Chapter 13: The paired *t*-test: Comparing two related sets of measurements

TXCL7565/PHSC7565

What This Chapter Covers

- Recognizing when data are paired
- Paired *t*-test vs. two-sample *t*-test
- Performing a paired *t*-test
- What determines significance in a paired t-test
- Power of the paired t-test
- Choice of experimental design
- Requirement for applying paired t-test
- Sample sizes

RECOGNIZING WHEN DATA ARE PAIRED

Paired Data

Two samples are paired when observations from the first sample are paired in some meaningful way with observations in the second sample.

Examples:

- A variable is measured in each subject before and after an intervention.
- Subjects are recruited as pairs and matched for variables such as age, postal code, or diagnosis. One of each pair receives one intervention, whereas the other receives an alternative treatment.
- Twins or siblings are recruited as pairs. Each receives a different treatment.
- Each run of a laboratory experiment has a control and treated preparation handled in parallel.
- A part of the body on one side is treated with a control treatment and the corresponding part of the body on the other side is treated with the experimental treatment (e.g., right and left eyes).

PAIRED *t*-TEST VS. TWO-SAMPLE *t*-TEST

Paired t-test

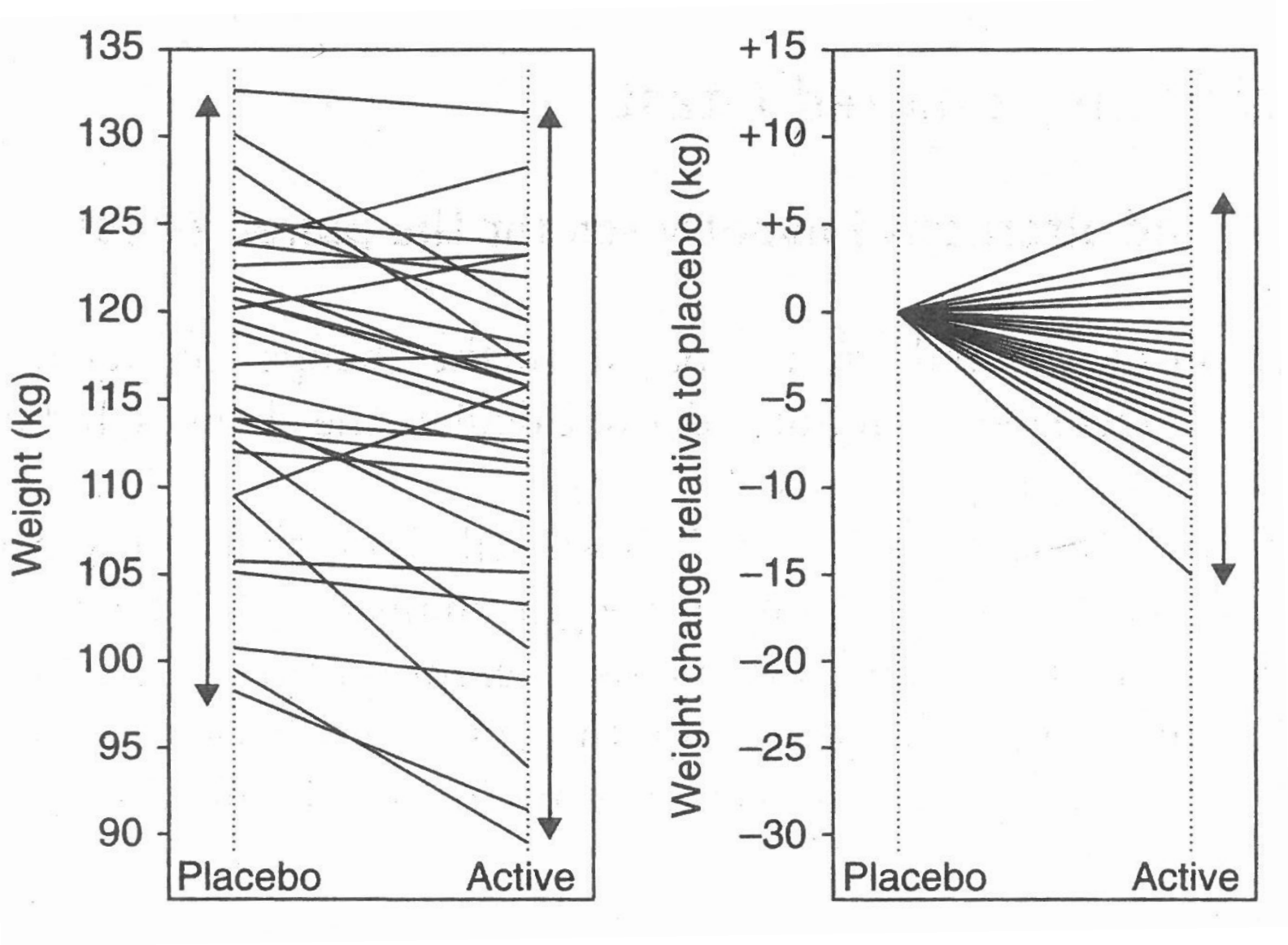
When observations in two samples are paired in a meaningful way, statistical inference is based on the *differences between the responses of the sample pairs*.

- paired t-test - mean of the differences
- two-sample t-test - difference of the means

weight example

Sample	Weight after placebo (kg)	Weight after active (kg)	Change in weight (active - placebo) (kg)
1	115.4	112.7	-2.7
2	118.9	113.8	-5.1
…			
30	113.8	103.3	-7.2
mean	116.08	112.47	-3.62
SD	8.94	10.43	4.78

Variability



Changes are less variable than the actual values

Where a series of individuals have widely differing values for an endpoint, we may yet find that a treatment induces a relatively constant change in all individuals. In such a case, the actual values will have large SDs but the SD among the changes will be smaller.

GRAPHING PAIRED DATA IN PRISM

Welcome to GraphPad Prism

GraphPad PRISM 7.0a

New table & graph

XY

Column

Grouped

Contingency

Survival

Parts of Whole

Existing file

Open a File

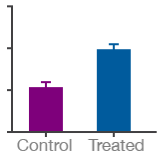
LabArchives

Clone a Graph

Graph Portfolio

Column tables have one grouping variable, with each group defined by a column

	A	B
	Control	Treated
	Y	Y
1		
2		



? Learn more

Enter/import data:

- Enter replicate values, stacked into columns
- Enter paired or repeated measures data - each subject on a separate row
- Enter and plot error values already calculated elsewhere

Enter: Mean, SD, N

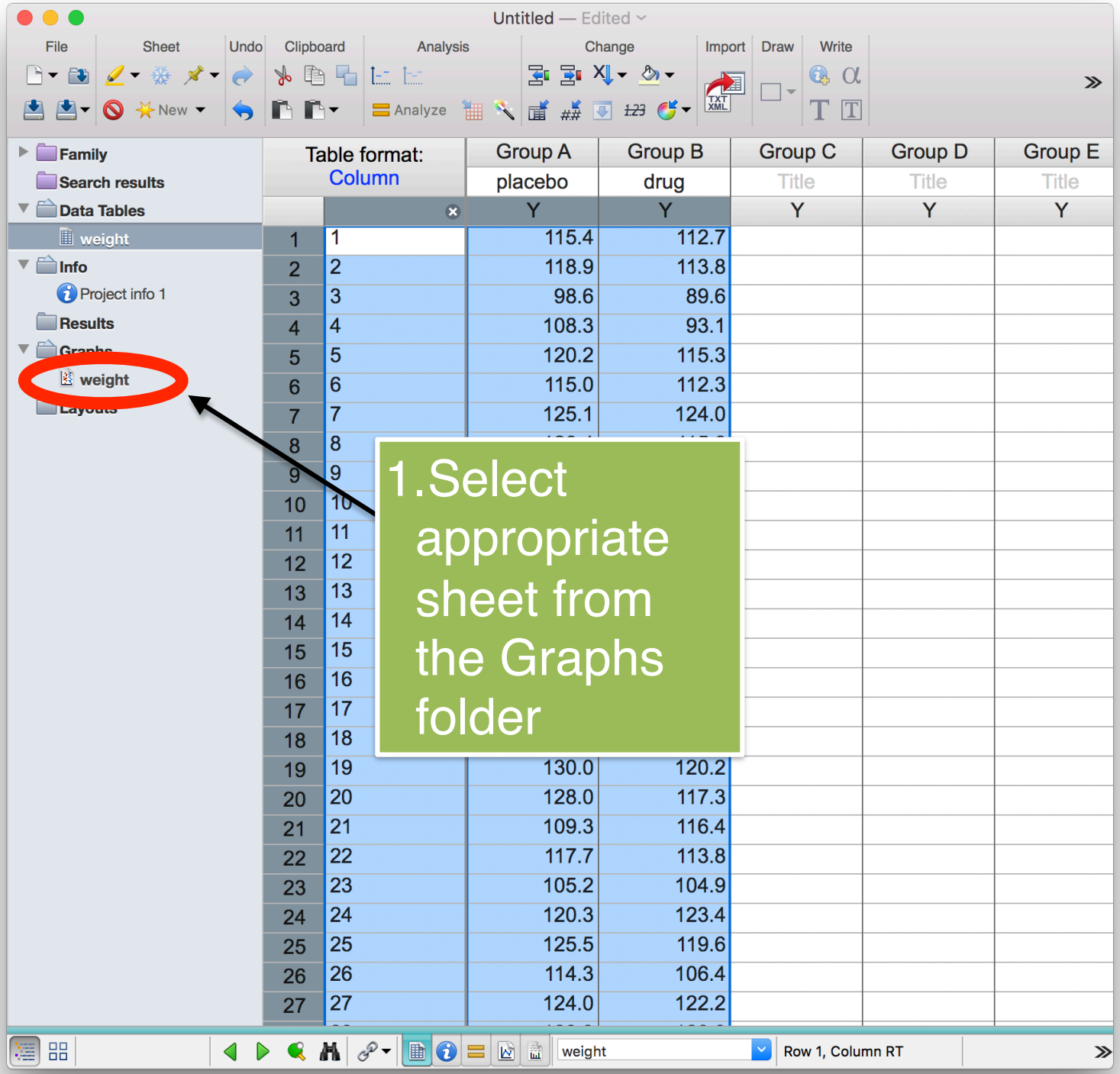
Use tutorial data:

- Frequency distribution data and histogram
- t test - Unpaired
- t test - Paired
- t test - One sample
- One-way ANOVA - Ordinary
- One-way ANOVA - Repeated measures
- Analyze a stack of P values
- More tutorial data...

