

## LAB 2 INSTRUCTIONS

### ONE-WAY ANALYSIS OF VARIANCE

The One-Way ANOVA procedure in SPSS performs a one-way analysis of variance for a quantitative dependent variable by a single independent variable. The one-way analysis of variance F-test is used to test the hypothesis that several group means are equal. However, the F test, even if successful in rejecting the hypothesis is not able to tell which means are different from which other means. This analysis can be conducted by a priori contrasts (set up before running the experiment) and post-hoc tests run after the experiment. Analysis of variance is usually abbreviated as ANOVA.

We will demonstrate the one-way ANOVA tool using a simple example.

**Example:** An environmental scientist monitors the amount of contaminants present in the output of smokestacks from three steel mills. The amount of contaminant is measured at ten randomly selected times at each factory. The results are saved in the SPSS file *contam.sav* that can be downloaded by clicking on the link below. The measurements for each factory were made in the order provided in the data file. Here is the description of the variables in the data file:

<u>Column</u>	<u>Variable Name</u>	<u>Description of Variable</u>
1	AMOUNT	Amount of air contaminants,
2	FACTORY	Factory number (1, 2, or 3).

#### [DOWNLOAD DATA](#)

You will use analysis of variance to determine whether the factories produce significantly different mean levels of contaminants. You will also determine which means are different from which other means.

#### 1. Checking the Assumptions

The ANOVA inferences are valid under the assumption that the populations of interest are normal, observations are independent within and across groups and the populations have equal standard deviations. The assumptions of normality and equal standard deviations are not critical but the assumption of independence is.

Of the all assumptions, independence is the most crucial. If this assumption is violated, the inferences based on ANOVA are invalid. The independence assumption means that there is no association between the observations in the different groups and between the observations in the same group. In order to test independence, it is recommended to obtain a plot of the response variable versus the order in which the observations were obtained. If a pattern emerges, it is likely that the requirement is violated.

The normality assumption in ANOVA can be checked by obtaining histograms or normal probability plots for each of the groups. The analysis of variance is not heavily dependent on the normality assumption; ANOVA is robust to minor departures from normality. There is no reason for concern as long as the data are not extremely skewed. It is worthy to mention that the ANOVA F-test is especially resistant to departures from normality when the sample sizes are equal which holds for our data. In general, extremely skewed distributions with different sample sizes may make the outcome of the F test invalid. In general, with small sample sizes of 10 or fewer observations it's very difficult to check distributional assumptions like normality.

The assumption of equal standard deviations is crucial but not critical. If the number of observations in each group is the same (which is the case for our data), inferences about means assuming a common

variance are not seriously affected by unequal population variances. The assumption of equal standard deviations (variances) can be checked by examining the spread of the observations in the boxplots.

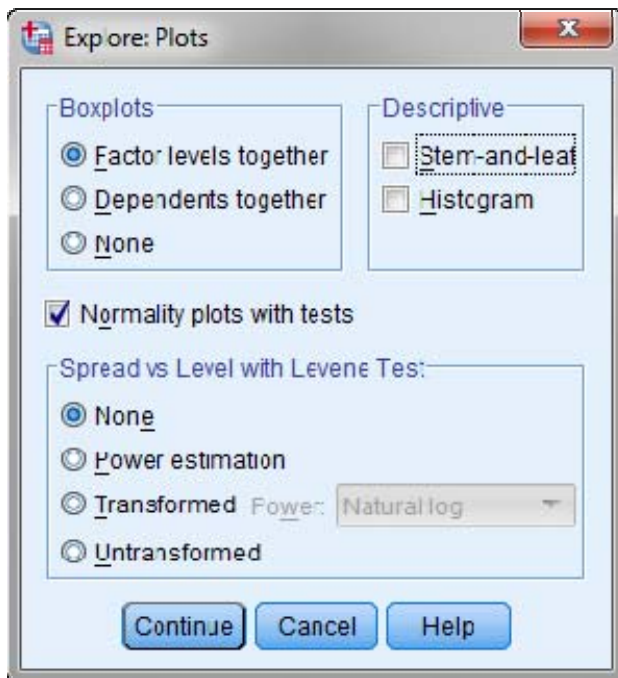
The assumption of equal standard deviations can also be verified with one of several available tests specifically designed for this purpose. However, they all are based on the assumption of normality and share lack of robustness against non-normality. The Levene's test already discussed in *Lab 1 Instructions* is least affected by non-normality. The null hypothesis for the Levene's test is that the standard deviations or equivalently the variances are homogeneous (identical). High p-value of the Levene's test (larger than the threshold value of 0.05) indicates that the data do support the assumption of equal standard deviations.

Because ANOVA procedures are not extremely sensitive to unequal standard deviations, it is not necessary to carry out a formal test of equality of standard deviations as a preliminary to the ANOVA. Instead, the spread of observations in the boxplots can be examined and the following rule of thumb may be used: If the ratio of the largest sample standard deviation to the smallest sample standard deviation is less than 2, the assumption of equal standard deviations is plausible.

