

**John MacInnes**

[An Introduction to Secondary Data Analysis with IBM SPSS Statistics](#)  
(Sage, Dec. 2017)

## 5.1 [Chapter 5 video tutorials](#) (direct link to companion website)

[NB: All video tutorials for chapter 5 are on the same web page and cannot (yet) be disaggregated]

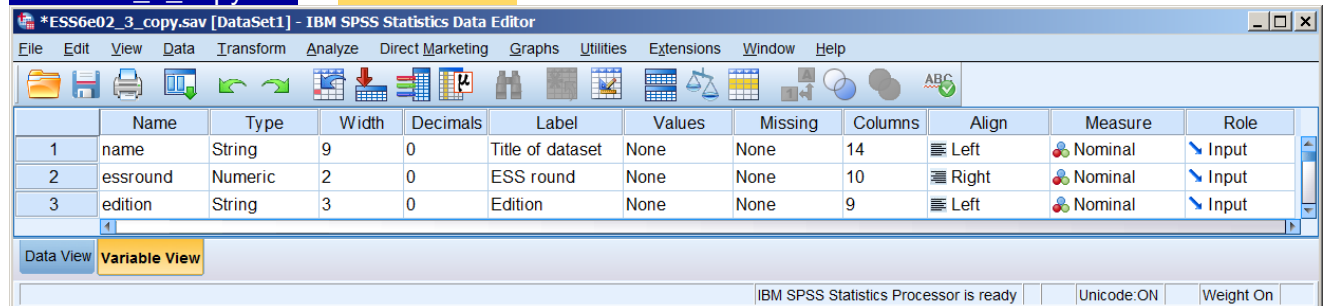
**Video tutorial 5.1.4:** Selecting cases and using the **COUNT** command (3'57")

**Task 1:** Select all countries except Albania

(**Task 2:** Create a new variable using **COUNT** is on page 4)

**Exemplar:** European Social Survey 2012  
**SPSS file:** **ESS6e02\_3\_copy.sav**<sup>1</sup>  
**SPSS command:** **FILTER**  
**Variable used:** country

**ESS6e02\_3\_copy.sav** in **Variable View**

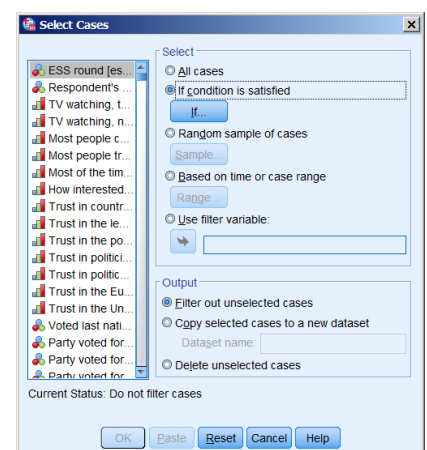
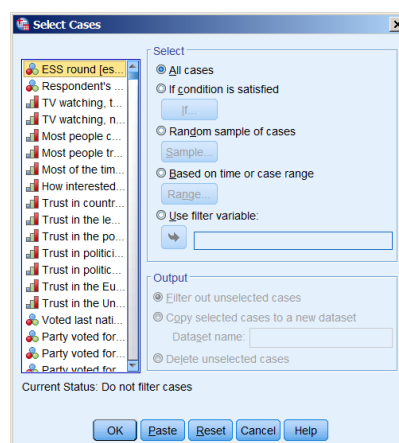
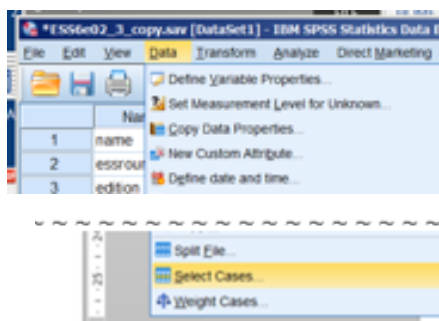
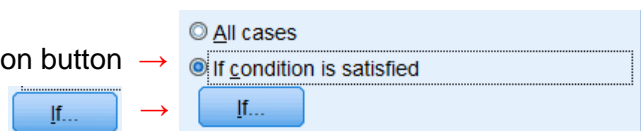


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	name	String	9	0	Title of dataset	None	None	14	Left	Nominal	Input
2	essround	Numeric	2	0	ESS round	None	None	10	Right	Nominal	Input
3	edition	String	3	0	Edition	None	None	9	Left	Nominal	Input

