INDEPENDENT SAMPLE T-TESTS

Richard Lee Rogers

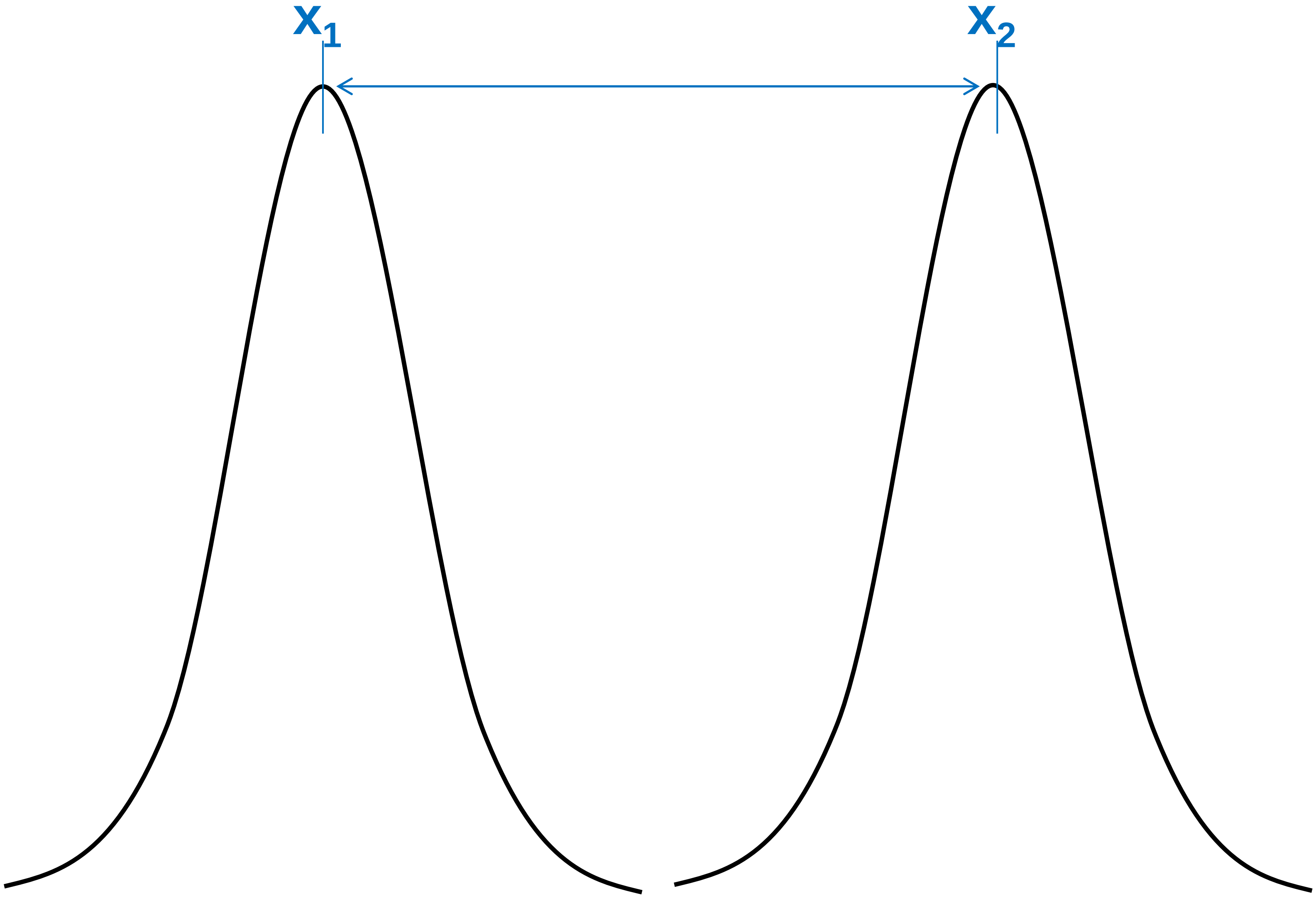
What is a t-test

- T-tests are used when one variable is numeric and the other variable has two categories. The test determines whether the distance between the two means is statistically significant.

