3.3 Multiple response

3.3.3.2 Analysing multiple response exercise 2 - More values than fields

Previous tutorial: 3.3.3.1 Analysing multiple response exercise 1 - One field per value

Research question: How prejudiced were the British in 1986?[1]

Exemplar: British Social Attitudes survey (1986 wave)

Raw data set: bsa86.txt[2]   (3.2 mb)

SPSS saved file   bsa86b.sav[3] (1.98 mb)

Task 1:   a) Read in raw data for (interviewer observed) ethnic group of respondent [Q.69]; whether R considers self to be prejudiced against other races; if so, which. [Q.70e and Q.70f].
  b) Produce frequency counts for Q.69, Q.70e and all three variables in Q.70f.
 c) Save your *.sps and *.sav files.

Task 2:   a) Complete the data dictionary with missing values, variable labels and value labels.
  b) Produce frequency counts for Q.69, Q.70e and all three variables in Q.70f.
  c) Save your *.sps and *.sav files again.

Task 3:   a) Define a multiple response group for the range of valid codes 01 - 10 for Q.70f
  b) Produce a grouped frequency count for Q.70f.

---

[1]   As the late Professor Sir Roger Jowell used to say, "The verb *to be prejudiced* has no 1st person singular!"
    We shall see!

[2]   Obtained under licence from UK Data Archive (UKDA) as a ***.dat** file in **Times New Roman** proportional font: reformatted to **Courier New** fixed-width font to align the columns and can be downloaded and saved with **[CTRL]+S**.

[3]   bsa86b.sav is my version of the mother file, containing data for all 3100 respondents. This version has been modified from the original (by Prof John Curtice et al at Strathclyde). The variable names have all been changed from **mnemonic** to **positional**, but most of the **variable labels** and **value labels** are still in **UPPER CASE**. However, the question number has been moved to the beginning of the variable label to make it easier to navigate the file. As I work through the tutorials I am gradually changing variable and value labels to **Mixed case**. **MISSING VALUES** have been declared, but I have had to change some of them: you should **always check your variables** before launching into any analyses. Once downloaded (In Windows 7 only) click on Save and open and this file will be opened immediately in SPSS.
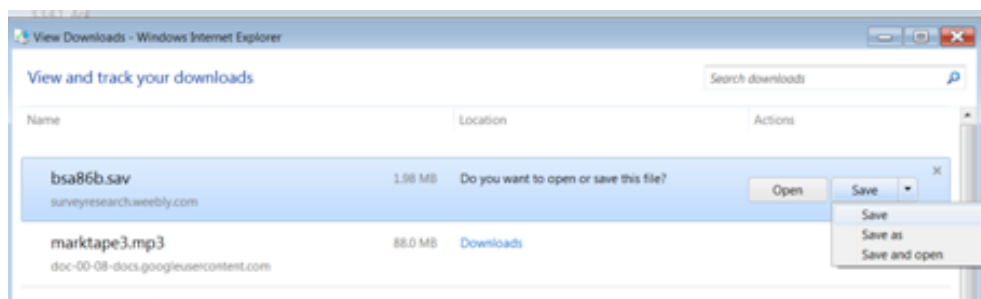
In bsa86.txt the data for a single case are spread over 23 x 80-column ASCII records, eg:

```
1020501010310721870611          020512      02
1020502210322023112111121321  31225112112211113442113  05
1020503
1020504           12  012  2004  3121
1020505                 030113121255111112 21221221   2113501  111111
1020506010205111311111  2111113102001      010400  1111122      411 2    5
1020507  5   11111833124225  2211121222222     23668656531112221230801030722
1020508222111 2      2       1 2      2       2   6        312
    09
1020510
1020511
1020512
1020513
1020514                          02  2                      2
102051502115212501                          22 011
1020516            9721    063010807327408  001010131 112112        103
102051713601080832340801981015102021      4    03  11 11310501870300486111
1020518
1020519
102052033323222343123443212222222332323321111221222122242224411211112122221 212
102052133333342312322222244133223221211111223222222221222241422244222
10205222222242442223132222
1020523  00.6666 04 4 1 02 2 4 06 2 07 08 6 5 4 4 4 4 2 1 08 04 3 6 1 1 1 5
```

Extracts from questionnaire and technical report

Q.69 Ethnic group of respondent [interviewer observed]

69.	INTERVIEWER: CODE FROM OBSERVATION FOR ALL RESPONDENTS	(570)
	White/European	1
	Indian/East African Asian/Pakistani/Bangladeshi/Sri Lankan	2
	Black/African/West Indian	3
	Other (inc. Chinese)	4

Q70e How prejudiced against other races?   [Filter question]

e)	How would you describe yourself: ... READ OUT ...	(575)	
	... as very prejudiced against people of other races,	1	f)
	a little prejudiced,	2	
	or - not prejudiced at all?	3	Q.71
	Other answer (SPECIFY) _____	7	

Q.70e: How would you describe yourself? [Single code in record 5 col 75]

a)	How would you describe yourself: ... READ OUT ...	(575)	
	... as very prejudiced against people of other races,	1	f)
	ONE CODE ONLY a little prejudiced,	2	
	8 = D/K or - not prejudiced at all?	3	Q.71
	9 = N/A		
	Other answer (SPECIFY) Recode if possible otherwise final wst with s.no	7	

2

[Open ended: post-coded, asked only if code 1 or 2 at Q70e]
Q70f Groups prejudiced against

Q.70f: Up to three verbatim replies, post-coded, 13 possible codes of which 10 valid (01 – 10) and three missing (97 - 99) but only three fields (cols 08-09, 10-11, 12-13 of record 6)

Five variables are to be read in: as usual a small table helps:

Question	Record	Column(s)	Positional name	Missing values	Label
Q.69	5	70	v570	8,9	Q.69 R's ethnic group (Interviewer observed)
Q.70e	5	75	v575	8,9	Q.70e How would you describe yourself..?
Q.70f [1st]	6	8-9	v608	96-99	Q.70f: 1st group prejudiced against
Q.70f [2nd]	6	10-11	v610	96-99	Q.70f: 2nd group prejudiced against
Q.70f [3rd]	6	12-13	v612	96-99	Q.70f: 3rd group prejudiced against

Task 1: a) Read in raw data for (interviewer observed) ethnic group of respondent [Q.69]; whether R considers self to be prejudiced against other races; if so, which. [Q.70e and Q.70f].
          b) Produce frequency counts for Q.69, Q.70e and all three variables in Q.70f.
          c) Save your *.sps and *.sav files.

**Step 1:** Go to your desktop and **o**pen folder **mybsa**:
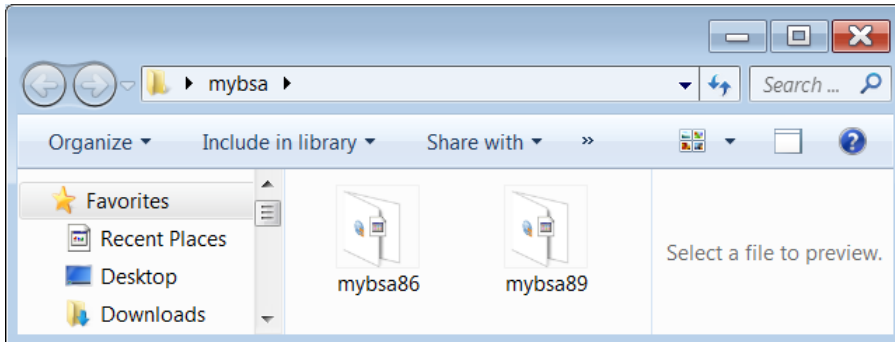

. . . and then open folder **mybsa86**:
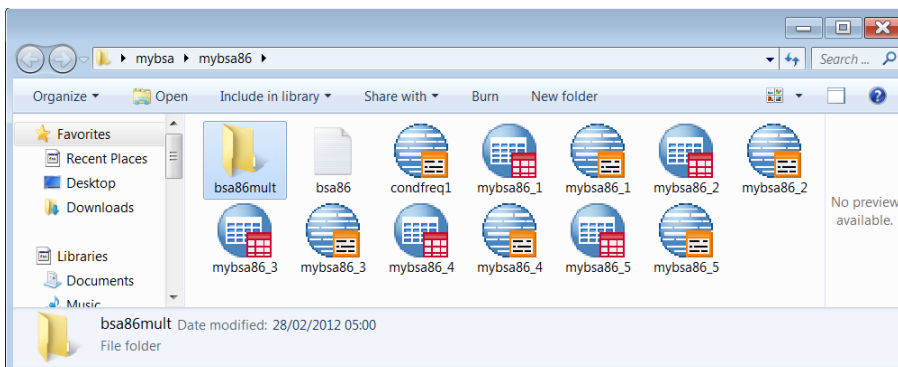
If you don't already have the raw data file **bsa86.txt** (second from left in top row) in your folder **mybsa86** you will need to copy the raw data file bsa86.txt into it from this site before you start.

