Chi-square Test of Independence & McNemar Test

Dr Daniel Boduszek

d.boduszek@hud.ac.uk

Outline

- Chi-square test for association
  - SPSS procedure
  - Interpretation of SPSS output
  - Reporting
- McNemar test
  - SPSS procedure
  - Interpretation of SPSS output
  - Reporting

Chi-Square test for association

- The chi-square test for independence, also called Pearson's chi-square test or the chi-square test of association

- It is used to discover if there is a relationship between **two categorical** variables

Chi-Square test for association

- Assumptions
  - 2 variables should be measured at an ordinal or nominal level
  - variables should consist of two or more categorical, independent groups.
    - This example include offending type (2 categories: violent and non-violent offenders), age (e.g., 3 groups: young, middle-age, and older)

Expected Frequencies versus Observed Frequencies

- The chi-square test of independence plugs the observed frequencies and expected frequencies into a formula which <u>computes how the pattern of observed frequencies differs from the pattern of expected frequencies</u>.

- Probabilities for the test statistic can be obtained from the chi-square probability distribution so that we can test hypotheses.

Hypotheses

- The **research hypothesis** states that the two variables are dependent or related. This will be true if the observed counts for the categories of the variables in the sample are different from the expected counts.

- The **null hypothesis** is that the two variables are independent. This will be true if the observed counts in the sample are similar to the expected counts.

Computing the Test Statistic

- Conceptually, the chi-square test of independence statistic is computed by summing the difference between the expected and observed frequencies for each cell in the table divided by the expected frequencies for the cell.

- We identify the value and probability for this test statistic from the SPSS statistical output.

Decision and Interpretation

- If the probability of the test statistic is less than or equal to the probability of the alpha error rate, we reject the null hypothesis and conclude that our data supports the research hypothesis. We conclude that there is a relationship between the variables.

- If the probability of the test statistic is greater than the probability of the alpha error rate, we fail to reject the null hypothesis. We conclude that there is no relationship between the variables, i.e. they are independent.

Which Cell or Cells Caused the Difference

- We are only concerned with this procedure if the result of the chi-square test is statistically significant.

- One of the problems in interpreting chi-square tests is the determination of which cell or cells produced the statistically significant difference. Examination of percentages in the contingency table and expected frequency table can be misleading.

- The **residual**, or the difference, between the observed frequency and the expected frequency is a more reliable indicator, especially if the residual is converted to a z-score and compared to a critical value equivalent to the alpha for the problem.

Standardized Residuals

- SPSS prints out the standardized residual (converted to a z-score) computed for each cell.
    - It does not produce the probability or significance.

- Without a probability, we will compare the size of the **standardized residuals** to the critical values that correspond to an **alpha of 0.05 (+/-1.96)** or an alpha of 0.01 (+/-2.58).
    - This is equivalent to testing the null hypothesis that the actual frequency equals the expected frequency for a specific cell versus the research hypothesis of a difference greater than zero.

- There can be 0, 1, 2, or more cells with statistically significant standardized residuals to be interpreted.

Interpreting Standardized Residuals

- Standardized residuals that have a <u>positive</u> value mean that the cell was <u>over-represented</u> in the actual sample, compared to the expected frequency, i.e. there were more subjects in this category than we expected.

