

The Kruskal-Wallis One-way Analysis of Variance by Ranks — Analysis of k-Between-Group Data with a Quantitative Response Variable

Application: To compare the distributions of scores on a quantitative response variable (which is either ordinal, not normally distributed or from a small sample) obtained from 2 or more groups. For k independent groups designs the Kruskal-Wallis is often used as a nonparametric substitute for the between groups ANOVA.

Research Hypothesis: The researcher hypothesized that Coop stores would have the most fish on display, chain stores would display the least, and private pet stores would display an intermediate amount.

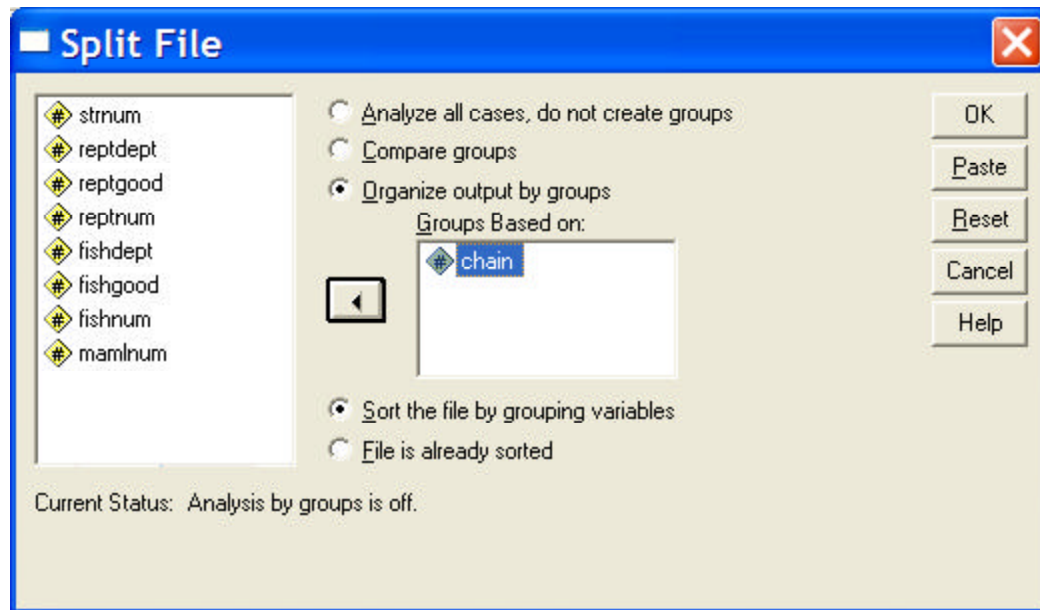
H₀: The three different types of pet shops have the same distribution of the number of fish displayed

Step 1 -- Getting the Univariate Statistics

SPSS has no convenient way of getting nonparametric univariate stats for separate groups. What we have to do is to “split” the file into two subfiles, based on the grouping variable. Then we can get the nonparametric univariate statistics for each group. Then we have to “un-Split” the file.

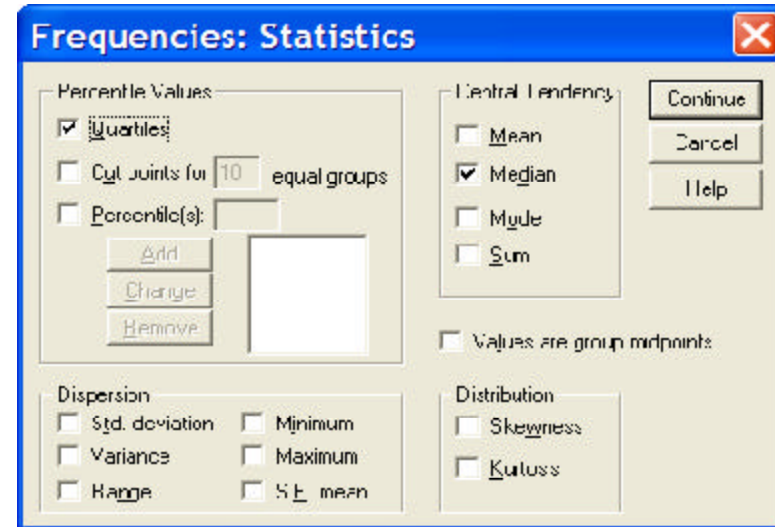
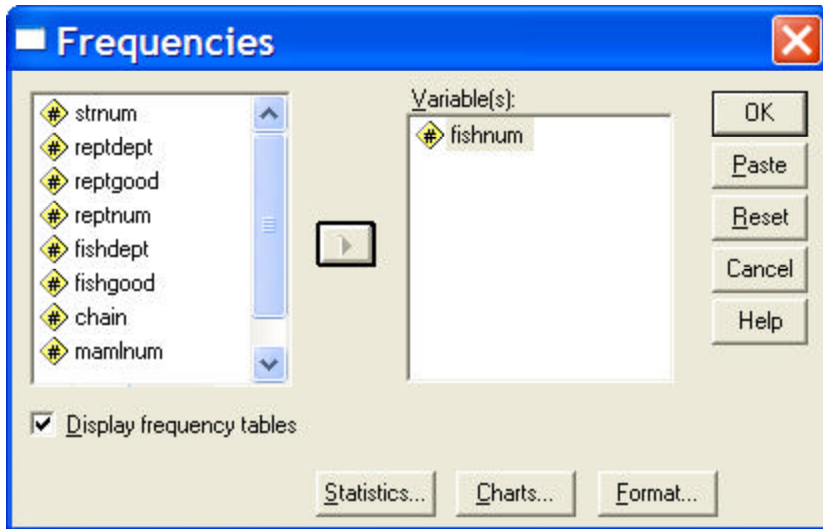
Data → Split File