You should also have a folder **bsa86mult** (top left corner) which you created in the previous tutorial to hold the files you create in this and subsequent multiple response exercises.
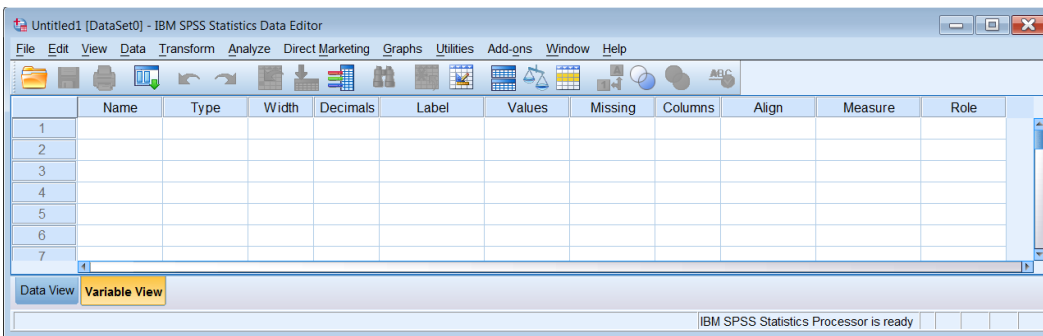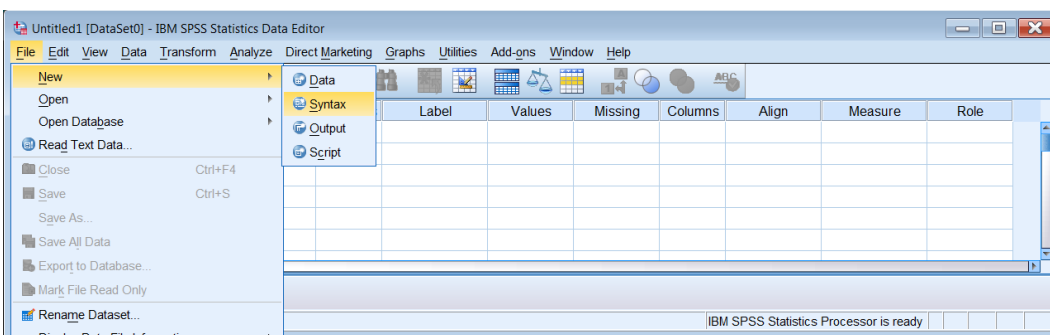
Step 2: Open SPSS
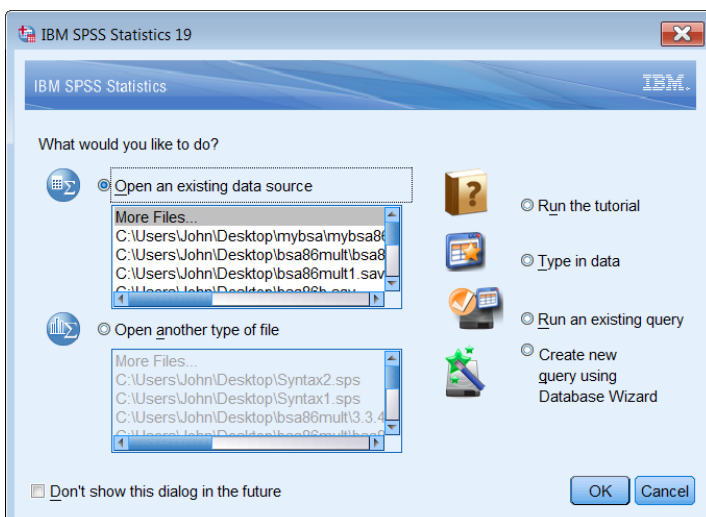
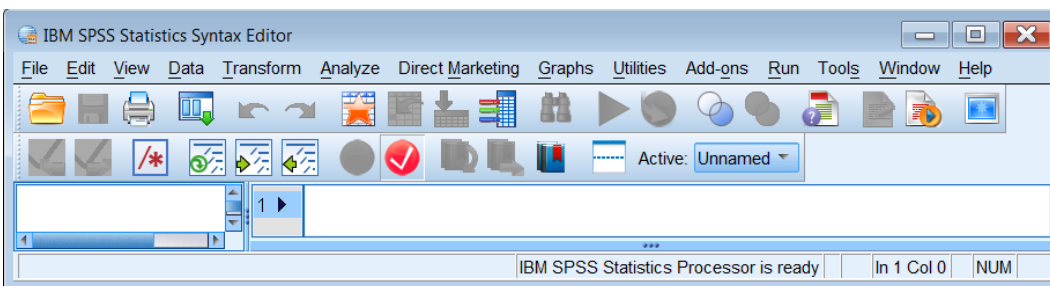Click on **Cancel** to reveal a blank data editor superimposed by a blank syntax editor:
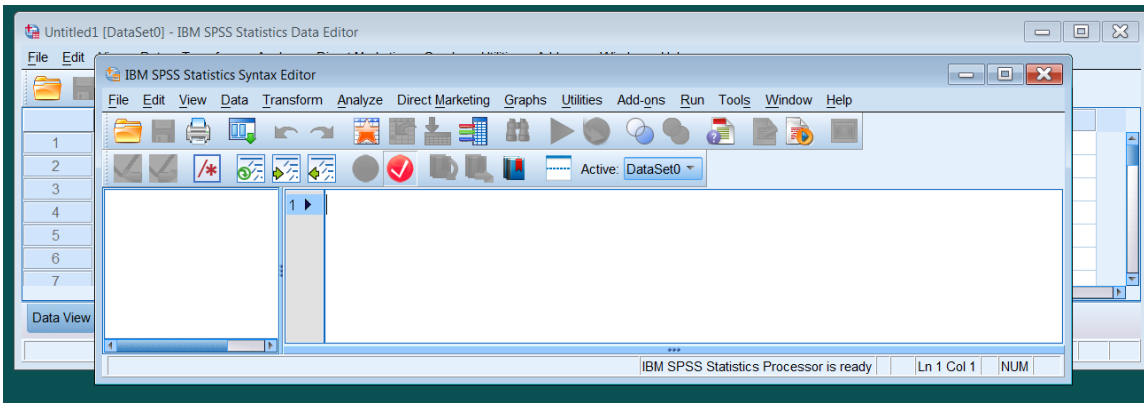
If your default SPSS settings don't automatically open the syntax editor, only a blank data editor will be displayed:

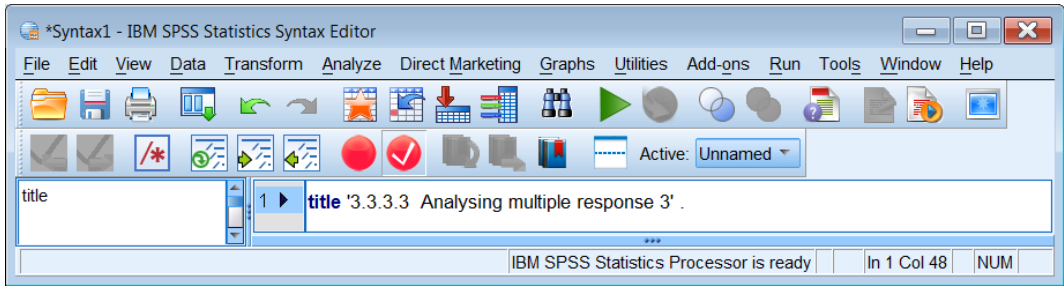Click on File **>** New **>** Syntax

. . . to open a blank syntax editor:

Step 3: Read in the raw data

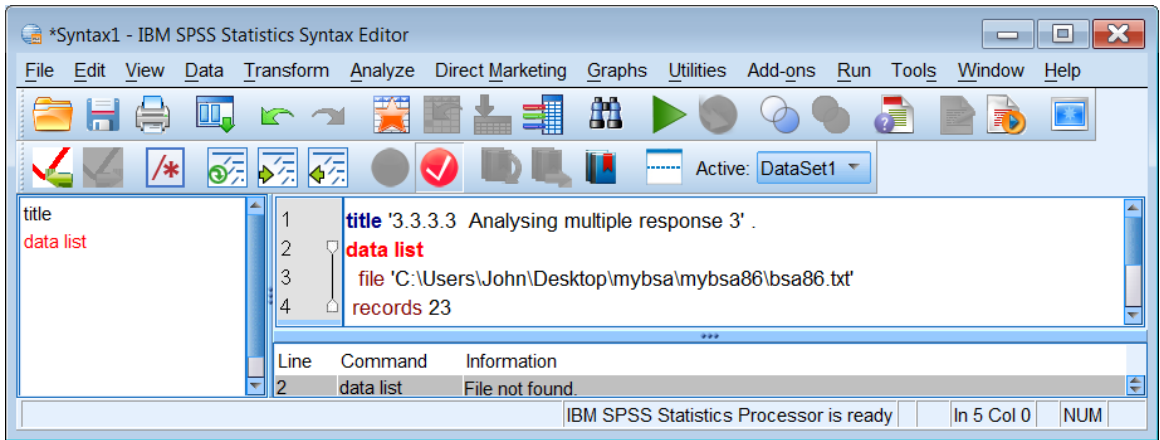Give your job a title (Don't forget the primes or the full stop.)

title '3.3.3.3 Analysing multiple response 3' .

Write out the first part of your **DATA LIST** command (inset subsequent lines by at least one space) [NB: Your pathway for the data file will be different from mine]
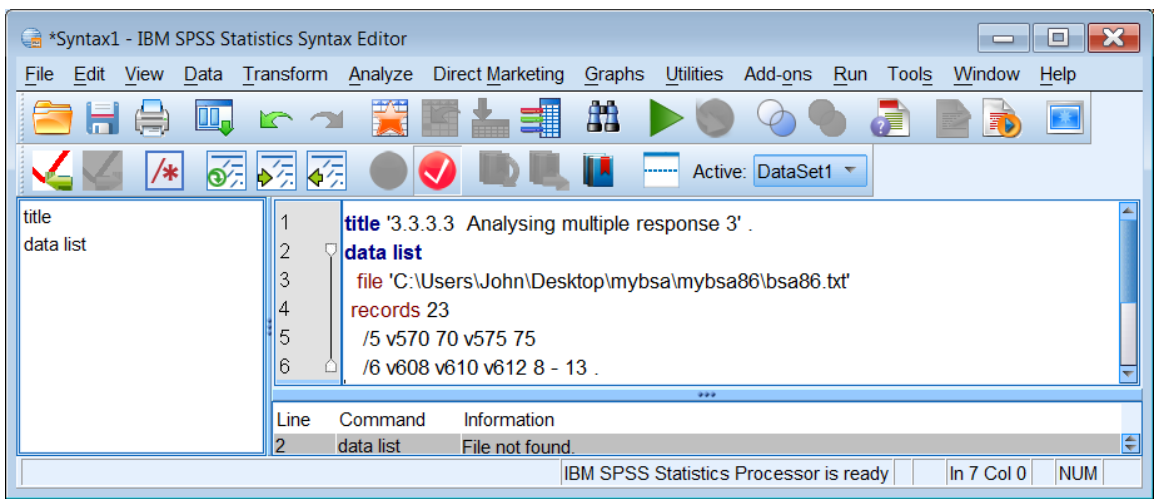
data list
      file 'C:\Users\John\Desktop\mybsa\mybsa86\bsa86.txt'
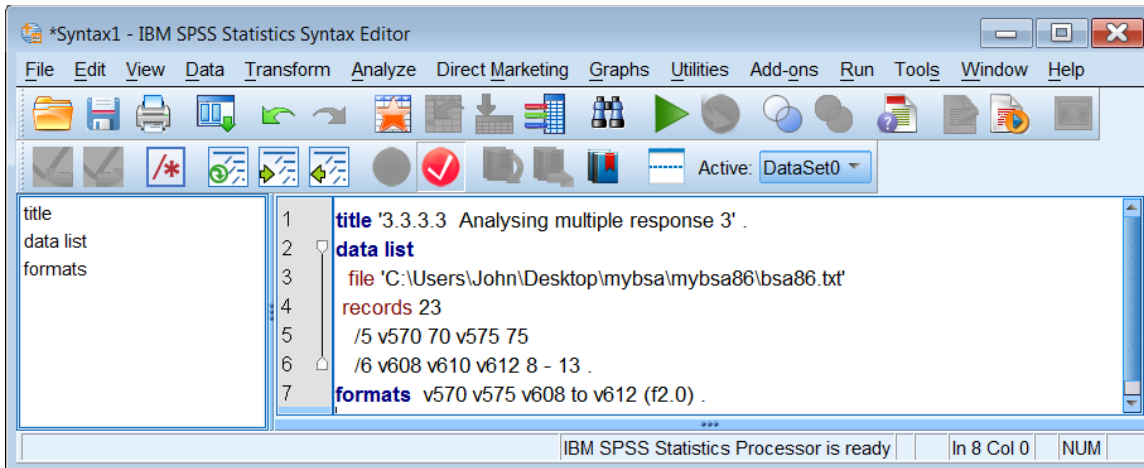   **records 23**