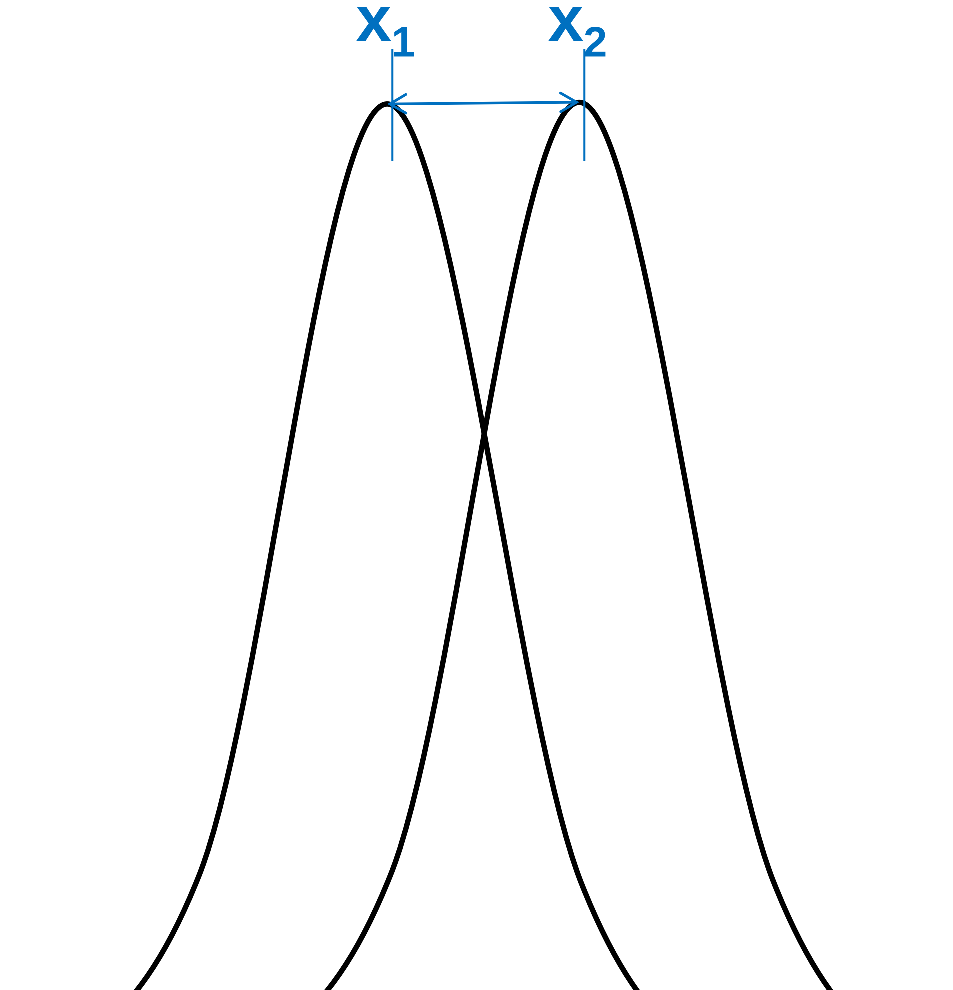
- t is the same Student’s t used in determining a confidence interval

- Elements of inference
  - Magnitude is expressed as an absolute number and is not standardized.
  - Direction not present in a traditional sense, but we can speak of one group being “larger” or “higher” than another group.
  - Statistical significance of the distance is present.

