

# **Data Manipulation with JMP**

## **Introduction**

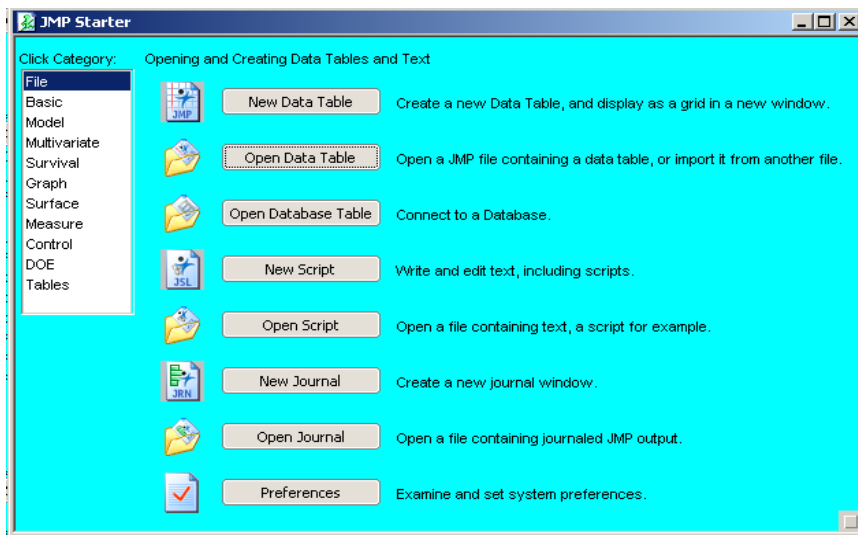
JMP was introduced in 1989 by its Principle developer, John Sall at SAS. It is a richly graphic environment and very useful for both preliminary data exploration and more advanced statistics. It is user friendly and provides both a series of menus and more advanced statistical techniques. For those wishing to go beyond the routine, a scripting language is available and many pre-scripted routines are available to ease the burden of coding.

It is assumed that all students have at least a cursory knowledge of the menu offerings and have used JMP, however sparingly, in the past. We will be using ver 9 of JMP for this lecture, and all manipulations described here may be found in two JMP manuals: Introductory Guide and the User's Guide.

## **Background: Getting the data into JMP**

There are two methods of doing this: simple cut-and-paste and importation via the Open Data Table menu of the JMP Starter box, or the File/Open command from the main menu bar. We will assume that your data comes from an EXCEL file to simplify the operations.

When doing a simple cut-and-paste from EXCEL, two drawbacks are immediately apparent: 1) with large data sets this can be exceedingly tedious and 2) with previous versions of JMP, the headers didn't come in with the data (and did you really want to re-type all that?!). The data can be readily imported from EXCEL as an .xls file merely by using the Open Data box in the JMP Starter dialog (note: if JMP Starter is not visible from your screen, just use the View/JMP Starter command from the toolbar). With this tool you can browse to the proper file and open it in JMP. Also, with the new Edit/Paste with Column Names' command, this is easily done in version 9.



JMP will easily import files as .xls, .txt, .csv, .dat, SAS and of course, .jmp. Version 7 & 8, unlike previous versions, will correctly place the column headers, even if they are not in the first row of an imported EXCEL spreadsheet.

### Use of the Data Table Window

Now let's bring in a typical data table by selecting a file and hitting the Open button on the Open Data File dialog.

File Edit Tables Rows Cols DOE Analyze Graph Tools View Window Help

GLJUG\_Data Manipulation

Windows

JMP Starter

GLJUG\_Data Manipul

glucose\_Feb\_08

GLJUG\_Data Manipulat

Columns (9/0)

Date

Job Dispo

Run Dispo

Inst. #

Job Lot #

Rep

Low Control

Mid Control

High Control

Rows

All rows 27




Selected 0

Excluded 0

Hidden 0

Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control	
1	1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
2	1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
3	1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
4	1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
5	1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
6	1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
7	2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
8	2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
9	2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288
10	2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
11	2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035
12	2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193
13	1	ADJREQ	ADJREQ	2	2	1	1.48	4.913	16.918
14	1	ADJREQ	ADJREQ	2	2	2	1.645	4.011	11.235
15	1	ADJREQ	ADJREQ	2	2	3	1.391	5.479	16.020
16	1	ADJREQ	ADJREQ	2	2	1	1.171	4.606	10.986
17	1	ADJREQ	ADJREQ	2	2	2	1.415	3.635	12.373
18	1	ADJREQ	ADJREQ	2	2	3	1.95	4.592	15.635
19	2	ADJREQ	ADJREQ	3	2	1	1.695	3.213	17.953
20	2	ADJREQ	ADJREQ	3	2	2	2.247	3.944	14.551
21	2	ADJREQ	ADJREQ	3	2	3	1.292	4.648	11.484
22	2	ADJREQ	ADJREQ	3	2	1	1.022	4.354	18.583
23	2	ADJREQ	ADJREQ	3	2	2	1.53	5.808	14.963
24	2	ADJREQ	ADJREQ	3	2	3	1.992	4.281	13.176
25	2	ADJREQ	ADJREQ	3	2	1	1.673	3.301	15.630
26	2	ADJREQ	ADJREQ	3	2	2	1.311	6.249	16.458
27	2	ADJREQ	ADJREQ	3	2	3	1.795	4.562	18.147

Before proceeding there are several items that need mention and emphasis. You will notice that some screenshots display far more items than others. You may add as many as you like depending upon your needs by using the View/Show Toolbars choice.

**Big Emphasis:** Once data is imported into JMP, it should be examined for accuracy and completeness. Try checking the first, middle, and last rows to ensure that they match the original EXCEL dataset. Check the area to the left of the table to verify that the expected number of rows and columns has been transferred. Also, be aware that when you proceed to do a hypothesis test or certain manipulations, JMP will be very picky as to the **Type** of the variables that you have imported. For certain operations the data must be continuous, and for others categorical (i.e. nominal or ordinal). For example your instrument data is usually continuous , the instruments and lots may be nominal , and dates are ordinal . We will change these later in this presentation

Another nice feature in JMP that allows data integrity checking is the Cols/Recode feature. Using this, the analyst gets a quick check of the number of items in each group and can quickly compare these with those in an original data sheet.

As an aside, you will often see large dots in the worksheet cells, and these merely indicate missing data. When zeroes are mistakenly imported and dots are desired, the change may be implemented in the List Check feature of the Column Properties dropdown found under Column Info (and accessed by right-clicking any column header).

Once the data is in JMP, and has been formatted to suit your needs, always remember to name and **save the table** with the <File/Save as> command. After the initial save, as you modify the data table, you may wish to use <File/Save> quite often.

Opening a data table window will allow the user to:

- change a column width
- hide/delete columns
- add/delete rows and rearrange their order
- add columns and rearrange their order
- select and subset rows for further analysis
- sort and combine tables

**The red, downward pointing arrow menu icons are a great help, click on them in the row/column selection area as below:**

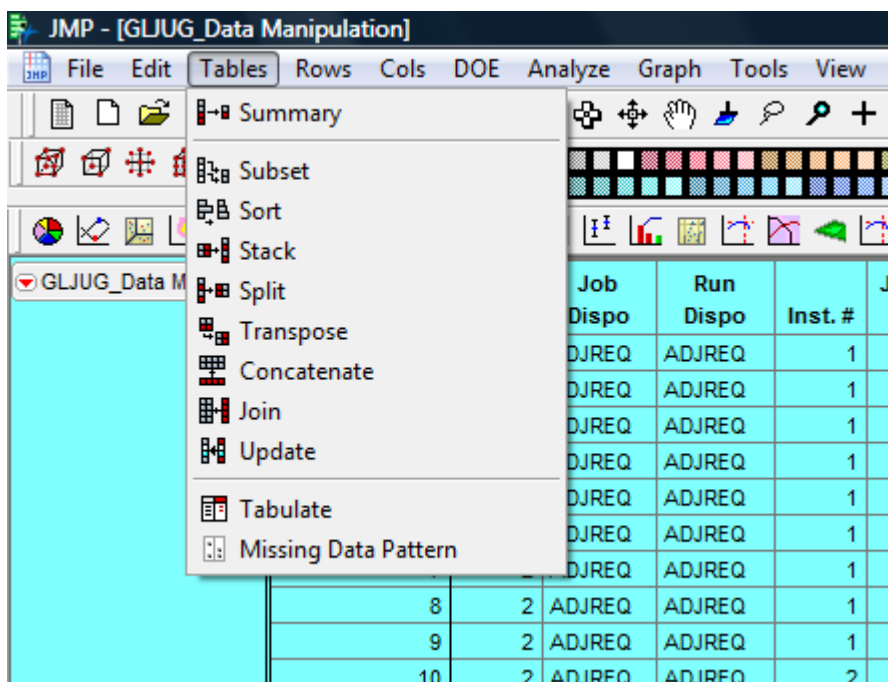
[illegible]

[illegible]

You will quickly find that these are the most useful tools imaginable for selecting and tagging data. In this course we will try several of these operations, alone and in combination to demonstrate their utility. You will find that they will assist in not only tagging and subsetting data, but in visualizing trends, as points in various groups can be labeled with many colors and shapes to make apparent any trends or aberrancies that might exist in the data. Also for merging with other data files and a quick look at any given subset of the data, the tools are invaluable. As with any other tool the secret is practice, practice, and more practice!

