

Research Methods

Chi-square for Independence

This worksheet will take you through the process to practice the chi-square test for independence. This test is used to discover if there is a relationship between two categorical variables.

There are a number of additional data sets at the end of this worksheet for you to practice this type of analysis. For each data set, you should input the data in the format shown on the following page, with a column for each of the variables. The first example depicts data collected from a study exploring the association between political party preference and Brexit vote.

Remember! – you are using categorical variables, so the first thing you will need to do is create your variables accordingly using numerical values.

	Labour	Conservative	Other	Total
Leave	5	10	5	20
Remain	12	1	7	20
Total	17	11	12	N = 40

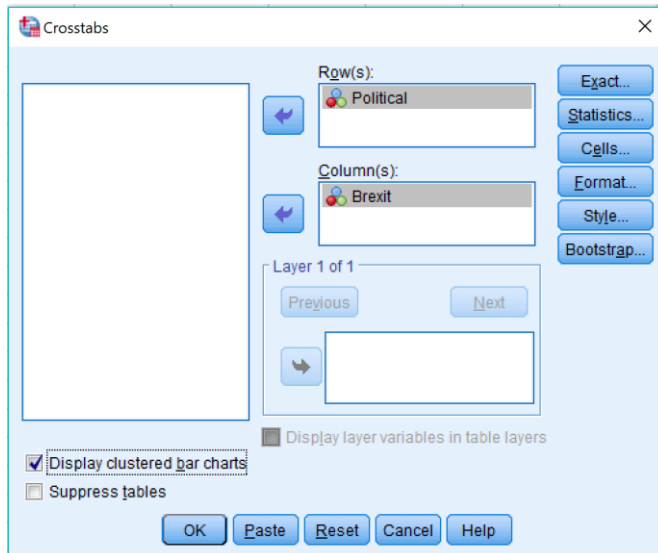
	Political	Brexit			
1	1.00	2.00	21	1.00	1.00
2	1.00	2.00	22	3.00	2.00
3	2.00	1.00	23	2.00	1.00
4	2.00	1.00	24	1.00	2.00
5	2.00	1.00	25	2.00	1.00
6	1.00	2.00	26	3.00	2.00
7	3.00	1.00	27	1.00	2.00
8	1.00	1.00	28	2.00	1.00
9	2.00	1.00	29	1.00	1.00
10	3.00	1.00	30	3.00	2.00
11	2.00	1.00	31	2.00	1.00
12	1.00	2.00	32	1.00	2.00
13	3.00	2.00	33	3.00	1.00
14	1.00	2.00	34	3.00	1.00
15	3.00	2.00	35	2.00	1.00
16	3.00	2.00	36	1.00	1.00
17	1.00	2.00	37	1.00	2.00
18	1.00	2.00	38	3.00	2.00
19	2.00	2.00	39	3.00	1.00
20	1.00	1.00	40	1.00	2.00

Now, we can carry out the statistical test to make a judgement about how 'significant' any relationship is between the variables.

- Select 'Analyze' on the top menu bar
- Go to the 'Descriptive statistics' section
- Select the 'Crosstabs' option

You will be presented with the following **Crosstabs** dialogue box:

The screenshot shows the 'Crosstabs' dialog box in SPSS. On the left, a list of variables includes 'Political' and 'Brexit'. 'Political' is moved to the 'Row(s):' field, and 'Brexit' is moved to the 'Column(s):' field. Below these fields, there is a 'Layer 1 of 1' section with 'Previous' and 'Next' buttons. At the bottom, there are checkboxes for 'Display clustered bar charts' (checked) and 'Display layer variables in table layers' (checked). On the right side, there are buttons for 'Exact...', 'Statistics...', 'Cells...', 'Format...', 'Style...', and 'Bootstrap...'. At the bottom of the dialog, there are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

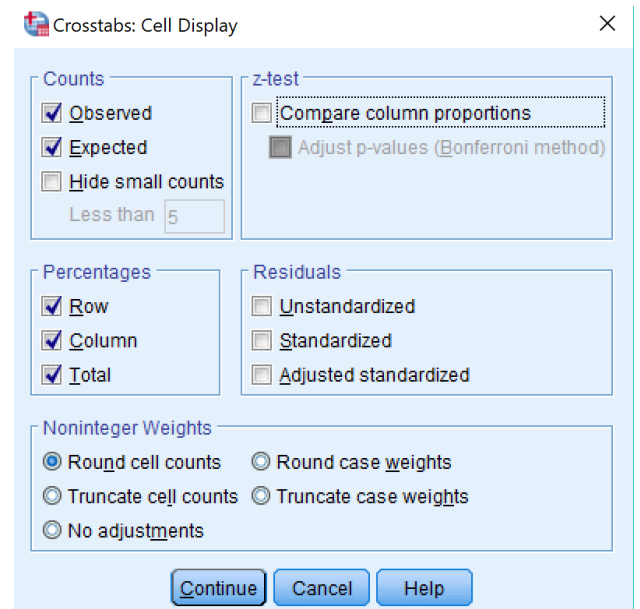


Move one of the variables into the **Rows** box and the other variable into the **Columns** box. In this example, Political has been transferred into the rows box, and Brexit into the columns box.

Check the box that says **Display clustered bar charts**.

Click on the **Statistics** button and select **Chi-Square** and **Phi and Cramer's V** options. Then click continue.

Next, click the **Cells** button and make sure the following are selected:



Click on **Continue** and **OK** to run the analysis.

The SPSS output will provide a table labelled **Chi-Square Tests**. When reading this table, you are interested in the results of the Pearson Chi-Square row. We can see here that:

$$\chi^2 (2, N = 40) = 10.58, p = .005$$

This tells us that there is a statistically significant association between participant's political party preferences and their Brexit vote. But now we need to interpret this finding appropriately, and we use the table labelled **Political * Brexit Crosstabulation** to do this.

Political * Brexit Crosstabulation

Political	labour		Brexit		Total
			leave	remain	
		Count	5	12	17
		Expected Count	8.5	8.5	17.0
		% within Political	29.4%	70.6%	100.0%
		Adjusted Residual	-2.2	2.2	
	conservative	Count	10	1	11
		Expected Count	5.5	5.5	11.0
		% within Political	90.9%	9.1%	100.0%
		Adjusted Residual	3.2	-3.2	
	other	Count	5	7	12
		Expected Count	6.0	6.0	12.0
		% within Political	41.7%	58.3%	100.0%
		Adjusted Residual	-.7	.7	
Total	Count		20	20	40
	Expected Count		20.0	20.0	40.0
	% within Political		50.0%	50.0%	100.0%

The row percentages describe the association we have established with our (significant) chi-square test...

Total leave = 50%

Total Conservative * Leave = 90.9%

This tells us that Conservative supporters are highly overrepresented in those who voted to Leave.

The opposite pattern holds true for labour voters – 70.6% of them voted to Remain

Write the full results of this analysis below:

Additional data-sets for practice:

1)

	Smoker	Non-smoker	Total
Male	9	1	10
Female	4	6	10

2)

	Political	Non-political	Total
18-25	2	8	10
26-40	4	6	10
40+	7	3	10

3)

	Pass	Fail	Total
Attended	15	5	20
Skipped	7	13	20