Kruskal-Wallis –Test

A graduation research submitted to the council Department of Mathematics in the college of Education for pure sciences/ University of Babylon, as part of the requirements obtaining a bacheloris drgree in mathematics

BY

Baneen Mujeeb Abd

Supervised By

Rawasy Adnan Hameed

بِسْمِ اللهِ الرَّحْمَنِ الرَّحِيمِ

((يَرْفَعِ اللَّهُ الَّذِينَ ءَامَنُواْ مِنكُمْ وَالَّذِينَ أُوتُواْ الْعِلْمَ دَرَجَاتٍ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ))

(صدق الله العلي العظيم)

(سورة المجادلة :الاية ١١)

الإهداء

الى رجل الكفاح ، إلى من زرع القِيم والمبادئ الإسلامية ، إلى من أفنى زهرة شبابه فَي تربِية أبنائه ... والدي الحبيب

إلى القلب النابض ، إلى رمز الحنان و الحب والتضحُية ، إلى من كانت دعواتها الصادقة سر نجاحَي ... أمي الغالية

اهدي اليهم هذا الجهد المتواضع ، سائلا الله العلي القدير ان ينفع به ، أنه سميع مجيب.

الشكر والتقدير

Abstract

The Kruskal Wallis test is the non-parametric alternative to the One Way ANOVA. Non parametric means that the test doesn't assume your data comes from a particular distribution. The H test is used when the assumptions for ANOVA aren't met (like the assumption of normality.

It is sometimes called the one-way ANOVA on ranks, as the ranks of the data values are used in the test rather than the actual data points.

The test determines whether the medians of two or more groups are different. Like most statistical tests, you calculate a test statistic and compare it to a distribution cut-off point. The test statistic used in this test is called the H statistic. The hypotheses for the test are:

H0: population medians are equal.

H1: population medians are not equal.

The Kruskal Wallis test will tell you if there is a significant difference between groups. However, it won't tell you which groups are different. For that, you'll need to run a Post Hoc test.

Table of contents

Introduction:

The Kruskal–Wallis test by ranks, Kruskal–Wallis test (named after William Kruskal and W. Allen Wallis), or one-way ANOVA on ranks is a non-parametric method for testing whether samples originate from the same distribution. It is used for comparing two or more independent samples of equal or different sample sizes. It extends the Mann–Whitney U test, which is used for comparing only two groups. The parametric equivalent of the Kruskal–Wallis test is the one-way analysis of variance (ANOVA).

Asignificant Kruskal–Wallis test indicates that at least one sample stochastically dominates one other sample. The test does not identify where this stochastic dominance occurs or for how many pairs of groups stochastic dominance obtains. For analyzing the specific sample pairs for stochastic dominance, Dunn's test, pairwise Mann–Whitney tests with Bonferroni correction, or the more powerful but less well known Conover–Iman test are sometimes used.

It is supposed that the treatments significantly affect the response level and then there is an order among the treatments: one tends to give the lowest response, another gives the next lowest response is second, and so forth. Since it is a nonparametric method, the Kruskal–Wallis test does not assume a normal distribution of the residuals,

unlike the analogous one-way analysis of variance. If the researcher can make the assumptions of an identically shaped and scaled distribution for all groups, except for any difference in medians, then the null hypothesis is that the medians of all groups are equal, and the alternative hypothesis is that at least one population median of one group is different from the population median of at least one other group. Otherwise, it is impossible to say, whether the rejection of the null hypothesis comes from the shift in locations or group dispersions. This is the same issue that happens also with the Mann-Whitney test. If the data contains potential outliers, if the population distributions have heavy tails, or if the population distributions are significantly skewed, the Kruskal-Wallis test is more powerful at detecting differences among treatments than ANOVA F-test. On the other hand, if the population distributions are normal or are light-tailed and symmetric, then ANOVA F-test will generally have greater power which is the probability of rejecting the null hypothesis when it indeed should be rejected.

1-What is it Kruskal-Wallis Test?

The Kruskal-Wallis Test was developed by Kruskal and Wallis (1952).

Jointly and is named after them. The Kruskal-Wallis test is

a nonparametric (distribution free) test, and is used when the assumptions of one way ANOVA are not met. They both assess for significant differences on a continuous dependent variable by a grouping independent variable (with three or more groups). In the ANOVA, we assume that distribution of each group is normally distributed and there is approximately equal variance on the scores for each group. However, in the Kruskal-Wallis Test, we do not have any of these assumptions.