The **data list** command is still **red** because the command is not yet complete. Referring to the table above, carefully write out the rest of the data layout needed:

        /5 v570 70 v575 75
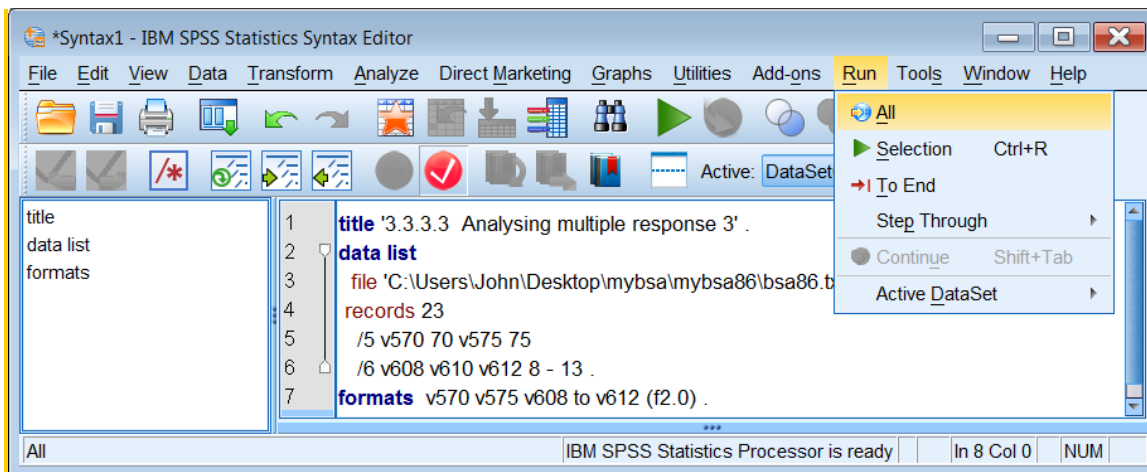   **/6 v608 v610 v612 8 - 13 .**

6

Now add a **formats** command to set all variables to type integer (no decimal points):

formats v570 v575 v608 to v612 (f2.0) .

Click on Run > All

The **DATA LIST** command produces the following output in the viewer:

```
    title '3.3.3.3  Analysing multiple response 3' .

»  3.3.3.3  Analysing multiple response 3

    data list
            file 'C:\Users\John\Desktop\mybsa\mybsa86\bsa86.txt'
            records 23
                    /5 v570 70 v575 75
                    /6 v608 v610 v612 8 - 13 .

    Data List will read 23 records from C:\Users\John\Desktop\mybsa\mybsa86\bsa86.txt

    Variable          Rec    Start     End   Format
    v570                5       70      70   F1.0
    v575                5       75      75   F1.0
    v608                6        8       9   F2.0
    v610                6       10      11   F2.0
    v612                6       12      13   F2.0
    formats   v570 v575 v608 to v612 (f2.0) .
```

7

As you can see, using **positional** rather than **mnemonic** variable names makes it much easier to check that the data for each variable have been read from the correct record(s) and column(s).
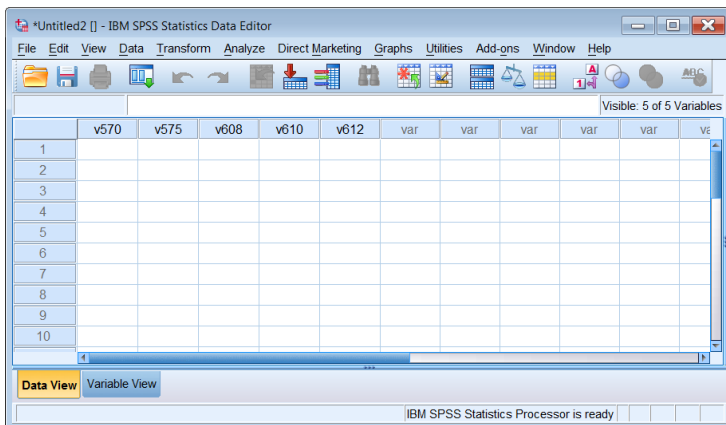
The data editor in Variable View has now filled up:
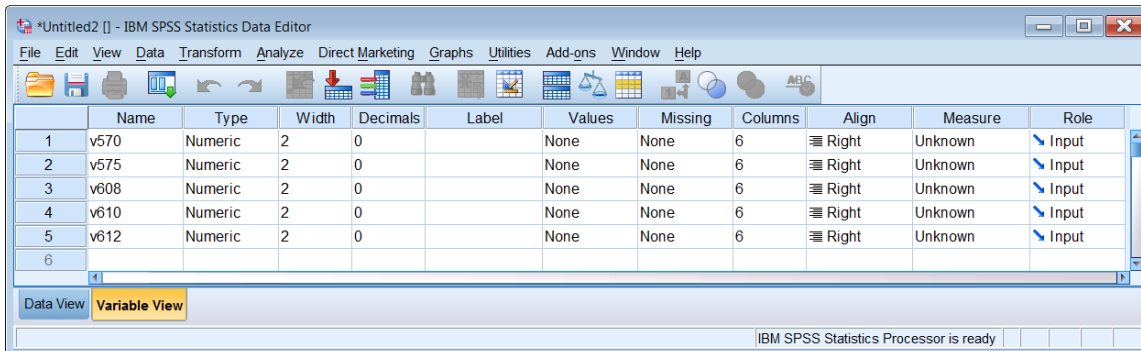
The title of the data editor has changed from *Untitled1 to *Untitled 2 (one of SPSS's little quirks). Levels of measurement are **Unknown** because the syntax so far is only a specification: SPSS has not yet made a pass through the data and the data editor in Data View will be empty.

Step 4: Check your data

As a first check, **LIST** the contents of the first 10 cases. [Command not available from menus]

list   v570 to v612 / cases 10 .
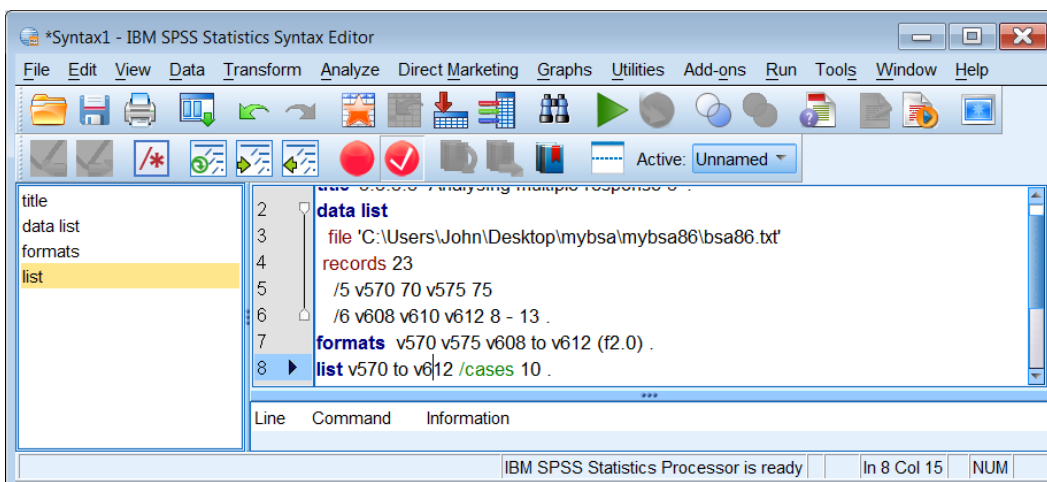
Click on ▶ to get a listing of the contents of the first 10 cases:

```
v570 v575 v608 v610 v612

   1    1    3    .    .
   1    3    .    .    .
   1    1    1    2    5
   1    3    .    .    .
   1    3    .    .    .
   1    3    .    .    .
   1    2   10    .    .
   1    2    9    .    .
   1    3    .    .    .
   1    3    .    .    .


Number of cases read:  10    Number of cases listed:  10
```
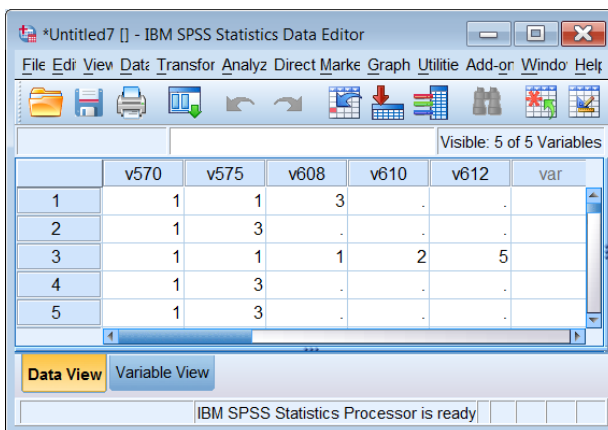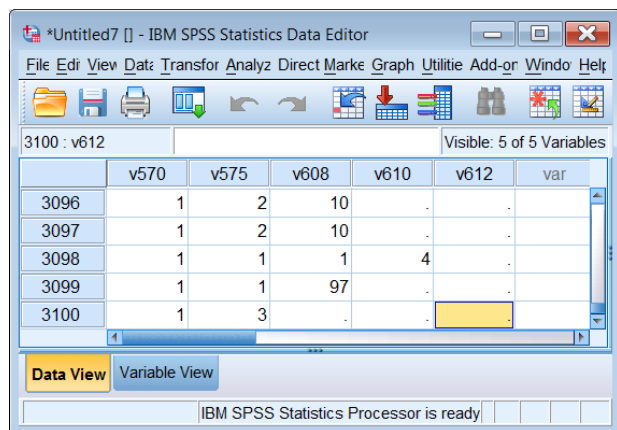
The **list** command forces a pass through the data, so the data editor has now filled up in Data View

