Chapter 14: Analyses of variance: Going beyond *t*-tests

TXCL7565/PHSC7565

What This Chapter Covers

- Factors and levels
- One-way analysis of variance
- Two-way analysis of variance
- Fixed and random factors

When we want to compare interval scale data between more than 2 groups, we use an **Analysis of Variance** (**ANOVA**).

FACTORS AND LEVELS

Factors and levels

- ***Factor*** - something that can take two or more categorical values.
  - either spontaneously varies or can be manipulated as part of an experiment
  - want to know if the different values of the factor cause different values of the outcome
- ***Level*** - individual values of a factor

Sex is a factor.

'Male' and 'female' are values of that factor.

ONE-WAY ANALYSIS OF VARIANCE

One-way ANOVA

One-way ANOVA - the independent (predictor) variable is single experimental factor with multiple levels and the dependent (outcome) variable is an interval variable

Table 14.1 Effect catalyst on yield (percentage of theoretical maximum)

Platinum	Palladium	Iridium	Palladium/ Iridium	Rhodium
11.3	15.4	12.1	13.1	12.0
10.7	17.0	12.2	13.7	11.6
9.8	18.4	13.1	13.5	9.1
10.4	17.5	11.8	14.0	11.9
11.5	18.8	10.4	14.2	11.3

Multiple *t*-tests. vs ANVOA

Could do several t-tests for all the comparisons that we are interested in…

- Multiple testing
    - For each comparison there is a 5% risk of a false positive. If there are 5 comparisons, the risk of at least one producing a significant results when none are truly different is …
    - (probability of not making a mistake)^5

$$1 - (1 - \alpha)^5 = 1 - 0.95^5 = 0.23$$

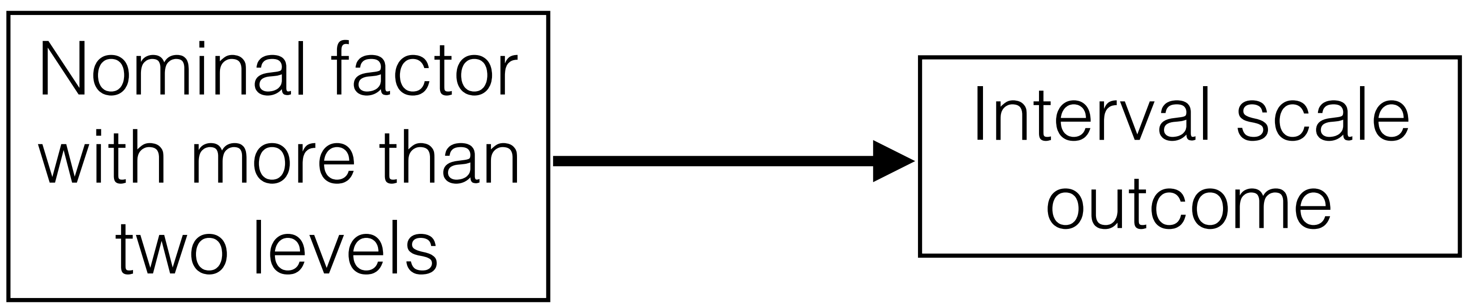
- Estimate of SD
    - In t-test, estimating within group SD from 10 samples. In ANOVA estimating within group SD from 25 samples.

Two-sample t-test vs. One-way ANOVA

- ***Two sample t-test*** - Is the difference between these two samples means greater than can reasonably be accounted for by random sampling error?

- ***One-way ANOVA*** - Are the differences among these various sample means greater than can reasonably be accounted for by random sampling error?

Diagrammatic representation of one-way ANOVA



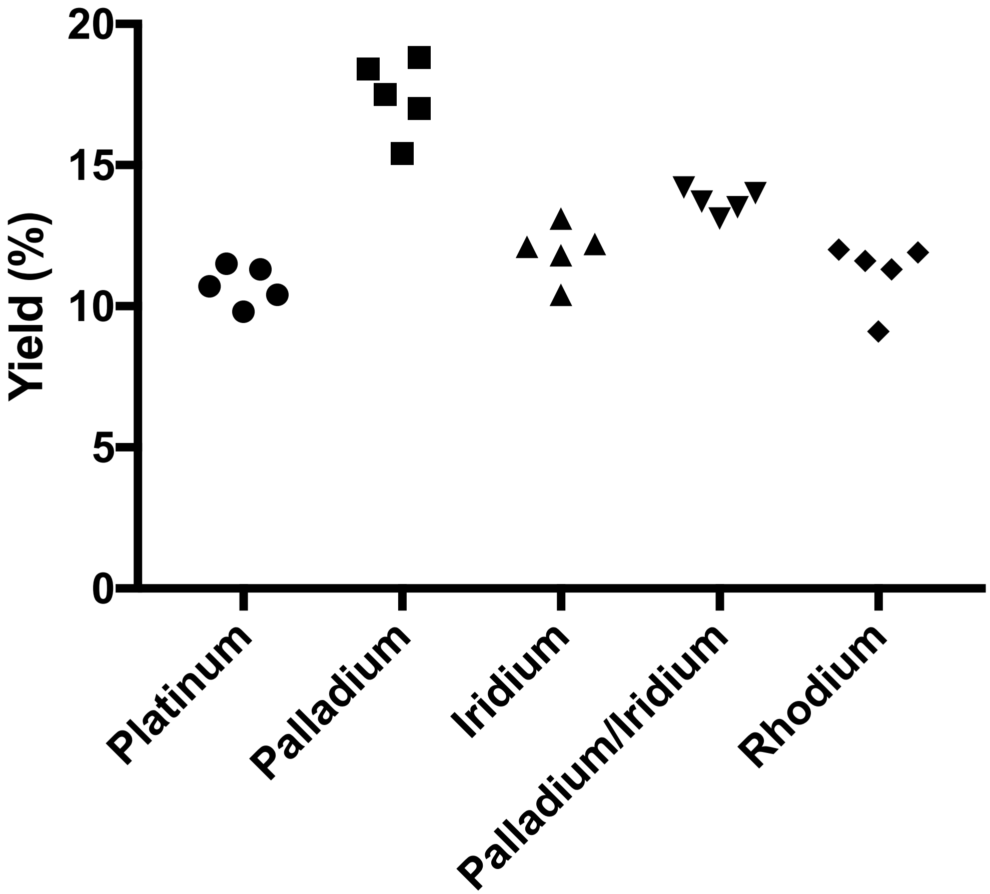
Null and alternative hypothesis

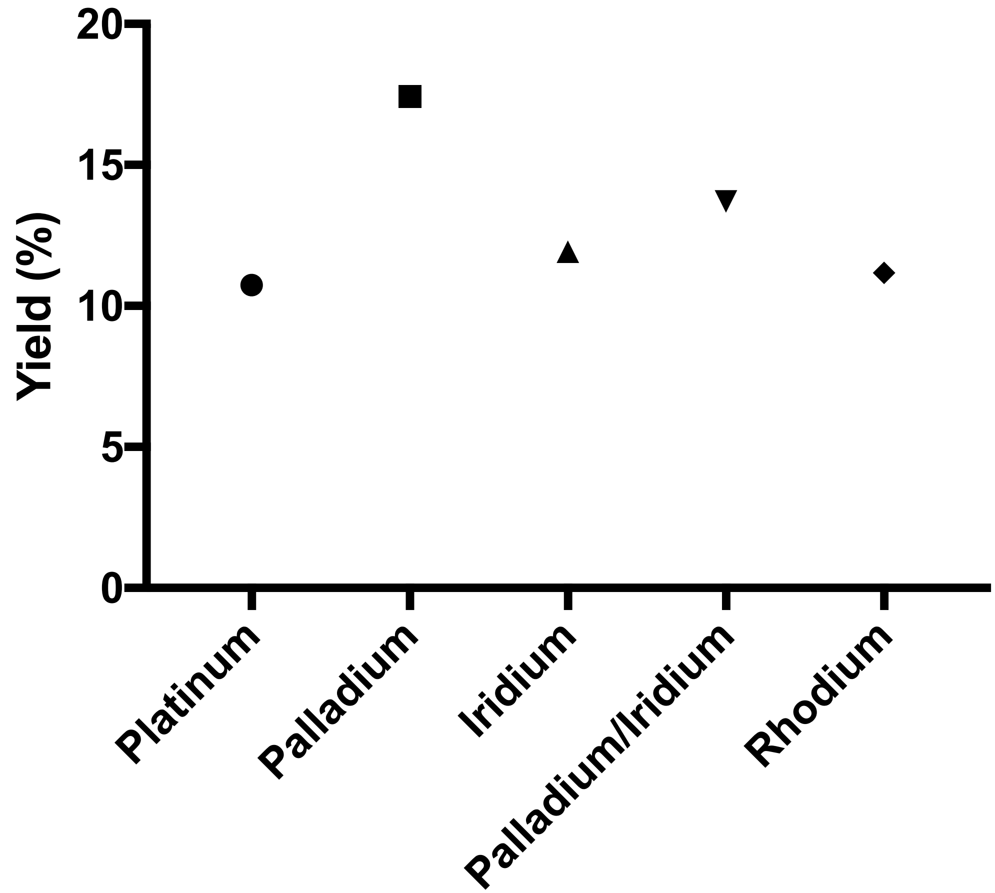
- ***Null hypothesis*** - the population means for all levels of the factor are the same

- ***Alternative hypothesis*** - at least one of the population means for a level of the factor is different from the rest of the population means for the remaining factors

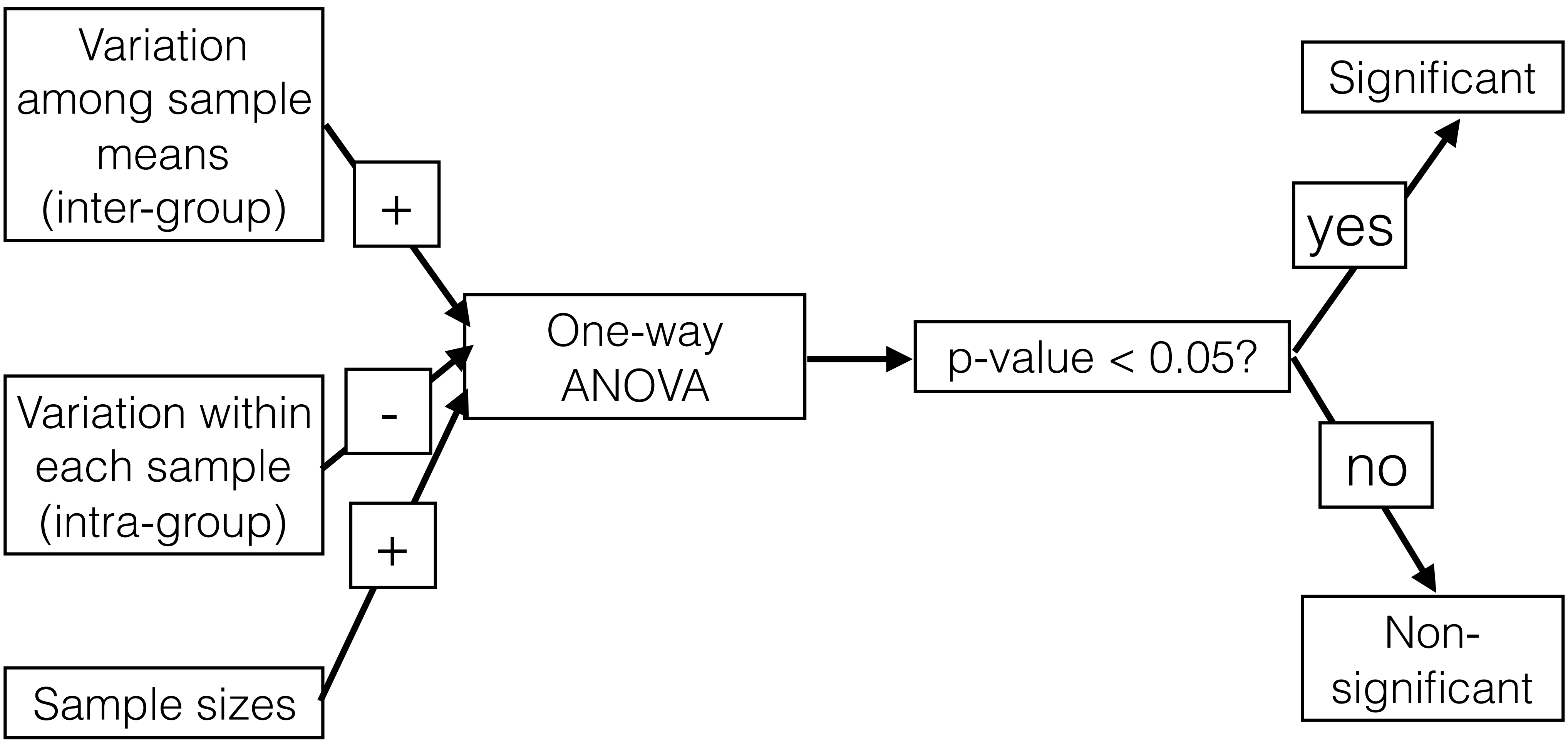
$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n$$
$$H_a : \text{at least one } \mu \text{ is different from the rest}$$