One way ANOVA is a statistical data analysis technique that is used to test the equality of the mean of three or more independent variable.

It is observed that selecting distinctive features through feature ranking significantly influences the performance of ML models. The maximum accuracy of fault classification is achieved with top 15 features ranked by the Kruskal–Wallis test for all the classifiers. The accuracies attained are 98.6, 99.4, 96.9, and 97.8% for Decision Tree, SVM, KNN, and ANN, respectively, with CWRU dataset. Similar trend is obtained through PU dataset giving corresponding maximum accuracies of 95.0, 97.8, 97.2, and 95.0%. The Kruskal–Wallis test outperforms the One-way ANOVA for both the data sources.

1-1: The Kruskal-Wallis H test (sometimes also called the "one-way (ANOVA on ranks"):

Is a rank-based nonparametric test that can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable. The Kruskal-Wallis one way ANOVA is a non-parametric method for comparing k independent samples. It is roughly equivalent to a parametric one way ANOVA with the data replaced by their ranks. Since ranking is conditional upon your observed values, so is this test. The null and alternative hypothesis in this case are

- Null hypothesis **:**The samples are from identical populations.

- Alternative hypothesis: The sample comes from different populations.

1-2: Kruskal-Wallis Tast-Assumptions:

- The assumptions of the kruskal-wallis test are similar those for the Mann-Whitney test.

- Samples are random samples,or allocation to treatment group is random.

- The tow samples are mutually independent.(independence within each sample and mutual independence among samples).

• The measurement scale is least ordinal ,and the variable is continuous.

• If the test is used as test of dominance,it has no distributional assumptions.If it used to compare medians,the distributions must be similar apart from their locations.

Example 1

Step by step example of the Kruskal-Wallis test:

Does physical exercise alleviate depression? We find some depressed people and check that they are all equivalently depressed to begin with. Then we allocate each person randomly to one of three groups: no exercise; 20 minutes of jogging per day; or 60 minutes of jogging per day. At the end of a month, we ask each participant to rate how depressed they now feel, on a Likert scale that runs from 1 ("totally miserable") through to 100 (ecstatically happy").

The appropriate test here is the Kruskal-Wallis test. We have three separate groups of participants, each of whom gives us a single score on a rating scale. Ratings are examples of an *ordinal* scale of measurement, and so the data are not suitable for a parametric test.

The Kruskal-Wallis test will tell us if the differences between the groups are so large that they are unlikely to have occurred by chance. Here are the data:

Ring on depression scale:

	No exercise	Jogging for 20 minutes	Jogging for 60 minutes
	23	22	59
	26	27	66
	51	39	38
	49	29	49
	58	46	56
	37	48	60
	29	49	56
	44	65	62
mean rating (SD):	**39.63** **(12.85)**	**40.63** **(14.23)**	**55.75** **(8.73)**

Step 1: Rank all of the scores, ignoring which group they belong to. The procedure for ranking is as follows: the lowest score gets the lowest rank. If two or more scores are the same then they are "tied". "Tied" scores get the average of the ranks that they *would* have obtained, had they not been tied. Here's the scores again, now with their ranks in brackets:

	C1 (No exercise)	C2 (Jogging for 20 minutes)	C3 (Jogging for 60 minutes)
	23 (2)	22 (1)	59 (20)
	26 (3)	27 (4)	66 (24)
	51 (16)	39 (9)	38 (8)
	49 (14)	29 (5.5)	49 (14)
	58 (19)	46 (11)	56 (17.5)
	37 (7)	48 (12)	60 (21)
	29 (5.5)	49 (14)	56 (17.5)
	44 (10)	65 (23)	62 (22)
mean rank	9.56	9.94	18.00
(SD)	(6.25)	(6.84)	(5.09)
sum of ranks (R_I)	76.5	79.5	144

In detail, this is how the ranks are arrived at for these scores.

(a) "22" is the lowest score. This gets a rank of 1.

(b) "23" is the next lowest score. This gets a rank of 2.

(c) "26" is the next lowest score. This gets a rank of 3.