## Use of the Tables Menu: Overview

The 'Tables' menu selection allows the user to modify or create new tables from one or more existing JMP tables. It has the following commands:



There are 11 choices on the menu and perhaps the best way of familiarizing ourselves with the use and functions are by using them all on a single problem. Our example comes from the diagnostics industry.

The instrument development and QA scientists have a control dataset that they wish to do some preliminary analyses on, to examine the distribution for outliers, compare within and between run variability, and further explore for possible important relationships between subgroups of the data.

We have data for three controls from 2 lots, done on 3 different instruments over 2 days with 3 repetitions per run. By stepping through subsets of the data with various menu commands, we may gain greater insight into class differences. This strategy can be generalized to most data and we usually use subject matter of familiarity to the target audience. Please note that in the following narrative, anything in < > are the JMP menu items and choices.

Step 1: The dataset consists of 27 readings (as rows) and data for 3 controls (as columns) plus six descriptor columns:

Data Manipulation 1		Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	CTRL_H	CTRL_L	CTRL_M
	1	1	ADJREQ	ADJREQ	1	1	1	22.012	0.983	5.831
	2	1	ADJREQ	ADJREQ	1	1	2	23.131	1.007	6.204
	3	1	ADJREQ	ADJREQ	1	1	3	23.948	1	5.962
	4	1	ADJREQ	ADJREQ	1	1	1	21.275	0.982	5.38
	5	1	ADJREQ	ADJREQ	1	1	2	20.462	0.956	5.021
	6	1	ADJREQ	ADJREQ	1	1	3	21.878	0.888	5.015
	7	2	ADJREQ	ADJREQ	1	1	1	23.108	1.095	6.087
	8	2	ADJREQ	ADJREQ	1	1	2	22.256	0.974	5.709
	9	2	ADJREQ	ADJREQ	1	1	3	22.729	1.023	5.921
	10	2	ADJREQ	ADJREQ	2	1	1	21.45	0.956	5.627
	11	2	ADJREQ	ADJREQ	2	1	2	21.555	1.013	5.87
	12	2	ADJREQ	ADJREQ	2	1	3	23.182	0.978	5.808
	13	1	ADJREQ	ADJREQ	2	2	1	22.275	0.977	5.794
	14	1	ADJREQ	ADJREQ	2	2	2	21.618	0.916	5.712
	15	1	ADJREQ	ADJREQ	2	2	3	23.678	0.97	5.866
	16	1	ADJREQ	ADJREQ	2	2	1	21.586	0.897	5.805
	17	1	ADJREQ	ADJREQ	2	2	2	22.283	0.876	5.461
	18	1	ADJREQ	ADJREQ	2	2	3	21.648	0.889	5.409
	19	2	ADJREQ	ADJREQ	3	2	1	21.77	1.024	5.664
	20	2	ADJREQ	ADJREQ	3	2	2	21.23	1.011	5.516
	21	2	ADJREQ	ADJREQ	3	2	3	21.6	1.094	6.113
	22	2	ADJREQ	ADJREQ	3	2	1	23.07	0.995	6.291
	23	2	ADJREQ	ADJREQ	3	2	2	22.944	1.045	5.805
	24	2	ADJREQ	ADJREQ	3	2	3	22.839	1.047	5.785
	25	2	ADJREQ	ADJREQ	3	2	1	23.753	0.982	5.857
	26	2	ADJREQ	ADJREQ	3	2	2	22.364	1.017	5.519
	27	2	ADJREQ	ADJREQ	3	2	3	23.012	0.994	4.999
Columns (9/0)										
Date										
Job Dispo										
Run Dispo										
Inst. #										
Job Lot #										
Rep										
CTRL_H										
CTRL_L										
CTRL_M										
Rows										
All rows	27									
Selected	0									

**[End of First Part of WebEx]**

## Use of the Row Selection Commands

As a first pass at data exploration, the chemists want a summary of all of the data for each control level. (This is done under the Analyze/Distribution menu item). Upon examining the results, they observe the data clustering in unexpected ways and wish to examine a specific control and lot, or several control/lots separately. We do this by the following commands:

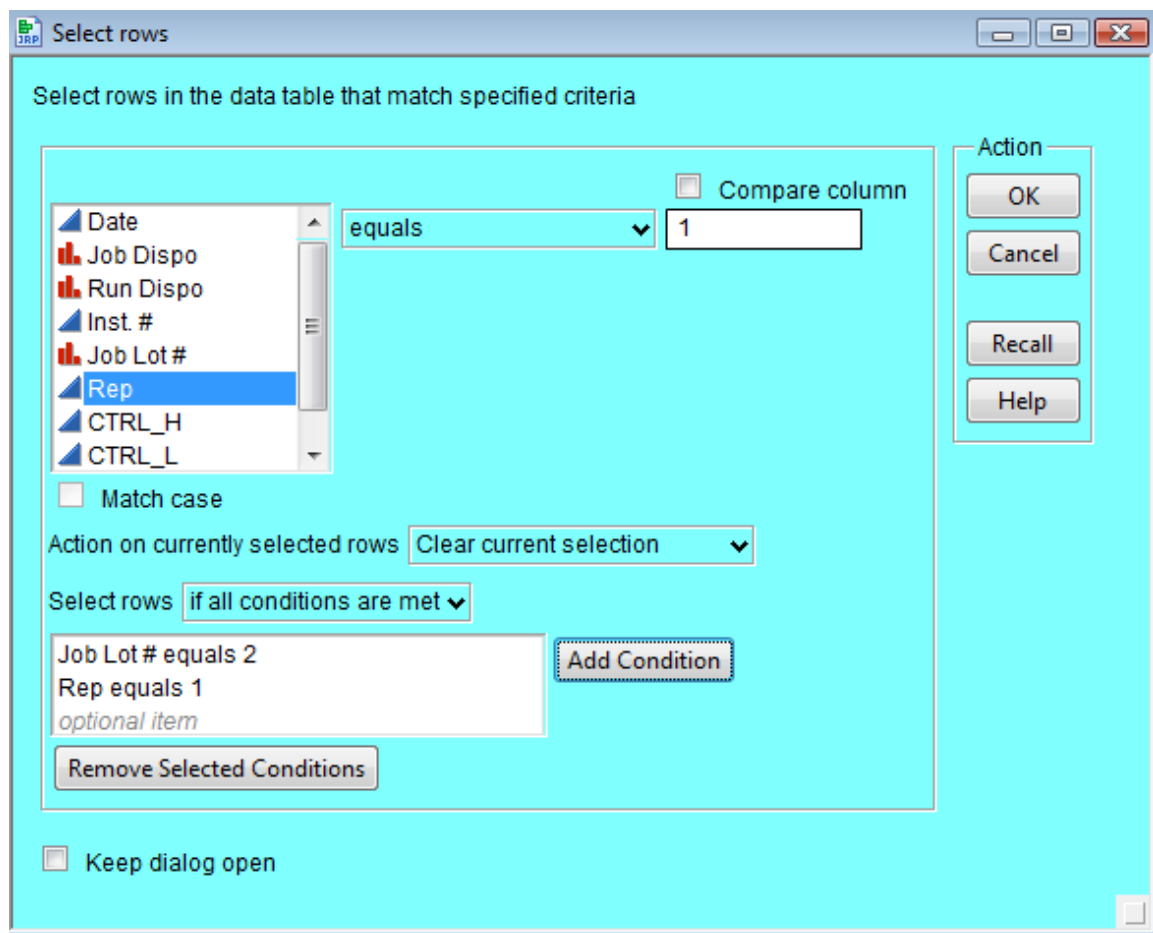


**Left-click on the red arrow above the row numbers column OR  
Right-click anywhere in the blank space (containing the red arrow) above the row  
numbers column**