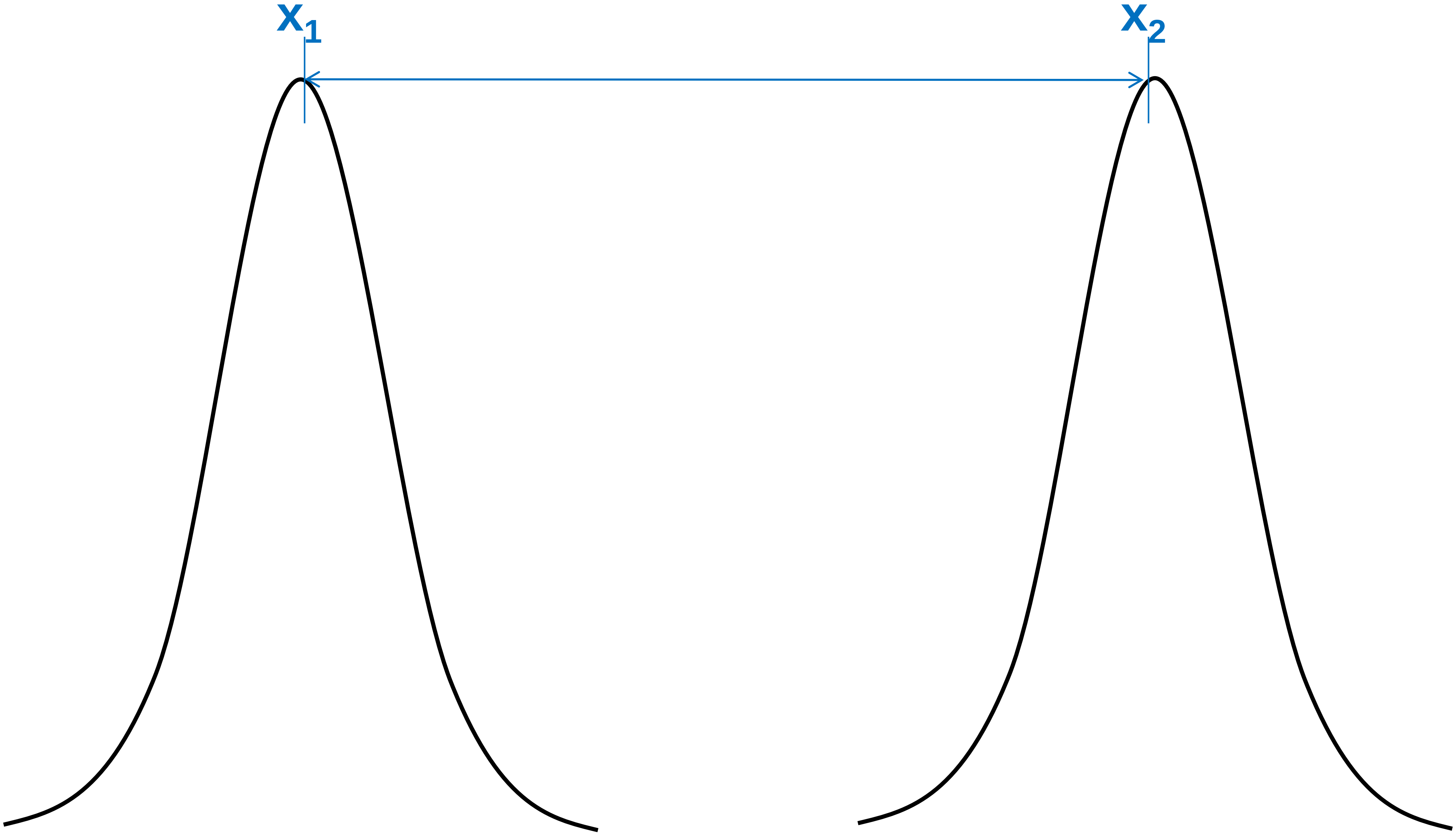
The T-test



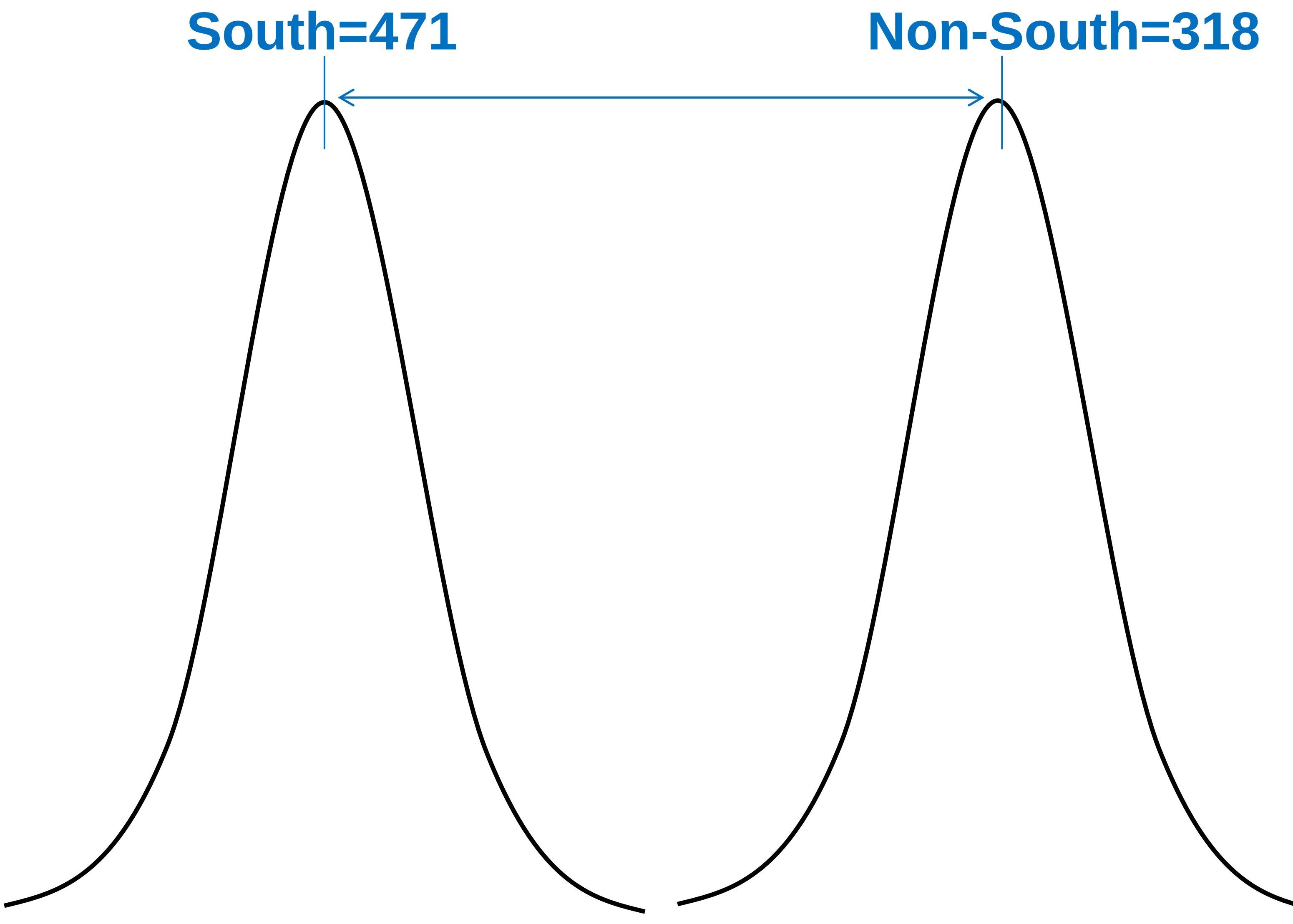
The T-test: A Small Difference