Prism Tips

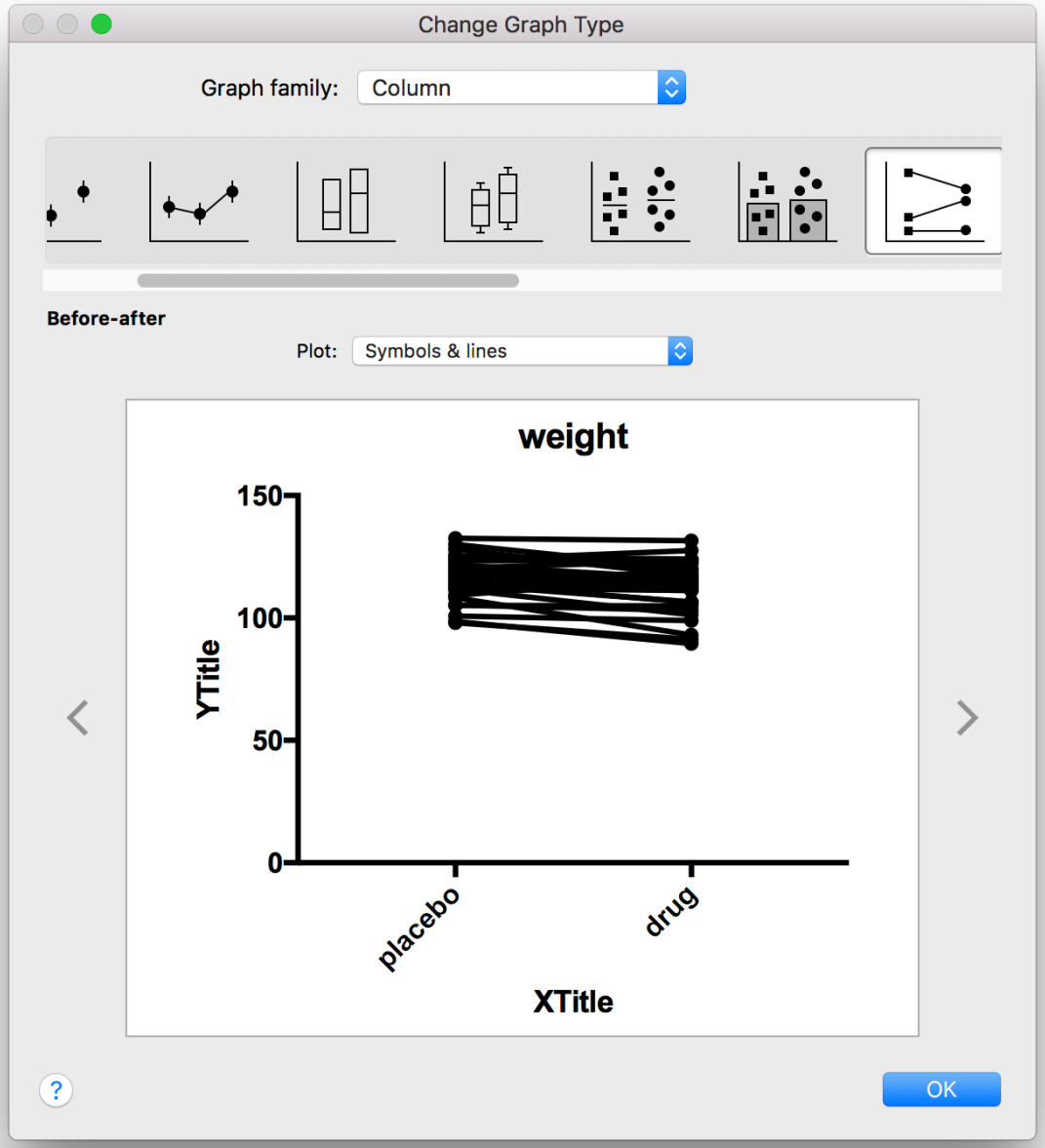
Cancel

Create

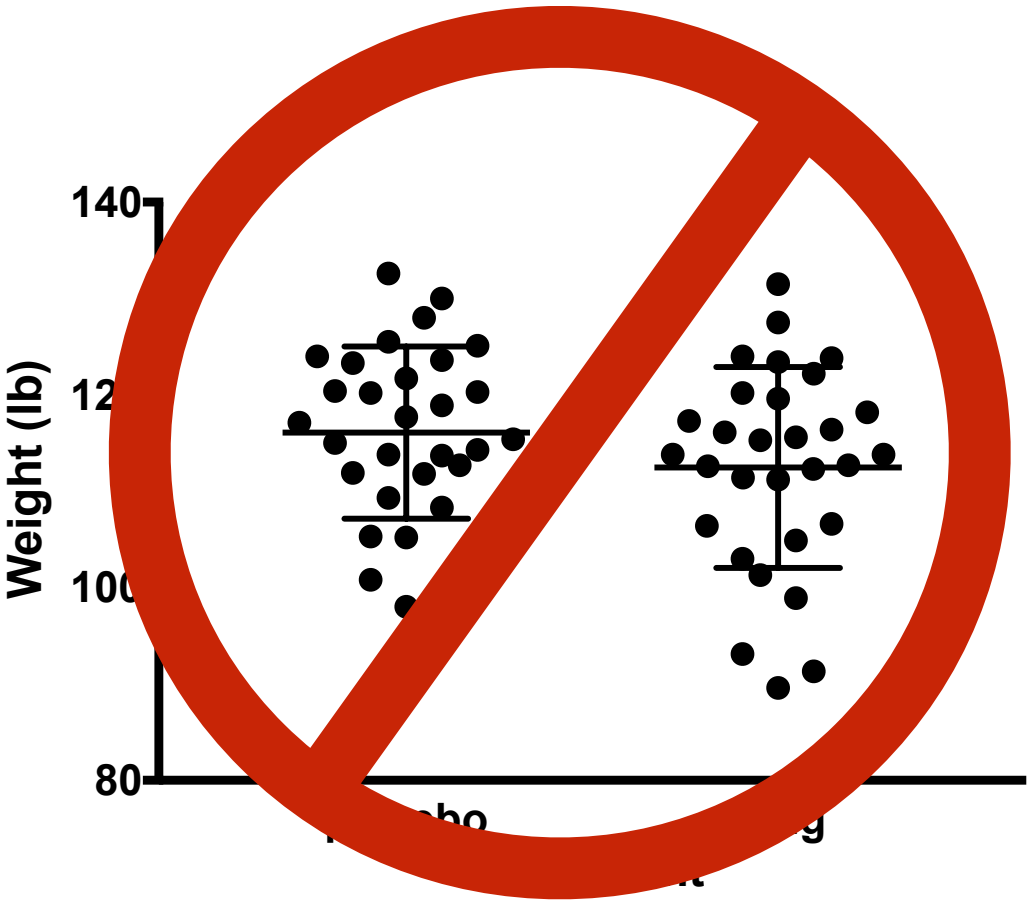


		Group A placebo	Group B drug	Group C Title	Group D Title	Group E Title
		Y	Y	Y	Y	Y
1	1	115.4	112.7			
2	2	118.9	113.8			
3	3	98.6	89.6			
4	4	108.3	93.1			
5	5	120.2	115.3			
6	6	115.0	112.3			
7	7	125.1	124.0			
8	8					
9	9					
10	10					
11	11					
12	12					
13	13					
14	14					
15	15					
16	16					
17	17					
18	18					
19	19	130.0	120.2			
20	20	128.0	117.3			
21	21	109.3	116.4			
22	22	117.7	113.8			
23	23	105.2	104.9			
24	24	120.3	123.4			
25	25	125.5	119.6			
26	26	114.3	106.4			
27	27	124.0	122.2			

Data from the same pair are places in the same row.



If you indicated that the data were paired when you initiated your data set. The 'Before-after' layout will be automatically chosen, if you did not indicate that the data were paired when you initiated this data set, you can still choose this graphic layout.



It is not appropriate to graph the data as two separate groups. More on this later…

PERFORMING A PAIRED *t*-TEST

Null and alternative hypotheses

$$H_0\text{: } \mu_{difference} = 0$$
$$H_a\text{: } \mu_{difference} \neq 0$$

Null hypothesis - The mean difference is equal to zero.

Alternative hypothesis - The mean difference is NOT equal to zero.

weight example

$$H_0\text{: } \mu_{drug-placebo} = 0$$
$$H_a\text{: } \mu_{drug-placebo} \neq 0$$

Test statistic

$$t = \frac{\bar{D}}{s_{diff}/\sqrt{N}}$$

$\bar{D}$ = sample mean of the difference
s_{diff} = sample standard deviation
N = number of differences, i.e., the number of pairs

2-sample *t*-test vs. paired *t*-test

In the two-sample t-test the two steps were:

1. Calculation of the mean for each sample

2. Calculation of the 95% CI for the difference between the two means

Means then difference of the means

For the paired t-test the order is reversed:

1. Calculation of the change that has occurred within each pair of results.

2. Calculation of the 95% CI for the mean among the changes.

Differences then mean of the differences

2 ways to perform a paired *t*-test

1. Enter pairs of data into software and request a paired *t*-test

2. Calculate the ‘difference’ for each pair and then request a one-sample *t*-test

paired *t*-test in Prism

Welcome to GraphPad Prism

GraphPad PRISM 7.0a

New table & graph

XY

Column

Grouped

Contingency

Survival

Parts of Whole

Existing file

Open a File

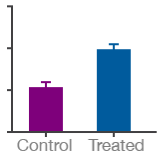
LabArchives

Clone a Graph

Graph Portfolio

Column tables have one grouping variable, with each group defined by a column

	A	B
	Control	Treated
	Y	Y
1		
2		



? Learn more

Enter/import data:

- Enter replicate values, stacked into columns
- Enter paired or repeated measures data - each subject on a separate row
- Enter and plot error values already calculated elsewhere

Enter: Mean, SD, N

Use tutorial data:

