使用Prism-7绘制统计图

Alex / 2016-10-15 / free_learner@163.com / learning-archive.org

更新于2023-05-01，主要是文字排版上的更新，内容基本保持不变。

一、背景

前一段时间，简单学习了一下使用Prism画统计图，包括相关、单因素方差分析和双因素方差分析。

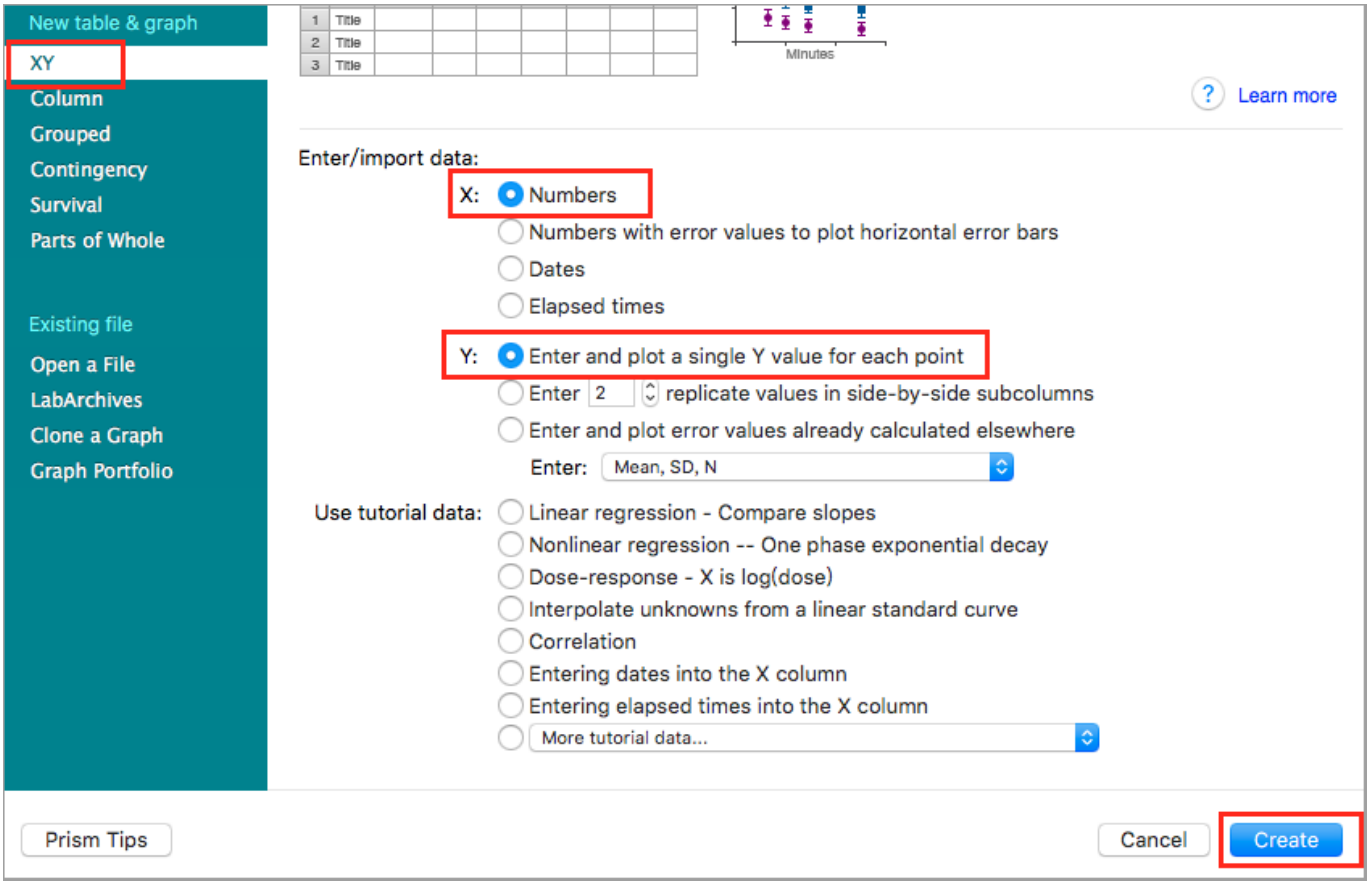
2023-05-01 当初学习Prism的原因是觉得可视化效果还不错，现在都是用R画图，R更方便、更强大、不花钱。