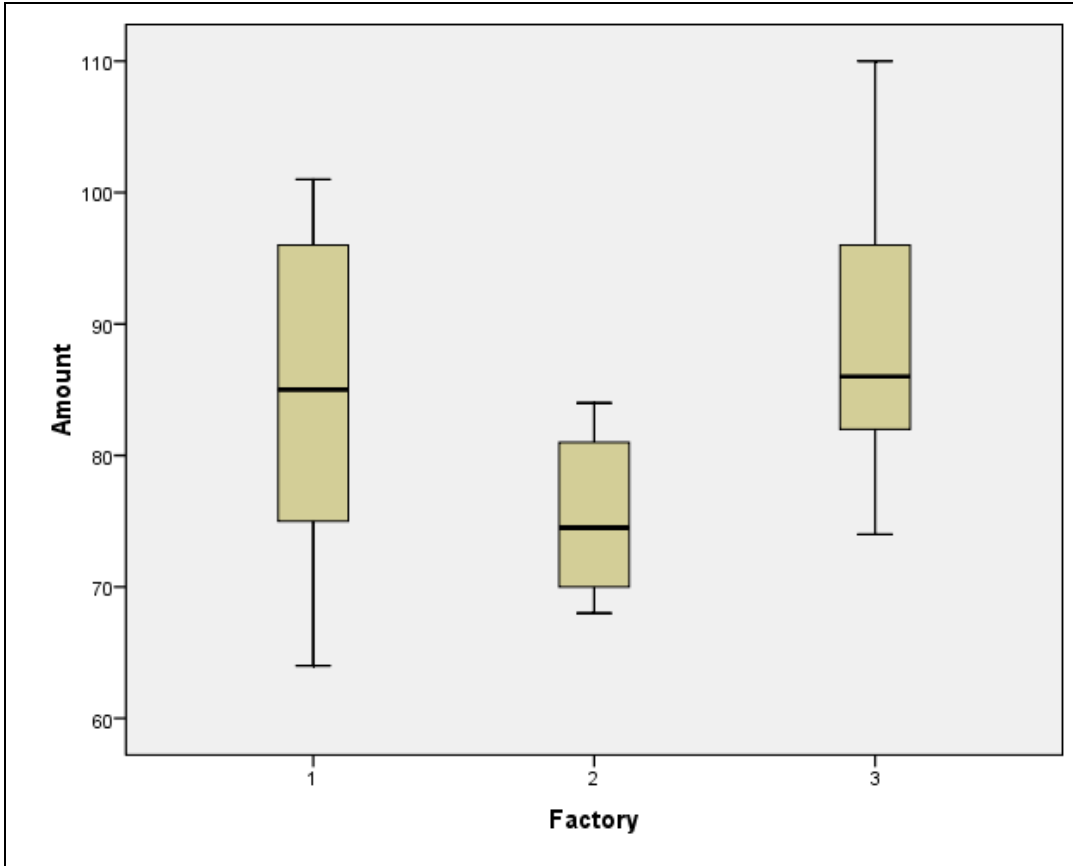
As a preliminary step in checking the assumptions of normality and equal standard deviations, we may obtain side-by-side boxplot. In case the boxplot shows that the distributions are extremely skewed with very unequal spread, further analysis with Q-Q plots may not be necessary.

The *Explore* procedure in SPSS allows obtaining the summary statistics (to check the ratio of the largest and smallest standard deviations), side-by-side boxplot for the data (to check the assumption of equal standard deviations and to check for any substantial departures from the normality assumption), Q-Q plots for the three groups and carry out the Levene's test.

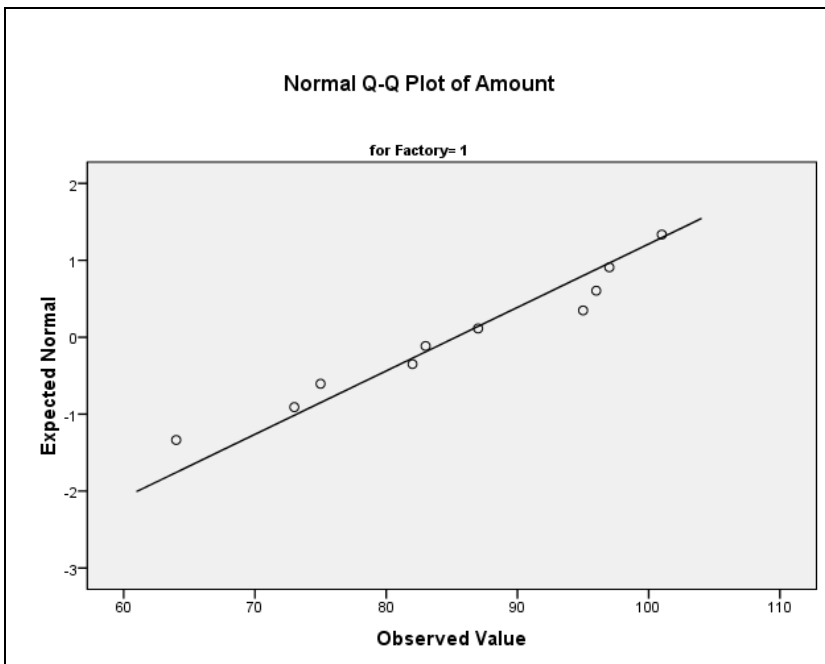
Click on *Analyze* in the main menu, *Descriptive Statistics* in the pull-down menu, and on *Explore*. Move the *Amount* variable in the variables list in the left panel to the *Dependent List* and the *Factory* variable to the *Factor List*. Click on *Plots* button and fill out its dialog box as follows:

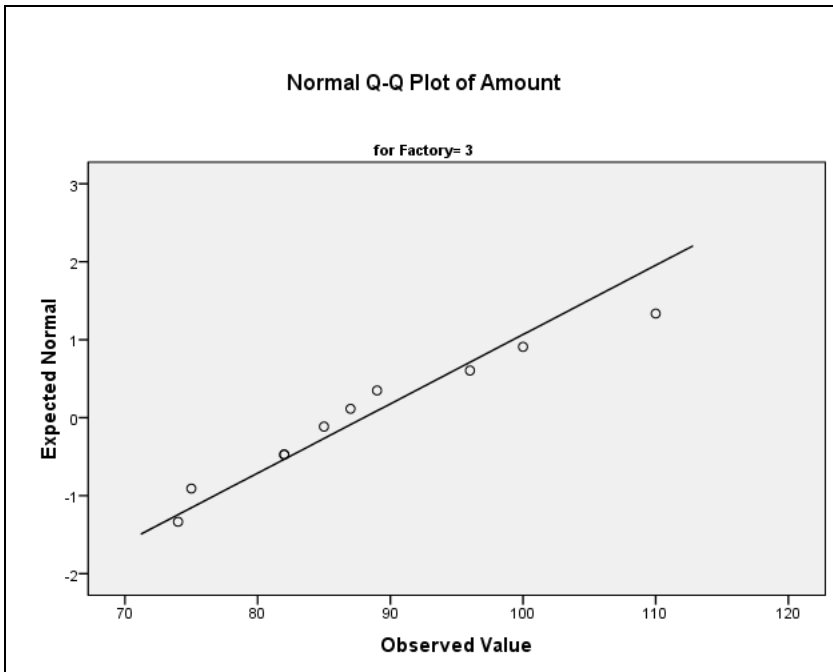
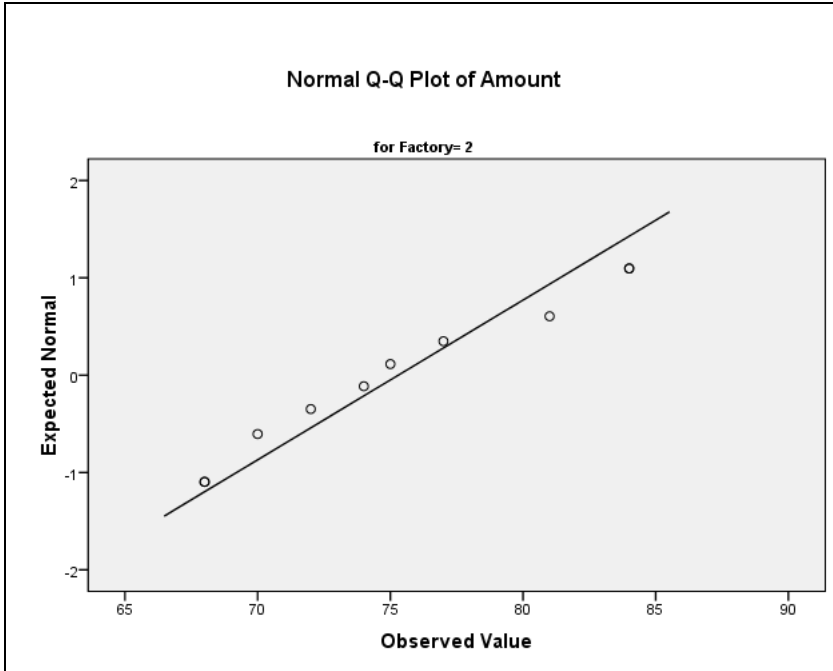


Click on *Continue* and *OK*. The *Explore* output is obtained. One of the items in the SPSS output is the side-by-side boxplot shown on the next page:



The boxplots shown above do not indicate that any of the three distributions might be extremely skewed. In order to get a better insight into the normality assumption for the three distributions, we will obtain Q-Q normality plots for the three factories. The normality plots for the three distributions are shown below:





The each of the above three Q-Q plots, most of the points lie close to a straight line, indicating that the normality assumption is not violated. Note that a single point of two points slightly separated from the linear pattern of the remaining observations do not provide evidence against normality. It is worthy to emphasize that small sample sizes do not allow us to make strong claims about normality for the data.