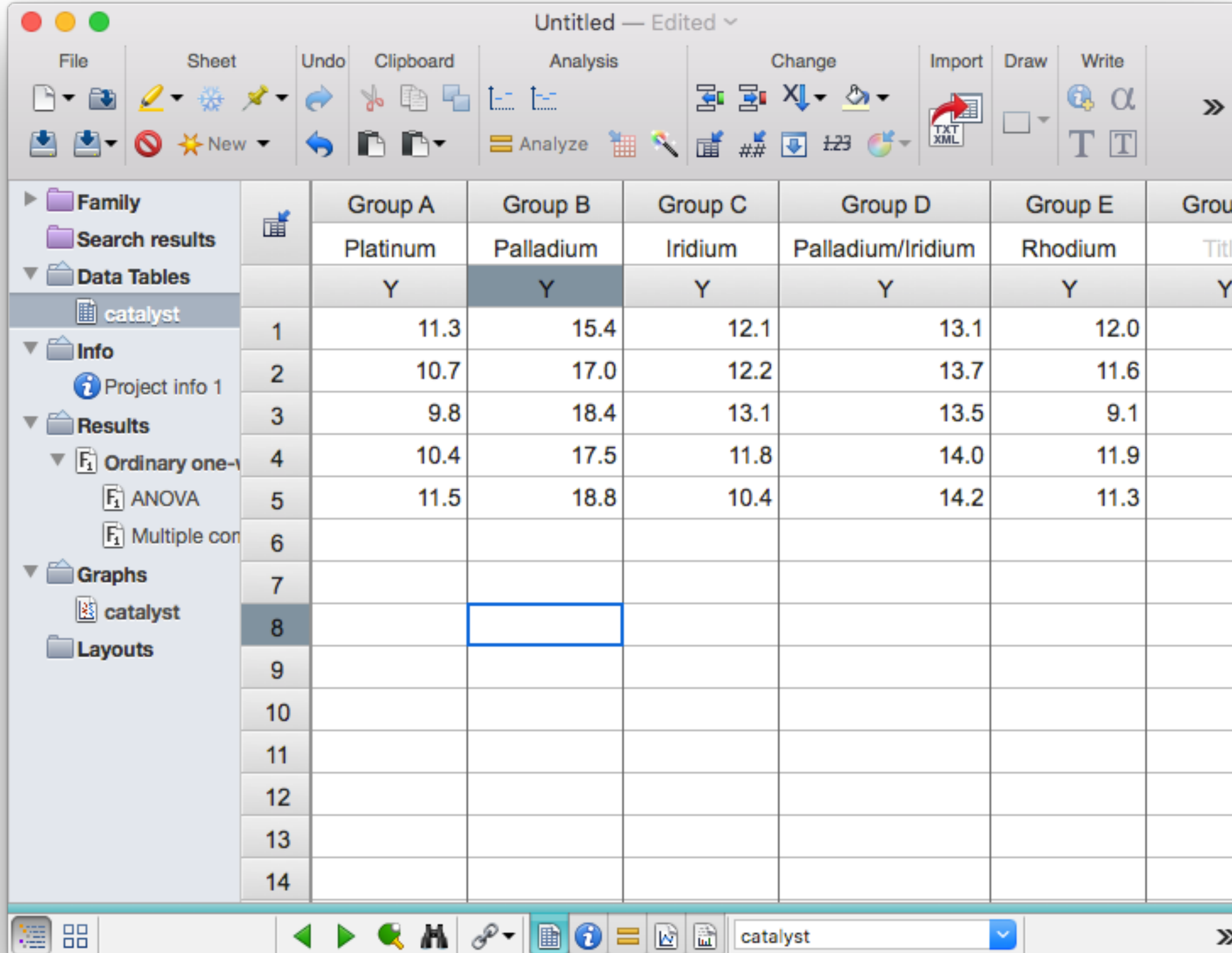
Intra-group variation - variation within a group of replicates
Inter-group variation - variation between groups





What governs whether a significant result will emerge?

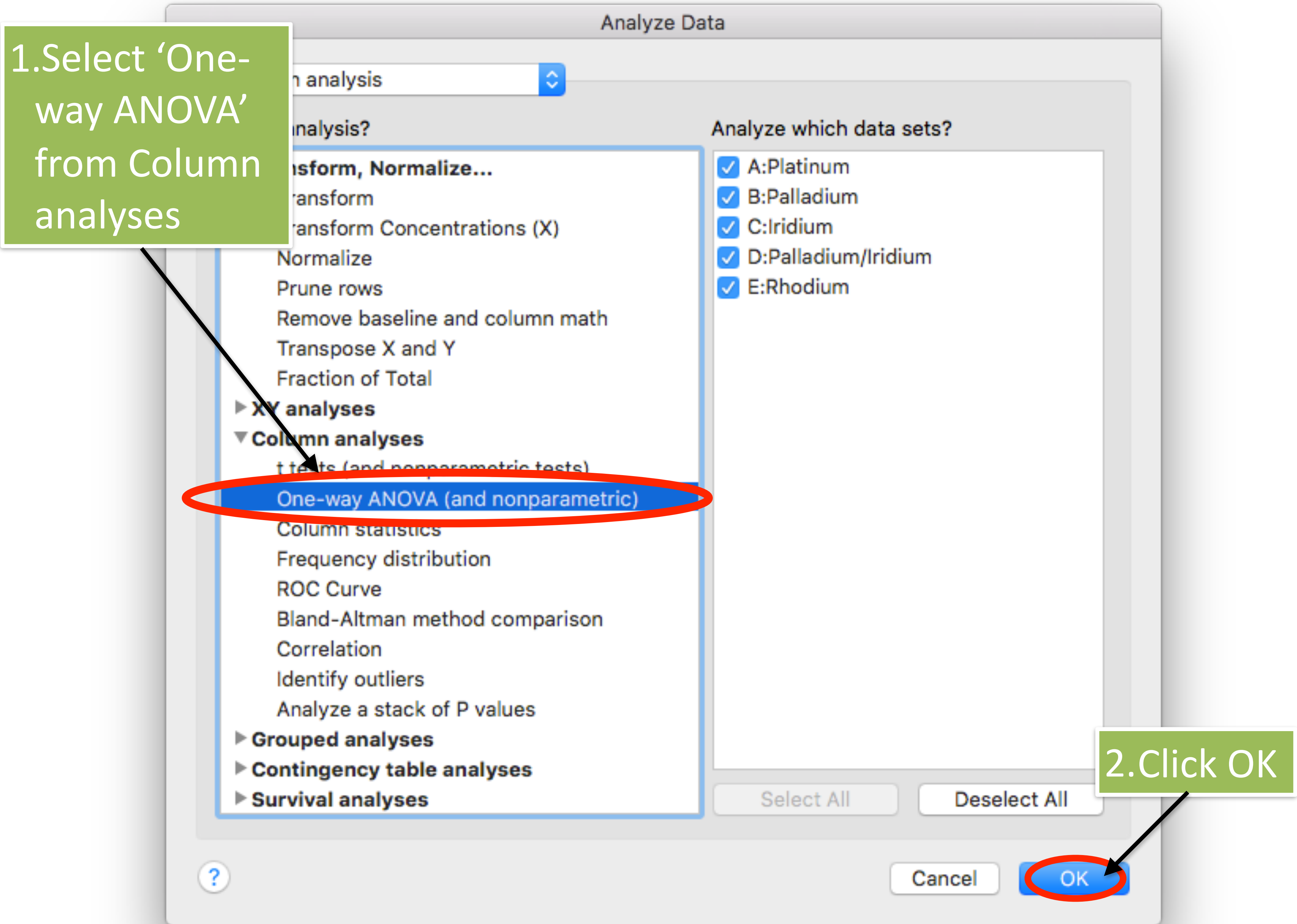




Untitled — Edited
File
Sheet
Undo
Clipboard
Analysis
Change
Import
Draw
Write
New
Analyze
TXT XML
Family
Search results
Data Tables
catalyst
Info
Project info 1
Results
Ordinary one-v
ANOVA
Multiple con
Graphs
catalyst
Layouts

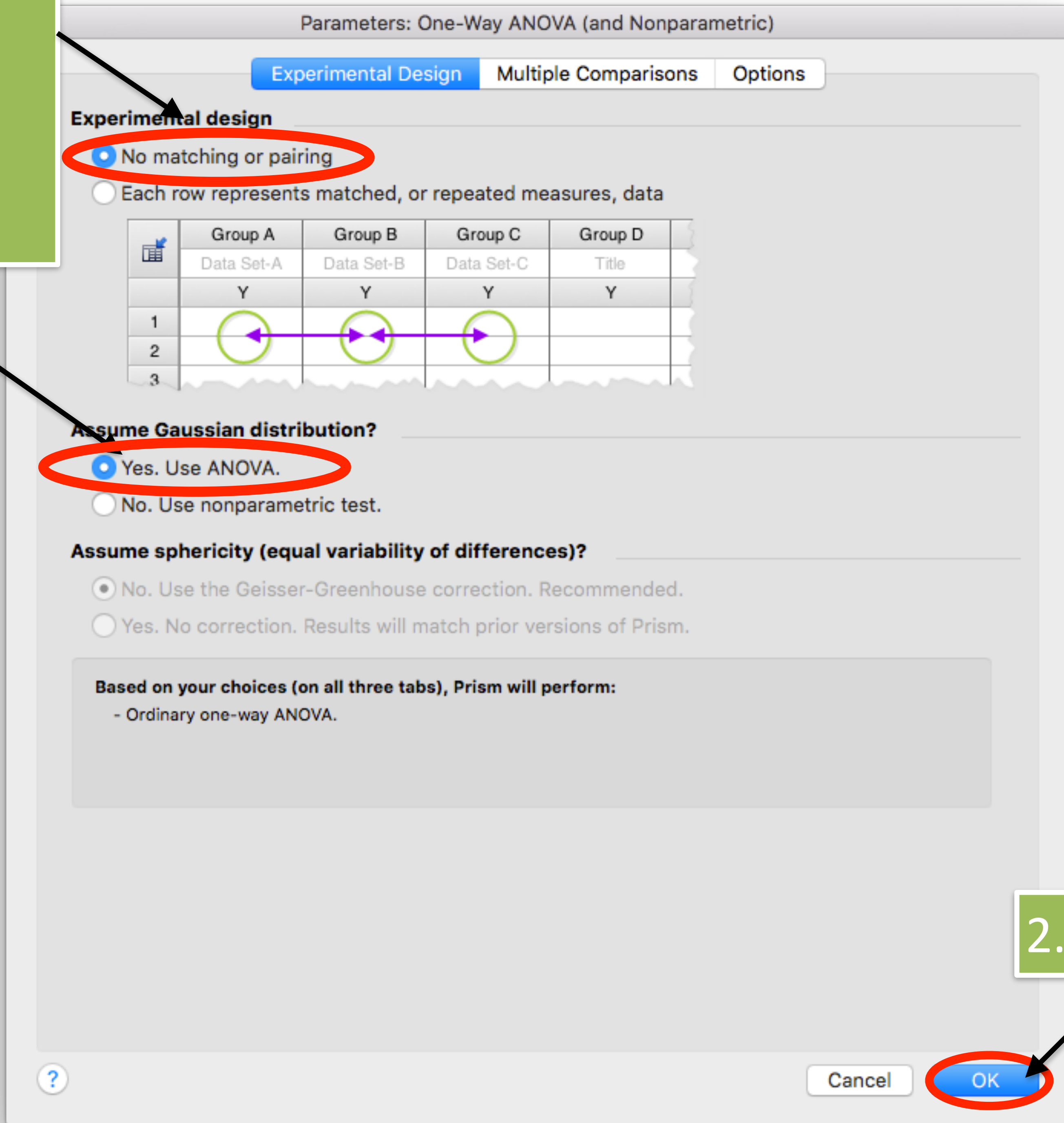
	Group A	Group B	Group C	Group D	Group E	Grou
	Platinum	Palladium	Iridium	Palladium/Iridium	Rhodium	Title
	Y	Y	Y	Y	Y	Y
1	11.3	15.4	12.1	13.1	12.0	
2	10.7	17.0	12.2	13.7	11.6	
3	9.8	18.4	13.1	13.5	9.1	
4	10.4	17.5	11.8	14.0	11.9	
5	11.5	18.8	10.4	14.2	11.3	
6						
7						
8						
9						
10						
11						
12						
13						
14						

