General One-Factor Tutorial

Part 1 – The Basics

Introduction

In this tutorial you will build a general one-factor design using Design-Expert® software. This type of design is very useful for simple comparisons of categorical treatments, such as:

- Who will be the best supplier,
- Which type of raw material should be selected,
- What happens when you change procedures for processing paperwork.

If you are in a hurry, skip the boxed bits—these are sidebars for those who want to spend more time and explore things.

☒ ***Explore response surface methods***: If you wish to experiment on a continuous factor, such as time, which can be adjusted to any numerical level, consider using response surface methods (RSM) instead. This is covered in a series of tutorials presented later in the Design-Expert User's Guide.

The data for this example come from the Stat-Ease bowling league. Three bowlers (Pat, Mark, and Shari) are competing for the last team position. They each bowl six games in random order – ideal for proper experimentation protocol. Results are:

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

Being a good experimenter, the team captain knows better than to simply pick the bowler with the highest mean score. The captain needs to know if the average scores are significantly different, given the variability in individual games. Maybe it's a fluke that Mark's score is highest.

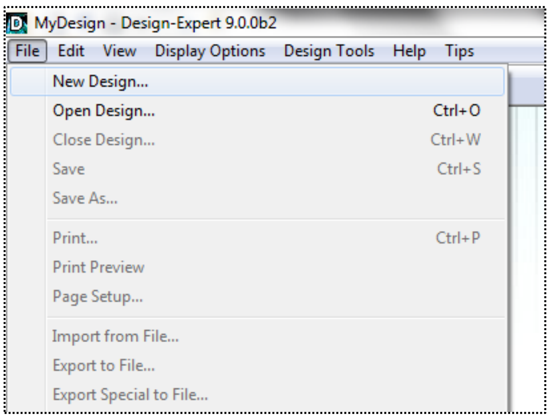
This one-factor case study provides a good introduction to the power of simple comparative design of experiments (DOE). It exercises many handy features found in Design-Expert software.

☒ ***Explore other resources***: We won't explain all features displayed in this current exercise because most will be covered in later tutorials. Many other features and outputs are detailed only in the help system, which you can access by clicking Help in 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 your computer's graphical user interface and your mouse. Start the program by double clicking the Design-Expert icon. You will then see the main menu and icon bar.

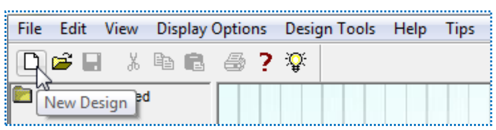
Click on **File** in the main menu. Unavailable items are dimmed. (If you prefer using your keyboard, press the Alt key and underlined letter simultaneously, in this case Alt F.)



File menu

Select the **New Design** item with your mouse.

☒ ***Explore optional ways to select a new design***: The blank-sheet icon 🗋 on the left of the toolbar is a quicker path to this screen. To try this, press Cancel to re-activate the tool bar.

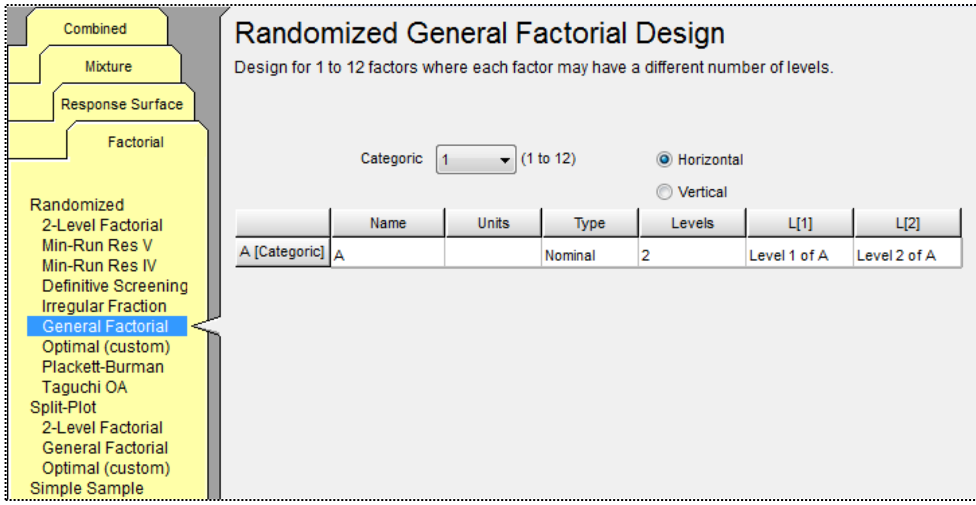


Opening a new design with the blank sheet icon

Using either path, you now see four yellow tabs on the left of your screen. 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 temperature, then you would use the One Factor option under the Response Surface tab.)

☒ ***Explore what the program tells you in its annotations***: Note the helpful description: "Design for 1 to 12 factors where each factor may have a different number of levels."

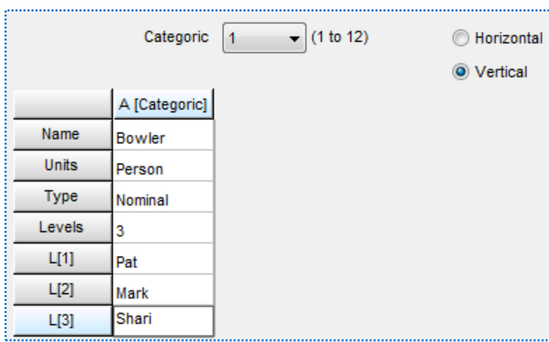
P.S. If any of your factors are quite hard to control, that is, not easily run at random levels, then consider using the Split-Plot General Factorial design. However, restricting randomization creates big repercussion on the power of your experiment, so do your best to allow all factors to vary run-by-run as chance dictates. (Design-Expert by default will lay out your design in a randomized run order.)



General Factorial design

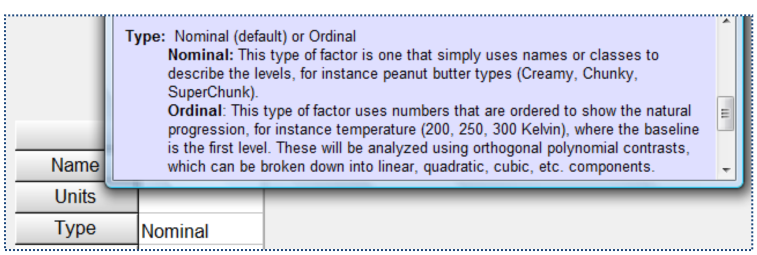
Enter the Design Parameters

Leave the number of factors at its default level of **1** but click the entry format **Vertical** (easier than Horizontal for multiple levels). Enter **Bowler** as the name of the factor. Tab down to the **Units** field and enter **Person**. Next tab to **Type**. Leaving Type at its default of **Nominal**, tab down to the **Levels** field and enter **3**. Now tab to **L(1)** (level one) and enter **Pat**. Type **Mark**, and **Shari** for the other two levels (L2 and L3).



General Factorial design-builder dialog box – completed

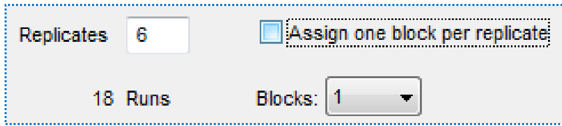
☒ ***Explore screen tips***: For details on the options for factor type, click the light bulb icon ( ) in the toolbar to access our context-sensitive screen tips.



Screen tips on factor Type

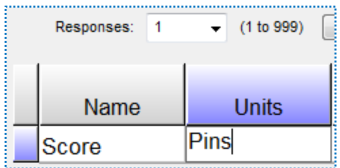
Press **Continue** to specify the remaining design options. In the **Replicates** field, which becomes active by default, type **6** (each bowler rolls six games). **Tab** to the

"Assign one block per replicate" field but leave it unchecked. Design-Expert now recalculates the number of runs for this experiment: 18.



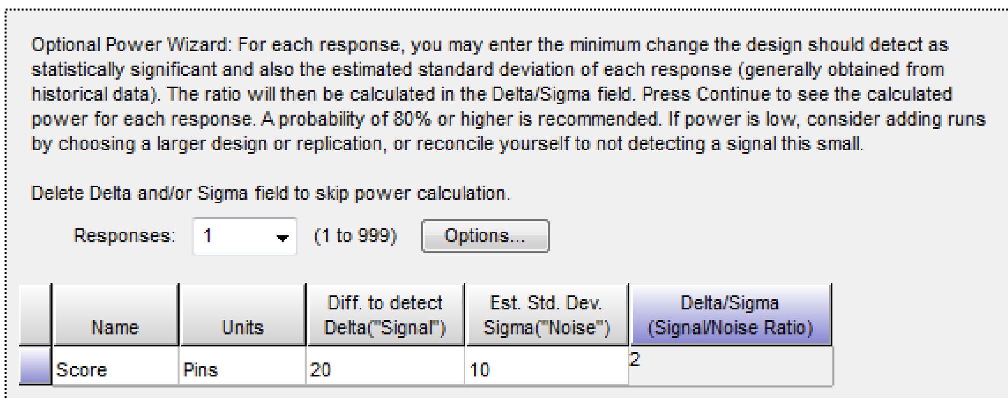
Design options entered

Press **Continue**. Let's do the easy things first. Leave the number of **Responses** at the default of **1**. Now click on the **Name** box and enter **Score**. **Tab** to the **Units** field and enter **Pins**.



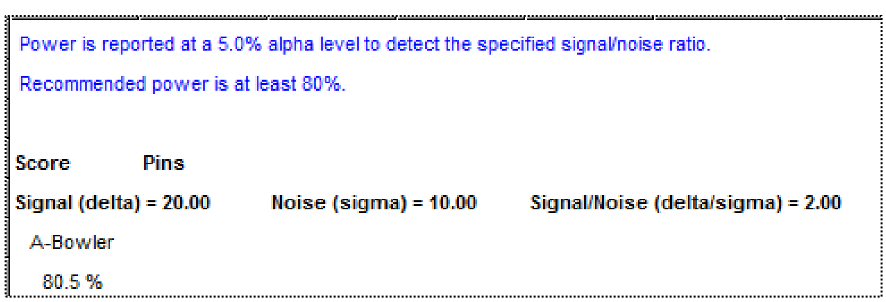
Response name dialog box – completed

At this stage you can skip the remainder of the fields and continue on. However, it is good to gain an assessment of the power of your planned experiment. In this case, as shown in the fields below, enter the value **20** because the bowling captain does not care if averages differ by fewer than **20** pins. Then enter the value **10** for standard deviation (derived from league records as the variability of a typical bowler). Design-Expert then compute a signal-to-noise ratio of **2** (10 divided by 5).