Summarizing, given lack of evidence for non-normality observed in the above three Q-Q plots, we do expect that the conclusions about equal standard deviations obtained with the Leven's test should be valid. The output for the Leven's test obtained with the *Explore* procedure is displayed below:

Test of Homogeneity of Variance					
		Levene Statistic	df1	df2	Sig.
Amount	Based on Mean	2.132	2	27	.138
	Based on Median	1.937	2	27	.164
	Based on Median and with adjusted df	1.937	2	21.489	.168
	Based on trimmed mean	2.075	2	27	.145

For this data the Levene's test is not significant (the p-value in each of the four versions of the test is much larger than the threshold value of 0.05), indicating the null hypothesis cannot be rejected, that is, the data support the assumption of equal variances.

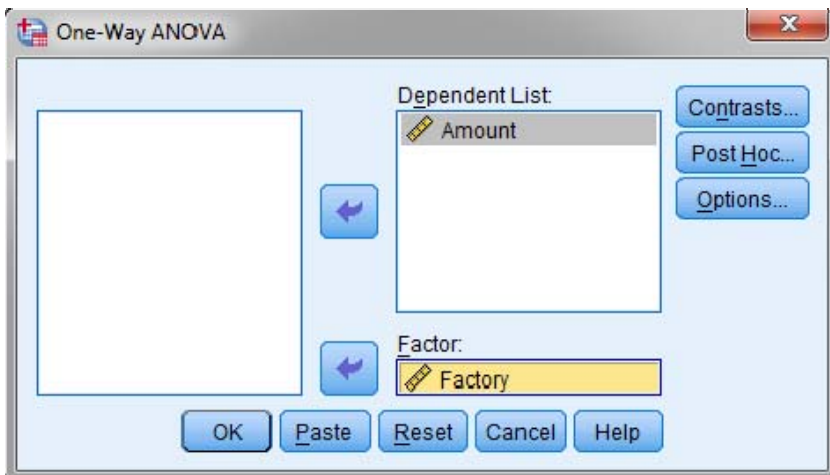
Notice also that based on the summaries provided in the output generated by *Explore*, the ratio of the largest and the smallest standard deviations is  $12.12/6.093 < 2$ , which also supports the assumption of equal standard deviations (variances).

## 2. One-Way ANOVA

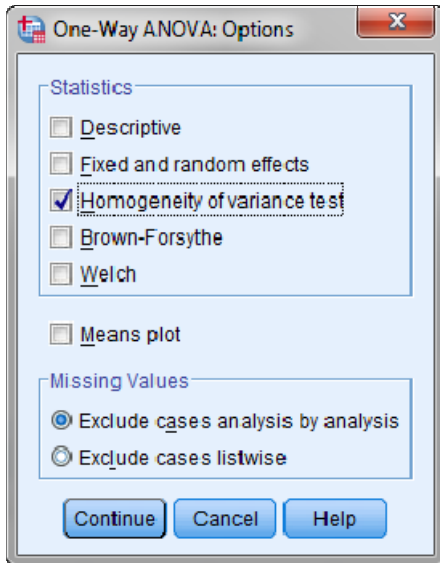
Suppose we wish to determine whether the three factories produce significantly different mean levels of contaminants. As there is no evidence that the assumptions of normality or equal standard deviations are violated for our data, we may apply One-way ANOVA procedure in SPSS to check for the differences.

Testing the null hypothesis in ANOVA is formulated as a problem of comparing two models, the full model and the reduced model. The full model has a separate mean for each group,  $\mu_1$ ,  $\mu_2$ , and  $\mu_3$ . The reduced model is obtained from the full model by assuming that the null hypothesis of equal means is true, i.e.  $\mu_1 = \mu_2 = \mu_3 = \mu$ . In order to determine which of the two models is more plausible given the data, we compare the sum of squares of residuals for the full and reduced models. Large differences in the size of residuals from the two models supply evidence against the hypothesis of equal means and in favour of a model with different means. The value of the test statistic F is obtained by dividing the change in the residual sum of squares for the full and reduced models (divided by the corresponding change in degrees of freedom) by the estimate of the variance based on the full model. The larger the value of F statistic, the stronger the evidence against the null hypothesis of equal population means

To access one-way ANOVA in SPSS, click on *Analyze* in the main menu bar, and then on *Compare Means* option from the pull-down menu. Clicking on *One-Way ANOVA* option opens the *One-Way ANOVA* dialog box.



