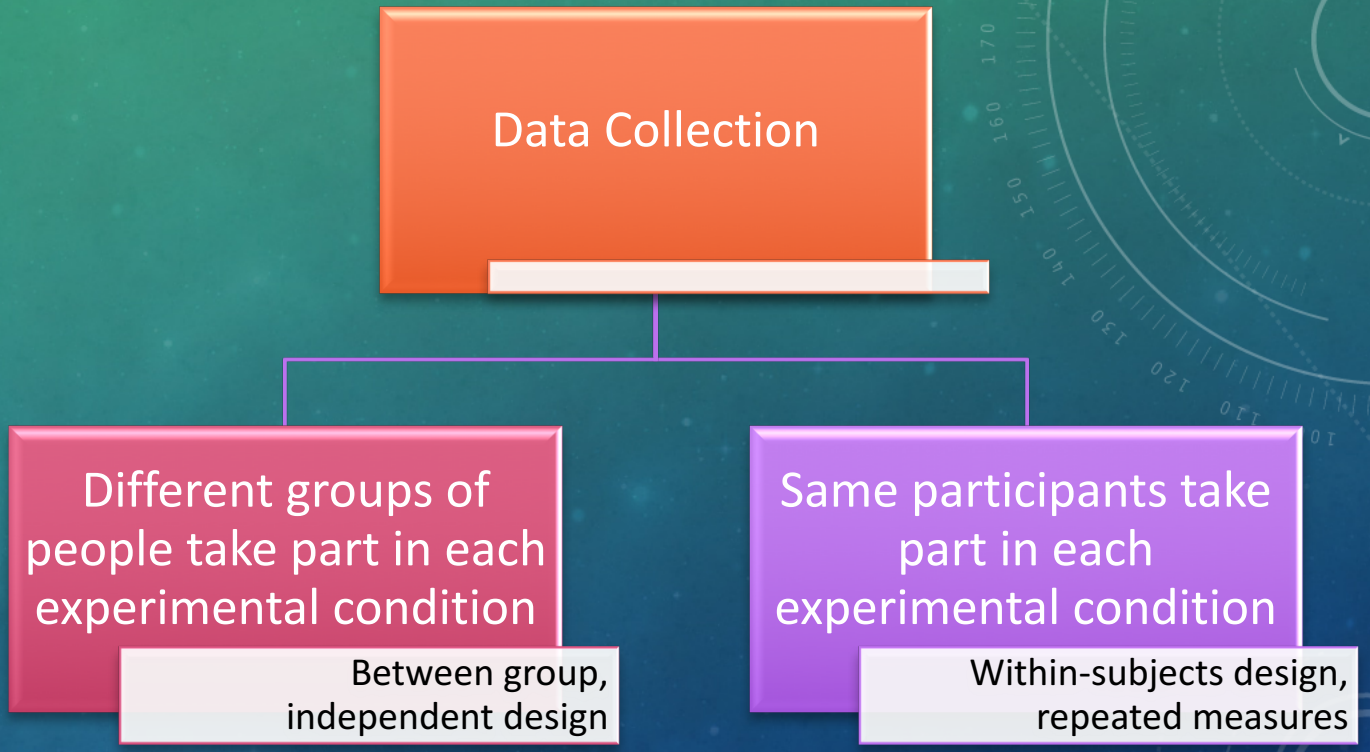
WEEK 9

COMPARING TWO MEANS: THE STUDENT T TEST

- Simplest scenario: Comparing two means..