二、下载并安装Prism 7

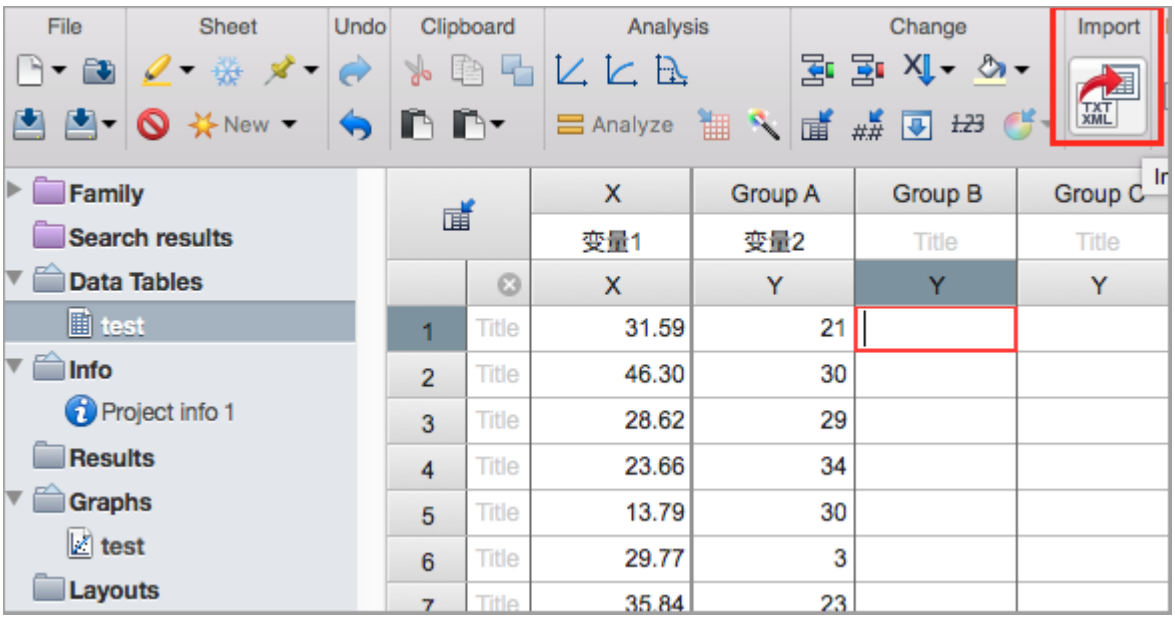
Prism是收费的软件，试用期30天。

三、两个变量间的相关关系

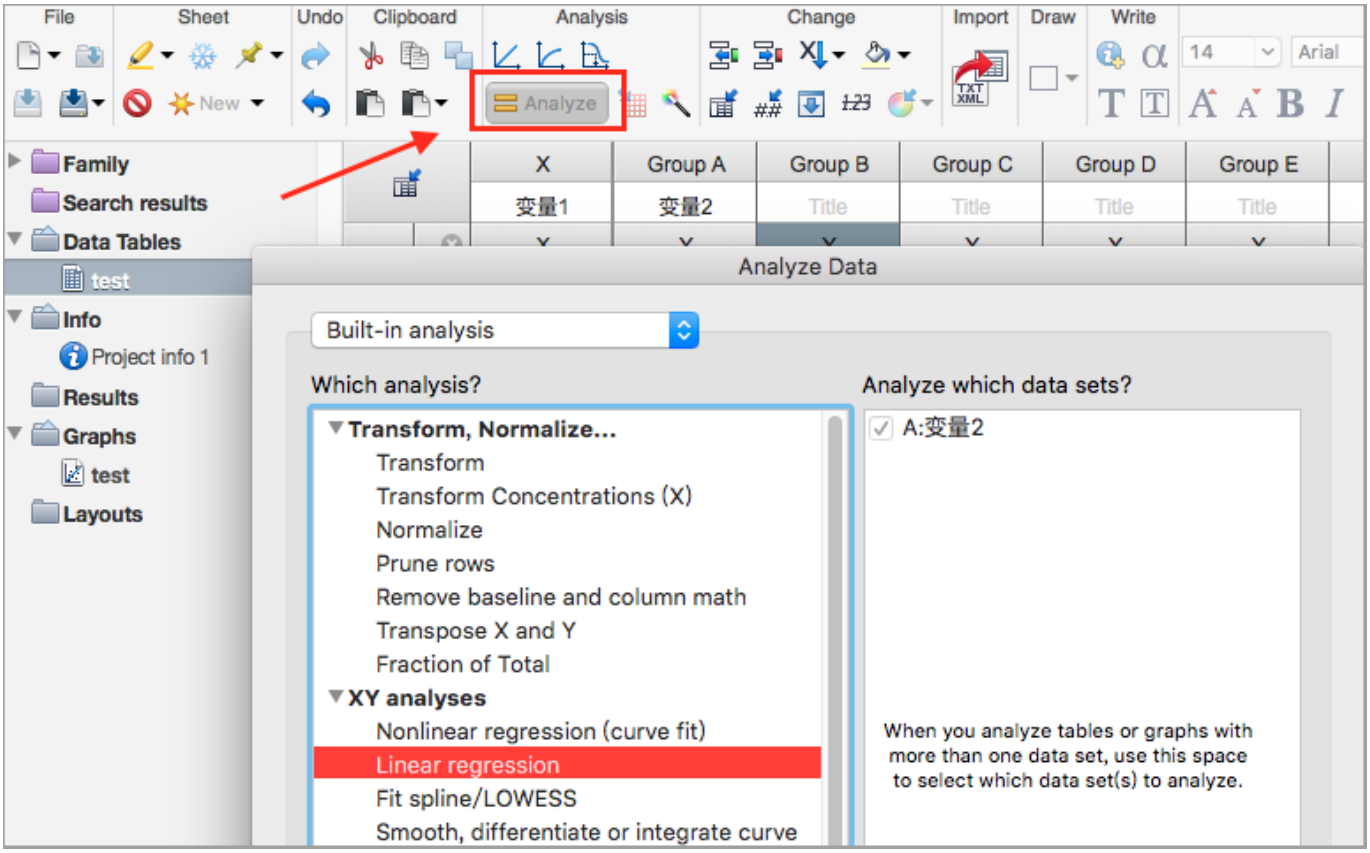
1. 打开Prism 7，左侧有六种数据组织形式可供选择，选择XY；右侧选择数据的类型，变量X选择Numbers，变量Y选择Enter and plot a single Y value for each point；最后点击Create。



