3,5,5,2 Data checks 1 – Status quo

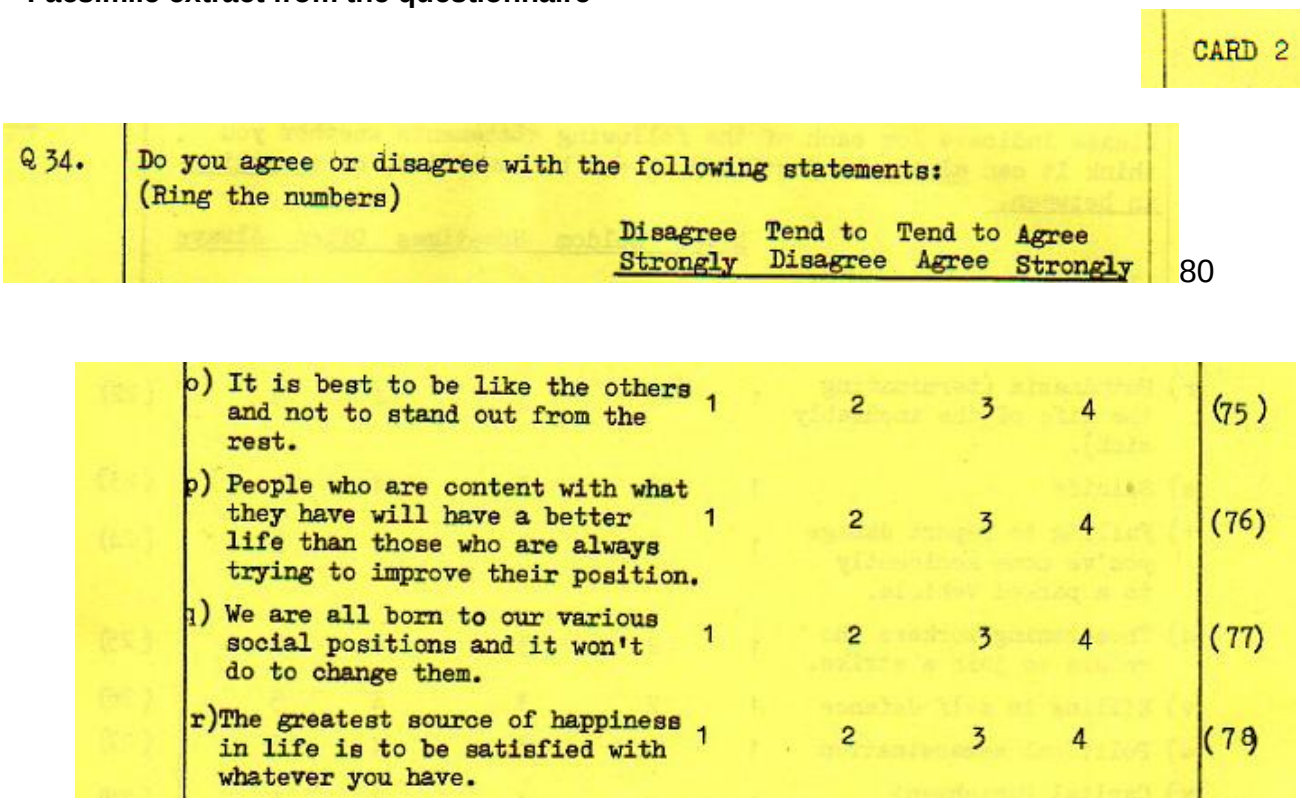
[Screenshots: SPSS15; last modified: 24 August 2011; previous version: 24 September 2009]

Exemplar: Fifth form survey fifthx.sav
Variable to be derived: Attachment to status quo
Source variables: v275 to v278 (Q34 o, p, q ,r)
Previous tutorial: 3.5.2.1 COUNT and COMPUTE - Preliminary notes

Attachment to status quo

Four statements in question Q.34 (items o, p, q ,r) are replicated from a scale developed by Himmelweit to measure "attachment to status quo" among teenagers.