1: **Data** >> **Select Cases**

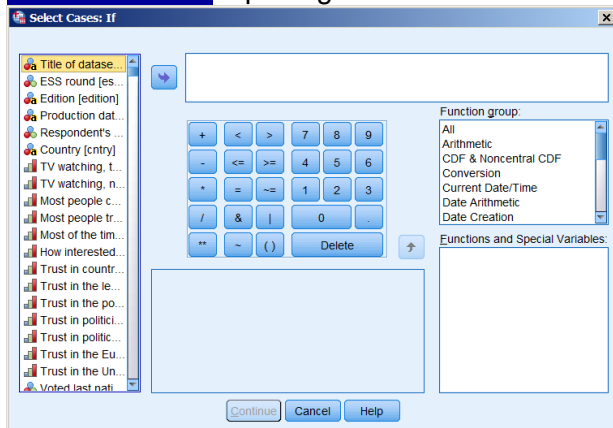
2: click on button →

3: click

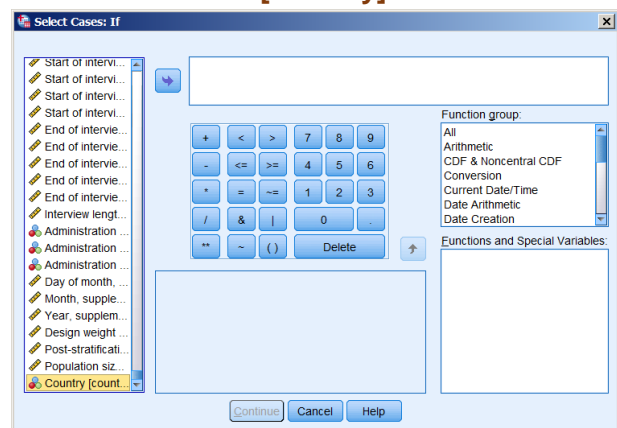



<sup>1</sup> The file used in this commentary is **ESS6e02\_3\_copy.sav**: in the video, JM uses **ESS6e02\_1.sav**, an earlier edition

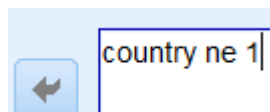
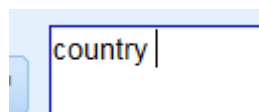
## Select Cases: If opening window



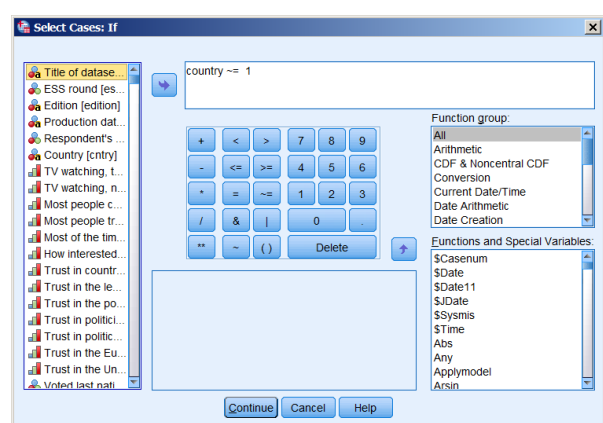
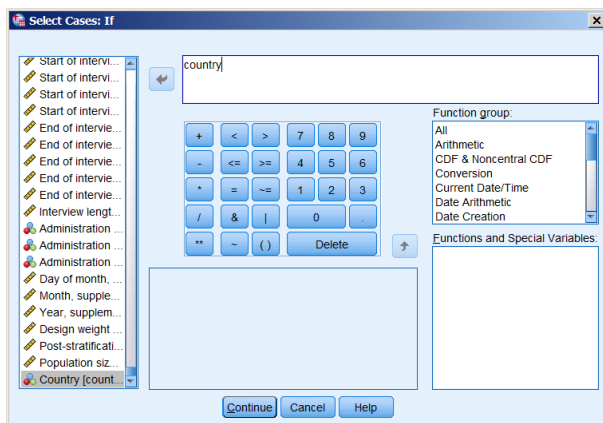
## Scroll down to find **[country]** at the end



Use  to move **[country]** to the box . . and manually add "**ne 1**" (not equal to 1)

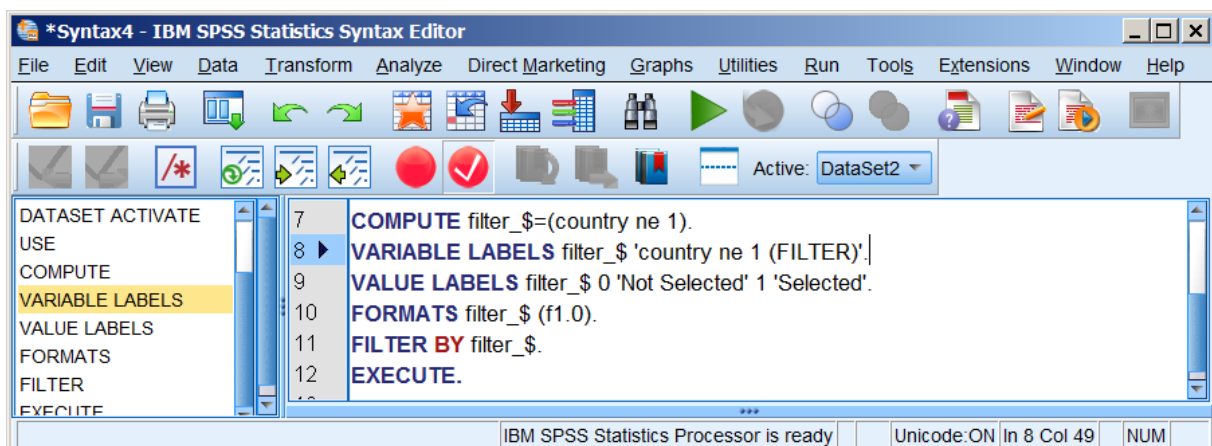


There doesn't seem to be a symbol  $\neq$  for "not equal to": if you enter **<** and **>** via the dialog box there will be two spaces in between the characters **<** **>**: the spaces need to be deleted to read **<>**.