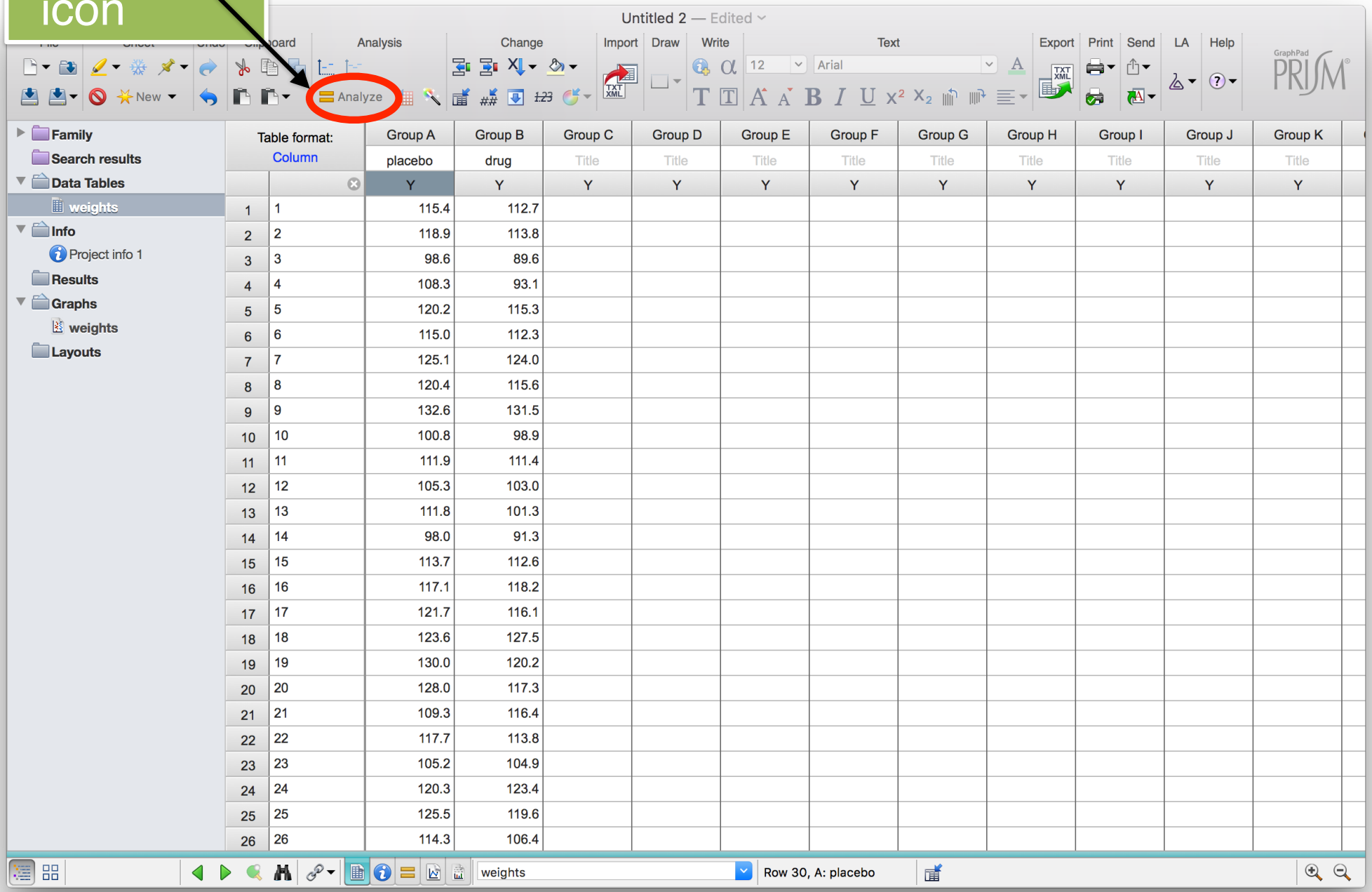
- Frequency distribution data and histogram
- t test - Unpaired
- t test - Paired
- t test - One sample
- One-way ANOVA - Ordinary
- One-way ANOVA - Repeated measures
- Analyze a stack of P values
- More tutorial data...

