Lab#10.

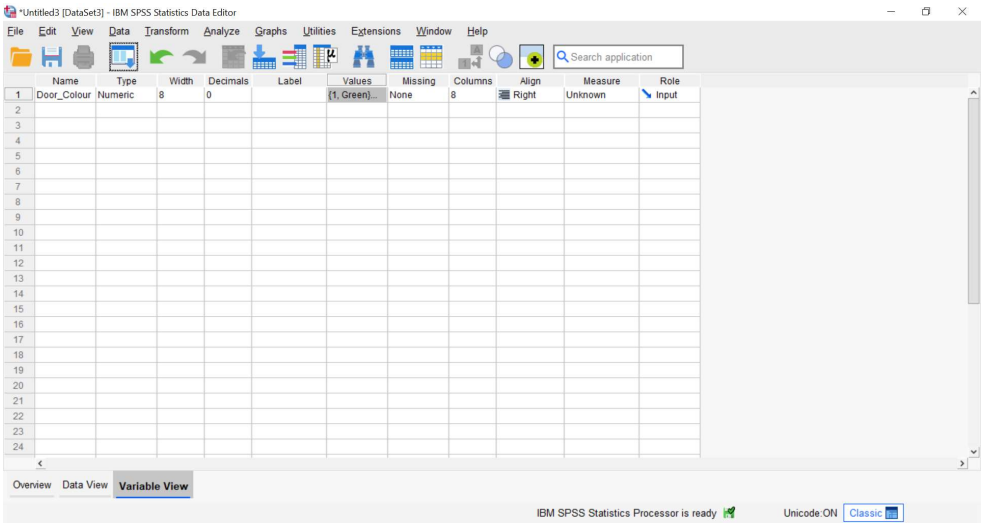
I. Chi-squared test for the Multinomial Distribution

A researcher designs an experiment in which a rat is attracted to the end of a ramp that divides, leading to doors of 3 different colours. The rat is sent down the ramp 90 times and its choice of door colour is observed. The results are: green door chosen 20 times, red door chosen 39 times and blue door chosen 31 times. Do these data provide evidence to suggest that the rat prefers one door colour over another? Use $\alpha = 0.05$.

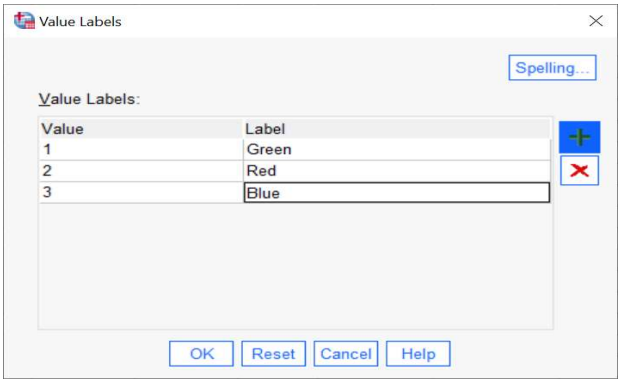
	Door Colour			
	Gteen	Red	Blue	Total
Observed	20	39	31	90
Expected	30	30	30	90

Once you open SPSS, in the **New Files** field choose **New Dataset** and click **OK**