Click on 

Click on 



[NB: This **filter** does not produce any output. **EXECUTE** forces a data pass and the filter remains in force for all subsequent commands unless SPSS encounters a **USE ALL** or **FILTER OFF** command.]

Once you have finished the current analysis, and to avoid being caught unawares by a filter lurking in SPSS, it's a good idea to switch it off<sup>2</sup> using:

**FILTER OFF .**

An alternative filter would be:

**SELECT IF**<sup>3</sup> (country **ne** 1) .

or

**SELECT IF** (country **<>** 1) .

**Warning:** this selection remains in force for all subsequent commands. To avoid inadvertently making a permanent change to the data, it's better to use a **TEMPORARY**<sup>4</sup> command first.

**temporary.**

**select if** (country **ne** 1) .

**freq** country .

or

**temporary.**

**select if** (country **<>** 1) .

**freq** country .

**Table: Countries excluding Albania**  
(Unweighted)

		country Country			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 Belgium	921	1.6	1.6	1.6
	3 Bulgaria	635	1.1	1.1	2.7
	4 Switzerland	676	1.2	1.2	3.9
	5 Cyprus	72	.1	.1	4.1
	6 Czech Republic	896	1.6	1.6	5.7
	7 Germany	7101	12.5	12.5	18.2
	8 Denmark	459	.8	.8	19.0
	9 Estonia	113	.2	.2	19.2
	10 Spain	3917	6.9	6.9	26.1
	11 Finland	451	.8	.8	26.9
	12 France	5319	9.4	9.4	36.3
	13 United Kingdom	5217	9.2	9.2	45.5
	14 Hungary	852	1.5	1.5	47.0
	15 Ireland	359	.6	.6	47.7
	16 Israel	563	1.0	1.0	48.7
	17 Iceland	25	.0	.0	48.7
	18 Italy	5229	9.2	9.2	57.9
	19 Lithuania	256	.5	.5	58.4
	20 Netherlands	1383	2.4	2.4	60.8
	21 Norway	406	.7	.7	61.6
	22 Poland	3272	5.8	5.8	67.3
	23 Portugal	898	1.6	1.6	68.9
	24 Russian Federation	12139	21.4	21.4	90.4
	25 Sweden	790	1.4	1.4	91.8
	26 Slovenia	176	.3	.3	92.1
	27 Slovakia	457	.8	.8	92.9
	28 Ukraine	3892	6.9	6.9	99.8
	29 Kosovo	135	.2	.2	100.0
	Total	56611	100.0	100.0	

For the next SPSS command, Albania will be restored to the data set

<sup>2</sup> If you don't switch the filter off the later variable **[missdep]** will display only two values (0, 1) not the full range (0 – 8).

<sup>3</sup> For an introduction to the **SELECT IF** command, see: [3.4.1 Tutorial - Conditional transformations](#)

<sup>4</sup> With **TEMPORARY**, the case selection remains in force only for the next command, after which the data revert to all cases: it has the same effect as **USE ALL**.

**Task 2:** Create a new variable using **COUNT**<sup>5</sup>

**Target variable:** missdep "Number of missing values in the 8-item depression inventory"  
**Source variables:** fltdpr flteeff slprl wrhpp fltlnl enjlf fltsd cldgng  
**Values counted:** (7, 8, 9)  
**SPSS command:** **COUNT**

**CARD 25**

*I will now read out a list of the ways you might have felt or behaved during the past week.  
 Using this card, please tell me how much of the time during the past week . . .*

	None or almost none of the time	Some of the time	Most of the time	All or almost all of the time	(Don't know)	
D5 . . . .you felt depressed?	1	2	3	4	8	[flteeff]
D6 . . . .you felt that everything you did was an effort?	1	2	3	4	8	[slprl]
D7 . . . .your sleep was restless?	1	2	3	4	8	[wrhpp]
D8 . . . .you were happy?	1	2	3	4	8	[enjlf]
D9 . . . .you felt lonely?	1	2	3	4	8	[fltlnl]
D10. . . .you enjoyed life?	1	2	3	4	8	[enjlf]
D11. . . .you felt sad?	1	2	3	4	8	[fltsd]
D12. . . .you could not get going?	1	2	3	4	8	[cldgng]

The associated variables are in rows **197 -204** of the **Data Editor**:

Name	Label
197 fltdpr	Felt depressed, how often past week
198 flteeff	Felt everything did as effort, how often past week
199 slprl	Sleep was restless, how often past week
200 wrhpp	Were happy, how often past week
201 fltlnl	Felt lonely, how often past week
202 enjlf	Enjoyed life, how often past week
203 fltsd	Felt sad, how often past week
204 cldgng	Could not get going, how often past week

**Values**

Value Labels

Value:

Label:

1 = "None or almost none of the time"  
 2 = "Some of the time"  
 3 = "Most of the time"  
 4 = "All or almost all of the time"  
 7 = "Refusal"  
 8 = "Don't know"  
 9 = "No answer"

Add Change Remove

OK Cancel Help

The example below uses **ESSe02\_3\_copy.sav**.