	v570	v575	v608	v610	v612	var
1	1	1	3	.	.	
2	1	3		.	.	
3	1	1	1	2	5	
4	1	3		.	.	
5	1	3		.	.	
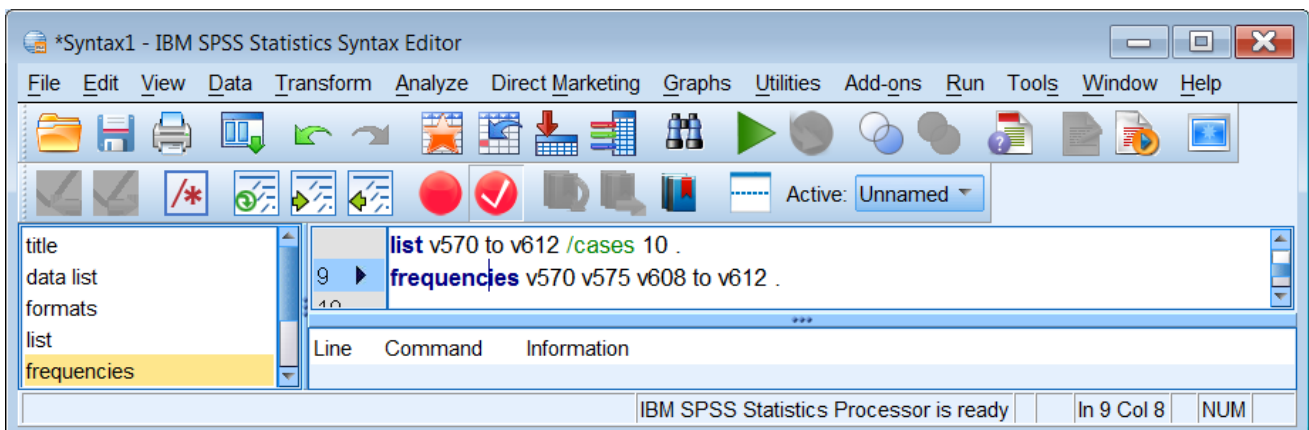
3100 : v612

	v570	v575	v608	v610	v612	var
3096	1	2	10	.	.	
3097	1	2	10	.	.	
3098	1	1	1	4		
3099	1	1	97			
3100	1	3	.			

As a quick check you can scroll up and down to ensure that respondents coded 3 ("Not at all prejudiced") in **v575** don't have any data in **v608** to **v612** (Groups prejudiced against)'

Step 5: Check your data again, this time to check which values are present for each variable.

frequencies   v570 v575 v608 to v612 .

*Syntax1 - IBM SPSS Statistics Syntax Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Run Tools Window Help

Active: Unnamed

```
title
data list
formats
list
frequencies
```

```
9  ▶   list v570 to v612 /cases 10 .
       frequencies v570 v575 v608 to v612 .
10
```

Line   Command   Information

IBM SPSS Statistics Processor is ready   In 9 Col 8   NUM

Click on ▶

The **FREQUENCIES** command is repeated in the viewer:

```
frequencies v570 v575 v608 to v612 .
```

. . and the following tables are produced

Statistics

		v570	v575	v608	v610	v612
N	Valid	3100	3100	1115	159	36
	Missing	0	0	1985	2941	3064

v570

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2976	96.0	96.0	96.0
	2	55	1.8	1.8	97.8
	3	49	1.6	1.6	99.4
	4	12	.4	.4	99.7
	9	8	.3	.3	100.0
	Total	3100	100.0	100.0	

v575

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	116	3.7	3.7	3.7
	2	999	32.2	32.2	36.0
	3	1953	63.0	63.0	99.0
	7	20	.6	.6	99.6
	8	4	.1	.1	99.7
	9	8	.3	.3	100.0
	Total	3100	100.0	100.0	

v608

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	153	4.9	13.7	13.7
	2	55	1.8	4.9	18.7
	3	44	1.4	3.9	22.6
	4	94	3.0	8.4	31.0
	5	22	.7	2.0	33.0
	6	1	.0	.1	33.1
	7	68	2.2	6.1	39.2
	8	11	.4	1.0	40.2
	9	27	.9	2.4	42.6
	10	567	18.3	50.9	93.5
	97	64	2.1	5.7	99.2
	98	2	.1	.2	99.4
	99	7	.2	.6	100.0
	Total	1115	36.0	100.0	
Missing	System	1985	64.0		
Total		3100	100.0		

v610

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	16	.5	10.1	10.1
	2	25	.8	15.7	25.8
	3	7	.2	4.4	30.2
	4	30	1.0	18.9	49.1
	5	25	.8	15.7	64.8
	6	3	.1	1.9	66.7
	7	26	.8	16.4	83.0
	8	13	.4	8.2	91.2
	9	7	.2	4.4	95.6
	10	4	.1	2.5	98.1
	97	3	.1	1.9	100.0
	Total	159	5.1	100.0	
Missing	System	2941	94.9		
Total		3100	100.0		

v612

|		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	4	.1	11.1	11.1
	2	1	.0	2.8	13.9
	3	2	.1	5.6	19.4
	4	5	.2	13.9	33.3
	5	4	.1	11.1	44.4
	7	4	.1	11.1	55.6
	8	4	.1	11.1	66.7
	9	10	.3	27.8	94.4
	10	2	.1	5.6	100.0
	Total	36	1.2	100.0	
Missing	System	3064	98.8		
Total		3100	100.0		

All variables have values in the expected ranges, so it's **good practice** to save the first edition of your work. The most important is the syntax editor as you can always run it again to generate the data editor. If you lose the syntax editor, you'll have to start again from the beginning!
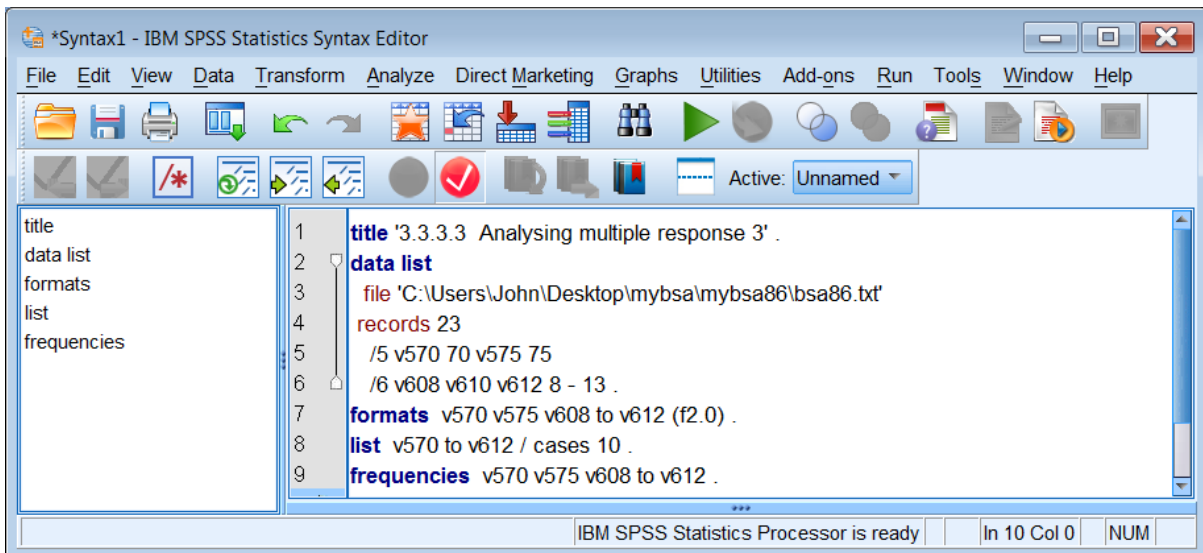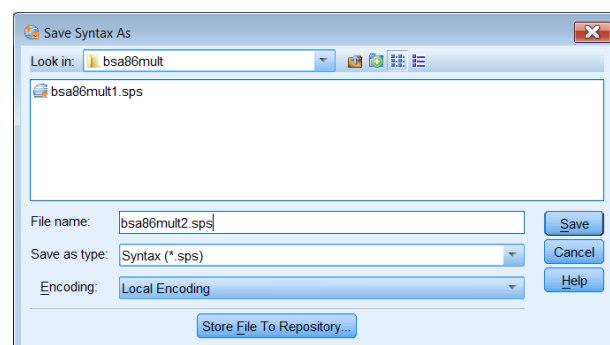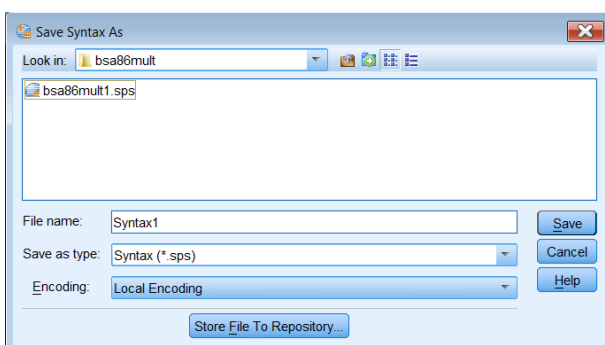