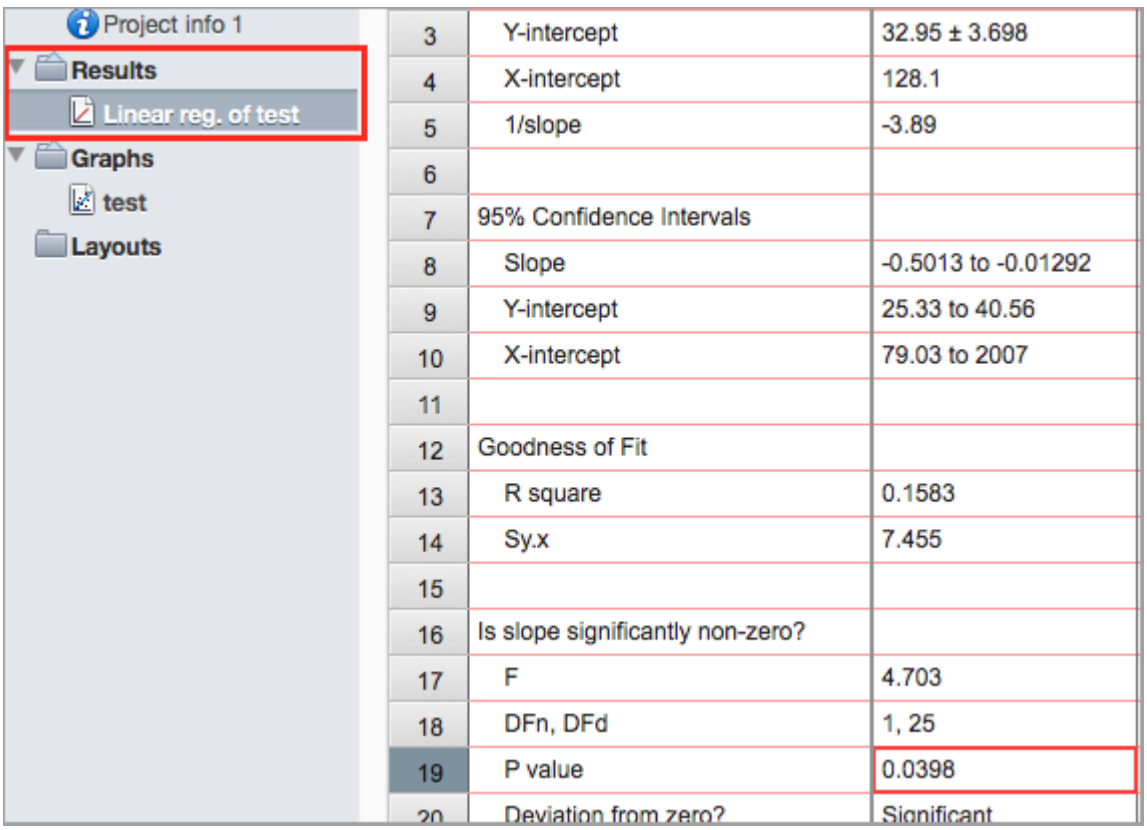
2. 导入两个变量所对应的两列数据，可以使用复制粘贴的方式，也可以是使用Import选项。



3. 点击Analyze，在弹出的对话框里选择Linear regression。之所以选择Linear regression而不是Correlation，是因为前者可以产生一条拟合曲线（详见Prism官方文档）。