Parametric test assumptions met: → Student t test (Two-sample t test) (Independent sample t test)

Paired sample t test

Parametric test assumptions violated: → Mann Whitney U test

Wilcoxon testi

ASSUMPTIONS OF TWO SAMPLE T TEST

- The two samples must be *independent*
- *(Ideally)* Observations should be chosen by random selection
- The variable of interest should be approximately *Normally* distributed in each population from which the samples are taken.
- The variability of the observations in each group, as measured by the two *variances*, should be approximately *equal*

Parametric test assumptions

REMEMBER THE STEPS FOR HYPOTHESIS TESTS

Step 1

- Specify the null and alternative hypothesis

Step 2

- Calculate the test statistics

Step 3

- Obtain P- value by referring the calculated value of test statistic (ignoring its sign) to the table of related distribution (which will be provided later)

Step 4

- Make a decision whether or not to reject the null hypothesis

Step 5

- Interpret the result

Calculating test statistic for two sample t test:

Equal variances

Unequal variances

Test statistic formula =>

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_o^2}{n_1} + \frac{S_o^1}{n_2}}}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^1}{n}}}$$

$$S_O^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

EXAMPLE:

- Suppose that you want to investigate the effect of a new prebiotic supplement on body weight gain of broilers in your farm. You divided a total of 200 broilers into two groups: Control and Treatment. You followed your standard feeding procedure for the control group and added the prebiotic supplement for the treatment group. After 42 days, you slaughtered and weigh the chickens. Assuming that your data follows normal distribution and equal variances, does the new prebiotic supplement significantly increase growth performance of your broilers?

	Control	Treatment
n	100	100
Arithmetic Mean	2790	2840
Standard Deviation	116	130

SOLUTION

- Step 1: Establish your hypothesis
  - *H0= Body weight gains for two groups are equal*
  - *H1= Body weight gain for two groups are unequal*
- Step 2= Calculate the test statistics

$n_1 = 100$, $n_2 = 100$
$S_1 = 116$, $S_2 = 130$
$X_1 = 2790$, $X_2 = 2840$

$$S_O^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_o^2}{n_1} + \frac{S_o^1}{n_2}}}$$

t= 2,869

- Step 3: Obtain the P value referring the calculated test statistic
  - *Refer to 2,869 with 198 degrees of freedom; P value is between= 0.01- 0.001*
- Step 4: Make a decision whether or not to reject the null hypothesis
  - *Considering the P value is <0.05; the null hypothesis, that there is no difference in the mean body weight gains in the two groups, is unlikely to be true. We therefore reject the null hypothesis in favour of the alternative hypothesis that there is a difference in body weight gains.*
- Step 5: Interpret the results
  - *The prebiotic supplement significantly improves body weight gain in broiler chickens (p<0.01).*

EXAMPLE

	Typeofbirth	Weight
1	Twin Birth	30
2	Twin Birth	35
3	Twin Birth	45
4	Twin Birth	40
5	Twin Birth	50
6	Twin Birth	35
7	Twin Birth	55
8	Twin Birth	25
9	Twin Birth	30
10	Twin Birth	45
11	Twin Birth	40
12	Twin Birth	50
13	Single Birth	40
14	Single Birth	35
15	Single Birth	50
16	Single Birth	55
17	Single Birth	65

- **Suppose that a researcher wants to evaluate the effect of type of birth on weight gain of sheeps at the end of shearing season?**

Hypothesis?