Optional power calculator – necessary inputs entered

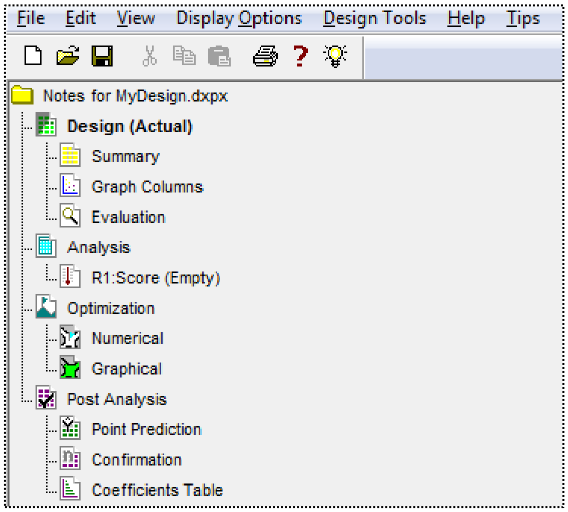
Press **Continue** to view the happy outcome – power that exceeds 80 percent probability of seeing the desired difference.



Results of power calculation

Click on **Continue** for Design-Expert to create the design and take you to the design layout window.

☒ ***Explore the program interface***: Before moving on, take a look at the unique branching interface provided by Design-Expert for the design and analysis of experiments and resulting optimization.

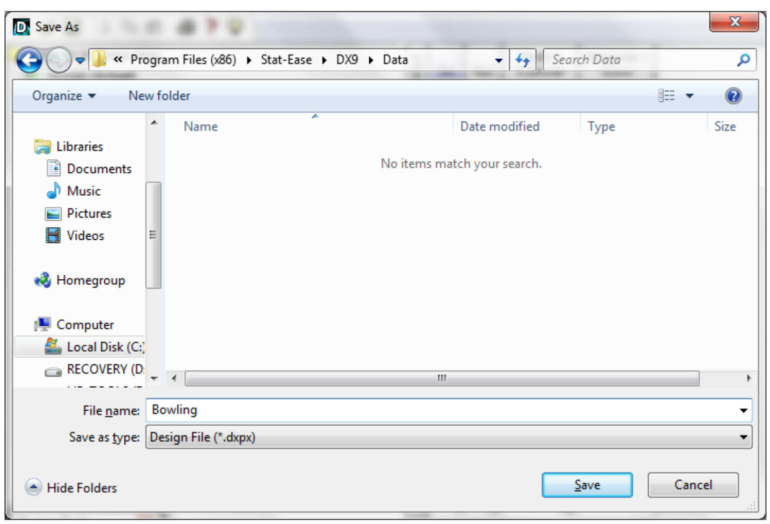


Design-Expert software's easy-to-use branching interface

You will explore some branches in this series of tutorials and others if you progress to more advanced features, such as response surface methods for process optimization.

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 (for this tutorial, we suggest **Bowling**) for your data file, which is saved as a *.dxpx type.

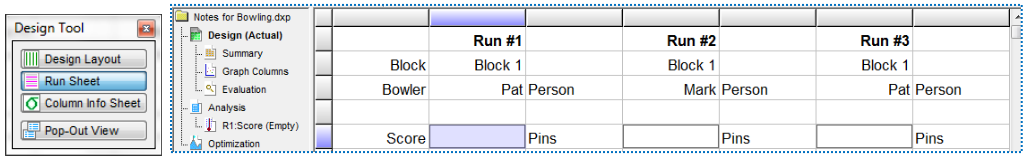


Save As dialog box

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

Create a Data Entry Form

In the floating **Design Tool** click **Run Sheet** (or go to the View menu and select Run Sheet) to produce a recipe sheet for your experiment with your runs in randomized order. A printout provides space to write down the responses. (Note: this view of the data does not allow response entry. To type results into the program you must switch back to the home base – the Design Layout view.)

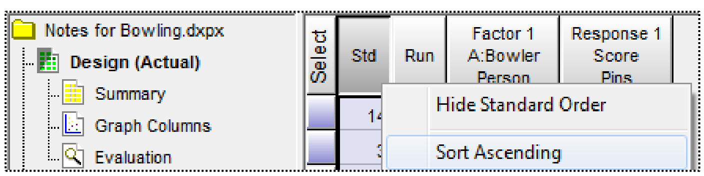


Run Sheet view (your run order may differ)

☒ ***Explore printing features***: It's not necessary for this tutorial, but if you have a printer connected, you can select **File**, **Print**, and **OK** (or click the printer icon) 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 performing your own experiments, you will need to go out and collect the data. Simulate this by clicking **File**, **Exit**. Click on Yes if you are prompted to Save. Now re-start Design-Expert and use **File**, **Open Design** or click the open file icon 🗁 on the toolbar)) to open your data file (**Bowling.dxpx**). You should now see your data tabulated in the randomized layout. For this example, you must enter your data in the proper order to match the correct bowlers. To do this, right-click the **Std** column header and choose **Sort Ascending**. This is the standard order you will see in DOE textbooks.



Sort runs by standard (std) order

Now enter the responses from the table on page one, or use the following screen. Except for run order, your design layout window must look like that shown below.

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

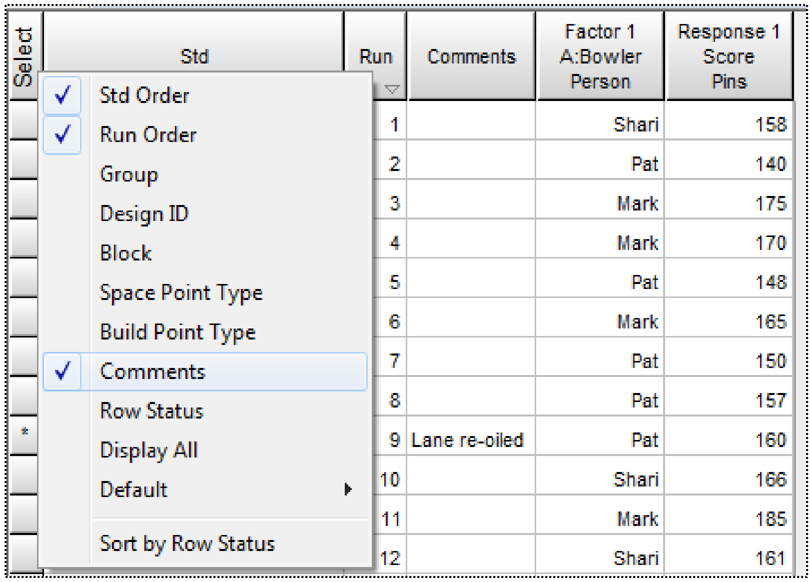
Design Layout in standard order with response data entered

When you conduct your own experiment, be sure to do the runs and enter the response(s) in randomized order. Standard order should only be used as a convenience for entering pre-existing design data.

> ☒ ***Explore advantages of being accurate on the actual run order***: If you are a real stickler, replace (type over) your run numbers with the ones shown above, thus preserving the actual bowlers' game sequence. Bowling six games is taxing but manageable for any serious bowler. However, short and random breaks while bowling six games protects against time-related effects such as learning curve (getting better as you go) and/or fatigue (tiring over time).

Save your data by selecting **File**, **Save** from the menu (or via the save icon 🖫 on the toolbar). Now you're backed up in case you mess up your data. This backup is good because now we'll demonstrate many beneficial procedures Design-Expert features in its design layout.

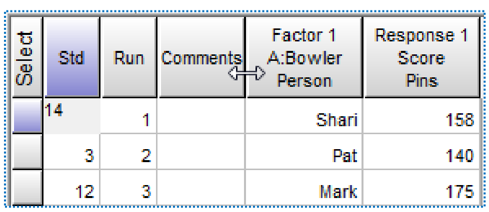
For example, right click the **Select** button. This allows you to control what Design-Expert displays. For this exercise, choose **Comments**.



Select button for choosing what you wish to display in the design layout

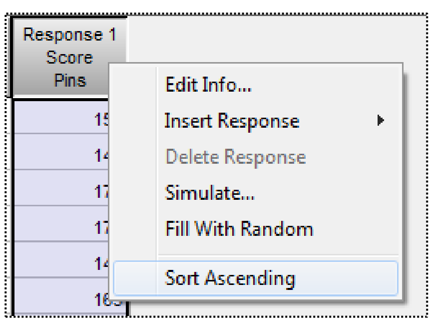
In the comments column above we added a notation that after run 8, the bowling alley proprietor re-oiled the lane – for what that was worth. Seeing Pat's scores, the effect evidently was negligible. ;)

☒ ***Explore entering comments***: Try this if you like. If comments exceed allotted space, move the cursor to the right border of the column header until it turns into a double-headed arrow (shown below). Then, just double-click for automatic column re-sizing.



Adjusting column size

Now, to better grasp the bowling results, order them from low-to-high as shown below by <u>right-clicking</u> the **Response** column header and selecting **Sort Ascending**.

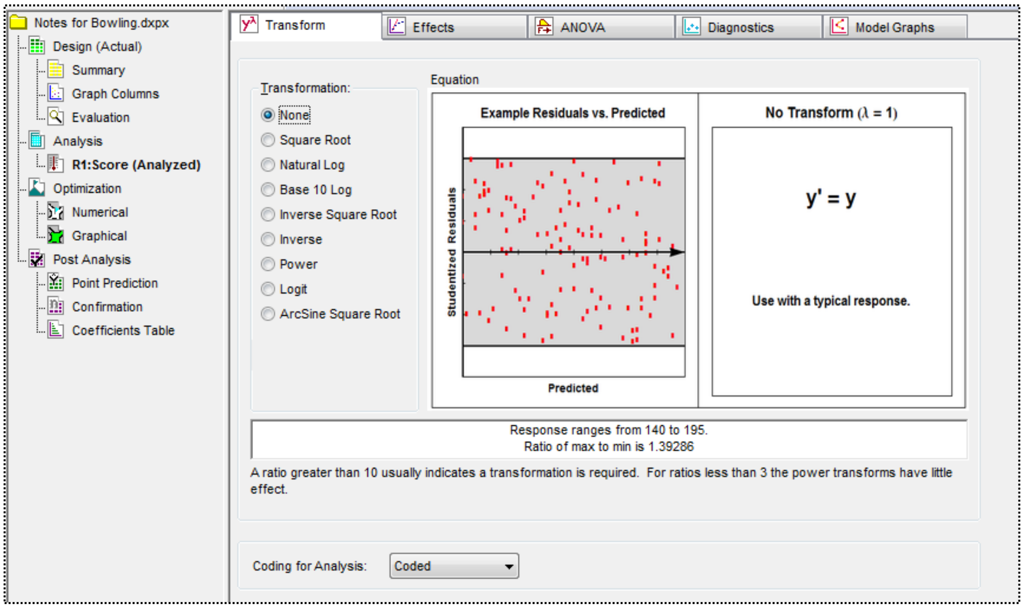


Sorting a response column (also works in the factor column)

You'll find sorting a very useful feature. It works on factors as well as responses. In this example, you quickly see that Mark bowled almost all the highest games.