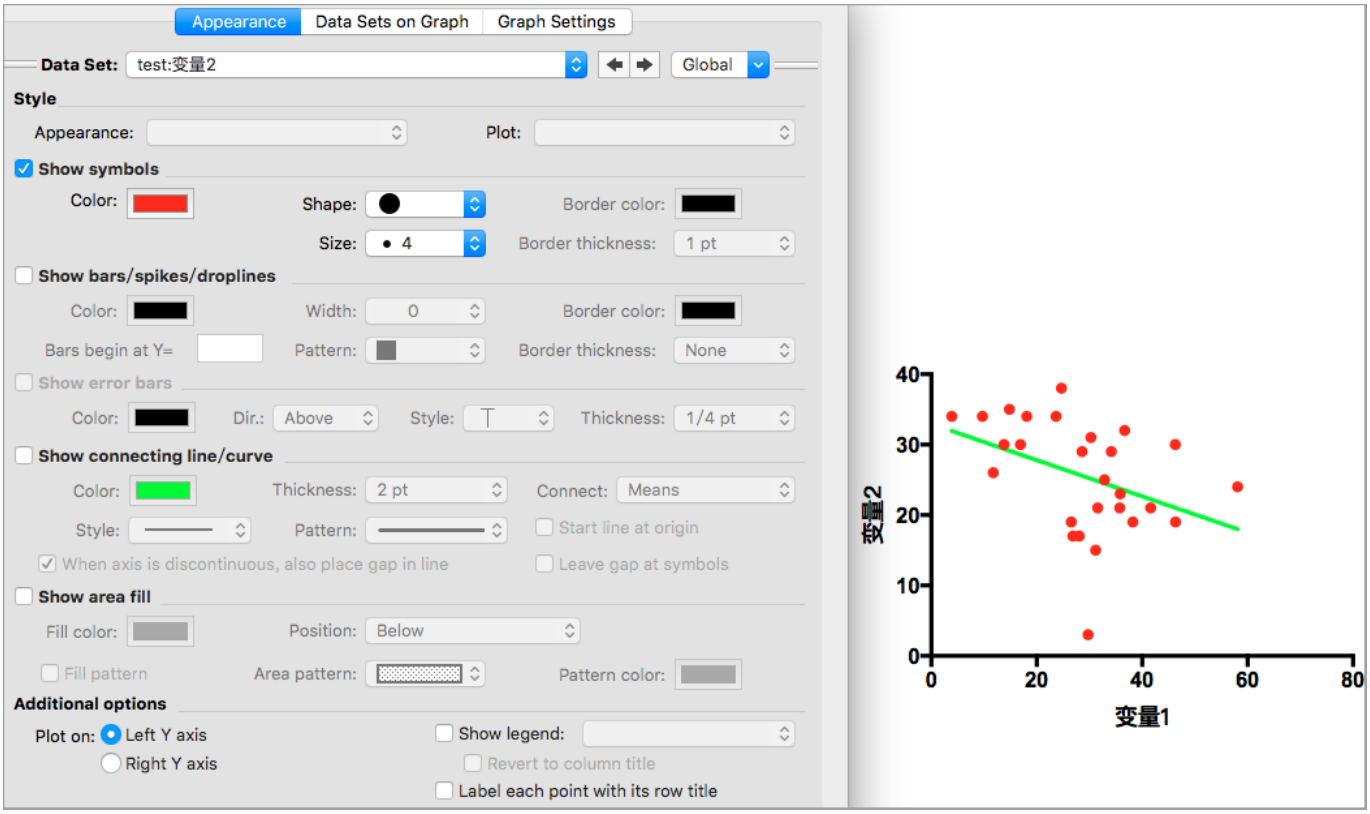


4. 选择Results，可以查看回归分析的结果，比如p值为0.039，小于0.05，说明这两个变量间存在显著的线性相关关系。



3	Y-intercept	32.95 ± 3.698
4	X-intercept	128.1
5	1/slope	-3.89
6		
7	95% Confidence Intervals	
8	Slope	-0.5013 to -0.01292
9	Y-intercept	25.33 to 40.56
10	X-intercept	79.03 to 2007
11		
12	Goodness of Fit	
13	R square	0.1583
14	Sy.x	7.455
15		
16	Is slope significantly non-zero?	
17	F	4.703
18	DFn, DFd	1, 25
19	P value	0.0398
20	Deviation from zero?	Significant

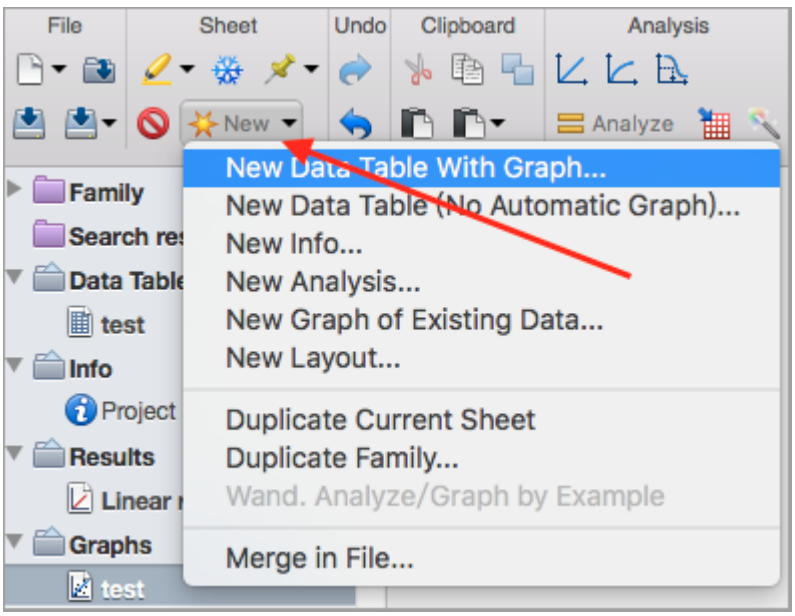
5. 选择Graphs，查看生成的图像。双击图中任意位置，在弹出的对话框里可以进一步修改图的形式。