H_0: No difference in weight of sheeps according to type of birth

H_A: There is a difference in weight of sheeps according to type of birth

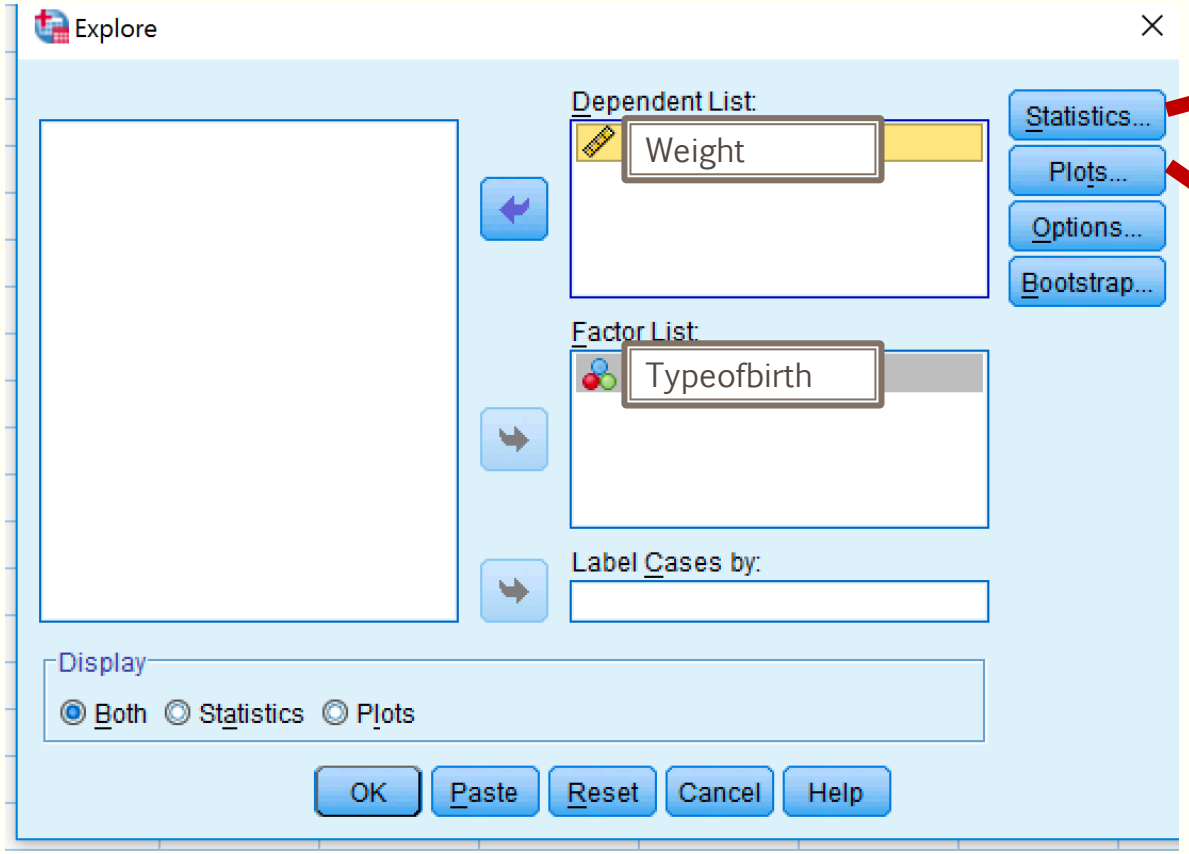
Dependent variable? → Weight

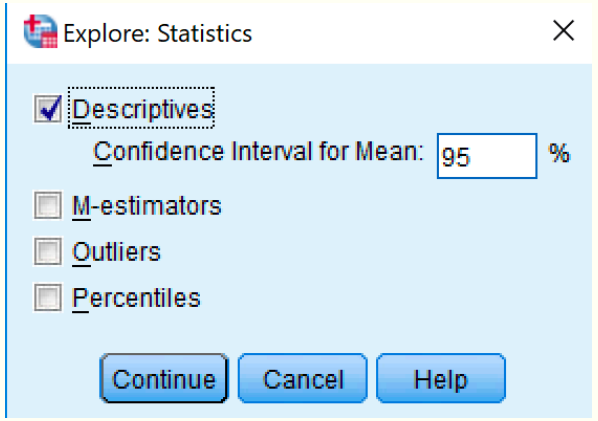
Independent variable? → Type of birth

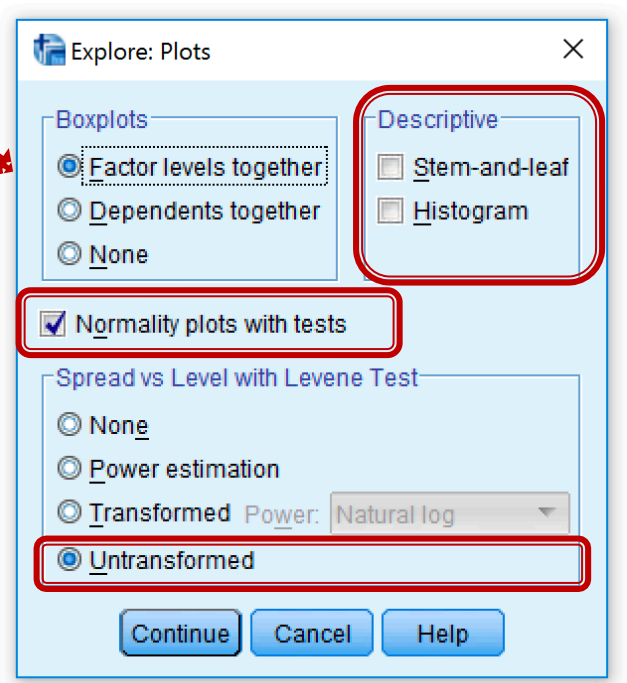
Dataset› Sheepweight.sav

Step 1: Testing the assumptions…

- **Analyze > Descriptive Statistics > Explore**







Testing the assumptions: a)Normality & b) Homogeneity of variances

a)Normality assumption:

H_0= The data follow a normal distribution

H_1= The data do not follow a normal distribution