catalyst

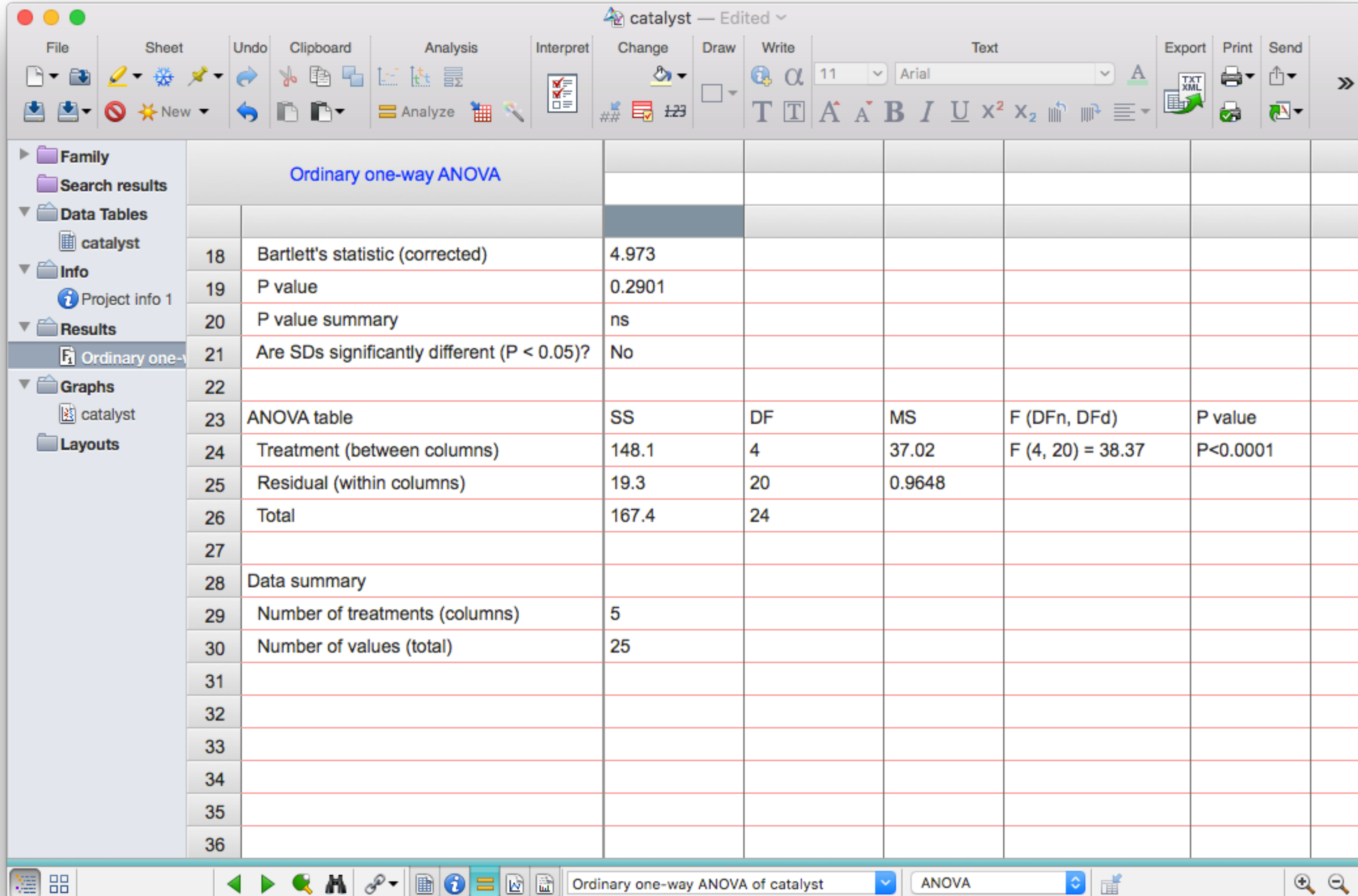


Analyze Data
1.Select 'One-way ANOVA' from Column analyses
n analysis
nalysis?
nsform, Normalize...
ransform
ransform Concentrations (X)
Normalize
Prune rows
Remove baseline and column math
Transpose X and Y
Fraction of Total
XY analyses
Column analyses
t tests (and nonparametric tests)
One-way ANOVA (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:Platinum
B:Palladium
C:Iridium
D:Palladium/Iridium
E:Rhodium
Select All
Deselect All
2.Click OK
?
Cancel
OK

1.Select ‘No matching or pairing’ and ‘Yes. Use ANOVA’



2.Click OK



18	Bartlett's statistic (corrected)	4.973				
19	P value	0.2901				
20	P value summary	ns				
21	Are SDs significantly different (P < 0.05)?	No				
22						
23	ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
24	Treatment (between columns)	148.1	4	37.02	F (4, 20) = 38.37	P<0.0001
25	Residual (within columns)	19.3	20	0.9648		
26	Total	167.4	24			
27						
28	Data summary					
29	Number of treatments (columns)	5				
30	Number of values (total)	25				
31						
32						
33						
34						
35						
36						

ANOVA Table

ANOVA Table	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	148.1	4	37.02	F (4,20) = 38.37	<0.0001
Residual (within columns)	19.3	20	0.9648		
Total	167.4	24			

Follow up tests - Interpreting significance

- p-value < 0.05 indicates that we have evidence against the null hypothesis
- don't know the nature of the differences

FOLLOW-UP TESTS

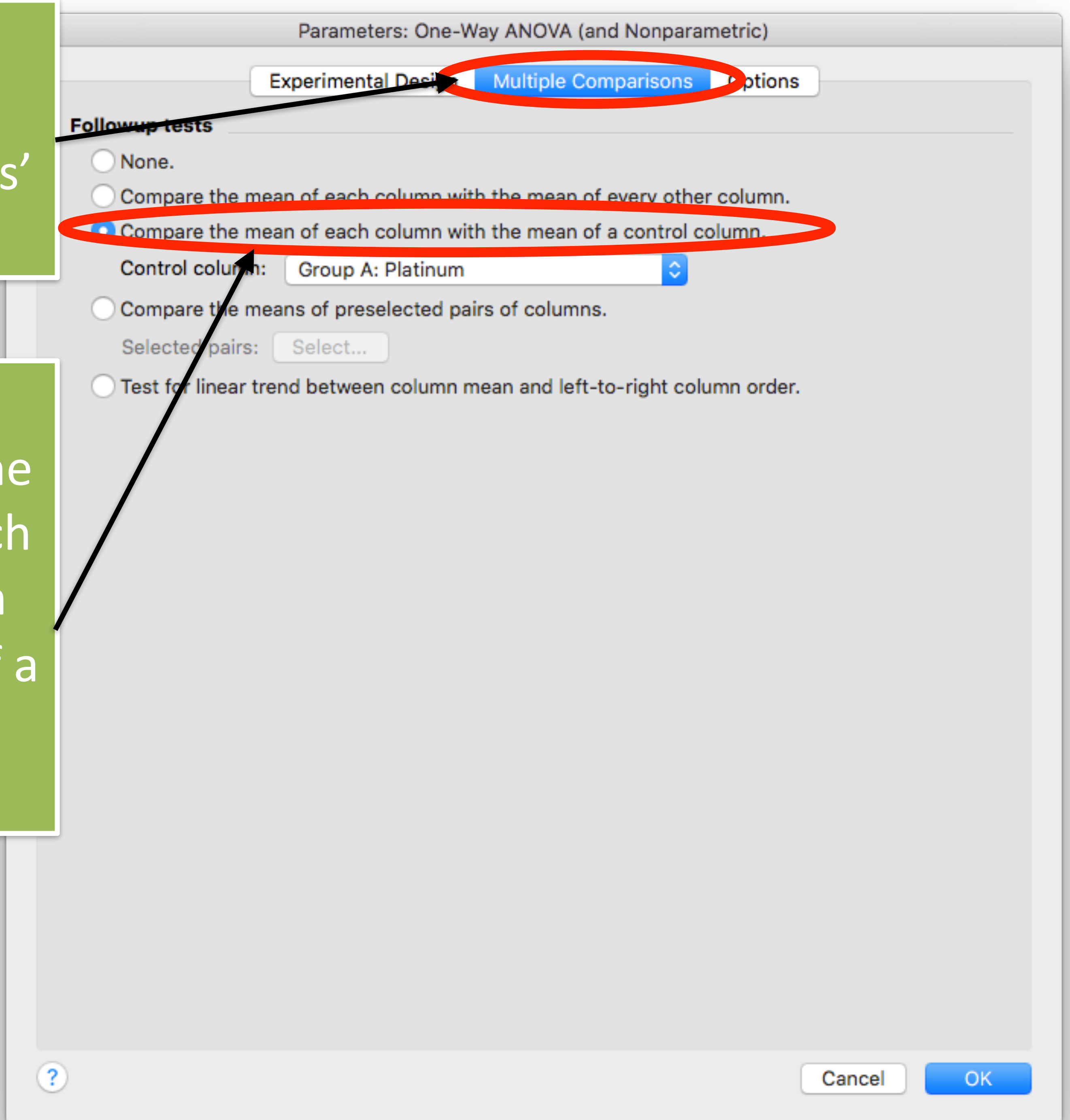
Dunnett’s test

- Select one treatment as a ‘control’ or ‘reference’ group. All other groups are then compared against the control.

- Confidence intervals and p-values are automatically adjusted to control ‘test-wide error rate’ to 5%