(d) "27" is the next lowest score. This gets a rank of 4.

(e) There are two instances of "29". This is a "tie". They both get the average of the ranks that they would have been allocated, had they been different from each other. So the next two ranks are 5 and 6. The average of 5 and 6 is 11/2 = 5.5. Both instances of "29" therefore get a rank of 5.5.

(f) "37" is the next lowest score. This gets a rank of 7 (because we've just "used up" ranks 5 and 6).

(g) " 38" is the next lowest score, and it gets a rank of 8.

(h) "39" is the next lowest score, and it gets a rank of 9.

(i) "44" gets a rank of 10, "46" gets a rank of 11, and "48" gets a rank of 12.

(j) There are three instances of "49", so this is another tie. They each get the average of the next three unused ranks ((13+14+15) / 3 = 14).

(k) "51" is the next lowest score, and it gets the next "unused" rank, which is 16.

(l) There are two instances of "56", so they get the average of the next two unused ranks ((17+18) /2 = 17.5).

(m) "58" gets the next unused rank, which is 19.

(n) "59" gets a rank of 20, "60" gets 21, "62" gets 22, "65" gets 23, and 66 gets 24.

This is all tedious, but really not difficult to do once you've practiced it a couple of times!

Step 2: Find "R_I", the total of the ranks for each group. Just add together all of the ranks for each group in turn.

Here, R_I1 (the rank total for the "no exercise" group) is 76.5. R_I2 (the rank total for the "20 minutes" group) is 79.5.

R_I3 (the rank total for the "60 minutes" group) is 144.

Step 3: **Find "H".**

$$H = \frac{12}{n(n+1)} \sum \frac{R_i^2}{ni} - 3(n+1)$$

N is the total number of participants (all groups combined).*

We have 24 participants (3 groups of 8).

R_I is the rank total for each group. $R_I 1 = 76.5$, $R_I 2 = 79.5$, and $R_I 3 = 144$.

Ni is the number of participants in each group. Here, ni1 = 8, ni2 = 8 and ni3 = 8.

For our data,

$$H = \left[\frac{12}{n(n+1)} * \sum \frac{R\,i^2}{ni} \right] - 3 * (24 + 1)$$

$$\sum \frac{R_i^2}{ni} \quad means\ the\ following:$$

First, take each group's rank total, square it and then divide the result by the number of participants in that group. Then, add these numbers together.

$$\sum \frac{R_i^2}{ni} = \frac{76.5^2}{8} + \frac{79.5^2}{8} + \frac{144^2}{8}$$

$$\sum \frac{R_i^2}{ni} = 721.5313 + 790.0313 + 2592.0000 = 4113.5625$$

H=$\left[\frac{12}{600} * 4113.5625\right] - 75$

$H = [0.02 * 4113.5625] - 75$

$H = 82.27125 - 75$

$H = 7.27$

Step 4: the degrees of freedom is the number of groups minus one. Here we have three groups, and so we have 2 d.f.

Step 5:

Assessing the significance of H depends on the number of participants and the number of groups.

If you have three groups, with five or fewer participants in each group, then you need to use the special table for small sample sizes (which is on my website).

If you have more than five participants per group, then treat H as Chi-Square. H is statistically significant if it is equal to or larger than the critical value of Chi- Square for your particular d.f. (The table of Chi-Square values is also on my website).

Here, we have eight participants per group, and so we treat H as Chi-Square. H is 7.27, with 2 d.f. Here's the relevant part of the Chi-Square table:

Table of critical Chi-Square values:

df	p = .05	p = .01	p = .001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27

Look along the row that corresponds to your number of degrees of freedom. So in this case, we look along the row for 2 d.f.

We compare our obtained value of H to each of the critical values in that row of the table, starting on the lefthand side and stopping once our value of H is no longer equal to or larger than the critical value.

So here, we start by comparing our H of 7.27 to 5.99. With 2 degrees of freedom, a value of Chi-Square as large as 5.99 is likely to occur by chance only 5 times in a hundred: i.e. it has a p of .05. Our obtained value of 7.27 is even larger than this, and so this tells us that our value of H is even *less* likely to occur by chance. Our H will occur by chance with a probability of *less* than 0.05.