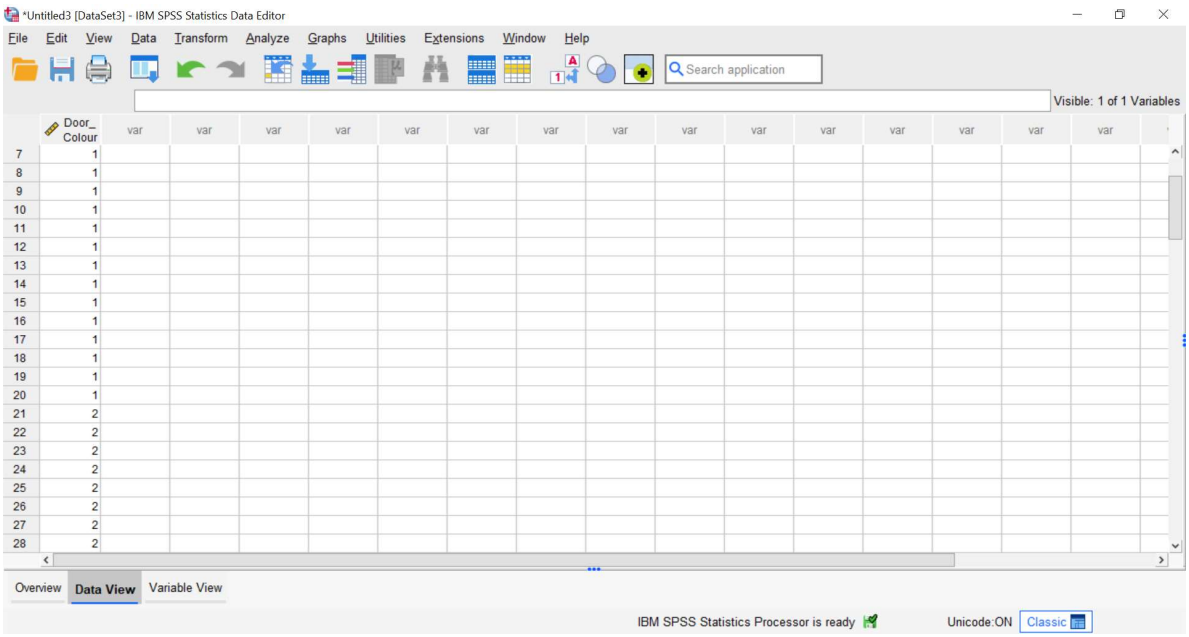
In **Variable View** create your variables:

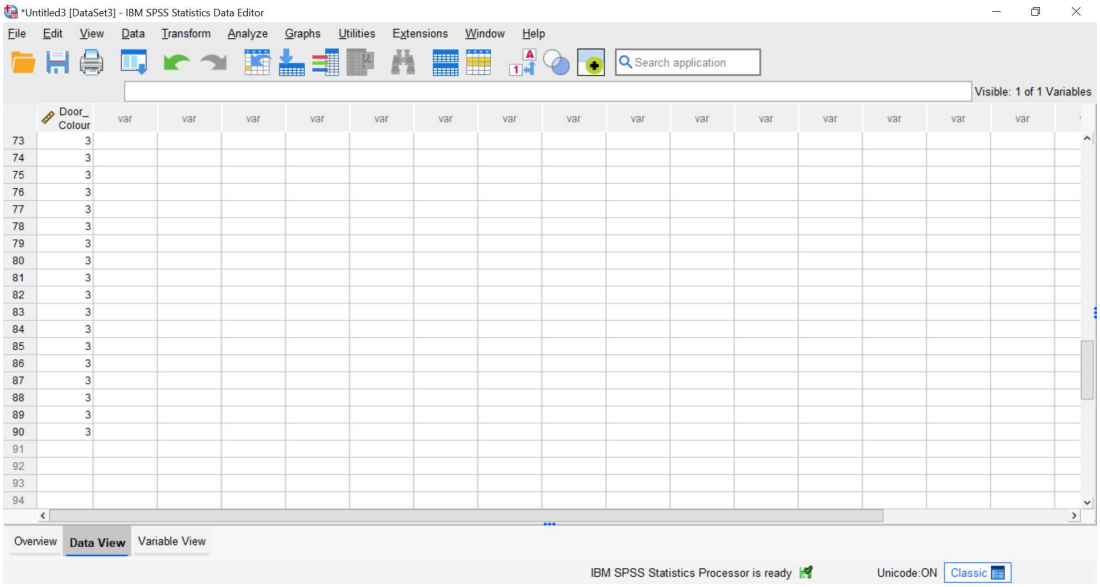


For “Green Door” we assign value 1, for “Red Door” value 2 and for”Blue Door” value 3 (see below)



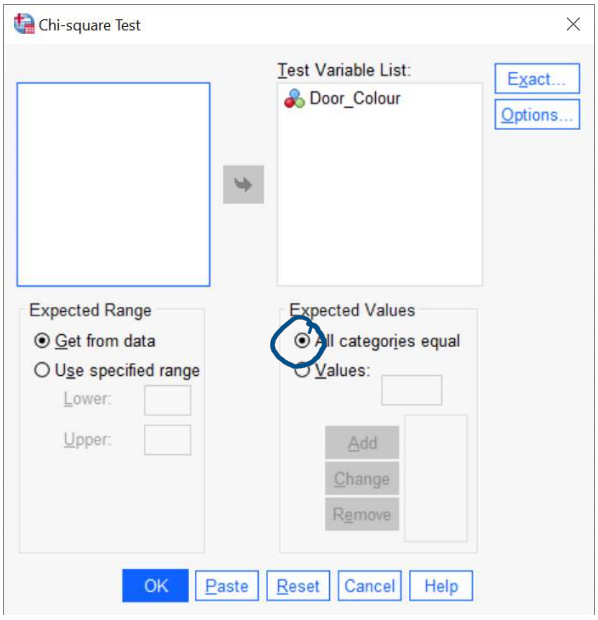
To enter the data, we switch to **Data View** tab and start entering the values. Please note, you will need to enter Green Door 20 times, Red Door 39 times and Blue Door 31 times. All together we will have 90 entries:





Obtaining the Chi-squared test for Multinomial Distribution

From the main menu on top, choose '**Analyze**' → '**Non-parametric Tests**' → '**Legacy Dialogs**' → '**Chi-square**'. You will get the following window:



Here, we select <u>equal expected values</u> (i.e. Ho: p1=p2=p3=1/3).

If, we did not expect to get equal expected values (e.g. Ho: p1=0.1, p2=0.6, p3=0.3), then we need to click on values and enter the values in an alphabetical order of the names of the categories, in this case the alphabetical order would be blue, green and red, i.e. 27, 9 and 54).

Remember, $(E_i = np_i)$

Click 'OK" and you will get following output:

Door_Colour

	Observed N	Expected N	Residual
Green	20	30.0	-10.0
Red	39	30.0	9.0
Blue	31	30.0	1.0
Total	90		

Test Statistics

	Door_Colour
Chi-Square	6.067[a]
df	2
Asymp. Sig.	.048

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 30.0.

χ^2 – statistics

p-value $< \alpha = 0,05$

∴ reject Ho

II. Test of Independence (r xc Contingency Table)

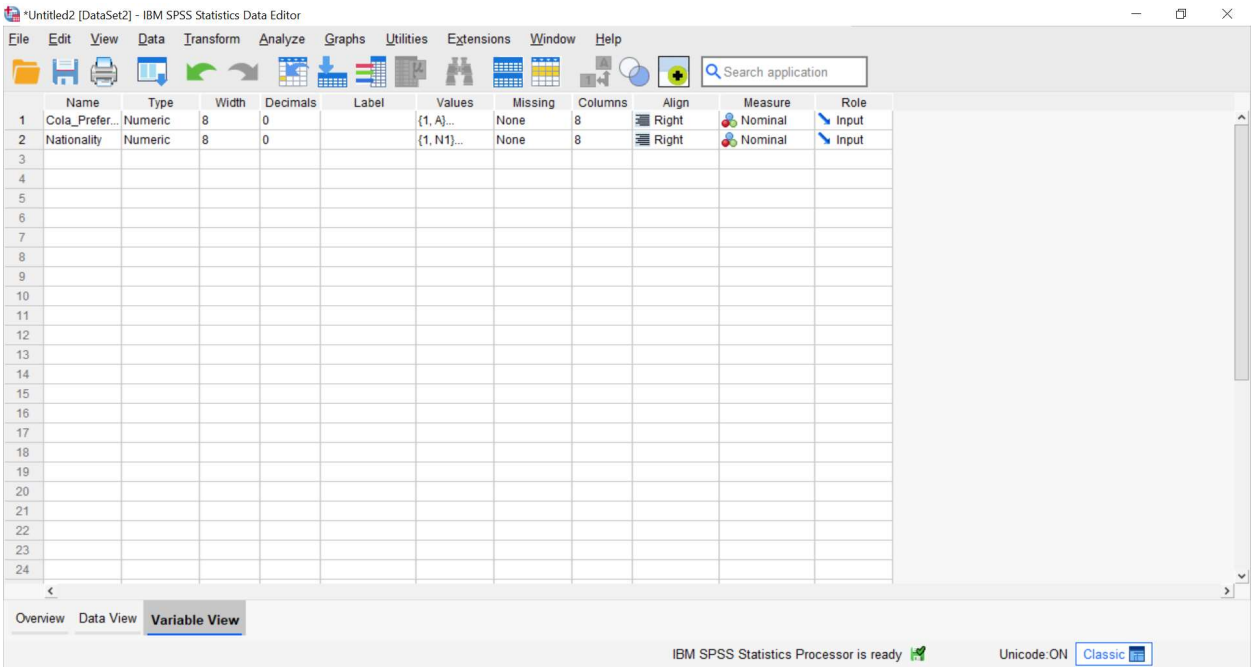
A certain cola company sells 4 types of cola throughout North America. To help determine if the same marketing approach used in the United States can be used in Canada and Mexico, one of the firm's marketing analysts wants to ascertain if there is an association between the type of cola preferred and the nationality of the consumer. A random sample of 250 cola drinkers from the 3 countries was interviewed and then classified according to the type of cola preferred and nationality. The observed frequency of drinkers falling into each of the possible cells is shown below. Is there an evidence of an association between cola preference and nationality? Use $\alpha = 0.01$.