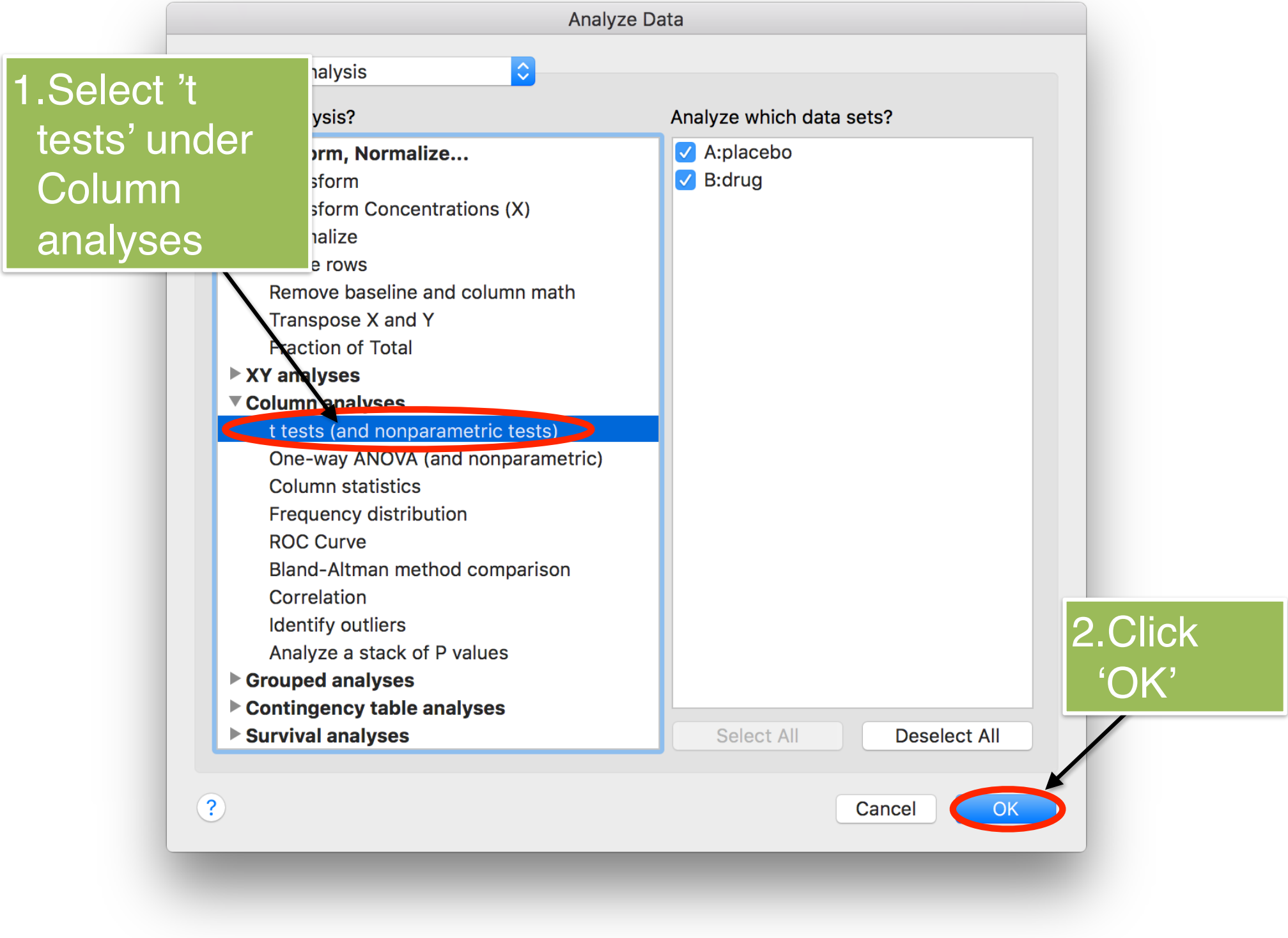
Prism Tips

Cancel

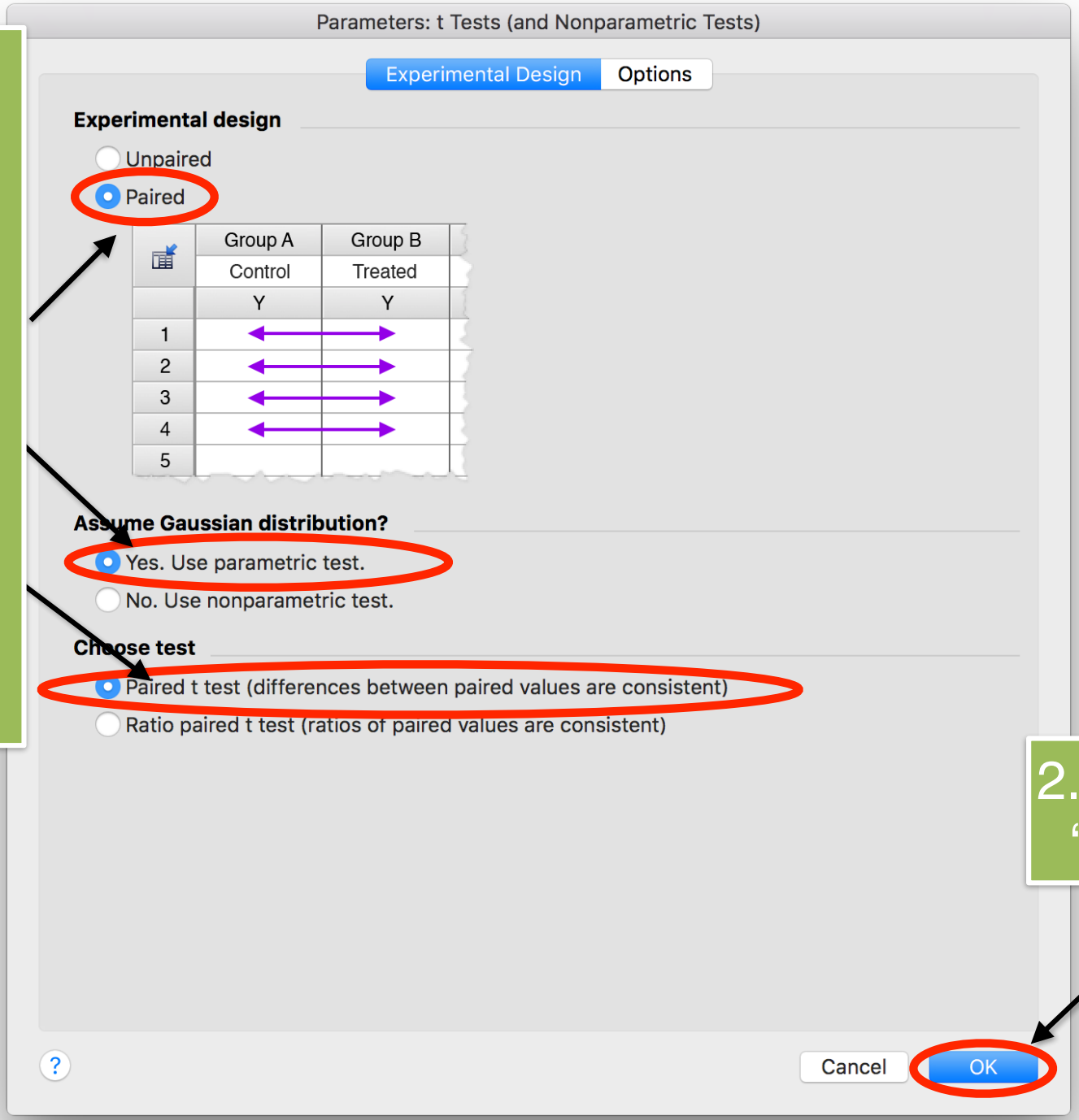
Create

1.Click ‘Analyze icon

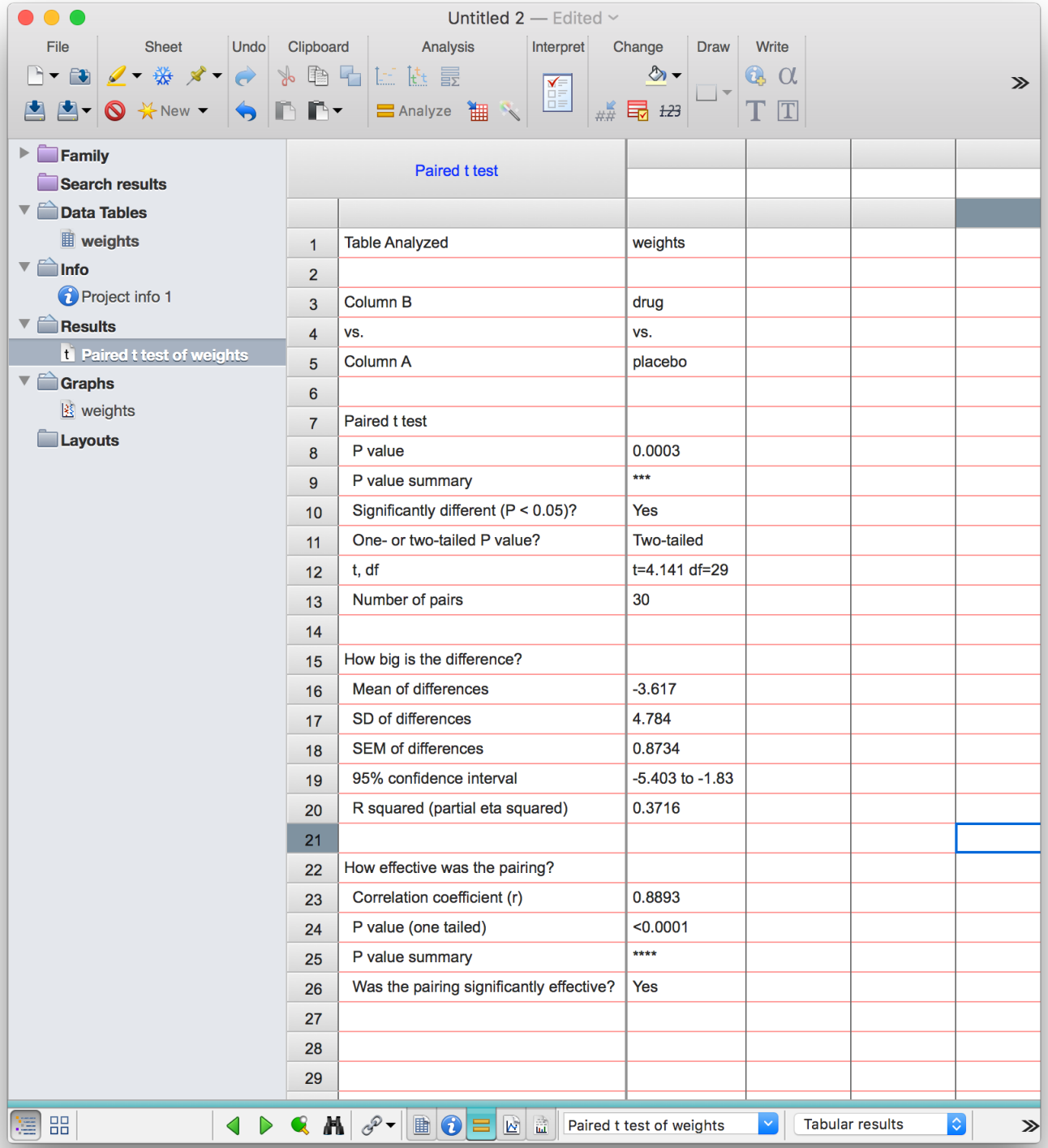



1.Select 't tests' under Column analyses
Analyze Data
nalysis
ysis?
Analyze which data sets?
A:placebo
B:drug
orm, Normalize...
sform
sform Concentrations (X)
nalize
e rows
Remove baseline and column math
Transpose X and Y
Fraction of Total
XY analyses
Column analyses
t tests (and nonparametric tests)
One-way ANOVA (and nonparametric)
Column statistics
Frequency distribution
ROC Curve
Bland-Altman method comparison
Correlation
Identify outliers
Analyze a stack of P values
Grouped analyses
Contingency table analyses
Survival analyses
Select All
Deselect All
Cancel
OK
2.Click 'OK'

1. Select a ‘Paired’ Experimental design; assume a Gaussian distribution; choose a paired t test based on differences



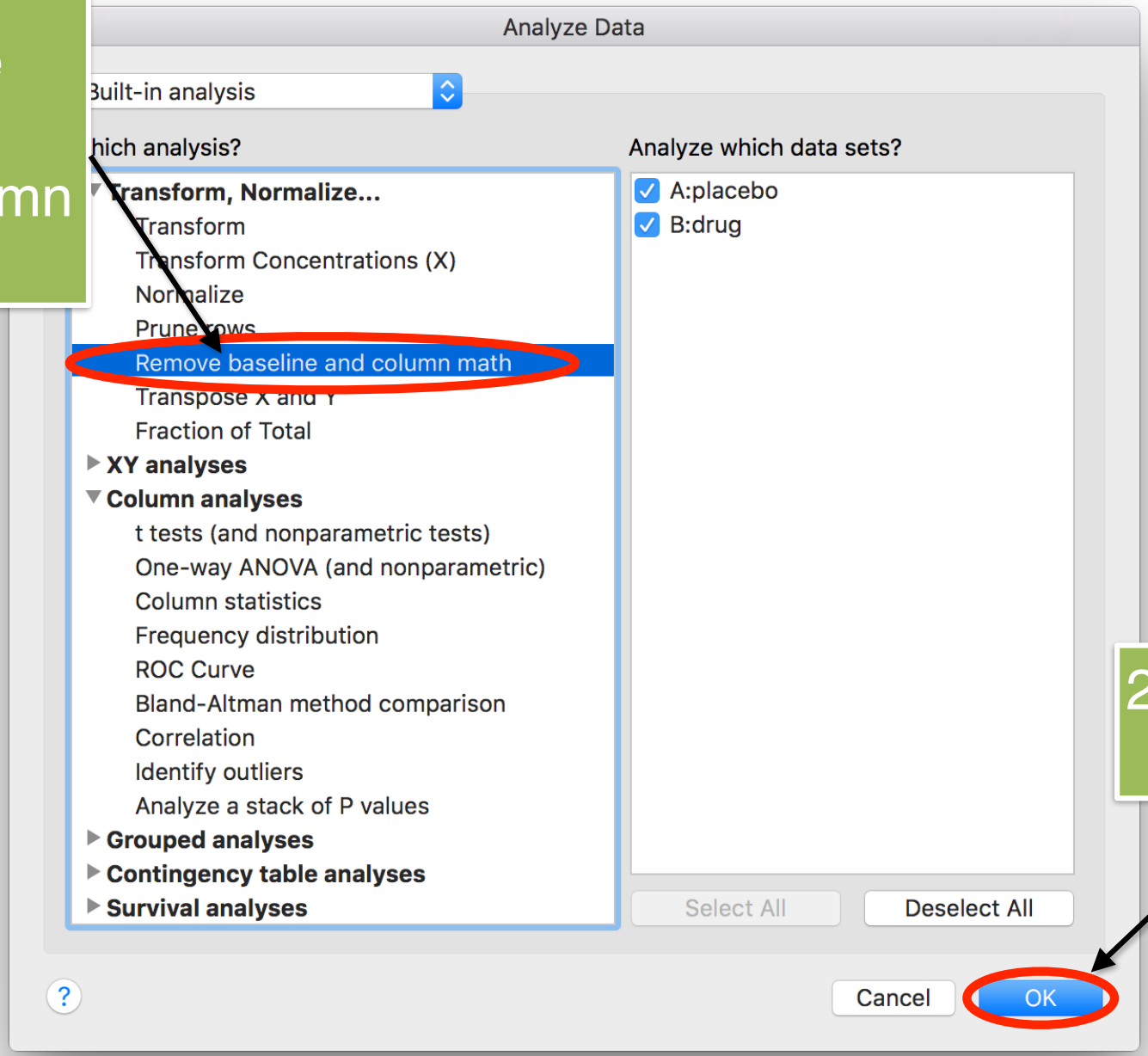
2. Click ‘OK’


Untitled 2 — Edited
File
Sheet
Undo
Clipboard
Analysis
Interpret
Change
Draw
Write
New
Analyze
Family
Search results
Data Tables
weights
Info
Project info 1
Results
Paired t test of weights
Graphs
weights
Layouts
Paired t test
1 Table Analyzed weights
2
3 Column B drug
4 vs. vs.
5 Column A placebo
6
7 Paired t test
8 P value 0.0003
9 P value summary ***
10 Significantly different (P < 0.05)? Yes
11 One- or two-tailed P value? Two-tailed
12 t, df t=4.141 df=29
13 Number of pairs 30
14
15 How big is the difference?
16 Mean of differences -3.617
17 SD of differences 4.784
18 SEM of differences 0.8734
19 95% confidence interval -5.403 to -1.83
20 R squared (partial eta squared) 0.3716
21
22 How effective was the pairing?
23 Correlation coefficient (r) 0.8893
24 P value (one tailed) <0.0001
25 P value summary ****
26 Was the pairing significantly effective? Yes
27
28
29
Paired t test of weights
Tabular results