Facsimile extract from the questionnaire



CARD 2

Q 34. Do you agree or disagree with the following statements: (Ring the numbers)

	Disagree Strongly	Tend to Disagree	Tend to Agree	Agree Strongly	80
o) It is best to be like the others and not to stand out from the rest.	1	2	3	4	(75)
p) People who are content with what they have will have a better life than those who are always trying to improve their position.	1	2	3	4	(76)
q) We are all born to our various social positions and it won't do to change them.	1	2	3	4	(77)
r) The greatest source of happiness in life is to be satisfied with whatever you have.	1	2	3	4	(78)

From the questionnaire extract above we can see (in the right hand margin) that the data were to be entered, one column per response, in columns 75 to 78 on card (record) 2. They were read into SPSS using the **positional**[1] variable names **v275 v276 v277** and **v278**.

[1] (See 1.3.1 Conventions for Naming Variables in SPSS for an explanation of this variable naming convention and the reasons for preferring it over the use of mnemonics).

Here is an extract from the user manual showing the original frequency counts for the responses to these items from Q.34.

34 Do you agree or disagree with the following statements: (Ring the numbers).

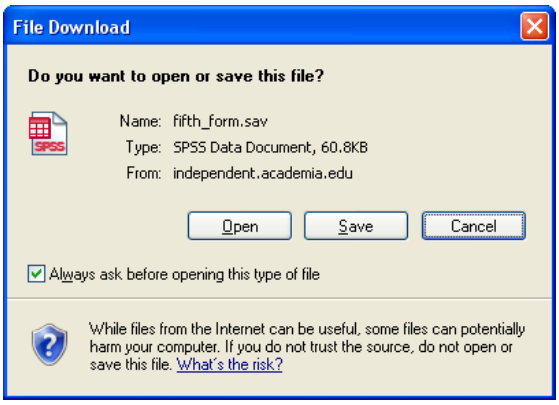
	Code:	Disagree Strongly 1	Tend to Disagree 2	Tend to Agree 3	Agree Strongly 4	Blank -1*	
o)	It is best to be like the others and not to stand out from the rest	23	60	34	1	24	(V275)
p)	People who are content with what they have will have a better life than those who are always trying to improve their position	16	52	38	12	24	(V276)
q)	We are all born to our various social positions and it won't do to change them	25	49	38	4	26	(V277)
r)	The greatest source of happiness in life is to be satisfied with whatever you have	5	15	62	40	20	(V278)

In the following exercise we are going to use the SPSS **COUNT** command to create a new, **derived variable** containing, for each pupil, a count of the **agree** or **agree strongly** responses across these four variables.

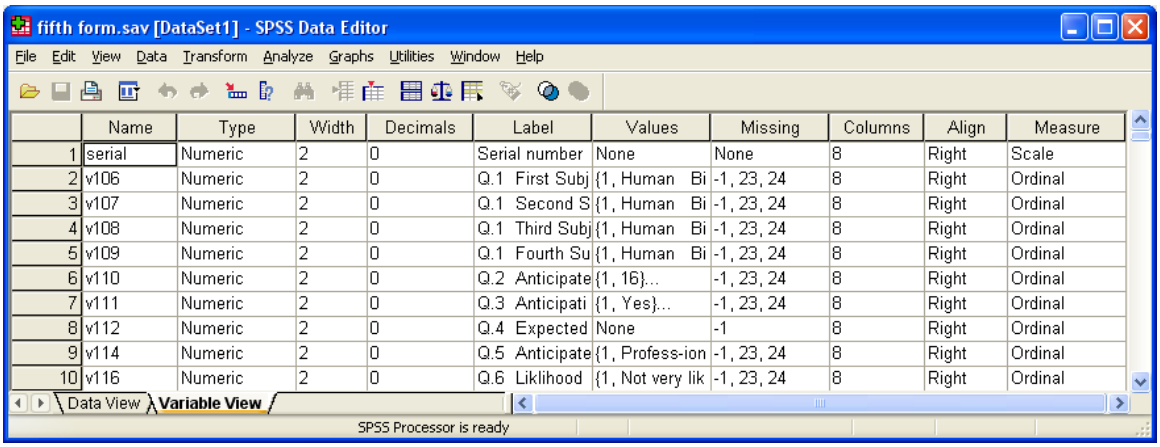
First, however, it is essential to check the original data, and a lot safer, especially in situations like this when working with batteries in which some items may need values to be reversed before generating scores.