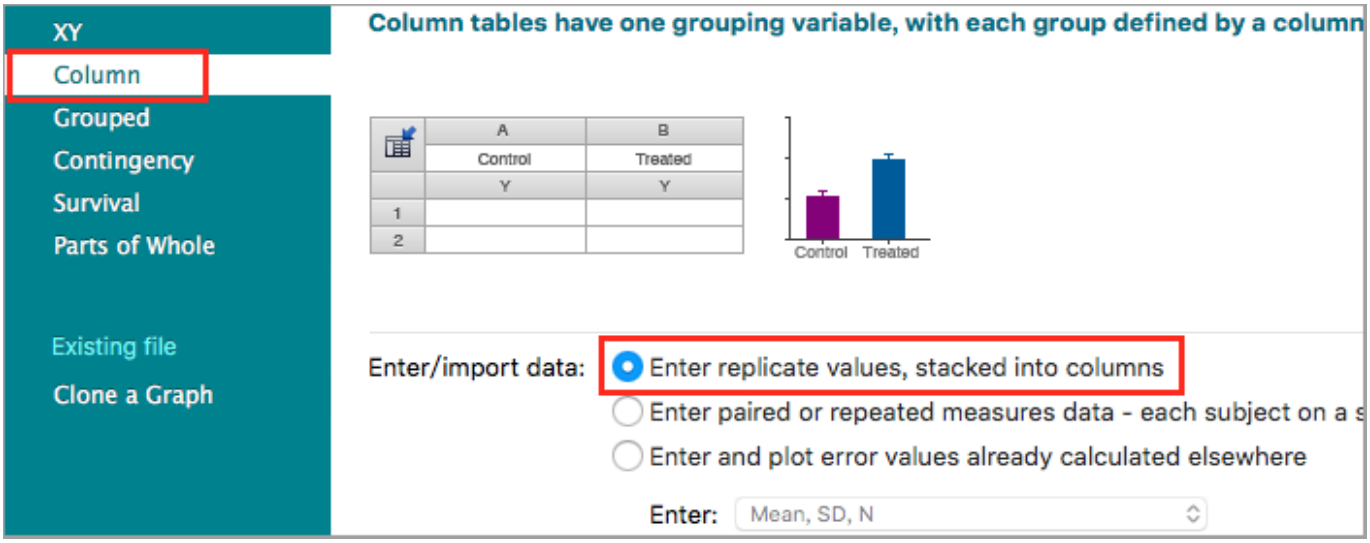


四、一个分类变量不同水平的比较

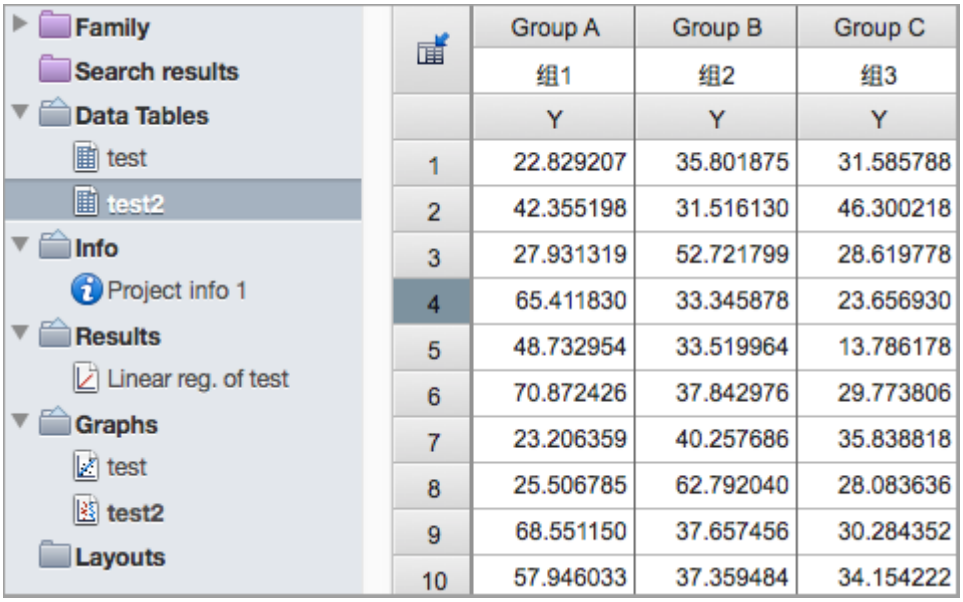
1. 假设该分类变量具有三个水平。

2. 点击New，选择New Data Table With Graph来添加新的数据表，在弹出的对话框中选择Column；输入数据选择Enter replicate values, stacked into columns；最后点击Create。