\*ESSe02\_3\_copy.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	name	String	9	0	Title of dataset	None	None	14	Left	Nominal	Input
2	essround	Numeric	2	0	ESS round	None	None	10	Right	Nominal	Input
3	edition	String	3	0	Edition	None	None	9	Left	Nominal	Input

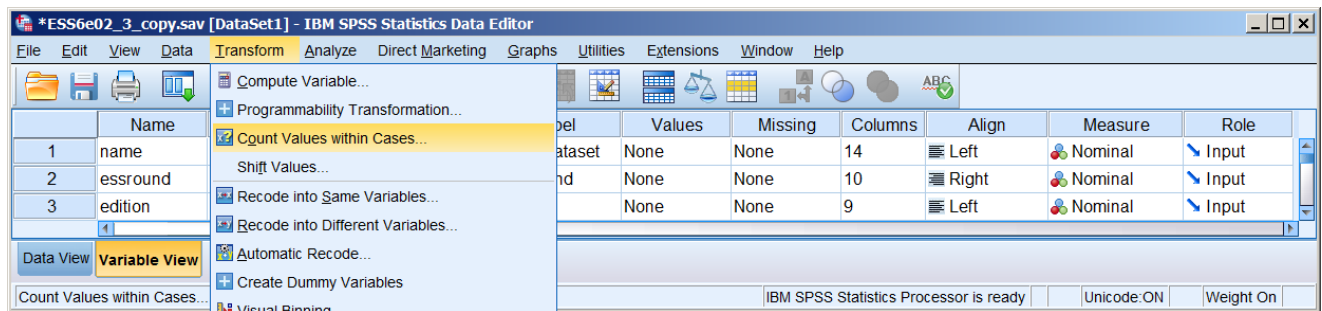
Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Weight On

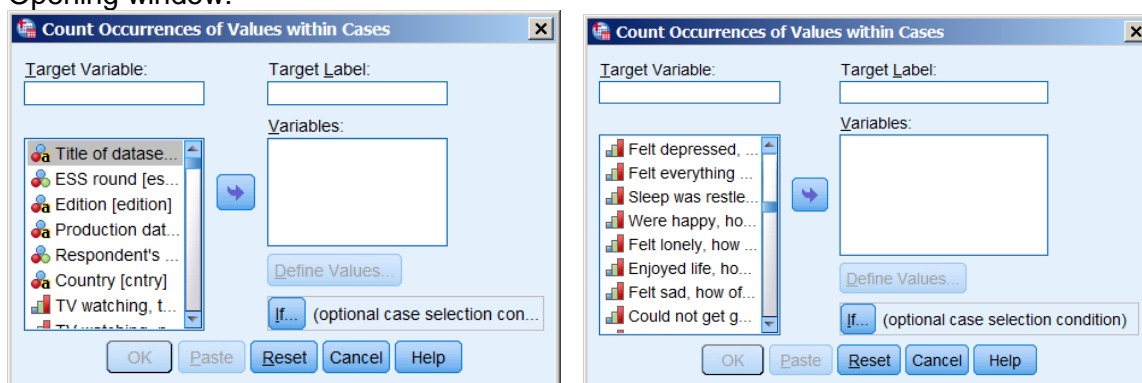
<sup>5</sup> For an introduction to the **COUNT** command, see: [3.5.1 An introduction to COUNT and COMPUTE](#)

Note that on the video JM is still using file **ESS6e02\_1.sav**. In fact his syntax has already been displayed in the video before he gets round to demonstrating how he obtained it.

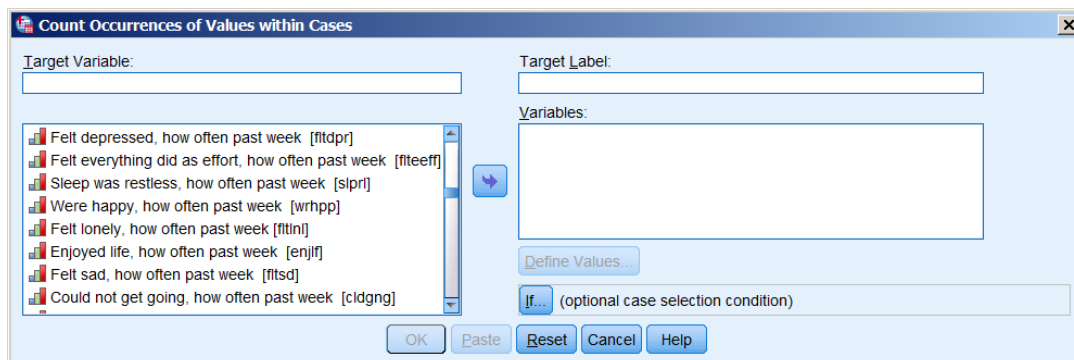
## Transform >> Count Values within Cases