Analyze the Results

Now we'll begin data analysis. Under the **Analysis** branch of the program (on the left side of your screen), click the **Score** node. ***Transform*** options appear in the main window of Design-Expert on a progressive tool bar. You'll click these buttons from left to right and perform the complete analysis. It's a very easy process. The ***Transform*** option gives you the opportunity to select a transformation for the response. This may improve the analysis' statistical properties.



Transformation button – the starting point for the statistical analysis

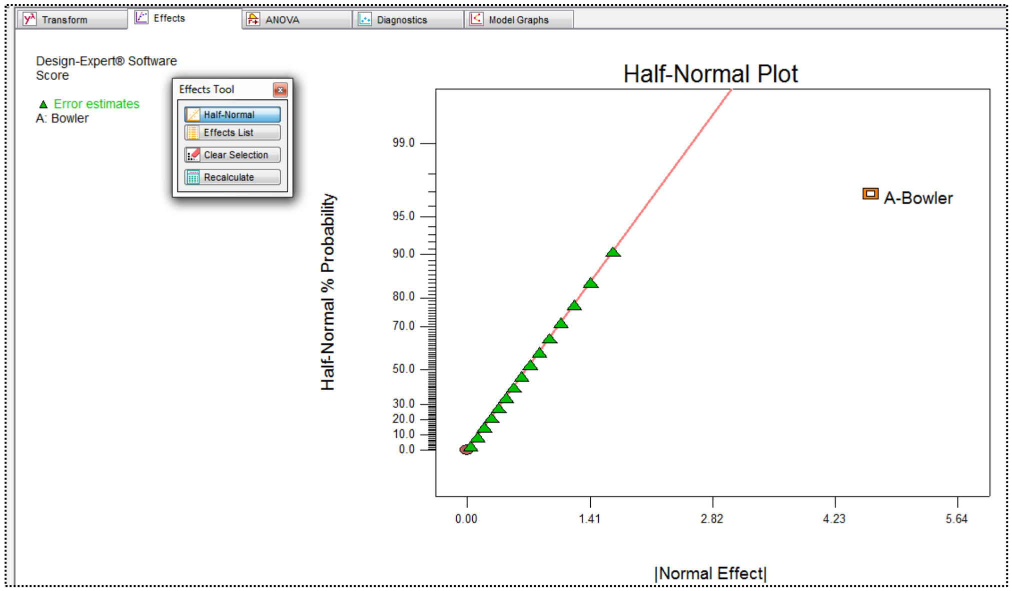
☒ ***Explore details on transformations***: If you need some background on transformations, first try Tips. For complete details, go to the Help command on the main menu. Click the Search tab and enter "transformations."

As shown at the bottom of the Transform screen above, the program provides data-sensitive advice, so press ahead with the default of None by clicking the ***Effects*** tab.

Examine the Analysis

By necessity, the tutorial now turns a bit statistical. If this becomes intimidating, we recommend you attend a basic class on regression, or better yet, a DOE workshop such as Stat-Ease's computer-intensive Experiment Design Made Easy.

Design-Expert now pops up a very specialized plot that highlights factor A—the bowlers—as an emergent effect relative to the statistical error, that is, normal variation, shown by the line of green triangles.

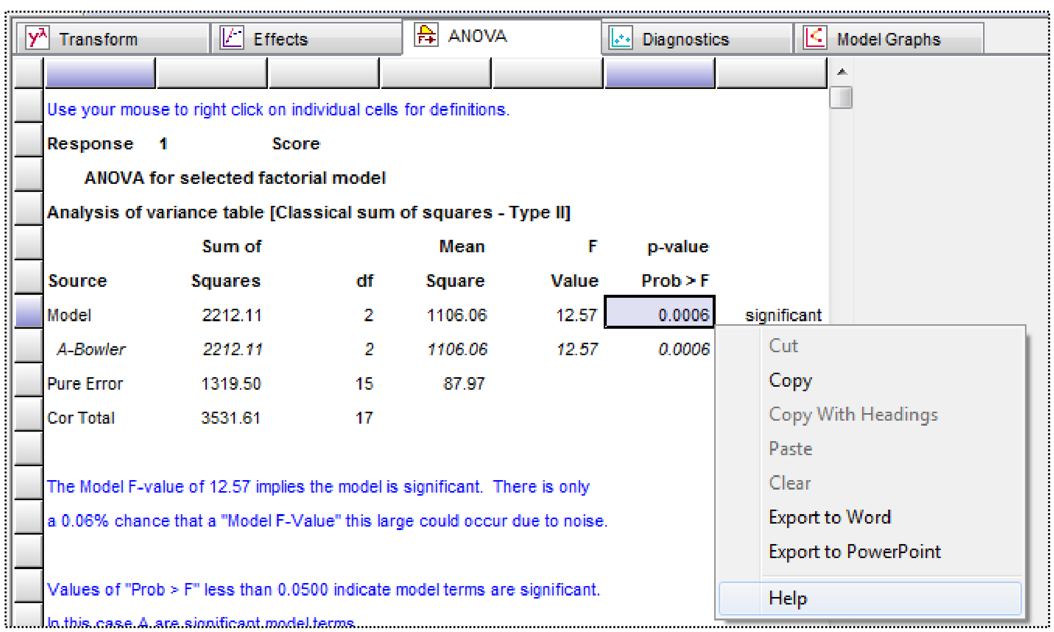


Initial view of the effect of Bowler

That is good! It supports what was obvious from the raw results—who bowls <u>does</u> matter.

☒ ***Explore half-normal plots***: If you want to learn more about half-normal plots of effects, work through the Two-Level Factorial Tutorial.

To get the statistical details, press the ***ANOVA*** (Analysis of Variance) tab. Notice to the far right side of your screen that Design-Expert verifies that the results are significant.



ANOVA results (annotated), with context-sensitive Help enabled via right-click menu

☒ ***Explore the ANOVA report***: Now select **View**, **Annotated ANOVA** from the menu atop the screen and uncheck (✓) this option. Note that the blue textual hints and explanations disappear so you can make a clean printout for statistically savvy clients. Re-select **View, Annotated ANOVA** to 'toggle' back all the helpful hints. Before moving on, try the first hint shown in blue: "Use your mouse to right click on individual cells for definitions." For example, perform this tip on the p-value of 0.0006 as shown above (select Help at the bottom of the pop-up menu). There's a wealth of information to be brought up from within the program with a few simple keystrokes: Take advantage!

Now click the 'floating' (moveable) **R-squared** Bookmark button (or press the scroll-down arrow at the bottom right screen) to see 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-Square	0.4620
PRESS	1900.08	Adeq Precisior	6.442

The "Pred R-Squared" of 0.4620 is in reasonable agreement with the "Adj R-Squared" of 0.5766; i.e. the difference is less than 0.2.

"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

☒ ***Explore the post-ANOVA statistics***: The annotations reveal the gist of what you need to know, but don't be shy about clicking on a value and getting online Help via a right-click (or try the F1 key). In most cases you will access helpful advice about the particular statistic.

Now click the **Coefficients** Bookmark button to view the output illustrated below.

Term	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High
Intercept	162.72	1	2.21	158.01	167.43
A[1]	-9.06	1	3.13	-15.72	-2.39
A[2]	15.61	1	3.13	8.95	22.27

Coefficient estimates

Here you see statistical details such as coefficient estimates for each model terms and their confidence intervals ("CI"). The intercept in this simple one-factor comparative experiment is simply the overall mean score of the three bowlers. You may wonder why only two terms, A1 and A2, are provided for a predictive model on three bowlers. It turns out that the last model term, A3, is superfluous because it can be inferred once you know the mean plus the averages of the other two bowlers.

Now let's move on to the next section within this screen: "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 means

Here are the averages for each of the three bowlers. As you can see below, these are compared via pair-wise t-tests in the following part of the ANOVA report.

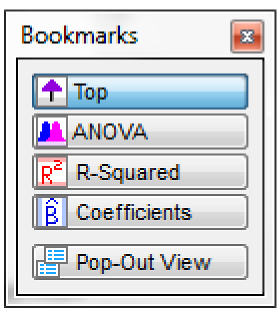
Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-24.67	1	5.41	-4.56	0.0004
1 vs 3	-2.50	1	5.41	-0.46	0.6509
2 vs 3	22.17	1	5.41	4.09	0.0010

Treatment means

You can conclude from the treatment comparisons that:

- Pat differs significantly (24.67 pins worse!) when compared with Mark (1 vs 2)
- The 2.5 pins mean difference between Pat and Shari (1 vs 3) is not significant (nor is it considered important by the bowling team's captain – recall in the design specification for power that a 10-pin difference was the minimum of interest)
- Mark differs significantly (22.17 pins better!) when compared with Shari (2 vs 3).

☒ ***Explore the Top feature***: Before moving ahead, press **Top** on the floating Bookmark. This is a very handy way of moving through long reports, so it's worth getting in the habit of using it.



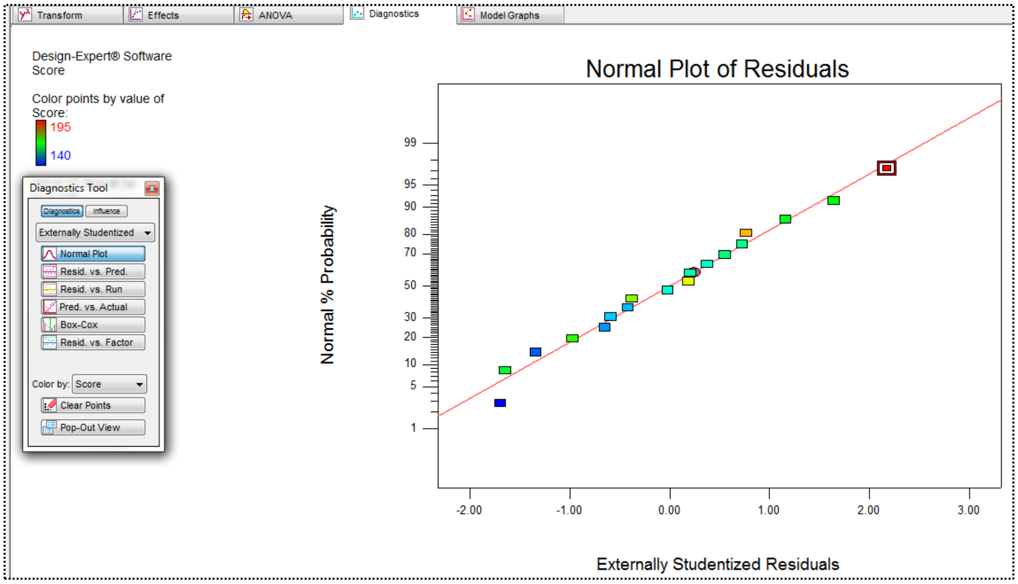
Back to the top

Analyze Residuals

Click the ***Diagnostics*** tab to bring up the normal plot of residuals. Ideally this will be a straight line, indicating no outlying abnormalities.