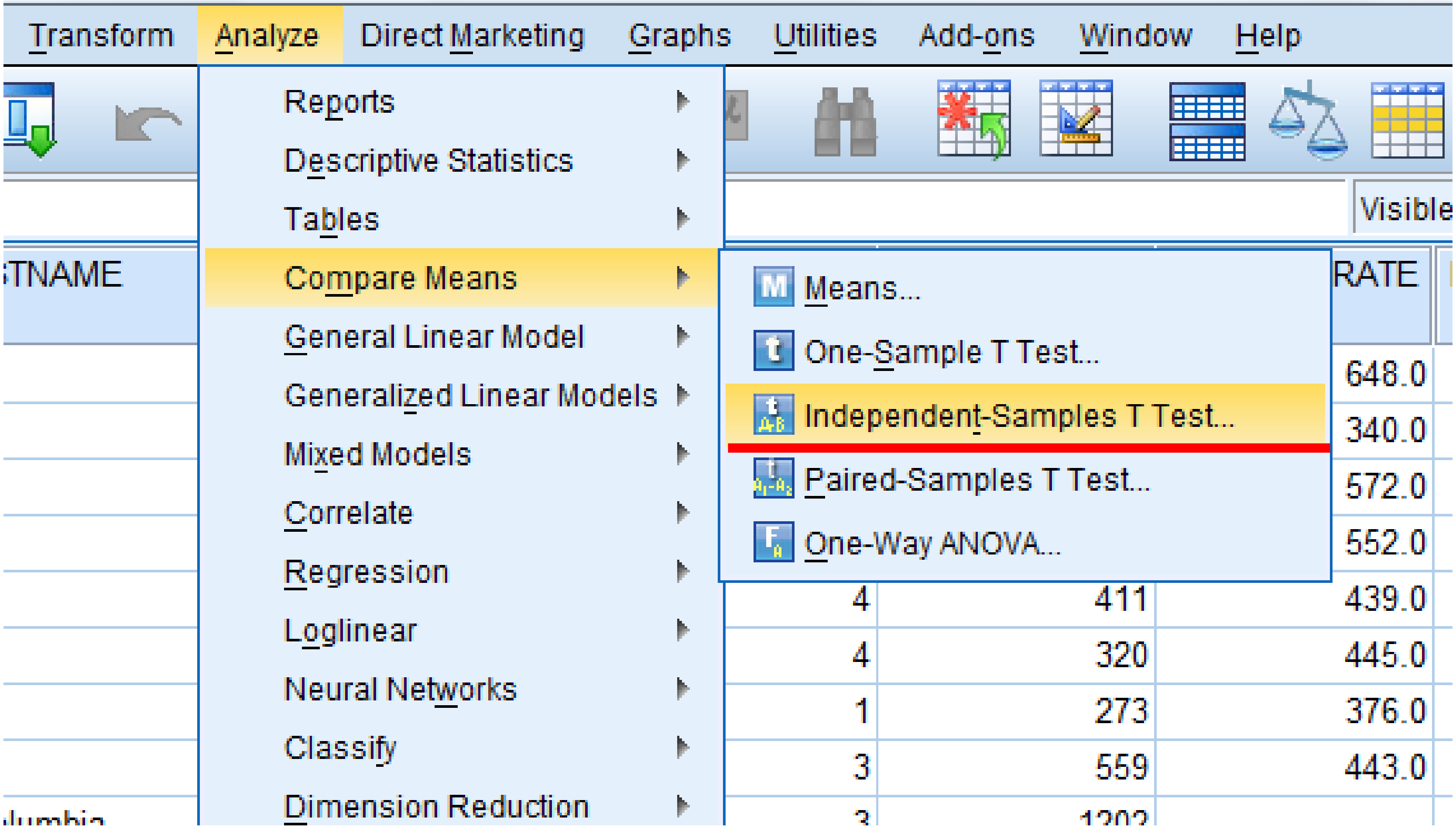
The T-test: A Large Difference



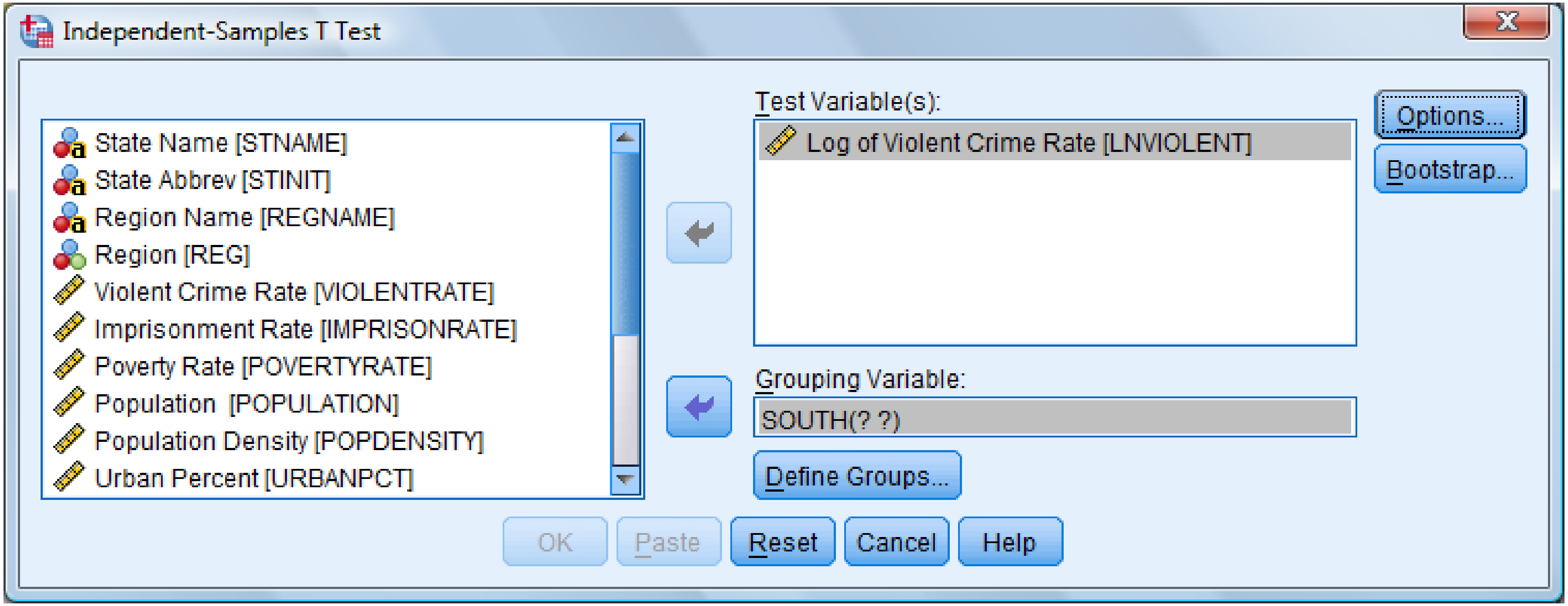
Example: Are Violent