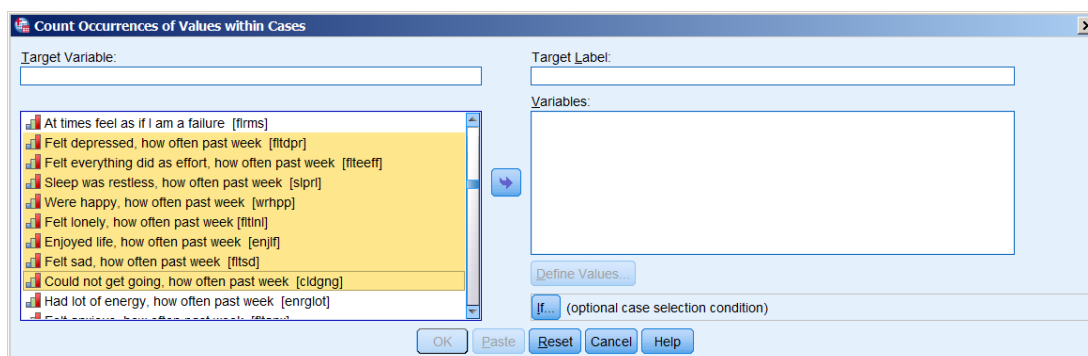
Opening window:



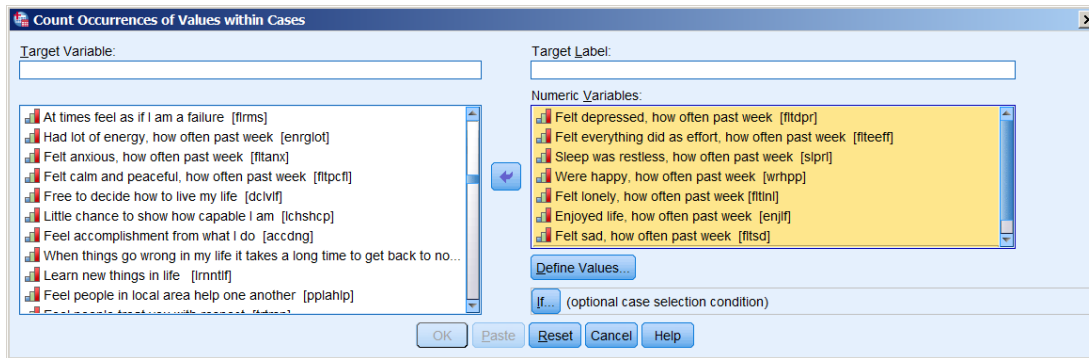
Oops! The file is set to display **Labels**. Scrolling down to find the right variables using **Names** is bad enough, but finding them with **Labels** is nigh on impossible. You can drag the right edge out to see more text:



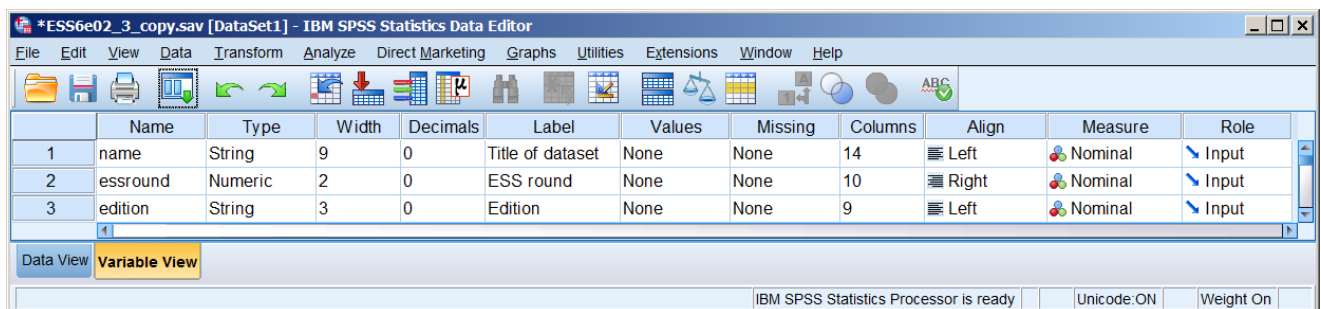
When you eventually find them, highlight the variable labels:



Click on the blue arrow  to transfer the variables to the right hand pane:



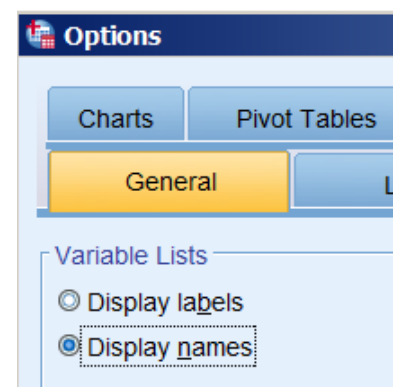
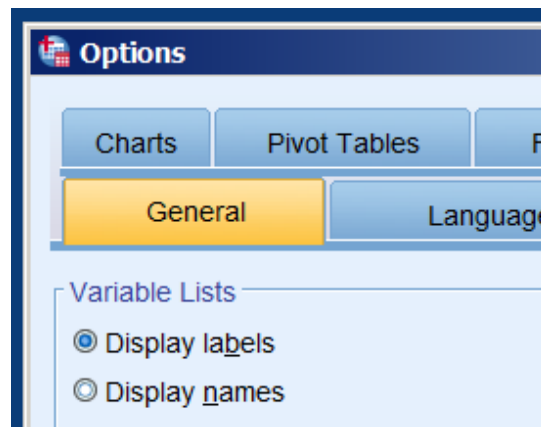
Hang on, there's an easier way to do this. To replicate this exercise more easily, users need to change the display settings back to **Names**