- Standardized residuals that have a <u>negative</u> value mean that the cell was <u>under-represented</u> in the actual sample, compared to the expected frequency, i.e. there were fewer subjects in this category than we expected.
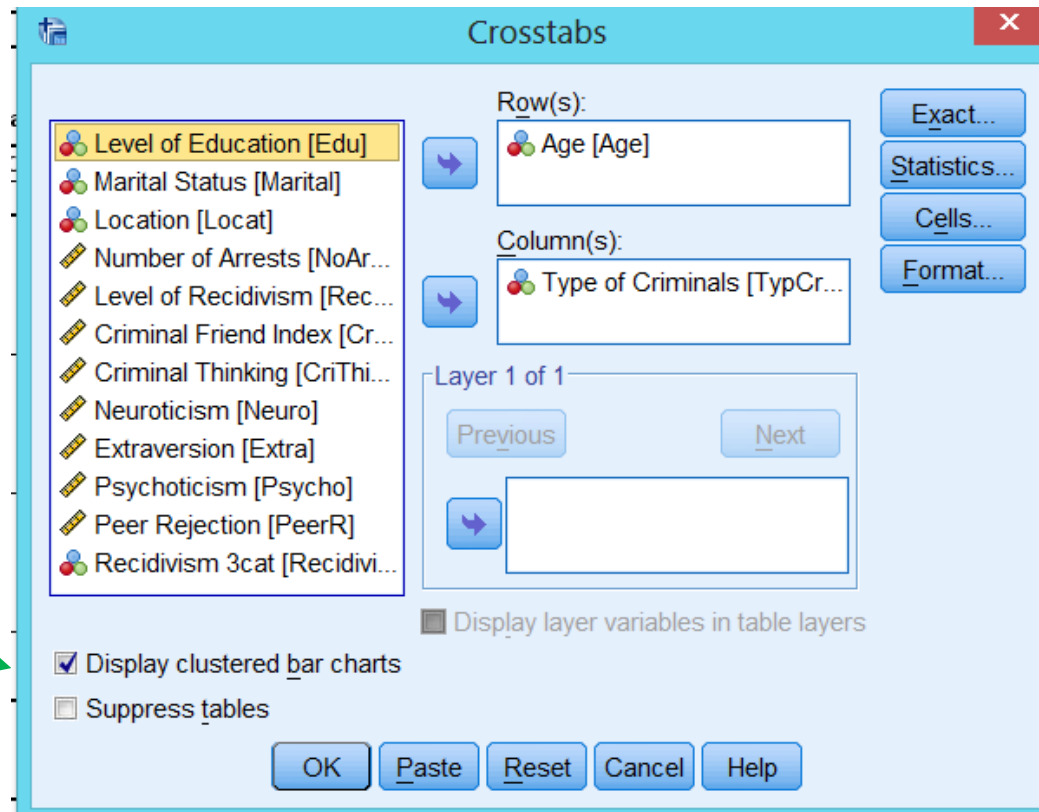
SPSS Procedure

- Click **Analyze** > **Descriptive Statistics** > **Crosstabs.**
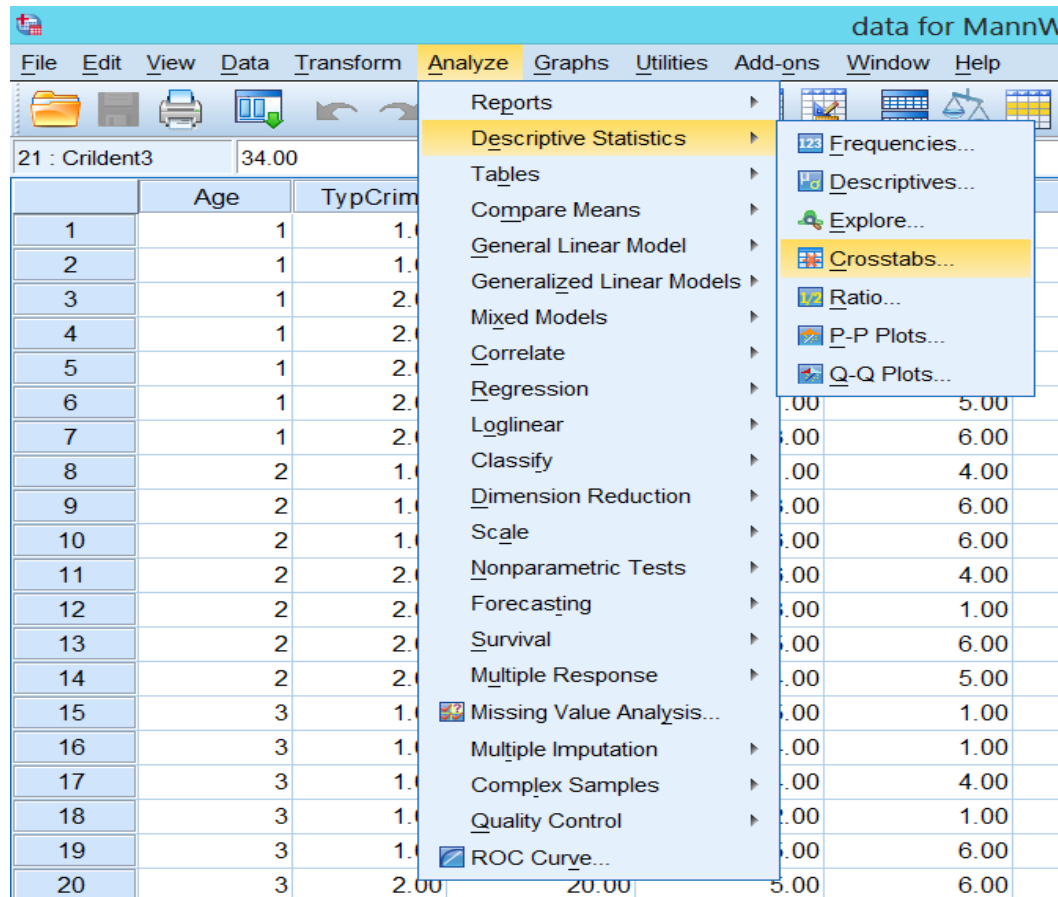
SPSS Procedure