Tests of Normality							
		Kolmogorov-Smirnov			Shapiro-Wilk		
	Group	Statistic	df	Sig.	Statistic	df	Sig.
Weight	single	0,121	12	,200*	0,965	12	0,852
	twin	0,191	12	,200*	0,949	12	0,621

P>0.05

H_0 is accepted

b) Homogeneity of variances assumption:

H_0= The population variances are equal

H_1= The population variances are not equal

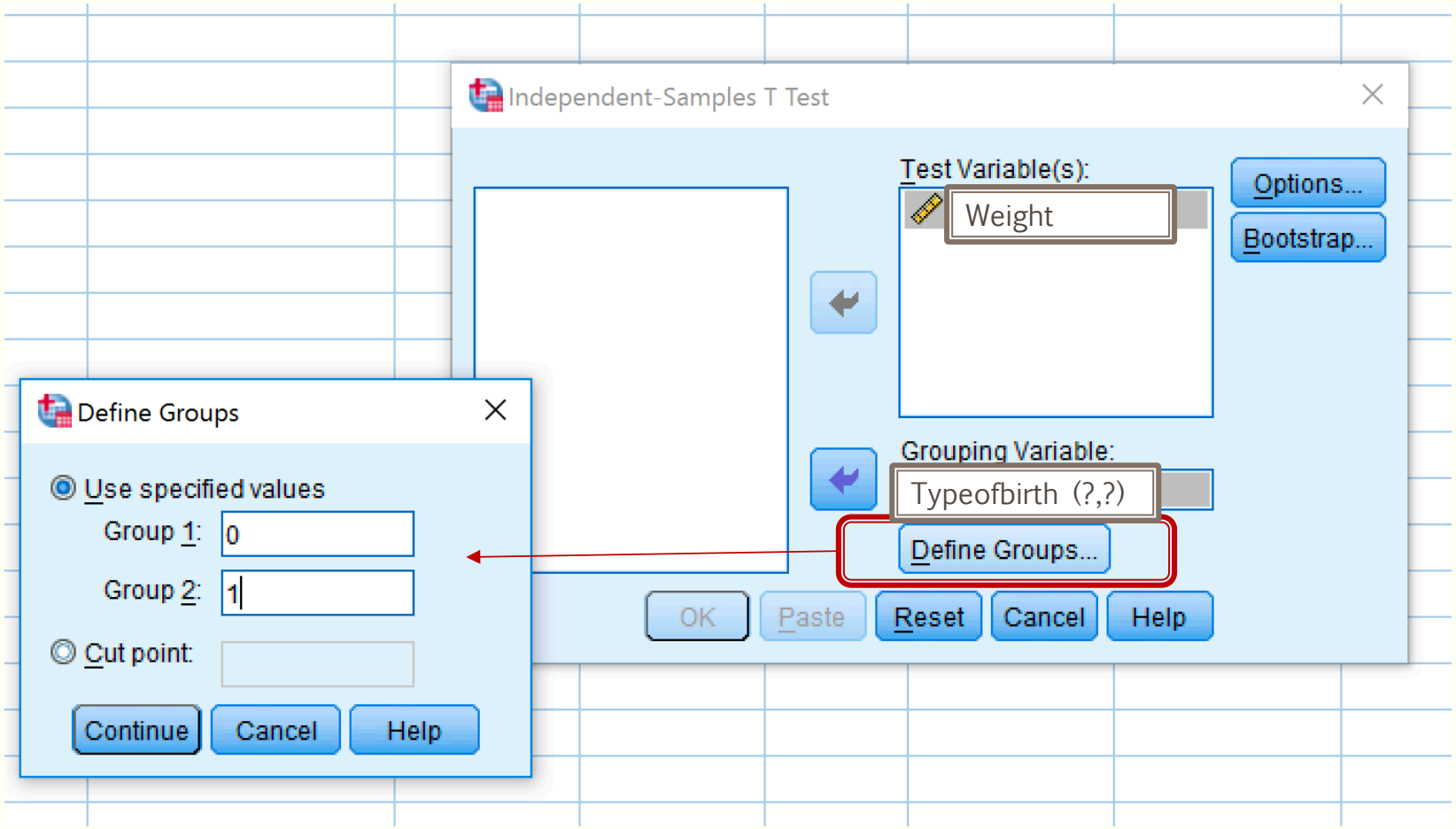
Test of Homogeneity of Variance					
		Levene Statistic	df1	df2	Sig.
Weight	Based on Mean	0,782	1	22	0,386
	Based on Median	0,299	1	22	0,59
	Based on Median and with adjusted df	0,299	1	20,194	0,59
	Based on trimmed mean	0,792	1	22	0,383