Move on, and compare our H to the next value in the row, 9.21. 9.21 will occur by chance one time in a hundred, i.e. with a p of .01. However, our H of 7.27 is less than 9.21, not bigger than it. This tells us that our value of H is not so large that it is likely to occur with a probability of 0.01.

Example 2

Three products received the following performance rating
by a panel of 20 customers. Use the Kurskal-Wallis test to determine whether
there is significant difference in the performance ratings for the product.

Here the hypothesis are

H_0: the performance ratings are same for the product.

H_1: the performance ratings are different for the product.

Product

A	B	C
25	60	50
70	20	70
60	30	60
85	15	80
95	40	90
90	35	70
80		75

ranking for pooled data

Rating	Product	Rank	Rating	Product	Rank
15	B	1	70	A	12
20	B	2	70	C	12
25	B	3	70	C	12
30	B	4	75	C	14
35	B	5	80	A	15.5
40	B	6	80	C	15.5
50	C	7	85	A	17
60	A	9	90	A	18.5
60	B	9	90	C	18.5
60	C	9	95	A	20

Product A	Rank	Product B	Rank	Product C	Rank
25	3	60	9	50	7
70	12	20	2	70	12
60	9	30	4	60	9
85	17	15	1	80	15.5
95	20	40	6	90	18.5
90	18.5	35	5	70	12
80	15.5			75	14
Total	R_1=95		R_2=27		R_3=88

Product B Rank Product C Rank

Here $n_1 = 7$, $n_2 = 6$, $n_3 = 7$, $\sum_{i=1}^{K} ni = 7+6+7 = 20$, k=3

R1 = 95, R2 = 27 and R3 = 88

$$\sum_{i=1}^{K} \frac{R_i^2}{ni} - \frac{95^2}{7} + \frac{27^2}{6} + \frac{88^2}{8} = 2.517.071$$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{K} \frac{R_i^2}{ni} - 3(n+1)$$

$$= \frac{12}{20(20-1)} \times 2507.071 - 3(20+1) = 8.916$$

The tabled value x_a^2 for a=0.05 and k=3-1 degrees of freedom is 5.99. Since the calculated value H> x_a^2 test is significant. So we reject H_0 and conclude that the performance ratings are different for the product.

-14-

Example 3

The marks of statistics for selected 5 students from three colleges are given in the table. Test whether the performance of the students of the three college are same or not.

• Here the hypothesis are

• H0: $M_1 = M_2 = M_3$

• H_1: some of the colleges differs in median marks.

Where M_i is the median mark of ith college.

College

A	B	C
50	80	60
62	95	45
75	98	30
48	87	58
65	90	57

ranking for pooled data:

Marks	College	Rank	Marks	College	Rank
30	C	1	65	A	9
45	C	2	75	A	10
48	A	3	80	B	11
50	A	4	87	B	12
57	C	5	90	B	13
58	C	6	95	B	14
60	C	7	98	B	15
62	A	8			

Marks College Rank Marks College Rank

College A	Rank	College B	Rank	College C	Rank
50	4	80	11	60	7
62	8	95	14	45	2
75	10	8/98	15	30	1
48	3	87	12	58	6
65	9	90	13	57	5
	$R_1=34$		$R_2=65$		$R_3=21$

$n_1=5$, $n_2=5$, $n_3=5$, $n_2=7$, n= $\sum_{i=1}^{k} ni = 5 + 5 + 5 = 15. k = 3$

$R_1=34$, $R_2=65$, and $R_3=21$

$$\sum_{i=1}^{k} \frac{R_1^2}{ni} = \frac{34^2}{5} + \frac{65^2}{5} + \frac{21^2}{5} = 1164.40$$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_1^2}{ni} - 3(n+1)$$

$$= \frac{12}{15(15+)} \times 1164.40 - 3(15+1) = 10.22$$

The tabled value x_a^2 for a=0.05 and k=3-1 degrees of freedom is 5.99 since the calculated value $H > x_a^2$ test is significant . So we reject H and conclude that the performance student are different for the three colleges.

Groups	Compare Colleges	R_i	U_i	U	Result
Group I	College A	15	25.00	0.00	Significant
	College B	40	0.00		
Group II	College A	34	6.00	6.00	Not Significant
	College C	21	19.00		
Group III	College B	40	0.00	0.00	Significant
	College C	15	25.00		
Critical $U_\alpha =2$					

Since the test is found significant, we have to determine which colleges are different from each other in median marks for this we conduct Mann Whitney U test.