Crime Rates Higher in the South?



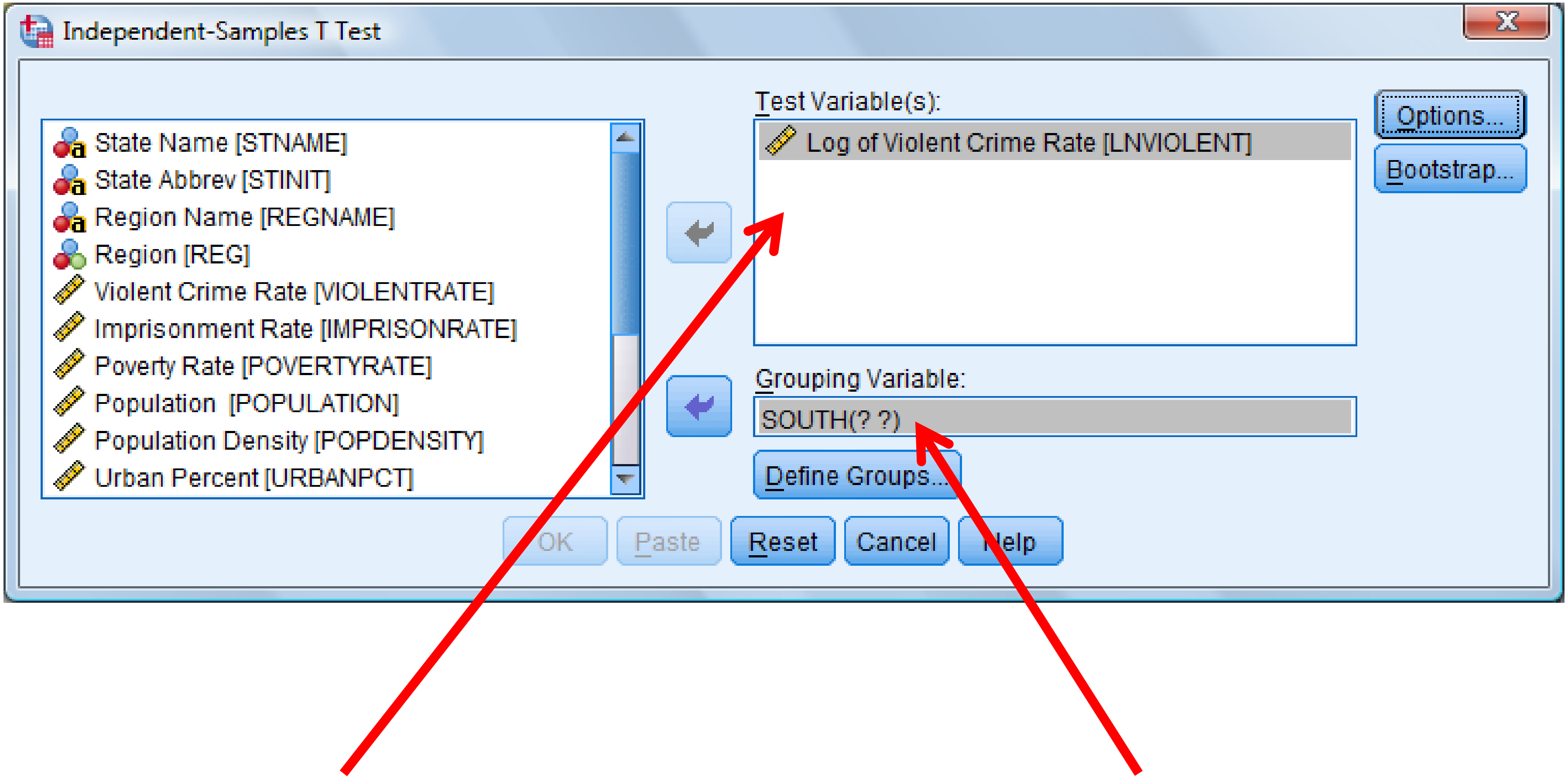
Analyze > Compare Means > Independent Samples T Test