From the resultant menu select <Row Selection/Select Where...>

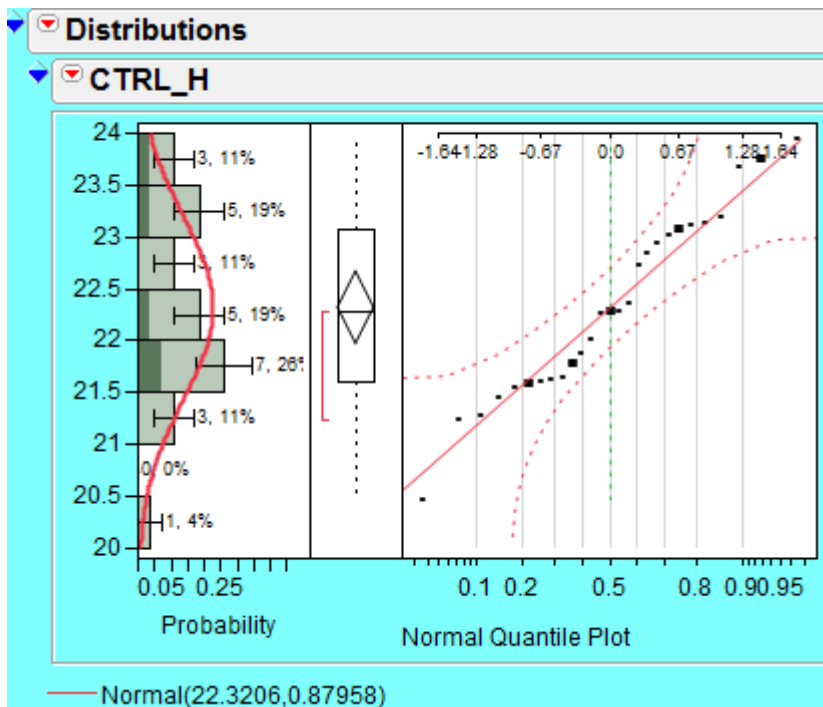
In the resultant dialog box: select 'Job Lot #' by clicking on it, select 'equals' from the drop down list (this is the default), type a '2' in the box (without the ' ') and click on <Add Condition>. Then select Rep, and type a '1' in the box (without the ' ') and click on <Add Condition>.

Your dialog box should look like the one below.



Now click <OK>

You will see that the rows on the data table for Job Lot #2, rep 1 have all been highlighted, and therefore are selected. Multiple selection of rows was a new feature added in version 6. By using the Distribution button on the toolbar , it is seen that the rows of interest are highlighted both on the histogram and quantile plot.





If the distribution of ONLY these data points were desired, they could be used by excluding all of the other points. This is done by:

Left-click on the red arrow above the row numbers column OR  
Right-click anywhere in the blank space (containing the red arrow) above the row numbers column

From the resultant menu select <Row Selection/Invert Row Selection>

Now all of the other points are selected and may be excluded from subsequent analysis by:

Left-click on the red arrow above the row numbers column OR Right-click anywhere in the blank space (containing the red arrow) above the row numbers column and select <Exclude/Unexclude> (an easier way to do this is with the Tables/Subset menu, and we will see its use shortly). Still another method of selection involves the Data Filter under the Rows menu.

All of the other points now have been labeled with a  and those data points will be excluded from any subsequent analysis. To remove the label, just go back to the same menu as before and select <Exclude/Unexclude>. The label is now removed and the data will be again included in subsequent calculations. Please note that if the intent was to merely observe a certain sub-group on any plot, while retaining all data for the statistical calculations, this could be done in the same manner, but using the <Hide/Unhide> choice from the Row menus. That symbol would appear as a mask .

These are particularly powerful features as they allow you to selectively utilize portions of the data without actually deleting any data.

If several labels (such as markers, colors, exclusions, etc.) have been used and you want to remove them all at once, merely

Left-click on the red arrow above the row numbers column OR

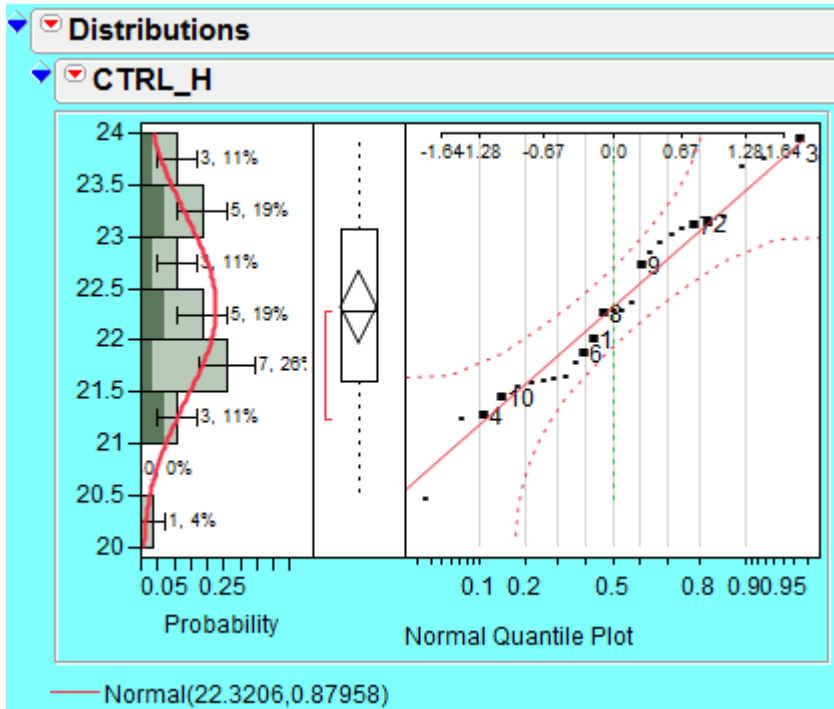
Right-click anywhere in the blank space (containing the red arrow) above the row numbers column and select <Clear Row States>.

Please note that while we employ labels to characterize the various groups, this is not the same as the Label function in JMP, which allows you to actually label points with the values in the data table. For example, if we labeled certain rows with the **Label** option, we would see:

Data Manipulation 1			Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	CTRL_ H	CTRL_ L	CTRL_ M
		1	1	ADJREQ	ADJREQ	1	1	1	22.012	0.983	5.831
		2	1	ADJREQ	ADJREQ	1	1	2	23.131	1.007	6.204
		3	1	ADJREQ	ADJREQ	1	1	3	23.948	1	5.962
		4	1	ADJREQ	ADJREQ	1	1	1	21.275	0.982	5.38
		5	1	ADJREQ	ADJREQ	1	1	2	20.462	0.956	5.021
		6	1	ADJREQ	ADJREQ	1	1	3	21.878	0.888	5.015
		7	2	ADJREQ	ADJREQ	1	1	1	23.108	1.095	6.087
		8	2	ADJREQ	ADJREQ	1	1	2	22.256	0.974	5.709
		9	2	ADJREQ	ADJREQ	1	1	3	22.729	1.023	5.921
		10	2	ADJREQ	ADJREQ	2	1	1	21.45	0.956	5.627
		11	2	ADJREQ	ADJREQ	2	1	2	21.555	1.013	5.87
		12	2	ADJREQ	ADJREQ	2	1	3	23.182	0.978	5.808
		13	1	ADJREQ	ADJREQ	2	2	1	22.275	0.977	5.794
		14	1	ADJREQ	ADJREQ	2	2	2	21.618	0.916	5.712

Note that the above is done by Left-clicking on the red arrow above the row numbers column OR Right-clicking anywhere in the blank space (containing the red arrow) above the row numbers column and select <Label/Unlabel>.

When we now do the <Analyze/Distribution> command we get:



Now the labeled points show up with the row numbers as labels. Alternatively we could use the more informative column labels so we could relate each point to a rep, run, or lot.

Now to see how easily data groups can be distinguished by these markers, let's select different ones for each rep.

Left-click on the red arrow above the row numbers column OR  
Right-click anywhere in the blank space (containing the red arrow) above the row numbers column


From the resultant menu select <Row Selection/Select Where...>  
And highlight Rep, select equals, and type in '1' then <OK>

Notice that all of the Rep1's have been selected. Now


Again, Left-click on the red arrow above the row numbers column OR  
Right-click anywhere in the blank space (containing the red arrow) above the row numbers column

Then click on 'Colors' and select a dark red.