- click “Organize output by groups”
- highlight the grouping variable (IV) and click the arrow to move it into the “Groups based on” window



Analyze/Statistics → Summarize → Frequencies

- highlight the response variable and click the arrow button
- Click “Statistics” — check “Quartiles” and “Median”



SPSS Output:

‘type of store’ = ‘chain store’

Statistics^a

‘number of fish at store’

N	Valid	5
	Missing	0
Median		20.00
Percentiles	25	12.00
	50	20.00
	75	21.50

a. ‘type of store’ = ‘chain store’

‘type of store’ = ‘privately owned’

Statistics^a

‘number of fish at store’

N	Valid	3
	Missing	0
Median		17.00
Percentiles	25	17.00
	50	17.00
	75	24.00

a. ‘type of store’ = privately owned

‘type of store’ = ‘coop’

Statistics^a

‘number of fish at store’

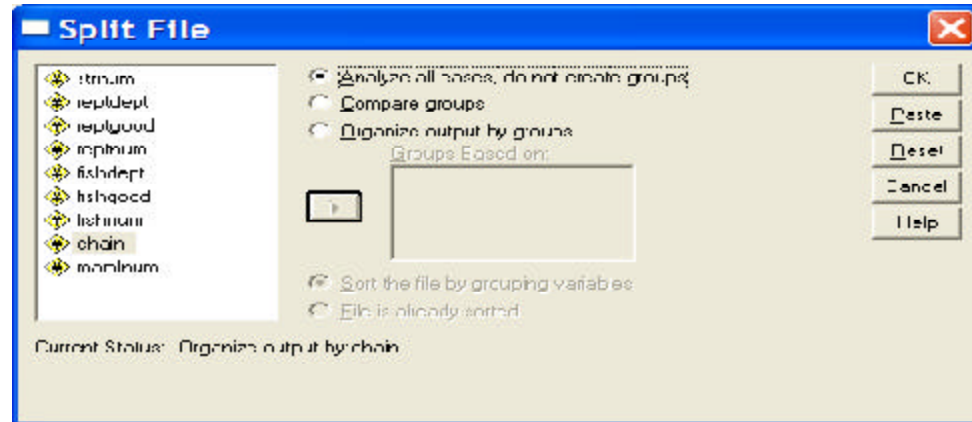
N	Valid	4
	Missing	0
Median		35.00
Percentiles	25	31.25
	50	35.00
	75	40.25

a. ‘type of store’ = coop

Step 2: Obtain the comparison of the groups.