Command Dialog Box



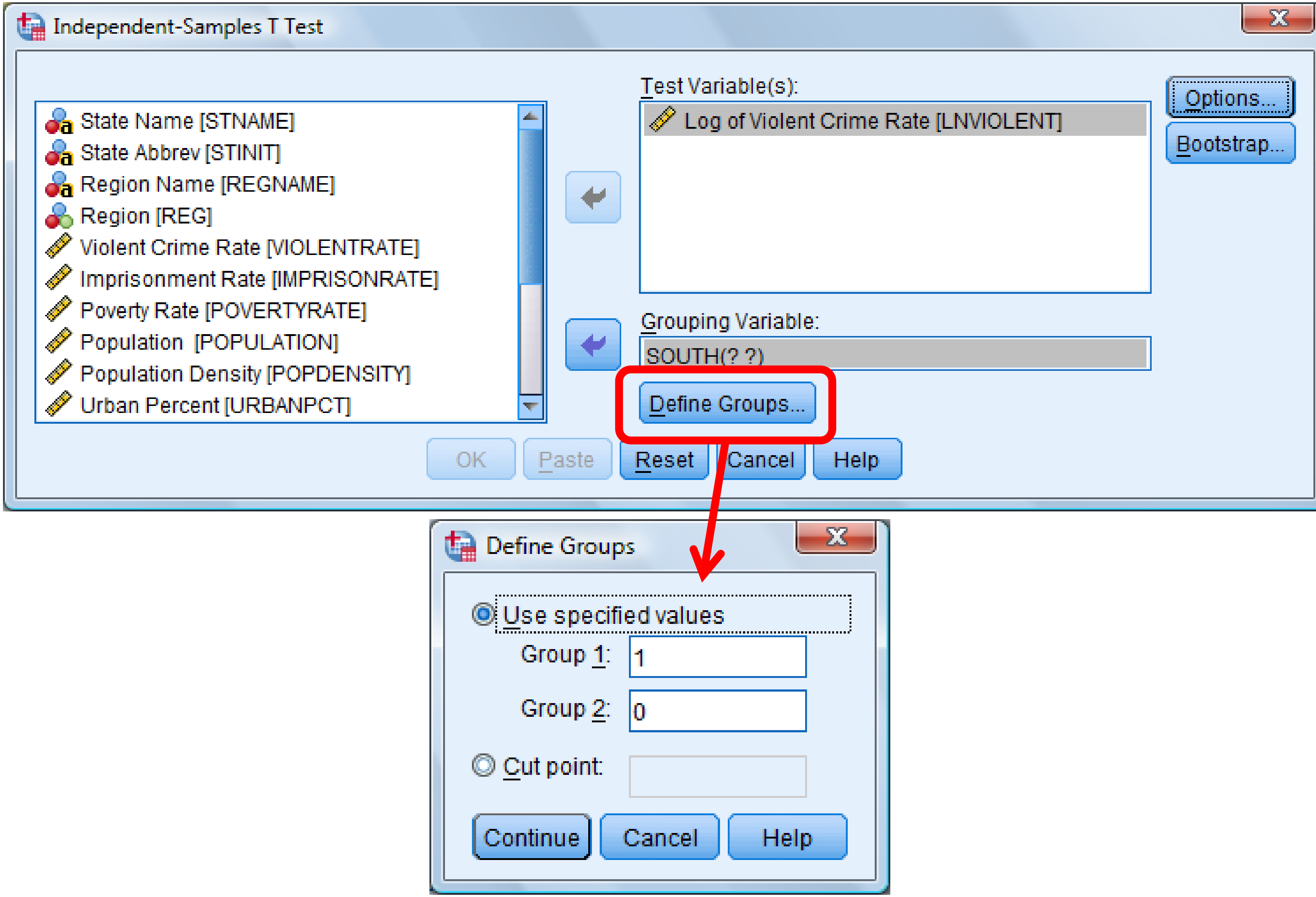
Command Dialog Box



Test Variable(s) = Numeric Variable(s)