1.Select the ‘Multiple Comparisons’ tab

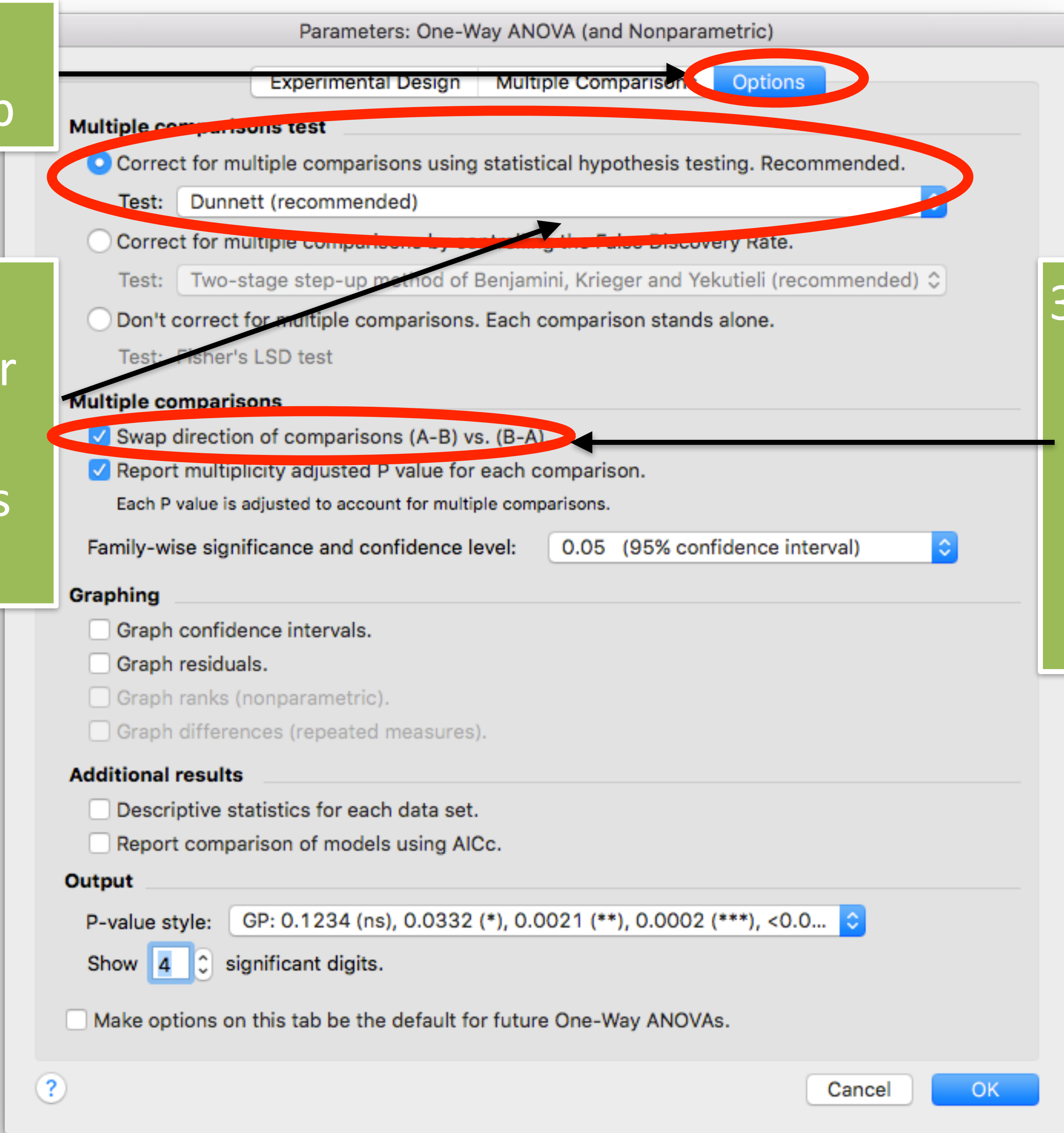
2.Select the ‘Compare the mean of each column with the mean of a control column’

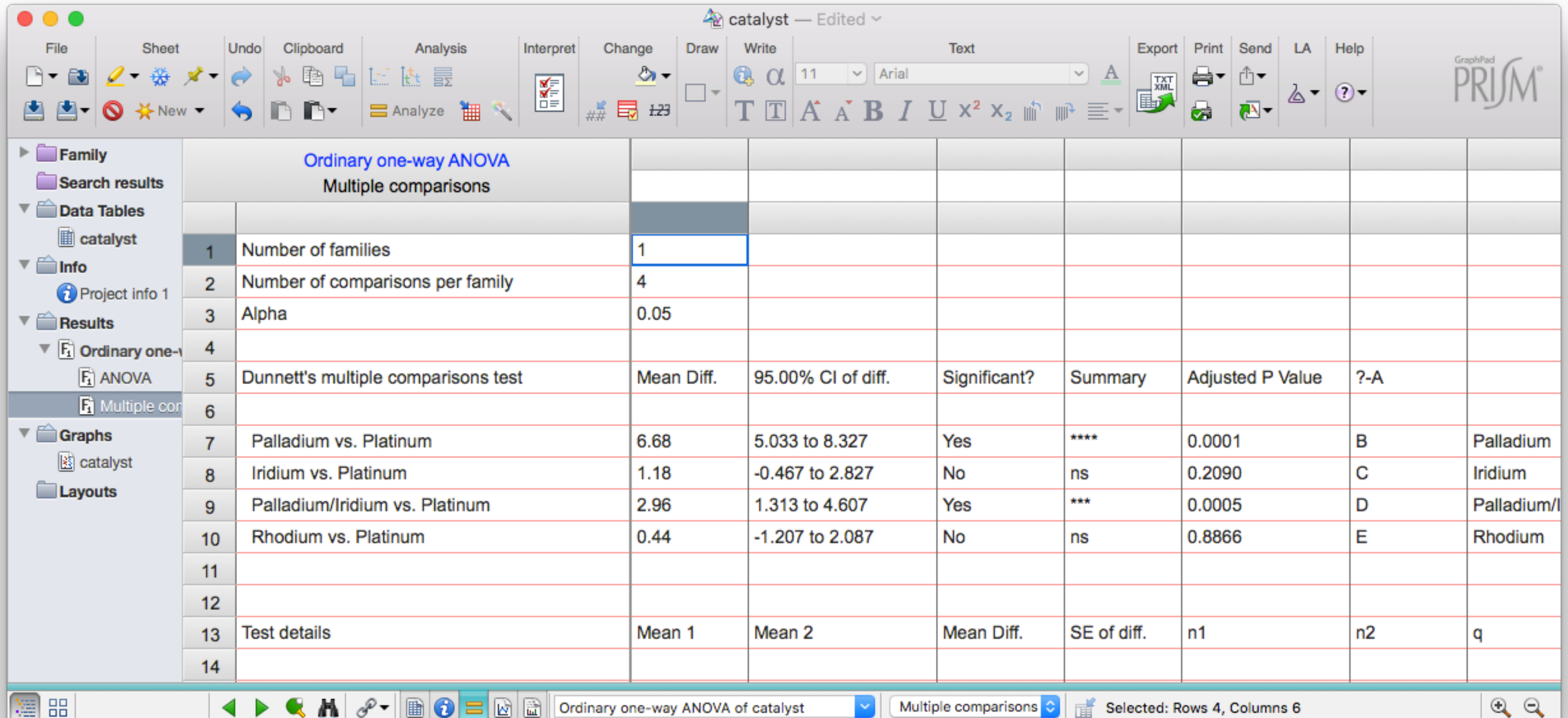


1.Select the 'Options' tab

2.Select 'Dunnett' for multiple comparisons test

3.Swap direction of comparison to match Table 14.5 results





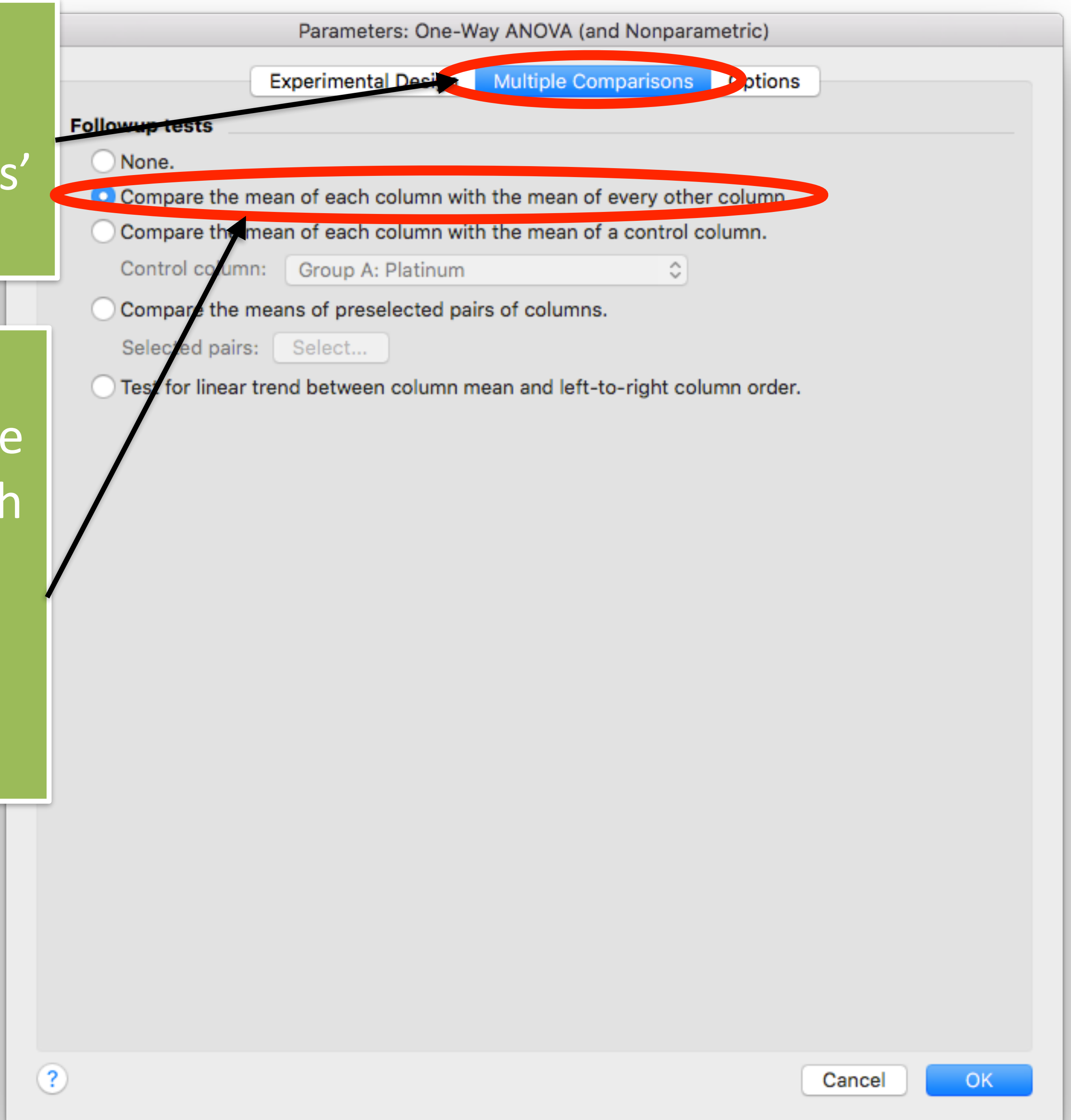
	Ordinary one-way ANOVA Multiple comparisons							
1	Number of families	1						
2	Number of comparisons per family	4						
3	Alpha	0.05						
4								
5	Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P Value	?-A	
6								
7	Palladium vs. Platinum	6.68	5.033 to 8.327	Yes	****	0.0001	B	Palladium
8	Iridium vs. Platinum	1.18	-0.467 to 2.827	No	ns	0.2090	C	Iridium
9	Palladium/Iridium vs. Platinum	2.96	1.313 to 4.607	Yes	***	0.0005	D	Palladium/I
10	Rhodium vs. Platinum	0.44	-1.207 to 2.087	No	ns	0.8866	E	Rhodium
11								
12								
13	Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	q
14								

