

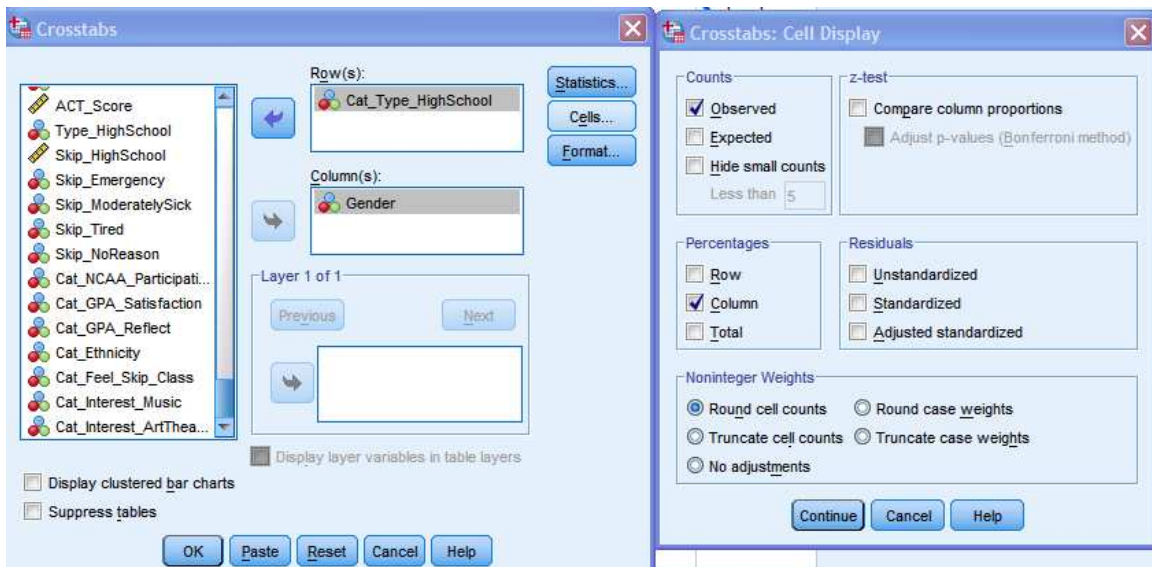
Comparing Two Proportions

We will use SPSS to compare the proportion of female Hope students that went to a private high school with the proportion of male students that went to a private high school.

Descriptive Statistics

Commands:

1. *Analyze>Descriptive Statistics>Crosstabs.*
2. Drag the explanatory variable into the "Column(s):" box.
3. Drag the response variable into the "Row(s):" box.
4. Under *Cells...* Check "Column" in the "Percentages" box.
5. Select *Continue.*
6. Select *OK.*



What type of HighSchool did you graduate from, private or public? * What is your gender? Crosstabulation

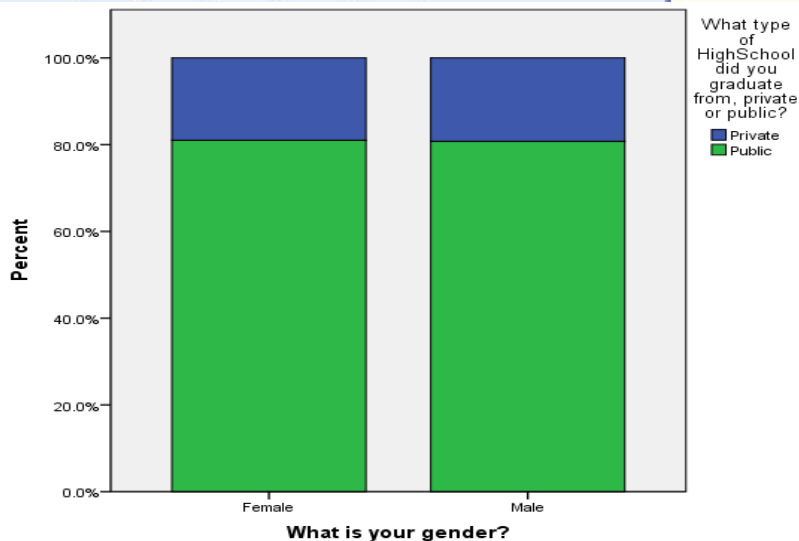
			What is your gender?		Total
			Female	Male	
What type of HighSchool did you graduate from, private or public?	Private	Count	117	47	164
		% within What is your gender?	19.0%	19.3%	19.1%
	Public	Count	499	197	696
		% within What is your gender?	81.0%	80.7%	80.9%
Total	Count	616	244	860	
	% within What is your gender?	100.0%	100.0%	100.0%	