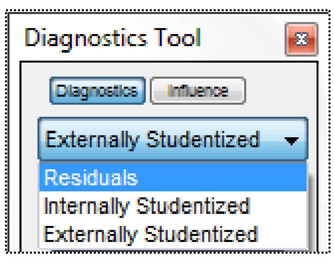
☒ ***Explore the 'pencil test'***: If you have a pencil handy (or anything straight), hold it up to the graph. Does it loosely cover up all the points? The answer is "Yes" in this example – it passes the "pencil test" for normality. You can reposition the thin red line by dragging it (place the mouse pointer on the line, hold down the left button, and move the mouse) or its "pivot point" (the round circle in the middle). However, we don't recommend you bother doing this – the program generally places the line in the ideal location automatically. If you need to re-set the line, simply double-click your left mouse button over the graph.

Notice that the points are coded by color to the level of response they represent – going from cool blue for lowest values to hot red for the highest. In this example, the red point is Mark's outstanding 195 game. Pat and Shari think Mark's 195 game should be thrown out because it's too high. Is this fair? Click this point so it will be highlighted on this and all the other residual graphs available via the Diagnostics Tool (the 'floating' palette on your screen).



Normal probability plot of residuals (195 game highlighted)

☒ ***Explore the Top feature***: Notice on the Diagnostics Tool that they are "studentized" by default. This converts raw residuals, reported in original units ('pins' of bowling in this example), to dimensionless numbers based on standard deviations, which come out in plus or minus scale. More details on studentization reside in Help. Raw residuals can be displayed by choosing it off the down-list on the Diagnostics Tool shown below. Check it out!



Other ways to display residuals

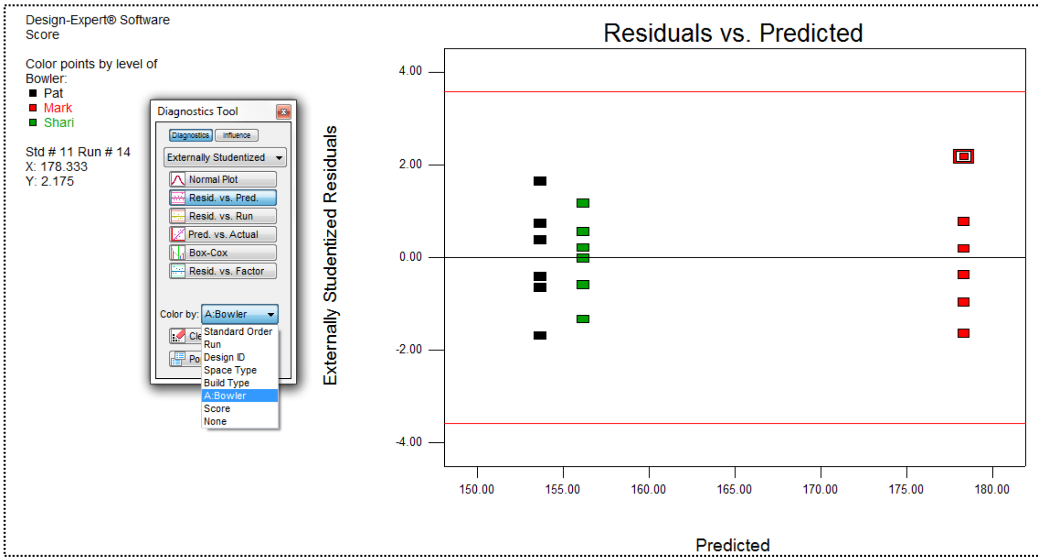
In any case, when runs have greater leverage (another statistical term to look up in Help), only the Studentized form of residuals produces valid diagnostic graphs. For example, if Pat and Shari succeed in getting Mark's high game thrown out (don't worry – they won't!), then each of Mark's remaining five games will exhibit a leverage of 0.2 (1/5) versus 0.167 (1/6) for each of the others' six games. Due to potential imbalances of this sort, we advise that you always leave the Studentized feature checked (as done by default). So if you are on Residuals now, go back to the original choice that came up by default (externally* studentized).

*P.S. Another aspect of how Design-Expert displays residuals by default is them being done "externally". This is explored in the Two-Level Factorial Tutorial. For now, suffice it to say that the program chooses this form of residual to provide greater sensitivity to statistical outliers. This makes it even more compelling not to throw out Mark's high game.

On the **Diagnostics Tool**, select **Resid. vs. Pred.** to generate a plot of residuals for each individual game versus what is predicted by the response model.

☒ ***Explore an apocryphal story***: Supposedly, "residuals" were originally termed "error" by statisticians, but the management people got upset at so many mistakes being made!

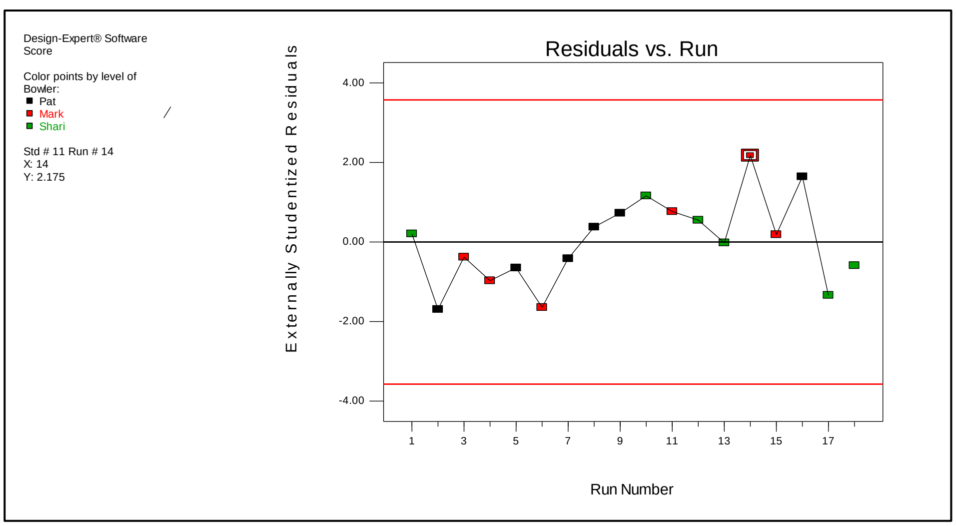
Let’s make it easier to see which residual goes with which bowler by pressing the down-list arrow for the **Color by** option in the **Diagnostics Tool** and selecting **A:Bowler**.



Residuals versus predicted values, colored by bowler

The size of the studentized residual should be independent of its predicted value. In other words, the vertical spread of the studentized residuals should be approximately the same for each bowler. In this case the plot looks OK. Don’t be alarmed that Mark’s games stand out as a whole. The spread from bottom-to-top is not out of line with his competitors, despite their protestations about the highest score (still highlighted).

Bring up the next graph on the Diagnostics Tool list – **Resid. vs Run** (residuals versus run number). *(<u>Note: your graph may differ due to randomization.</u>)*



Residuals versus run chart (<u>Note: your graph may differ due to randomization</u>)

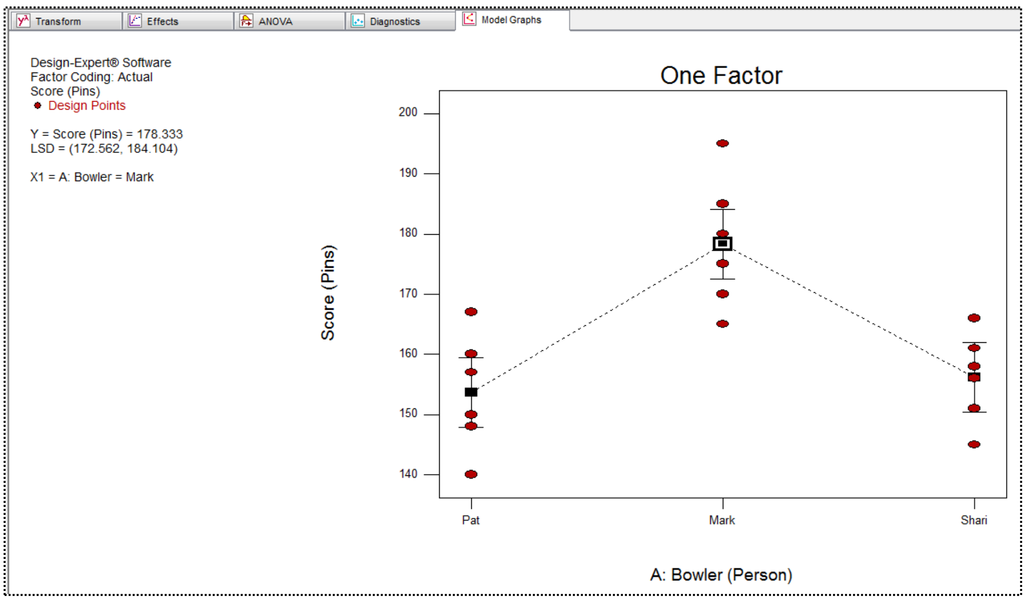
Here you might see trends due to changing alley conditions (the lane re-oiling, for example), bowler fatigue, or other time-related lurking variables.

☒ ***Explore repercussion of possible trends***: In this example things look relatively normal. However, even if you see a pronounced upward, downward, or shift change, it will probably not bias the outcome because the runs are completely randomized. To ensure against your experiment being sabotaged by uncontrolled variables, always randomize!

More importantly in this case, all points fall within the limits (calculated at the 95 percent confidence level). In other words, Mark's high game does not exhibit anything more than common-cause variability, so it should not be disqualified.

View the Means and Data Plot

Select the ***Model Graphs*** tab from the progressive tool bar to display a plot containing all 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. Note how conveniently Design-Expert scaled the Y axis from 140 to 200 pins in increments of 10.



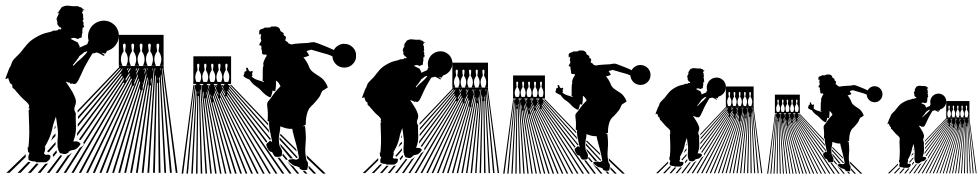
One-factor effects graph with Mark's predicted score (mean) highlighted

The squares in this effects plot represent predicted responses for each factor level (bowler). Click the square representing Mark's mean score as shown above. Notice that Design-Expert displays the prediction for this treatment level (reverting to DOE jargon) on the legend at the left of the graph. Vertical 'I-beam-shaped' bars represent the 95% least significant difference (LSD) intervals for each treatment. Mark's LSD bars don't overlap horizontally with Pat's or Shari's, so with at least 95% confidence, Mark's mean is significantly higher than the means of the other two bowlers.

