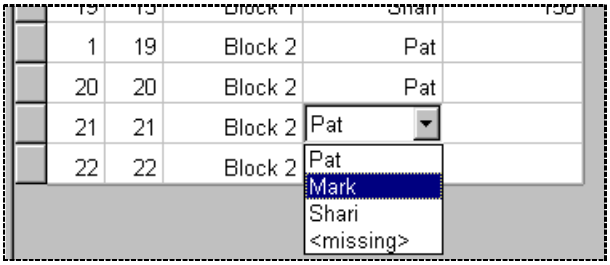
Now, the two new rows can now be seen at the bottom of the design layout. We need two new rows apiece for Shari and Mark. Let's just duplicate the new rows for Pat and change the bowler's name. Do this by first left-clicking the button (open square) to the left of the first new row for Pat, so it is highlighted. Then while holding down the **Shift** key, click the button to the left of the second new row for Pat. Now both rows should be highlighted. (This is a bit tricky, but it will save time.)

Now you can right-click on any button (open square) on the left-most part of the highlighted block and select **Duplicate**.



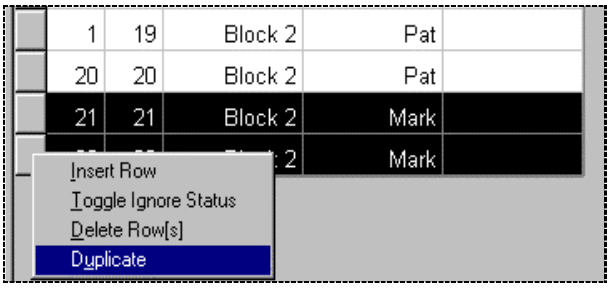
Duplicating a Block of Rows

In the first duplicated row, click on the field for **Bowler** and select **Mark**.



Changing Name of Bowler

Do the same for the last row. You now should have two new rows for both Pat and Mark. Click the button to the left of the first new row for Mark, so it is highlighted. Then while holding down the **Shift** key, click the button to the left of the second new row for Mark. Both rows should now be highlighted. Right-click on any button (open square) in the left-most part of the highlighted block and select **Duplicate**.



Duplicating Two More Rows

In the first duplicated row, click on the field for **Bowler** and select **Shari**. Do the same for the last row. You are now ready to enter the new data as shown below.

	19	13	Block 1	Shari	156
	1	19	Block 2	Pat	162
	20	20	Block 2	Pat	153
	21	21	Block 2	Mark	175
	22	22	Block 2	Mark	180
	23	23	Block 2	Shari	163
	24	24	Block 2	Shari	166

Data Entered for Second Block of Games

Go ahead now and re-analyze the data. It turns out that the added games cause no change in the overall conclusions as to who's the better bowler. Mark remains on top. It would now be appropriate to recover the low and high games for each bowler from block 1. Since this data was not deleted, only ignored, getting it back is simply a matter of right-clicking to the left of each of the six rows and re-selecting Toggle Ignore Status. Give this a try! Then re-analyze one last time.

By working through this exercise you now see how easy it is to manipulate data in Design-Ease's design layout. Keep this in mind should the need arise for editing your own data.