Graphs

Commands:

1. *Graphs>Chart Builder>Bar.*
2. Drag the stacked bar to the upper window.
3. Drag the explanatory variable to the “X-Axis” box.
4. Drag the response variable to the “Stack: set color” box.
5. Under *Statistic:* choose *Percentage()*.
6. Under *Set Parameters...* choose *Total for Each X-Axis Category.*
7. Select *Continue.*
8. Select *Apply.*
9. Select *OK.*

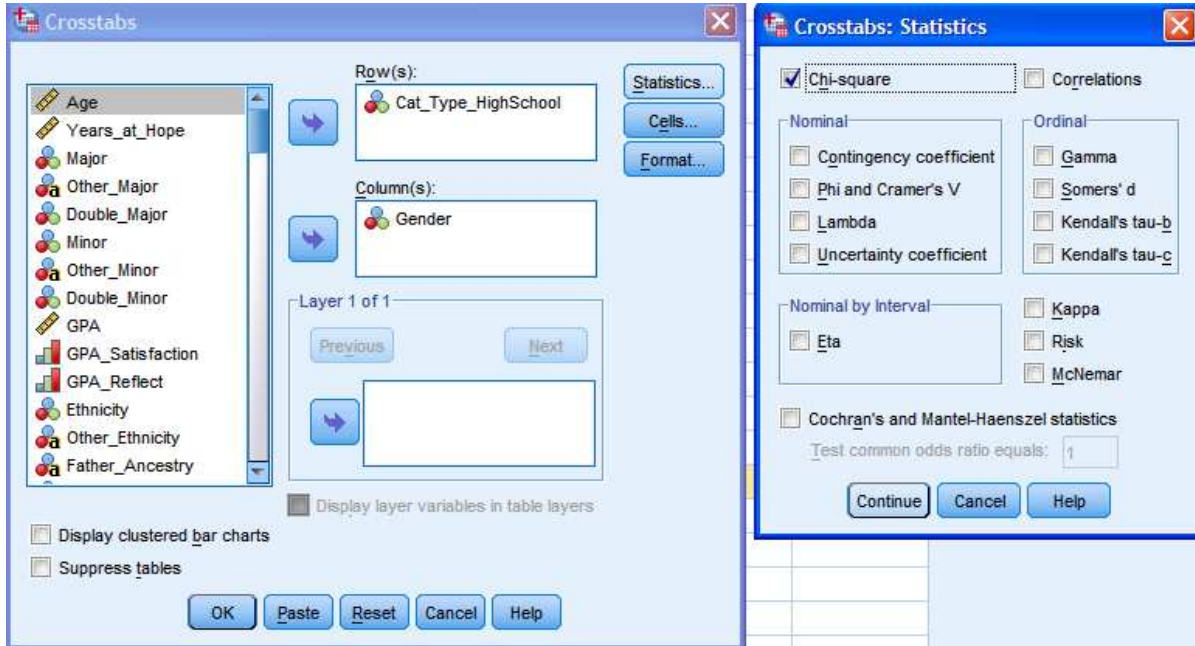
The screenshot shows the SPSS Chart Builder and Element Properties dialog boxes. The Chart Builder window displays a preview of a stacked bar chart with 'Gender' on the X-axis and 'Count' on the Y-axis. The 'Stack: set color' box contains 'Cat_Type_HighSchool'. The Element Properties dialog box shows the 'Bar1' element with 'Statistic' set to 'Percentage ()' and 'Display error bars' checked. The 'Error Bars Represent' section has 'Confidence intervals' selected with a level of 95%. The 'Bar Style' is set to 'Bar'.