Select the variable whose mean you want to test (*Amount*) and move it into the *Dependent List* box. You can move more than one variable into the *Dependent List* to test them all across the same set of groups. Now select the variable whose values define the groups (*Factory*) and move it into the *Factor* box. At this point please do not click on either *Contrasts* or *Post Hoc* tabs as the options will be discussed later. Click the *Options* button. The following *One-Way ANOVA: Options* dialog box opens:



Check the *Homogeneity of variance test* button. Click *Continue* and *OK* to run the procedure. The output is shown below:

**Test of Homogeneity of Variances**

Amount

Levene Statistic	df1	df2	Sig.
2.132	2	27	.138

**ANOVA**

Amount

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	895.267	2	447.633	4.322	.024
Within Groups	2796.200	27	103.563		
Total	3691.467	29			

Sum of squared residuals from fitting the full model is 2796.2

Sum of squared residuals from fitting the reduced model is 3691.467

Pooled estimate of the common variance is 103.563

According to the above output, the value of the F-statistic is 4.322 and the p-value of the test is 0.024. There is convincing evidence that there are significant differences among the three group means (mean amount of contaminants among the three factories).

### 3. Contrasts in SPSS

Suppose that before conducting the experiment, the researchers decided to compare the factory 1 and 2 with the factory 3. This comparison may be justified by differences in the contaminants produced by factories 1 and 2 and those produced by factory 3. The comparison can be done using the *Contrasts* feature in SPSS.

The contrast of interest is

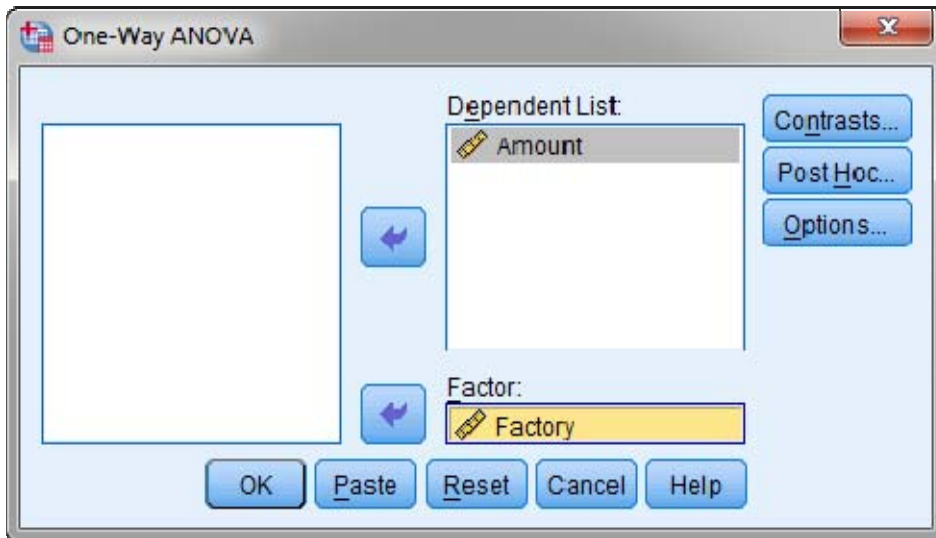
$$(\mu_1 + \mu_2)/2 - \mu_3 = 0$$

In general, in order to avoid converting fractions to decimals, it is recommended to express contrasts containing fractions into equivalent contrasts with integer coefficients. For example, the above contrast can be expressed in the equivalent form as

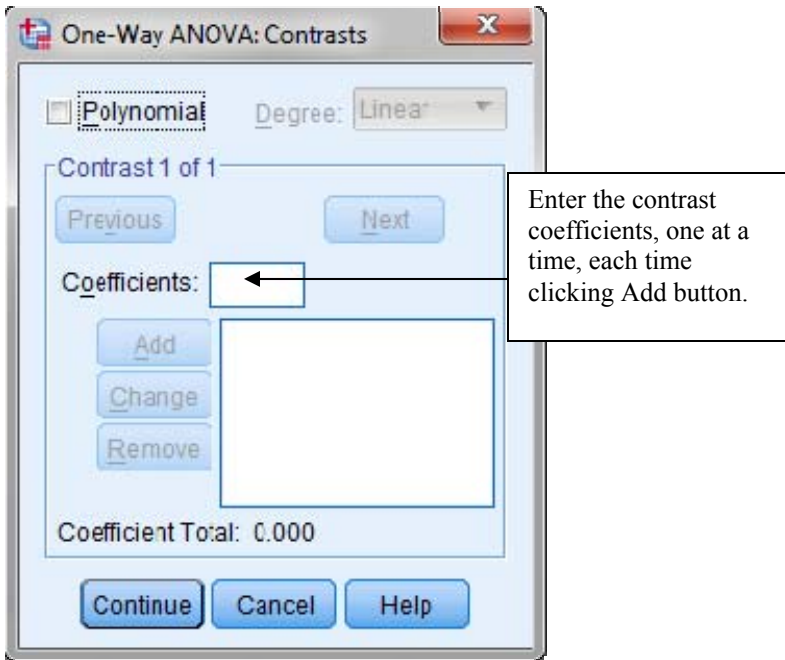
$$\mu_1 + \mu_2 - 2\mu_3 = 0$$

The operation also multiplies the value of the contrast and its standard deviation by 2, but does not affect the value of the test statistic and the p-value of the test. Replacing contrasts with fraction coefficients by equivalent contrasts with integer coefficients is especially justified when the fractions at the coefficients of the contrast have infinite decimal expansion i.e. 1/3, 1/6,...etc.

