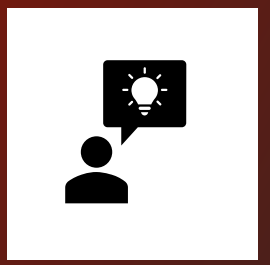
STATISTICAL TEST

Assoc. Prof. Dr. Nurulhuda Firdaus Bt Mohd Azmi
huda@utm.my

RECAP

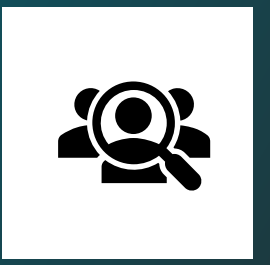
Feature	Parametric Tests	Non-Parametric Tests
Assumption	Assumes data follows a specific distribution (usually normal)	No strict distribution assumptions
Data Type	Works best with **continuous** and **normally distributed** data	Works with **ordinal, ranked, or non-normally distributed** data
Sample Size	Requires **larger** sample sizes (n > 30 for reliability)	Works well with **small** sample sizes
Test Type	Tests **means** and variance	Tests **medians** and ranks
Statistical Power	More **powerful** when assumptions hold	Less powerful but **robust** for non-normal data
Examples	**Paired t-test, ANOVA, Pearson correlation**	**Wilcoxon signed-rank test, Kruskal-Wallis, Spearman correlation**

SPSS for Statistical Test



One Sample Statistical Test: Parametric and Non-Parametric

Independent Group Statistical Test: Parametric and Non-Parametric



Dependent (Paired) Group Statistical Test: Parametric and Non-Parametric

First Skill | Second Skill | Third Skill

STATISTICAL HYPOTHESIS TEST

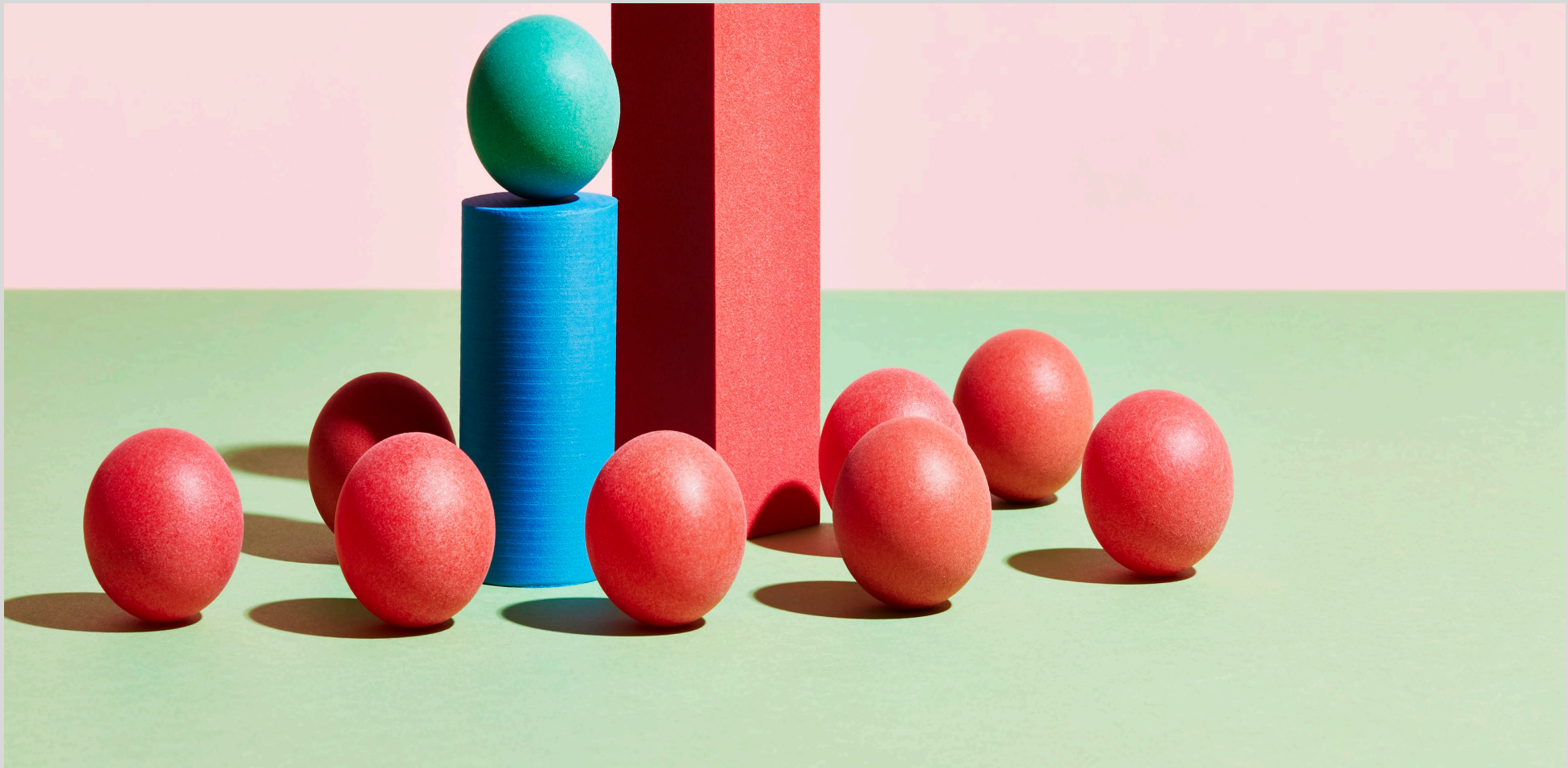
Single Population (one sample test):

- Uses the mean of a single sample to determine if the population mean is statistically different from a specified value.

Two Population (two sample test):

- Independent Test: Uses the mean of two samples to determine if the two populations are statistically different from each other. The group of these two samples is independent.
- Paired Test: Uses the difference of two paired samples to determined if the paired populations are statistically different. The group of these two samples is dependent.

PARAMETRIC STATISTICAL TEST

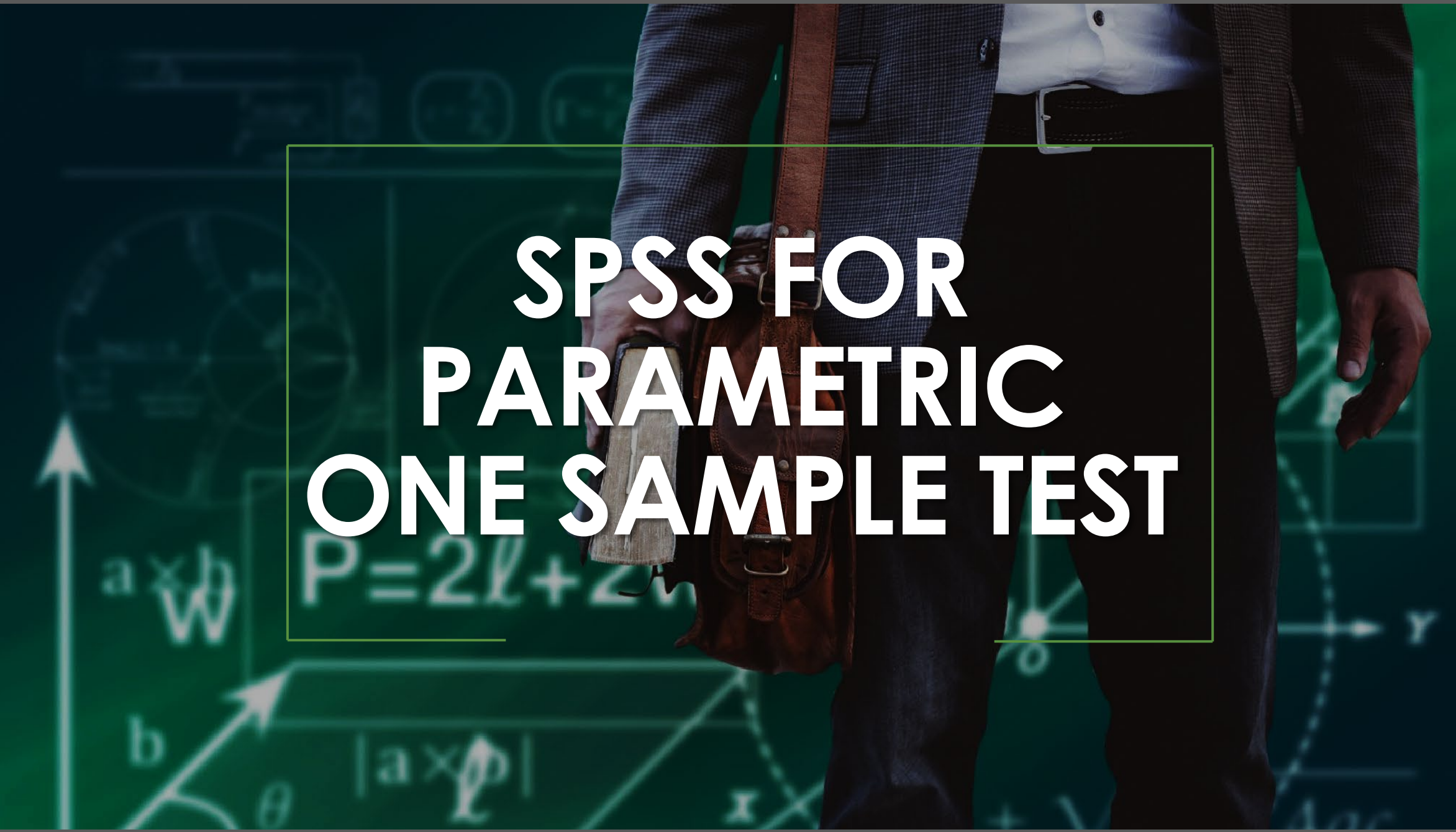


PARAMETRIC TEST

(SAMPLE DATA WHICH FOLLOW NORMAL DISTRIBUTION)

Variables which are Numerical (quantitative)

- ONE SAMPLE
  - T-TEST
- 2 INDEPENDENT SAMPLE
  - INDEPENDENT T-TEST
- 2 DEPENDENT (PAIRED) SAMPLE
  - PAIRED T-TEST


SPSS FOR
PARAMETRIC
ONE SAMPLE TEST

ONE GROUP SAMPLE TEST

Research Question:

Is the mean age population is the same with reported by the local authorities which is 45 years old?

Data Set: health survey.sav

Hypothesis Statement:

$H_0 : \mu = 45$ years (claim)

$H_A : \mu \neq 45$ years

This is two tailed hypothesis test

Checking Assumption:

1. Sample size?
2. Normal distribution or not?

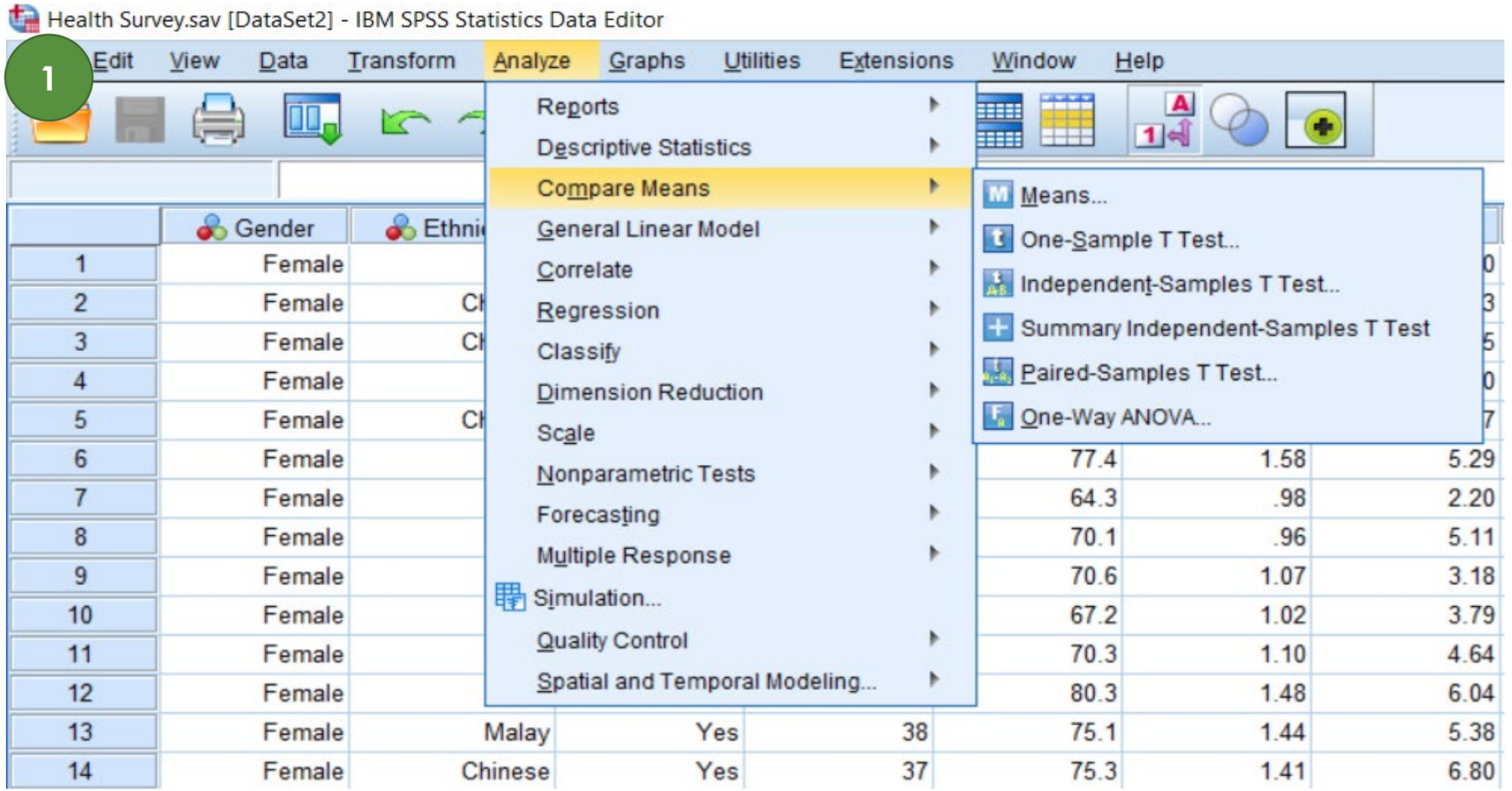
*** Notes: For large sample size (say, N > 30), the normality assumption not holding doesn't really affect the results.

Parametric One-Sample statistical Test: t-test

Data Set: Health Survey.sav

Notes: We have check normality (refer to previous notes. It shows that variable 'age' follows normal distribution. Therefore, for one sample test, we will use t-test.

1. Analyze > Compare Means > One-Sample T test

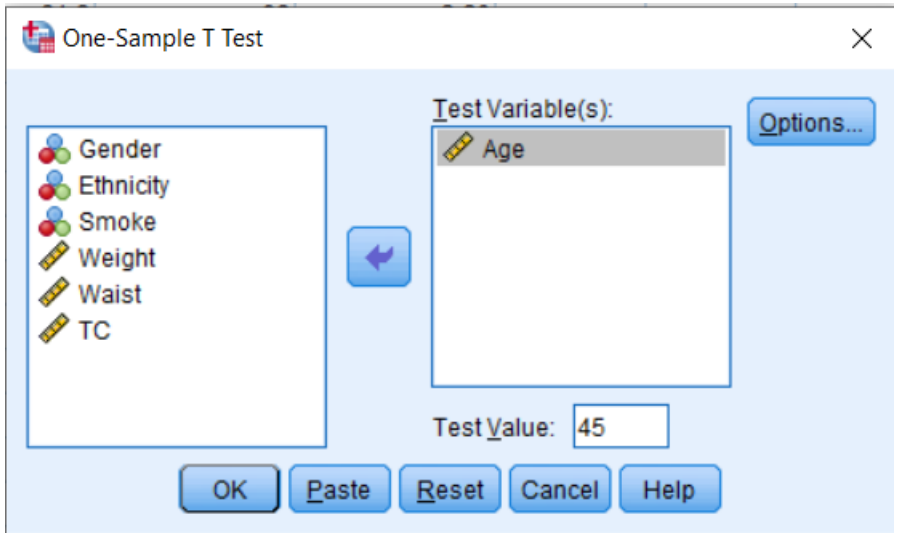


Parametric One-Sample statistical Test: t-test

Data Set: Health Survey.sav

2. Select variable Age as Test Variable(s).
   Put Test Values = 45. This is the hypothesized value.
   Click Option if you want to change the confidence interval (CI) percentage. By default the CI percentage is 95% (or $\alpha = 0.05$)

2



OUTPUT➔ Parametric One-Sample statistical Test: t-test

Data Set: Health Survey.sav

H_0 : μ = 45 years (claim)
H_A : μ ≠ 45 years
This is two tailed hypothesis test

1

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Age	350	40.00	6.708	.359

One-Sample Test

Test Value = 45

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Age	-13.944	349	.000	-5.000	-5.71	-4.29

To summarize:

- We reject the null hypothesis that the mean age in the sample is equal to the reported age.
- There is a significant difference in age mean between the sample and the overall reported.
- Thus, we conclude the average age of the sample is about 5 years younger than the overall reported.


LET US PRACTICE

A government health department is analyzing healthcare expenditure to determine whether patients in a specific region spend **less** on healthcare compared to the national average. According to recent reports, the **national average annual healthcare expenditure per patient is $5,000**.

A study was conducted using **a random sample of 137 patients** from the region to analyze their **annual healthcare expenditure**. The objective is to determine whether the **average healthcare expenditure in this region is significantly lower than the national average**. Data set is **Healthcare_Expenditure.xlsx**

Using the provided dataset, answer the following:

1.Perform a normality test to check if the data follows a normal distribution.