P>0.05

H_0 is accepted

Data analysis: Student t test

Analyze > Compare Means > Independent Samples t test



Results

	Group	n	Mean	Std. Deviation	Std. Error Mean
Weight	Twin birth	12	40	9,293	2,683
	Single birth	12	47	11,029	3,184

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Weight	Equal variances assumed	0,782	0,386	-1,681	22	0,107	-7	4,163	-15,634	1,634
	Equal variances not assumed			-1,681	21,385	0,107	-7	4,163	-15,649	1,649

Reporting the results of the analysis

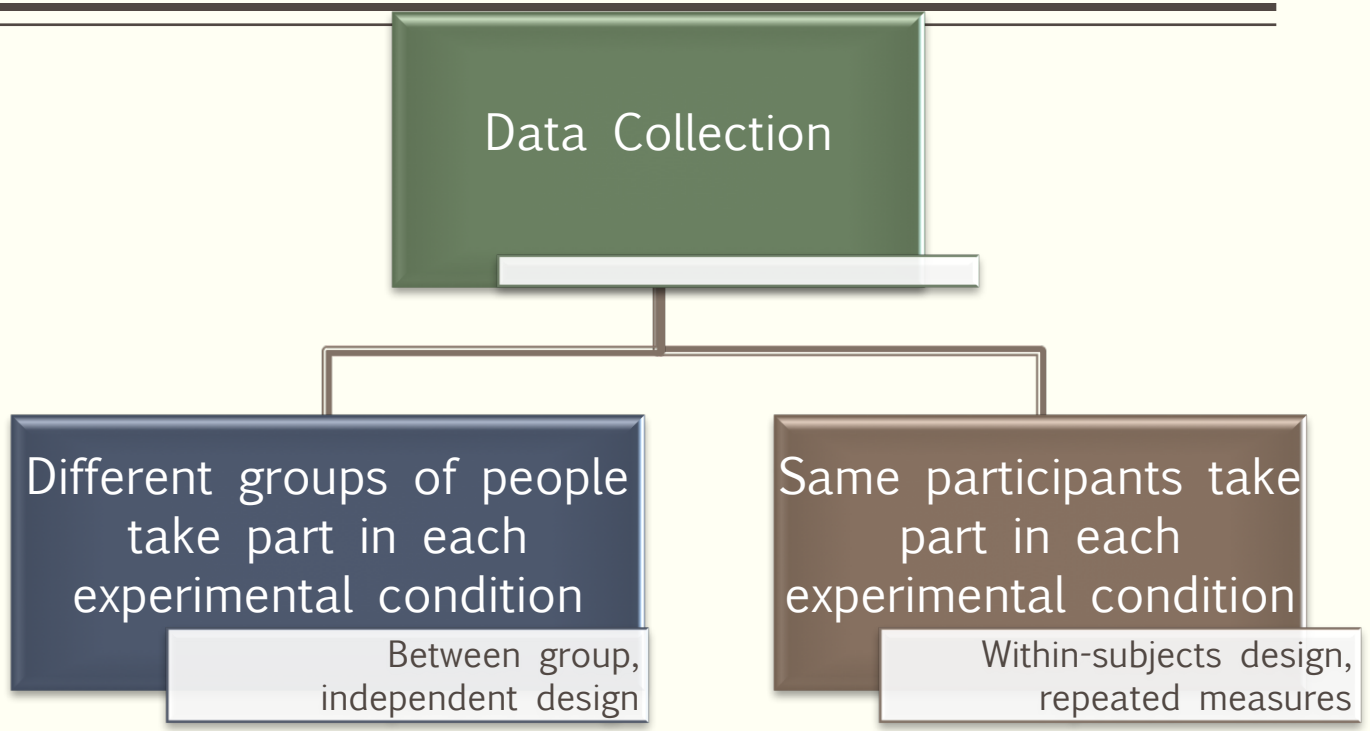
HEADLINE

Variable	Group	n	Mean ± SEM	P