Click on **File > Save As ...**
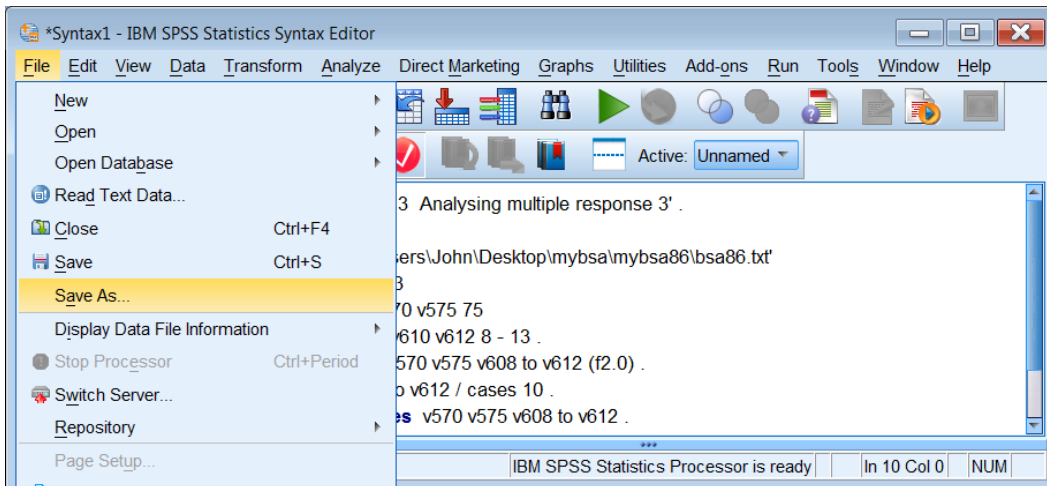
SPSS defaults to **Documents**, but we need folder **mybsa86mult**, so a bit of navigation is required:
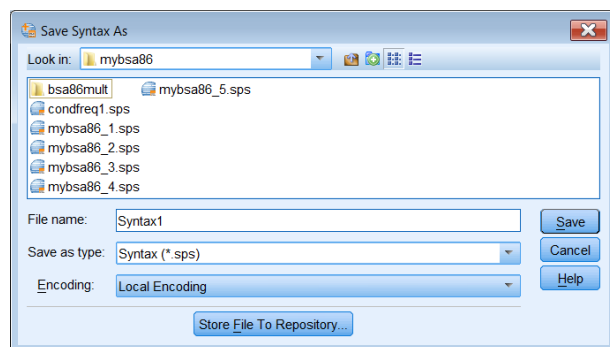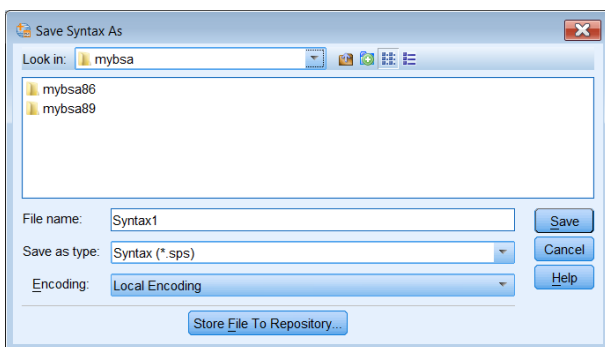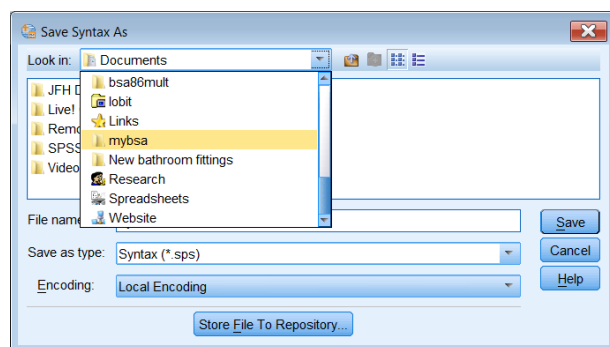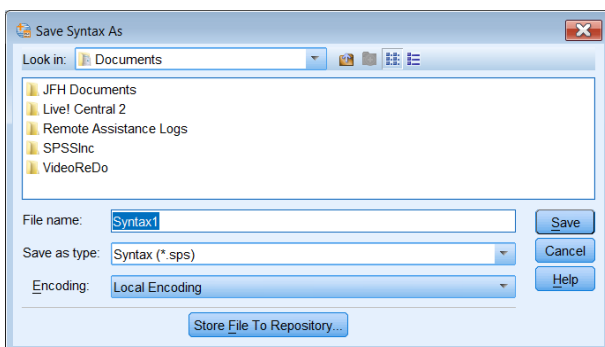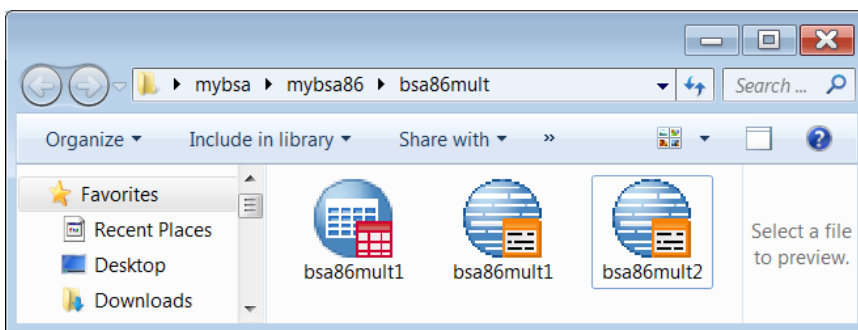

There is already a file **bsa86mult1.sps** which you created in the previous tutorial: there are no other files in the folder of type ***.sps**. In the File name: box, replace Syntax1 with mybsamult2.sps and press Save :
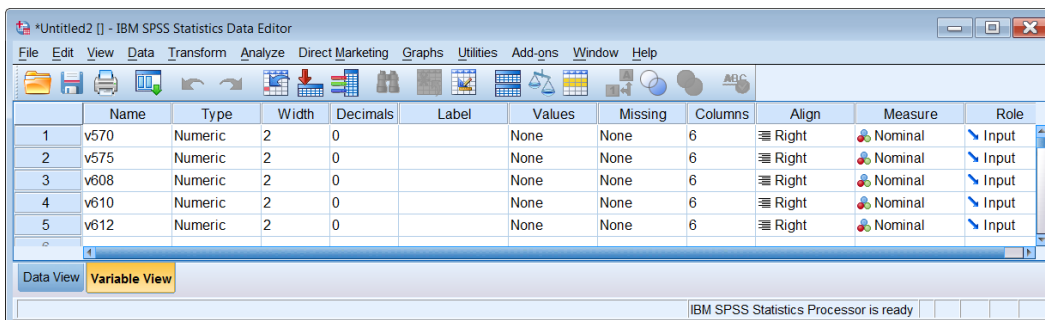
The syntax editor name changes from Syntax1 to bsamult2.sps
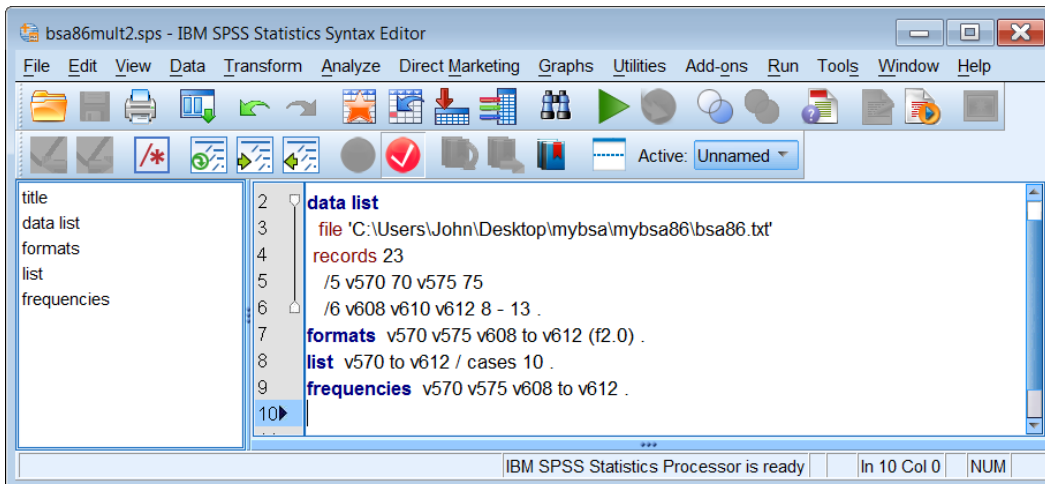
The new file **bsa86mult2.sps** is now dspliyed in folder **mybsa86mult**

You don't actually need to save the data editor yet (it can always wait until the data dictionary is complete) but we'll do it anyway. Go back to the data editor and repeat the process for saving:

File **>** Save As ...
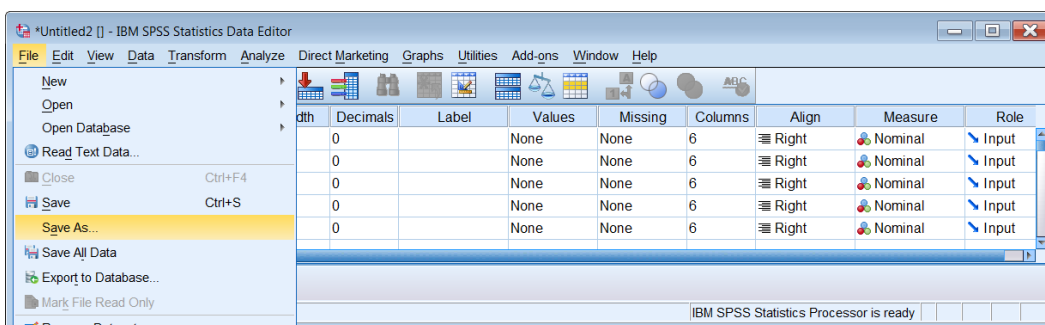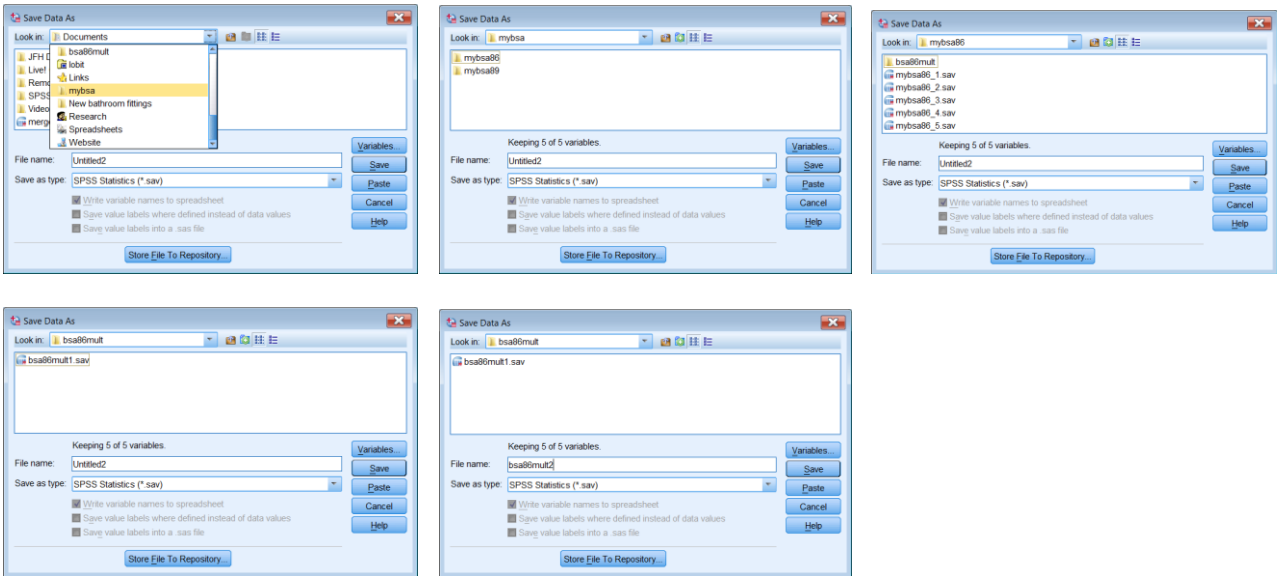
The name of the data editor changes from *Untitled2 to bsamult2.sav
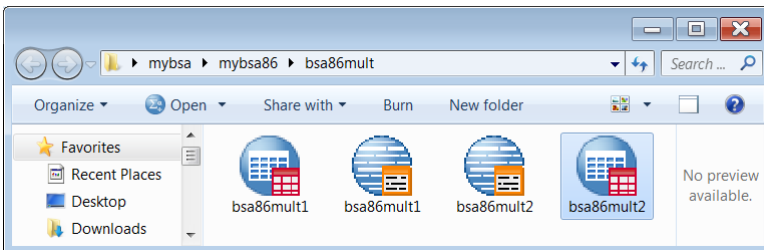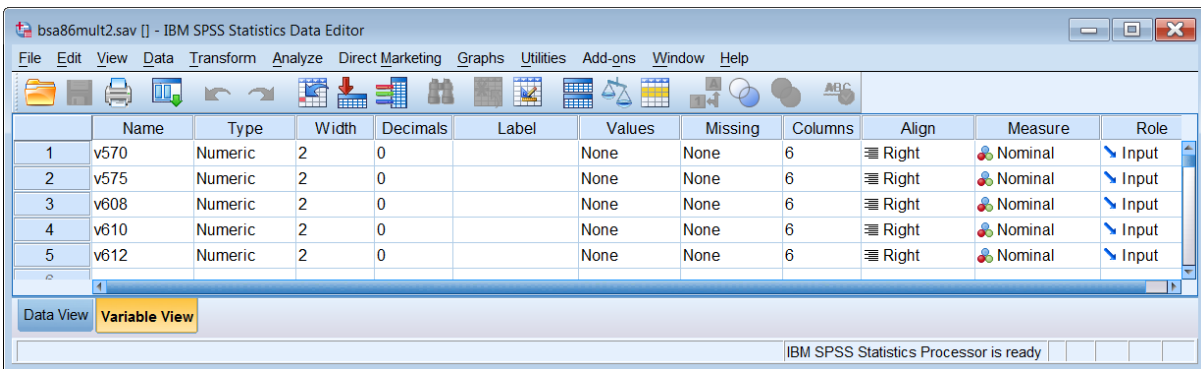
That's the basic job done: your variables have been defined and the data read in. If you like, you can now close SPSS and take a break, otherwise skip to page 16.

When you come back, just re-open folder **mybsa86mult**
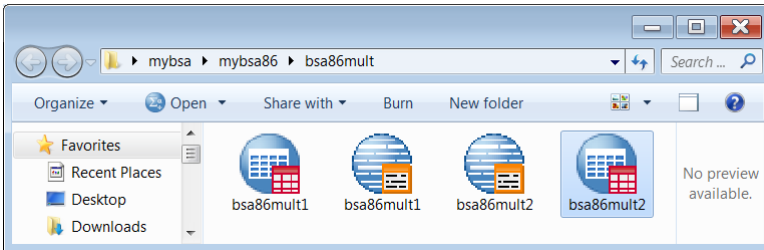
and double click on **bsamult2.sps**.

SPSS will open a new data editor and (if your settings are set for it) a new syntax editor, but your own **bsamult2.sps** file will be in front.
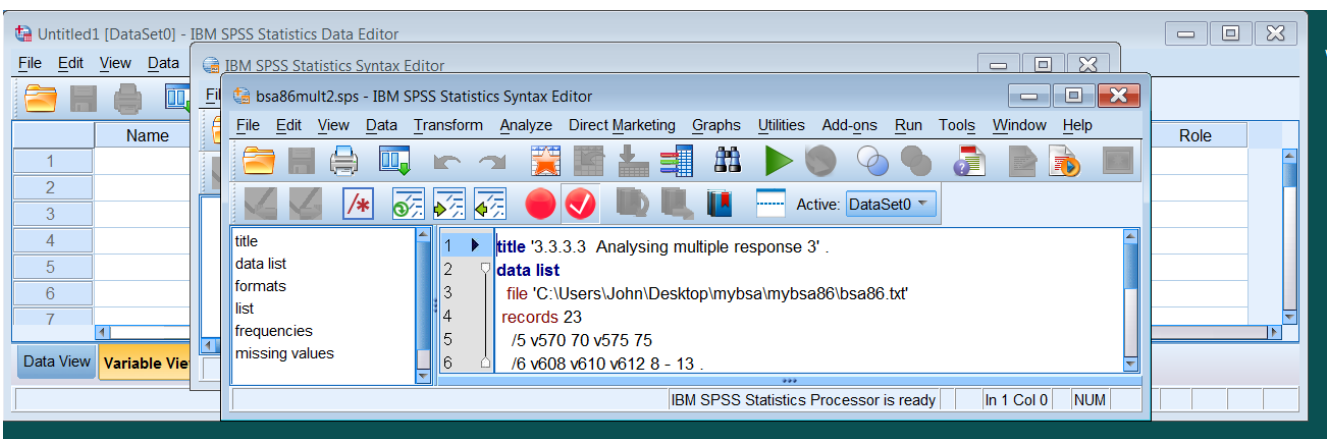
You'll also need to double click on **bsamult2.sav** otherwise SPSS will try to use the the blank data editor as the working file and you will get errors when it can't find the variable names!