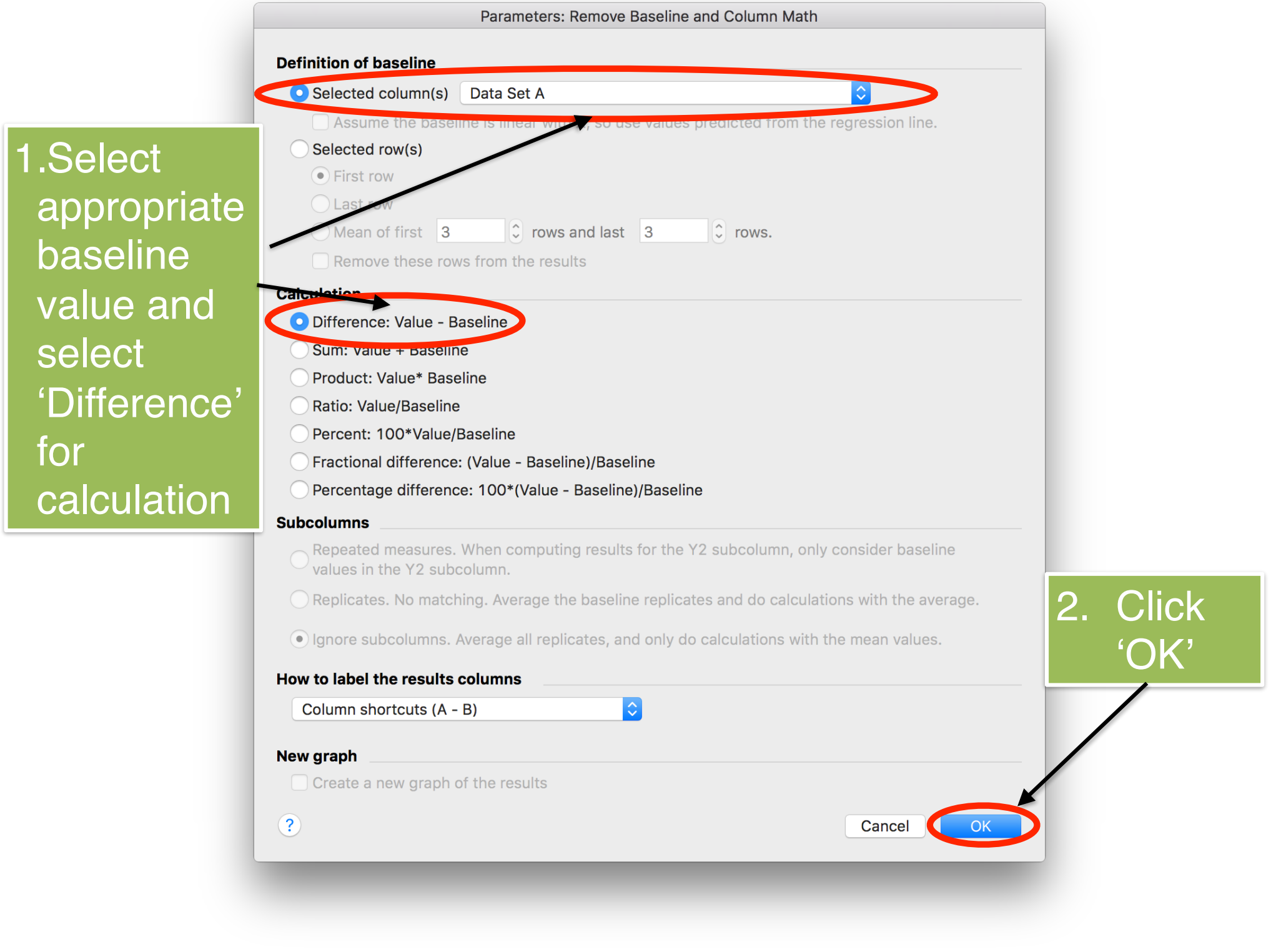
one sample *t*-test on difference in Prism

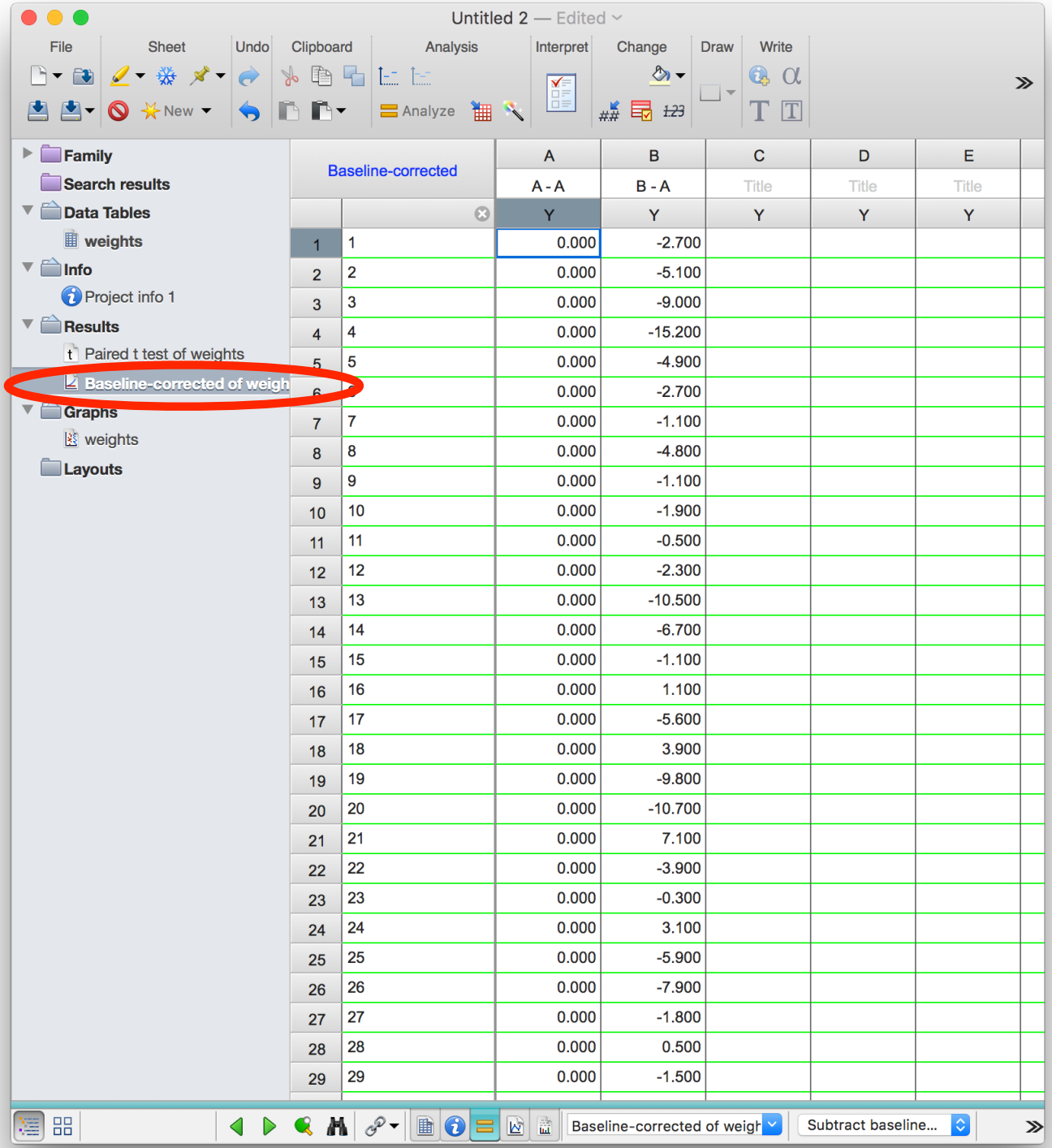
Calculate Difference Between Weights (Drug - Placebo)

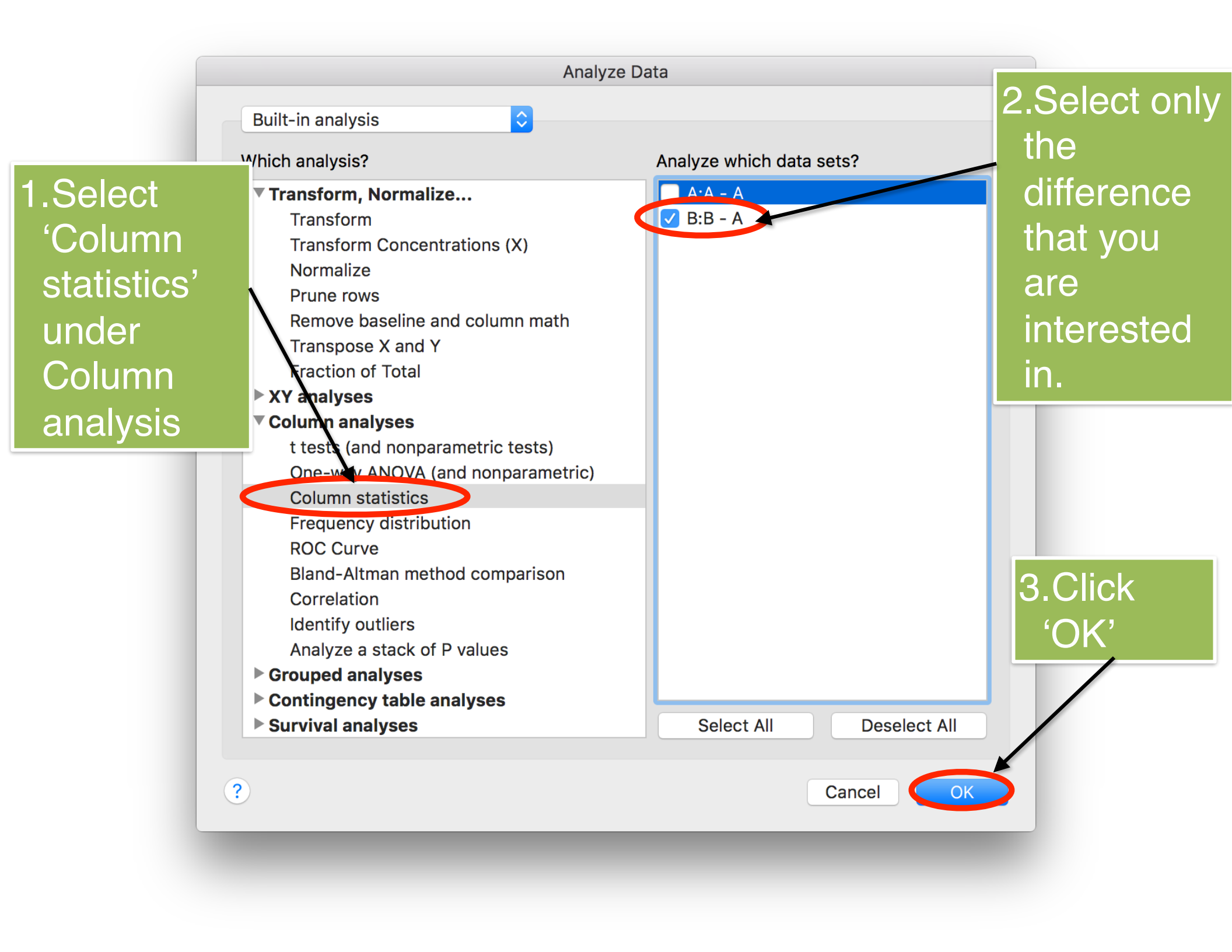
1. Select ‘Remove baseline and column math’