Inference

Commands:

1. *Analyze>Descriptive Statistics>Crosstabs.*
2. Drag the explanatory variable into the “Column(s):” box.
3. Drag the response variable into the “Row(s):” box.
4. Under *Statistics...Check “Chi-Square.”*
5. Select *Continue.*
6. Select *OK.*



This box gives us the p-value of .928 (shown in red). The p-value is in the “Pearson Chi-Square” row and “Asymp. Sig. (2-sided)” column. Note that it is 2 sided.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.008 ^a	1	.928		
Continuity Correction ^b	.000	1	1.000		
Likelihood Ratio	.008	1	.928		
Fisher's Exact Test				.924	.499
Linear-by-Linear Association	.008	1	.928		
N of Valid Cases	860				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 46.53.

b. Computed only for a 2x2 table

Applet

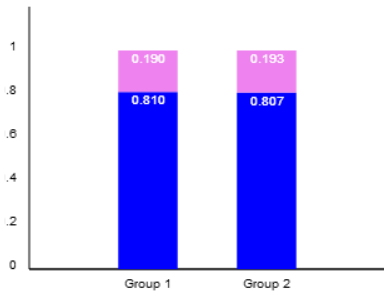
After you have found descriptive measures, you may take that information and use it in the Theory-based applet. In fact, SPSS does compute confidence intervals for the difference in proportions, so if you want this you should use the applet. Below are the graphs, p-value and confidence interval for our example.

Scenario: Two proportions

Paste Data

Group 1	Group 2
n: <input type="text" value="616"/>	n: <input type="text" value="244"/>
count: <input type="text" value="499"/>	count: <input type="text" value="197"/>
sample \hat{p} : <input type="text" value="0.8101"/>	sample \hat{p} : <input type="text" value="0.8074"/>

Sample Data



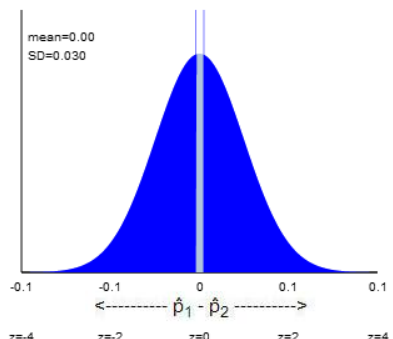
Group	Proportion
Group 1	0.810
Group 2	0.807

Theory-Based Inference

Test of significance

$H_0: \pi_1 - \pi_2 =$

$H_a: \pi_1 - \pi_2 \neq$



mean=0.00
SD=0.030

standardized statistic

p-value

Confidence interval

confidence level %

(-0.0611, 0.0557)