1. Is the assumption of normality met for conducting a **one-sample parametric test (t-test)?**

2.Conduct a one-sample t-test to test the following hypothesis:

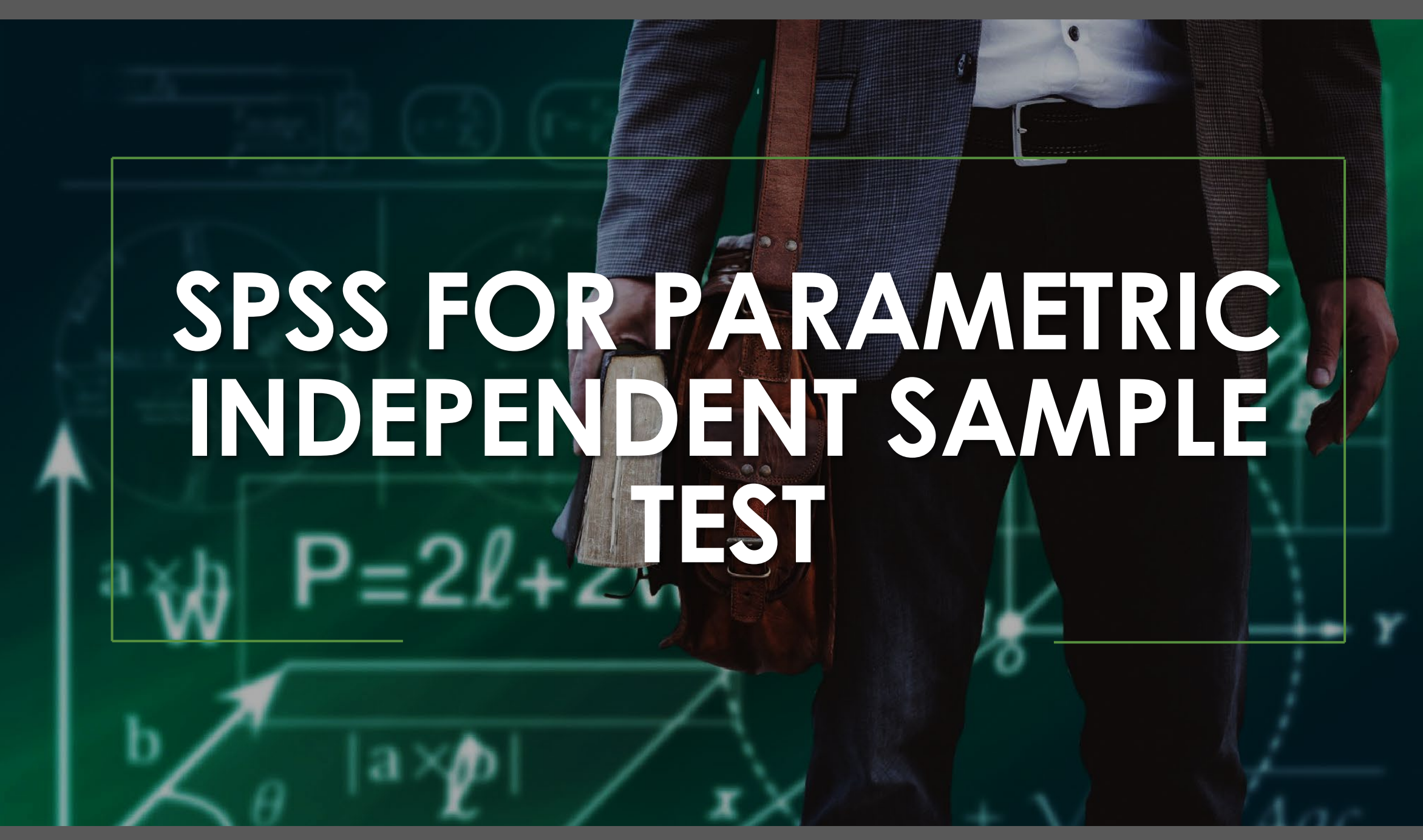
Null Hypothesis (H_0): The average annual healthcare expenditure per patient is **greater than or equal to $5,000**.

Alternative Hypothesis (H_1): The average annual healthcare expenditure per patient is **less than $5,000**.

3.Interpret the results:

1. What is the **p-value** and **t-statistic**?
2. Based on a **significance level of 0.05**, should the null hypothesis be rejected?
3. Can we conclude that the average healthcare expenditure in this region is **significantly lower than the national average**?

TWO GROUP SAMPLE TEST



SPSS FOR PARAMETRIC INDEPENDENT SAMPLE TEST

Research Question:

Is there any significant difference in the age among male and female in the study population?

Data Set: health survey.sav

Hypothesis Statement:

$$H_0 : \mu_{male} = \mu_{female}$$
$$H_A : \mu_{male} \neq \mu_{female} \text{ (claim)}$$

This is two tailed hypothesis test

Checking Assumption:

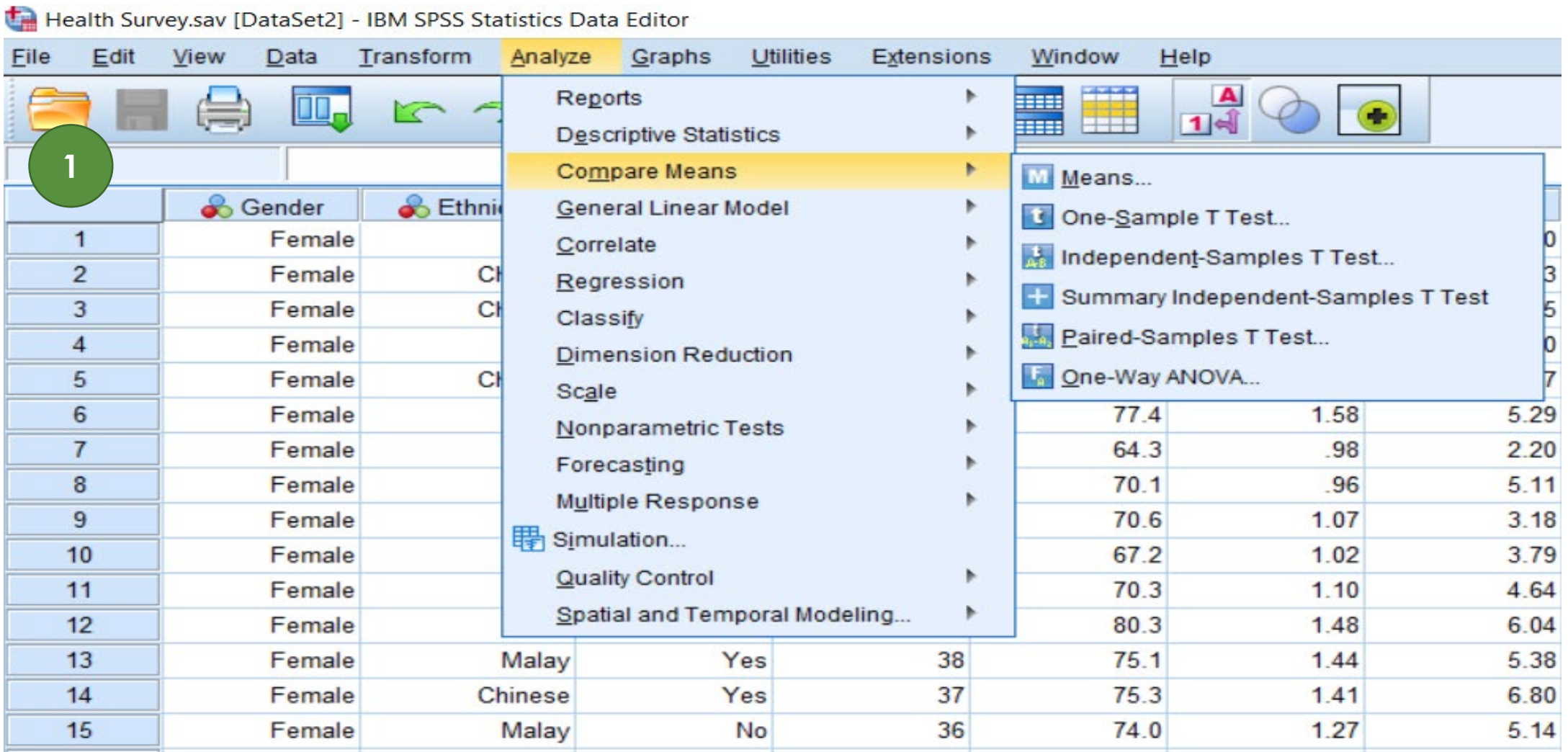
1. Dependent Variables should be in continuous scale.
2. Independent Variables should be in 2 group categorical only.
3. Observations are independent (not related or measured twice)
4. No significant outliers
5. Dependent should be normally distributed
6. Homogeneity of variances should be need. (Levene'sTest)

Parametric Independent Sample Test: Student's t test

Data Set: Health Survey.sav

Notes: We have check normality (refer to previous notes. It shows that variable 'age' follows normal distribution.

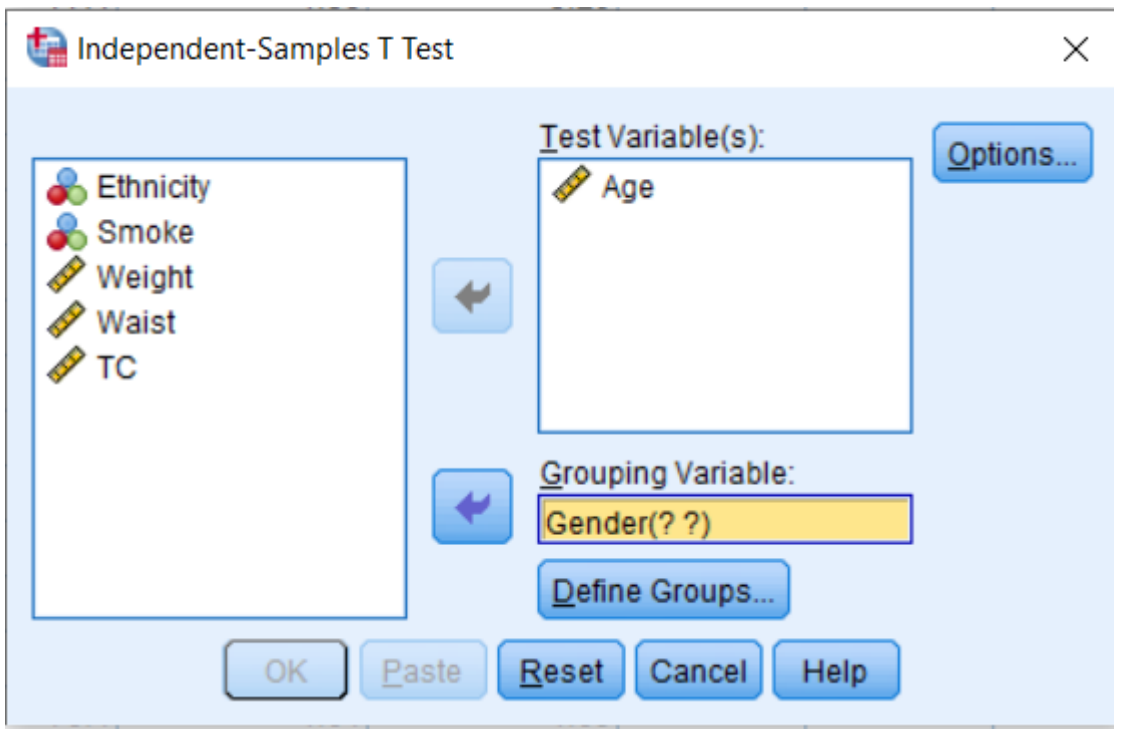
1. Analyze > Compare Means > Independent-Samples T test



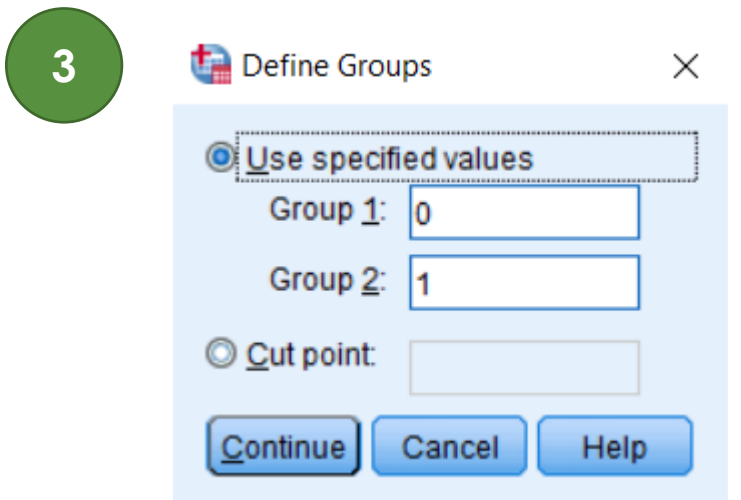
Parametric Independent Sample Test: Student's t test

Data Set: Health Survey.sav

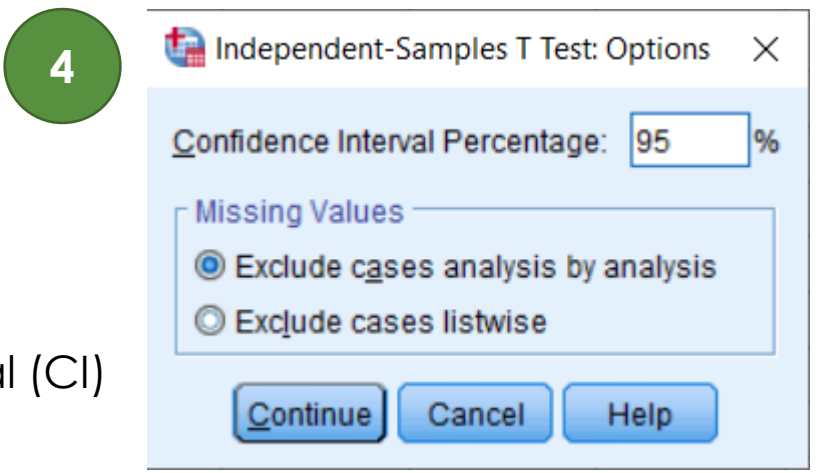
2. Select variable Age as Test Variable(s).
   Select variable Gender as grouping variable and click Define Groups..



3. Define the group variable. For Gender, 0 = female and 1 = male



4. Click Option if you want to change the confidence interval (CI) percentage. By default the CI percentage is 95% (or $\alpha = 0.05$)



OUTPUT➔ Parametric Independent Sample Test: Student's t test

Data Set: Health Survey.sav

Levine test for test variance homogeneity
H_0 : Variance is homogeneity
H_A : Variance is not homogeneity

Parametric Independent Sample Test
$H_0 : \mu_{male} = \mu_{female}$
$H_A : \mu_{male} \neq \mu_{female}$ (claim)

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Age	Female	234	38.21	6.112	.400
	Male	116	43.60	6.428	.597

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Age	Equal variances assumed	.855	.356	-7.630	348	.000	-5.387	.706	-6.776	-3.999
	Equal variances not assumed			-7.501	219.422	.000	-5.387	.718	-6.803	-3.972

To summarize:

- Check the homogeneity of variance through Levine's test: p-value = 0.356 thus, accept H_0 which indicate that the variance is homogeneity.
- P-value is small compare to α, and therefore we reject H_0 that there is a statistical significant differences in the mean age between male and female
- Thus, we conclude the average age of the sample is about 5.387 years younger (-6.776, -3.999) than the overall reported.


LET US PRACTICE

Research Question:

Is there any significant difference in weight among smokers and non smokers in the study population?

Data Set: health survey.sav

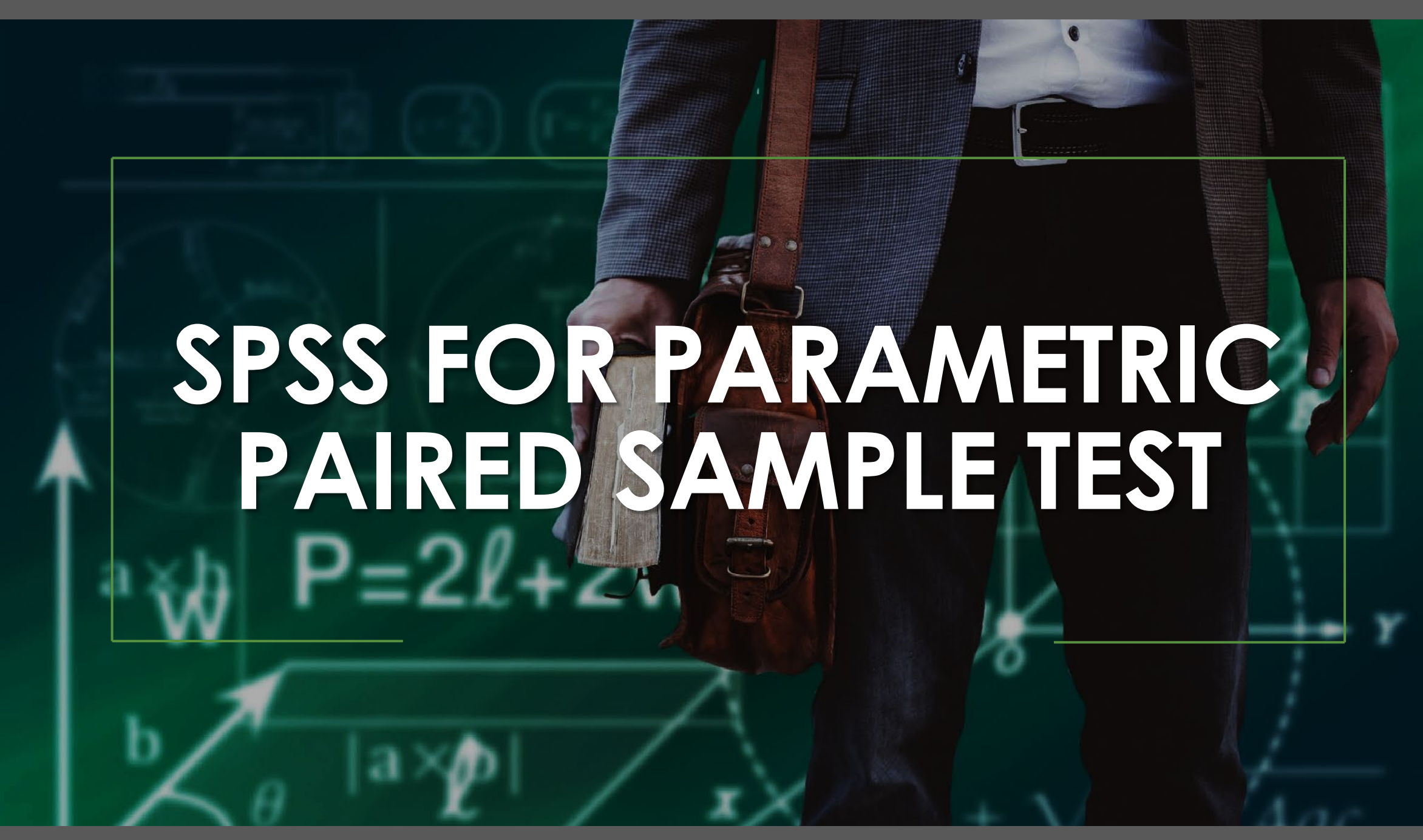
Hypothesis Statement:

$$H_0 : \mu_{smokers} = \mu_{nonsmokers}$$
$$H_A : \mu_{smokers} \neq \mu_{non\ smokers} \text{ (claim)}$$

This is two tailed hypothesis test

Checking Assumption:

1. Dependent Variables should be in continuous scale.
2. Independent Variables should be in 2 group categorical only.
3. Observations are independent (not related or measured twice)
4. No significant outliers
5. Dependent should be normally distributed
6. Homogeneity of variances should be need. (Levene'sTest)



SPSS FOR PARAMETRIC PAIRED SAMPLE TEST

Problem Statement

A hospital is conducting a clinical study to evaluate the effectiveness of a new antihypertensive drug. Researchers collected **systolic blood pressure readings from 40 patients** before and after taking the drug for 30 days. The goal is to determine whether the drug significantly reduces blood pressure using a **parametric paired sample t-test** at a **5% significance level**.