Again, Left-click on the red arrow above the row numbers column OR  
Right-click anywhere in the blank space (containing the red arrow) above the row numbers column

Then click on 'Markers' and select the spider 

The spreadsheet now appears as below:


	Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	CTRL_ H	CTRL_ L	CTRL_ M	
✖	1	1	ADJREQ	ADJREQ	1	1	1	22.012	0.983	5.831
	2	1	ADJREQ	ADJREQ	1	1	2	23.131	1.007	6.204
	3	1	ADJREQ	ADJREQ	1	1	3	23.948	1	5.962
✖	4	1	ADJREQ	ADJREQ	1	1	1	21.275	0.982	5.38
	5	1	ADJREQ	ADJREQ	1	1	2	20.462	0.956	5.021
	6	1	ADJREQ	ADJREQ	1	1	3	21.878	0.888	5.015
✖	7	2	ADJREQ	ADJREQ	1	1	1	23.108	1.095	6.087
	8	2	ADJREQ	ADJREQ	1	1	2	22.256	0.974	5.709
	9	2	ADJREQ	ADJREQ	1	1	3	22.729	1.023	5.921
✖	10	2	ADJREQ	ADJREQ	2	1	1	21.45	0.956	5.627
	11	2	ADJREQ	ADJREQ	2	1	2	21.555	1.013	5.87
	12	2	ADJREQ	ADJREQ	2	1	3	23.182	0.978	5.808
✖	13	1	ADJREQ	ADJREQ	2	2	1	22.275	0.977	5.794
	14	1	ADJREQ	ADJREQ	2	2	2	21.618	0.916	5.712
	15	1	ADJREQ	ADJREQ	2	2	3	23.678	0.97	5.866
✖	16	1	ADJREQ	ADJREQ	2	2	1	21.586	0.897	5.805
	17	1	ADJREQ	ADJREQ	2	2	2	22.283	0.876	5.461
	18	1	ADJREQ	ADJREQ	2	2	3	21.648	0.889	5.409
✖	19	2	ADJREQ	ADJREQ	3	2	1	21.77	1.024	5.664
	20	2	ADJREQ	ADJREQ	3	2	2	21.23	1.011	5.516
	21	2	ADJREQ	ADJREQ	3	2	3	21.6	1.094	6.113
✖	22	2	ADJREQ	ADJREQ	3	2	1	23.07	0.995	6.291
	23	2	ADJREQ	ADJREQ	3	2	2	22.944	1.045	5.805
	24	2	ADJREQ	ADJREQ	3	2	3	22.839	1.047	5.785
✖	25	2	ADJREQ	ADJREQ	3	2	1	23.753	0.982	5.857
	26	2	ADJREQ	ADJREQ	3	2	2	22.364	1.017	5.519
	27	2	ADJREQ	ADJREQ	3	2	3	23.012	0.994	4.999

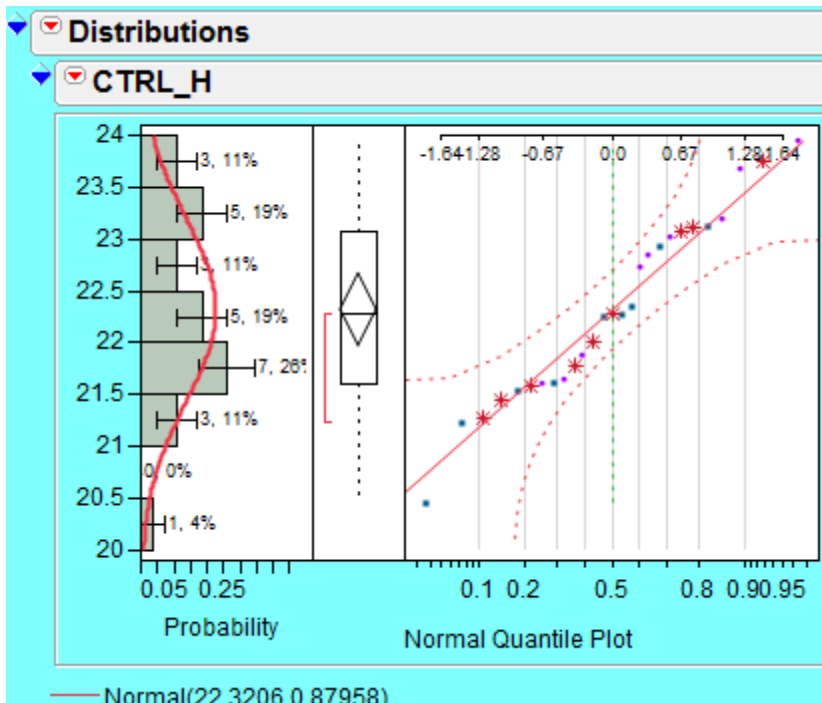
Notice that by left-clicking in the small region above the row numbers, you can remove the selection highlighting.

		Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	CTRL_ H	CTRL_ L	CTRL_ M
*	1	1	ADJREQ	ADJREQ	1	1	1	22.012	0.983	5.831
	2	1	ADJREQ	ADJREQ	1	1	2	23.131	1.007	6.204
	3	1	ADJREQ	ADJREQ	1	1	3	23.948	1	5.962
*	4	1	ADJREQ	ADJREQ	1	1	1	21.275	0.982	5.38
	5	1	ADJREQ	ADJREQ	1	1	2	20.462	0.956	5.021
	6	1	ADJREQ	ADJREQ	1	1	3	21.878	0.888	5.015
*	7	2	ADJREQ	ADJREQ	1	1	1	23.108	1.095	6.087
	8	2	ADJREQ	ADJREQ	1	1	2	22.256	0.974	5.709
	9	2	ADJREQ	ADJREQ	1	1	3	22.729	1.023	5.921
*	10	2	ADJREQ	ADJREQ	2	1	1	21.45	0.956	5.627
	11	2	ADJREQ	ADJREQ	2	1	2	21.555	1.013	5.87
	12	2	ADJREQ	ADJREQ	2	1	3	23.182	0.978	5.808
*	13	1	ADJREQ	ADJREQ	2	2	1	22.275	0.977	5.794
	14	1	ADJREQ	ADJREQ	2	2	2	21.618	0.916	5.712
	15	1	ADJREQ	ADJREQ	2	2	3	23.678	0.97	5.866
*	16	1	ADJREQ	ADJREQ	2	2	1	21.586	0.897	5.805
	17	1	ADJREQ	ADJREQ	2	2	2	22.283	0.876	5.461
	18	1	ADJREQ	ADJREQ	2	2	3	21.648	0.889	5.409
*	19	2	ADJREQ	ADJREQ	3	2	1	21.77	1.024	5.664
	20	2	ADJREQ	ADJREQ	3	2	2	21.23	1.011	5.516
	21	2	ADJREQ	ADJREQ	3	2	3	21.6	1.094	6.113
*	22	2	ADJREQ	ADJREQ	3	2	1	23.07	0.995	6.291
	23	2	ADJREQ	ADJREQ	3	2	2	22.944	1.045	5.805
	24	2	ADJREQ	ADJREQ	3	2	3	22.839	1.047	5.785
*	25	2	ADJREQ	ADJREQ	3	2	1	23.753	0.982	5.857
	26	2	ADJREQ	ADJREQ	3	2	2	22.364	1.017	5.519
	27	2	ADJREQ	ADJREQ	3	2	3	23.012	0.994	4.999

Now repeat the process above for all Rep 2's and 3's choosing new colors and markers.  
The table then appears as:

		Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	CTRL_ H	CTRL_ L	CTRL_ M
✖	1	1	ADJREQ	ADJREQ	1	1	1	22.012	0.983	5.831
■	2	1	ADJREQ	ADJREQ	1	1	2	23.131	1.007	6.204
●	3	1	ADJREQ	ADJREQ	1	1	3	23.948	1	5.962
✖	4	1	ADJREQ	ADJREQ	1	1	1	21.275	0.982	5.38
■	5	1	ADJREQ	ADJREQ	1	1	2	20.462	0.956	5.021
●	6	1	ADJREQ	ADJREQ	1	1	3	21.878	0.888	5.015
✖	7	2	ADJREQ	ADJREQ	1	1	1	23.108	1.095	6.087
■	8	2	ADJREQ	ADJREQ	1	1	2	22.256	0.974	5.709
●	9	2	ADJREQ	ADJREQ	1	1	3	22.729	1.023	5.921
✖	10	2	ADJREQ	ADJREQ	2	1	1	21.45	0.956	5.627
■	11	2	ADJREQ	ADJREQ	2	1	2	21.555	1.013	5.87
●	12	2	ADJREQ	ADJREQ	2	1	3	23.182	0.978	5.808
✖	13	1	ADJREQ	ADJREQ	2	2	1	22.275	0.977	5.794
■	14	1	ADJREQ	ADJREQ	2	2	2	21.618	0.916	5.712
●	15	1	ADJREQ	ADJREQ	2	2	3	23.678	0.97	5.866
✖	16	1	ADJREQ	ADJREQ	2	2	1	21.586	0.897	5.805
■	17	1	ADJREQ	ADJREQ	2	2	2	22.283	0.876	5.461
●	18	1	ADJREQ	ADJREQ	2	2	3	21.648	0.889	5.409
✖	19	2	ADJREQ	ADJREQ	3	2	1	21.77	1.024	5.664
■	20	2	ADJREQ	ADJREQ	3	2	2	21.23	1.011	5.516
●	21	2	ADJREQ	ADJREQ	3	2	3	21.6	1.094	6.113
✖	22	2	ADJREQ	ADJREQ	3	2	1	23.07	0.995	6.291
■	23	2	ADJREQ	ADJREQ	3	2	2	22.944	1.045	5.805
●	24	2	ADJREQ	ADJREQ	3	2	3	22.839	1.047	5.785
✖	25	2	ADJREQ	ADJREQ	3	2	1	23.753	0.982	5.857
■	26	2	ADJREQ	ADJREQ	3	2	2	22.364	1.017	5.519
●	27	2	ADJREQ	ADJREQ	3	2	3	23.012	0.994	4.999

Now by merely selecting <Analyze/Distribution> on the menu bar or just the distribution button on the toolbar ,



We can readily visualize/verify that our reps are evenly distributed throughout the range. Also note the following column feature:

We can mark any column as a point label on the graphics by left-clicking on a column name to highlight it, then left-clicking the red arrow in the column area to the left of the column names and selecting 'Label/Unlabel'. A small tag will appear next to the column name and on any graphic; the value of the point label (say rep or lot number) will appear by any point with a simple mouse-over.

Please note that while we had originally saved the data table, we have NOT saved any analysis! To save the analysis use the Edit/Journal then 'File/Save as' menu choices.

**[End of Second Part of WebEx]**

## Use of the Tables Selections for Data Manipulation

Suppose that we would like to quickly **summarize the data** by groups. This could be done by selecting only those rows in a given group and using the distribution platform. This would give much more output than we would like to see so a simple summary might suffice. To do a quick summary:

<Tables/Summary>

Highlight columns of interest (do NOT hit the group button during these operations)

Choose the statistics of interest by left-clicking on the Statistics box AND WHILE HOLDING DOWN THE LEFT MOUSE BUTTON, select each statistic of interest (only one at a time) and releasing the button between selections.



Request Summary Statistics by Grouping Columns.

Select Columns

- Date
- Job Dispo
- Run Dispo
- Inst. #
- Job Lot #
- Rep
- Low Control
- Mid Control
- High Control

Statistics

- N
- Mean
- Std Dev
- Min
- Max
- Range
- % of Total
- N Missing
- Sum
- Sum Wgt
- Variance
- Std Err
- CV
- Median
- Quantiles

Action

OK

Cancel

Remove

Recall

Help

☐ Include marginal statistics

For quantile statistics, enter value (%)

25

statistics column name format

stat(column)

Output table name:

This is rather inconvenient as the summary statistics are strung out to the right in a single row.

	N Rows	Mean(CTRL_H)	Std Dev(CTRL_H)	Variance(CTRL_H)	Std Err(CTRL_H)
1	27	22.3205926	0.87958099	0.77366271	0.16927544

Of course we could always **transpose the data** <Tables/Transpose> , select all of the columns and hit the <Transpose Columns> to get a better format:

	Label	Row 1
1	N Rows	27
2	Mean(CTRL_H)	22.3205926
3	Std Dev(CTRL_H)	0.87958099
4	Variance(CTRL_H)	0.77366271
5	Std Err(CTRL_H)	0.16927544

Notice what the Transpose does; we are now interchanging our columns and rows so our data or headers in the columns from left to right, are now in the rows from top to bottom.

Now that we have had a little practice with Tables, let's try the very powerful **sort** command.

Many times we wish to make group comparisons, and in JMP that means → stacked data (one data column and one column for each group descriptor such as day, instrument, rep, run, etc.).

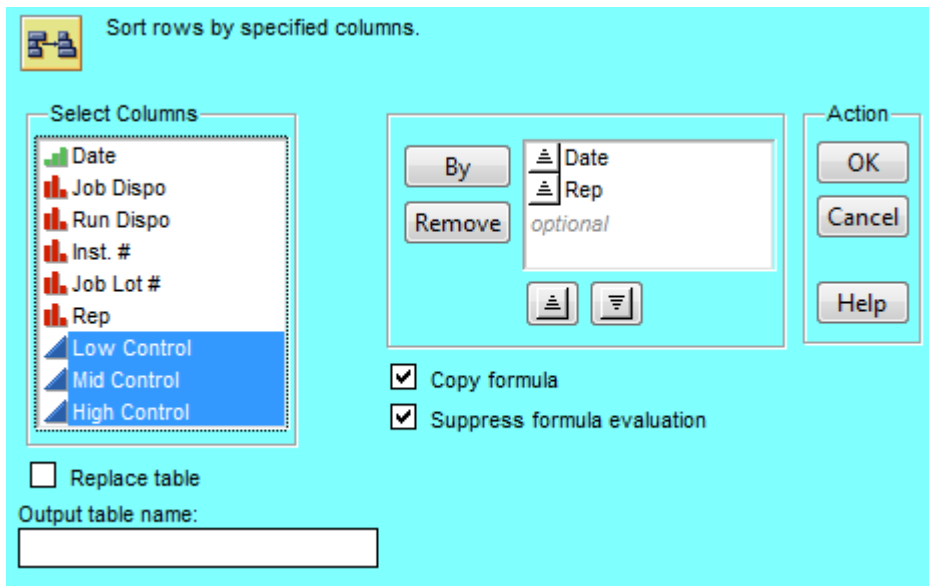
We are used to having our data in separate columns so this mandates a rearrangement. In JMP we may implement multiple “By” statements automatically and sort by descriptor1, then 2 within 1, then 3 within 2, etc. In our data set we will invoke

<Tables/sort>

Select Date and hit <By>






















Select Rep and hit <By>

Highlight the 3 controls



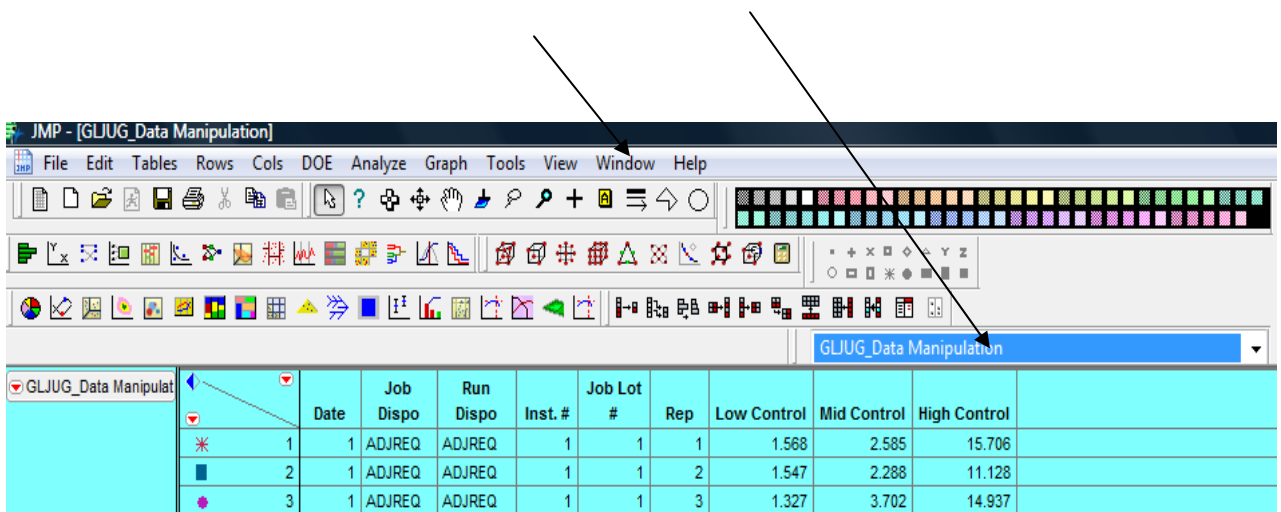
Notice that I have changed the modeling type for Date to Ordinal and Job Dispo, Run Dispo, Inst #, Job Lot #, and Rep to Nominal. This is not so much for the data manipulation steps but for any successive analysis that is done. We will briefly touch on modeling types at the end of this session and in future presentations as we actually analyze the data.