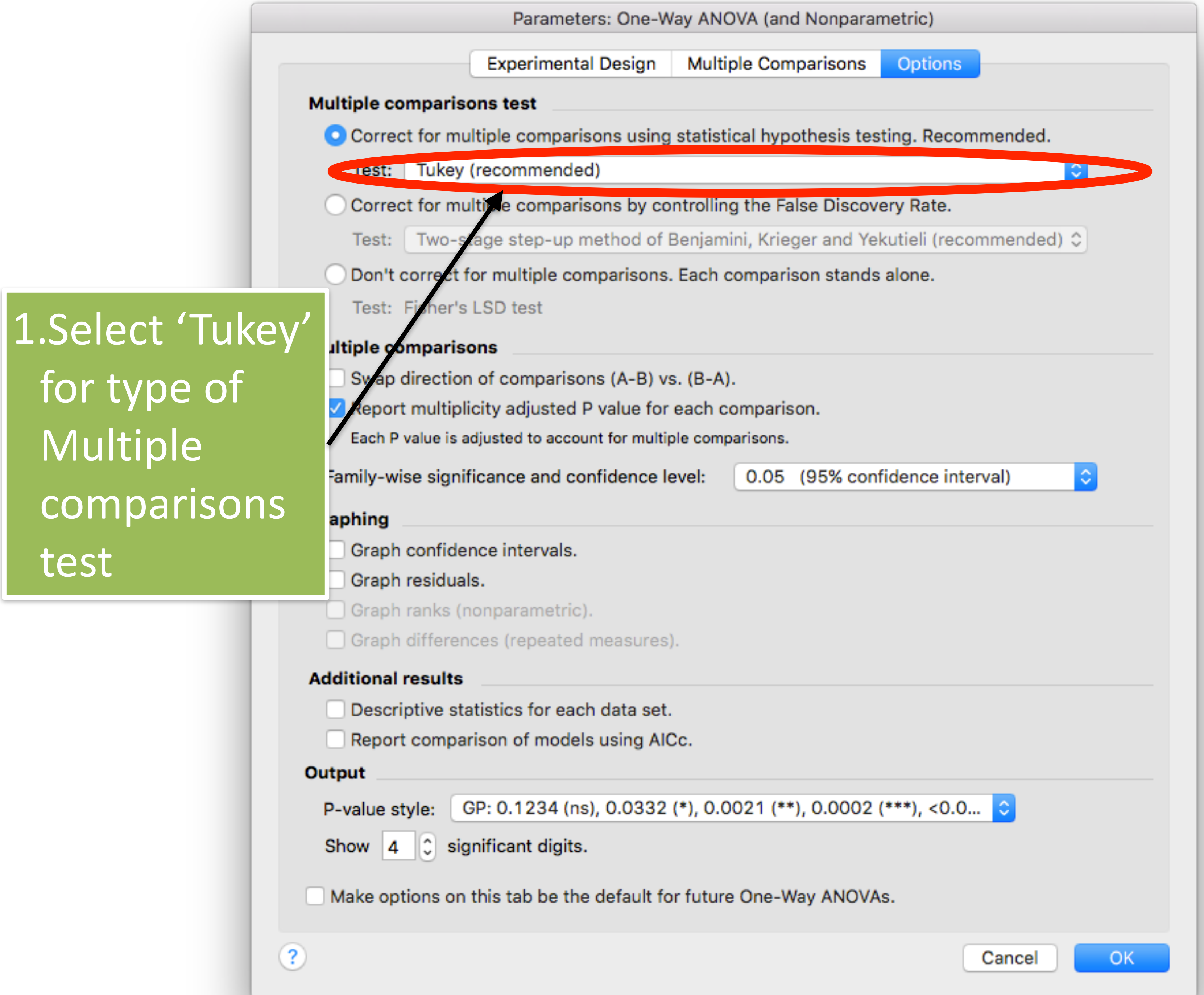
Tukey's test

- All pairwise comparisons between levels are made.
- Confidence intervals and p-values are automatically adjusted to control 'test-wide error rate' to 5%

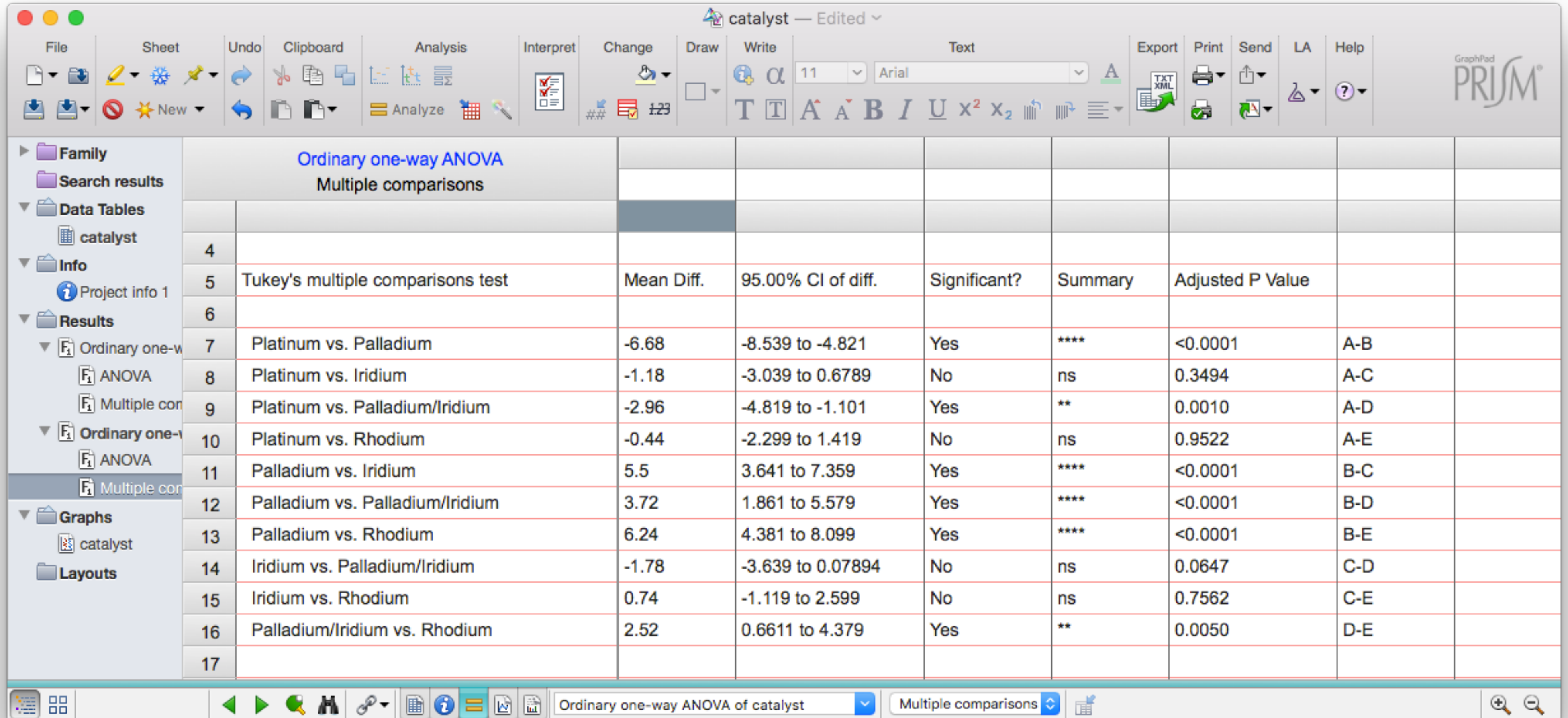
1. Select the 'Multiple Comparisons' tab

2. Select the 'Compare the mean of each column with the mean of every other column'





Parameters: One-Way ANOVA (and Nonparametric)
Experimental Design
Multiple Comparisons
Options
Multiple comparisons test
Correct for multiple comparisons using statistical hypothesis testing. Recommended.
Test: Tukey (recommended)
Correct for multiple comparisons by controlling the False Discovery Rate.
Test: Two-stage step-up method of Benjamini, Krieger and Yekutieli (recommended)
Don't correct for multiple comparisons. Each comparison stands alone.
Test: Fisher's LSD test
Multiple comparisons
Swap direction of comparisons (A-B) vs. (B-A).
Report multiplicity adjusted P value for each comparison.
Each P value is adjusted to account for multiple comparisons.
Family-wise significance and confidence level: 0.05 (95% confidence interval)
Graphing
Graph confidence intervals.
Graph residuals.
Graph ranks (nonparametric).
Graph differences (repeated measures).
Additional results
Descriptive statistics for each data set.
Report comparison of models using AICc.
Output
P-value style: GP: 0.1234 (ns), 0.0332 (*), 0.0021 (**), 0.0002 (***), <0.0...
Show 4 significant digits.
Make options on this tab be the default for future One-Way ANOVAs.
?
Cancel
OK
1.Select 'Tukey' for type of Multiple comparisons test



Ordinary one-way ANOVA
Multiple comparisons

4							
5	Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P Value	
6							
7	Platinum vs. Palladium	-6.68	-8.539 to -4.821	Yes	****	<0.0001	A-B
8	Platinum vs. Iridium	-1.18	-3.039 to 0.6789	No	ns	0.3494	A-C
9	Platinum vs. Palladium/Iridium	-2.96	-4.819 to -1.101	Yes	**	0.0010	A-D
10	Platinum vs. Rhodium	-0.44	-2.299 to 1.419	No	ns	0.9522	A-E
11	Palladium vs. Iridium	5.5	3.641 to 7.359	Yes	****	<0.0001	B-C
12	Palladium vs. Palladium/Iridium	3.72	1.861 to 5.579	Yes	****	<0.0001	B-D
13	Palladium vs. Rhodium	6.24	4.381 to 8.099	Yes	****	<0.0001	B-E
14	Iridium vs. Palladium/Iridium	-1.78	-3.639 to 0.07894	No	ns	0.0647	C-D
15	Iridium vs. Rhodium	0.74	-1.119 to 2.599	No	ns	0.7562	C-E
16	Palladium/Iridium vs. Rhodium	2.52	0.6611 to 4.379	Yes	**	0.0050	D-E
17							

Balanced data

- ***Balanced data*** - the same number of sample in each group

- The power of the ANOVA will be greatest with a balanced data set

- The power for Tukey's tests will be greatest with a balanced data set.

- The only circumstance where power will be greater with an imbalanced data set is where a Dunnett's test is planned (more samples in 'control' group = more power).

Requirements for performing analysis of variance

- Randomly sampled from a population
- Normally distributed values within a level
- Equal SD in each group

Assumption of equal SD

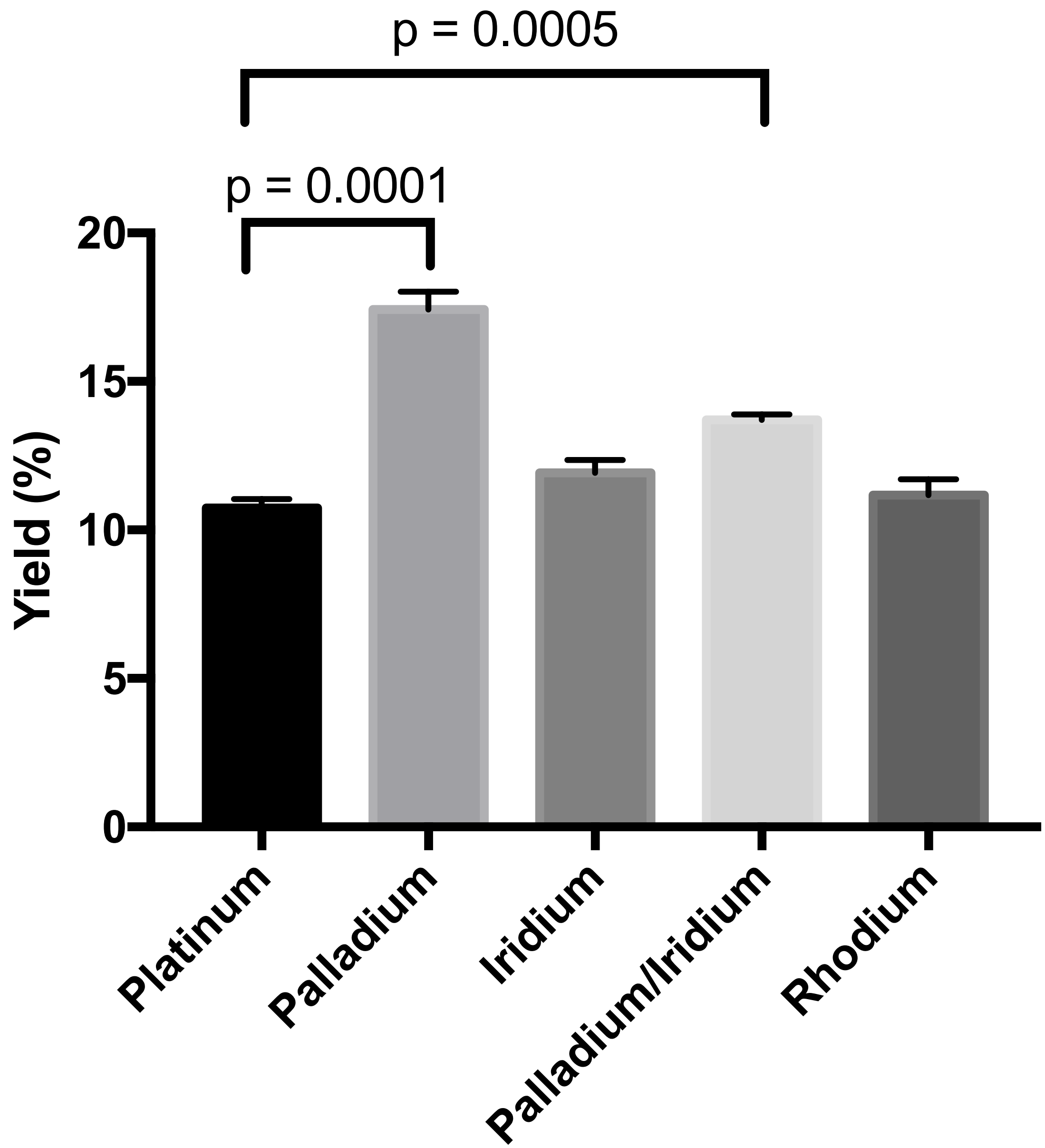
- Brown-Forsythe Test - statistical test for differences in SD between groups