Task 2: a) Complete the data dictionary with missing values, variable labels and value labels.
        b) Define a multiple response group for the range of valid codes 01 - 10 for Q.70f
        c) Produce a grouped frequency count.

```
bsa86mult2.sps - IBM SPSS Statistics Syntax Editor

File  Edit  View  Data  Transform  Analyze  Direct Marketing  Graphs  Utilities  Add-ons  Run  Tools  Window  Help

                                    Active: Unnamed

title              2     data list
data list          3       file 'C:\Users\John\Desktop\mybsa\mybsa86\bsa86.txt'
formats            4      records 23
list               5       /5 v570 70 v575 75
frequencies        6       /6 v608 v610 v612 8 - 13 .
                   7     formats  v570 v575 v608 to v612 (f2.0) .
                   8     list  v570 to v612 / cases 10 .
                   9     frequencies  v570 v575 v608 to v612 .
                  10

                                                    IBM SPSS Statistics Processor is ready    In 10 Col 0    NUM
```
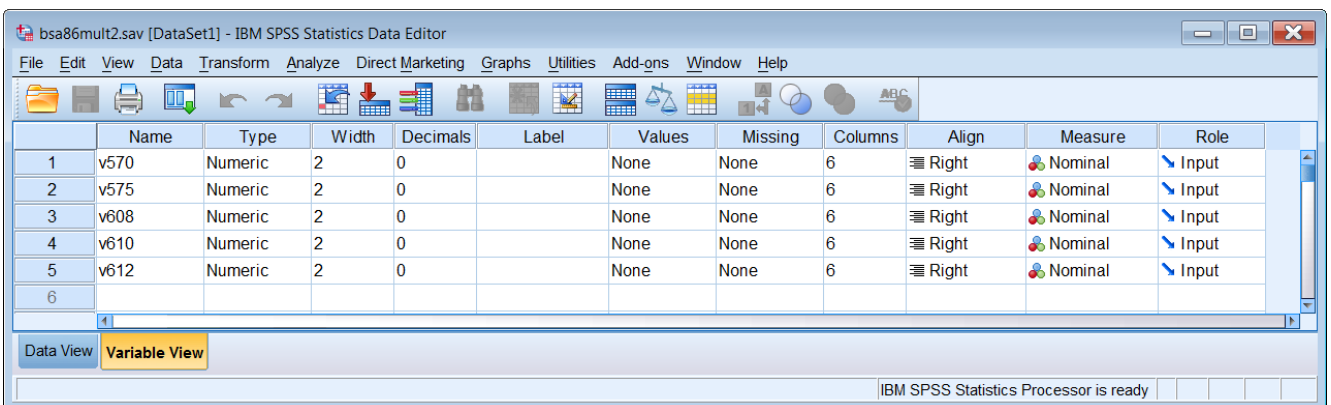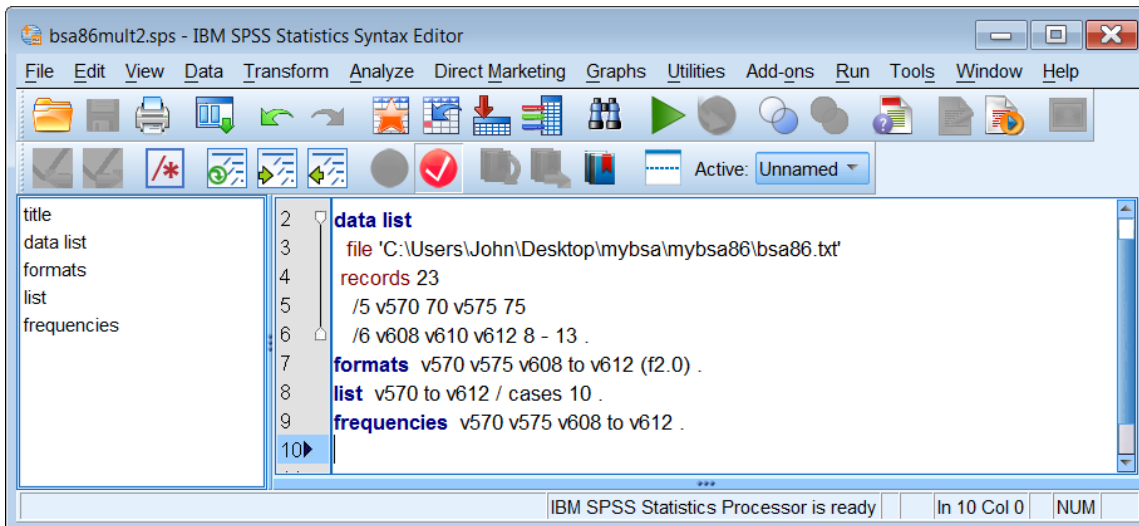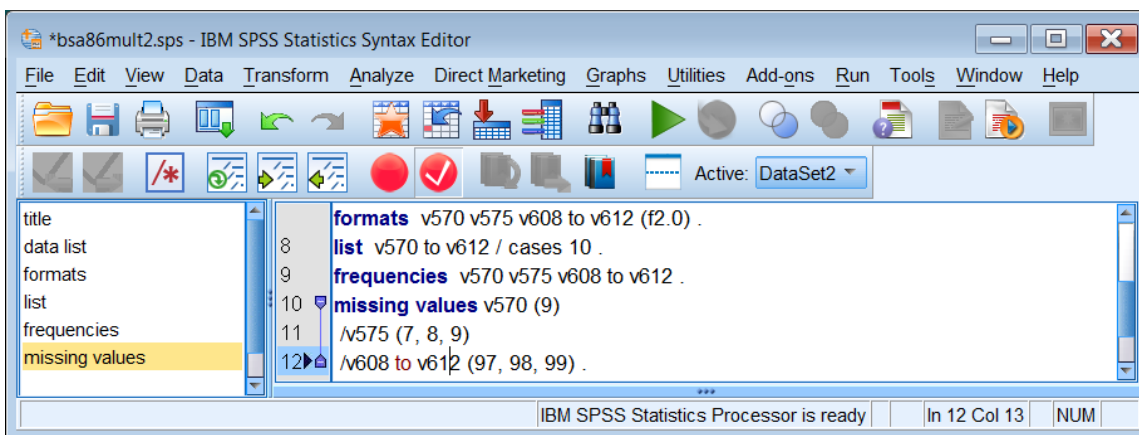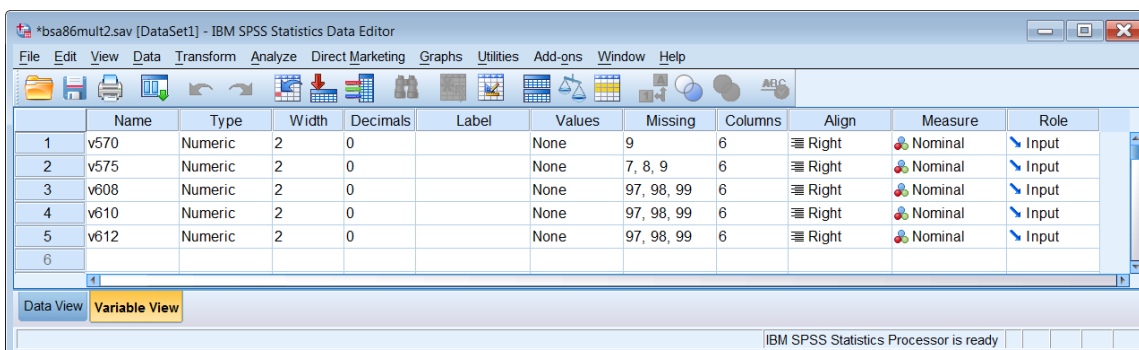
Step 7: Specify missing values

missing values v570 (9)
        /v575 (7, 8, 9)
        /v608 to v612 (97, 98, 99) .

```
*bsa86mult2.sps - IBM SPSS Statistics Syntax Editor

File  Edit  View  Data  Transform  Analyze  Direct Marketing  Graphs  Utilities  Add-ons  Run  Tools  Window  Help

                                    Active: DataSet2

title                  formats  v570 v575 v608 to v612 (f2.0) .
data list         8    list  v570 to v612 / cases 10 .
formats           9    frequencies  v570 v575 v608 to v612 .
list             10    missing values v570 (9)
frequencies      11     /v575 (7, 8, 9)
missing values   12     /v608 to v612 (97, 98, 99) .

                                                    IBM SPSS Statistics Processor is ready    In 12 Col 13    NUM
```

Click on

```
*bsa86mult2.sav [DataSet1] - IBM SPSS Statistics Data Editor

File  Edit  View  Data  Transform  Analyze  Direct Marketing  Graphs  Utilities  Add-ons  Window  Help

        Name     Type      Width  Decimals  Label    Values   Missing    Columns  Align     Measure     Role
1       v570     Numeric   2      0                   None     9          6        Right     Nominal     Input
2       v575     Numeric   2      0                   None     7, 8, 9    6        Right     Nominal     Input
3       v608     Numeric   2      0                   None     97, 98, 99 6        Right     Nominal     Input
4       v610     Numeric   2      0                   None     97, 98, 99 6        Right     Nominal     Input
5       v612     Numeric   2      0                   None     97, 98, 99 6        Right     Nominal     Input
6

Data View  Variable View
                                                    IBM SPSS Statistics Processor is ready
```
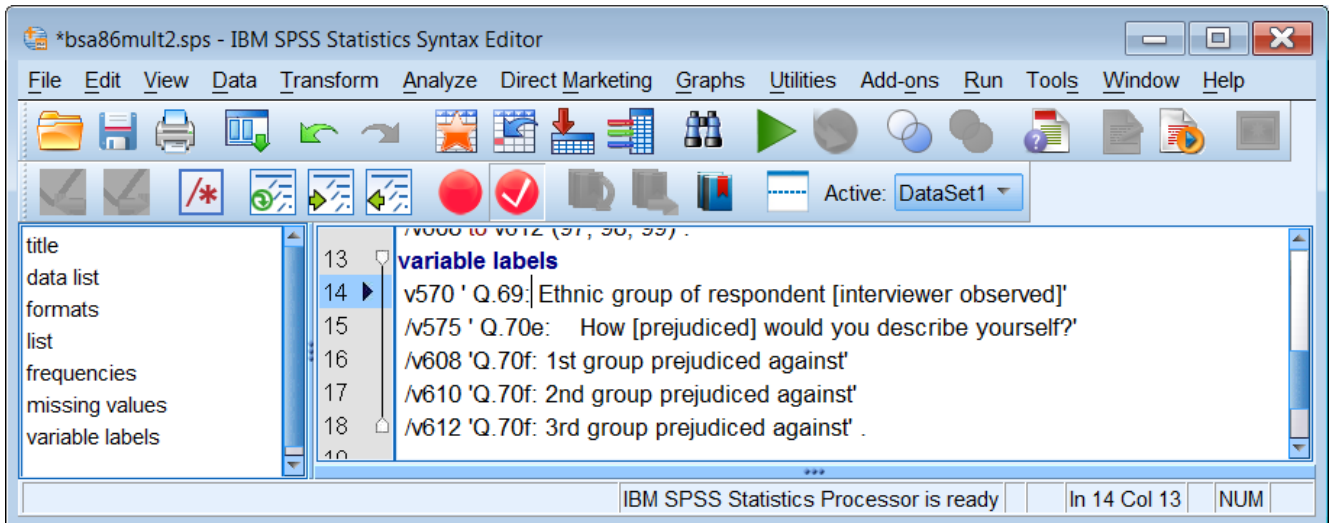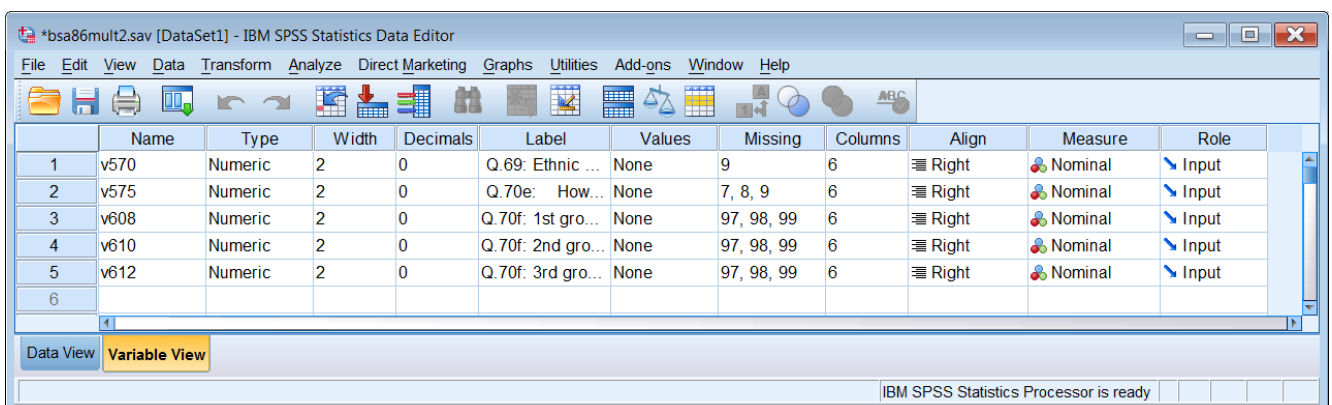
16

Step 8: Add variable labels
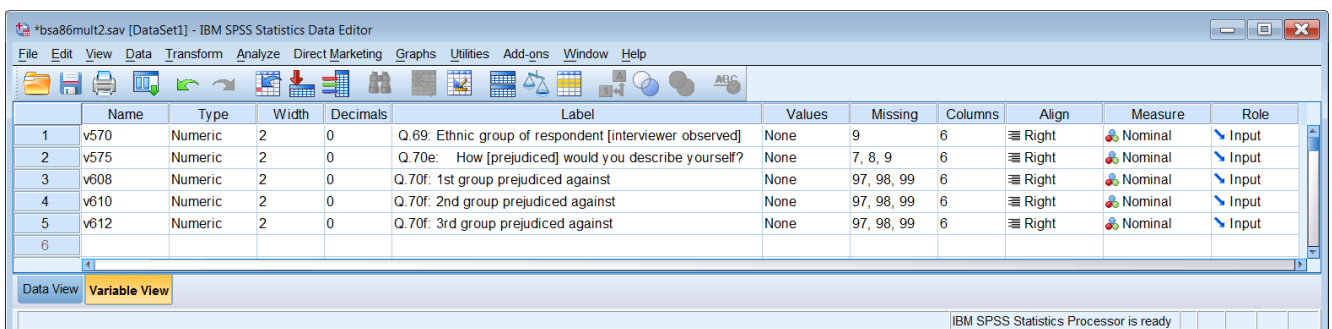
variable labels
v570 ' Q.69 Ethnic group of respondent [interviewer observed]'
/v575 ' Q.70e: How [prejudiced] would you describe yourself?'
/v608 'Q.70f: 1st group prejudiced against'
/v610 'Q.70f: 2nd group prejudiced against'
/v612 'Q.70f: 3rd group prejudiced against' .

