

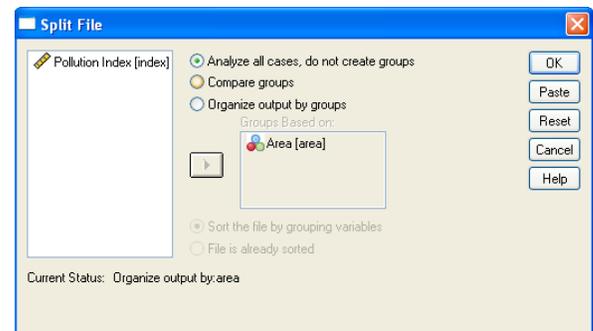
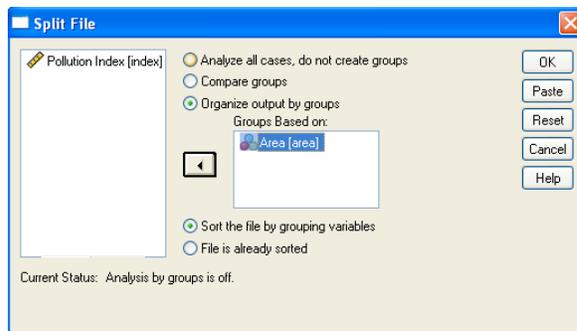
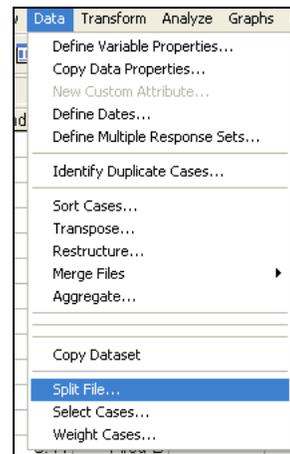
Two-Sample T-Tests in SPSS

The table below shows the observed pollution indexes of air samples in two areas of a city. Test the hypothesis that the mean pollution indexes are the same for the two areas. (Use $\alpha = 0.05$.)

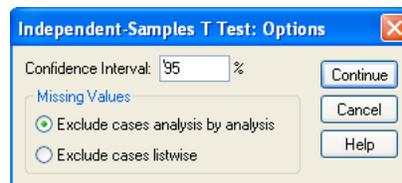
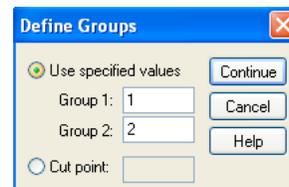
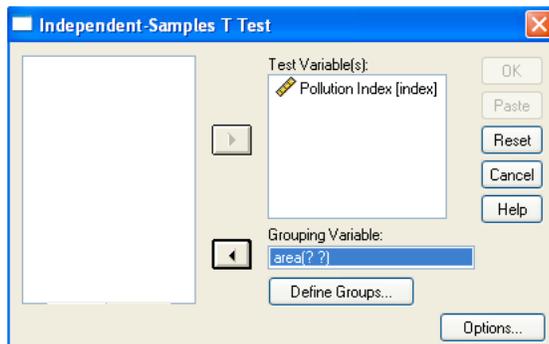
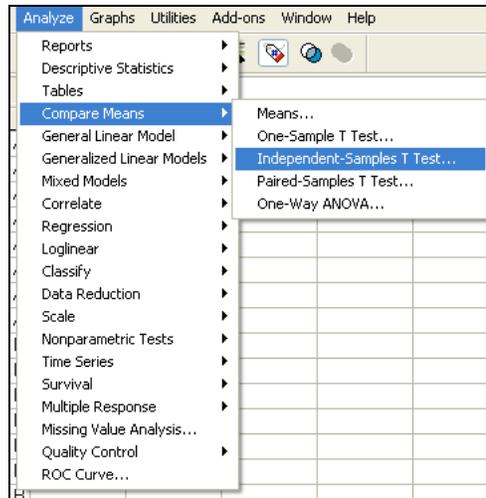
	Area A		Area B	
	2.92	4.69	1.84	3.44
	1.88	4.86	0.95	3.69
	5.35	5.81	4.26	4.95
	3.81	5.55	3.18	4.47

1. Enter the values into one variable and the corresponding sample number (1 for Area A, 2 for Area B) into another variable (*see upper-left figure, below*). Be sure to code your variables appropriately. Now it is time to check the normality assumption. Select “Split File” from the “Data” menu so that we can tell SPSS that we want separate Q–Q Plots for each group (*see upper-right figure, below*). Select “Organize output by groups” and enter “area” as the variable that groups are based upon (*see lower-left figure, below*). Now create Normal Q–Q Plots to assess the normality of each group (*see separate handout on Normal Q–Q Plots*). Once you’ve created your Q–Q Plots and determined that your groups are approximately normally distributed, select “Split File” from the “Data” menu and then select “Analyze all cases, do not create groups” in order to return SPSS to its normal data analysis mode (*see lower-right figure, below*).