☒ ***Explore other points on the model graph***: Oh, by the way, maybe you noticed that the numerical value for the height of the LSD bar appeared when you clicked Mark's square. You can also click on any round point to see the actual scores. Check it out!

Pat and Shari's LSD bars overlap horizontally, so we can't say which of them bowls better. It seems they must spend a year in a minor bowling league and see if a year's worth of games reveals a significant difference in ability. Meanwhile, Mark will be trying to live up to the high average he exhibited in the tryouts and thus justify being chosen for the Stat-Ease bowling team.

That's it for now. Save your results by going to **File**, **Save**. You can now **Exit** Design-Expert if you like, or keep it open and go on to the next tutorial – part two for general one-factor design and analysis. It delves into advanced features via further adventures in bowling.



General One-Factor Tutorial

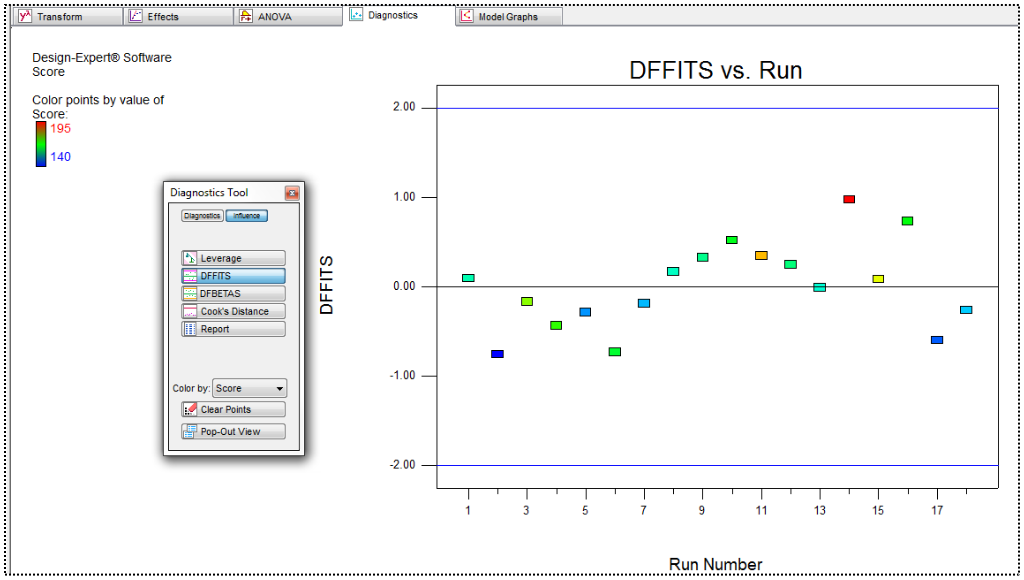
(Part 2 – Advanced Features)

Digging Deeper Into Diagnostics

(Caution: Only the more daring new users should press ahead from here—those who like to turn over every rock to see what's underneath, that is—the types who are curious to know everything there is to know. If that's not you, skip the rest and go on to another tutorial if it offers feature you need for your particular experiment.)

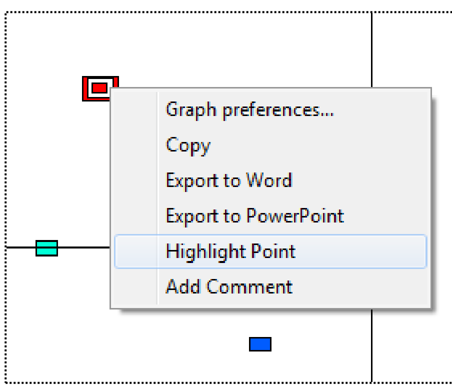
If your bowling data is active in Design-Expert® software from Part 1 of this tutorial, continue on. If you exited the program, re-start it and use **File**, **Open Design** to open data file (**Bowling.dxpx**). Otherwise, set up this data file as instructed above in our General One-Factor Tutorial (Part 1 – The Basics). Then, under the **Analysis** branch (you may already be here) click the **Score** node and press the ***Diagnostics*** tab.

We're now going to look at a new graph in the **Diagnostics Tool**. Click the **Influence** option on the Diagnostics Tool palette. Then click on **DFFITS**. This statistic, which stands for difference in fits, measures the change in each predicted value that occurs when that response is deleted. The larger the absolute value of DFFITS, the more it influences the fitted model. (For more details on this statistic and the related deletion diagnostic, DFBETAS, see our program Help or refer to Raymond Myers' *Classical and Modern Regression with Applications, 2nd Edition* (PWS Pub. Co., 1990).)



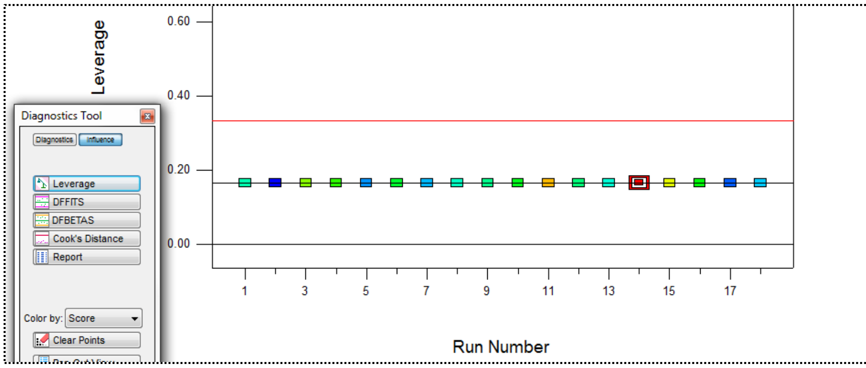
DFFITS graph (your graph may differ due to random runs)

Notice that one point lies above the rest. (The pattern on your graph may differ from what we show here due to randomized run order, but this isn't a concern in this discussion.) The top-most point is Mark's high game, which earlier created controversy, particularly among competitors Pat and Shari. Mark's point falls far below a relatively conservative high benchmark of plus-or-minus two for the DFFITS. So, taking all other diagnostics into consideration, we don't advise that this particular run be investigated further. Nevertheless, for purposes of learning how to use new Design-Expert software features, right-click Mark's top point with your mouse and select **Highlight Point** as shown below.



Highlighting a point

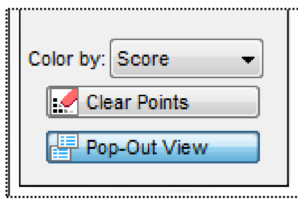
Myers demonstrates mathematically that the DFFITS statistic is really the externally studentized residual multiplied by high leverage points. Click the **Leverage** button and you'll see that all runs exhibit equal leverage here because an equal number of runs were made at each treatment level (all three bowlers rolled six games each).



Leverages

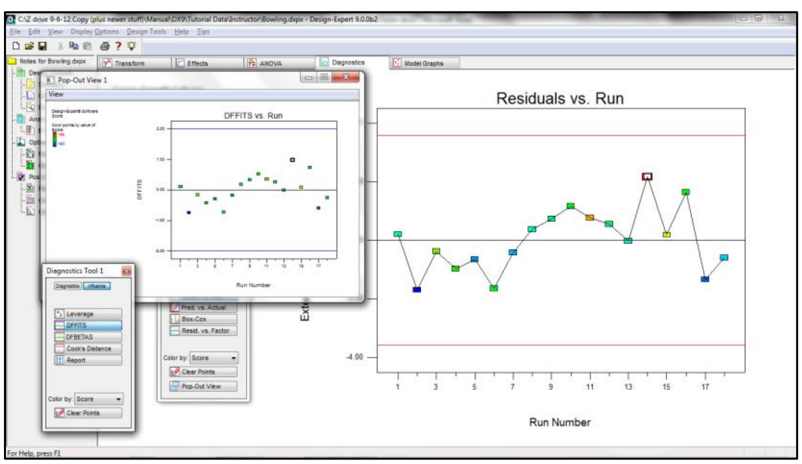
Therefore, this DFFITS exhibits a pattern identical to that shown on the externally studentized residual graph, which you studied in the preceding tutorial. The reason we're reviewing this is to set the stage for what you'll do later in this tutorial – unbalance the leverages to make this session more significant for diagnostic purposes.

☒ ***Explore the Pop-Out View feature***: Now is a good time to go back to the **DFFITS** plot and press the **Pop-Out View** button the very bottom of the **Diagnostics Tool**.



Pop-Out View button

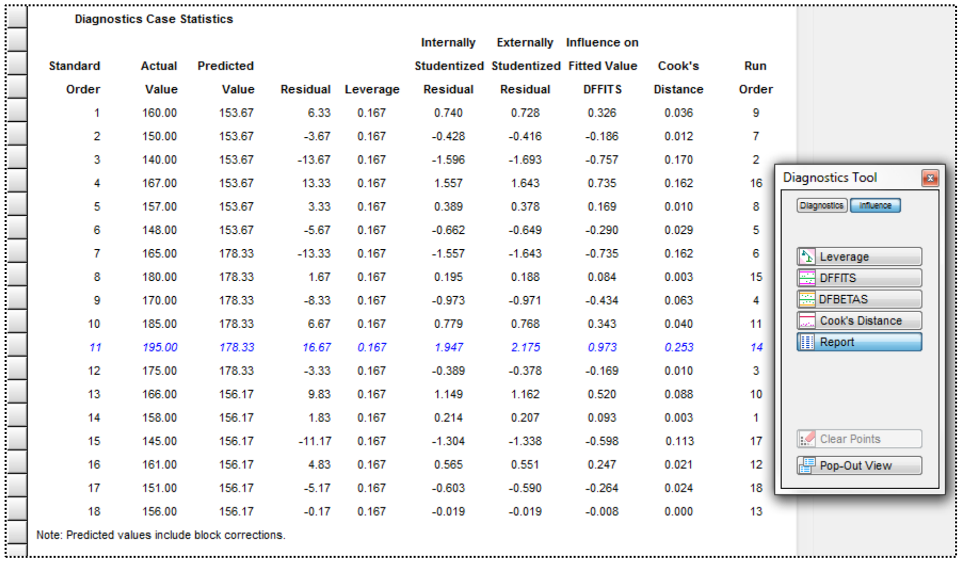
Next go back via the **Diagnostics** button to the **Resid. vs Run** plot and verify the statement above that in this case these two plots (DFFITs and residuals versus run) exhibit the same pattern. (You may need to press Alt-Tab to get the windows you want on the same screen.)



Demonstration of pop-out view to see two plots side by side

You'd best now close out the pop-out view by pressing the X at the upper right corner. Otherwise your screen will get too messy.