Group Variable(s)= Categorical Variable

Command Dialog Box



Command Dialog Box after Define Groups

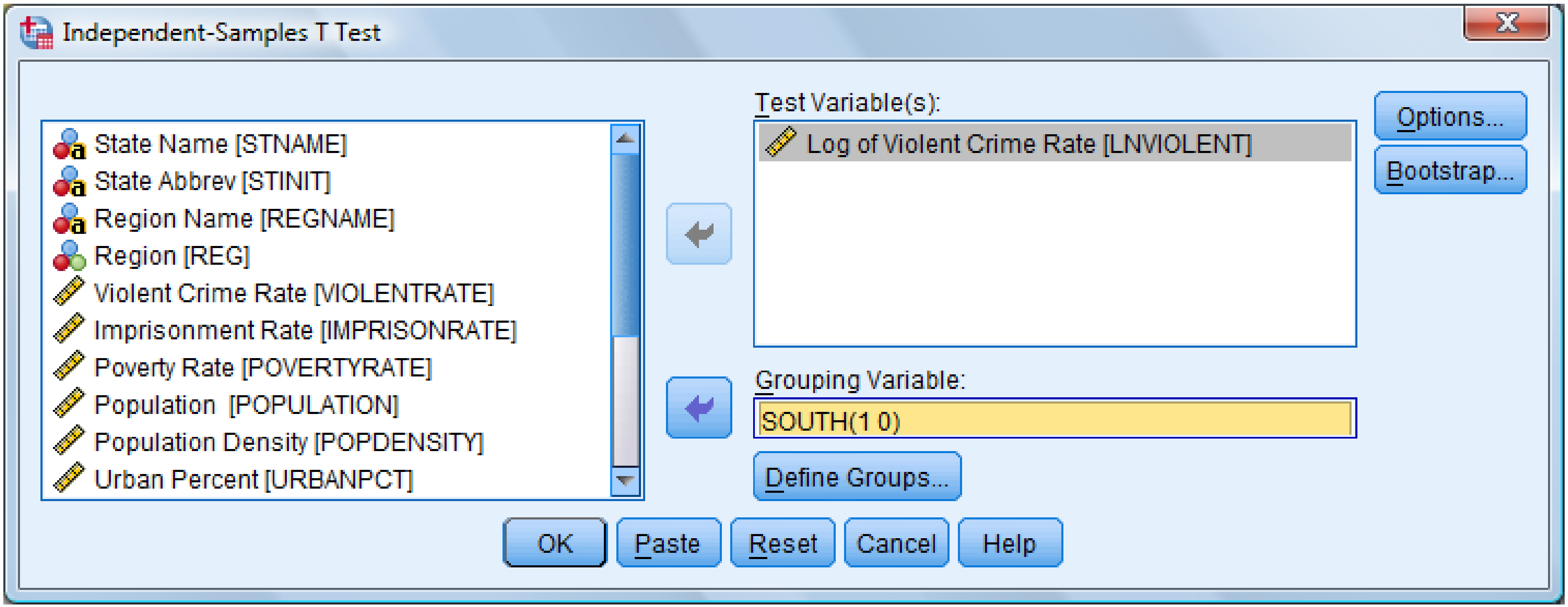


Table 1: Descriptive Statistics

Group Statistics

	South	N	Mean	Std. Deviation	Std. Error Mean
Log of Violent Crime Rate	South	17	6.0694	.41800	.10138
	Non-South	34	5.6961	.37603	.06449

Table 2: Independent Samples Test

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
Log of Violent Crime Rate	Equal variances assumed	.022	.883	3.221	49	.002	.37334	.11592	.14040	.60629
	Equal variances not assumed			3.107	29.246	.004	.37334	.12015	.12769	.61900

Levine's Test

		Levene's Test for Equ[illegible]y of Variances	
		F	Sig.
Log of Violent Crime Rate	Equal variances assumed	.022	.883
	Equal variances not assumed		

p ≥ .05, then the variances equal

p < .05, then the variances not equal

T-test for Equality of Means

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

T-test for Equality of Means: t

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

T-test for Equality of Means: df

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

T-test for Equality of Means: Significance

statistical significance of the test

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

T-test for Equality of Means: Mean Difference

magnitude (not standardized)