Weight	Twin	12	40 ± 2,683	0,107
	Single	12	47 ± 3,184	

H_0: No difference in weight of sheeps according to type of birth
H_A: There is a difference in weight of sheeps according to type of birth

RESULTS
Although mean weight was higher in single births, this difference was not statistically significant (P>0.05). ➔ H_0 is accepted.

What if the parametric test assumptions are violated?



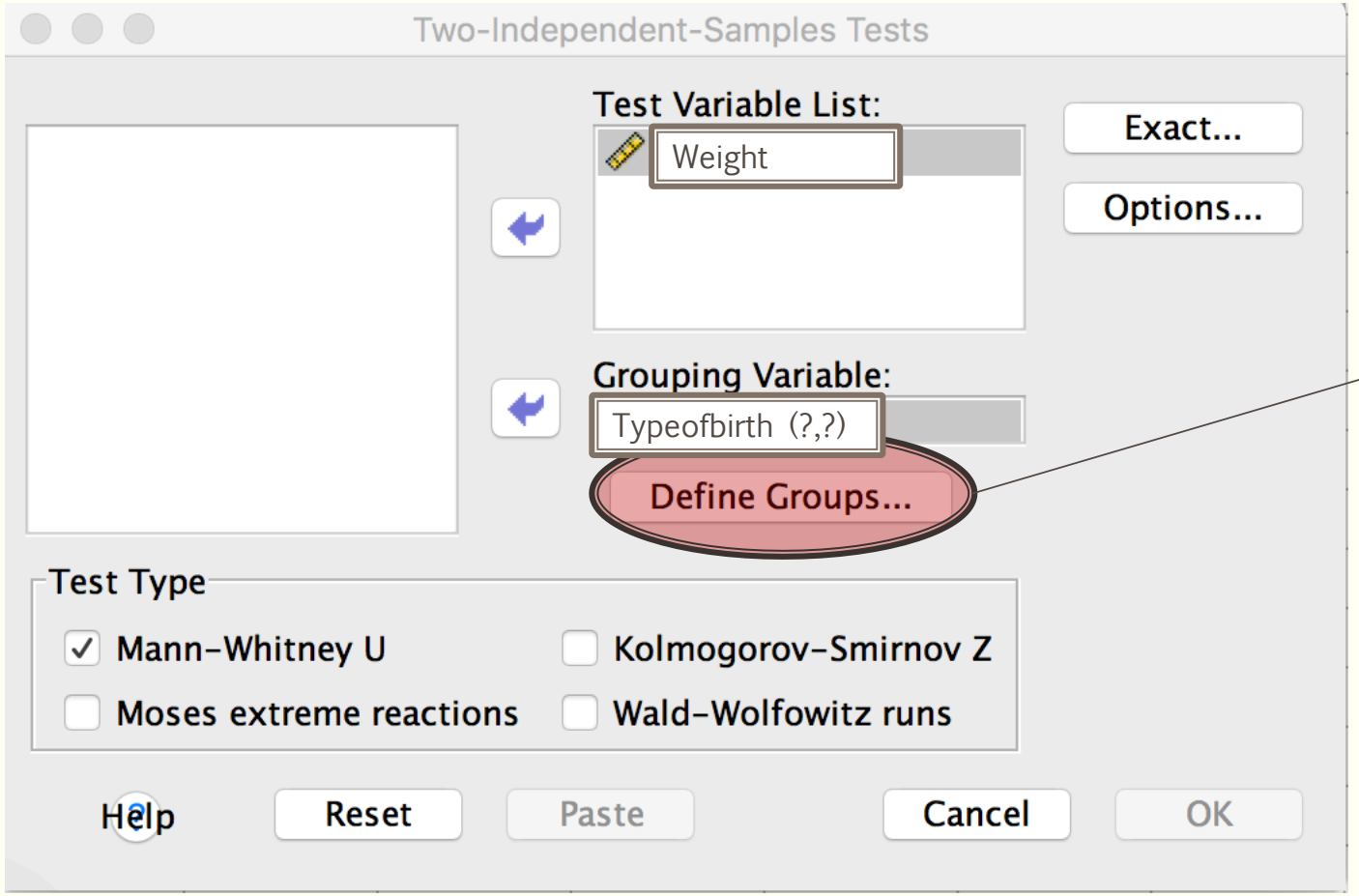
Parametric test assumptions met: → Student t test (Two-sample t test) (Independent sample t test) — Paired sample t test