[NB: You won't actually be able to do the next bit without immediate direct access to SPSS: don't worry just follow the rest of the tutorial anyway.**]**

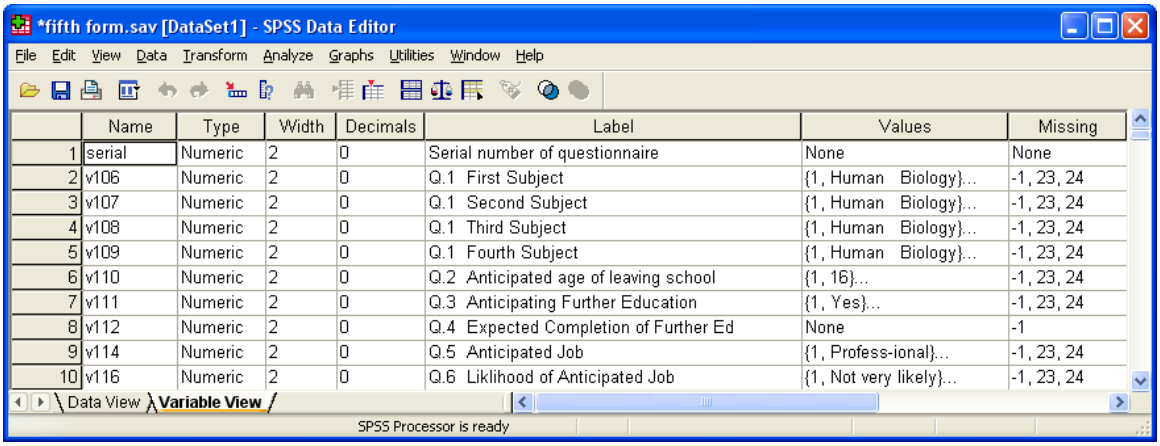
Start by downloading the SPSS saved file fifthx.sav



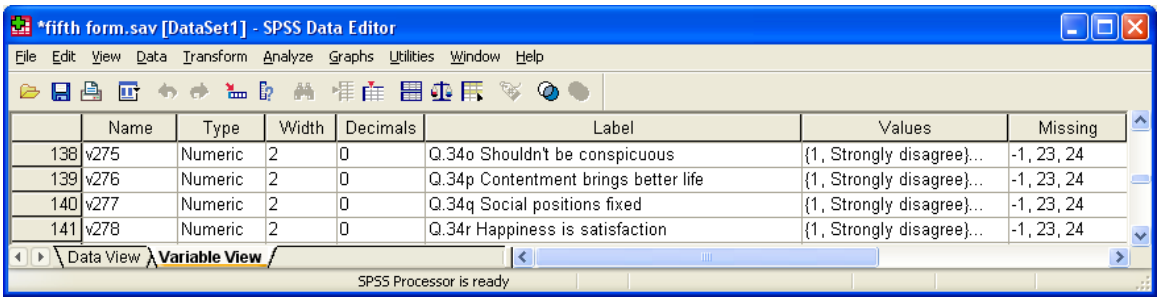
Click on **Open** to get [2]



For now we don't really need **Columns**, **Align** or **Measure**. Use the mouse to drag the column separators for **Labels** and **Values** over to the right to make the display easier to read, then adjust the size of the pane to suit.



Scroll down to find variables **v275** to **v278** containing the data for items Q34o to Q34r, then reduce the pane vertically to show only these four variables and make it easier to work with:



Now you know why positional variable names and informative variable labels are so important! They are so much easier to find in the Data Editor than mnemonics and you (and, more important, others) can work straight from the original questionnaire to the data set and *vice versa*. If you insist on mnemonics, using question numbers as variable names helps, but don't leave the full stops in (ie use **Q_34** instead of **Q.34**).

[2] The size of the Data Editor displayed will depend on how it was left when last used. You can drag the corners of the pane to reduce or enlarge it, and the edges up or down and left or right, using the left mouse button. This pane is adjusted to display only the first few variables in the file.

First, let's check the distribution of values for these four variables.