Research Question

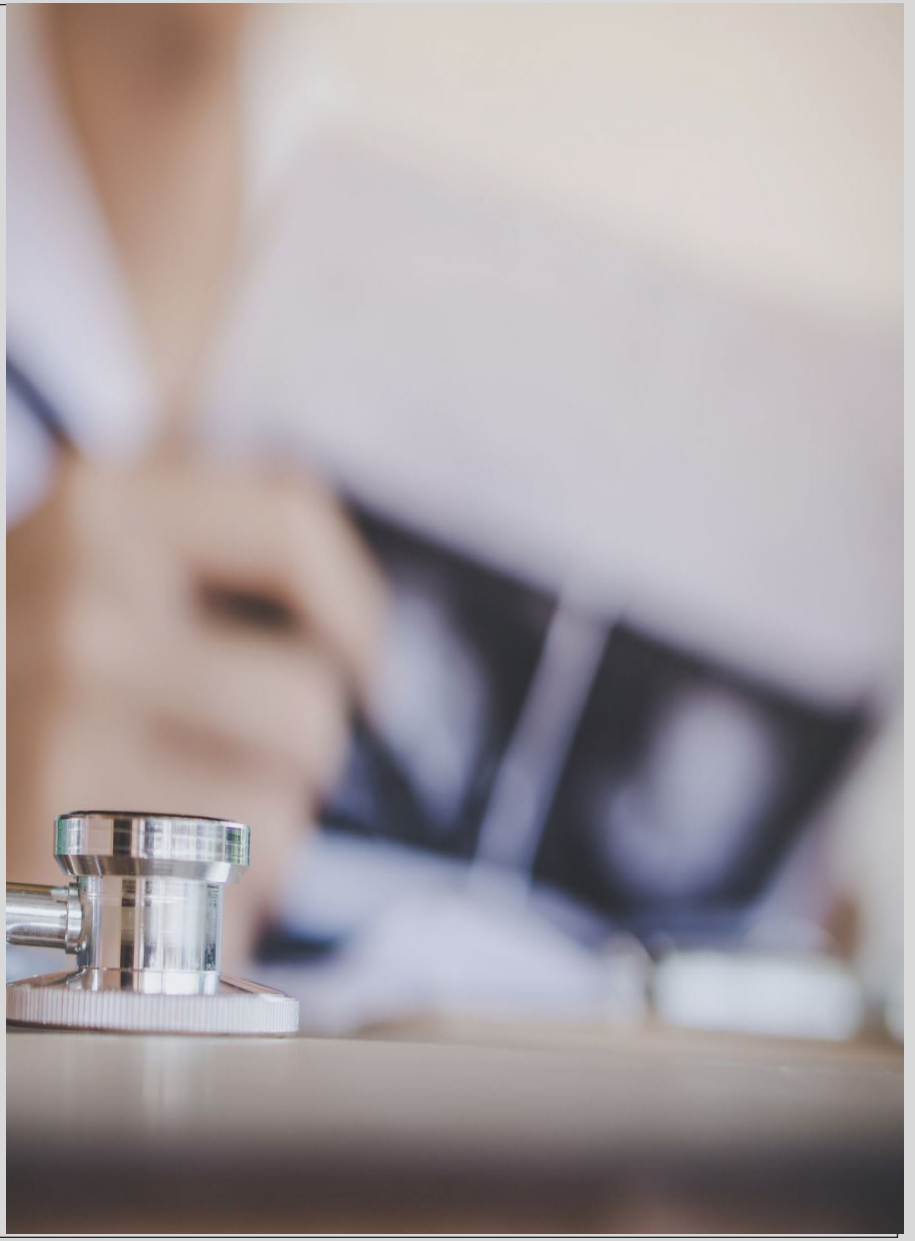
"Does the new antihypertensive drug significantly reduce systolic blood pressure after 30 days of use?"

Data Set: large_blood_pressure.xlsx

Hypothesis Statement:

H_0 : *There are **no differences** on systolic blood pressure reading before and after the treatment*

H_A : *There are **differences** on systolic blood pressure reading before and after the treatment*

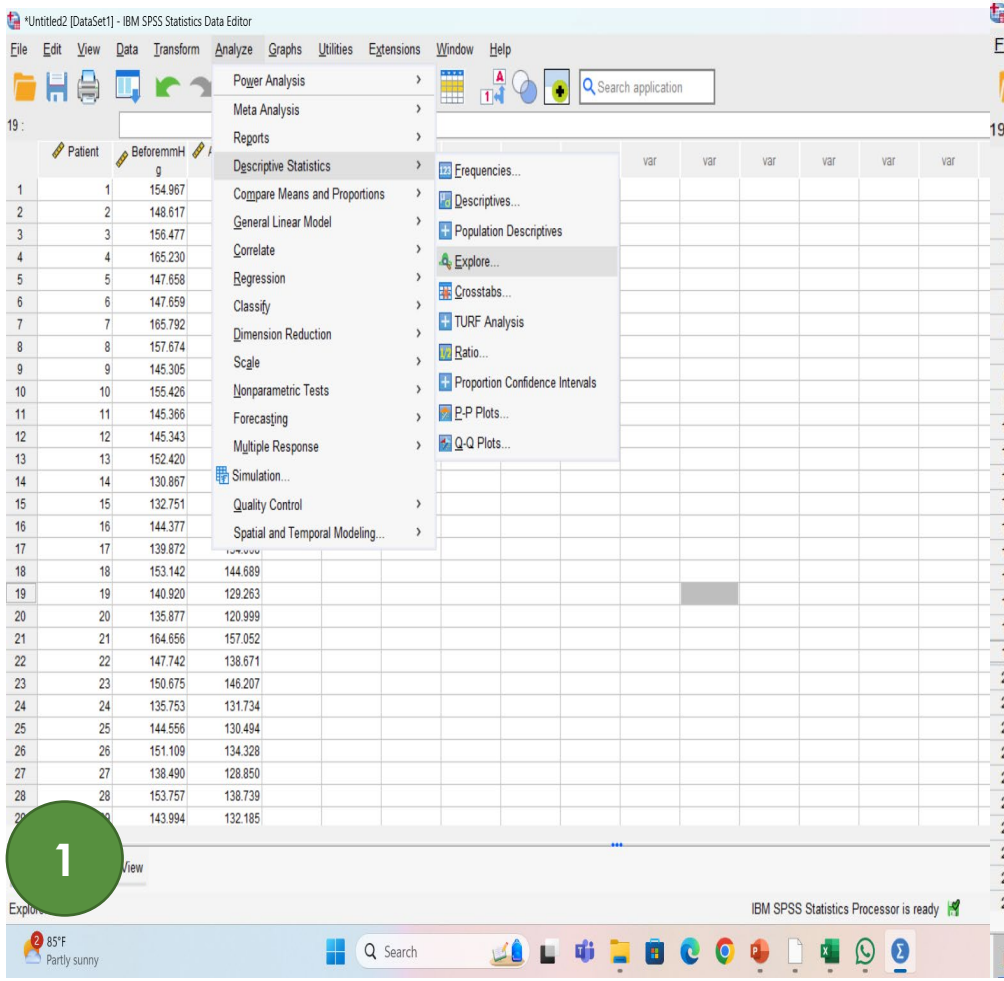


Parametric Paired Sample Test: Student's t test

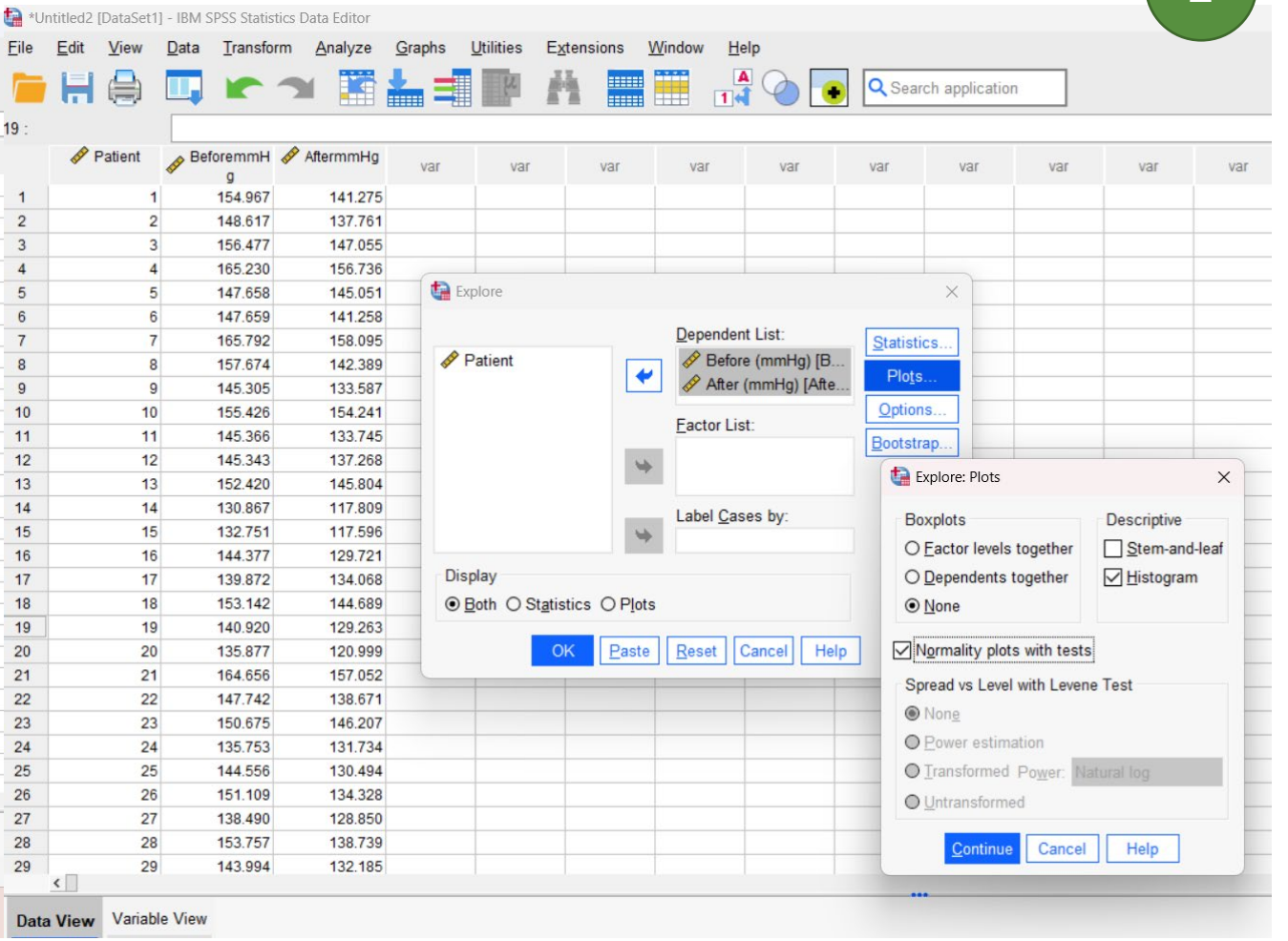
Data Set: large_blood_pressure.xlsx

Normality Test Assessment

1



2



Parametric Paired Sample Test: Student's t test

Data Set: large_blood_pressure.xlsx

Normality Test Assessment

Hypothesis Statement for normality test
H_0 : Variable is normally distributed
H_A : Variable is Not Normally distributed

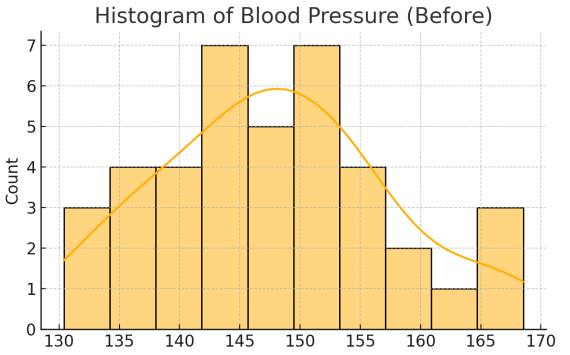
Tests of Normality

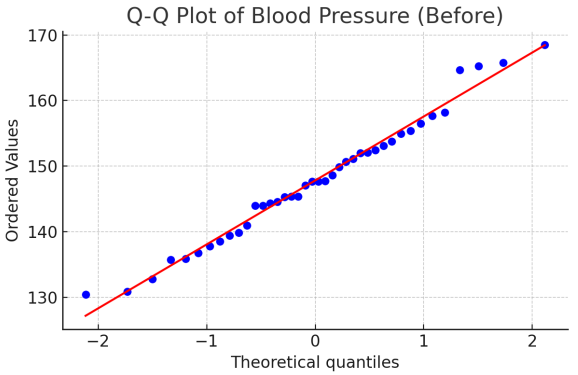
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Before (mmHg)	.069	40	.200*	.979	40	.660
After (mmHg)	.075	40	.200*	.978	40	.620

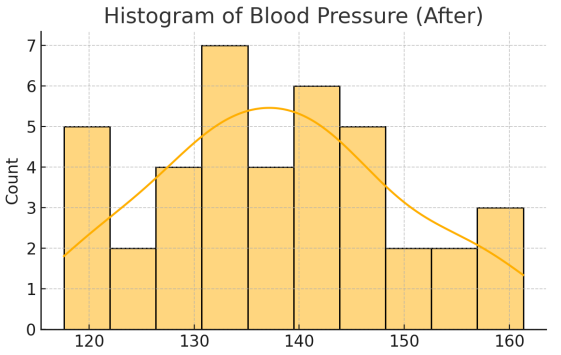
*. This is a lower bound of the true significance.

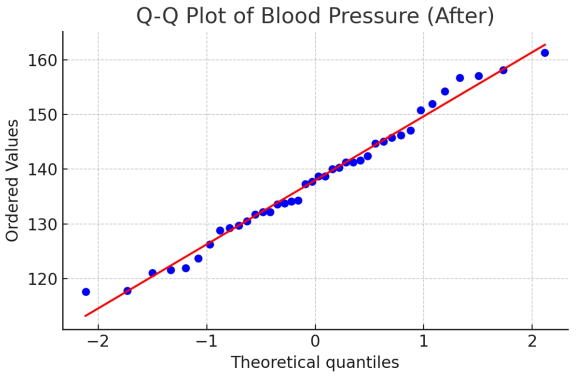
a. Lilliefors Significance Correction

From the result of Shapira Wilk (due to sample size < 50) since both **p-values > 0.05**, we **fail to reject the null hypothesis of normality**, indicating that both datasets **follow a normal distribution**.





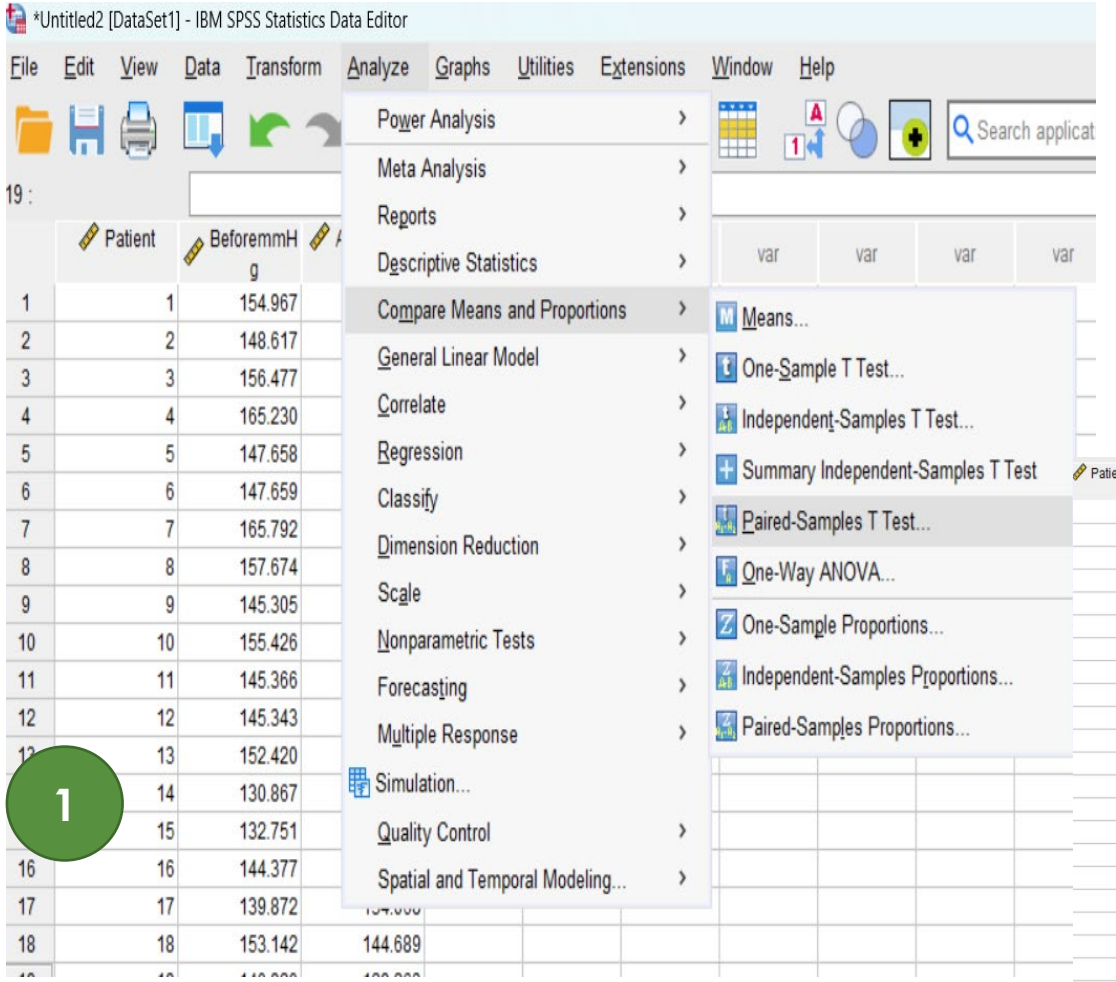




Parametric Paired Sample Test: Student's t test

Data Set: large_blood_pressure.xlsx

Analyze > Compare Means > Paired-Samples T Test

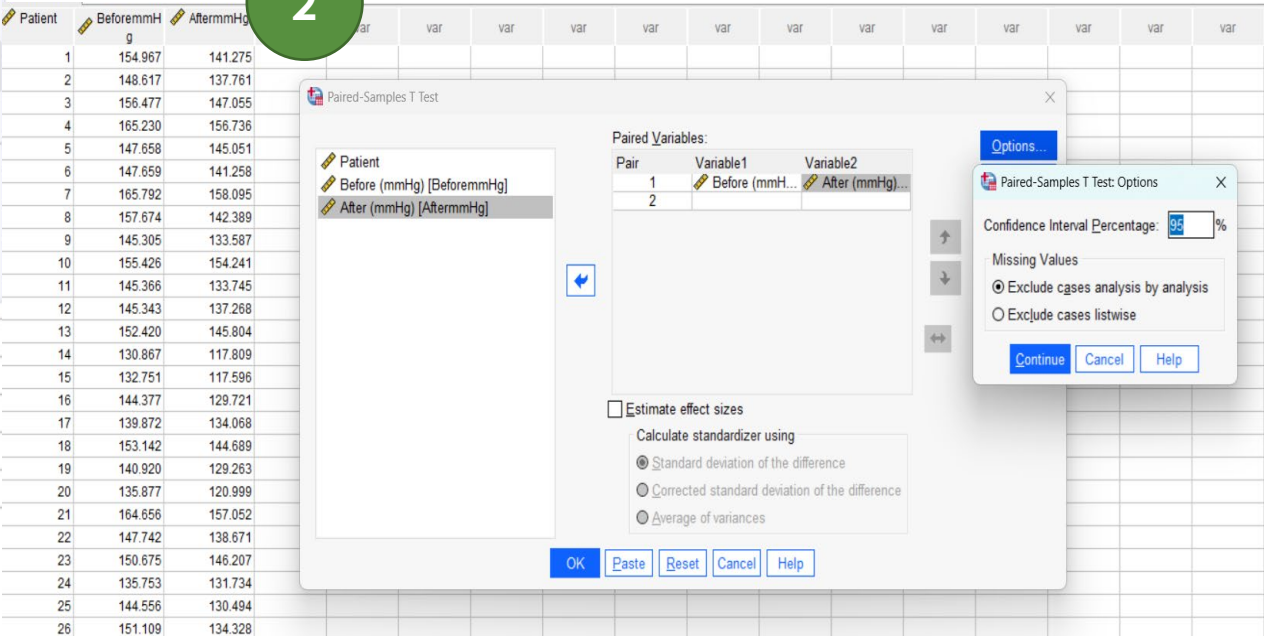


1

Select the variable as Pair -> Option (to change the confidence level) -> OK

2: Option if you want to change the significance level, else by default it will be 95% significance level ($\alpha = 0.05$)