The result presented in the following table. The result indicate that the college A and College B, College B and College C differs significantly. But college A and College C does not differ with each other.

1-3: The difference between Mann Whitney and Kruskal-Wallis Tests:

• The major difference between the Mann-Whitney U and the Kruskal-Wallis H is simply that the latter can accommodate more than two groups.

• That is Mann-Whitney compare the median or distribution between two groups or two populations.
• Kruskal-Wallis compare the median or distribution between three or more groups or populations.

2- When to use The Kruskal-Wallis test?

This test is appropriate for use under the following circumstances:

(a) you have three or more conditions that you want to compare;

(b) each condition is performed by a different group of participants; i.e. you have an independent-measures design with three or more conditions.

(c) the data do not meet the requirements for a parametric test. (i.e. use it if the data are not normally distributed; if the variances for the different conditions are markedly different; or if the data are measurements on an ordinal scale). If the data meet the requirements for a parametric test, it is better to use a one-way independent-measures Analysis of Variance (ANOVA) because it is more powerful than the Kruskal-Wallis test.

3 : Application on Kruskal-Wallis Test:

  In real world, data is coming from various domains. You may need to obtain meaningful inferences from data. In statistics, various statistical techniques are available for making inferences. Mainly, two approaches named as parametric and non parametric can be introduced. Parametric approaches are based on strong assumptions on distribution of data whereas non parametric approaches are based on weak assumptions about nature of data. Although parametric tests are more powerful than non parametric tests, mostly non parametric approaches are appropriate to apply for real world data because they are free from distributional assumptions.

When considering non parametric statistical method, it should satisfy at least one of the criteria pointed out below.

• The method is applicable on Nominal scale data (Examples for Nominal scale data are Gender (i.e.: Male, Female) and Hair color (i.e.: Brown, Black, White)).
• The method is applicable on Ordinal scale data (Examples for Ordinal scale data are time of day (i.e.: Morning, noon, afternoon, evening, night) and Agreement (i.e.: strongly agree, moderately agree, agree, neutral , disagree , moderately disagree, strongly disagree)).

•The method is applicable on data with an Interval or Ratio scale of measurement (Examples for Interval scale data are Celsius Temperature, Fahrenheit Temperature and examples for Ratio scale data are Age, Height and Weight).
Accordingly, it is essential to identify the nature of the data prior to apply statistical test.

4- Kruskal-Wallis Tests in SPSS

Three teaching methods were tested on a group of 19 students with homogeneous backgrounds in statistics and comparable aptitudes. Each student was randomly assigned to a method and at the end of a 6-week program was given a standardized exam. Because of classroom space and group size, the students were not equally allocated to each method. The results are shown in the table below. Test for a difference in distributions (medians) of the test scores for the different teaching methods using the Kruskal-Wallis test.

|							
Method 1	94	87	90	74	86	97	
Method 2	82	85	79	84	61	72	80
Method 3	89	68	72	76	69	65	

1. Enter the time values into one variable and the corresponding teaching method number (1 for Method 1, 2 for Method 2, 3 for Method 3) into another variable *(see figure, below)*. Be sure to code your variables appropriately.