The syntax way

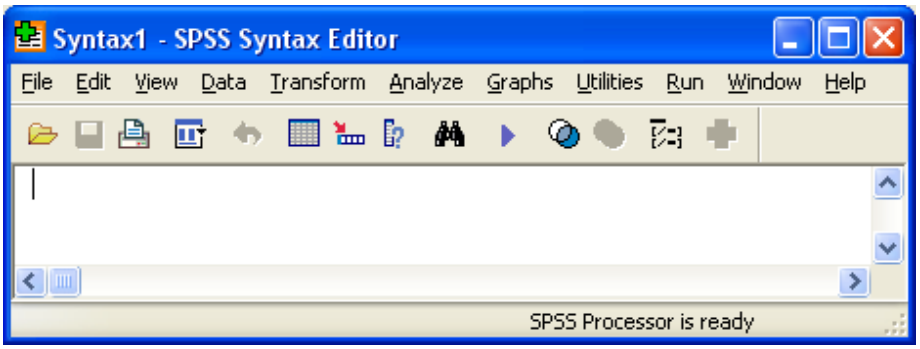
We've done frequencies before. The simplest general format is **FREQUENCIES <variable list>** so for a small job like this, we can write **FREQUENCIES v258 to v261.**

Tip: Commands and specifications are case insensitive and SPSS only reads the first 3 or 4 characters of most of them anyway. You soon get used to writing in lower case and using abbreviated commands and it makes things a lot quicker. Variable names can be typed in lower case, but will always be displayed on output in upper case.

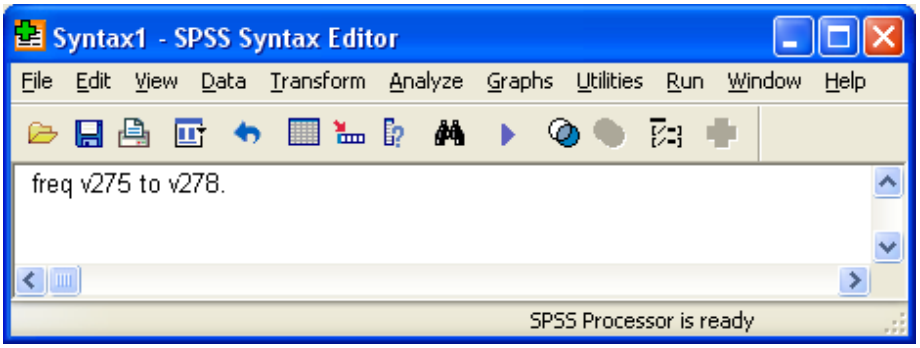
When preparing a long or complex setup file, especially for initial data loading and specification of missing values and variable and value labels (and also to save initialising SPSS) it's usually better to write out the SPSS setup file in Word first. When you're happy with it, start up SPSS, copy the text into a new syntax file and then run the job.

Open a new SPSS syntax file with:

File
…New
…Syntax



…and write **freq v275 to v278.** in the box. **Don't forget the full stop!**



Press **[CTRL]+R** to run the job.

SPSS will produce the following tables:

Statistics

		Q.34o Shouldn't be conspicuous	Q.34p Contentment brings better life	Q.34q Social positions fixed	Q.34r Happiness is satisfaction
N	Valid	118	118	116	122
	Missing	24	24	26	20

This is a summary table for all four variables showing the number of valid and missing cases for each variable. It is followed by separate frequency tables, one for each variable:

Q.34o Shouldn't be conspicuous

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	23	16.2	19.5	19.5
	Disagree	60	42.3	50.8	70.3
	Agree	34	23.9	28.8	99.2
	Agree strongly	1	.7	.8	100.0
	Total	118	83.1	100.0	
Missing	-1	24	16.9		
Total		142	100.0		

Q.34p Contentment brings better life

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	16	11.3	13.6	13.6
	Disagree	52	36.6	44.1	57.6
	Agree	38	26.8	32.2	89.8
	Agree strongly	12	8.5	10.2	100.0
	Total	118	83.1	100.0	
Missing	-1	24	16.9		
Total		142	100.0		

Q.34q Social positions fixed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	25	17.6	21.6	21.6
	Disagree	49	34.5	42.2	63.8
	Agree	38	26.8	32.8	96.6
	Agree strongly	4	2.8	3.4	100.0
	Total	116	81.7	100.0	
Missing	-1	26	18.3		
Total		142	100.0		

Q.34r Happiness is satisfaction

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	5	3.5	4.1	4.1
	Disagree	15	10.6	12.3	16.4
	Agree	62	43.7	50.8	67.2
	Agree strongly	40	28.2	32.8	100.0
	Total	122	85.9	100.0	
Missing	-1	20	14.1		
Total		142	100.0		

From these tables we can see that there are 142 cases in the file, that '**-1**' has been declared as a missing value for all four variables and that there are no values outside the range of 1 to 4.