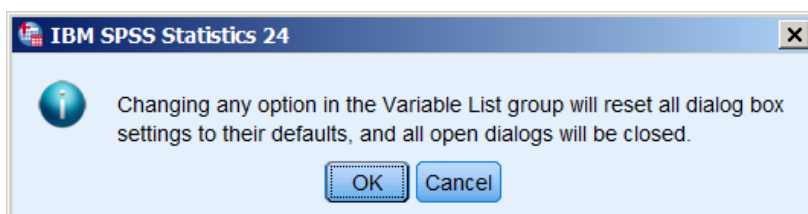
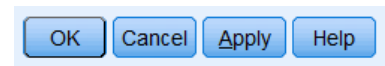
Edit >> Options

Change Display labels

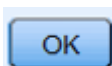
to Display names



Clicking on **Apply** gets an SPSS warning message:



Click on

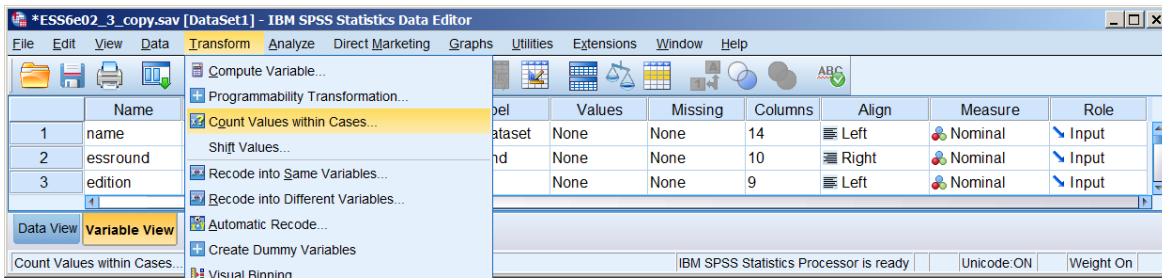


... but you can ignore it.

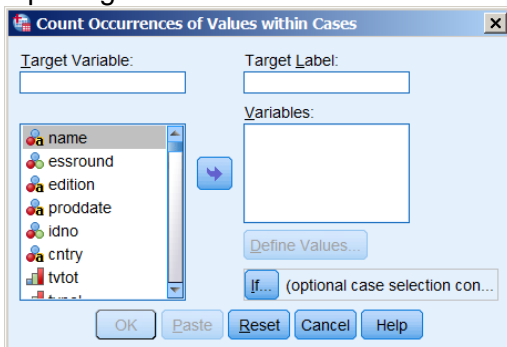


Now try again.

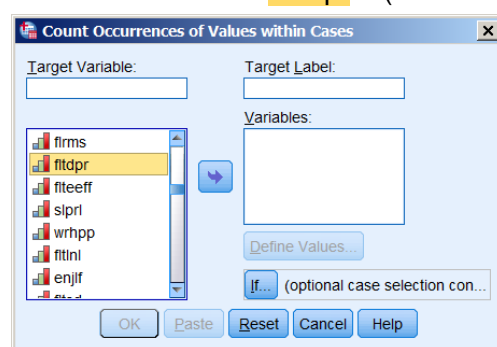
**Transform** >> **Count Values within Cases**



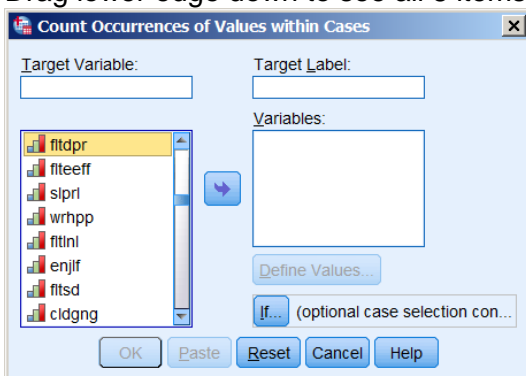
Opening window



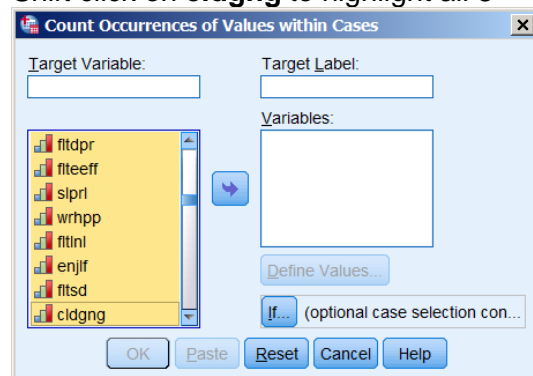
Scroll down to find **feldtpr** (it's on line 197)