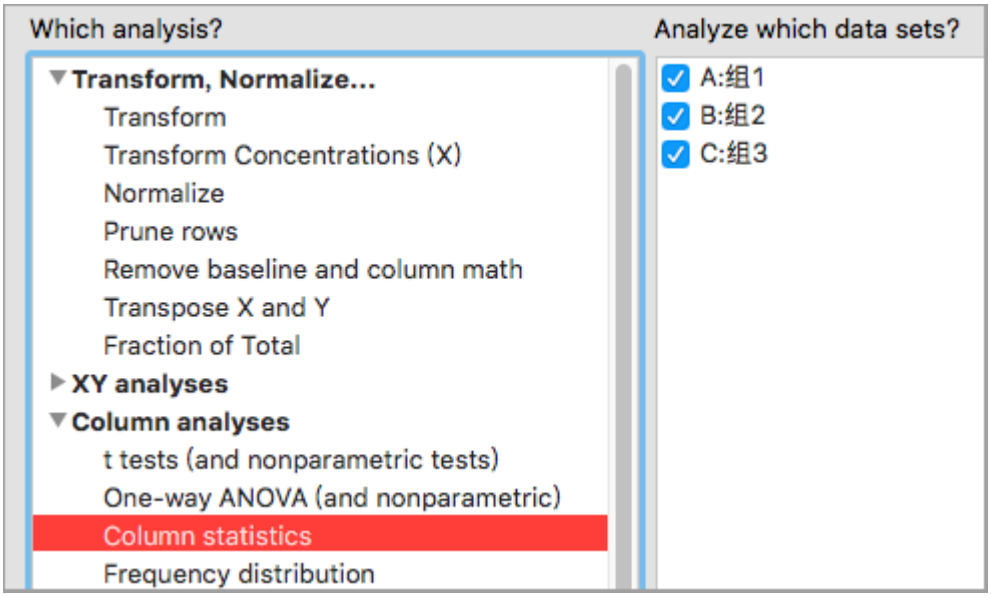


3. 按照前面相同的方式导入数据，每一列对应一个水平。

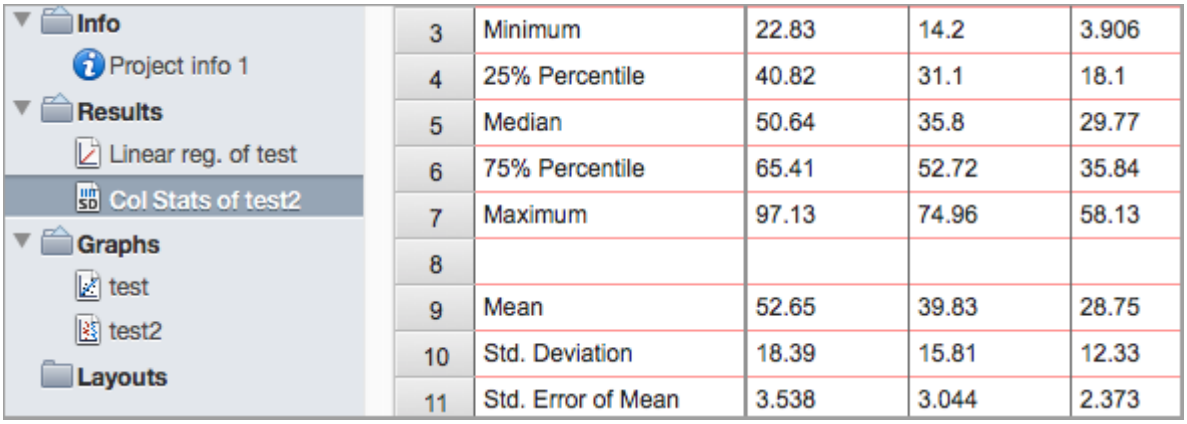


	Group A	Group B	Group C
	组1	组2	组3
	Y	Y	Y
1	22.829207	35.801875	31.585788
2	42.355198	31.516130	46.300218
3	27.931319	52.721799	28.619778
4	65.411830	33.345878	23.656930
5	48.732954	33.519964	13.786178
6	70.872426	37.842976	29.773806
7	23.206359	40.257686	35.838818
8	25.506785	62.792040	28.083636
9	68.551150	37.657456	30.284352
10	57.946033	37.359484	34.154222

4. 点击Analyze，在弹出的对话框中选择Column statistics。因为我已经用其他软件做了组间比较，这里我只需要计算统计量用于图示。

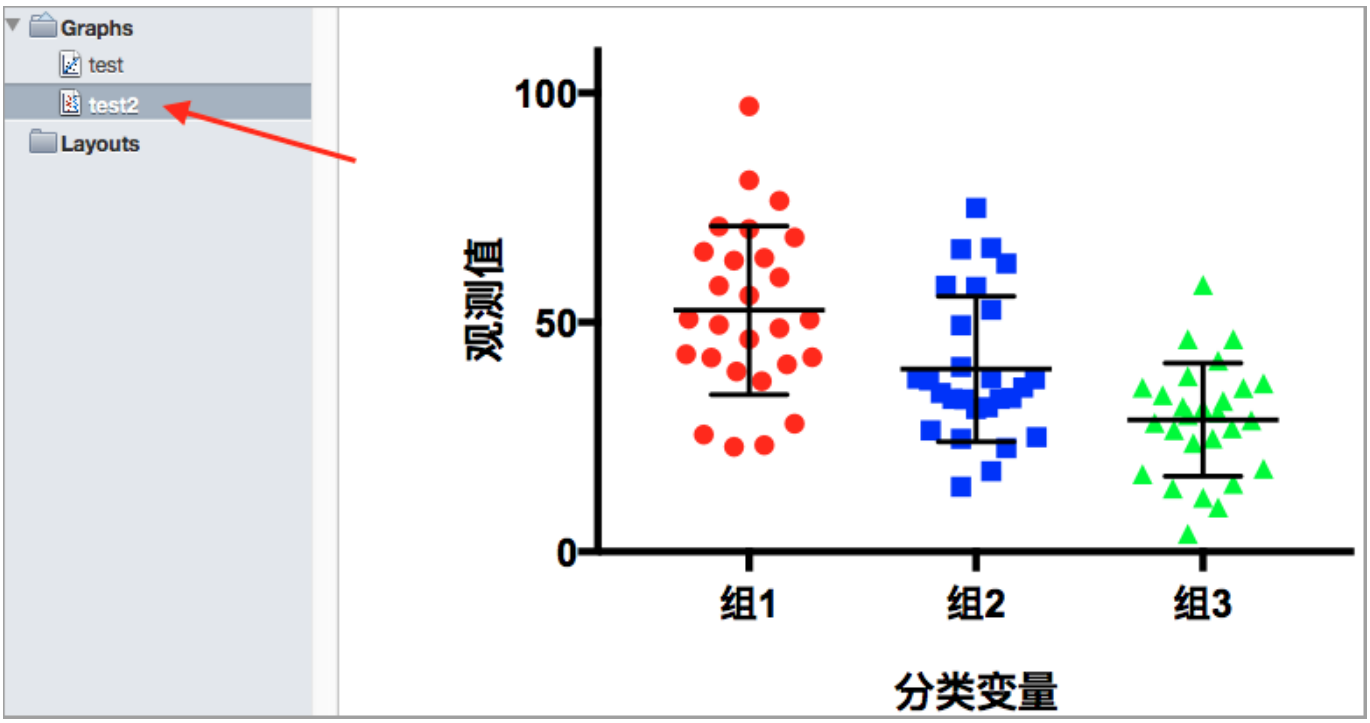


5. 在Results选项下面可以查看分析结果。



3	Minimum	22.83	14.2	3.906
4	25% Percentile	40.82	31.1	18.1
5	Median	50.64	35.8	29.77
6	75% Percentile	65.41	52.72	35.84
7	Maximum	97.13	74.96	58.13
8				
9	Mean	52.65	39.83	28.75
10	Std. Deviation	18.39	15.81	12.33
11	Std. Error of Mean	3.538	3.044	2.373