- Transfer one of the variables into the **Row(s)**: box and the other variable into the **Column(s)**: box.
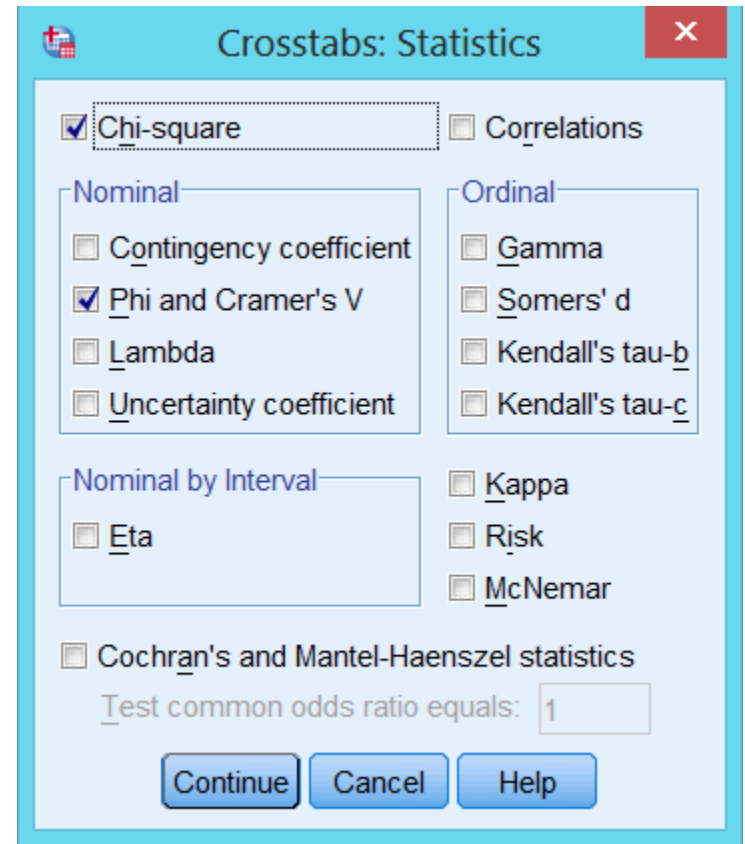
If you want to display clustered bar charts, make sure that **Display clustered bar charts** checkbox is ticked
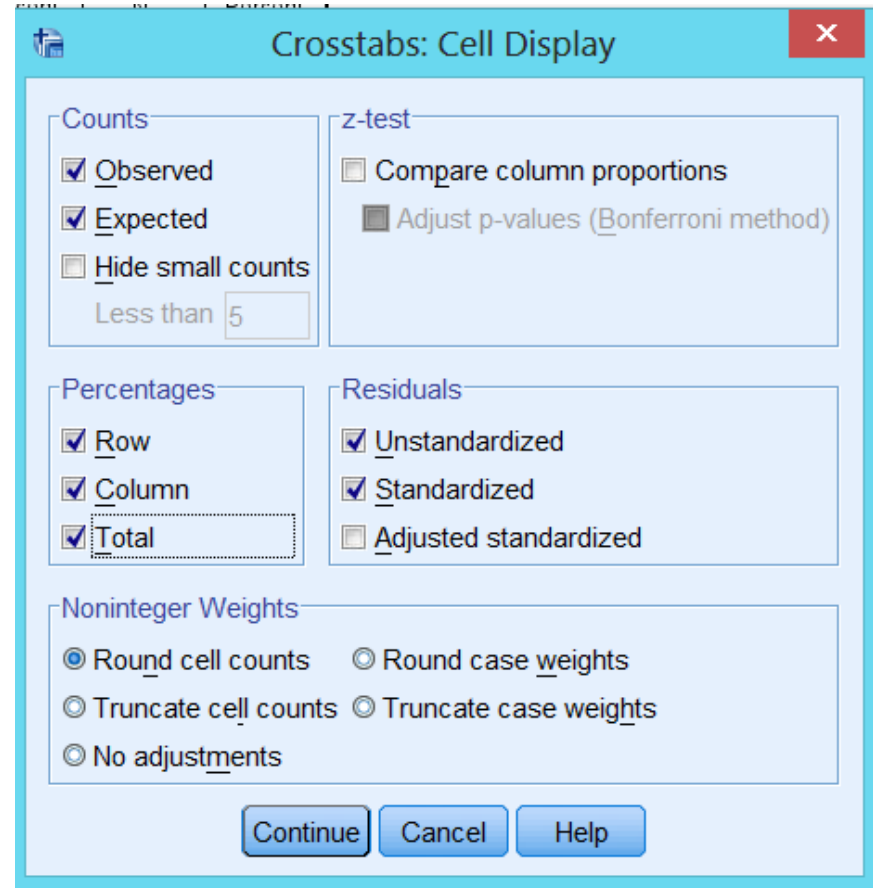
SPSS Procedure