You can also produce a single summary table using SPSS syntax, but it looks very complicated for beginners so it's easier for now to do it with the menus. Here's the syntax generated by SPSS.

```
TABLES
  /FORMAT BLANK MISSING('.')
  /TABLES  (LABELS)  BY  ( v275 + v276 + v277 + v278 )
  /STATISTICS COUNT ((F5.0) 'Count' ) .
```

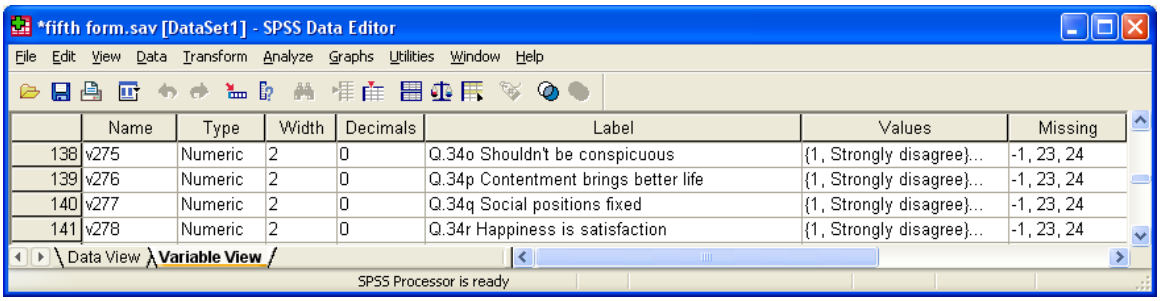
…which produces a summary table for all four variables, but no information on missing values.

	Q.34o Shouldn't be conspicuous	Q.34p Contentment brings better life	Q.34q Social positions fixed	Q.34r Happiness is satisfaction
	Count	Count	Count	Count
Strongly disagree	23	16	25	5
Disagree	60	52	49	15
Agree	34	38	38	62
Agree strongly	1	12	4	40

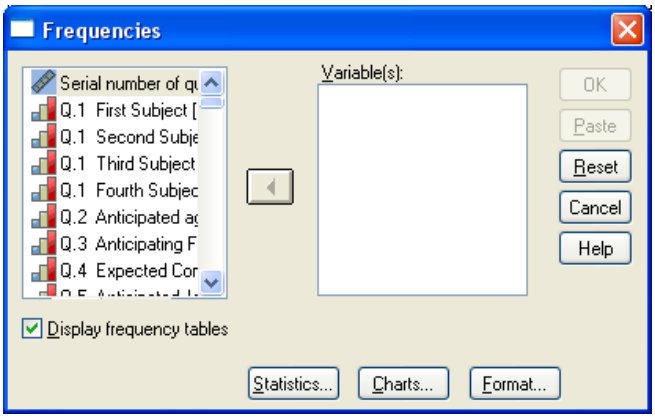
The point-and-click way

You can produce all the above tables using the drop-down menus, but this takes a lot longer and is potentially confusing and frustrating.

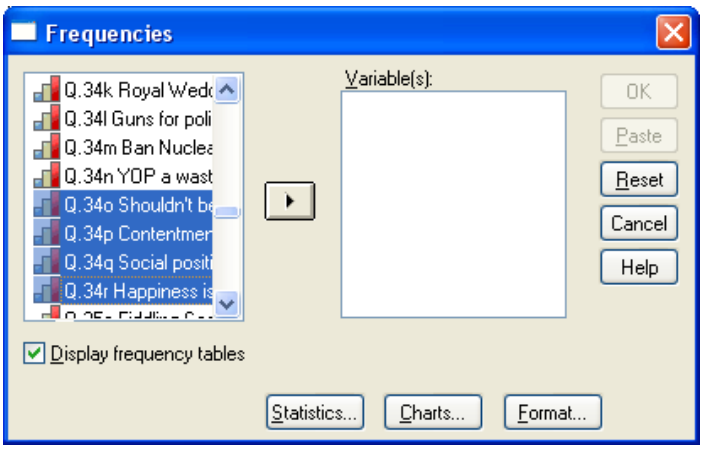
Go back to the Data Editor:



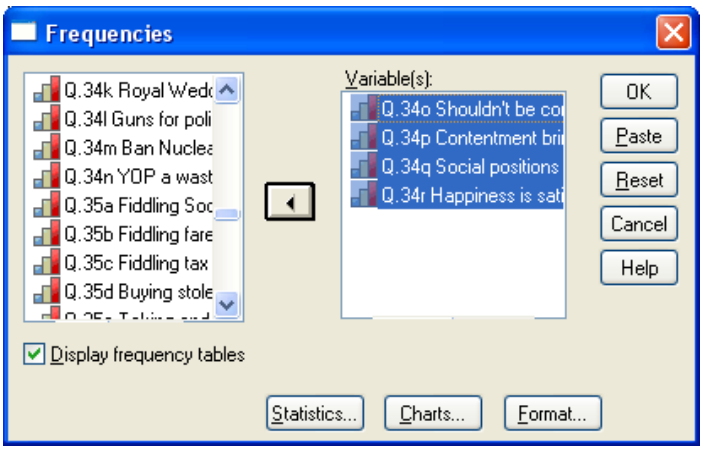
Analyze
…Descriptive statistics
…Frequencies



Scroll down to find the variables we need



… and drag them over to the **Variable(s)** box

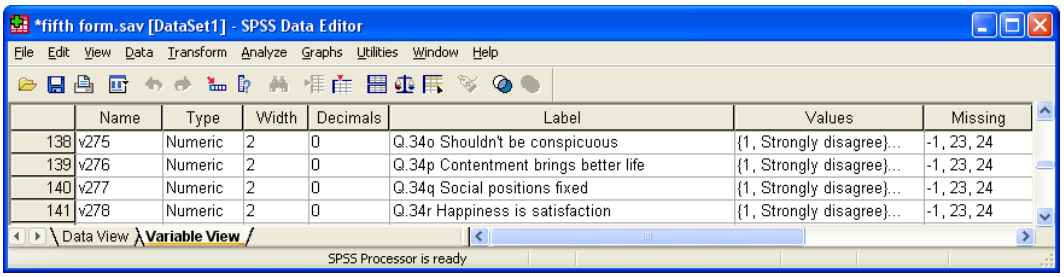


Click on **OK** to run the job.

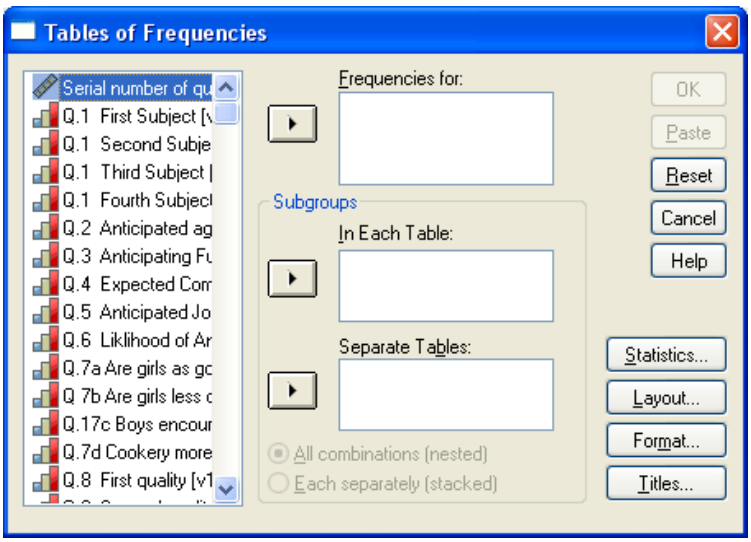
This method produces exactly the same tables as the abbreviated **freq** command in syntax mode, but generates syntax in unabbreviated form and upper case, also adding an extra line, and displays it in the SPSS Viewer.

```
FREQUENCIES
  VARIABLES=v275 v276 v277 v278
  /ORDER=  ANALYSIS .
```