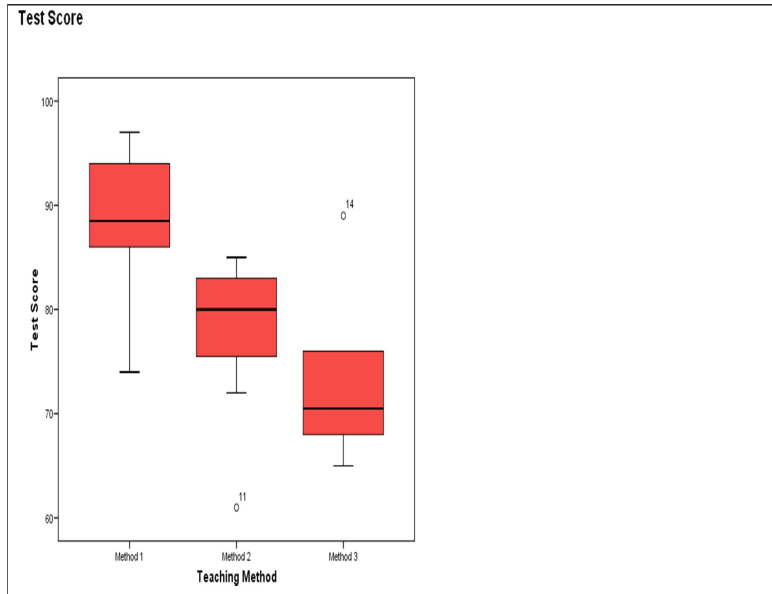|	method	score
1	Method 1	94
2	Method 1	87
3	Method 1	90
4	Method 1	74
5	Method 1	86
6	Method 1	97
7	Method 2	82
8	Method 2	85
9	Method 2	79
10	Method 2	84
11	Method 2	61
12	Method 2	72
13	Method 2	80
14	Method 3	89
15	Method 3	68
16	Method 3	72
17	Method 3	76
18	Method 3	69
19	Method 3	65

2. Now it is time to check the assumption of same shape for the treatment distributions. Select Graphs → Legacy Dialogs → Boxplot… (be sure the type is "Simple" and the data in the chart are "Summaries for groups of cases") with

 "Test Score" as the variable measured and "Teaching Method" as the category axis variable *(see figures, below)*. Click "OK". You may double-click the resulting boxplot to access the chart editor and adjust any display settings as desired (background color, bar width, bar color, etc.). As long as the boxes have approximately the same shape, you may continue with the Kruskal-Wallis

 procedure.
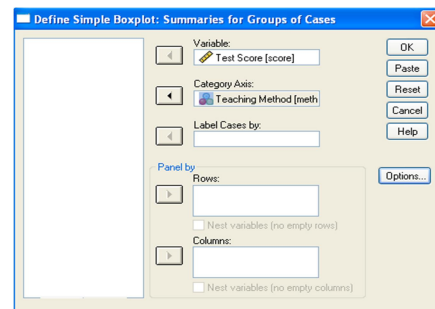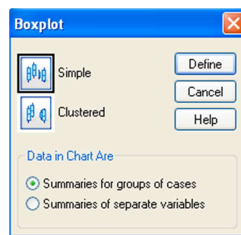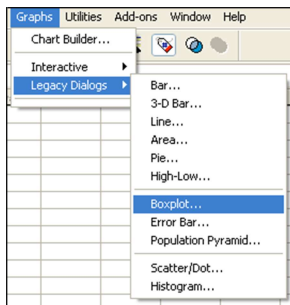
3. Select Analyze →Nonparametric Tests → K Independent Samples… (see upper-left figure, below). Select "Test Score" as the test variable, select "Teaching Method" as the grouping factor, and select "Kruskal-Wallis H" as the test type (see upper-right figure, below). Click "Define Range…" and enter the minimum value (1) as well as the maximum value (3) (see lower-left figure, below). Click "Continue" to close the range definitions and then click "OK" (see lower-right figure, below).
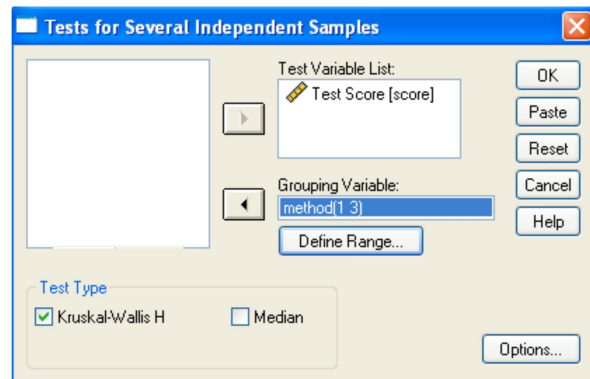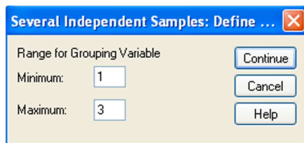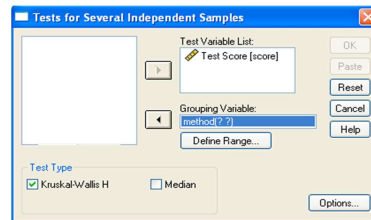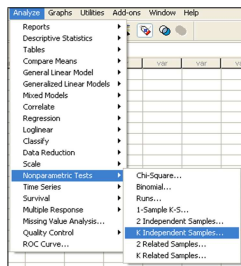
4. Your output should look like this.