2



Parametric Paired Sample Test: Student's t test

Data Set: large_blood_pressure.xlsx

To summarize:

- Since the **p-value (1.13×10^{-15})** is much smaller than the significance level **$\alpha = 0.05$**, we **reject the null hypothesis**. This confirms that the drug has a **statistically significant** effect in reducing systolic blood pressure after 30 days.

Confidence Interval Analysis (Paired t-test)

The **95% confidence interval (CI) for the mean difference** in systolic blood pressure before and after taking the drug is:

(8.31,11.40)(8.31, 11.40)(8.31,11.40)

Interpretation

•We are **95% confident** that the **true mean reduction** in systolic blood pressure due to the drug lies **between 8.31 mmHg and 11.40 mmHg**.

•Since the **confidence interval does not include 0**, it further supports the conclusion that the drug has a **statistically significant effect** in reducing blood pressure.

Paired Samples Test

		Paired Differences					t	df	Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1	Before (mmHg) - After (mmHg)	9.854783	4.823838	.762716	8.312044	11.397521	12.921	39	<.001	<.001


LET US PRACTICE

Problem Statement:

A healthcare research institute is evaluating the effectiveness of a 12-week weight loss program. The institute recruited 50 participants and measured their Body Mass Index (BMI) before and after the program. The goal is to determine whether the weight loss program significantly reduces BMI. Test the hypothesis at a 5% significance level.

Research Question:

"Does the 12-week weight loss program significantly reduce participants' BMI?"

Data Set: BMI_Weight_Loss_Program.xlsx

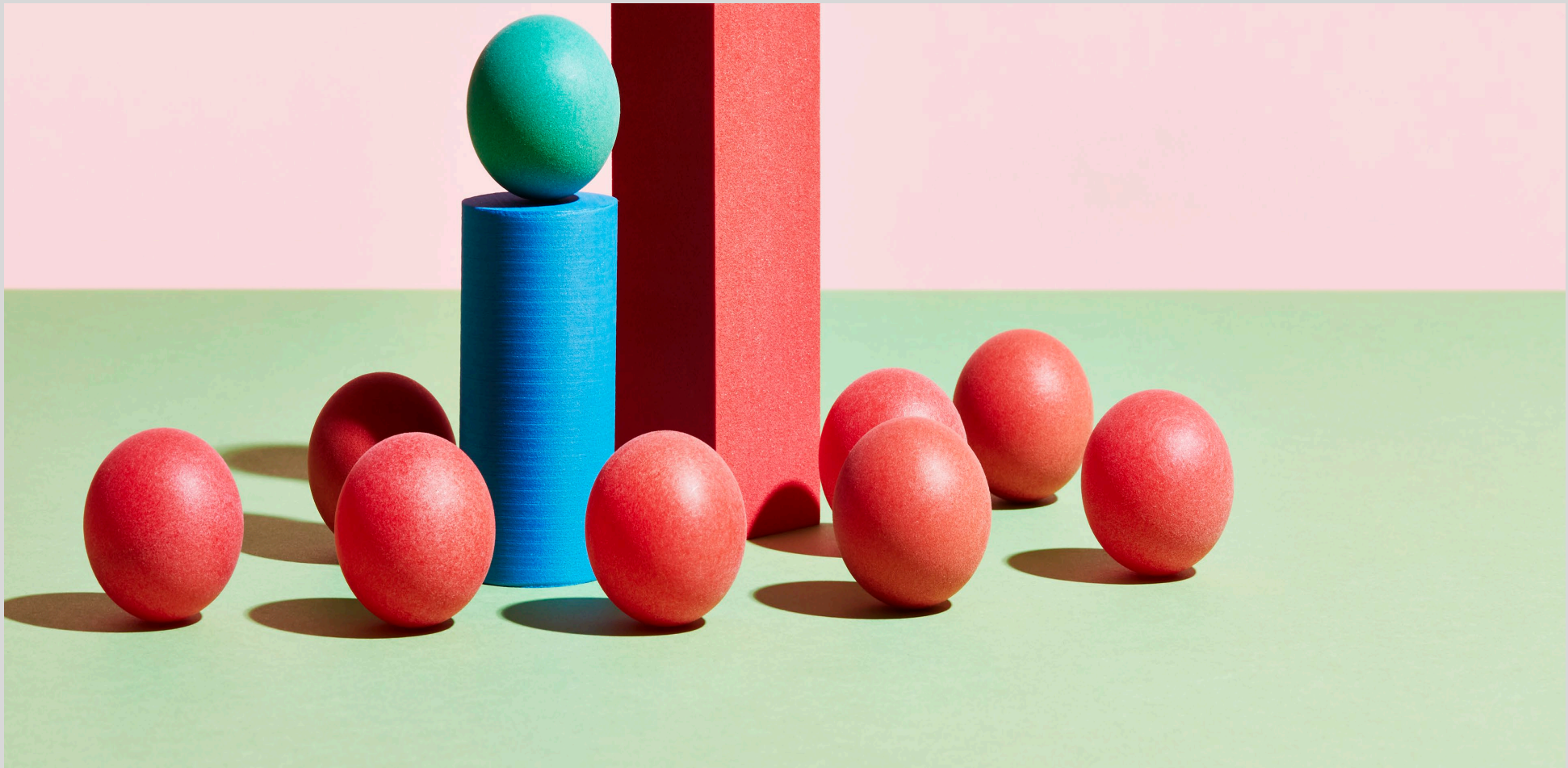
Hypothesis Statement:

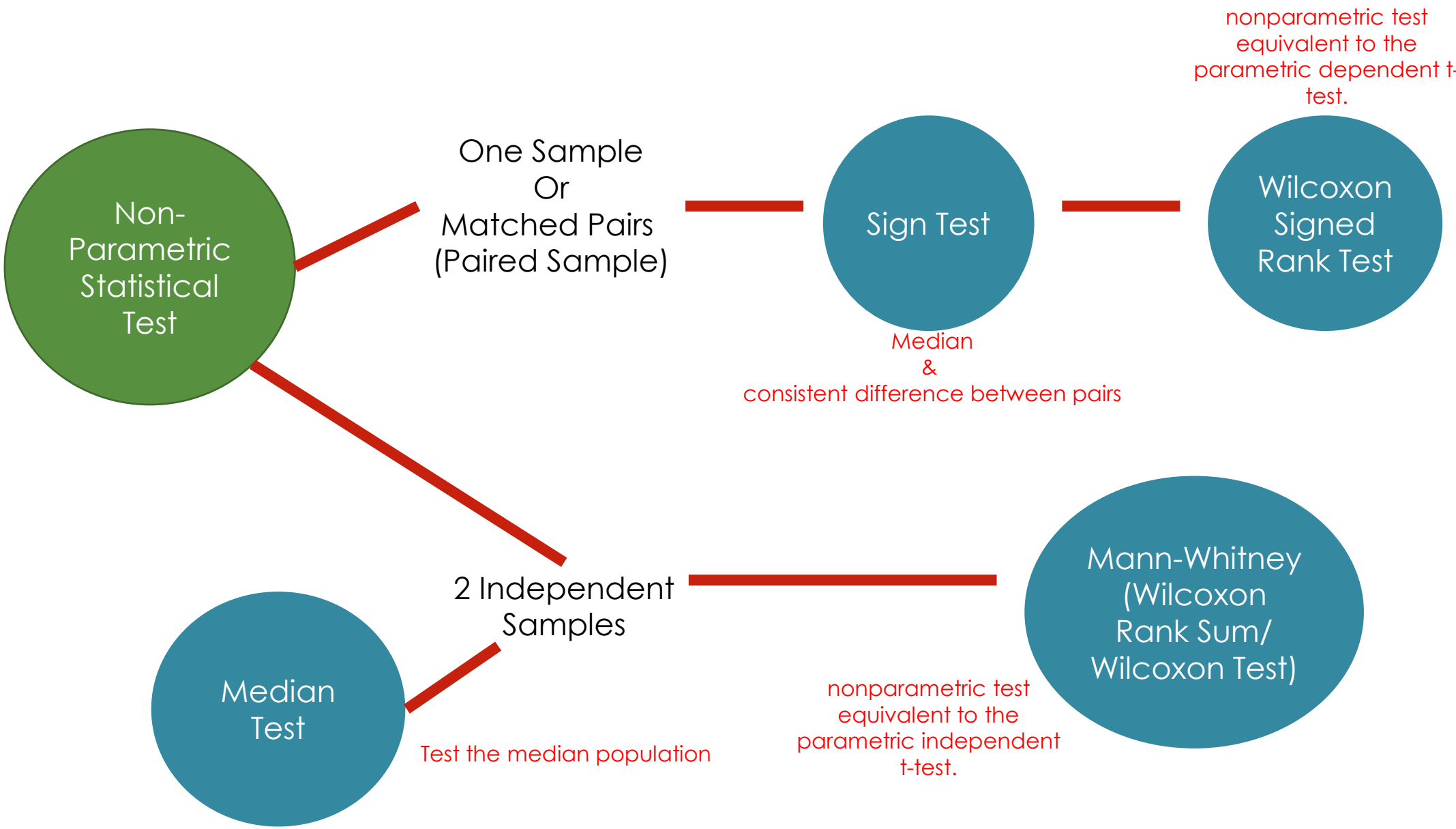
H_0 : *There are **no differences** on BMI reading before and after the treatment*

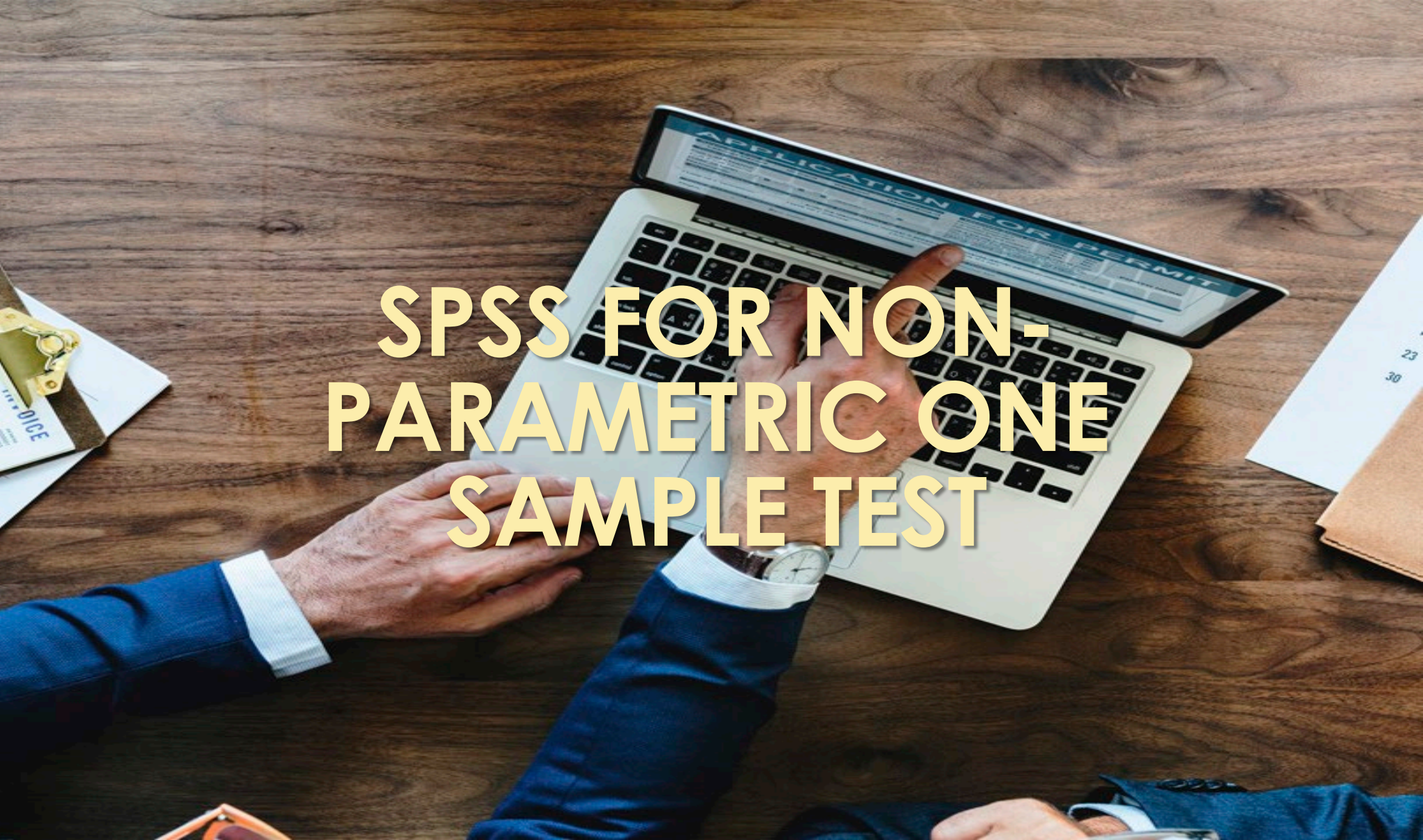
H_A : *There are **differences** on BMI reading before and after the treatment*

This is two tailed hypothesis test

NON-PARAMETRIC STATISTICAL TEST




Non-Parametric Statistical Test
One Sample Or Matched Pairs (Paired Sample)
Sign Test
Wilcoxon Signed Rank Test
nonparametric test equivalent to the parametric dependent t-test.
Median & consistent difference between pairs
2 Independent Samples
Mann-Whitney (Wilcoxon Rank Sum/ Wilcoxon Test)
nonparametric test equivalent to the parametric independent t-test.
Median Test
Test the median population



SPSS FOR NON-PARAMETRIC ONE SAMPLE TEST

Problem Statement:

A hospital wants to evaluate whether a **new pain relief drug** is effective. They collected **pain scores from 18 patients** after receiving the drug. The pain score is measured on a **scale from 0 (no pain) to 100 (extreme pain)**. A pain relief treatment is considered effective if the **median pain score** after treatment is **less than 40**.

Data Set: Pain_Scores.xlsx

Research Question:

"Does the new pain relief drug result in a median pain score lower than 40?"

H_0: The population median pain score is ≥ 40.

H_a: The population median pain score is < 40.

Checking Assumption:

- Sample size? n = 13
- Normal distribution or not?

Descriptive Statistics and Normality Test

Data Set: pain_scores.xlsx

Normality Test Assessment

Hypothesis Statement for normality test

H_0 : Variable is normally distributed

H_A : Variable is Not Normally distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pain Score (After Drug)	.132	18	.200*	.964	18	.679

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From the result of Shapira Wilk (due to sample size < 50) since both **p-values > 0.05**, we **fail to reject the null hypothesis of normality**, indicating that variable pain_scores **follow a normal distribution**.

However, since the sample size (n= 18) is small, therefore using parametric test is not suitable.