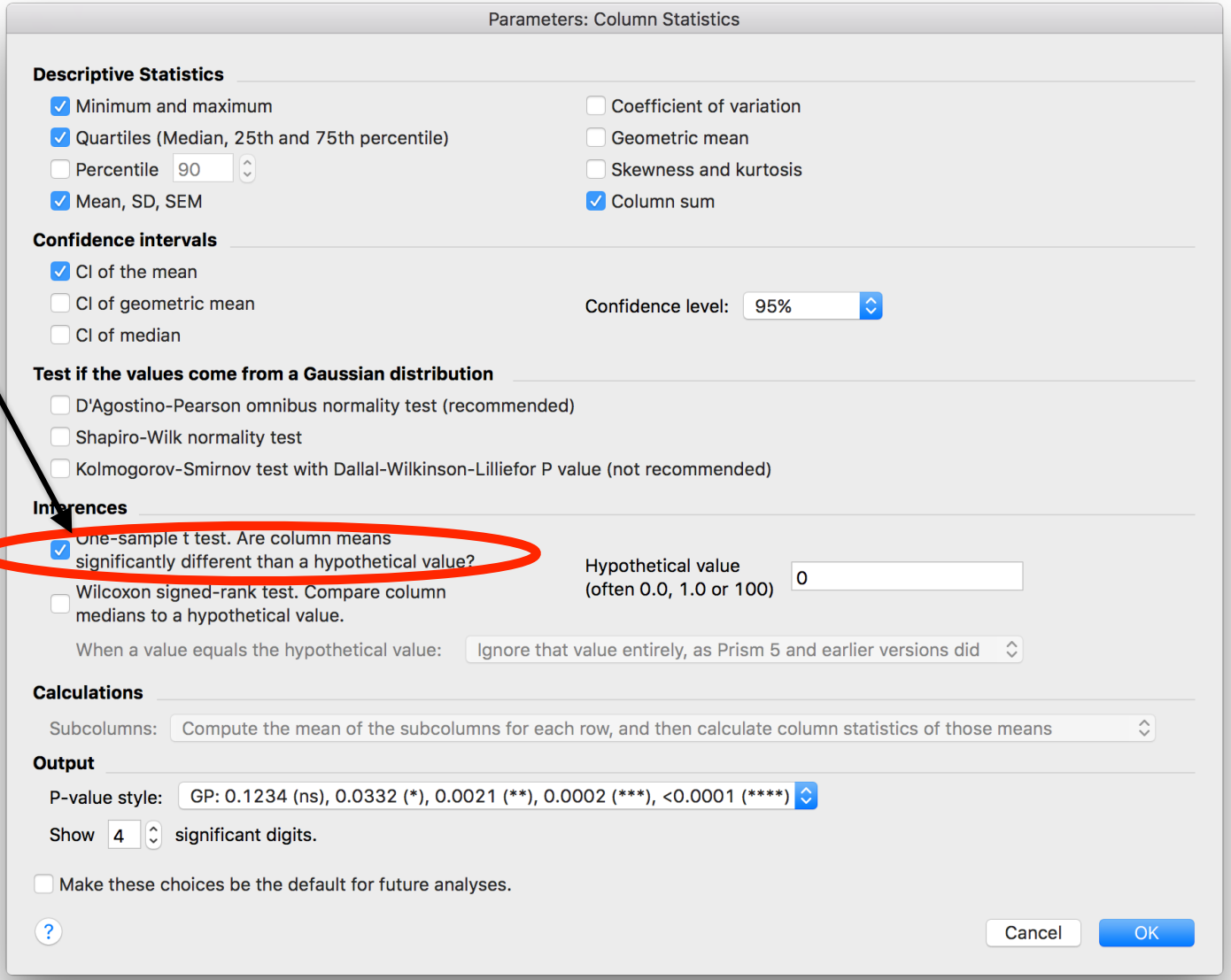
2. Click ‘OK’

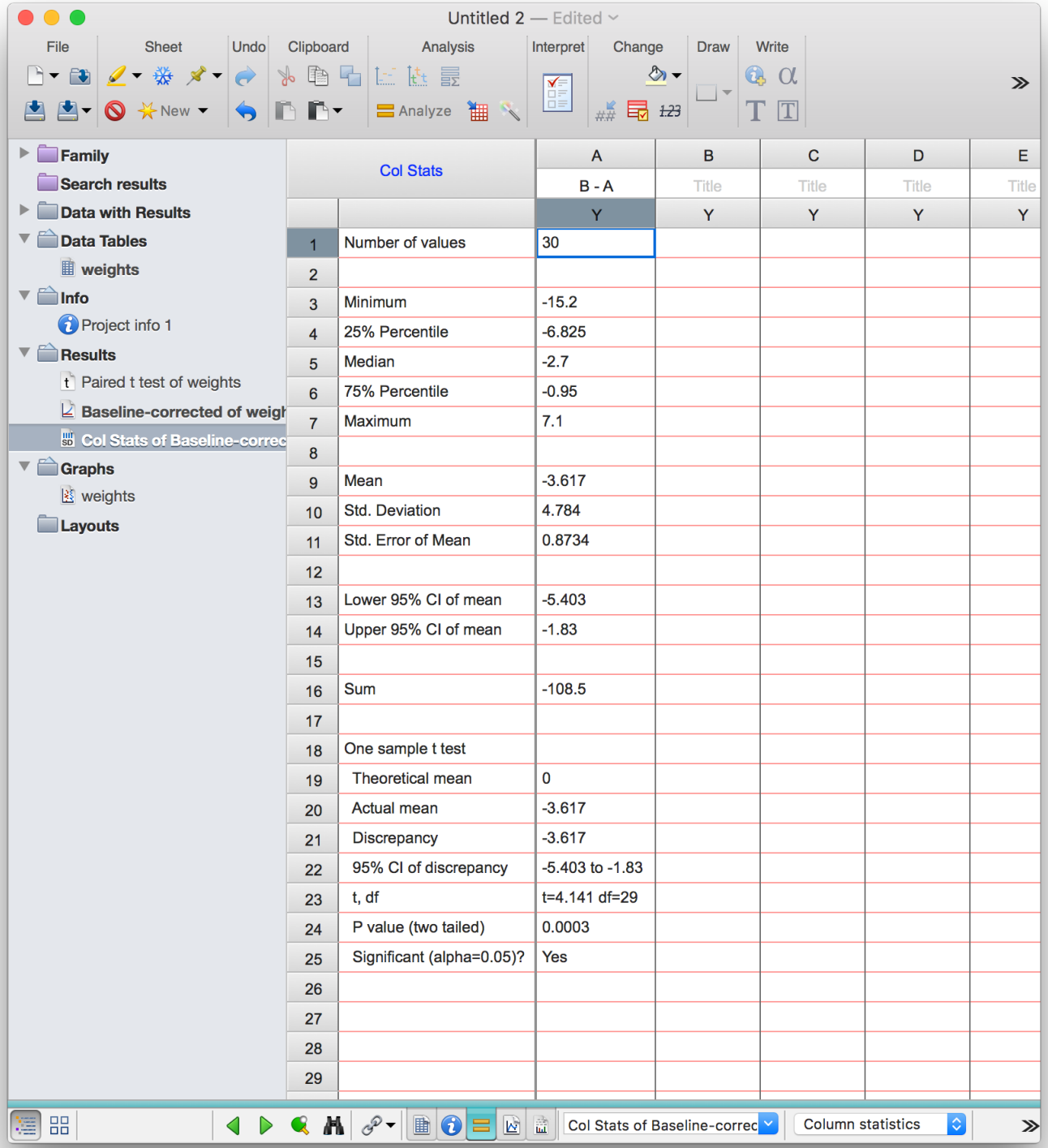

Parameters: Remove Baseline and Column Math
Definition of baseline
Selected column(s)
Data Set A
Assume the baseline is linear with
Selected row(s)
First row
Last row
Mean of first
3
rows and last
3
rows.
Remove these rows from the results
Calculation
Difference: Value - Baseline
Sum: Value + Baseline
Product: Value* Baseline
Ratio: Value/Baseline
Percent: 100*Value/Baseline
Fractional difference: (Value - Baseline)/Baseline
Percentage difference: 100*(Value - Baseline)/Baseline
Subcolumns
Repeated measures. When computing results for the Y2 subcolumn, only consider baseline values in the Y2 subcolumn.
Replicates. No matching. Average the baseline replicates and do calculations with the average.
Ignore subcolumns. Average all replicates, and only do calculations with the mean values.
How to label the results columns
Column shortcuts (A - B)
New graph
Create a new graph of the results
?
Cancel
OK
1.Select appropriate baseline value and select ‘Difference’ for calculation
2. Click ‘OK’