In the above contrast, the values 1, 1, and -2 are the coefficients for  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  respectively. In order to carry out contrasts in SPSS, it is necessary to click on the *Contrasts* button in the *One-Way ANOVA* dialog box.



This opens the *One-Way ANOVA: Contrasts* dialog box shown on the next page:



Check the *Polynomial* check box and enter the contrast coefficients, one at the time. Click Add button after each entry. In our case, enter the following coefficients 1, 1, and -2 for the contrast. Each new value is added to the bottom of the coefficient list. To specify other contrasts, click *Next* button. Then click *Continue* and *OK* to run the procedure. The contrast output is displayed as shown below:

Contrast Coefficients			
Contrast	Factory		
	1	2	3
1	1	1	-2

Contrast Tests							
		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Amount	Assume equal variances	1	-15.40	7.883	-1.954	27	.061
	Does not assume equal variances	1	-15.40	8.311	-1.853	15.352	.083

The p-value of the contrast (two-sided alternative) is 0.061 (equal variances assumed) or 0.083 (equal variances not assumed). Thus there is no sufficient evidence to claim that factories 1 and 2 differ from factory 3 in their levels of contaminants.

#### 4. Range Tests and Multiple-Comparison Procedures in SPSS

The ANOVA uses the F test to determine whether there exists a significant difference among population means. In this sense it is a preliminary test that informs us if we should continue the investigation of the data.

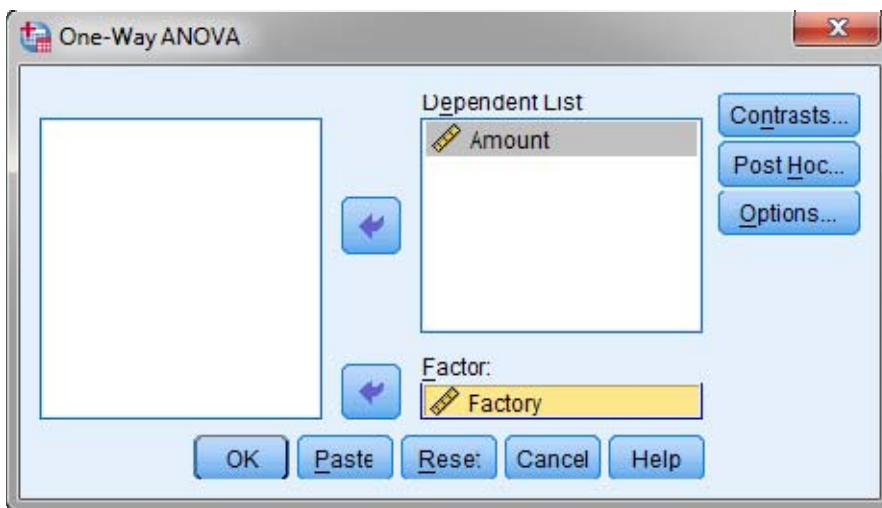


If the ANOVA null hypothesis is not rejected, there is an implication that no relation exists between the factor levels and the response. Thus, there is not much we can learn from the data, and we are finished with the analysis.

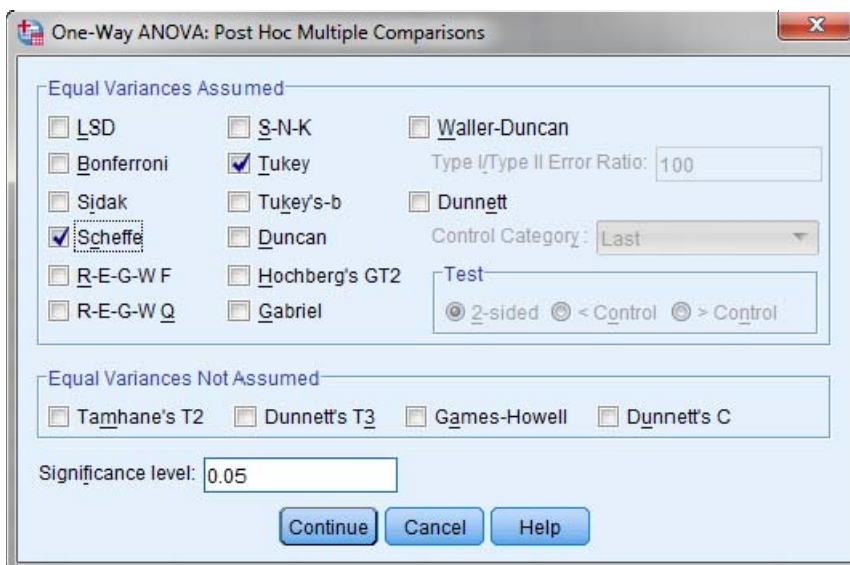
On the other hand, the significant ANOVA result indicates that there are differences in the means, but it does not tell you which means are different from which others, e.g. group 1's mean might be different than group 2's mean but not different from group 3's mean. Post hoc range tests and pairwise multiple comparisons can determine which means differ.