- Click on the **Statistics** button.

- Select the **Chi-square** and **Phi and Cramer's V** options

- Click **Continue**

SPSS Procedure

- Click the **Cells**

- Select Observed and Expected from the -**Counts**- area, and Row, Column and Total from the -**Percentages**- area

- Also Unstandardized and Standardized from **Residuals** (post-hoc)

- Click **Continue** and **OK**

SPSS Output

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	34.277[a]	2	.000
Likelihood Ratio	36.052	2	.000
Linear-by-Linear Association	3.440	1	.064
N of Valid Cases	309		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 24.92.

In the table Chi-Square Tests result, SPSS also tells us that "0 cells have expected count less than 5 and the minimum expected count is 24.92".

The sample size requirement for the chi-square test of independence is satisfied.

SPSS Output

- We can see here that Chi-square (2) = 34.277, $p < 0.05$. This tells us that there is statistically significant association between violent/non-violent offending and age groups.

The probability of the chi-square test statistic (chi-square=34.277) was p=0.000, less than the alpha level of significance of 0.05.

The research hypothesis that differences in "violent offending" are related to differences in "age" is supported by this analysis.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	34.277[a]	2	.000
Likelihood Ratio	36.052	2	.000
Linear-by-Linear Association	3.440	1	.064
N of Valid Cases	309		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 24.92.

SPSS Output

- **Phi** and **Cramer's V** are both tests of the strength of association.

- We can see that the strength of association between the variables is moderate (0.33)

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.333	.000
	Cramer's V	.333	.000
N of Valid Cases		309	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Post-hoc test

Age * Type of Criminals Crosstabulation

			Type of Criminals .00	Type of Criminals NonV	Total
Age	18 - 25	Count	39	31	70
		Expected Count	45.1	24.9	70.0
		% within Age	55.7%	44.3%	100.0%
		% within Type of Criminals	19.6%	28.2%	22.7%
		Residual	-6.1	6.1	
		Std. Residual	-.9	1.2	
	26 - 35	Count	107	22	129
		Expected Count	83.1	45.9	129.0
		% within Age	82.9%	17.1%	100.0%
		% within Type of Criminals	53.8%	20.0%	41.7%
		Residual	23.9	-23.9	
		Std. Residual	2.6	-3.5	
	36 and more	Count	53	57	110
		Expected Count	70.8	39.2	110.0
		% within Age	48.2%	51.8%	100.0%
		% within Type of Criminals	26.6%	51.8%	35.6%
		Residual	-17.8	17.8	
		Std. Residual	-2.1	2.9	
Total		Count	199	110	309
		Expected Count	199.0	110.0	309.0
		% within Age	64.4%	35.6%	100.0%
		% within Type of Criminals	100.0%	100.0%	100.0%

The residual is the difference between the actual frequency and the expected frequency (107-83.1=23.9).