  - $p<0.05$ indicates the intra-group variances are significantly different

- Bartlett's Test - statistical test for differences in SD between groups; sensitive to non-normality

  - $p<0.05$ indicates the intra-group variances are significantly different and/or data are NOT normally distributed

Common graphics for ANOVA



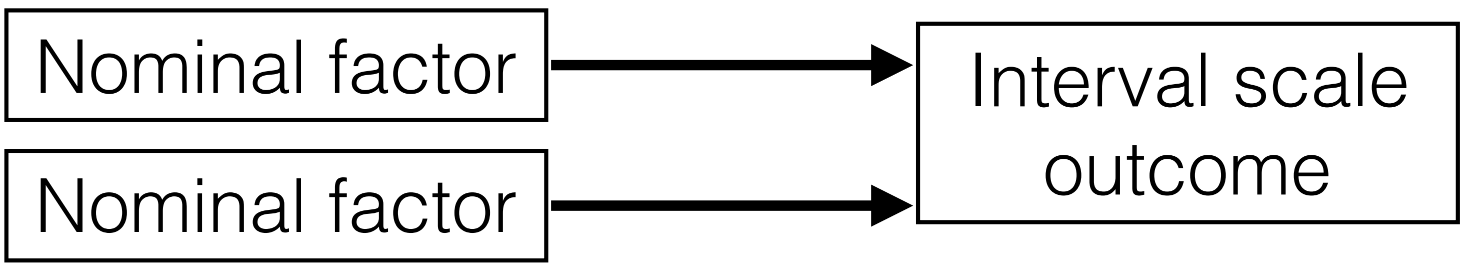
TWO-WAY ANALYSIS OF VARIANCE

Investigating two experimental factors simultaneously

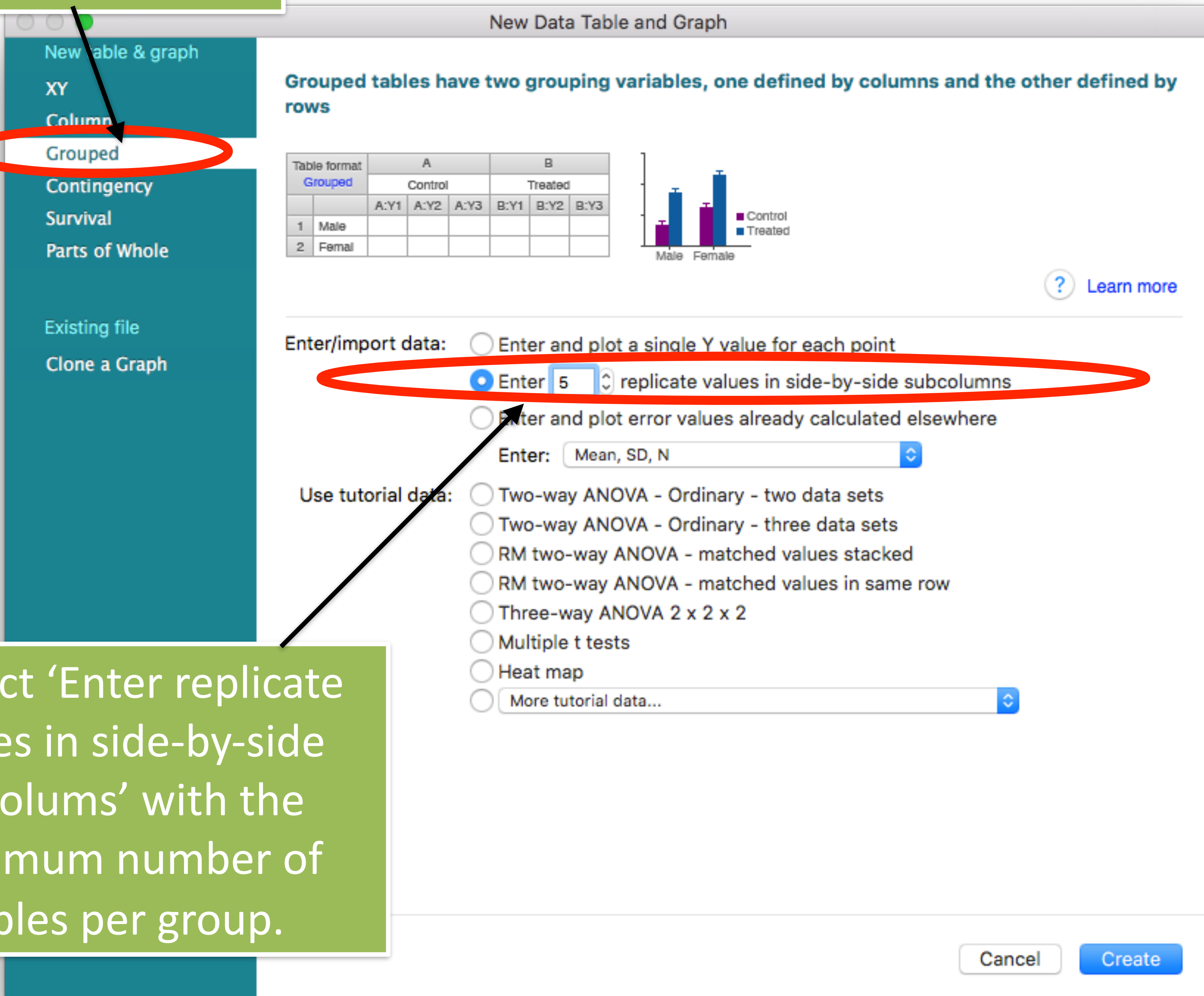
Two-way analysis of variance is used when:

- The endpoint is a measured value (generally on the interval scale)
- There are two experimental factors
- All possible combinations of the two factors have been studied (full factorial experiment)

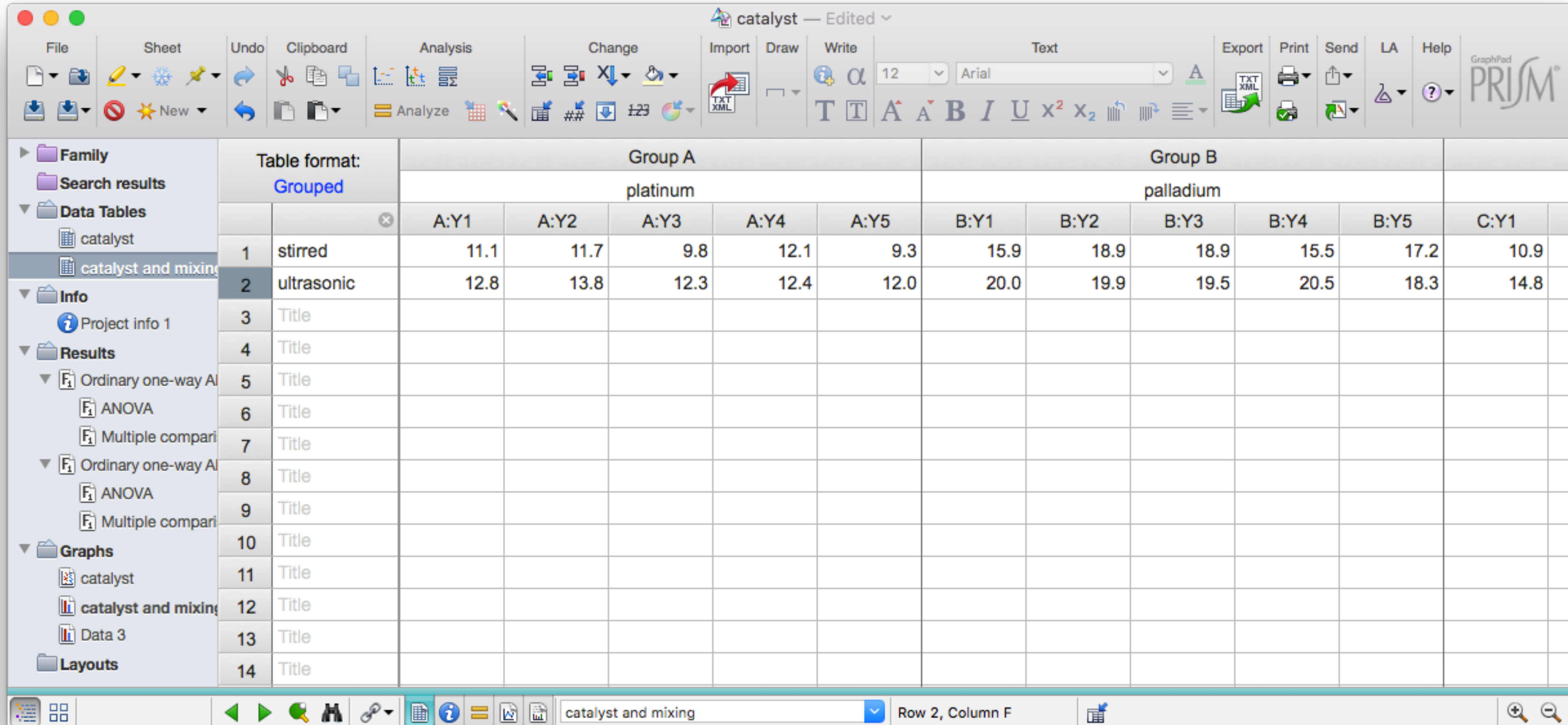
Diagrammatic representation of two-way ANOVA



1.Select ‘Grouped’ for new table.



2.Select ‘Enter replicate values in side-by-side subcolums’ with the maximum number of samples per group.



	Group A					Group B					
	platinum					palladium					
	A:Y1	A:Y2	A:Y3	A:Y4	A:Y5	B:Y1	B:Y2	B:Y3	B:Y4	B:Y5	C:Y1
1 stirred	11.1	11.7	9.8	12.1	9.3	15.9	18.9	18.9	15.5	17.2	10.9
2 ultrasonic	12.8	13.8	12.3	12.4	12.0	20.0	19.9	19.5	20.5	18.3	14.8