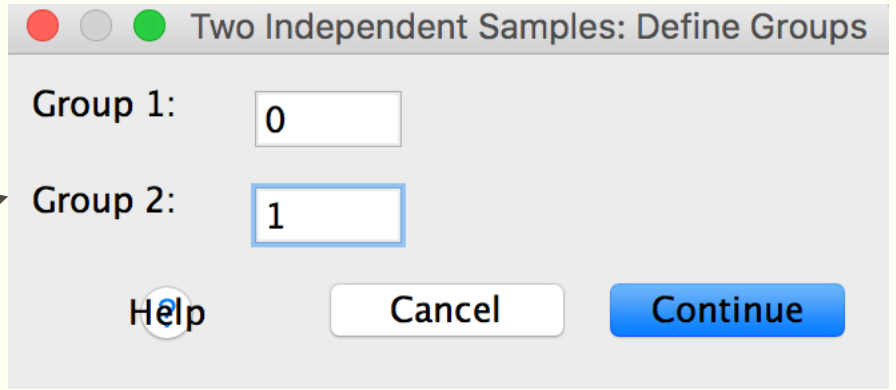
Parametric test assumptions violated: → Mann Whitney U test — Wilcoxon testi

Let's use the same dataset and assume that the assumptions are violated

Mann Whitney U testi

- Analyze > Non-Parametric Tests > Legacy Dialogs > 2 Independent Samples





RESULT

Test Statistics[a]

	Kaygı_Düzeyi
Mann-Whitney U	46,000
Wilcoxon W	124,000
Z	-1,515
Asymp. Sig. (2-tailed)	,130
Exact Sig. [2*(1-tailed Sig.)]	,143[b]

a. Grouping Variable: Grup

b. Not corrected for ties.

INTERPRETATION?

Why do you think the p value obtained from student t test is different from the one we obtain from Mann Whitney U test?

How should we report the results of a non parametric test?

Type of Birth	Weight		P
	n	Median (Min – Max)	
Twin Birth	12	40 (25 - 55)	0,13
Single birth	12	50 (30 - 65)	