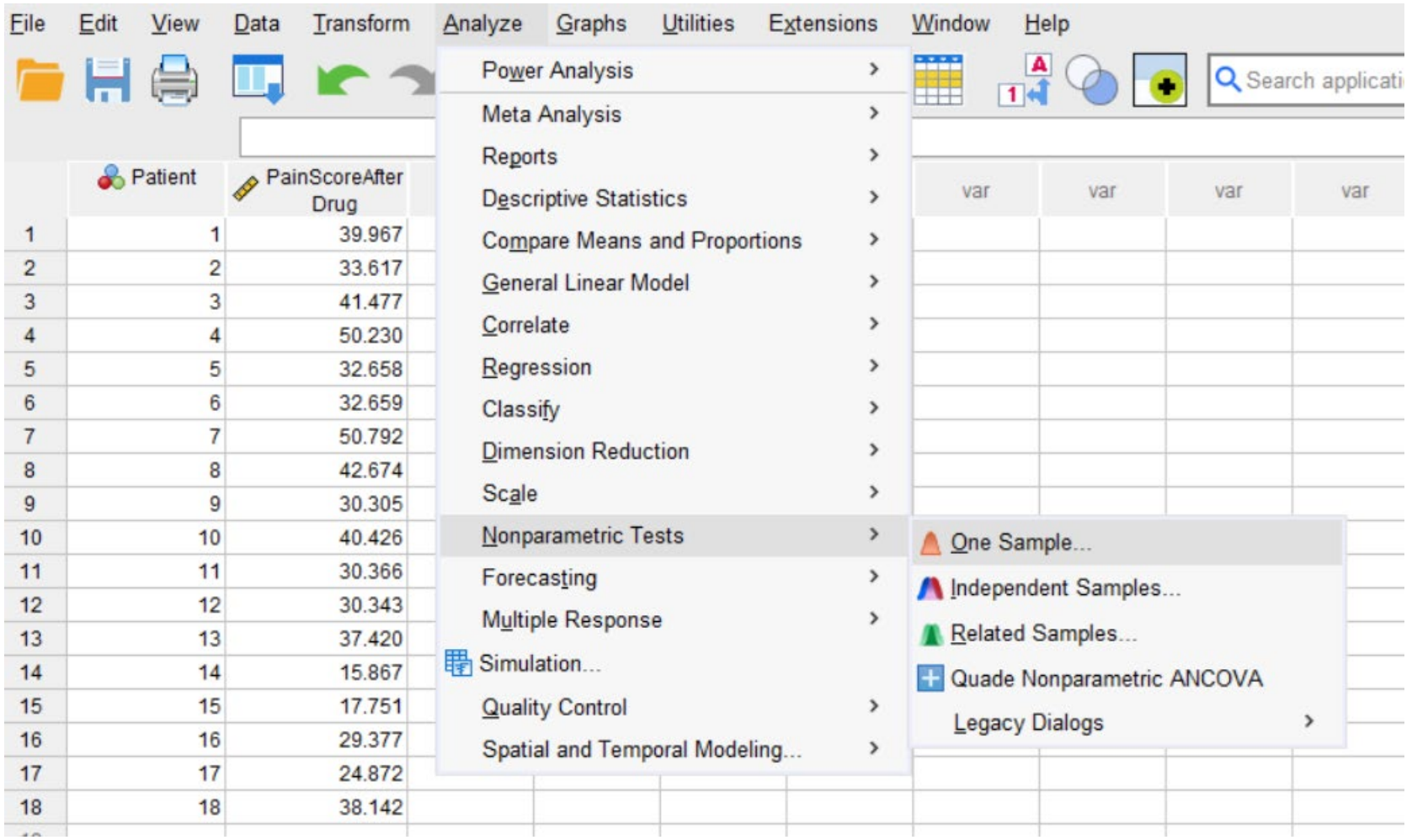
Small sample size, n (< 30), non-parametric tests are still widely used as they are **robust and do not rely on strict assumptions**.

Descriptive Statistics and Normality Test

Data Set: pain_scores.xlsx

1. Analyze > Nonparametric Tests > One Sample..

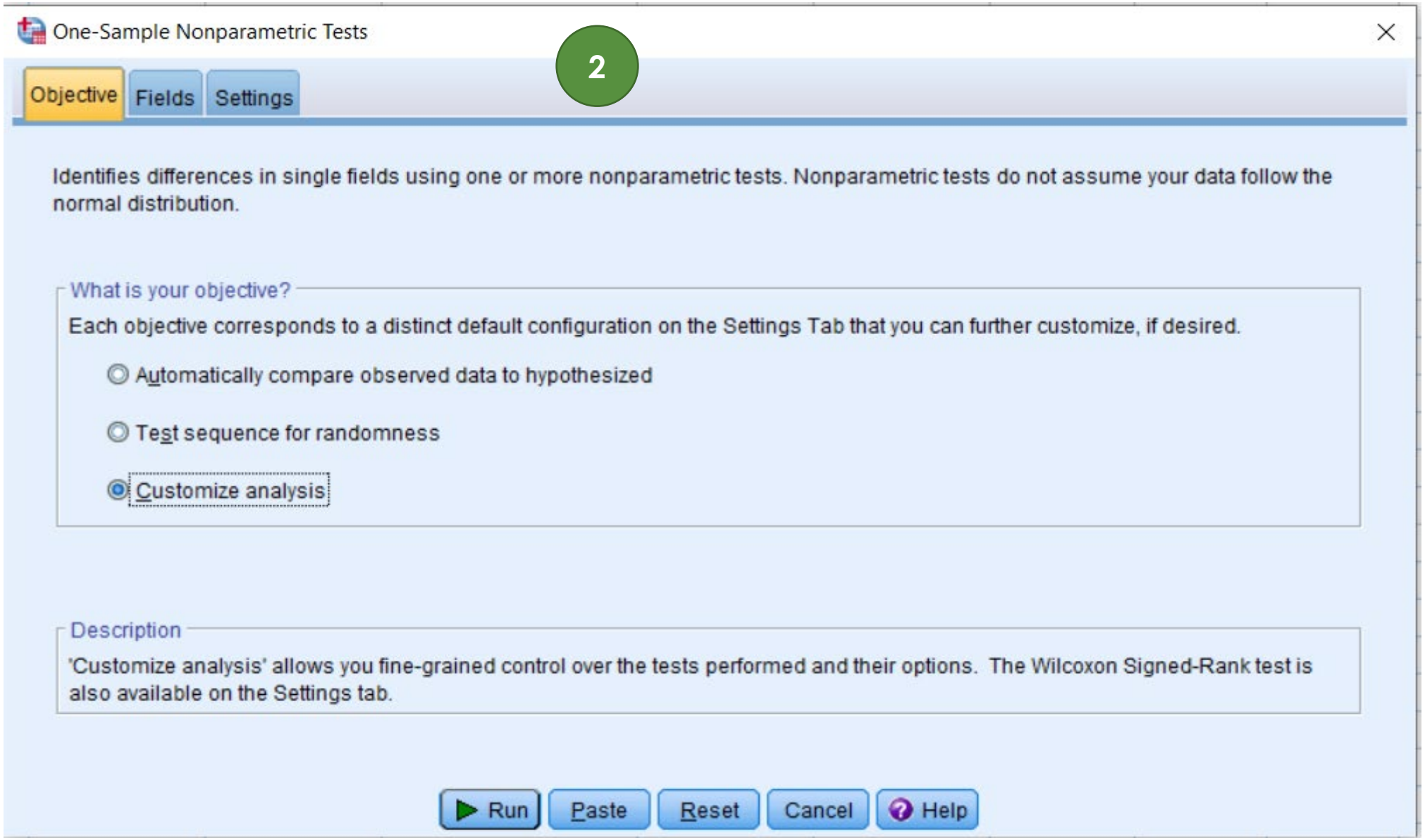
1



Descriptive Statistics and Normality Test

Data Set: pain_scores.xlsx

2. Objective tab, checked Customize Analysis

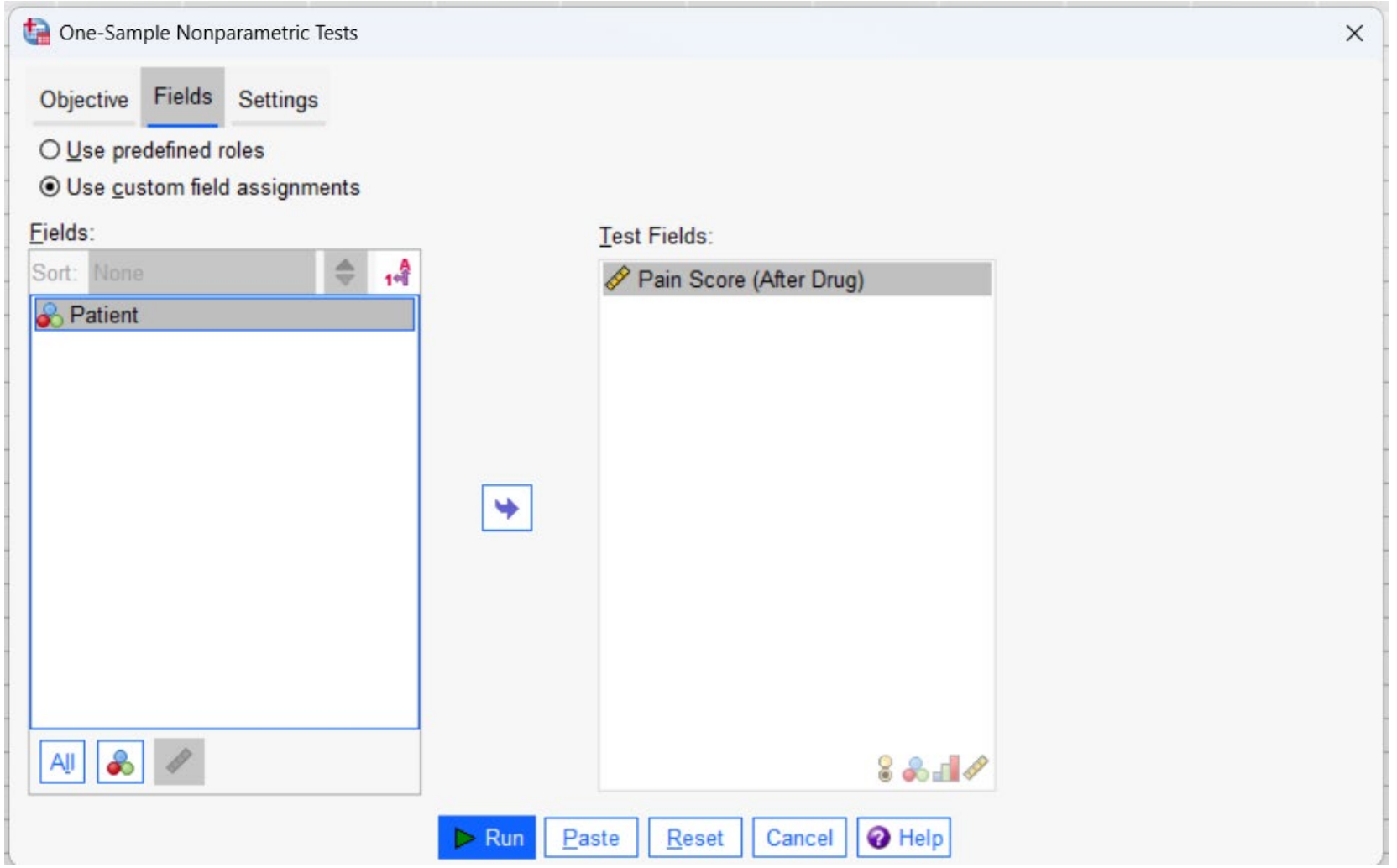


Descriptive Statistics and Normality Test

Data Set: pain_scores.xlsx

3

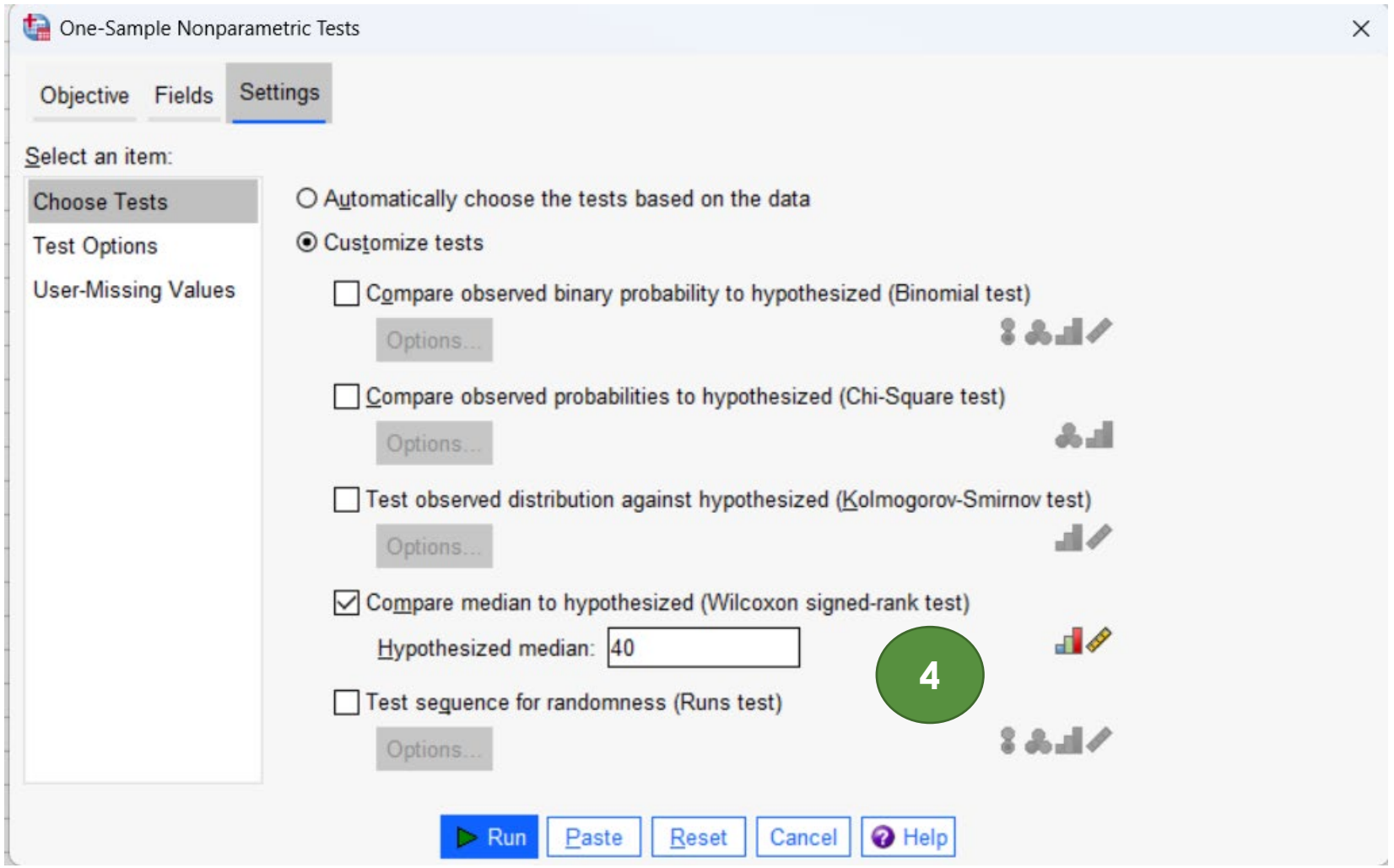
3. Fields tab, checked Use custom field assignments. Then select Rating family car as Test Field



Descriptive Statistics and Normality Test

Data Set: pain_scores.xlsx

4. Settings tab, checked Customize Tests. Then select compare median to hypothesized. The hypothesized median is 40 (as stated in the hypothesis).
5. Click Run



Descriptive Statistics and Normality Test

Data Set: pain_scores.xlsx

Hypothesis Statement:

H_0: The population median pain score is ≥ 40.

H_a: The population median pain score is < 40.

Hypothesis Test Summary

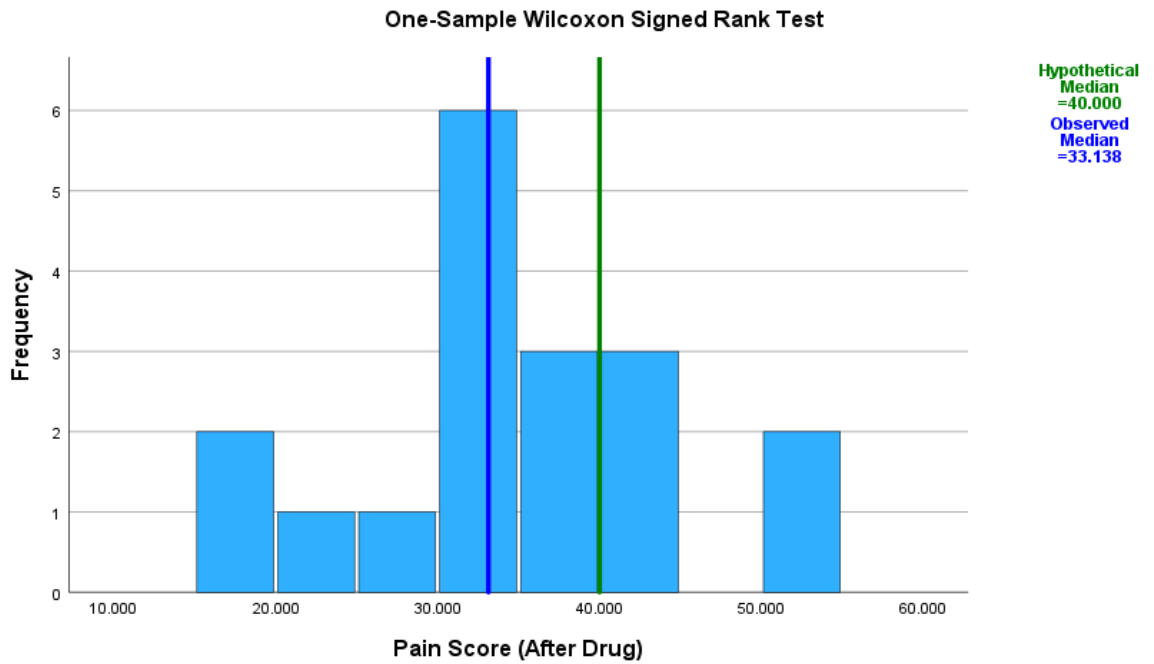
	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The median of Pain Score (After Drug) equals 40.000.	One-Sample Wilcoxon Signed Rank Test	.043	Reject the null hypothesis.

a. The significance level is .050.

b. Asymptotic significance is displayed.