When converted to a z-score, the standardized residual (2.6) was greater than the critical value (1.96), supporting a specific finding that among prisoners who were violent offenders, there were more who reported their age 26-35 than would be expected.

McNemar test

- Chi-square test for within-subjects designs is called McNemar's chi-square.

- As with the paired t-test or the within-subjects ANOVA, the McNemar test is used whenever the same individuals are measured (or surveyed) twice, matched on some variable

- **This example:** attitudes towards rape victim before after exposure to victim's story (responses coded yes/no)
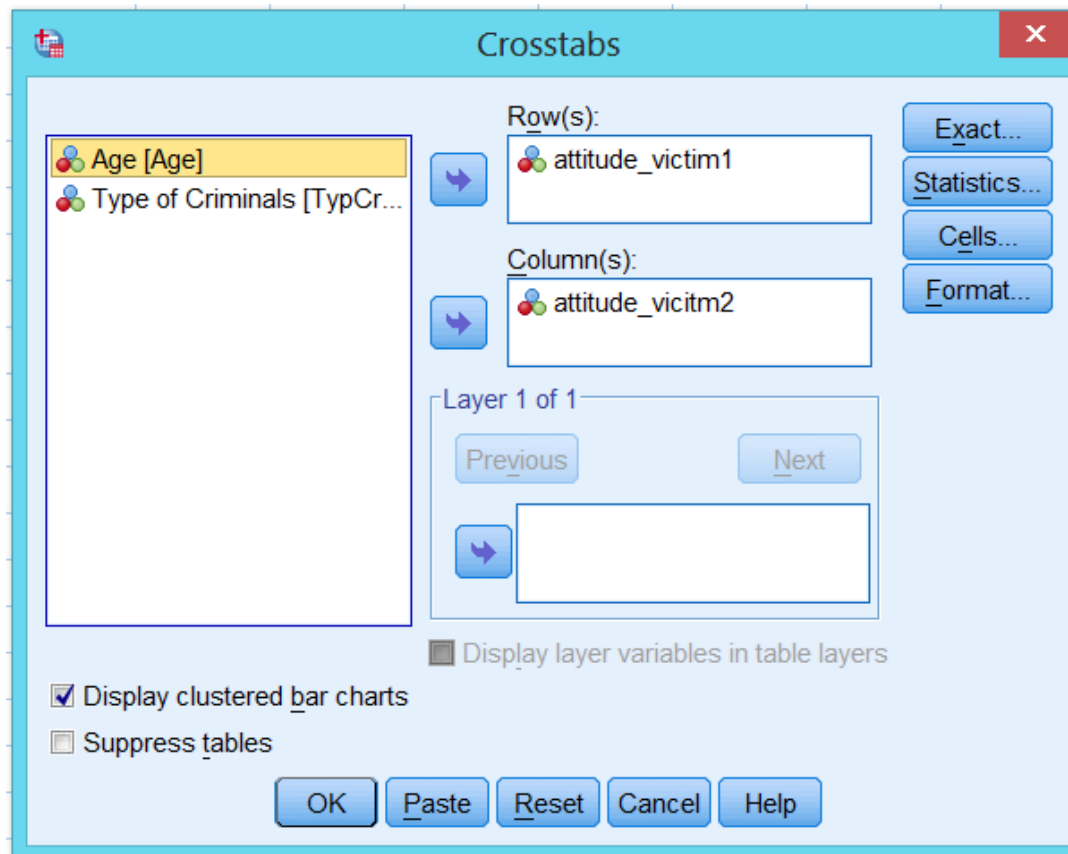
SPSS Procedure