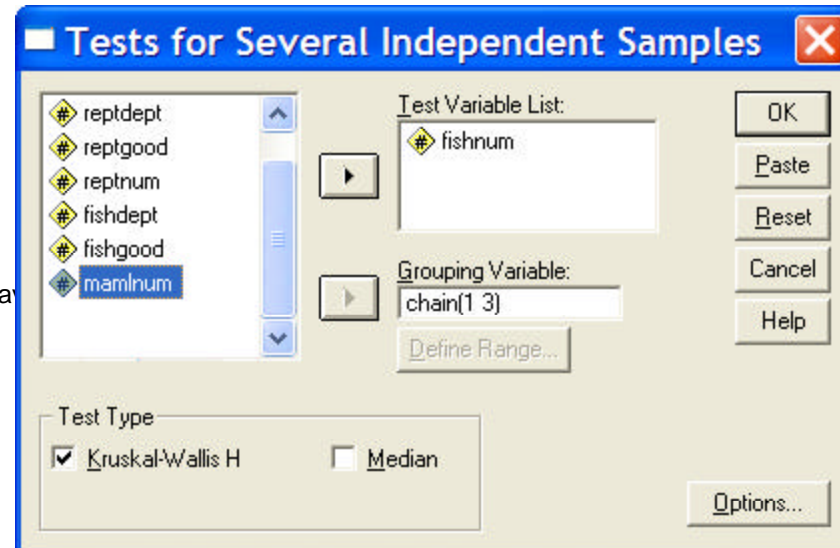
Data → Split File

- click “Analyze all cases. Do not create groups.”



Analyze/Statistics → Nonparametric Tests → K Independent Samples

- highlight the quantitative response variable and click the arrow to move it to the “Test Variable List” window
- highlight the grouping variable (IV) and click the arrow to move it to the “Grouping Variable” window
 - click the “Define Groups” button -- enter the lowest and highest value for the groups and click “continue”
- be sure the “Kruskal-Wallis H” is checked



Output:

Ranks

	'type of store'	N	Mean Rank
'number of fish at stor	'chain store'	5	4.20
	privately owned	3	5.00
	coop	4	10.50
	Total	12	

Test Statistics^{a,b}

	'number of fish at store'
Chi-Square	7.503
df	2
Asymp. Sig.	.023

- a. Kruskal Wallis Test
 b. Grouping Variable: 'type of store'

There is a significant effect. However, since there are three groups, we do not know the specific pattern of that effect until we complete pairwise comparisons among those three groups.

Pairwise follow-up analyses use the Kruskal Wallis to compare each pair of conditions.

Comparing groups 1 and 2 (chain vs. private)

Analyze/Statistics → Nonparametric Tests → K Independent Samples

- highlight "Grouping Variable" window
- click the "Define Groups" button -- change the "3" in the "maximum" window to "2"

Test Statistics^{a,b}

	'number of fish at store'
Chi-Square	.202
df	1
Asymp. Sig.	.653

a. Kruskal Wallis Test

b. Grouping Variable: 'type of store'

Comparing groups 2 and 3 (private vs. coop)

Analyze/Statistics → Nonparametric Tests → K Independent Samples

- highlight "Grouping Variable" window
- click the "Define Groups" button -- enter "2" as the minimum and "3" as the maximum

Test Statistics^{a,b}

	'number of fish at store'
Chi-Square	4.582
df	1
Asymp. Sig.	.032

a. Kruskal Wallis Test

b. Grouping Variable: 'type of store'

No difference between chain (1) and private (2) stores.

There is a difference between private (2) and coop (3) stores.

There is a difference between chain (1) and coop (was 3, now is 0) stores.

Comparing groups 1 and 3 (chain vs. coop)

In order to compare groups 1 & 3, we have to first recode group 3 (coop stores) to have a value of 0. Then ask for a comparison between groups 1 and 0 (which is really 3 - coop stores).

Data → Select Cases → K Independent Samples

- click "If condition is satisfied" button
- click the "If" button to open the "Select Cases: If" window
 - highlight the grouping variable and click the arrow to move it into the working window
 - type "≠ 2" in the window after the name of the grouping variable and press "continue" (this tells the computer to use only cases with a grouping variable score that is not 2)

Analyze/Statistics → Nonparametric Tests → K Independent Samples

- highlight "Grouping Variable" window
- click the "Define Groups" button -- enter "1" as the minimum and "3" as the maximum (since there is no cases in group "2" this will make a comparison of groups 1 and 3)

Test Statistics^{a,b}

	'number of fish at store'
Chi-Square	6.000
df	1
Asymp. Sig.	.014

a. Kruskal Wallis Test

b. Grouping Variable: 'type of store'

Reporting the Results:

The number of fish displayed at each type of store is summarized in Table 1. There were significantly different distributions of the number of fish displayed among the three types of stores (using Kruskal-Wallis, $X^2 = 7.03$, $p = .024$). Pairwise comparisons using the Kruskal-Wallis test ($p < .05$) revealed that, as hypothesized, Coop stores displayed the most fish. However, contrary to the hypothesis, there was no difference between the number of fish displayed by Chain and Private pet stores.

Table 1.

Quartile values for the number of fish displayed by each of the three types of stores

Number of Fish displayed	Store Type		
	Chain Store	Private Owned	Coop Store
Median	20.00	17.00	35.00
1st Quartile	12.00	17.00	31.25
3rd Quartile	21.50	22.50	40.25