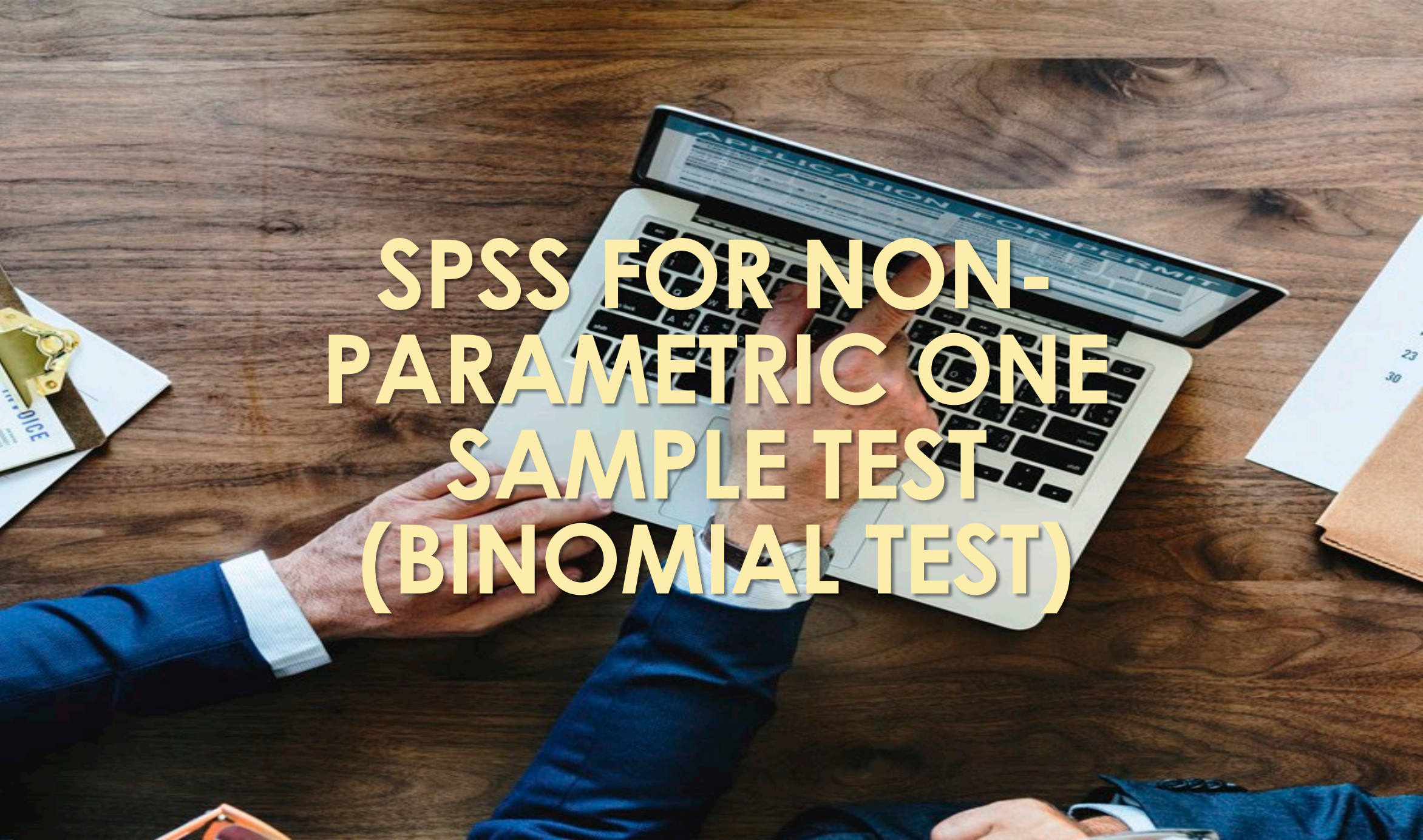
One-Sample Wilcoxon Signed Rank Test Summary

Total N	18
Test Statistic	39.000
Standard Error	22.962
Standardized Test Statistic	-2.025
Asymptotic Sig.(2-sided test)	.043



To summarize:

Since the **p-value (0.0216) < 0.05**, we **reject the null hypothesis** at a **5% significance level**. This indicates that the **pain relief drug significantly reduces pain scores, with a median score lower than 40**.



SPSS FOR NON-PARAMETRIC ONE SAMPLE TEST (BINOMIAL TEST)

Problem Statement:
A public health department claims that a new **vaccine prevents infection in at least 80%** of vaccinated individuals. To test this claim, researchers **randomly select 20 vaccinated individuals** and record whether they contracted the infection (**Yes/No**).

Research Question:
"Is the proportion of vaccinated individuals who remain infection-free lower than 80%?"

Data Set: Vaccine_Effectiveness_Outcomes.xlsx

Hypothesis Statement:

H_0: The population proportion of vaccinated individuals who remain infection-free is ≥ 80%.

H_a: The population proportion of vaccinated individuals who remain infection-free is < 80%.

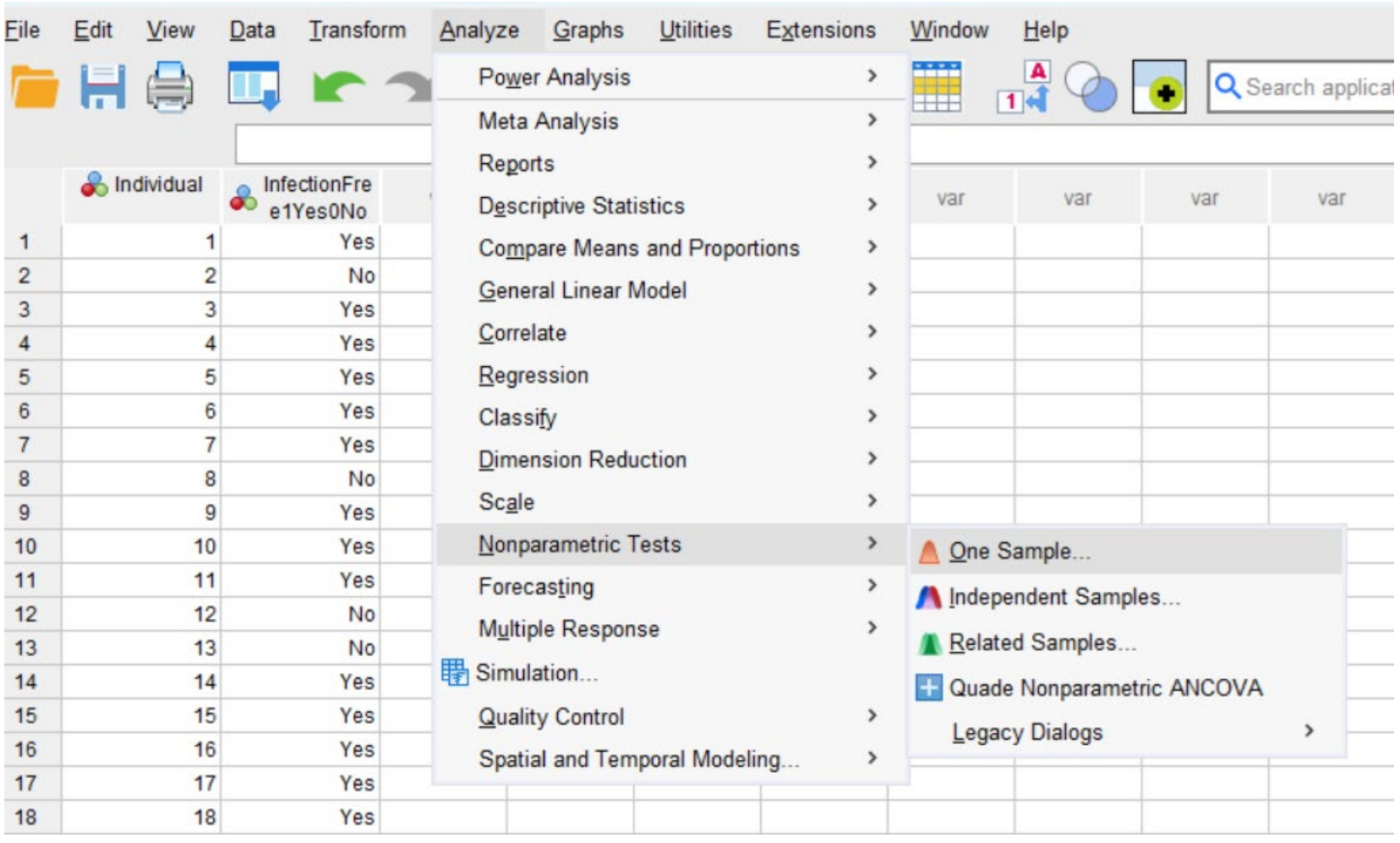
This is one tailed hypothesis test

Non Parametric One-Sample Population Proportion Statistical Test: Binomial test

Data Set: Vaccine_Effectiveness_Outcomes.xlsx

1. Analyze > Nonparametric Tests > Legacy Dialogs > Binomial

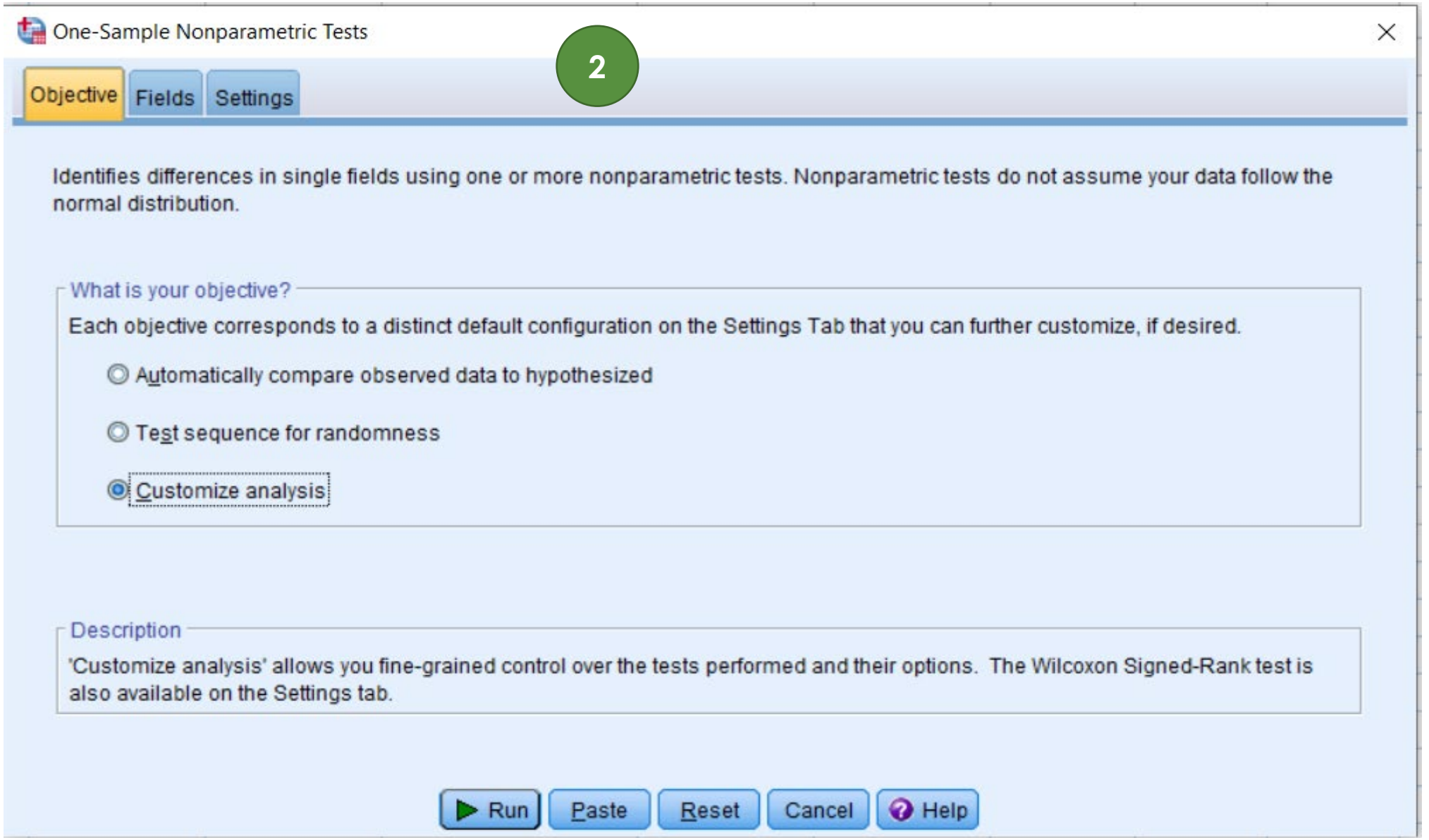
1



Non Parametric One-Sample Population Proportion Statistical Test: Binomial test

Data Set: Vaccine_Effectiveness_Outcomes.xlsx

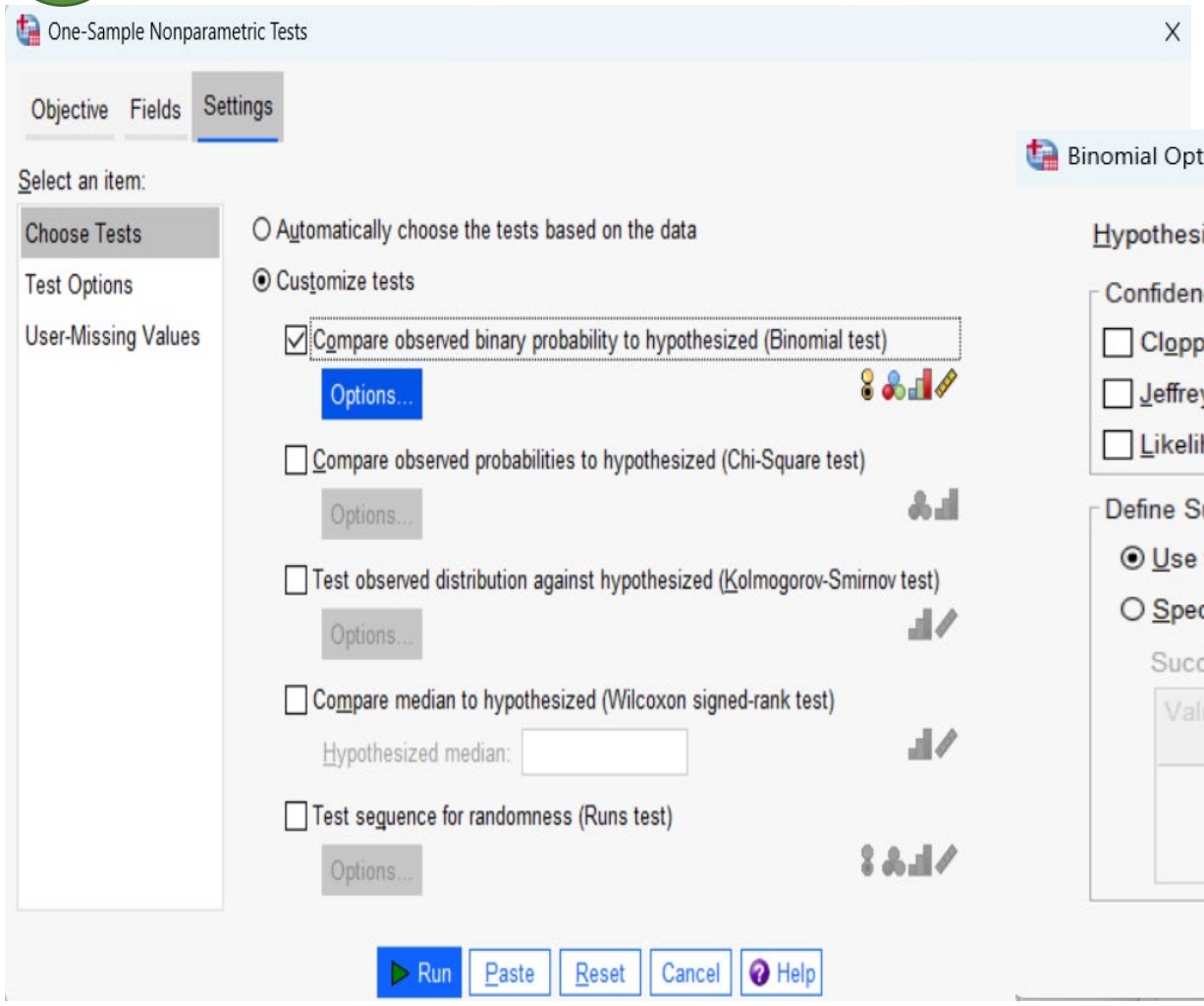
2. Objective tab, checked Customize Analysis



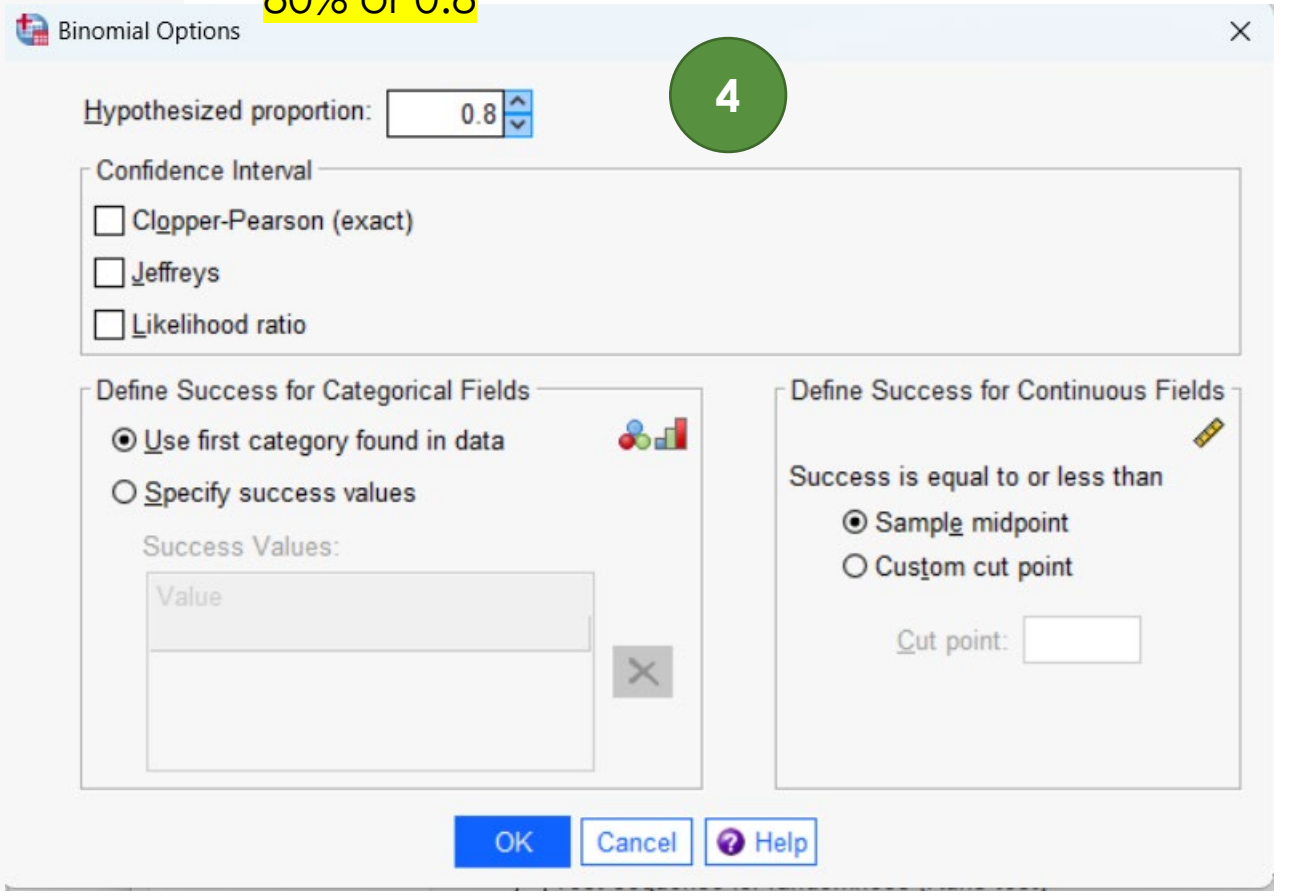
Non Parametric One-Sample Population Proportion Statistical Test: Binomial test

Data Set: Vaccine_Effectiveness_Outcomes.xlsx

3



4. Change the hypothesized proportion to 80% or 0.8



Non Parametric One-Sample Population Proportion Statistical Test: Binomial test

Data Set: Vaccine_Effectiveness_Outcomes.xlsx

H_0: The population proportion of vaccinated individuals who remain infection-free is ≥ 80%.