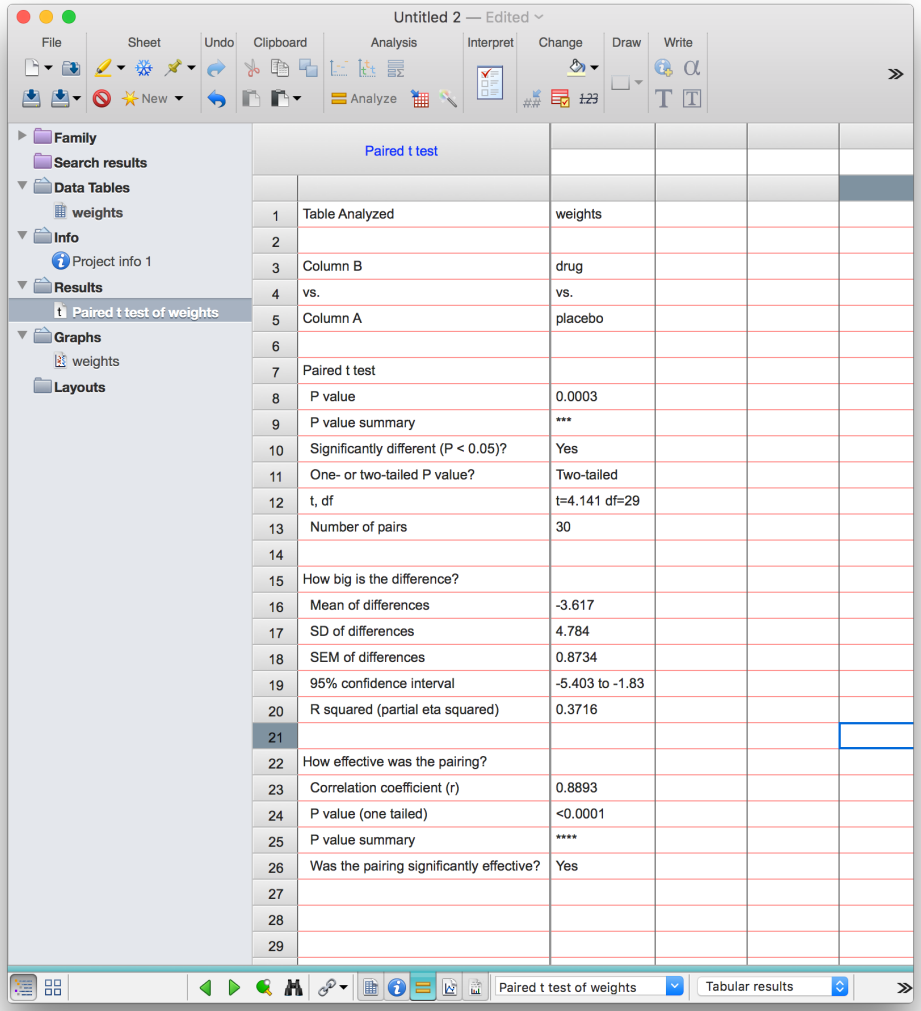

Untitled 2 — Edited
File
Sheet
Undo
Clipboard
Analysis
Interpret
Change
Draw
Write
New
Analyze
Family
Search results
Data Tables
weights
Info
Project info 1
Results
Paired t test of weights
Baseline-corrected of weigh
Graphs
weights
Layouts
Baseline-corrected
A
B
C
D
E
A - A
B - A
Title
Title
Title
Y
Y
Y
Y
Y
1 1 0.000 -2.700
2 2 0.000 -5.100
3 3 0.000 -9.000
4 4 0.000 -15.200
5 5 0.000 -4.900
6 0.000 -2.700
7 7 0.000 -1.100
8 8 0.000 -4.800
9 9 0.000 -1.100
10 10 0.000 -1.900
11 11 0.000 -0.500
12 12 0.000 -2.300
13 13 0.000 -10.500
14 14 0.000 -6.700
15 15 0.000 -1.100
16 16 0.000 1.100
17 17 0.000 -5.600
18 18 0.000 3.900
19 19 0.000 -9.800
20 20 0.000 -10.700
21 21 0.000 7.100
22 22 0.000 -3.900
23 23 0.000 -0.300
24 24 0.000 3.100
25 25 0.000 -5.900
26 26 0.000 -7.900
27 27 0.000 -1.800
28 28 0.000 0.500
29 29 0.000 -1.500
Baseline-corrected of weigh
Subtract baseline...

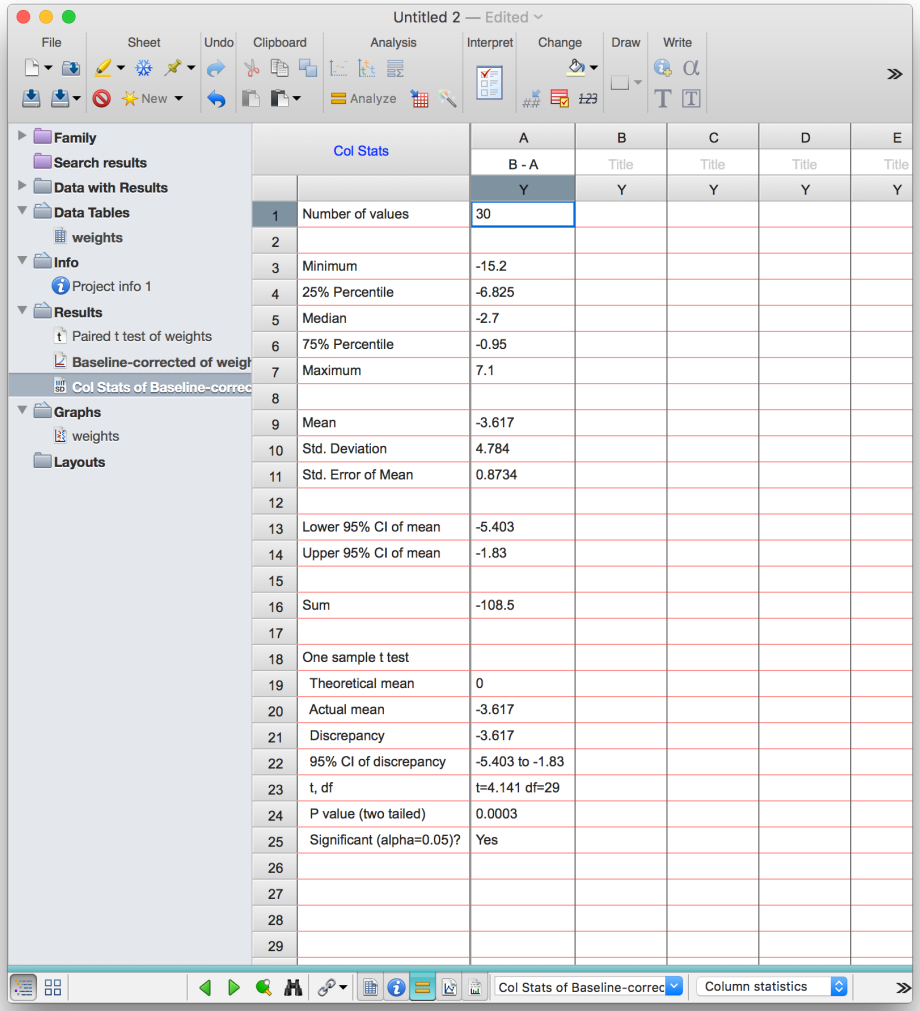

Analyze Data
Built-in analysis
Which analysis?
Transform, Normalize...
Transform
Transform Concentrations (X)
Normalize
Prune rows
Remove baseline and column math
Transpose X and Y
Fraction of Total
XY analyses
Column analyses
t tests (and nonparametric tests)
One-way ANOVA (and nonparametric)
Column statistics
Frequency distribution
ROC Curve
Bland-Altman method comparison
Correlation
Identify outliers
Analyze a stack of P values
Grouped analyses
Contingency table analyses
Survival analyses
Analyze which data sets?
A-A - A
B:B - A
Select All
Deselect All
Cancel
OK
1.Select 'Column statistics' under Column analysis
2.Select only the difference that you are interested in.
3.Click 'OK'

1. Select ‘One-sample t test’ under Inferences




Untitled 2 — Edited
File
Sheet
Undo
Clipboard
Analysis
Interpret
Change
Draw
Write
New
Analyze
Family
Search results
Data with Results
Data Tables
weights
Info
Project info 1
Results
Paired t test of weights
Baseline-corrected of weigh
Col Stats of Baseline-correc
Graphs
weights
Layouts
Col Stats
A
B
C
D
E
B - A
Title
Title
Title
Title
Y
Y
Y
Y
Y
1 Number of values 30
2
3 Minimum -15.2
4 25% Percentile -6.825
5 Median -2.7
6 75% Percentile -0.95
7 Maximum 7.1
8
9 Mean -3.617
10 Std. Deviation 4.784
11 Std. Error of Mean 0.8734
12
13 Lower 95% CI of mean -5.403
14 Upper 95% CI of mean -1.83
15
16 Sum -108.5
17
18 One sample t test
19 Theoretical mean 0
20 Actual mean -3.617
21 Discrepancy -3.617
22 95% CI of discrepancy -5.403 to -1.83
23 t, df t=4.141 df=29
24 P value (two tailed) 0.0003
25 Significant (alpha=0.05)? Yes
26
27
28
29
Col Stats of Baseline-correc
Column statistics


Untitled 2 — Edited
File
Sheet
Undo
Clipboard
Analysis
Interpret
Change
Draw
Write
New
Analyze
Family
Search results
Data Tables
weights
Info
Project info 1
Results
Paired t test of weights
Graphs
weights
Layouts
Paired t test
1 Table Analyzed weights
3 Column B drug
4 vs. vs.
5 Column A placebo
7 Paired t test
8 P value 0.0003
9 P value summary ***
10 Significantly different (P < 0.05)? Yes
11 One- or two-tailed P value? Two-tailed
12 t, df t=4.141 df=29
13 Number of pairs 30
15 How big is the difference?
16 Mean of differences -3.617
17 SD of differences 4.784
18 SEM of differences 0.8734
19 95% confidence interval -5.403 to -1.83
20 R squared (partial eta squared) 0.3716
22 How effective was the pairing?
23 Correlation coefficient (r) 0.8893
24 P value (one tailed) <0.0001
25 P value summary ****
26 Was the pairing significantly effective? Yes
Paired t test of weights
Tabular results