- Click **Analyze** > **Descriptive Statistics** > **Crosstabs.**
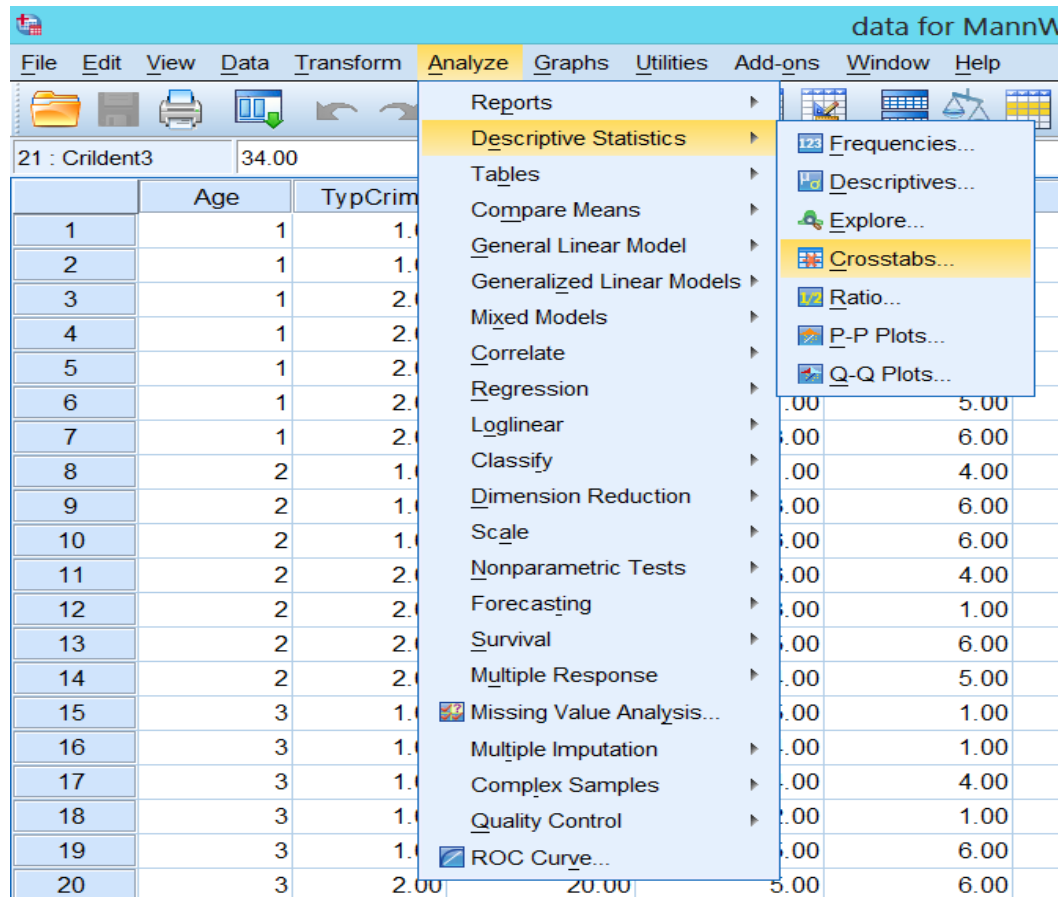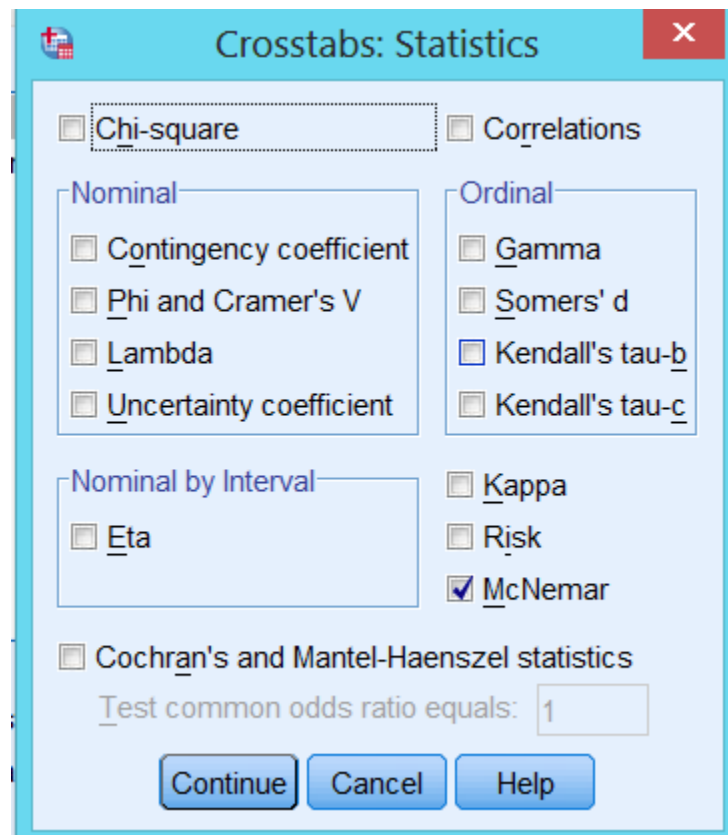
SPSS Procedure