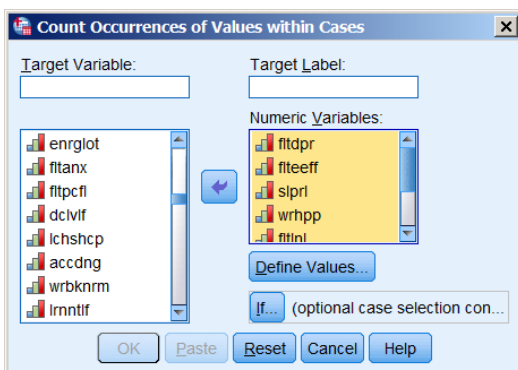
Drag lower edge down to see all 8 items



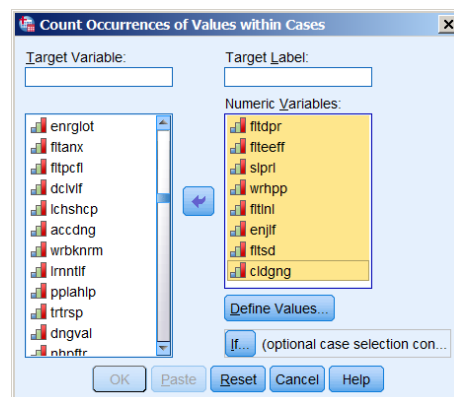
Shift click on **cldgng** to highlight all 8



Click on  to transfer all 8 to the Variables pane:



Drag lower edge down to see all 8

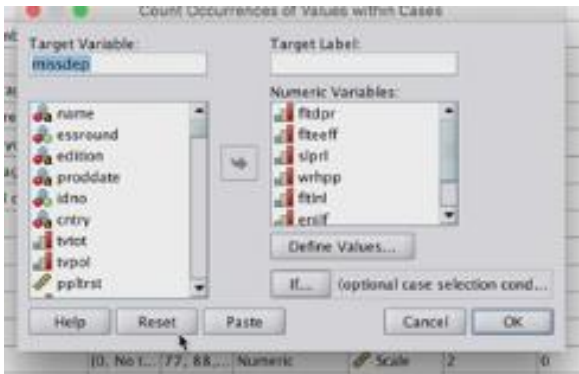


JM doesn't show the scrolling, so we don't know how long he actually takes to find the correct variables. He knows where to find them: we don't

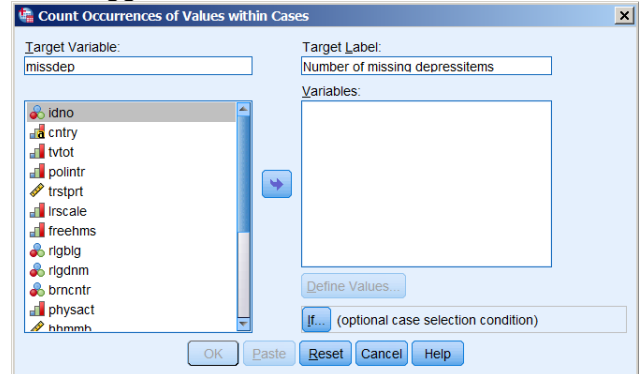
His next step, naming a **target variable**, is not shown: **missdep** is already written in the Target Variable box. JM has obviously done a trial run: he does **Reset** to clear the dialog, enters a name and label for the target variable and clicks on: **Define Values...**

He is now using variable **Names** not **Labels** and has already transferred the variables in the depression inventory to the right hand pane. However, it would make more sense, and be more logical, to choose a name and label for the target variable **before** defining the specifications.