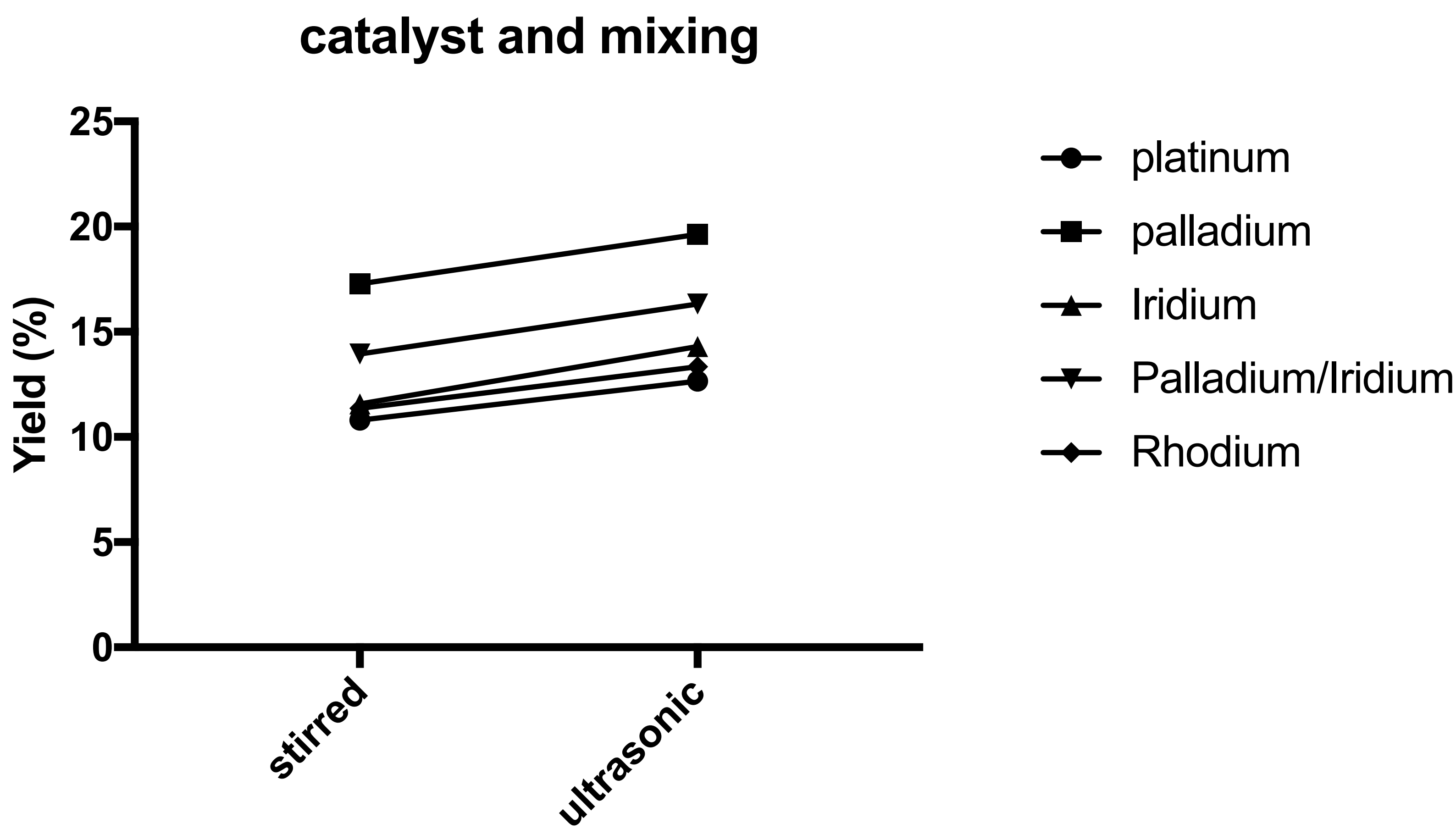
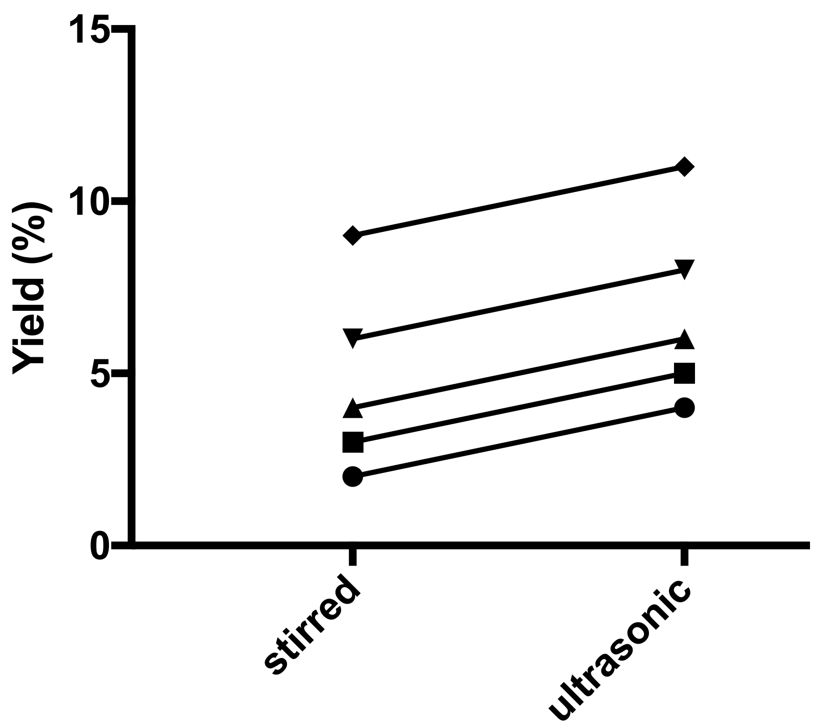


Table 14.8 Mean yield for each combination of catalyst and stirring method (%)

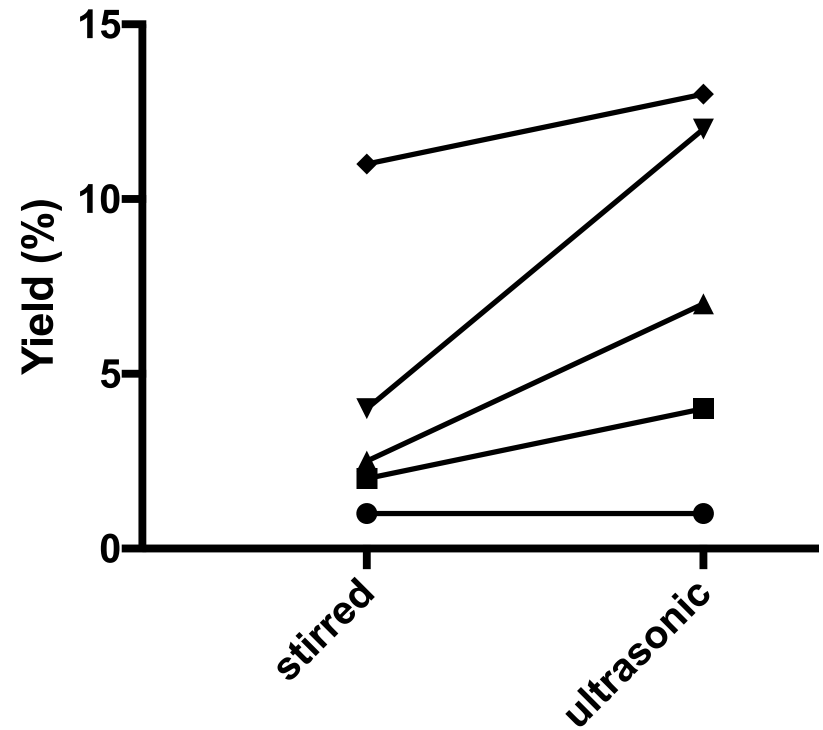
	Pt	Pd	IR	Pd/Ir	Rh
Stirred	10.80	17.28	11.58	13.94	11.36
Ultrasonic	12.66	19.64	14.30	16.32	13.34

Interactions

Interaction is present when the effect produced by changing the level of one factor is dependent upon the level of another factor.



Parallel lines =
No interaction



Non-parallel lines =
Interaction present

Null hypotheses

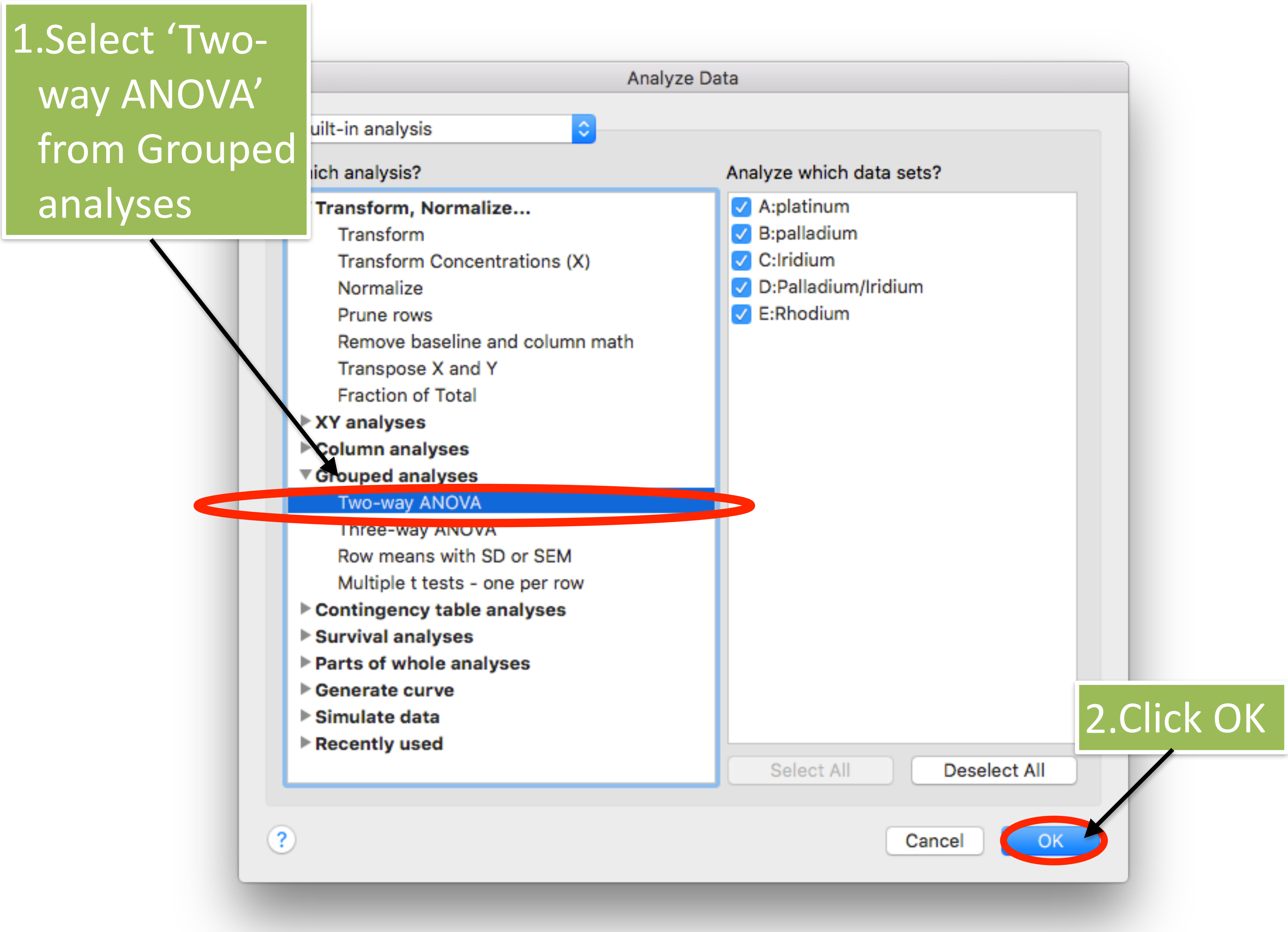
In this case, we will be testing three null hypotheses:

1. Population means among the levels of the first factors are the same

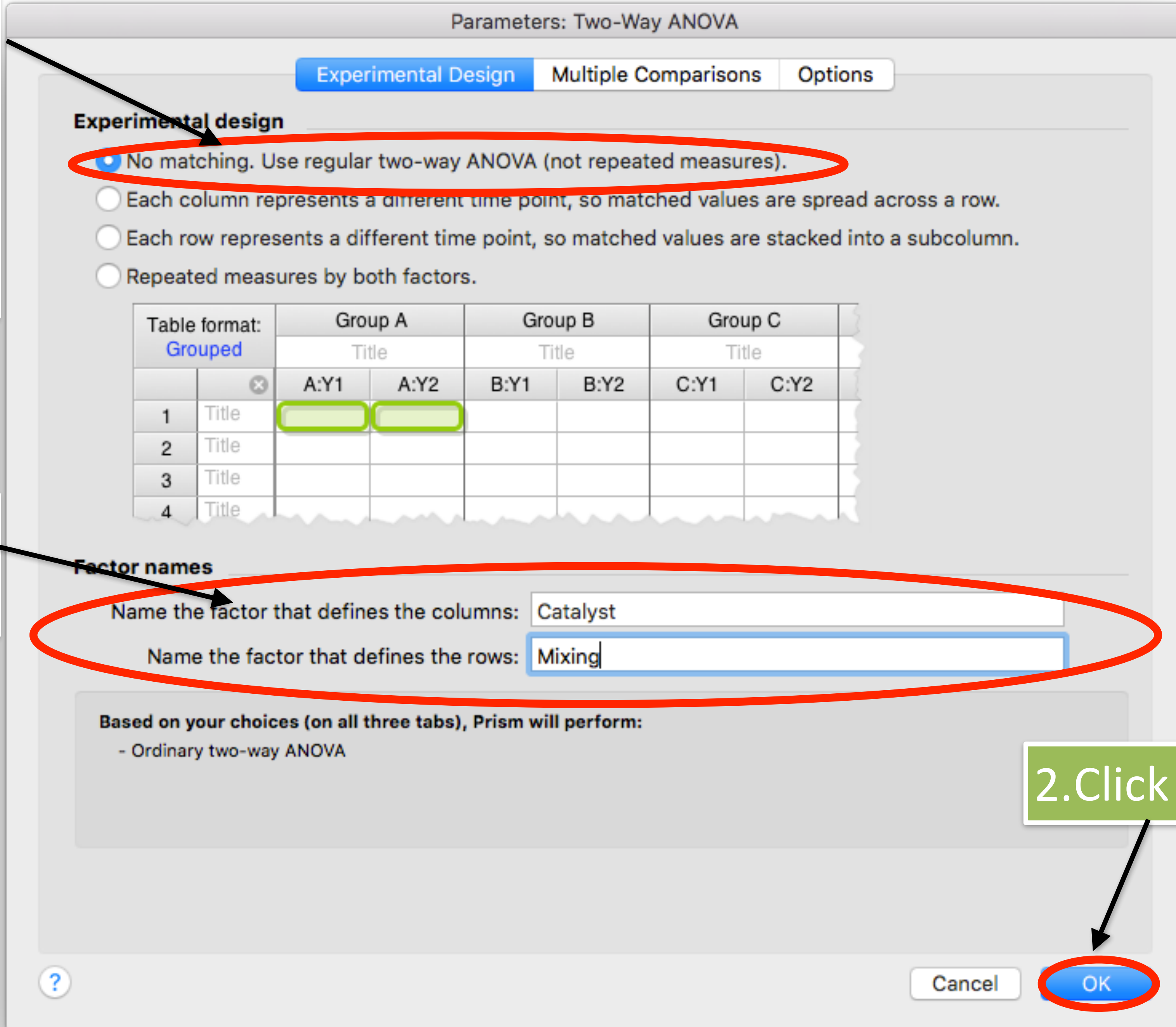
2. Population means among the levels of the second factors are the same

3. There is not interaction between the two factors

These are independent hypotheses. We could end up accepting them all, rejecting them all, or accepting some and rejecting others.