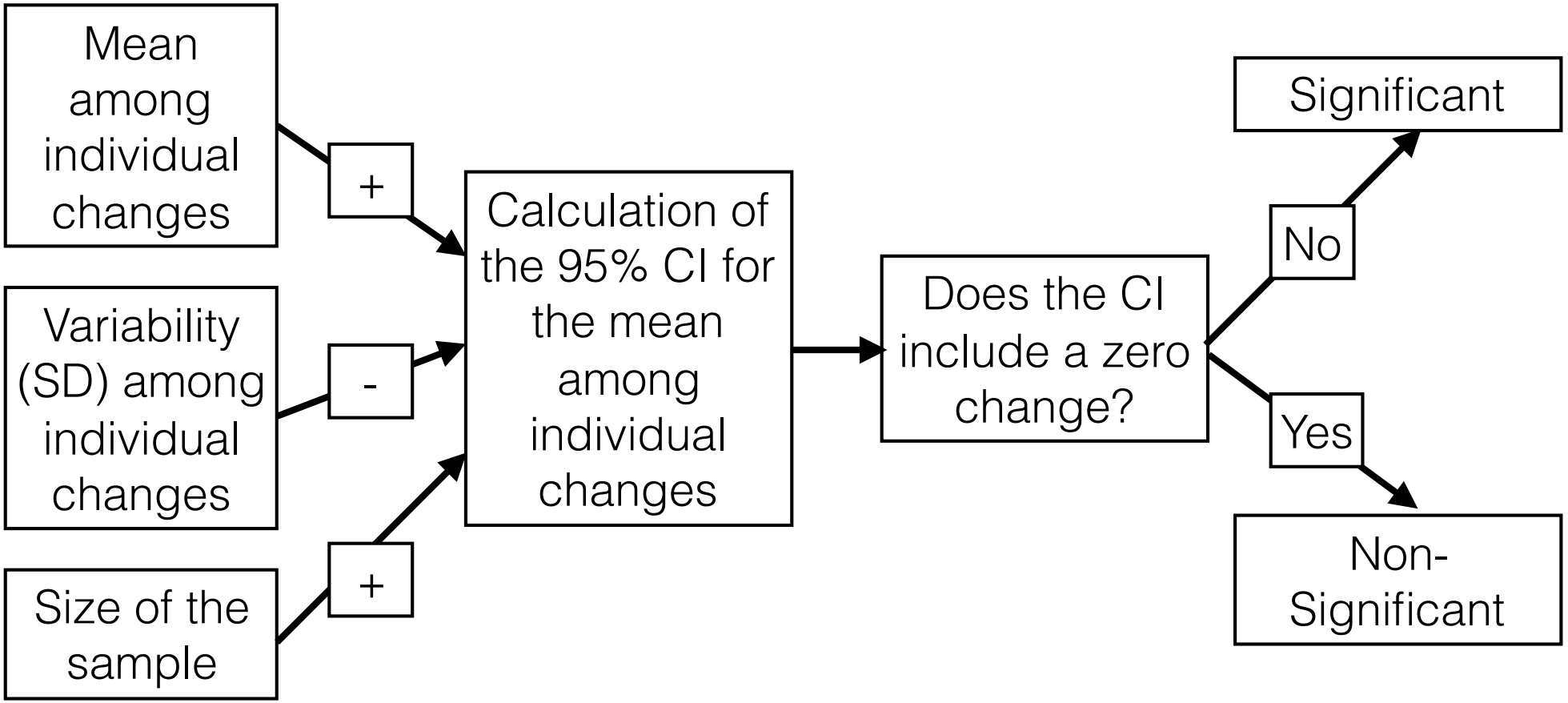

Untitled 2 — Edited
File
Sheet
Undo
Clipboard
Analysis
Interpret
Change
Draw
Write
New
Analyze
Family
Search results
Data with Results
Data Tables
weights
Info
Project info 1
Results
Paired t test of weights
Baseline-corrected of weigh
Col Stats of Baseline-correc
Graphs
weights
Layouts
Col Stats
A B - A
B C D E
Title
Y
1 Number of values 30
3 Minimum -15.2
4 25% Percentile -6.825
5 Median -2.7
6 75% Percentile -0.95
7 Maximum 7.1
9 Mean -3.617
10 Std. Deviation 4.784
11 Std. Error of Mean 0.8734
13 Lower 95% CI of mean -5.403
14 Upper 95% CI of mean -1.83
16 Sum -108.5
18 One sample t test
19 Theoretical mean 0
20 Actual mean -3.617
21 Discrepancy -3.617
22 95% CI of discrepancy -5.403 to -1.83
23 t, df t=4.141 df=29
24 P value (two tailed) 0.0003
25 Significant (alpha=0.05)? Yes
Col Stats of Baseline-correc
Column statistics

WHAT DETERMINES SIGNIFICANCE IN A PAIRED *t*-TEST

What determines whether a paired *t*-test will be significant?

- How far from zero is the mean value for the individual changes?
- How wide is the interval?

Factors influencing the outcome of a paired *t*-test



POWER OF THE PAIRED *t*-TEST

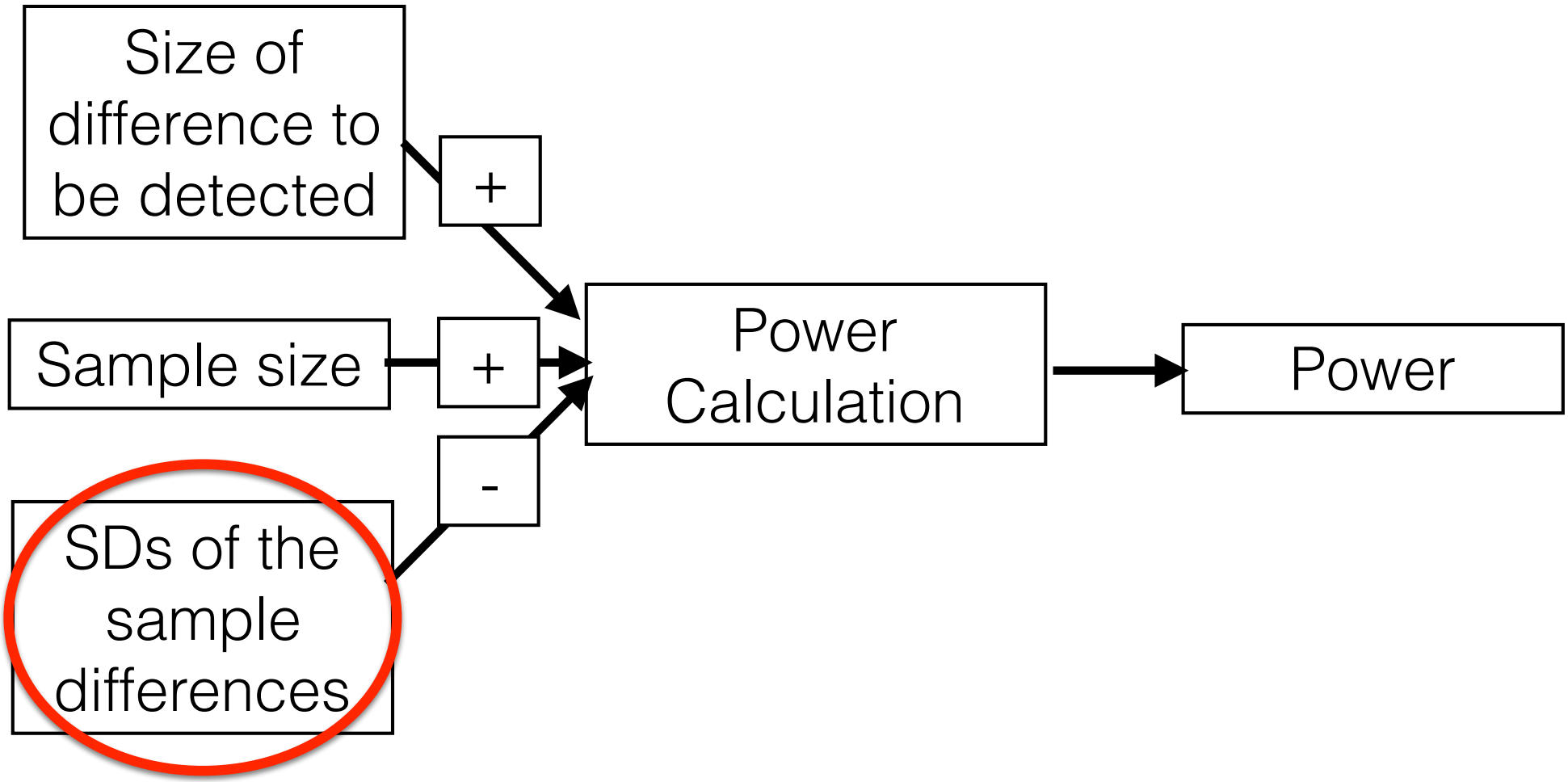
Greater power of the paired *t*-test

Where data form natural pairs, it is often the case that individuals who have the highest values under the first set of experimental condition also have the highest values under the alternative conditions.

As a result, the variability in the two original sets of data is greater than that among the individual changes that occur.

➡ The paired *t*-test is often ***more powerful*** than the two-sample *t*-test.

calculating power



CHOICE OF EXPERIMENTAL DESIGN

Choice of Experimental Design

Pros for paired t-test:

- Greater power

Cons for paired t-test:

- Greater practical difficulties
- Greater problems in the case of data loss
- Paired design may be logistically impossible

REQUIREMENTS FOR APPLYING A PAIRED *t*-TEST

Assumptions of a paired *t*-test

- Samples are randomly drawn from the population
- The individual **changes** must be consistent with a normal distribution (NOT the original observations)

SAMPLE SIZES

Sample size calculations

- **Size of minimum relevant difference**. Still based off of clinical relevance.
- **Variability of the change**. Not the variability of the original values within a group.
- **Power**. Doesn't change between two-sample *t*-test and paired *t*-test.

Typically, we feed these values into a sample size calculated for a one-sample t-test.

	Paired *t*-test	Two-sample *t*-test
Methodology	1. Calculate the difference for each pair of values. 2. Calculate C.I. for the mean of these differences.	1. Calculate the mean for each column. 2. Calculate C.I. for the difference between these means.
Use with unpaired data?	Unjustified and pointless	Correct procedure
Use with paired data?	Correct procedure	Poor choice
Use if unequal numbers of observations in the two columns	Data evidently are likely not to be paired	No problem

What did we learn?

- Paired data occur often in research and we can gain power by accounting for them in our analysis
- In general, the variability of individual differences is less then the variability of the original values
- Paired t-test - differences then mean of differences; two-sample t-test - means then difference of means
- Although paired t-test are more powerful, they aren't always easy to implement.
- Sample size calculations are based on the variability of the individual differences

thought questions

- Can I decide how the subjects should be matched after collecting the data?
- When computing the difference for each pair, in which order is the subtraction done?
- What if one of the values for a pair is missing?
- Can a paired t test be computed if you know only the mean and SD of the two treatments and the number of pairs?
- Can a paired t test be computed if you know only the mean and SD of the set of differences and the number of pairs?
- Does it matter whether the two groups are sampled from populations that are not Gaussian?

example

Does treatment change enzyme activity?

In enzyme.csv is example data that test whether treating cultured cells with a drug increases the activity of an enzyme. Ten different clones of the cell were tested. With each clone, control cells and treated cells were tested side by side.