Here's one final Design-Expert software feature for you before we leave the **Diagnostics Tool**: Click the **Report** button to get a table of statistics case-by-case in standard order for the entire experiment. For those of you who prefer numbers over pictures (statisticians for sure!), this should satisfy your appetite. Notice that Mark's high 195 game is highlighted in blue text as shown below.



Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Internally Studentized Residual	Externally Studentized Residual	Influence on Fitted Value DFFITS	Cook's Distance	Run Order
1	160.00	153.67	6.33	0.167	0.740	0.728	0.326	0.036	9
2	150.00	153.67	-3.67	0.167	-0.428	-0.416	-0.186	0.012	7
3	140.00	153.67	-13.67	0.167	-1.596	-1.693	-0.757	0.170	2
4	167.00	153.67	13.33	0.167	1.557	1.643	0.735	0.162	16
5	157.00	153.67	3.33	0.167	0.389	0.378	0.169	0.010	8
6	148.00	153.67	-5.67	0.167	-0.662	-0.649	-0.290	0.029	5
7	165.00	178.33	-13.33	0.167	-1.557	-1.643	-0.735	0.162	6
8	180.00	178.33	1.67	0.167	0.195	0.188	0.084	0.003	15
9	170.00	178.33	-8.33	0.167	-0.973	-0.971	-0.434	0.063	4
10	185.00	178.33	6.67	0.167	0.779	0.768	0.343	0.040	11
11	*195.00*	*178.33*	*16.67*	*0.167*	*1.947*	*2.175*	*0.973*	*0.253*	*14*
12	175.00	178.33	-3.33	0.167	-0.389	-0.378	-0.169	0.010	3
13	166.00	156.17	9.83	0.167	1.149	1.162	0.520	0.088	10
14	158.00	156.17	1.83	0.167	0.214	0.207	0.093	0.003	1
15	145.00	156.17	-11.17	0.167	-1.304	-1.338	-0.598	0.113	17
16	161.00	156.17	4.83	0.167	0.565	0.551	0.247	0.021	12
17	151.00	156.17	-5.17	0.167	-0.603	-0.590	-0.264	0.024	18
18	156.00	156.17	-0.17	0.167	-0.019	-0.019	-0.008	0.000	13

Note: Predicted values include block corrections.

Report with case statistics used for preceding diagnostics graphs

Remember, you can right-click any value in reports of this nature within Design-Expert software to view context-sensitive Help with statistical details.

Modifying the Design Layout

Design-Expert offers great flexibility when modifying data in its design layout. We'll see in this bowling scenario how our software allows you to modify an existing design with added blocks and factor levels.

The outcome of the bowling match appears to be definitive, especially from Mark's perspective. But Pat and Shari demand one more chance to prove themselves worthy of the team. They still think Mark's high 195 game was a fluke, even though this isn't supported by the diagnostic analysis. Mark objects and a dispute ensues.

Attempting compromise, the team captain decides to toss out the highest and lowest games for each of the three bowlers and replace them with two new scores each. But Ben, a newly hired programmer and avid bowler, arrives at the alley and is allowed to participate in this second block of runs. (Yes, this makes little sense, but it will add some interest to this tour of Design-Expert's flexibility for design and analysis of experiments – no matter how convoluted they become in actuality.)

It quickly becomes apparent that this new kid does things differently. He's a lefty with a huge hook that's hard to control. To aggravate this variability, Ben does something very different from other bowlers – he does not put his thumb in the ball's hole made for that purpose. When Ben's odd approach works, the pins go flying. But as likely as not, that ball slides off into the left gutter or careens over the edge on the right.

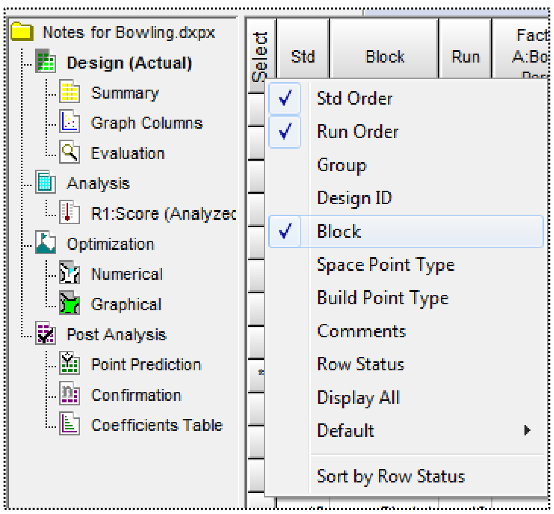
The results for Ben and the three original bowling team candidates are below.

Block	Game	Pat	Mark	Shari	Ben
1	1	160	~~165~~	~~166~~	NA
1	2	150	180	158	NA
1	3	~~140~~	170	~~145~~	NA
1	4	~~167~~	185	161	NA
1	5	157	~~195~~	151	NA
1	6	148	175	156	NA
2	1	162	175	163	200
2	2	153	180	166	130

Bowling scores with high and low games replaced by two new games (plus a new guy)

To enter this new data (and ignore some of the old), click the **Design** node near the upper left of your screen. You should now see the bowling data from the first tutorial. Mark's high 195 game remains highlighted in blue text (assuming you clicked on it as instructed on page 2 of this tutorial while performing the diagnostics).

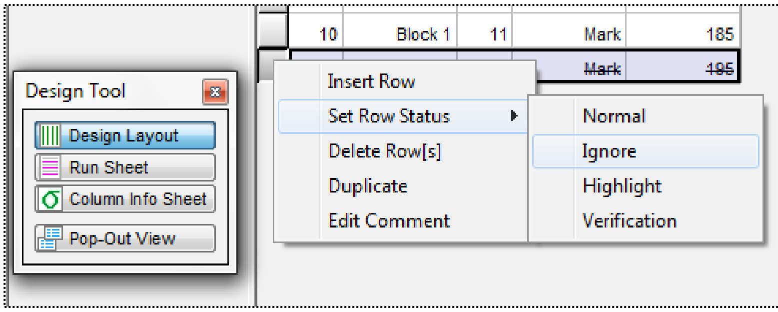
<u>Right click</u> the **Select** column header and click **Block.** This design attribute is now needed to accommodate the new bowler's (Ben's) incoming score data.



Selecting block to display it as a column in the design layout

<u>Right click</u> the **Response** column header and choose **Sort Ascending**. You did this before in Part 1 so you are now have this feature mastered…we hope. ;)

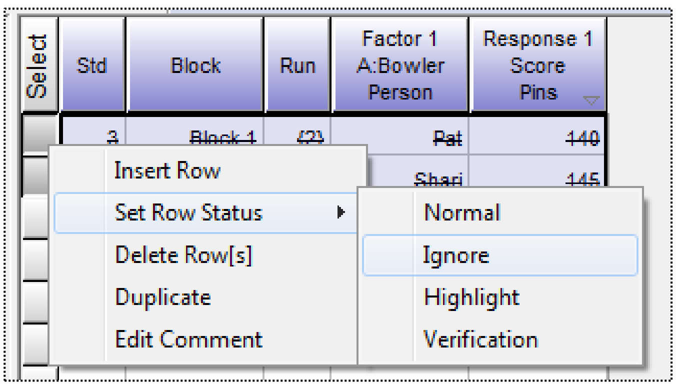
Mark's best game now drops to the very bottom. Let's single him out first to placate Pat and Shari. <u>Right-click</u> the square button at the left of the last row (Mark's 195 score). Click **Set Row Status**, then **Ignore** as shown below.



Ignoring Mark's high game

By the way, it's OK to change your mind when modifying your design layout: You can 'un-ignore' a row by clicking Set Row Status, Normal.

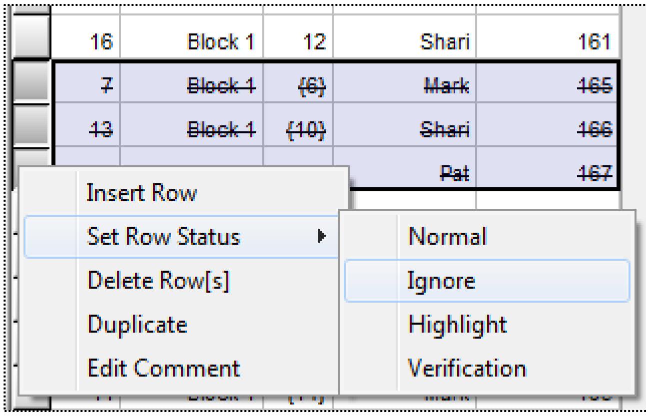
Now let's really get Pat's and Shari's hopes high by excluding their low games from consideration. Click the square button (in the Select column) to the left of the top row (Pat's low 140 game) and, <u>while pressing down the **Shift**</u> key, also click the button in the Select column's second row (Shari's low 145 game). Release the **Shift** key. Keep your mouse within the Select column's first or second row, <u>right-click</u> and choose **Set Row Status**, **Ignore** for these two low games, as shown below.



Ignoring the low games for Pat and Shari

Now move down a few rows and click the square button in the Select column's row showing Mark's low 165 game.

Notice the two rows below Mark's low 165 game – the high games for Shari (166) and Pat (167). It's now time for Shari and Pat to pay the price for complaining. While first <u>pressing and holding down the **Shift**</u> key, click the following two square buttons in the Select column's row: Shari's high 166 game and Pat's high 167 game. Release the **Shift** key. Three rows should now be highlighted in light blue as shown below. Keep your mouse within the Select column's highlighted three rows, <u>right-click</u> and choose **Set Row Status**, **Ignore**.



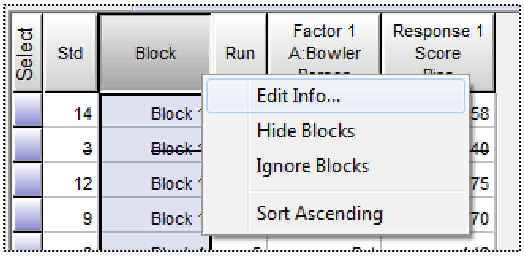
Ignoring Mark's low game and the high games for Shari and Pat

Now let's restore the original layout order. Right-click the **Std** column header, then choose **Sort Ascending**. Compare your screen with what we show below. If there are differences, fix them now to match this screenshot. However, remember that the run number is random, so you don't need to fix that.

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

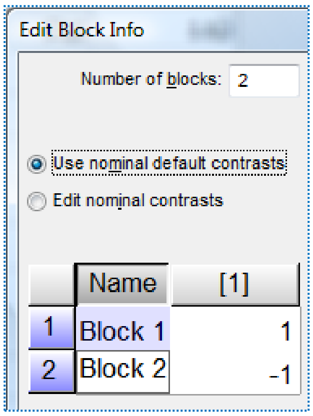
Back to standard order after low and high games ignored for each bowler

Now create a new block (needed for the second round of bowling) by right-clicking the **Block** column header and choosing **Edit Info** as shown below.