Range tests identify homogeneous subsets of means that are not different from each other. Pairwise multiple comparisons test the difference between each pair of means and produce a matrix where asterisks at each pair indicate significantly different group means at an  $\alpha$  level of 0.05.

The post hoc range tests and multiple comparisons in SPSS are incorporated into ANOVA test. In order to carry out post hoc analysis, click on *Post Hoc...* button in *One-Way ANOVA* dialog box.



The following *One-Way ANOVA: Post Hoc Multiple Comparisons* dialog box is obtained:



Tukey's honestly significant difference (Tukey HSD or briefly Tukey), and Scheffé methods are one of several multiple comparison tests and range tests to choose in the dialog box. Tukey HSD is used for testing the significance of unplanned pairwise comparisons. The test assumes equal sample sizes in all groups.

The Scheffé's method applies to the set of estimates of all possible contrasts among the factor level means, not just the pairwise differences considered by the Tukey method. Scheffé's method is typically quite conservative, and the experimental error rate will generally be much smaller than  $\alpha$ . In particular, the Scheffé's test is more conservative than Tukey's: whenever Scheffé's procedure detects a significant difference, the Tukey's procedure confirms the difference.

Select the *Scheffe* and *Tukey* check boxes. Then click *Continue* and *OK* to run the procedure. The post hoc tests results are given below.

## Post Hoc Tests

Multiple Comparisons							
Dependent Variable: Amount							
	(I) Factory	(J) Factory	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	10.000	4.551	.090	-1.28	21.28
		3	-2.700	4.551	.825	-13.98	8.58
	2	1	-10.000	4.551	.090	-21.28	1.28
		3	-12.700*	4.551	.025	-23.98	-1.42
	3	1	2.700	4.551	.825	-8.58	13.98
		2	12.700*	4.551	.025	1.42	23.98
Scheffe	1	2	10.000	4.551	.109	-1.79	21.79
		3	-2.700	4.551	.840	-14.49	9.09
	2	1	-10.000	4.551	.109	-21.79	1.79
		3	-12.700*	4.551	.033	-24.49	-.91
	3	1	2.700	4.551	.840	-9.09	14.49
		2	12.700*	4.551	.033	.91	24.49

\*. The mean difference is significant at the 0.05 level.

## Homogeneous Subsets

Amount				
Factory	N	Subset for alpha = 0.05		Sig.
		1	2	
Tukey HSD <sup>a</sup>	2	10	75.30	
	1	10	85.30	85.30
	3	10		88.00
				.090
Scheffe <sup>a</sup>	2	10	75.30	
	1	10	85.30	85.30
	3	10		88.00
				.109

Means for groups in homogeneous subsets are displayed.  
a. Uses Harmonic Mean Sample Size = 10.000.

The first table in the above output provides the pairwise confidence intervals for mean difference between two groups. Moreover, it provides the p-value of the corresponding test. If the two-sided p-value (denoted by “Sig.”) is less than or equal to 0.05, or equivalently the corresponding 95% confidence interval does not contain zero, there is a strong evidence to reject the null hypothesis of equal group means. In our case, both Tukey HSD and Scheffe’s methods provide strong evidence of a significant difference in the mean level of contaminants between factories 2 and 3.

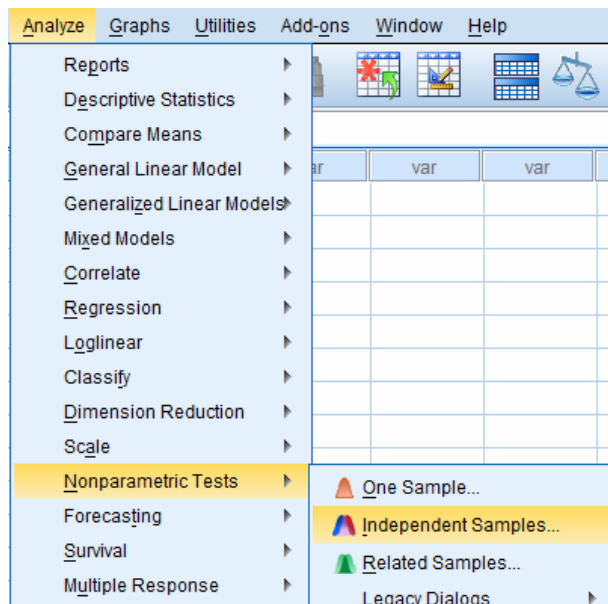
Based on the table above, we conclude that factory 1 and 2 are indistinguishable from each other (in terms of the mean level of contaminants as determined by the Tukey and Scheffe’s methods). There is no evidence that their average scores are any different. In addition, factories 1 and 3 are indistinguishable. However, the mean level for factory 3 is significantly higher than that of factory 2.