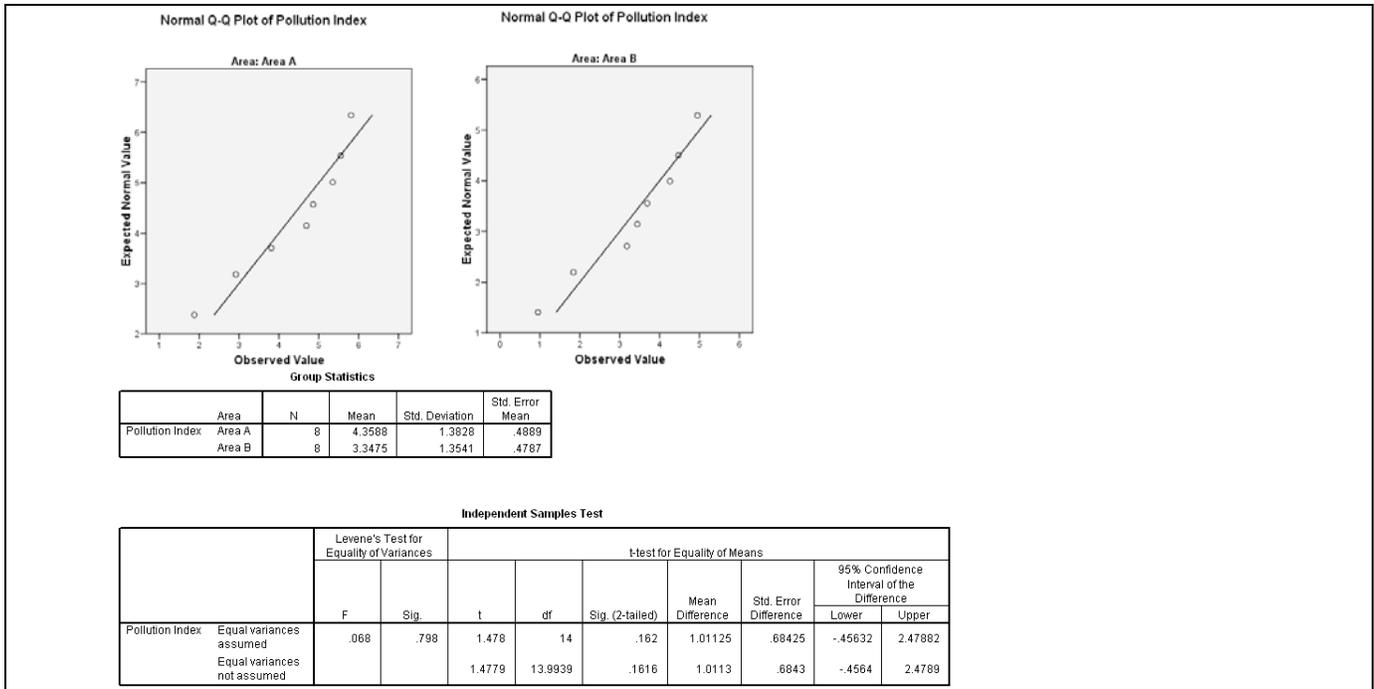
	index	area
1	2.92	Area A
2	1.88	Area A
3	5.35	Area A
4	3.81	Area A
5	4.69	Area A
6	4.86	Area A
7	5.81	Area A
8	5.55	Area A
9	1.84	Area B
10	.95	Area B
11	4.26	Area B
12	3.18	Area B
13	3.44	Area B
14	3.69	Area B
15	4.95	Area B
16	4.47	Area B



2. Select Analyze → Compare Means → Independent-Samples T Test... (see upper figure, below).
3. Select “Pollution Index” as the test variable and click “Define Groups...”. For Group 1, enter 1 (Area A), and for Group 2, enter 2 (Area B). Click the “Options...” button and enter the appropriate confidence level (95%), if needed. Click “Continue” to close the options and then click “OK” (see the 3 lower figures, below).



4. Your output should look like this.



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

Step 0: Check Assumptions

Since the points of each Q-Q Plot lie close to their respective diagonal lines, we conclude that each of the data groups is from an approximately normally distributed population.

Step 1: Hypotheses

$$H_0: \mu_A - \mu_B = 0$$

$$H_a: \mu_A - \mu_B \neq 0$$

Step 2: Significance Level

$$\alpha = 0.05$$

Step 3: Rejection Region

Since we don't know the population variances (σ_A^2 and σ_B^2) but think that they are not equal (air varies across different areas of the same city due to industrialization, vegetation, etc.), we'll use the non-pooled t -test ("Equal variances not assumed").

Reject the null hypothesis if p -value ≤ 0.05 .

Step 4: Test Statistic

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pollution Index	Equal variances assumed	.068	.798	1.478	14	.162	1.01125	.68425	-4.5632	2.47882
	Equal variances not assumed			1.4779	13.9939	.1616	1.0113	.6843	-4.564	2.4789

From the output, $T = 1.4779$ with 13.9939 degrees of freedom.

$$p\text{-value} = \text{Sig. (2-tailed)} = 0.1616$$

Step 5: Conclusion

Since p -value = 0.1616 $>$ 0.05 = α , we fail to reject the null hypothesis.

Step 6: State conclusion in words

At the $\alpha = 0.05$ level of significance, there is not enough evidence to conclude that the mean pollution indexes are the same for the two areas. [Since we failed to reject the null hypothesis, no confidence interval is needed.]