Cola Preference

Nationality	A	B	C	D	Total
N1	72	8	12	23	115
N2	26	10	16	33	85
N3	7	10	14	19	50
Total	105	28	42	75	250

Once you open SPSS, in the **New Files** field choose **New Dataset** and click **OK**

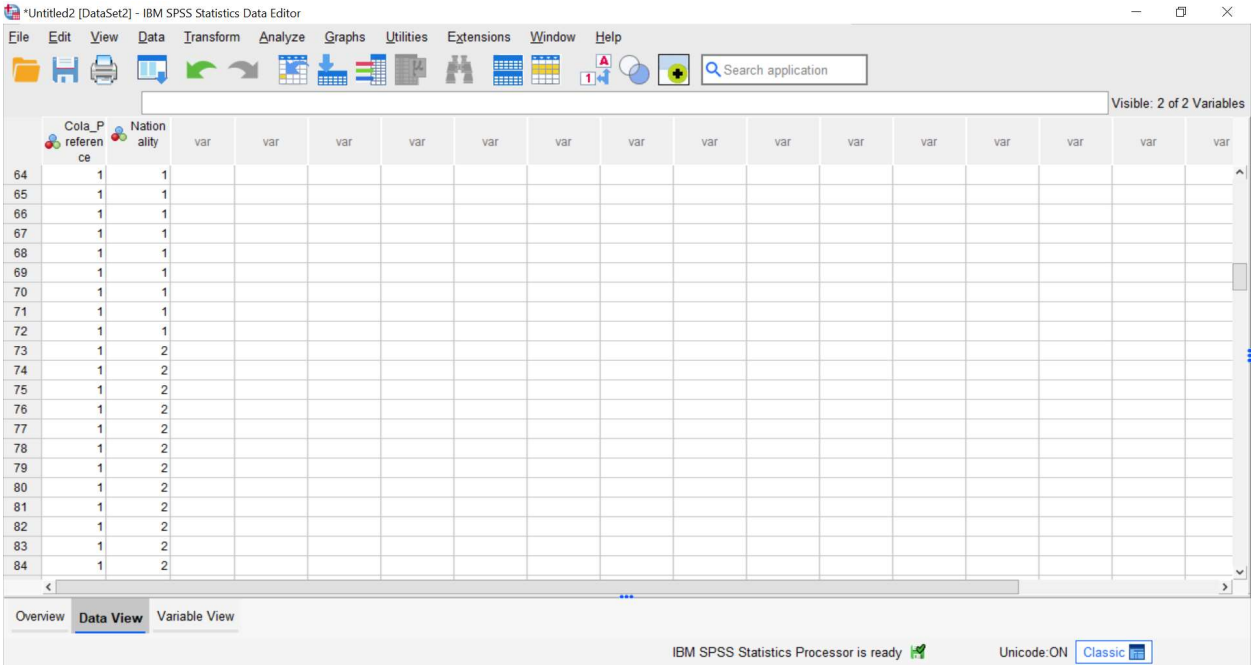
In **Variable View** create your variables:

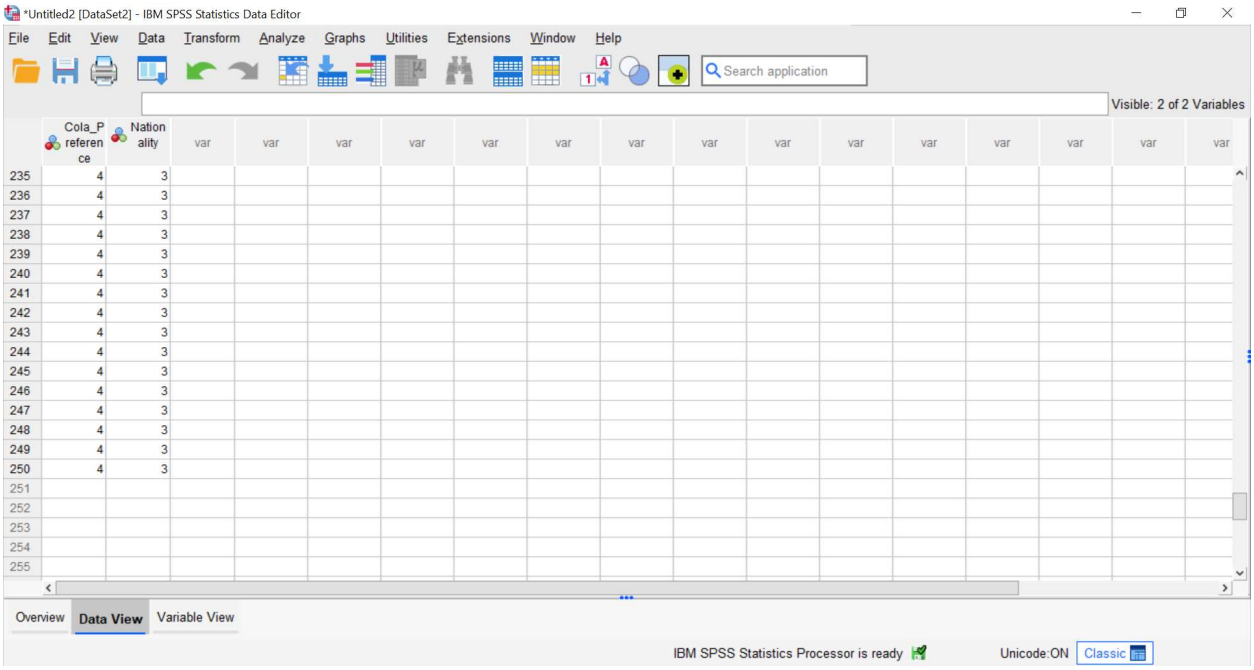


For "Cola Preference" and "Nationality" we will create values in a following pattern:

Cola A has value 1, Cola B value 2, etc. Similarly we assign values to the nationalities (see below)

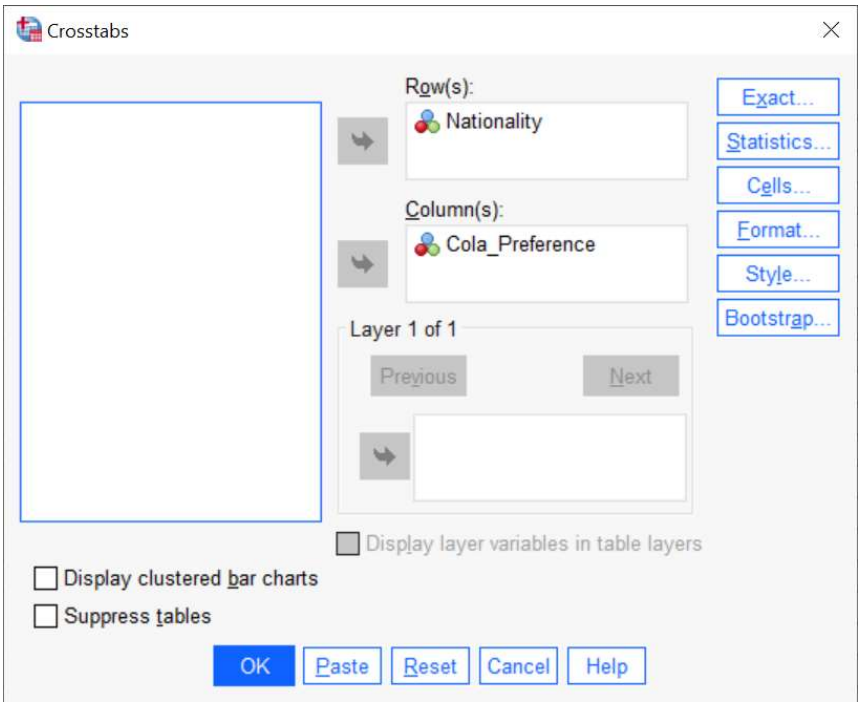
To enter the data, we switch to **Data View** tab and start entering the values. Please note, you will need to enter Cola A and Nationality N1 72 times, followed by Cola A and Nationality N2 26 times, etc. All together we will have 250 entries:



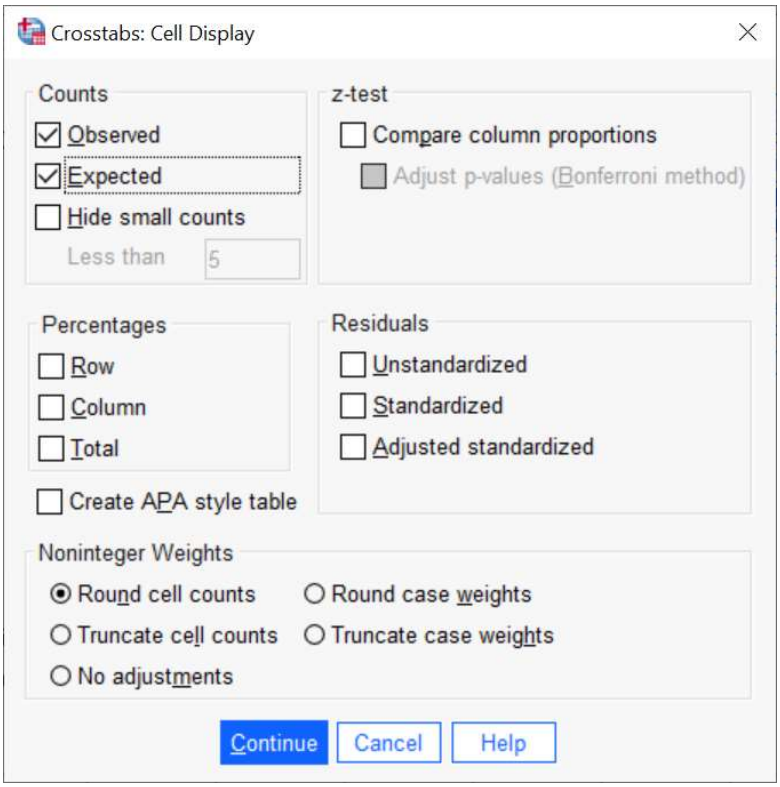


Obtaining the Chi-squared test for Independence

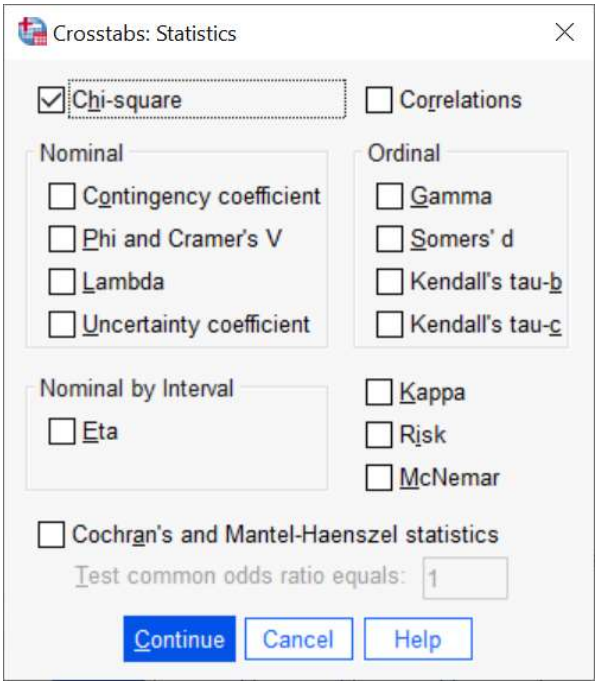
From the main menu on top, choose '**Analyze**' → '**Descriptive Statistics**' → '**Crosstabs**'. You will get the following window:



We place 'Cola Preference' and 'Nationality into Rows and Columns, then click on "**Cells**" and select both observed and expected counts, then click on 'continue'.



After click on “Statistics” and select ‘Chi-square’ and click on ‘Continue’



Then click **OK**. In the output window you will get the following outputs:

Nationality * Cola_Preference Crosstabulation

			Cola_Preference				
			A	B	C	D	Total
Nationality	N1	Count	72	8	12	23	115
		Expected Count	48.3	12.9	19.3	34.5	115.0
	N2	Count	26	10	16	33	85
		Expected Count	35.7	9.5	14.3	25.5	85.0
	N3	Count	7	10	14	19	50
		Expected Count	21.0	5.6	8.4	15.0	50.0
Total		Count	105	28	42	75	250
		Expected Count	105.0	28.0	42.0	75.0	250.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	42.748[a]	6	<.001
Likelihood Ratio	44.444	6	<.001
Linear-by-Linear Association	26.878	1	<.001
N of Valid Cases	250		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.60.