Now hit <OK> and observe the results:

 		Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control
	1	1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
	2	1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
	3	1	ADJREQ	ADJREQ	2	2	1	1.48	4.913	16.918
	4	1	ADJREQ	ADJREQ	2	2	1	1.171	4.606	10.986
	5	1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
	6	1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
	7	1	ADJREQ	ADJREQ	2	2	2	1.645	4.011	11.235
	8	1	ADJREQ	ADJREQ	2	2	2	1.415	3.635	12.373
	9	1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
	10	1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
	11	1	ADJREQ	ADJREQ	2	2	3	1.391	5.479	16.020
	12	1	ADJREQ	ADJREQ	2	2	3	1.95	4.592	15.635
	13	2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
	14	2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
	15	2	ADJREQ	ADJREQ	3	2	1	1.695	3.213	17.953
	16	2	ADJREQ	ADJREQ	3	2	1	1.022	4.354	18.583
	17	2	ADJREQ	ADJREQ	3	2	1	1.673	3.301	15.630
	18	2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
	19	2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035

Now we have the control data sorted by rep within date (reps 1, 2, and 3 within each date) and can easily utilize the 'Fit Y by X' platform to compare reps within days, or any other grouping for which we have pre-sorted.

At this point you may see several windows tiled on your screen. Some may be completely hidden behind others. We can rapidly **toggle between windows** with the Window main menu choice, or the dropdown Windows box.



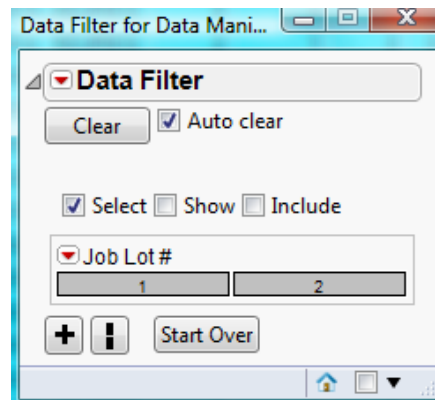
Note that another easy way of isolating data for further analysis is with the **Subset** key  used in conjunction with row selection.

## Review:

**Left-click on the red arrow above the row numbers column OR  
Right-click anywhere in the blank space (containing the red arrow) above the row numbers column**

**From the resultant menu select <Row Selection/Select Where...>**

In the resultant dialog box: select 'Job Lot #' by clicking on it, select equals from the drop down list (this is the default), type a '1' in the box (without the ') and click on <Add Condition, <OK> . There is also a more powerful feature to do row selections and that is the Data Filter (<Rows/Data Filter>). For non-continuous data this is very fast and easy:












Now that we have selected only Lot 1 data, subset it:

<Tables/Subset>

Put a "." in the 'Selected Rows' radio button

Hit <OK>

We now have a table of only Lot 1 data that can be analyzed by any means that we choose.

		Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control
	1	1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
	2	1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
	3	1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
	4	1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
	5	1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
	6	1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
	7	2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
	8	2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
	9	2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288
	10	2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
	11	2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035
	12	2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193

Before we leave the simple manipulations, let's take a look at the fast and easy way to **stack** the data for Y by X analysis. Starting with our basic data table, and using commands that we have already learned, we will isolate our needed columns, and then stack them for subsequent analysis.

Start by selecting the last 6 columns (instrument, lot, rep, and the 3 controls)  
<Tables/Subset> ('all rows' and 'selected columns' buttons should be filled) <OK>

In the new table:

	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control
✱	1	1	1	1.568	2.585	15.706
■	2	1	1	2	1.547	2.288
✱	3	1	1	3	1.327	3.702
✱	4	1	1	1	2.198	6.241
■	5	1	1	2	2.016	5.615
✱	6	1	1	3	0.997	4.711
✱	7	1	1	1	1.259	4.605
■	8	1	1	2	0.671	4.039
✱	9	1	1	3	1.929	3.533
✱	10	2	1	1	1.343	2.639
■	11	2	1	2	1.076	5.738
✱	12	2	1	3	1.679	4.715


<Tables/Stack> select the 3 controls <Stack Columns>  
**remove the check mark from the 'Stack by row' box**  
<OK>

The controls are now stacked into a single column, labeled as to instrument, Lot, and rep, and are ready for Y by X analyses.






●	9	1	1	3	Low Control	1.9291
✱	10	2	1	1	Low Control	1.3429
■	11	2	1	2	Low Control	1.0756
●	12	2	1	3	Low Control	1.6787
✱	13	1	1	1	Mid Control	2.5853
■	14	1	1	2	Mid Control	2.2885
●	15	1	1	3	Mid Control	3.702
✱	16	1	1	1	Mid Control	6.2414
■	17	1	1	2	Mid Control	5.6148
●	18	1	1	3	Mid Control	4.711
✱	19	1	1	1	Mid Control	4.6051
■	20	1	1	2	Mid Control	4.0385
●	21	1	1	3	Mid Control	3.5326
✱	22	2	1	1	Mid Control	2.6394
■	23	2	1	2	Mid Control	5.7378
●	24	2	1	3	Mid Control	4.715
✱	25	1	1	1	High Control	15.706
■	26	1	1	2	High Control	11.128
●	27	1	1	3	High Control	14.937
✱	28	1	1	1	High Control	11.888
■	29	1	1	2	High Control	13.229
●	30	1	1	3	High Control	10.606
✱	31	1	1	1	High Control	12.301
■	32	1	1	2	High Control	16.687
●	33	1	1	3	High Control	15.288

Occasionally the analyst may have a need to deal with the data in its original format, and in JMP it's a relatively easy matter to **split** data columns into logical groups.

<Tables/Split>

 Split unstacks multiple rows for each 'Split Column' into multiple columns as identified by a 'Split By' column.

Select Columns

-  Inst. #
-  Job Lot #
-  Rep
-  Label
-  Data

Remaining columns

☐ Keep All

☒ Drop All

☐ Select

Split By:

Split Columns:

Group:

Action

Output table name:

Select the column titled 'Label' then hit <Split By>

Select the Data column and hit <Split Columns>

If these are the only columns of interest, put a '.' in the Drop All radio button

<OK>

The column have been split back to the original multiple columns:

	High Control	Low Control	Mid Control
1	15.706	1.5677	2.5853
2	11.128	1.5467	2.2885
3	14.937	1.3273	3.702
4	11.888	2.198	6.2414
5	13.229	2.0165	5.6148
6	10.606	0.9967	4.711
7	12.301	1.2594	4.6051
8	16.687	0.6715	4.0385
9	15.288	1.9291	3.5326
10	10.287	1.3429	2.6394
11	13.035	1.0756	5.7378
12	14.193	1.6787	4.715

Occasionally we may wish to get around Excel's 64K row limitation. New data can easily be added to existing sets with the **Concatenate** feature. Assuming that the new set is in the same format (and now we know how to quickly reformat data in JMP!) this can be done by saving both data tables in JMP and merely using the Concatenate command.

Table 1 (the original data table)

	Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control
1	1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
2	1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
3	1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
4	1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
5	1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
6	1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
7	2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
8	2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
9	2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288
10	2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
11	2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035
12	2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193
13	1	ADJREQ	ADJREQ	2	2	1	1.48	4.913	16.918
14	1	ADJREQ	ADJREQ	2	2	2	1.645	4.011	11.235
15	1	ADJREQ	ADJREQ	2	2	3	1.391	5.479	16.020
16	1	ADJREQ	ADJREQ	2	2	1	1.171	4.606	10.986
17	1	ADJREQ	ADJREQ	2	2	2	1.415	3.635	12.373
18	1	ADJREQ	ADJREQ	2	2	3	1.95	4.592	15.635
19	2	ADJREQ	ADJREQ	3	2	1	1.695	3.213	17.953
20	2	ADJREQ	ADJREQ	3	2	2	2.247	3.944	14.551
21	2	ADJREQ	ADJREQ	3	2	3	1.292	4.648	11.484
22	2	ADJREQ	ADJREQ	3	2	1	1.022	4.354	18.583
23	2	ADJREQ	ADJREQ	3	2	2	1.53	5.808	14.963
24	2	ADJREQ	ADJREQ	3	2	3	1.992	4.281	13.176
25	2	ADJREQ	ADJREQ	3	2	1	1.673	3.301	15.630
26	2	ADJREQ	ADJREQ	3	2	2	1.311	6.249	16.458
27	2	ADJREQ	ADJREQ	3	2	3	1.795	4.562	18.147