- Move over the two variables to the row and column boxes (used rows for the pre-test and columns for the post-test)
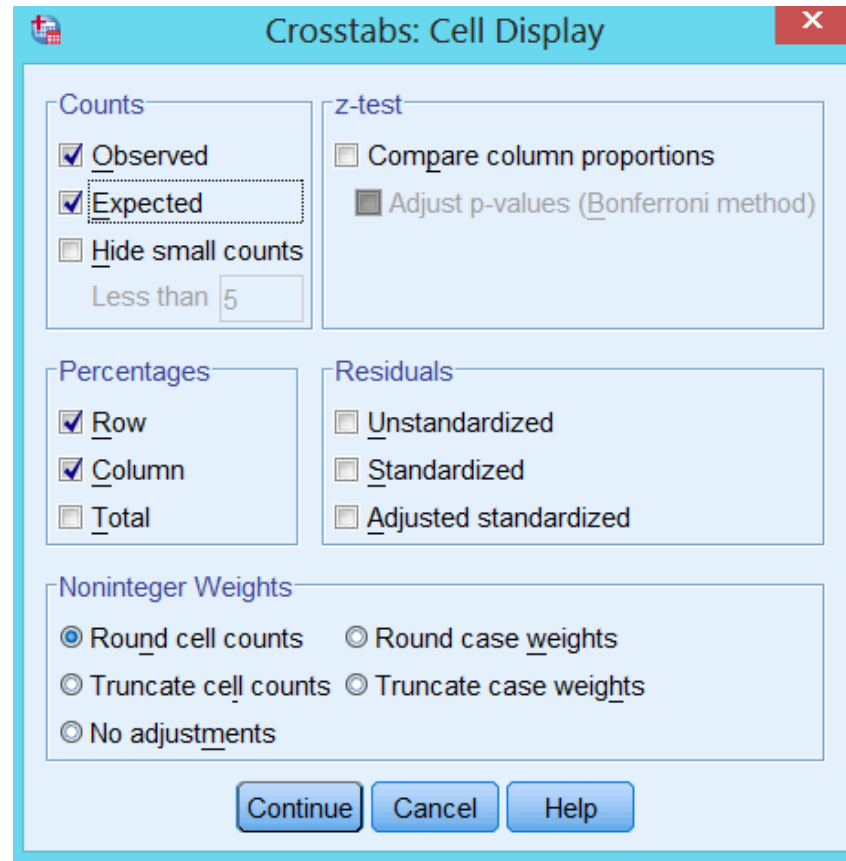
SPSS Procedure

- Click on **Statistics** and check **McNemar**, then click **Continue**

SPSS Procedure

- Click on **Cells** and then check **Row** and **Column** under **Percentages**, and **Observed** and **Expected** under **Counts**, then click **Continue** and **OK**

SPSS Outcome

attitude_victim1 * attitude_vicitm2 Crosstabulation

			attitude_vicitm2 no	attitude_vicitm2 yes	Total
attitude_victim1	no	Count	33	147	180
		Expected Count	58.3	121.7	180.0
		% within attitude_victim1	18.3%	81.7%	100.0%
		% within attitude_vicitm2	33.0%	70.3%	58.3%
	yes	Count	67	62	129
		Expected Count	41.7	87.3	129.0
		% within attitude_victim1	51.9%	48.1%	100.0%
		% within attitude_vicitm2	67.0%	29.7%	41.7%
Total		Count	100	209	309
		Expected Count	100.0	209.0	309.0
		% within attitude_victim1	32.4%	67.6%	100.0%
		% within attitude_vicitm2	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		.000[a]
N of Valid Cases	309	

a. Binomial distribution used.

- Attitudes toward the rape victims changed significantly **(p < .001)**. Participants were more likely to be more empathic after reading the victim's story (**67.6%**) than before (**32.4%**)

- **Note** that SPSS does not give the value of the McNemar chi-square, just p-value.

- Also, in this case, it makes more sense to use the marginal total percentages rather than the percentages within particular cells