#### JM Video

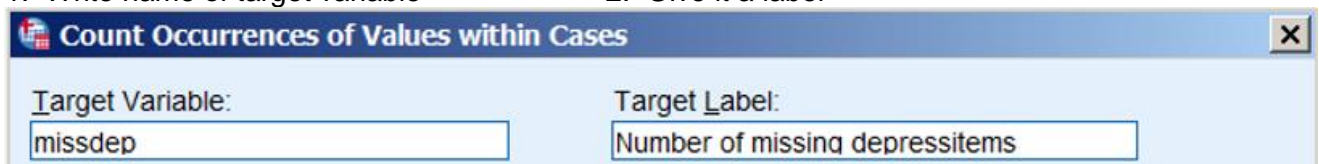



#### JFH suggestion

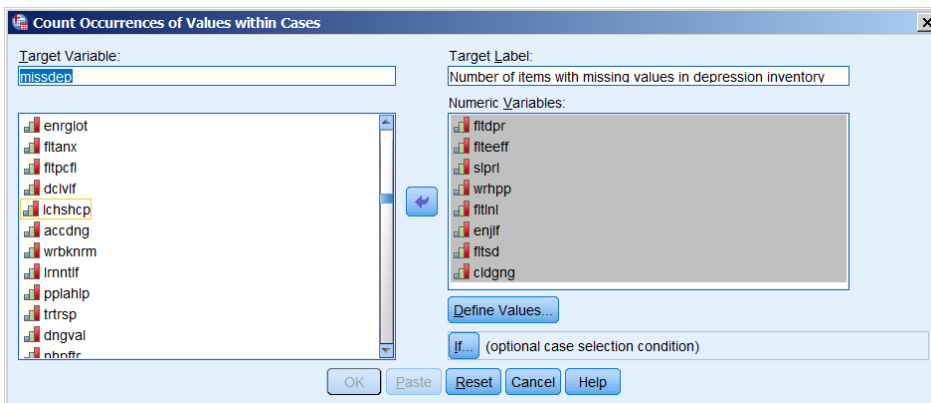


1: Write name of target variable

2: Give it a label



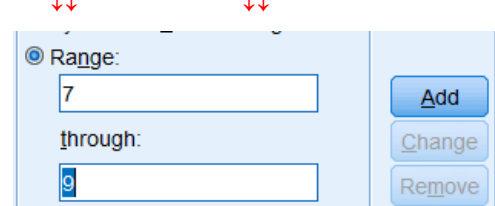
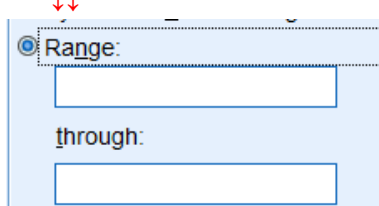
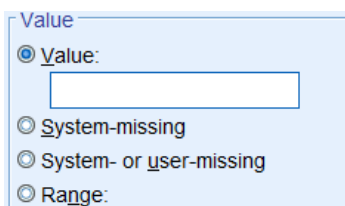
3: Find the eight depression items, highlight them and click on  to transfer them to the **Numeric Variables** box



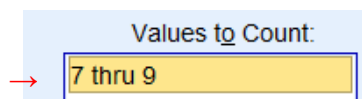
4: Click on **Define Values...**

5: Click on Range

6: Enter **7** in upper and **9** in lower box



7: Click **Add**



8: Click **Continue**



"Before I do that I need to give a name to the new variable . ." and writes **missdep** in the **Target Variable** box: he does not give it a label.

Syntax generated by [Paste](#)

```
1 DATASET ACTIVATE DataSet2.
2 COUNT missdep=fldpr flteeff slprl wrhpp fltinl enjlf fltsd cldgng(7 thru 9).
3 EXECUTE.
```

It's far quicker and easier to write syntax direct:

**COUNT** missdep = fldpr to cldgng (7,8,9) .

[NB: There's a mistake on line 21 in the syntax file on the companion website:

```
*alternative syntax.
COUNT missdep to cldgng (7 8 9).
```

should be:

```
*alternative syntax.
COUNT missdep = fldpr to cldgng (7 8 9).
```

JM doesn't actually run the command, but it would cause an error:

```
>Error # 4570 in column 15. Text: to
>On the COUNT command, the count variable was not followed by the required
>equals sign.
>Execution of this command stops.
```

There is no variable label for **[missdep]**, and weights are not applied.

For a quick check:

**freq** missdep .

Statistics		
missdep		
N	Valid	56835
	Missing	0

		missdep			Cumulative Percent
		Frequency	Percent	Valid Percent	
Valid	.00	53542	94.2	94.2	94.2
	1.00	2136	3.8	3.8	98.0
	2.00	641	1.1	1.1	99.1
	3.00	201	.4	.4	99.4
	4.00	88	.2	.2	99.6
	5.00	76	.1	.1	99.7
	6.00	26	.0	.0	99.8
	7.00	41	.1	.1	99.9
	8.00	84	.1	.1	100.0
	Total	56835	100.0	100.0	

When the **COUNT** command is executed, the new variable **missdep** appears in **Variable View** on line 628, but in the **Level** column has been classified as **Unknown** because SPSS has not yet made a data pass.

628	missdep	Numeric	8	2		None	None	10	Right	Unknown	Input
-----	---------	---------	---	---	--	------	------	----	-------	---------	-------

When **FREQUENCIES** is executed **missdep** is reclassified as **Nominal**, although it is in fact **Scale**

628	missdep	Numeric	8	2		None	None	10	Right	Nominal	Input
-----	---------	---------	---	---	--	------	------	----	-------	---------	-------

SPSS has only encountered nine values for **missdep** and has wrongly assigned its measurement level. There are also 2 superfluous decimal places. A much better specification would be:

**variable level** missdep (**scale**) .

**variable labels** missdep 'Number of items with missing values in depression inventory'.

**formats** missdep (f2.0) .

**frequencies** missdep.

**missdep Number of items with missing values in depression inventory**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	53542	94.2	94.2	94.2
	1	2136	3.8	3.8	98.0
	2	641	1.1	1.1	99.1
	3	201	.4	.4	99.4
	4	88	.2	.2	99.6
	5	76	.1	.1	99.7
	6	26	.0	.0	99.8
	7	41	.1	.1	99.9
	8	84	.1	.1	100.0
	Total	56835	100.0	100.0	

It is important to take account of missing values in statistical analysis, but all JM uses them for here is to discover that because of a CAPI problem all respondents in Albania were missing for **cldgng**. (whether they felt they "*could not get going*") It would have been far more interesting to tackle a substantive research question about the nature of depression and its constituent items, and to explore its relationship to demographic variables such as sex, age, class and education (possibly with reference to the literature and some comparison tables).

**End of:** 5.1.4 Selecting cases and using the **COUNT** command

**Back to:** [MacInnes \(2017\)](#)