Table 2 (new data)

Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control
1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288
2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035
2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193

<Tables/Concatenate>



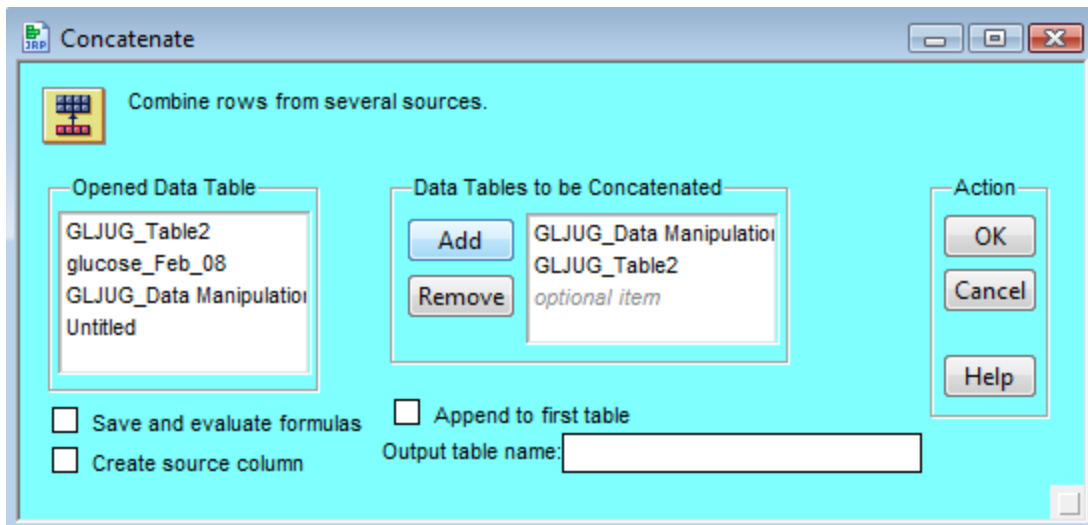


Table 1 appears in the 'Data Tables to be Concatenated' box. Select Table 2 and hit <ADD>, <OK> to get the concatenated table:

Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control
1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288
2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035
2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193
1	ADJREQ	ADJREQ	2	2	1	1.48	4.913	16.918
1	ADJREQ	ADJREQ	2	2	2	1.645	4.011	11.235
1	ADJREQ	ADJREQ	2	2	3	1.391	5.479	16.020
1	ADJREQ	ADJREQ	2	2	1	1.171	4.606	10.986
1	ADJREQ	ADJREQ	2	2	2	1.415	3.635	12.373
1	ADJREQ	ADJREQ	2	2	3	1.95	4.592	15.635
2	ADJREQ	ADJREQ	3	2	1	1.695	3.213	17.953
2	ADJREQ	ADJREQ	3	2	2	2.247	3.944	14.551
2	ADJREQ	ADJREQ	3	2	3	1.292	4.648	11.484
2	ADJREQ	ADJREQ	3	2	1	1.022	4.354	18.583
2	ADJREQ	ADJREQ	3	2	2	1.53	5.808	14.963
2	ADJREQ	ADJREQ	3	2	3	1.992	4.281	13.176
2	ADJREQ	ADJREQ	3	2	1	1.673	3.301	15.630
2	ADJREQ	ADJREQ	3	2	2	1.311	6.249	16.458
2	ADJREQ	ADJREQ	3	2	3	1.795	4.562	18.147
1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706
1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128
1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937
1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888
1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229
1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606
2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301
2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687
2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288
2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287
2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035
2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193

If we would ever want to add another type of data to the table, say a new column with data on a different control, where all of the other collection factors were the same, we can **Join** the new data to the old table.

New Data:

	Very High Control
1	22.606
2	34.852
3	30.991
4	28.276
5	34.65
6	30.48
7	37.5
8	35.698
9	29.485
10	26.741
11	31.157
12	29.043
13	30.665
14	29.475
15	34.399
16	35.234
17	38.95
18	26.58
19	36.441
20	20.182
21	23.053
22	35.817
23	34.421
24	33.053
25	23.102
26	22.869
27	29.013

Now use <Tables/Join> and select 'Table3' in the 'Join Table 1 with' box. Be sure that 'By Row Number' is selected as the Matching Specification. <OK>

Join rows from several sources by matching value.

Join 'GLJUG\_Data Manipulation' with

- GLJUG\_Table2
- glucose\_Feb\_08
- GLJUG\_Data Manipulation
- Untitled
- GLJUG\_Table3**
- Untitled 3

Options

- ☐ Update first table with data from second table
- ☐ Merge same name columns
- ☒ Copy formula
- ☒ Suppress formula evaluation

Matching Specification

By Row Number

Output Columns

- ☐ Select columns for joined table

Action

OK

Cancel

Remove

Recall

Help

Output table name:

Source Columns

GLJUG\_Data Manipulation

- Date
- Job Dispo
- Run Dispo
- Inst. #

GLJUG\_Table3

- Very High Control
- optional

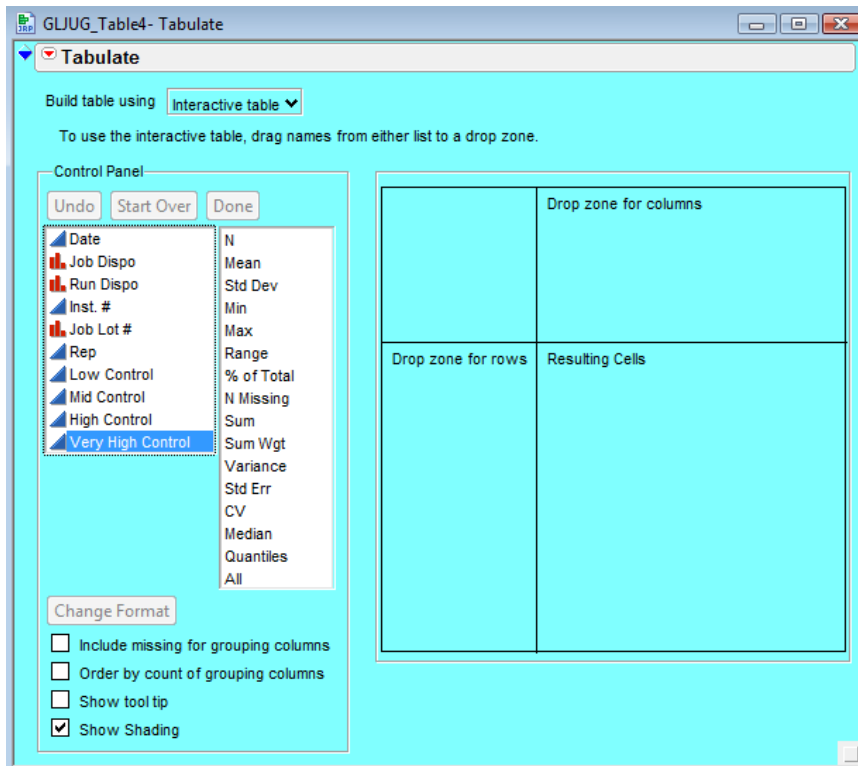
The resulting table now has a new column with the additional control:

	Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control	Very High Control
1	1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706	22.606
2	1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128	34.852
3	1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937	30.991
4	1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888	28.276
5	1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229	34.65
6	1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606	30.48
7	2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301	37.5
8	2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687	35.698
9	2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288	29.485
10	2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287	26.741
11	2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035	31.157
12	2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193	29.043
13	1	ADJREQ	ADJREQ	2	2	1	1.48	4.913	16.918	30.665
14	1	ADJREQ	ADJREQ	2	2	2	1.645	4.011	11.235	29.475
15	1	ADJREQ	ADJREQ	2	2	3	1.391	5.479	16.020	34.399
16	1	ADJREQ	ADJREQ	2	2	1	1.171	4.606	10.986	35.234
17	1	ADJREQ	ADJREQ	2	2	2	1.415	3.635	12.373	38.95
18	1	ADJREQ	ADJREQ	2	2	3	1.95	4.592	15.635	26.58
19	2	ADJREQ	ADJREQ	3	2	1	1.695	3.213	17.953	36.441
20	2	ADJREQ	ADJREQ	3	2	2	2.247	3.944	14.551	20.182
21	2	ADJREQ	ADJREQ	3	2	3	1.292	4.648	11.484	23.053
22	2	ADJREQ	ADJREQ	3	2	1	1.022	4.354	18.583	35.817
23	2	ADJREQ	ADJREQ	3	2	2	1.53	5.808	14.963	34.421
24	2	ADJREQ	ADJREQ	3	2	3	1.992	4.281	13.176	33.053
25	2	ADJREQ	ADJREQ	3	2	1	1.673	3.301	15.630	23.102
26	2	ADJREQ	ADJREQ	3	2	2	1.311	6.249	16.458	22.869
27	2	ADJREQ	ADJREQ	3	2	3	1.795	4.562	18.147	29.013