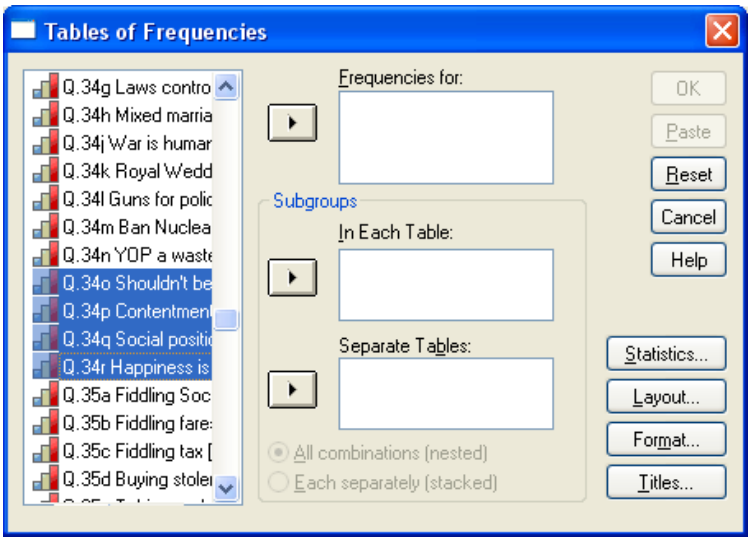
You can produce the summary table of frequencies for all four variables by:



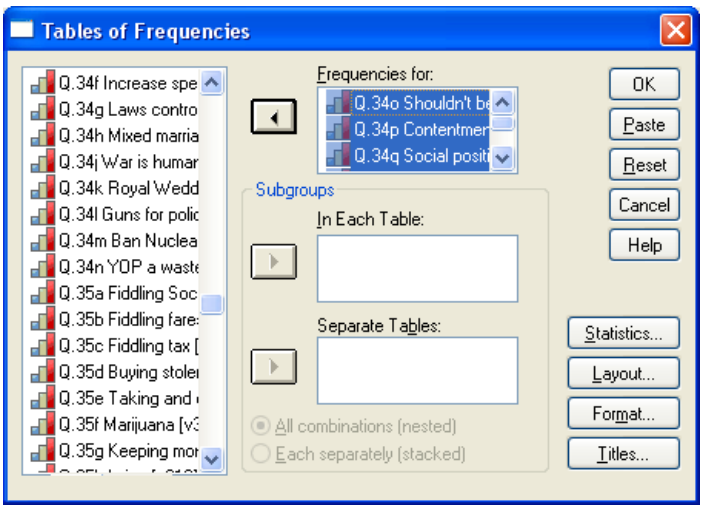
Analyze
…Tables
…Tables of frequencies



Scroll down to find items Q.34o to Q.34r



and drag across as before:



Click on **OK** to yield.

	Q.34o Shouldn't be conspicuous	Q.34p Contentment brings better life	Q.34q Social positions fixed	Q.34r Happiness is satisfaction
	Count	Count	Count	Count
Strongly disagree	23	16	25	5
Disagree	60	52	49	15
Agree	34	38	38	62
Agree strongly	1	12	4	40

For once the drop-down menus are simpler and quicker. For analytical purposes, to compare distributions, we'd be better off with % rather than n in the cells, but this time we're checking frequencies against the user manual.

By checking the frequencies for these variables against those in the user manual you will see that the variables in this data set are actually all still in their original form, but it's **essential** always to carry out such checks when creating derived variables as sometimes the original variables may have been changed or permanently recoded and any derived variables based on them may consequently be rendered meaningless.

For an instance of this actually happening, see the slide show for my repetition of the exercises from Julie Pallant's **SPSS Survival Manual** (2nd edition 2005) Exercises from Julie Pallant SPSS Survival Manual I or the full Old Dog, Old Tricks paper (pages 45-76)

Anyway, everything is OK with the data from the fifth form survey and we can begin constructing our derived variables to measure "attachment to status quo" and "sexism".

Next tutorial: 3.5.2.3 The COUNT command 1 - Attachment to status quo

Feedback on ease of understanding and use of tutorial, please, to: **johnfhall@orange.fr**