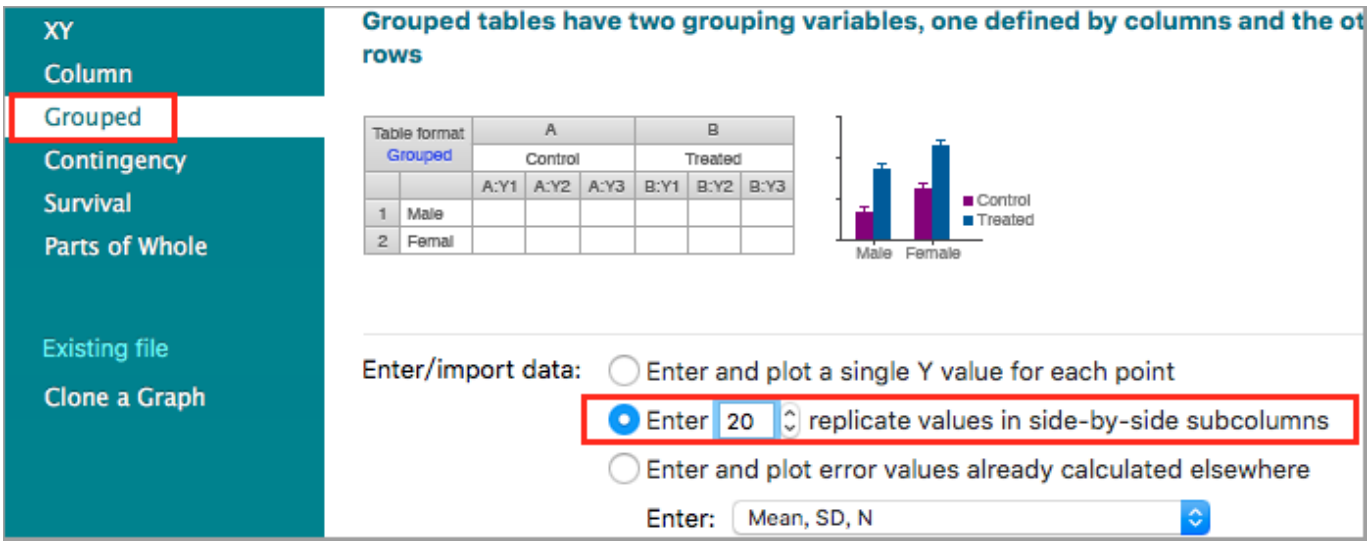
6. 在Graphs选项下查看生成的图像。双击图中任意位置，在弹出的对话框里可以进一步修改图的形式。



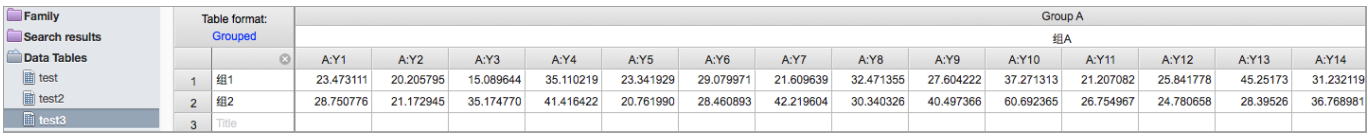
五、两个分类变量不同水平的比较

1. 假设一个分类变量具有三个水平，另一个分类变量具有两个水平，共有6种不同的组合。

2. 点击New，选择New Data Table With Graph来添加新的数据表，在弹出的对话框中选择Grouped；输入数据选择Enter ＊ replicate values in side-by-side subcolumns，其中＊表示不同组合下的样本数，如果不同组合下样本数不同，选择最大值；最后点击Create。