H_a: The population proportion of vaccinated individuals who remain infection-free is < 80%.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The categories defined by Infection-Free (1=Yes, 0=No) = Yes and No occur with probabilities .800 and .200.	One-Sample Binomial Test	.589[c]	Retain the null hypothesis.

a. The significance level is .050.

b. Asymptotic significance is displayed.

c. Exact significance is displayed for this test.

2

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)
Infection-Free (1=Yes, 0=No)	Group 1	Yes	16	.8	.8	.589[a]
	Group 2	No	4	.2		
	Total		20	1.0		

a. Alternative hypothesis states that the proportion of cases in the first group < .8.

4

3

1

To summarize:

1. N: 16 out of 20 cases are vaccinated individuals who remain infection-free
2. The observed proportion for vaccinated individuals who remain infection-free is 0.8 or 80%
3. The hypothesized test proportion is 0.8 or 80%
4. p(denoted as "Exact Significance (1-tailed)") = 0.589: We generally ACCEPT our null hypothesis that the population proportion of vaccinated individuals who remain infection-free is ≥ 80%.


LET US PRACTICE

Problem Statement:

A healthcare research institute is evaluating the effectiveness of a 12-week weight loss program. The institute recruited 50 participants and measured their Body Mass Index (BMI) before and after the program. The goal is to determine whether the weight loss program significantly reduces BMI. Test the hypothesis at a 95% significance level.

Research Question:

"Does the 12-week weight loss program significantly reduce participants' BMI?"

Data Set: BMI_Weight_Loss_Program.xlsx

Hypothesis Statement:

H_0 : *There are **no differences** on BMI reading before and after the treatment*

H_A : *There are **differences** on BMI reading before and after the treatment*

This is two tailed hypothesis test

SPSS FOR NON-PARAMETRIC INDEPENDENT SAMPLE TEST

Problem Statement:
A hospital is conducting a study to compare the effectiveness of two different pain management treatments for patients with chronic lower back pain. The study involves **two independent groups**:
Group A: Patients receiving a new non-invasive pain relief therapy
Group B: Patients receiving conventional physiotherapy
Each patient undergoes only **one** of the two treatments. After completing the treatment regimen, the pain level of each patient is measured using a standardized pain scale. **Is there a significant difference in post-treatment pain levels between the two independent groups?**

Data Set: pain_management.xlsx

Hypothesis Statement:

H_0 : The pain score between group A and group B are equal

H_A : The pain score between group A and group B are not equal

This is two tailed hypothesis test

Normality Test

Data Set: pain_management.xlsx

Normality Test Assessment

Hypothesis Statement for normality test

H_0 : Variable is normally distributed

H_A : Variable is Not Normally distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pain Score	.140	120	<.001	.948	120	<.001

a. Lilliefors Significance Correction

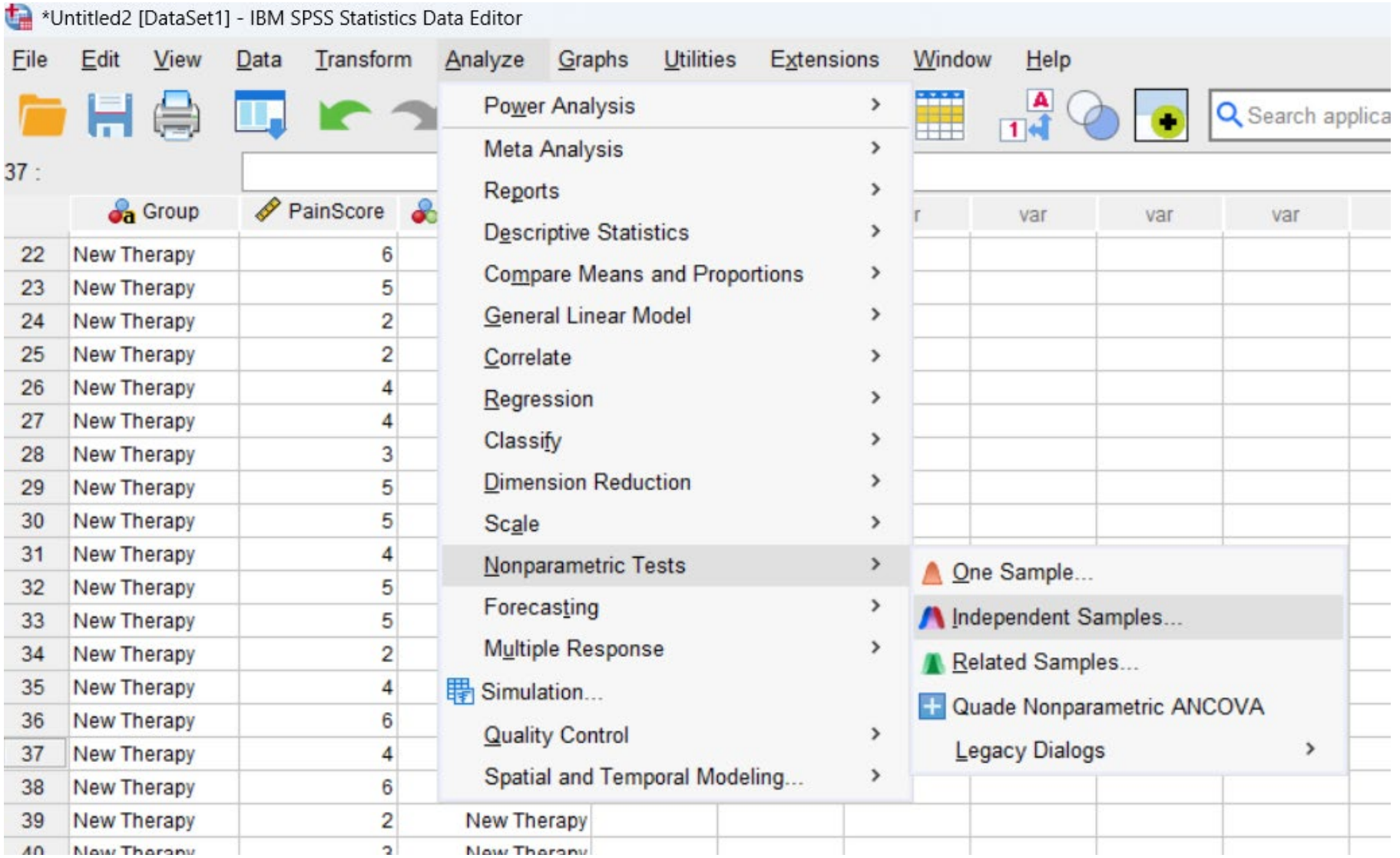
From the result of Kalmogorov Smirnov(due to sample size >50) since the **p-values < 0.05**, we **reject the null hypothesis of normality**, indicating that variable pain_scores **does not follow a normal distribution**.

Non Parametric Independent-Sample statistical Test: Mann-Whitney Test

Data Set: pain_management.xlsx

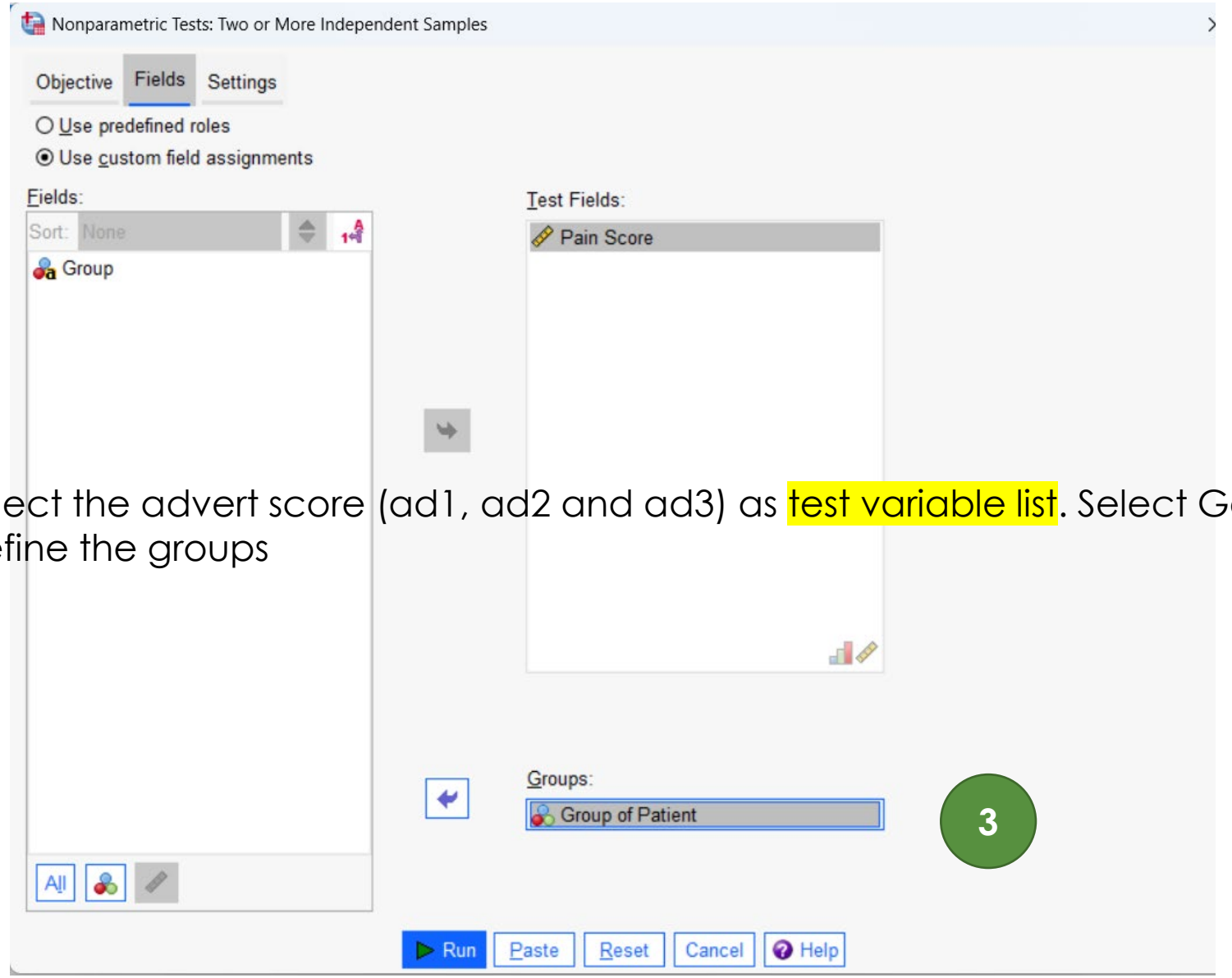
1. Analyze > Nonparametric Tests > 2 Independent Samples

1



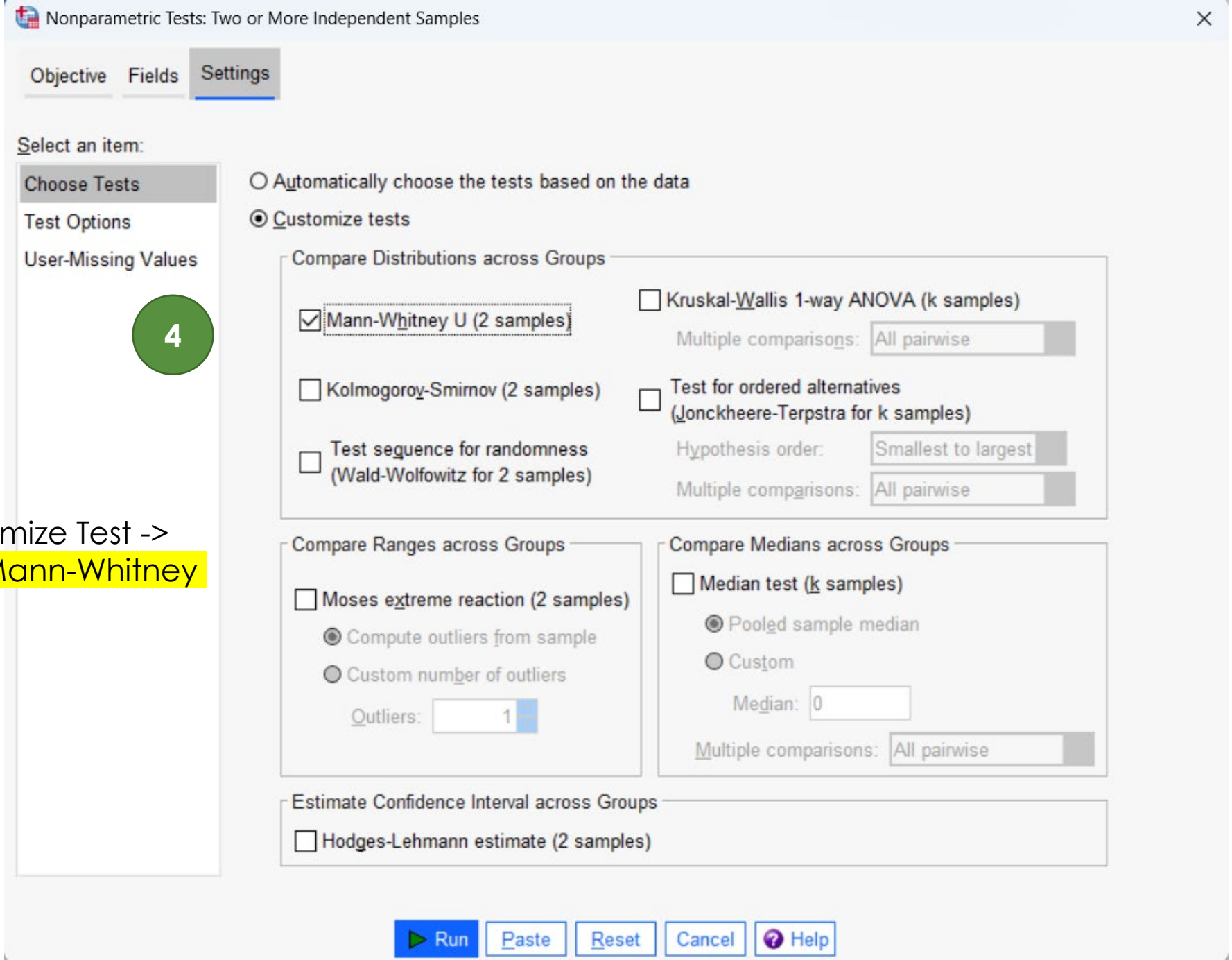
Non Parametric Independent-Sample statistical Test: Mann-Whitney Test

Data Set: adratings.sav



2. Select the advert score (ad1, ad2 and ad3) as test variable list. Select Gender as grouping variable.
3. Define the groups

4. Customize Test -> Select Mann-Whitney U .



OUTPUT→ Non Parametric Independent-Sample statistical Test: Mann-Whitney Test

Data Set: adratings.sav

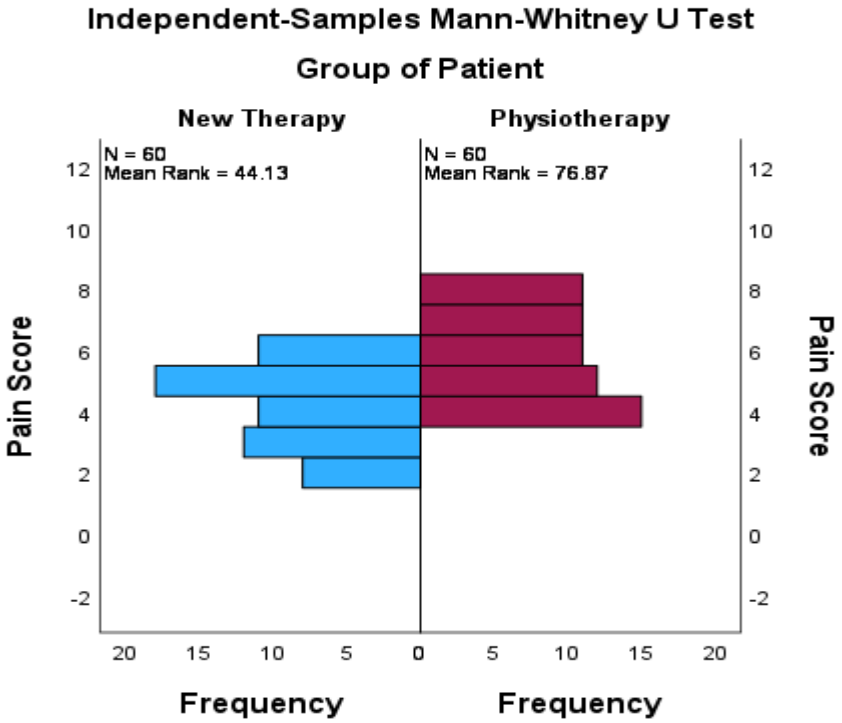
H_0 : The pain score between group A and group B are equal
H_A : The pain score between group A and group B are not equal

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.[a,b]	Decision
1	The distribution of Pain Score is the same across categories of Group of Patient.	Independent-Samples Mann-Whitney U Test	<.001	Reject the null hypothesis.

a. The significance level is .050.

b. Asymptotic significance is displayed.



Since the **p-value is extremely small (< 0.05)**, we reject the null hypothesis and conclude that **there is a significant difference in pain levels between the two groups**. This supports the claim that **the new therapy differs significantly from conventional physiotherapy in terms of pain reduction**.


LET US PRACTICE

Research Question:

Is there any difference for the BMI between men and women are equal?