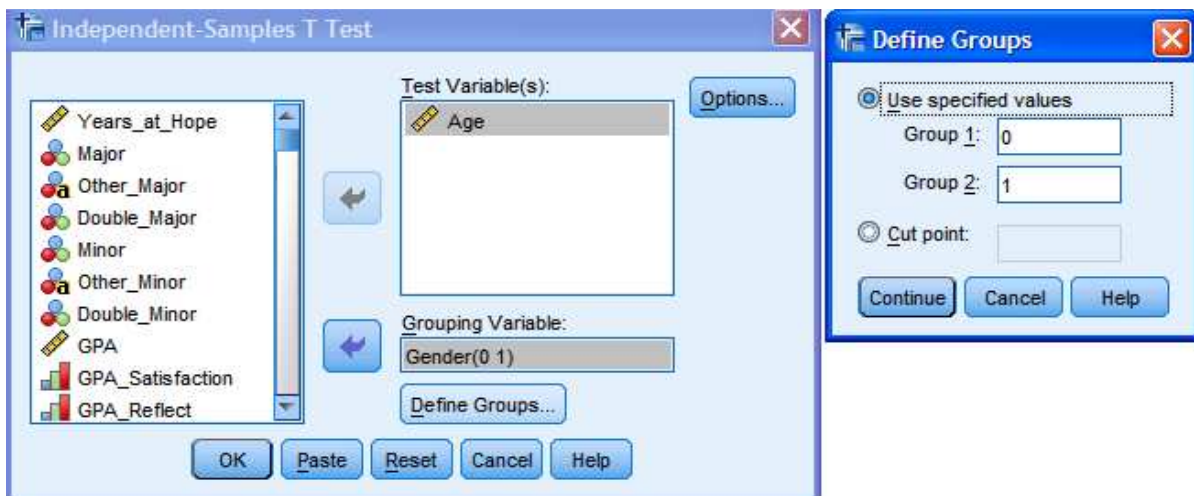
Comparing Two Means

We will use SPSS to compare the average age of female Hope students with the average age of male Hope students.

Descriptive Statistics and Inference

Commands:

1. *Analyze>Compare Means>Independent-Samples T Test...*
2. The "Test Variable" is the response variable (which should be quantitative).
3. The "Grouping Variable" is the explanatory variable (which should be categorical).
4. *Define Groups...* to indicate to SPSS which two groups are being compared. In the box "Group1," type in the variable value (as it appears in "Data View") for the first group (i.e.: "0"). In the box "Group 2," type in the variable value (as it appears in "Data View") for the second group (i.e.: "1"). Select *Continue*.



Group Statistics

What is your gender?		N	Mean	Std. Deviation	Std. Error Mean
What is your age?	Female	643	20.1213	1.67655	.06612
	Male	259	20.2355	1.30982	.08139

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
What is your age?	Equal variances assumed	.004	.948	-.982	900	.326	-.11421	.11629	-.34245	.11402
	Equal variances not assumed			-1.089	604.998	.276	-.11421	.10486	-.32015	.09172

In red above are the standardized statistic (t-statistic), the 2-sided p-value, our observed statistic (the difference in means) and a 95% confidence interval.

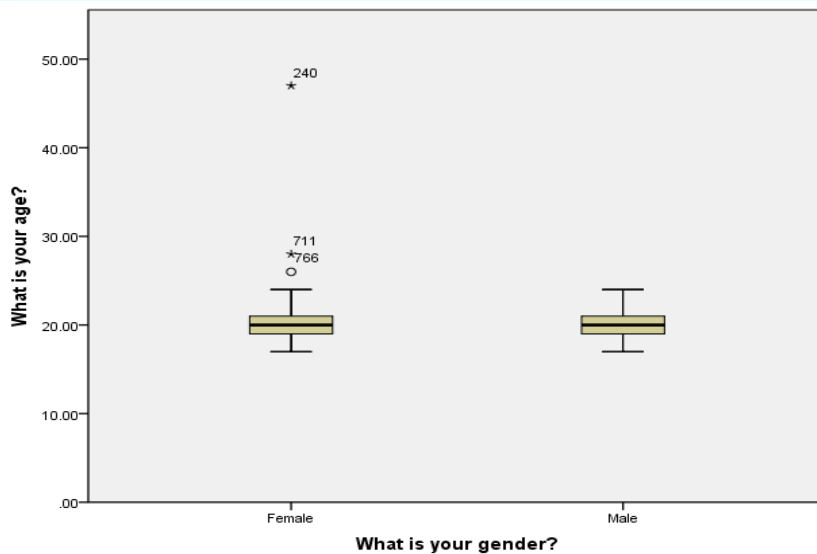
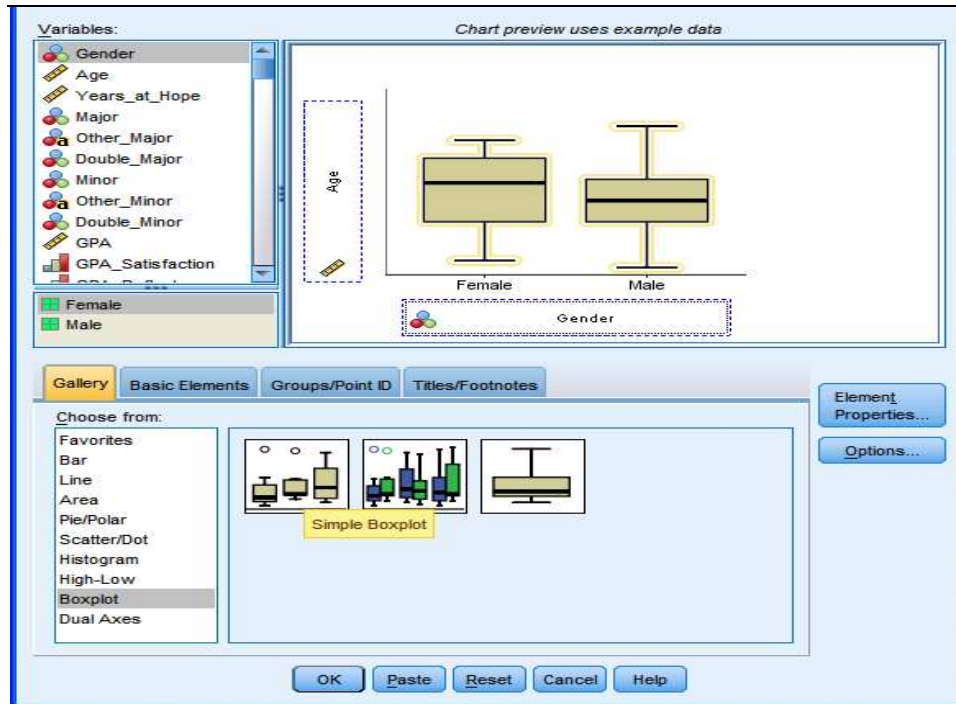