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

T-test for Equality of Means: Standard Error

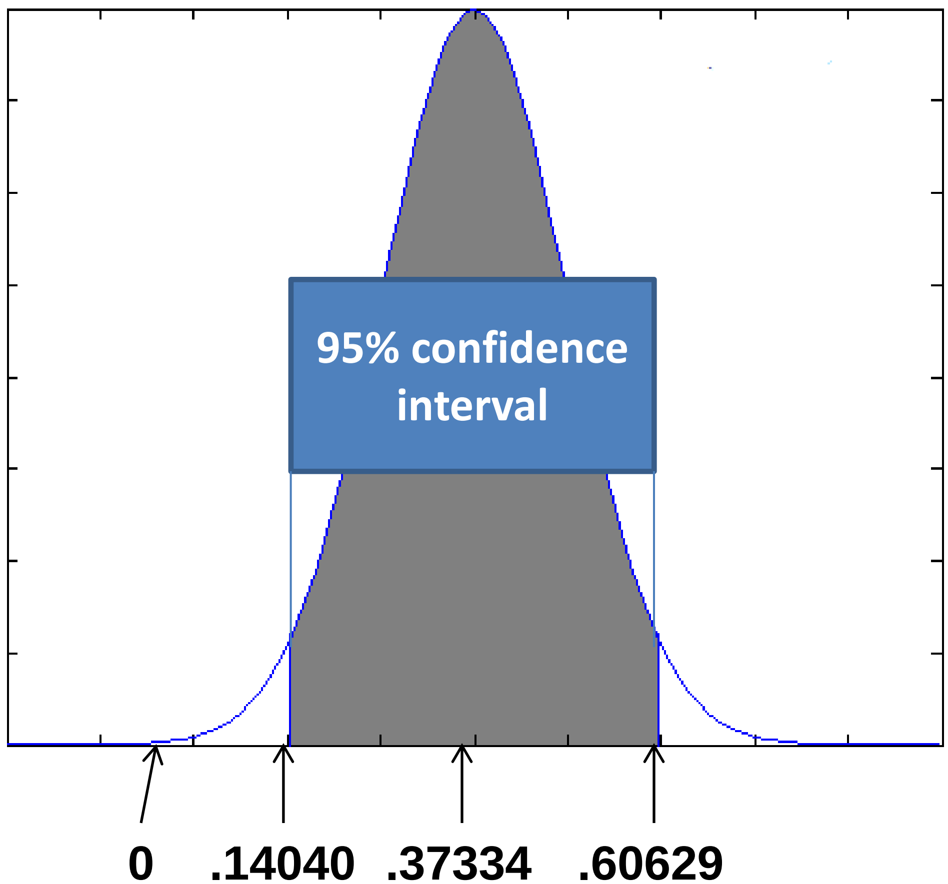
magnitude (not standardized)

t-test for Equality of Mea						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

T-test for Equality of Means: Confidence Interval

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
3.221	49	.002	.37334	.11592	.14040	.60629
3.107	29.246	.004	.37334	.12015	.12769	.61900

Null Hypothesis



Formula: Equal Variances

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{x_1 x_2}\sqrt{2/n}}$$

$$\text{where } S_{x_1 x_2} = \sqrt{\frac{(n_1 - 1)s^2_{x1} + (n_2 - 1)s^2_{x2}}{n_1 + n_2 - 2}}$$

Formula: Unequal Variances

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{x_1 x_2}\sqrt{\left(\frac{1}{n_1}\right) + \left(\frac{1}{n_2}\right)}}$$

$$\text{where } S_{x_1 x_2} = \sqrt{\frac{(n_1 - 1)s^2_{x1} + (n_2 - 1)s^2_{x2}}{n_1 + n_2 - 2}}$$

Levine’s Test Controversy

- Levine’s test is not universally accepted and not included in some statistical software.
- The issue is that the statistical significance number masks underlying aspects of the distribution of each group that should be studied more closely, much like the way we study skewness and kurtosis.
- Some statisticians assert that the Levine’s test should not substitute for one’s own judgment.
- If one cannot tell whether to assume equal or unequal variances, assume unequal.

Effect Size: Cohen's d

$$d = \frac{\text{Mean}_2 - \text{Mean}_1}{\text{SD}_{\text{pooled}}}$$

where the pooled standard deviation is

$$\text{SD}_{\text{pooled}} = \sqrt{((\text{SD}_1{}^2 + \text{SD}_2{}^2) / 2)}$$

Interpretation

.2 Weak

.5 Medium

.8 Strong

$$d = \frac{5.70 - 6.07}{.40} = \frac{.37}{.40} = .93$$

where the pooled standard deviation is $\sqrt{((.42^2 + .38_2{}^2) / 2)}$

The End

- T-tests are used when one variable is numeric and the other variable has two categories.

- The test determines whether the distance between the two means is statistically significant.

- A standardized effect size can be obtained using Cohen’s d.

- Direction does not exist in a traditional sense, though one group can be said to be “larger” or “higher” than another.