3. 根据不同组合情况导入样本数据。默认情况下，数据将按照列的方式导入（粘贴），这里数据需要按行的方式导入。

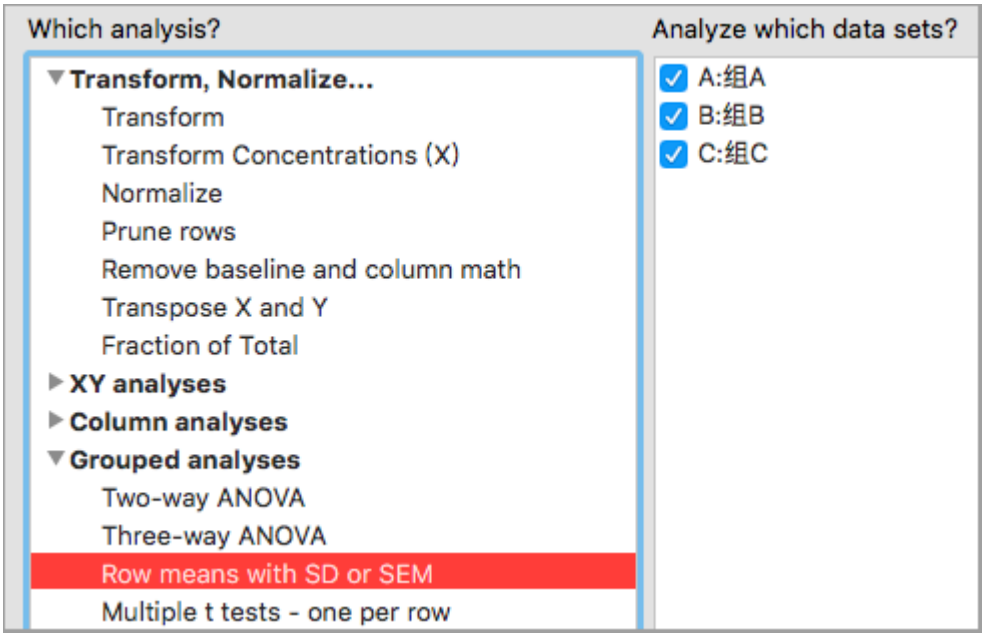


Family
Search results
Data Tables
test
test2
test3

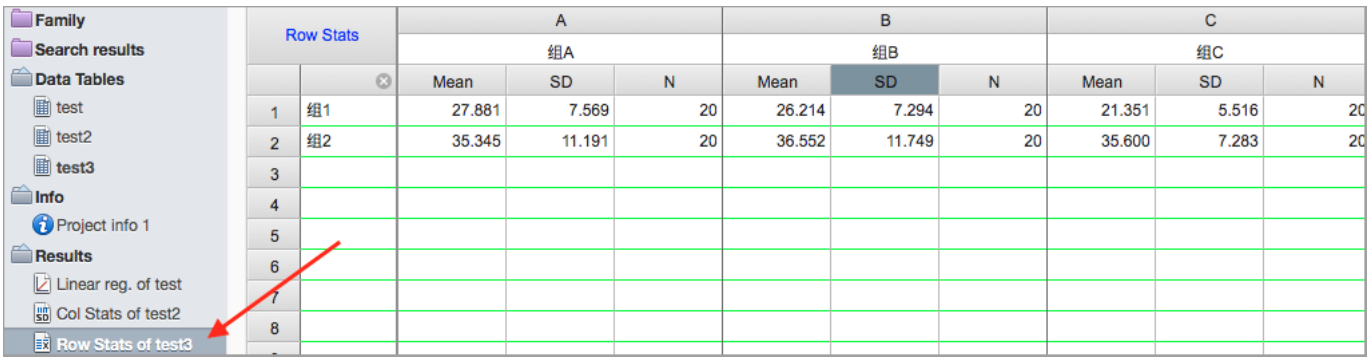
Table format: Grouped

		Group A													
		组A													
		A:Y1	A:Y2	A:Y3	A:Y4	A:Y5	A:Y6	A:Y7	A:Y8	A:Y9	A:Y10	A:Y11	A:Y12	A:Y13	A:Y14
1	组1	23.473111	20.205795	15.089644	35.110219	23.341929	29.079971	21.609639	32.471355	27.604222	37.271313	21.207082	25.841778	45.25173	31.232119
2	组2	28.750776	21.172945	35.174770	41.416422	20.761990	28.460893	42.219604	30.340326	40.497366	60.692365	26.754967	24.780658	28.39526	36.768981
3	Title														

4. 点击Analyze，在弹出的对话框中选择Row means with SD or SEM，原因如前所述。



5. 在Results选项下面可以查看分析结果，可以得到每种组合下的均值、方差和样本数。

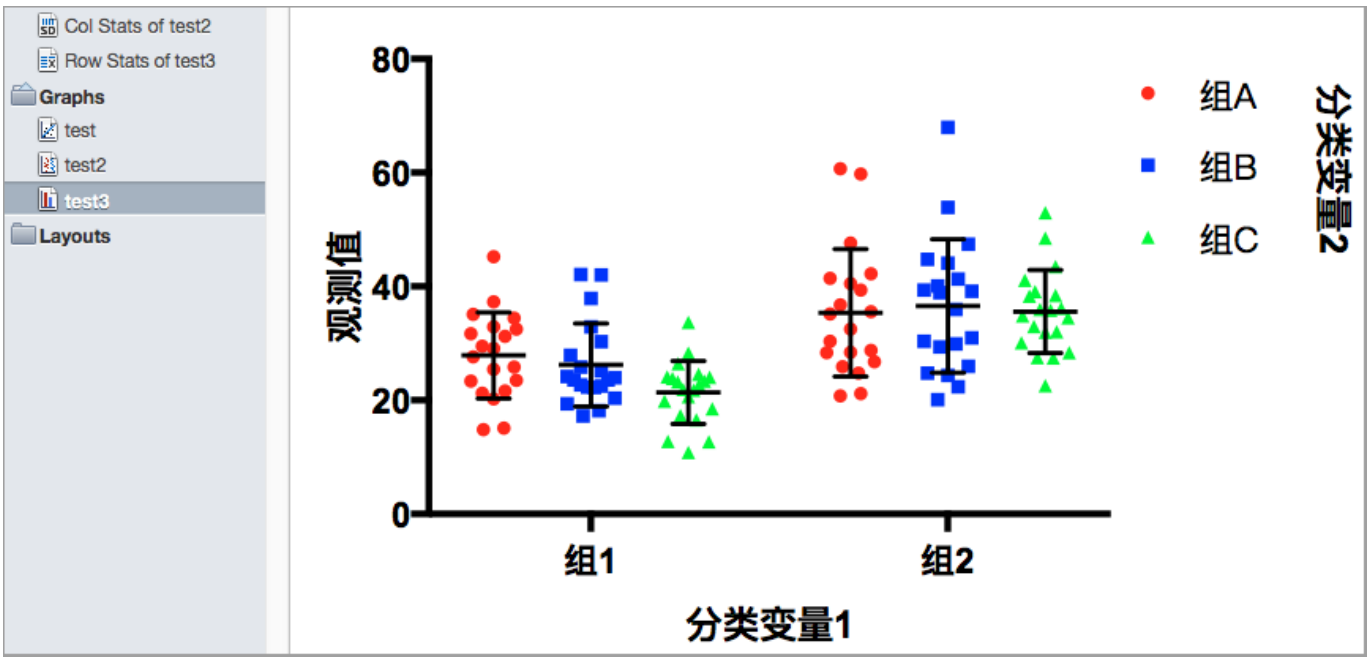


Family
Search results
Data Tables
test
test2
test3
Info
Project info 1
Results
Linear reg. of test
Col Stats of test2
Row Stats of test3

Row Stats

		A			B			C		
		组A			组B			组C		
		Mean	SD	N	Mean	SD	N	Mean	SD	N
1	组1	27.881	7.569	20	26.214	7.294	20	21.351	5.516	20
2	组2	35.345	11.191	20	36.552	11.749	20	35.600	7.283	20
3										
4										
5										
6										
7										
8										

6. 在Graphs选项下查看生成的图像。双击图中任意位置，在弹出的对话框里可以进一步修改图的形式。



五、将图像导出为不同格式

使用Export选项，可以将生成的图像导出为不同的格式，比如PDF格式。