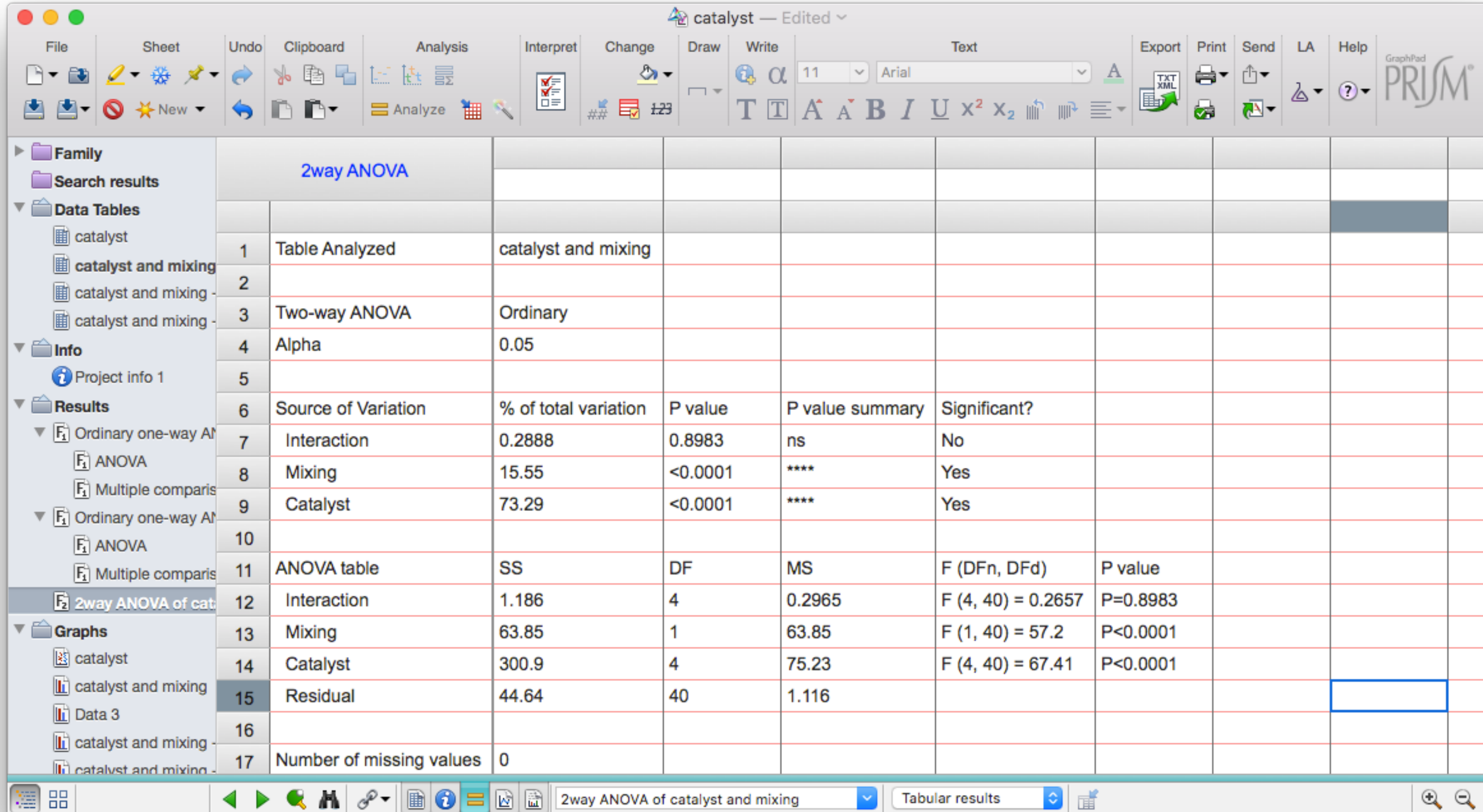

1.Select 'Two-way ANOVA' from Grouped analyses
Analyze Data
uilt-in analysis
ich analysis?
Transform, Normalize...
Transform
Transform Concentrations (X)
Normalize
Prune rows
Remove baseline and column math
Transpose X and Y
Fraction of Total
XY analyses
Column analyses
Grouped analyses
Two-way ANOVA
Three-way ANOVA
Row means with SD or SEM
Multiple t tests - one per row
Contingency table analyses
Survival analyses
Parts of whole analyses
Generate curve
Simulate data
Recently used
Analyze which data sets?
A:platinum
B:palladium
C:Iridium
D:Palladium/Iridium
E:Rhodium
Select All
Deselect All
Cancel
OK
2.Click OK

1.Select ‘No matching’ under Experimental design



2.Can change Factor names

2.Click OK



	2way ANOVA					
1	Table Analyzed	catalyst and mixing				
2						
3	Two-way ANOVA	Ordinary				
4	Alpha	0.05				
5						
6	Source of Variation	% of total variation	P value	P value summary	Significant?	
7	Interaction	0.2888	0.8983	ns	No	
8	Mixing	15.55	<0.0001	****	Yes	
9	Catalyst	73.29	<0.0001	****	Yes	
10						
11	ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
12	Interaction	1.186	4	0.2965	F (4, 40) = 0.2657	P=0.8983
13	Mixing	63.85	1	63.85	F (1, 40) = 57.2	P<0.0001
14	Catalyst	300.9	4	75.23	F (4, 40) = 67.41	P<0.0001
15	Residual	44.64	40	1.116		
16						
17	Number of missing values	0				

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Interaction	1.186	4	0.2965	F (4, 40) = 0.2657	P=0.8983
Mixing	63.85	1	63.85	F (1, 40) = 57.2	P<0.0001
Catalyst	300.9	4	75.23	F (4, 40) = 67.41	P<0.0001
Residual	44.64	40	1.116		

Interpretation of two-way ANOVA in the absence of interaction

- P-value for interaction effect = 0.898, non-significant interaction effect
- Since catalyst and mixing are significant but the interaction is not, there are ’additive’ effects of catalyst and mixing

Estimating the effect of type of mixing

Metal	Increase in yield (%)
Pt	1.86
Pd	2.36
Ir	2.72
Pd/Ir	2.38
Rh	1.98
MEAN	2.26%

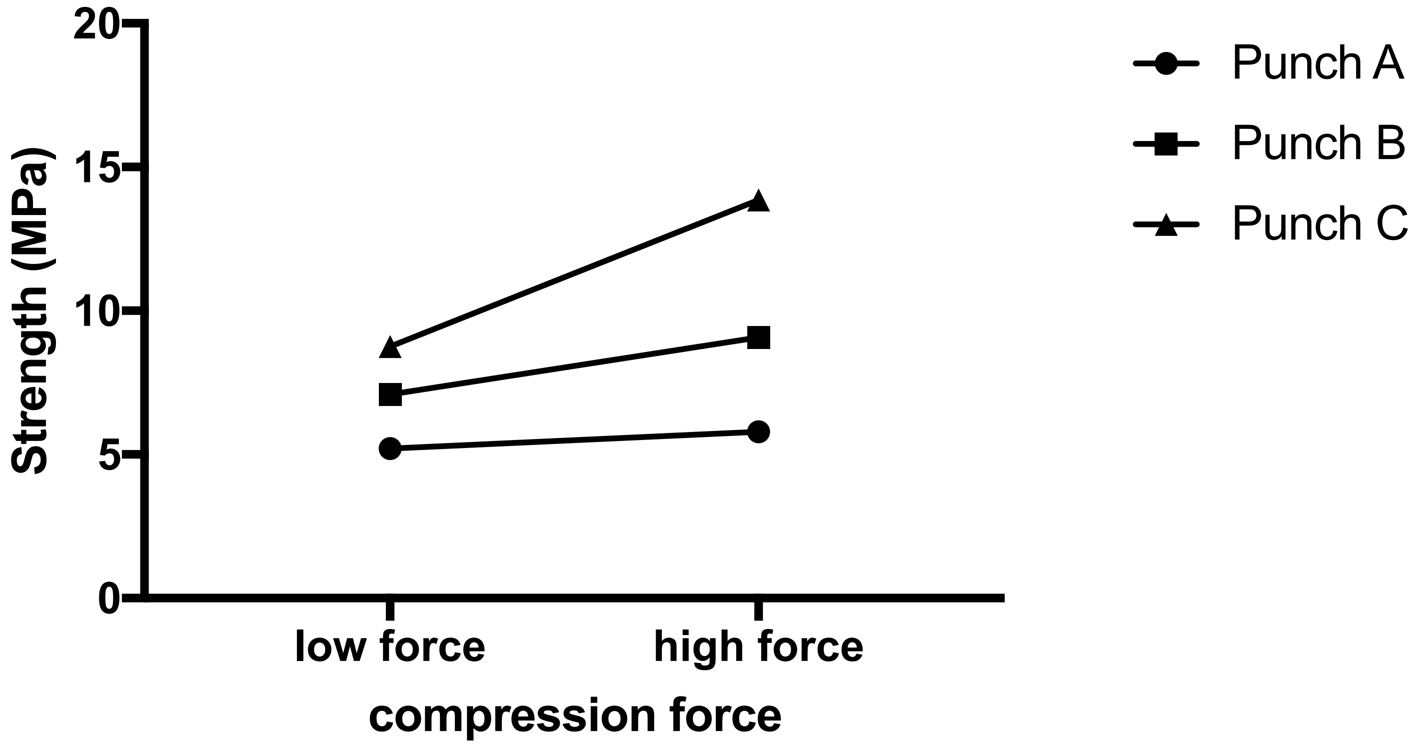
Summary of example two-way ANOVA

- There is no significant evidence of interaction between the two factors ($p=0.898$).

- There is significant evidence that yield varies according to the mixing method used ($p<0.001$). Whatever metal is used, ultrasonic mixing produces a yield around 2.3% greater than that achieved by stirring.

- There is significant evidence confirming that yield varies according to which catalyst is used ($p<0.001$).

Two-way ANOVA with Interaction



Summary of example two-way ANOVA with interaction

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Interaction	32.29	2	16.14	F (2, 30) = 20.35	P<0.0001
Force	58.85	1	58.85	F (1, 30) = 74.21	P<0.0001
Punch	203.4	2	101.7	F (2, 30) = 128.2	P<0.0001
Residual	23.79	30	0.7931		

- There is significant evidence of interaction between the two factors ($p<0.0001$).
- This is an example of quantitative interaction and should be followed up by specific comparisons.

What did we learn?

- One-way ANOVA is used when there are more than two groups to compare based on an interval outcome.

- When looking at pairwise comparisons of levels in a significant ANOVA model, you must 'adjust' your p-values for multiple testing, e.g., Tukey, Dunnett.

- Two-way ANOVA is used when there are two nominal/ ordinal factors that contribute to an interval outcome and all possible combinations of the two factors have been studied

- A significant interaction in a two-way ANOVA need to be interpreted with caution (graphics are extremely helpful).