Note: SPSS will always give the two-sided p-value. To get the one-sided p-value (assuming that the sample mean of Group 1 is larger than Group 2 and the alternative hypothesis is such that the population mean of Group 1 is larger than Group 2), divide the SPSS two-sided p-value in half.

Graphs

Commands:

1. *Graphs>Chart Builder>Boxplot.*
2. Drag the simple boxplot to the upper window.
3. Drag the explanatory variable to the “X-Axis” box.
4. Drag the response variable to the “Y-Axis” box.
5. Select *OK*.



Shown here is side-by-side boxplots are shown of ages for males and females. The open circles represent potential outliers that are above the upper quartile + 1.5(IQR) and the asterisks

Applet

After you have found descriptive measures, you may take that information and use it in the Theory-based applet. If you have your data in SPSS in columns right next to each other, you may also paste that in the applet. Below are the graphs, p-value and confidence interval for our example.

Scenario:

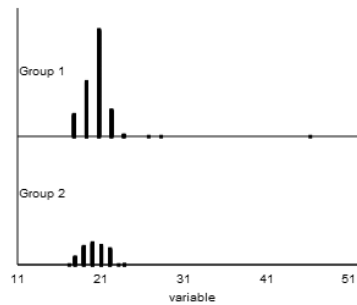
- Paste Data
- Stacked (Group# Value)
- Includes header

Sample Data:

0	19.00
0	20.00
1	21.00
0	21.00
1	22.00
1	20.00
0	19.00
0	21.00
1	20.00
0	21.00

Group 1	Group 2
n: <input type="text" value="643"/>	n: <input type="text" value="259"/>
mean, \bar{x} : <input type="text" value="20.121"/>	mean, \bar{x} : <input type="text" value="20.236"/>
sample sd, s: <input type="text" value="1.677"/>	sample sd, s: <input type="text" value="1.310"/>

Sample Data

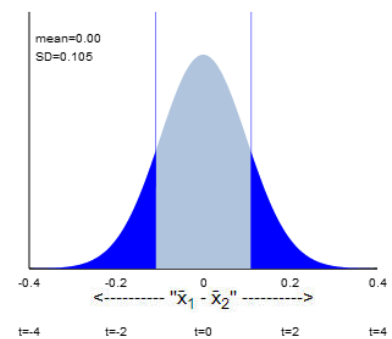


Theory-Based Inference

Test of significance

$H_0: \mu_1 - \mu_2 =$

$H_a: \mu_1 - \mu_2 \neq$



standardized statistic $t =$ $df =$

p-value

Confidence interval

confidence level %

(-0.3210, 0.0910)