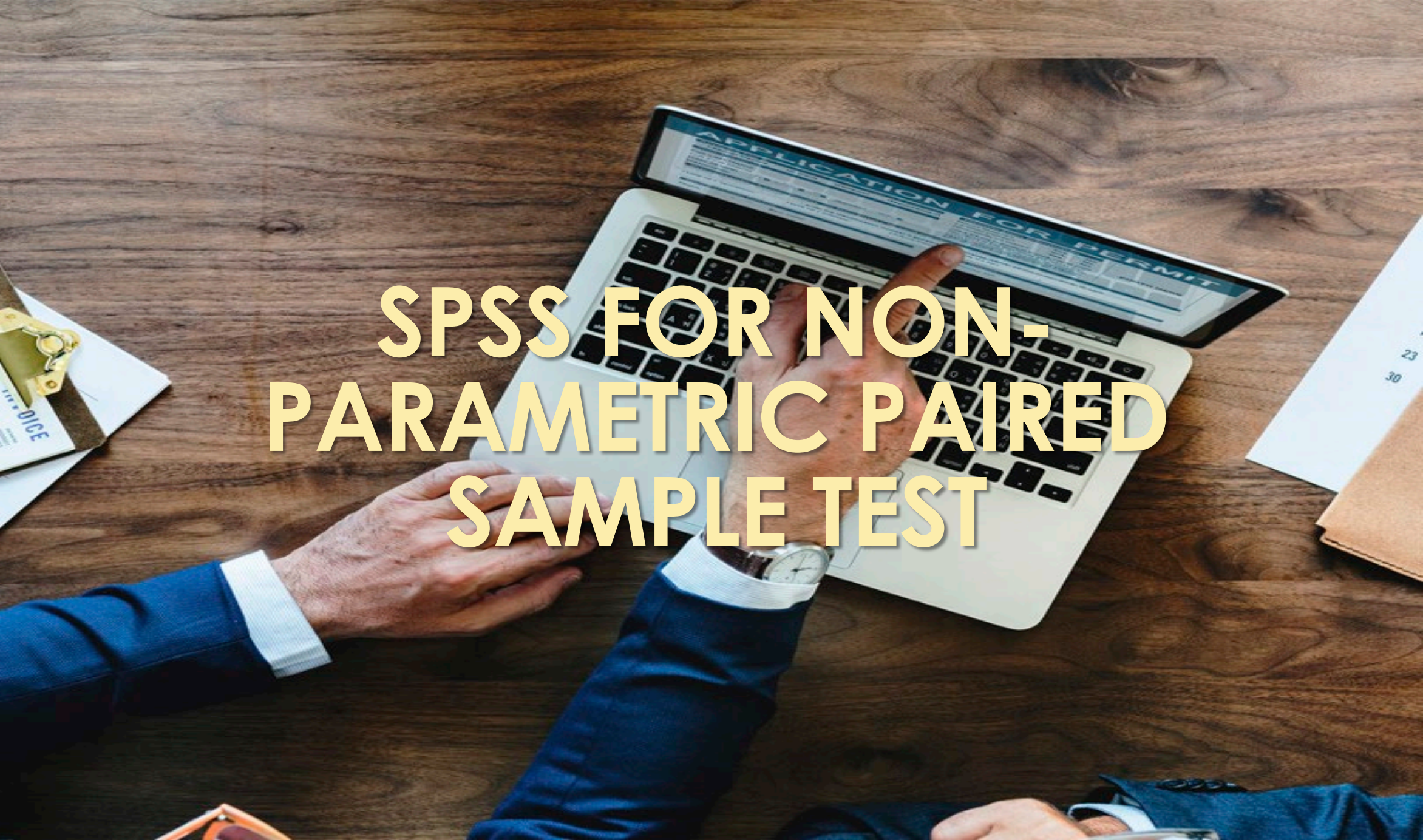
Data Set: BMI.sav

Hypothesis Statement:

H_0 : The BMI between men and women are equal

H_A : The BMI between men and women are not equal

This is two tailed hypothesis test



SPSS FOR NON-PARAMETRIC PAIRED SAMPLE TEST

Research Question:

Does arterial plague within blood vessel thickness before & after treatment shows a significantly difference?

Data Set: arterial plague experiment.sav

Hypothesis Statement:

H_0 : There are no significantly difference on median blood vessel thickness before & after treatment ($Median_{post-pre} = 0$)

H_A : There are significantly difference on median blood vessel thickness before & after treatment

($Median_{post-pre} \neq 0$)

This is two tailed hypothesis test

Normality test, Histogram and Descriptive Statistics

Data Set: arterial plague experiment.sav

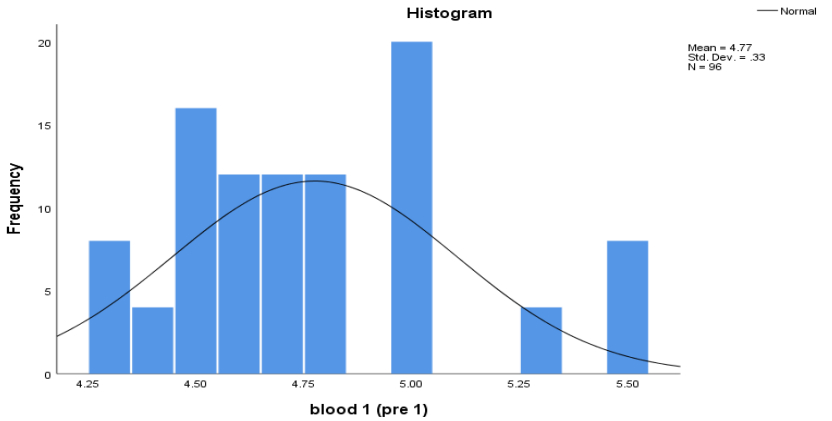
Tests of Normality

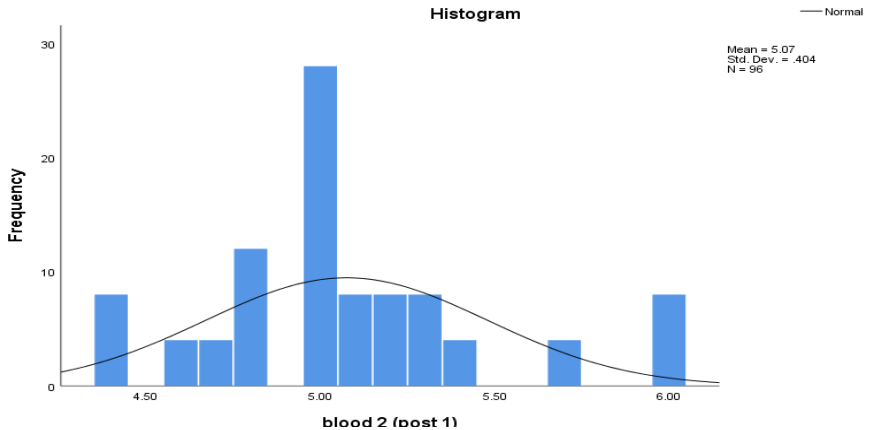
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
blood 1 (pre 1)	.136	96	.000	.916	96	.000
blood 2 (post 1)	.157	96	.000	.915	96	.000
Diff_pla	.179	96	.000	.901	96	.000

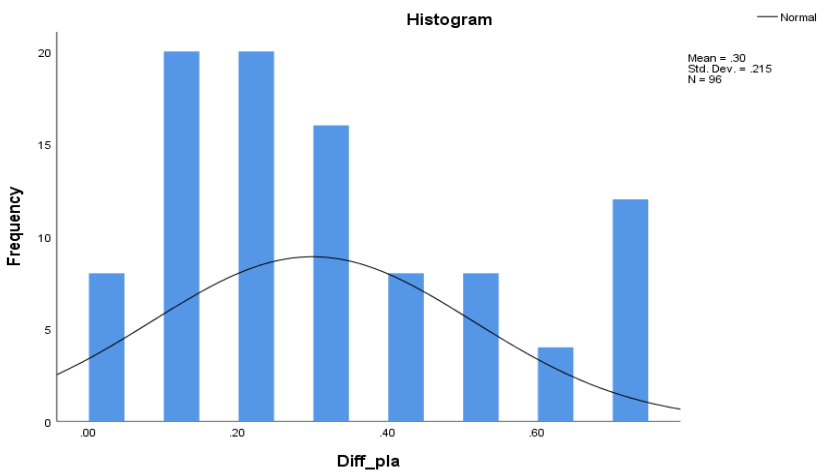
a. Lilliefors Significance Correction

Statistics

		blood 1 (pre 1)	blood 2 (post 1)	Diff_pla
N	Valid	96	96	96
	Missing	0	0	0
Median		4.7000	5.0000	.2500
Skewness		.714	.681	.595
Std. Error of Skewness		.246	.246	.246
Kurtosis		-.111	.484	-.735
Std. Error of Kurtosis		.488	.488	.488
Range		1.20	1.60	.70
Percentiles	25	4.5000	4.8000	.1000
	50	4.7000	5.0000	.2500
	75	5.0000	5.2750	.4750

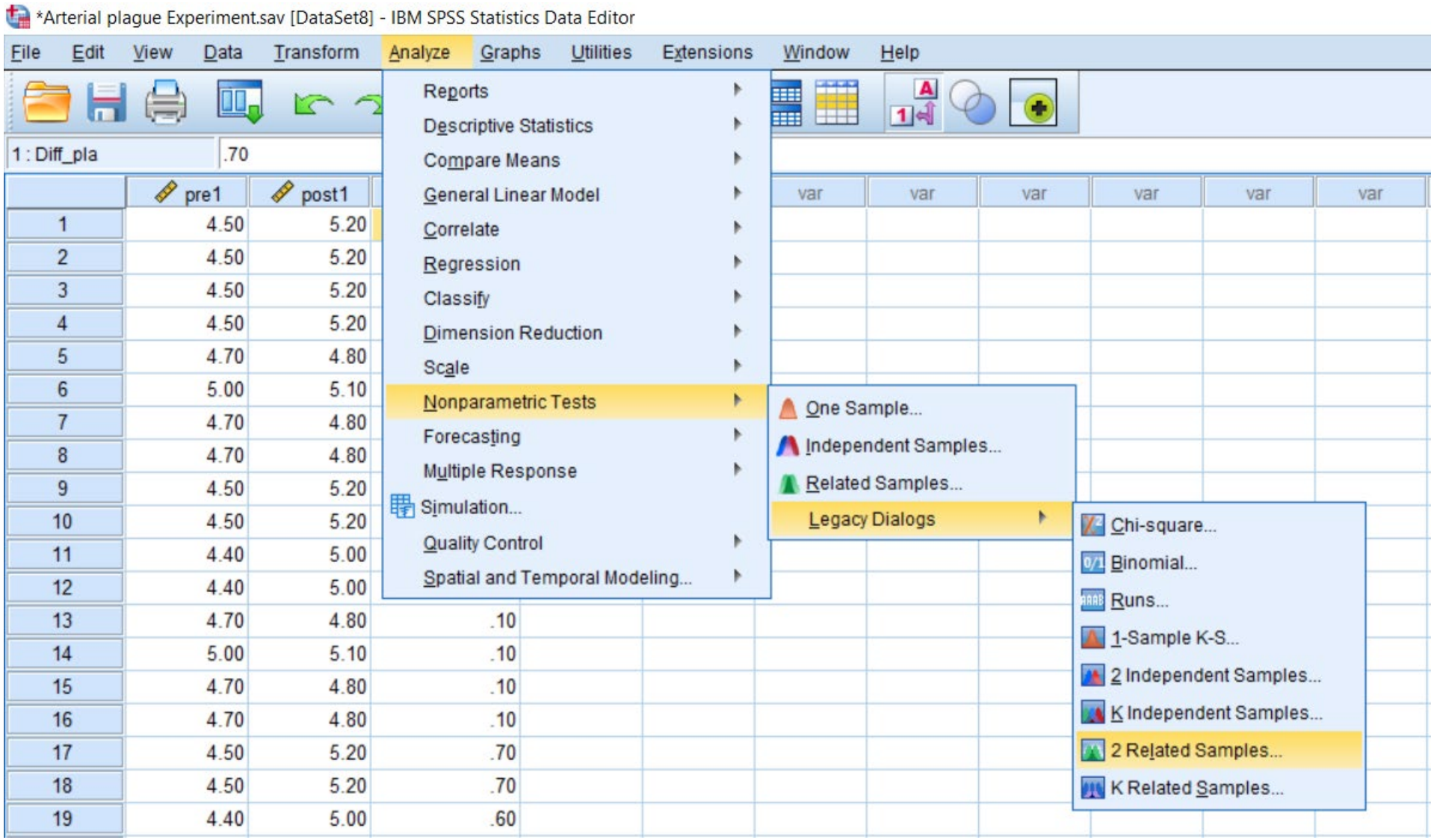






Non Parametric Paired-Sample statistical Test: Wilcoxon SR Test

Data Set: arterial plague experiment.sav



OUTPUT→ Non Parametric Paired-Sample statistical Test: Wilcoxon SR Test

Data Set: arterial plague experiment.sav

Positive rank = post (after) is higher than pre (before)
Negative rank = post (after) is lower than pre (before)
Ties = post (after) is similar with pre (before)

*** similar population distributions → when the signs (plus and minus) is distributed roughly evenly over ranks.

Ranks

		N	Mean Rank	Sum of Ranks
blood 2 (post 1) - blood 1 (pre 1)	Negative Ranks	0[a]	.00	.00
	Positive Ranks	88[b]	44.50	3916.00
	Ties	8[c]		
	Total	96		

a. blood 2 (post 1) < blood 1 (pre 1)

b. blood 2 (post 1) > blood 1 (pre 1)

c. blood 2 (post 1) = blood 1 (pre 1)

The sum of positive ranks is way larger (3916) than the sum of negative ranks. This indicate that the arterial plague within blood vessel thickness shows a difference after the treatment.

Test Statistics[a]

	blood 2 (post 1) - blood 1 (pre 1)
Z	-8.181[b]
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

$H_0 : Median_{post-pre} = 0$
$H_A : Median_{post-pre} \neq 0$

P-value is smaller than a = 0.05, therefore, H-null is rejected. This indicate that there is a difference on median blood vessel thickness before & after treatment


LET US PRACTICE

Problem Statement:
A hospital is conducting a study to assess the effectiveness of a new blood pressure medication in reducing systolic blood pressure (SBP) among patients with hypertension. The study involves a **dependent (paired) sample design**, where **the same group of patients** is measured **before and after** taking the medication for a specified period.

Research Question:
Is there a significant difference in patients' systolic blood pressure before and after taking the new medication?

Hypothesis Statement:

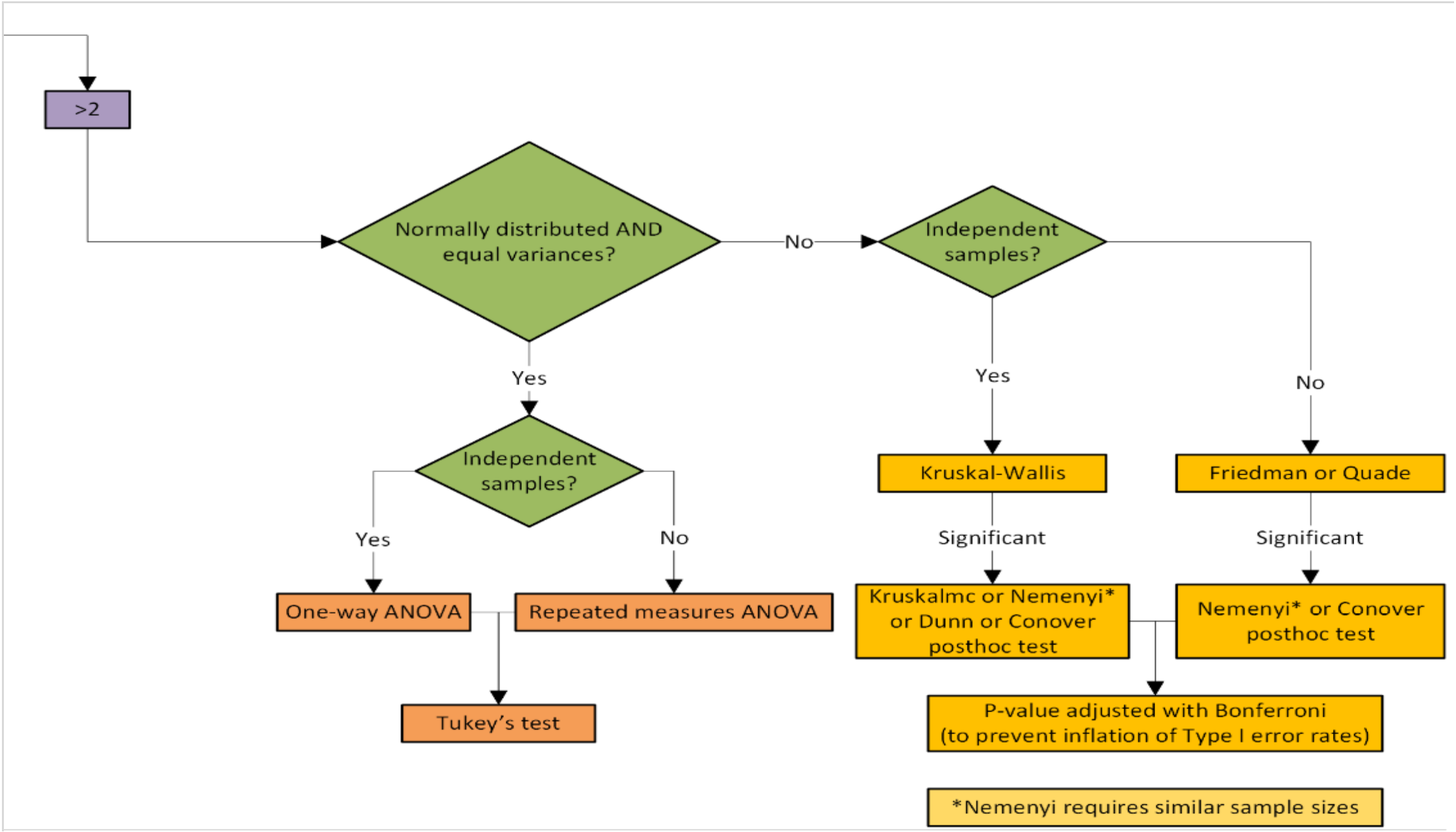
H_0 : There are no significantly difference on systolic blood pressure (SBP) before & after treatment ($Median_{post-pre} = 0$)

H_A : There are significantly difference on systolic blood pressure (SBP) before & after treatment ($Median_{post-pre} \neq 0$)

Data Set: Blood_Pressure_Study.xlsx

THE JOURNEY OF STATISTICAL TEST DOES NOT STOP HERE…

1. One Way ANOVA: to determine whether there are any statistically significant differences between the means of three or more independent (unrelated) groups.
2. Two-Way ANOVA: is used to estimate how the mean of a quantitative variable changes according to the levels of two categorical variables.
3. MANOVA: an ANOVA with several dependent variables.
4. ANCOVA: is a blend of analysis of variance (ANOVA) and regression. It is similar to factorial ANOVA


>2
Normally distributed AND equal variances?
No
Independent samples?
Yes
Yes
No
Independent samples?
Yes
No
Kruskal-Wallis
Friedman or Quade
Significant
Significant
One-way ANOVA
Repeated measures ANOVA
Kruskalmc or Nemenyi* or Dunn or Conover posthoc test
Nemenyi* or Conover posthoc test
Tukey's test
P-value adjusted with Bonferroni (to prevent inflation of Type I error rates)
*Nemenyi requires similar sample sizes

THANK YOU