Creating a new block

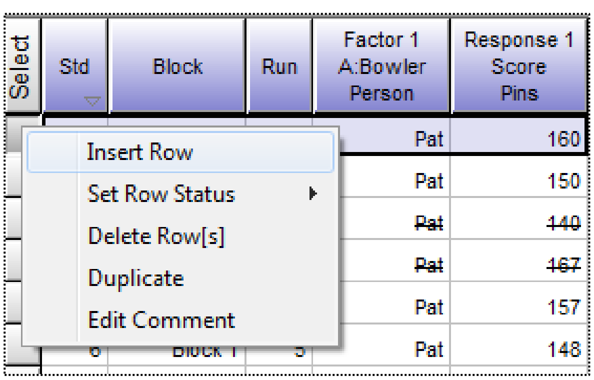
You'll see a form allowing you to assign names to the block(s). Don't bother doing this now. As shown below, change **Number of Blocks** at the top to **2**. Press the **Tab** key to see the change take effect. (If the name field truncates, click and move the right border of the column header to re-size it.)



Adding a second block of runs

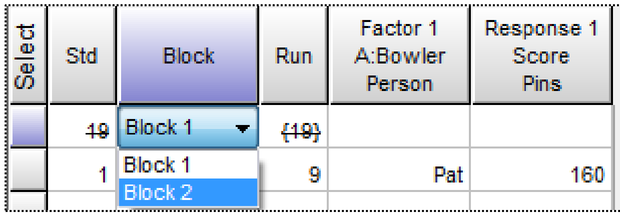
Click **OK**. It seems that nothing changed, but actually the program now knows that you will be conducting another block of runs.

Now you are ready to begin adding and/or duplicating rows. This can be accomplished in different ways, depending on your ingenuity. We'll follow routes revealing as many of the editing features as possible, although they may not demonstrate the most elegant approaches. As shown below, <u>right click</u> the Select column's square button at the <u>left</u> of the <u>first row</u> (Pat's 160 game) to bring up the editing menu. Click the first selection, **Insert Row**.



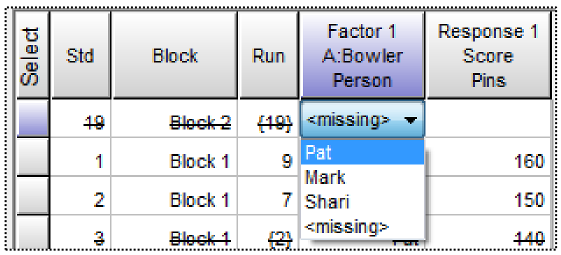
Inserting a new row

You now see a new row containing blanks for the bowler and the score. (Don't worry if it's being ignored – crossed out, that is – for the moment.) Click the first row's block cell directly below the block field header, then click the list arrow. Select **Block 2** as shown below.



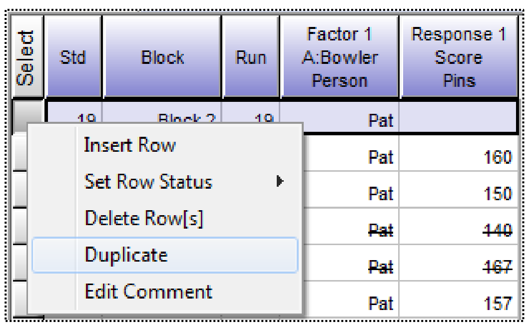
Changing the block number

Click the blank field for bowler and press the list arrow (⮟). Select **Pat**. (We're using categorical factors here, but if this were a numerical field, you'd enter a value.)



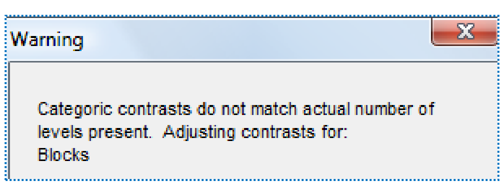
Entering a categorical value for factor

Again, right-click the Select column's square button at the left of the first row to bring up the editing menu as shown below. Click **Duplicate**.



Duplicating a row

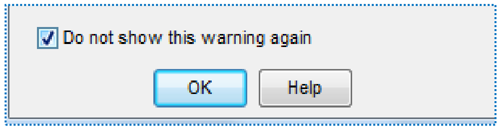
Design-Expert may pop up a warning like the one shown below.



Warning about categoric contrasts

The program is recognizing a potential problem here and is alerting you that only one bowler is in the second block. You need not worry at this stage because you will be adding others. Click the check option **Do not show this warning again.**

This will save you aggravation later. Don't worry – you will not be unprotected indefinitely. This warning will be re-enabled the next time you start the program.



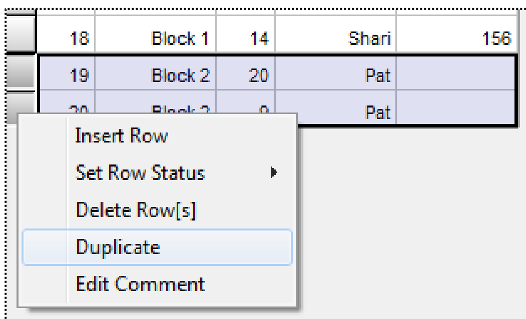
Turning off a warning (it will come back the next time you run the program)

Press **OK** to proceed.

Right-click the **Block** column header and choose **Sort Ascending**.

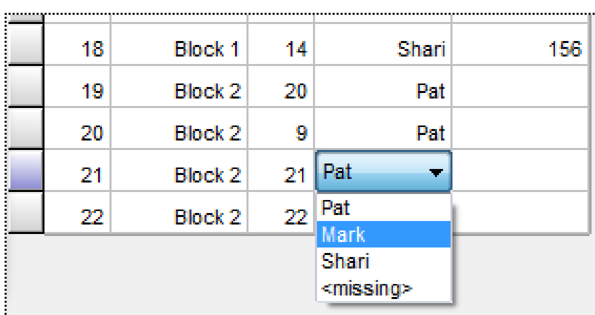
Two new rows are now seen at the bottom of your design layout. We need two new rows apiece for Shari and Mark. Let's simply duplicate Pat's two new rows and update the names. Do this by first clicking the Select column's square button at the left of Pat's first new row, so it is highlighted. Then while holding down the Shift key, click the Select column's square button at the left of Pat's second new row. Both rows should now be highlighted. (This is a bit tricky, but it saves time.)

Now right-click any Select column's square button at the left of the highlighted block and select **Duplicate**. (If the warning screen pops up again, click OK.)



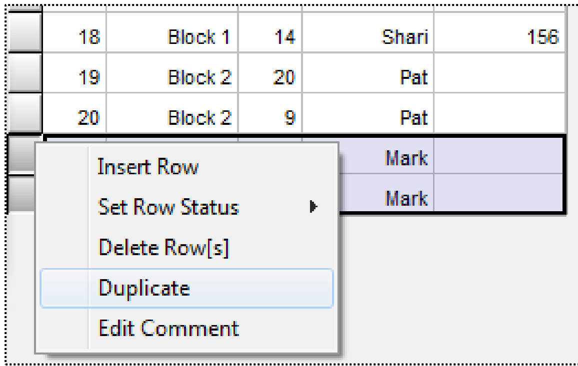
Duplicating a block of rows

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



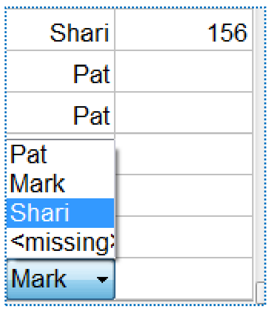
Changing name of bowler

Do the same for the last row. You now should have two new rows for both Pat and Mark. Click the Select column's square button at the left of Mark's first new row, so it is highlighted. Then while holding down the **Shift** key, click the Select column's square button at the left of Mark's second new row. Both rows should now be highlighted. As before, right-click any Select column's square button at the left of the highlighted block and select **Duplicate**.



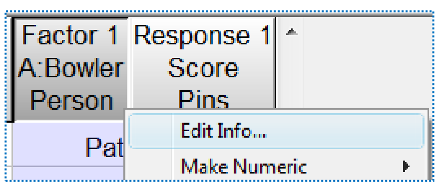
Duplicating two more rows

In the first duplicated row, click the field for **Bowler** and select **Shari**. Do the same for the last row.



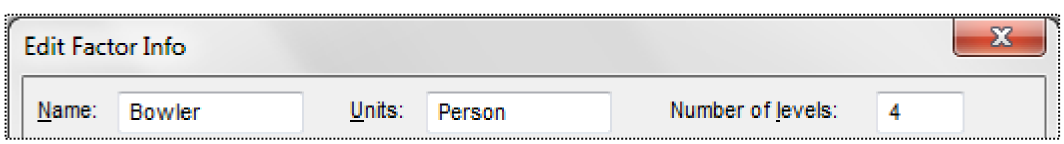
Completing lineup for block 2 – the second round of bowling

But what about the new kid – Ben? We need to identify him as a new competitor in this bowling contest. Do this by right-clicking the header for **Bowler** and selecting **Edit Info**.



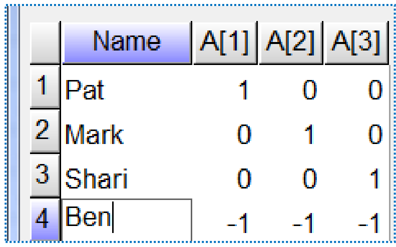
Getting ready to add a new level for the factor

Change **Number of Levels** to **4** (see below left).



Adding another bowler

Press **Tab** once. Click the field intersecting at **Name** column and row **4** (below right). Type the name **Ben**.

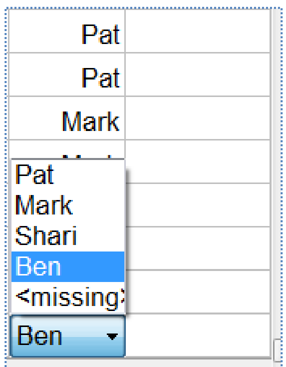


	Name	A[1]	A[2]	A[3]
1	Pat	1	0	0
2	Mark	0	1	0
3	Shari	0	0	1
4	Ben	-1	-1	-1

Entering the new bowler

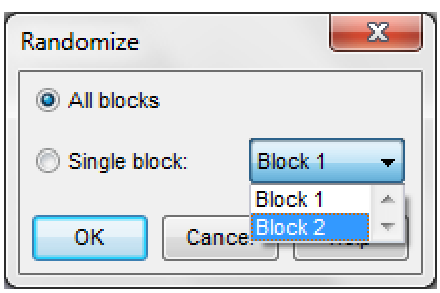
Press **OK**. Now duplicate two more rows by clicking the Select column's square button at the left of the first of Shari's two new games at the bottom of the list. While holding down the **Shift** key, click the Select column's square button at the left of the last run. Finally, right-click any Select column's square button at the left of the highlighted block and select **Duplicate.**

In both of these new duplicated rows, click the fields for **Bowler** and select **Ben**.