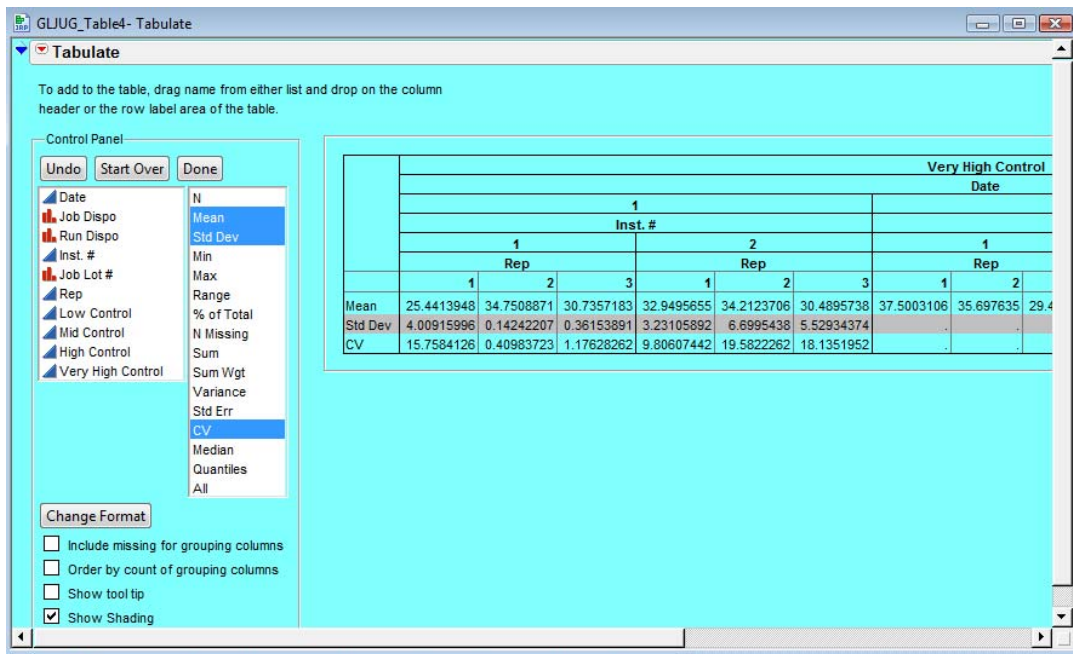
The **Update** function will actually change the data in your tables by replacing the old data with new, updated data. This would be useful for replacing older data in a time series collection, but as we rarely discard our data, it would be best to save the original data table, change the file name, and then replace the data under a new file name. In practice, we would prefer to just add, rather than replace data, so we would find more utility in the Join and Concatenate functions.

We do have some very useful additions that will quickly group and build compact data tables for immediate visualization of complex data structures. For example, if we wanted to show how the values of our High control varied with Rep, Lot, and Date, we could easily do so.

<Tables/**Tabulate**> Now to construct the new table, we start by selecting the lowest factor, in this case the actual data, dragging it from the control panel are to the drop zone, and letting JMP know whether we want it used for grouping or analysis. We then select each higher order nest of the design and add it in sequence to the area of the table just over the last factor. A gray area appears to let us know where this factor name will be dropped.



Repetitive series of these actions with higher order grouping terms will eventually lead to the completed table. (Note: there is also a 'Dialog' choice under the 'Build Table Using' area and some people find this method easier than using the 'Interactive Table').



Other statistics may be added to the bottom row by the simple drag-and-drop technique.

New in JMP 9 is the Format button that allows truncation of the table by limiting the field width and number of decimals displayed, a very useful feature!

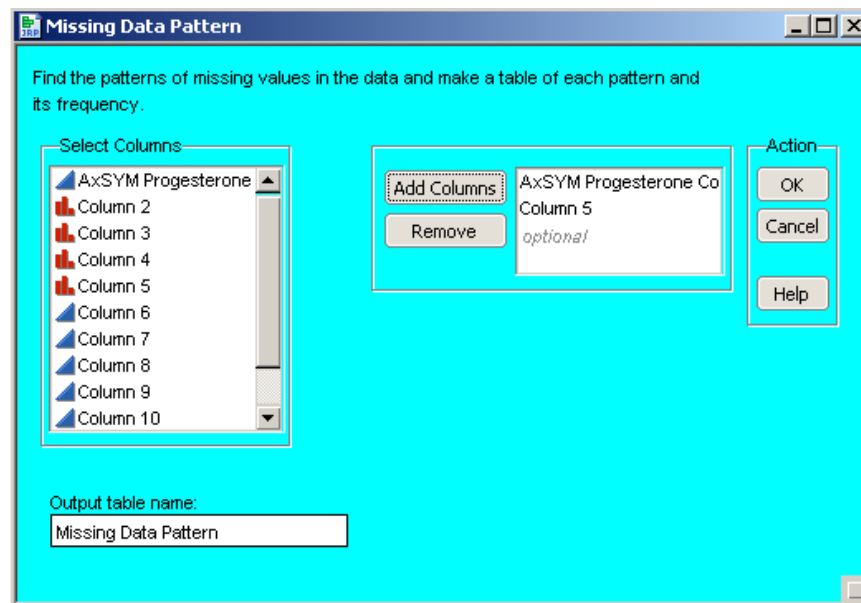
The final table may be saved with the Edit/Journal menu item and the output pasted into other documents.

JMP also has a new feature that will tally the amount of missing data, column by column.

If we have a table that looks like this:

	Date	Job Dispo	Run Dispo	Inst. #	Job Lot #	Rep	Low Control	Mid Control	High Control	Very High Control
1	1	ADJREQ	ADJREQ	1	1	1	1.568	2.585	15.706	22.606
2	1	ADJREQ	ADJREQ	1	1	2	1.547	2.288	11.128	34.852
3	1	ADJREQ	ADJREQ	1	1	3	1.327	3.702	14.937	30.991
4	.	.	.	.	.	.	.	.	.	.
5	1	ADJREQ	ADJREQ	1	1	1	2.198	6.241	11.888	28.276
6	1	ADJREQ	ADJREQ	1	1	2	2.016	5.615	13.229	34.65
7	1	ADJREQ	ADJREQ	1	1	3	0.997	4.711	10.606	30.48
8	.	.	.	.	.	.	.	.	.	.
9	2	ADJREQ	ADJREQ	1	1	1	1.259	4.605	12.301	37.5
10	2	ADJREQ	ADJREQ	1	1	2	0.671	4.039	16.687	35.698
11	2	ADJREQ	ADJREQ	1	1	3	1.929	3.533	15.288	29.485
12	2	ADJREQ	ADJREQ	2	1	1	1.343	2.639	10.287	26.741
13	.	.	.	.	.	.	.	.	.	.
14	.	.	.	.	.	.	.	.	.	.
15	.	.	.	.	.	.	.	.	.	.
16	2	ADJREQ	ADJREQ	2	1	2	1.076	5.738	13.035	31.157
17	2	ADJREQ	ADJREQ	2	1	3	1.679	4.715	14.193	29.043

but contains several thousand rows, we would like the software to quickly find and summarize the missing data and tell us in which columns this occurs. This is done with the <Tables/Missing Data Pattern> menu item.



Now highlight the columns of interest in the 'Select Columns' area and

Hit <Add Columns> <OK>

The tally appears as:

		Count	Number of columns missing	Patterns	Mid Control	High Control
1		12	0	00	0	0
2		5	2	11	1	1

This tells us how many rows have 0, 1, or both chosen columns containing missing data. Note that this is not really a pattern in the broadest sense as we don't know where the missing data occurs, just in which columns.

Finally please note JMP is not EXCEL but will offer nearly the complete menu of 'slice and dice' techniques offered by the spreadsheet. However, users will quickly find the use of such functions as sort and subset to be extremely useful when analyzing data. The others assist in a variety of conditions. Choose wisely!