Test Score

NPar Tests

Kruskal-Wallis Test

Ranks

	Teaching Method	N	Mean Rank
Test Score	Method 1	6	15.0000
	Method 2	7	8.7857
	Method 3	6	6.4167
	Total	19	

Test Statistics[a,b]

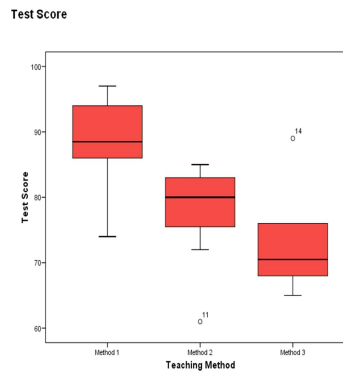|	Test Score
Chi-Square	7.5023
df	2
Asymp. Sig.	.0235

a. Kruskal Wallis Test

b. Grouping Variable: Teaching Method

5. You should use the output information in the following manner to answer the question.

Step 1: Check Assumptions

The samples were taken randomly and independently of each other.

The populations have approximately the same shapes since the boxplots are all about the same shape.

Test Score

All sample sizes are at least 6 if $k = 3$ (smallest in table is 6)

Ranks

	Teaching Method	N	Mean Rank
Test Score	Method 1	6	15.0000
	Method 2	7	8.7857
	Method 3	6	6.4167
	Total	19	

Step 2: Hypotheses

$H_0: M = M_2 = M_3$ (The median test scores are equal.)

H_a: Not all of the medians are equal.

Step 3: Significance Level

$a = 0.05$

Step 4: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05.

Step 5: Test Statistic

Test Statistics[a,b]

	Test Score
Chi-Square	7.5023
df	2
Asymp. Sig.	.0235

a. Kruskal Wallis Test

b. Grouping Variable: Teaching Method

Step 6:   Decision

Since p-value $= 0.0235 \leq 0.05 = $ a, we reject the null hypothesis.

Step 7:   State conclusion in words

At the $a = 0.05$ level of significance, there exists enough evidence to conclude that there is a difference in the median test scores (and, hence, the mean test scores) among the three teaching methods based.

Reference:

1- Berger, Paul D.; Maurer, Robert E.; Celli, Giovana B. (2018). Experimenta Design. Cham: Springer International Publishing. doi:10.1007/978-3-319 64583-4. ISBN 978-3-319-64582-7

2- Corder, G.W. & Foreman, D.I. (2010). Nonparametric Statistics for Non statisticians: A Step-by-Step Approach. Hoboken, NJ: Wiley

3- Elliott, A. C., & Hynan, L. S. (2011). A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal–Wallis analysis. Computer methods and programs in biomedicine, 102(1), 75-80.

4- Feir, B. J., & Toothaker, L. E. (1974). The ANOVA F-Test Versus The Kruskal-Wallis Test: A Robustness Study.

5- Gleason, J. (2013). Comparative power of the ANOVA, randomization ANOVA, and Kruskal-Wallis test.

6- Guo, S., Zhong, S., & Zhang, A. (2013). Privacy-preserving kruskal–walli test. Computer methods and programs in biomedicine, 112(1), 135-145.

7- Hinton, P. R., McMurray, I., & Brownlow, C. (2014). SPSS explained. Routledge.

8- Junaidi, J. (2010). Statistik uji kruskal-wallis. Jurnal Fakultas Ekonomi Universitas Jambi, 1-5.

9- MacFarland, T. W., Yates, J. M., MacFarland, T. W., & Yates, J. M. (2016). Kruskal–Wallis H-test for oneway analysis of variance (ANOVA) by ranks. Introduction to nonparametric statistics for the biological sciences using R, 177-211.

10- https://ssor.vcu.edu/media/statistics/pdf/SPSS.KruskalWallis.PC.pdf.

11-Kruskal Wallis H Test: https://esjeevanand.uccollege.edu.in/wp-content/uploads/sites/114/2020/08/NON-PARAMTERIC-TEST-9.pdf.