Ben now on the list as a bowler

An important aside: Always randomize your run orders for actual experiments. For purposes of this tutorial, this will just be a bother, so do this only if you wish to try it out, but it's very easy to do – simply right-click the **Run** column-header and do this for **Block 2** as shown.



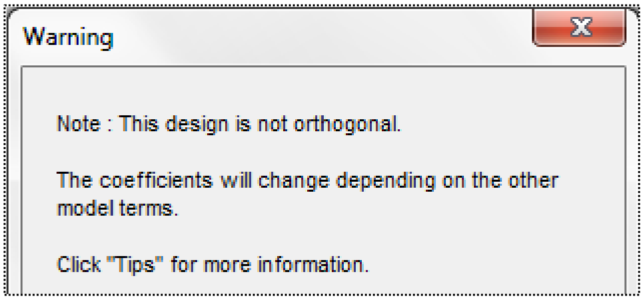
How to randomize the run order in the second block

To make it easier to enter the results, right-click the **Std** column and **Sort Ascending**. Now enter the eight new scores as shown below.

Pat	162
Pat	153
Mark	175
Mark	180
Shari	163
Shari	166
Ben	200
Ben	130

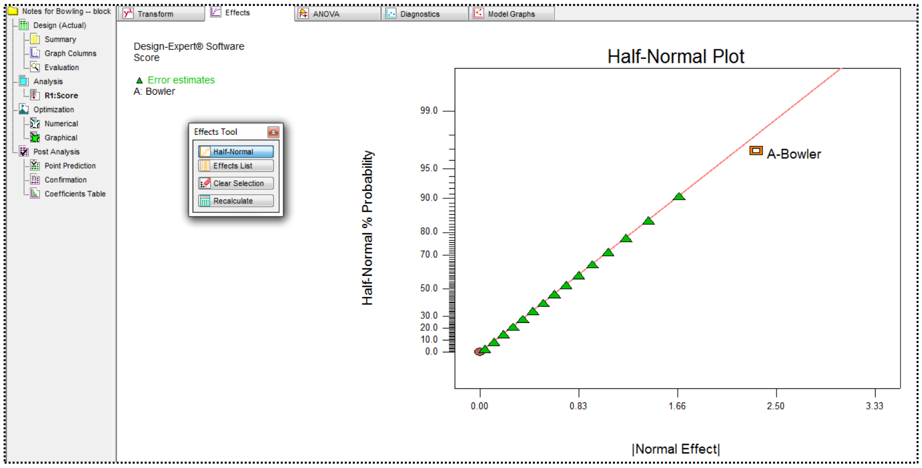
Data entered for second block of games

Go ahead now and re-analyze your data by clicking the **Score** node under **Analysis**. Move through ***Transform*** and click on the ***Effects*** tab. A warning pops up that the design is not "orthogonal."



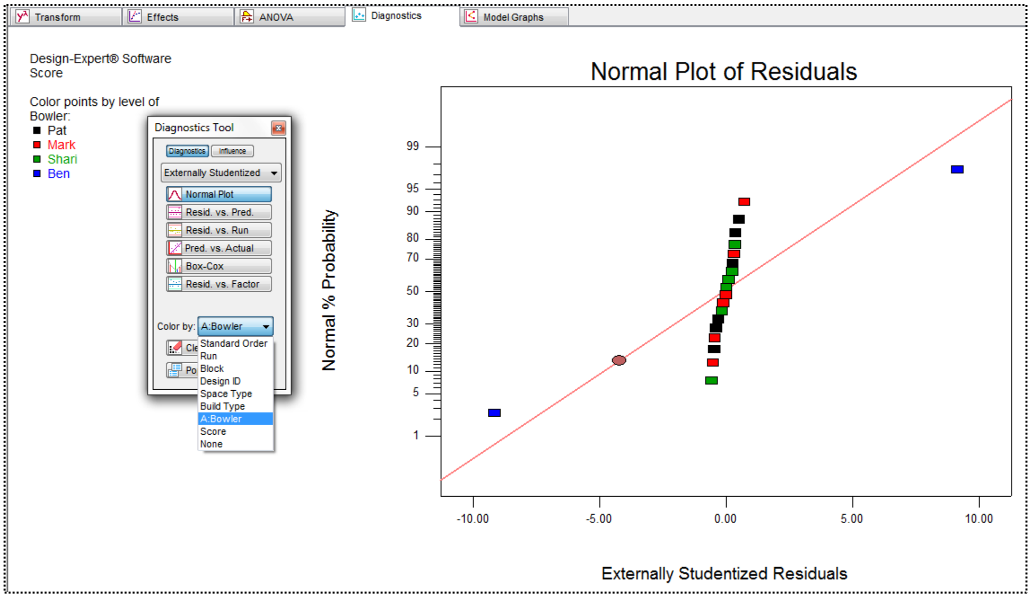
Warning about design now being non-orthogonal

This is a mathematical artifact of our ad hoc addition of runs in a second block. It will not create any material impact on the outcome so just press on via **OK** and click the square appearing at the end of the green triangles (error estimates) on the half-normal plot of effects. This puts A-Bowler in your model.



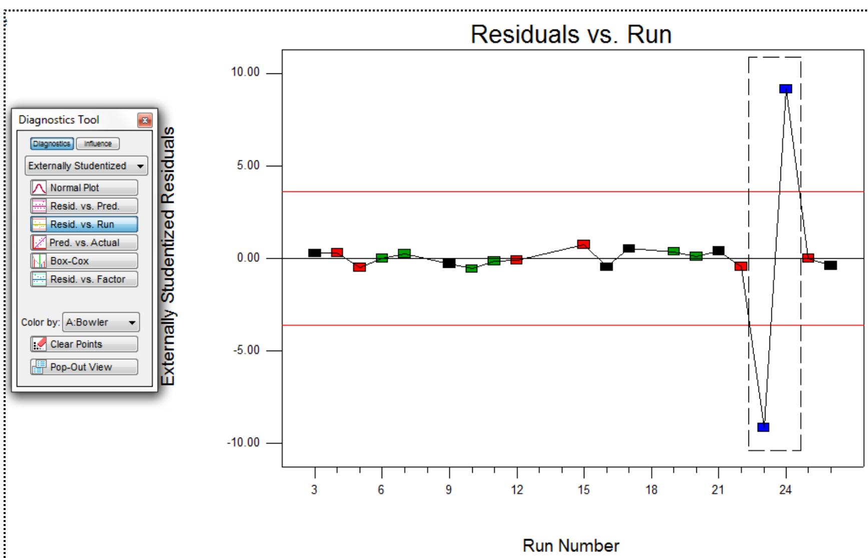
Bowler picked on half-normal plot of effects

Proceed to ***ANOVA*** (overlooking this model being not significant) and then to ***Diagnostics***. As you will see, something is abnormal about this data. Do you notice that the residuals now line up very poorly, especially at the extreme points as shown below? On the floating **Diagnostics Tool** change **Color by** to **A: Bowler**.



Diagnostics for bowling results – part two: Normal plot with poorly aligned residuals

Now (referring to the color key at the left of the plot) you see that the results from Ben do not fit with the others (his games are the two outliers – low and high). Considering his odd, unstable style of bowling, this should be no surprise. Click the **Resid. vs Run** button to bring up the externally studentized residuals – a good tool for detecting outliers. <u>Drag your mouse over Ben's residuals at the far right</u>.

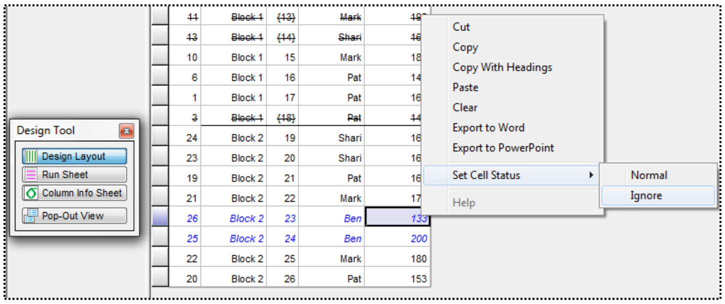


Ben's games highlighted for being outliers

Both points should now be highlighted. We must ignore or delete them. (Sorry Ben, odd behavior by programmers is considered normal at Stat-Ease, but not when it comes to bowling!)

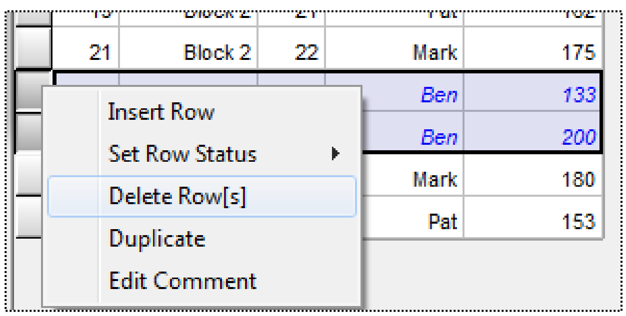
Click the **Design** node (upper left) to get back to the home base of the design layout. Notice that Ben's games are conveniently highlighted in blue text so they can easily be deleted.

☒ ***Explore an option for ignoring data***: It provides no advantage in this case, which features only one response measure, but you can ignore a specific result by right-clicking that cell and setting Set Cell Status to Ignore as shown below.



Ignoring a single cell – an option that's not recommended for this case

In this case you could ignore his entire runs (we explained how to do this earlier). Better yet, simply delete them altogether. No offense to Ben, but given that he only bowled two games and his unorthodox style creates such abnormal variability, it is best now to <u>click</u> the Select column's square button at the <u>left</u> of his <u>first</u> score of 200 (making him feel really bad ☹), <u>shift-click</u> the button below it for the <u>second</u> game of 130 (not so sorry to see this gone!), then without moving your mouse, <u>right-click</u> and select **Delete Row(s)**.



Deleting Ben's games

Click **Yes** on the warning that pops up about deleting rows (a safety precaution) and **OK** when asked to re-sequence the runs to fill the gap. Then go ahead and re-analyze the results.

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. Because 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 suspect rows and changing Set Row Status to Normal. (Or, if you're adept at manipulating lines of text or data with your mouse, do all rows at once using a click and shift-click.) Give this a try! Then re-analyze one last time.

By working through this exercise, you now see how easy it is to manipulate Design-Expert's design layout.

P.S. Still feeling bad about deleting Ben's scores? Don't worry – he gets to bowl with Pat and Shari in a lesser league. After bowling for an entire year (roughly 100 games), it will become clear whether Ben's crazy way of bowling will pay off by achieving a good average overall. After all, his 2 game average of 165 wasn't so bad, just inconsistent (high variability). With more data, his true ability will become more apparent.