## 5. The Kruskal-Wallis Test

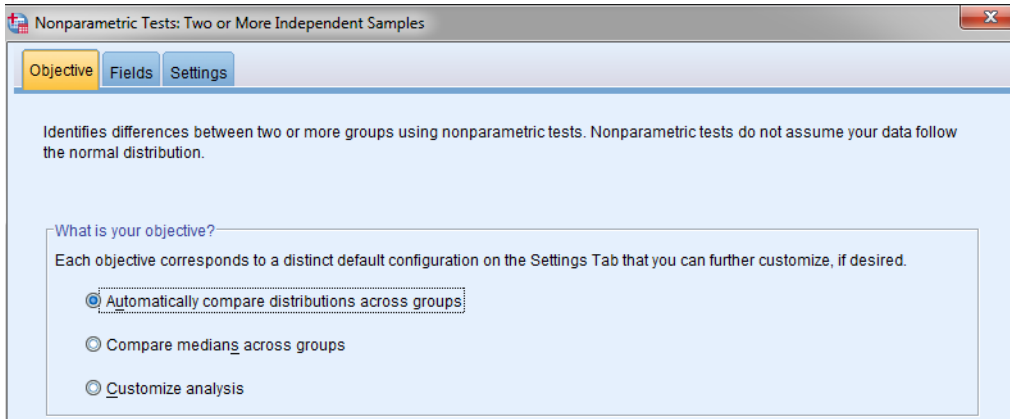
The F-test in ANOVA has the underlying assumptions of normality and equal variances. In case one of the assumptions or both of them are violated, natural logarithm or other data transformations may be tried to correct this problem. However, if none of the available transformations turns out to be successful (for example when data contain seriously outlying observations), the Kruskal-Wallis test can be applied to the data.

The Kruskal-Wallis test is a non-parametric test for finding if three or more independent samples come from populations having the same distribution. It is identical to a one-way analysis of variance with the data replaced by their ranks. Unlike the parametric independent group ANOVA (one way ANOVA), this non-parametric test makes no assumptions about the distribution of the data. Since this test does not make distributional assumptions, it is not as powerful as the ANOVA.

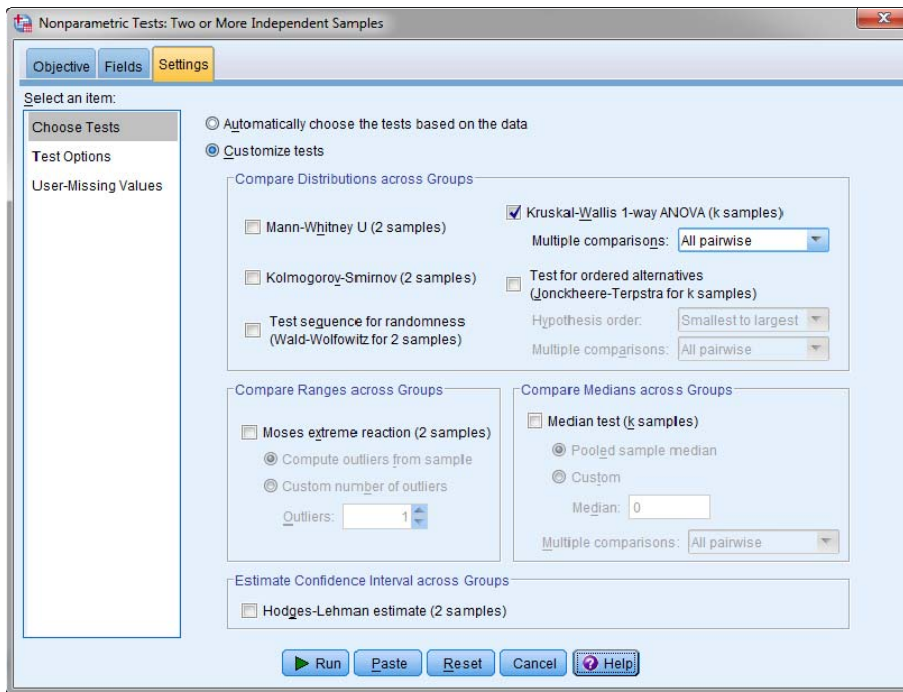
The Kruskal-Wallis test can be obtained by clicking Analyze, Nonparametric Tests and *Independent Samples....*



The *Nonparametric Tests: Two or More Independent Samples* dialog box is displayed. Make sure that *Automatically compare distributions across groups* radio button is checked.



Click *Fields* tab and select the variable *Amount* and move it into the *Test Fields* box. Then select the variable whose values define the groups (*Factory*) and move it into the *Groups* box. Finally click *Settings* tab, and check the *Customize tests* radio button. Make sure that the *Kruskal-Wallis 1-Way ANOVA* box is checked.



Click *Run*. The Kruskal-Wallis test output is displayed below.

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Amount is the same across categories of Factory.	Independent-Samples Kruskal-Wallis Test	.023	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

The p-value of the test is reported as 0.023 indicating strong evidence against the assumption that the three samples come from populations having the same distribution. This is consistent with the results obtained with the F-test.