Click on

Drag the Label | Values column separator to the right to see the full label:

Step 9: Add value labels

value labels
     v575
         1 'Very prejudiced'
         2 'A little prejudiced'
         3 'Not at all prejudiced'
    /v570
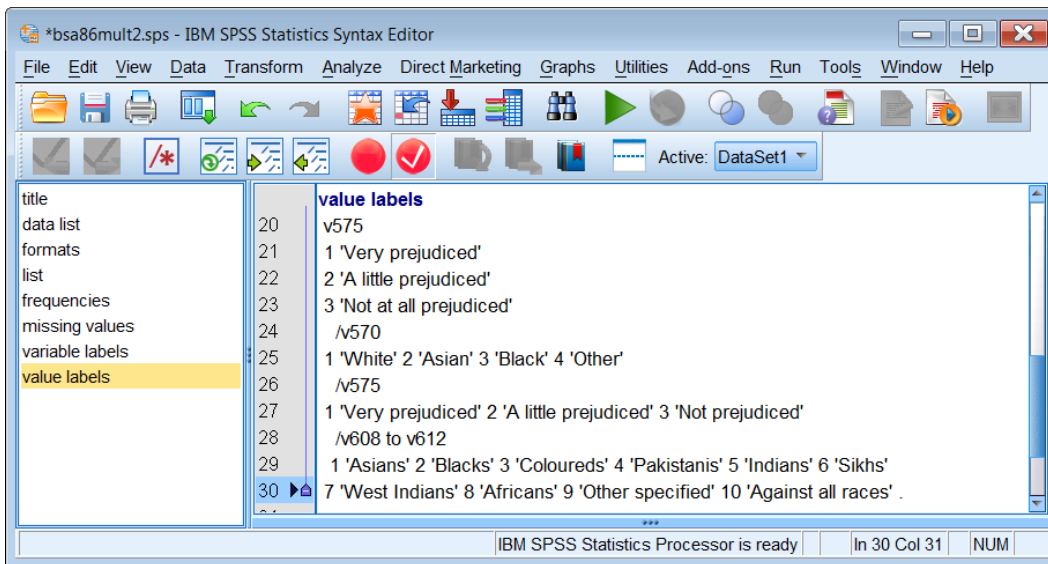   **1 'White' 2 'Asian' 3 'Black' 4 'Other'**
     /v575
         1 'Very prejudiced' 2 'A little prejudiced' 3 'Not prejudiced'
    /v608 to v612
         1 'Asians' 2 'Blacks' 3 'Coloureds' 4 'Pakistanis' 5 'Indians' 6 'Sikhs'
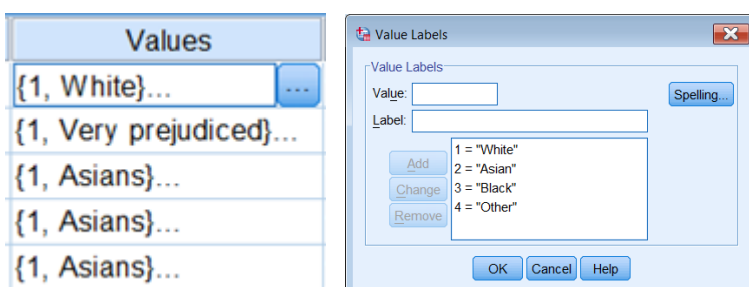   **7 'West Indians' 8 'Africans' 9 'Other specified' 10 'Against all races' .**

Click on

You can do quick checks on the value labels by clicking on a cell in the Values column and then on the blue square :

18

Step 10: **Save your work again**

This time you only have to click on ![save icon] in the tool bar at the top of each editor. The files are instantly updated and saved in folder **bsa86mult**. You only have to do all the previous work once. The data editor is now fixed and can be used again, and again, and again.

Step 11: Now run the frequency count again

frequencies v570 v575 v608 to v612 .

Click on ![green play button]

Statistics

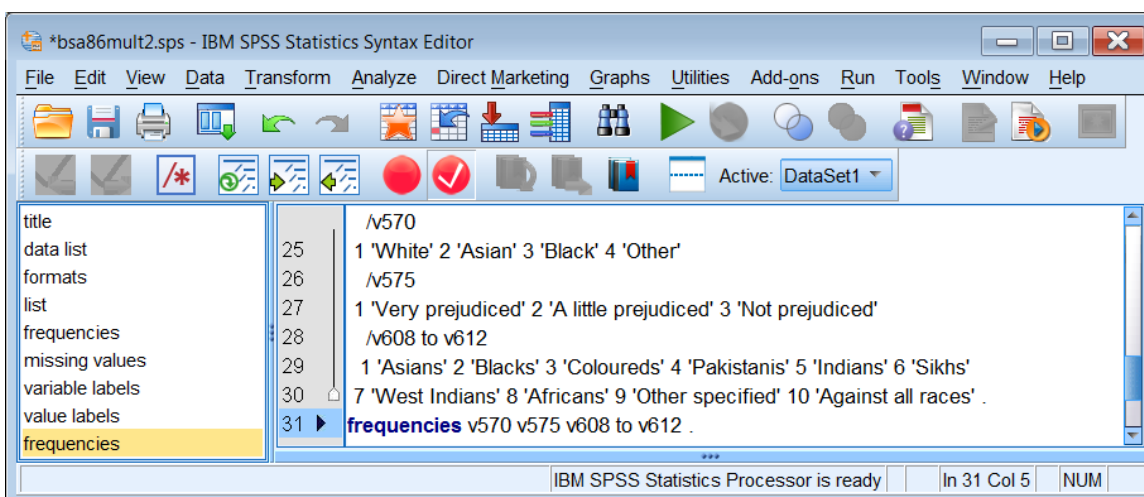|		Q.69 Ethnic group of respondent [interviewer observed]	Q.70e: How [prejudiced] would you describe yourself?	Q.70f: 1st group prejudiced against	Q.70f: 2nd group prejudiced against	Q.70f: 3rd group prejudiced against
N	Valid	3092	3068	1042	156	36
	Missing	8	32	2058	2944	3064

Q.69 Ethnic group of respondent [interviewer observed]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	White	2976	96.0	96.2	96.2
	Asian	55	1.8	1.8	98.0
	Black	49	1.6	1.6	99.6
	Other	12	.4	.4	100.0
	Total	3092	99.7	100.0	
Missing	9	8	.3		
Total		3100	100.0		

Q.70e: How [prejudiced] would you describe yourself?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very prejudiced	116	3.7	3.8	3.8
	A little prejudiced	999	32.2	32.6	36.3
	Not prejudiced	1953	63.0	63.7	100.0
	Total	3068	99.0	100.0	
Missing	7	20	.6		
	8	4	.1		
	9	8	.3		
	Total	32	1.0		
Total		3100	100.0		

Q.70f: 1st group prejudiced against

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Asians	153	4.9	14.7	14.7
	Blacks	55	1.8	5.3	20.0
	Coloureds	44	1.4	4.2	24.2
	Pakistanis	94	3.0	9.0	33.2
	Indians	22	.7	2.1	35.3
	Sikhs	1	.0	.1	35.4
	West Indians	68	2.2	6.5	41.9
	Africans	11	.4	1.1	43.0
	Other specified	27	.9	2.6	45.6
	Against all races	567	18.3	54.4	100.0
	Total	1042	33.6	100.0	
Missing	97	64	2.1		
	98	2	.1		
	99	7	.2		
	System	1985	64.0		
	Total	2058	66.4		
Total		3100	100.0		

Q.70f: 2nd group prejudiced against

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Asians	16	.5	10.3	10.3
	Blacks	25	.8	16.0	26.3
	Coloureds	7	.2	4.5	30.8
	Pakistanis	30	1.0	19.2	50.0
	Indians	25	.8	16.0	66.0
	Sikhs	3	.1	1.9	67.9
	West Indians	26	.8	16.7	84.6
	Africans	13	.4	8.3	92.9
	Other specified	7	.2	4.5	97.4
	Against all races	4	.1	2.6	100.0
	Total	156	5.0	100.0	
Missing	97	3	.1		
	System	2941	94.9		
	Total	2944	95.0		
Total		3100	100.0		

Q.70f: 3rd group prejudiced against

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Asians	4	.1	11.1	11.1
	Blacks	1	.0	2.8	13.9
	Coloureds	2	.1	5.6	19.4
	Pakistanis	5	.2	13.9	33.3
	Indians	4	.1	11.1	44.4
	West Indians	4	.1	11.1	55.6
	Africans	4	.1	11.1	66.7
	Other specified	10	.3	27.8	94.4
	Against all races	2	.1	5.6	100.0
	Total	36	1.2	100.0	
Missing	System	3064	98.8		
Total		3100	100.0		

Task 3: a) Define a multiple response group for the range of valid codes 01 - 10 for Q.70f
   b) Produce a grouped frequency count for Q.70f.

Step 11: Define your **mult response group variable** by giving it a group name and label and by specifying the list of variables to scan together with the lower and upper limits of the range of values to be tabulated:

mult response groups =
      q70f 'Q.70f: Groups prejudiced against'
      (v608 to v612 (1,10))
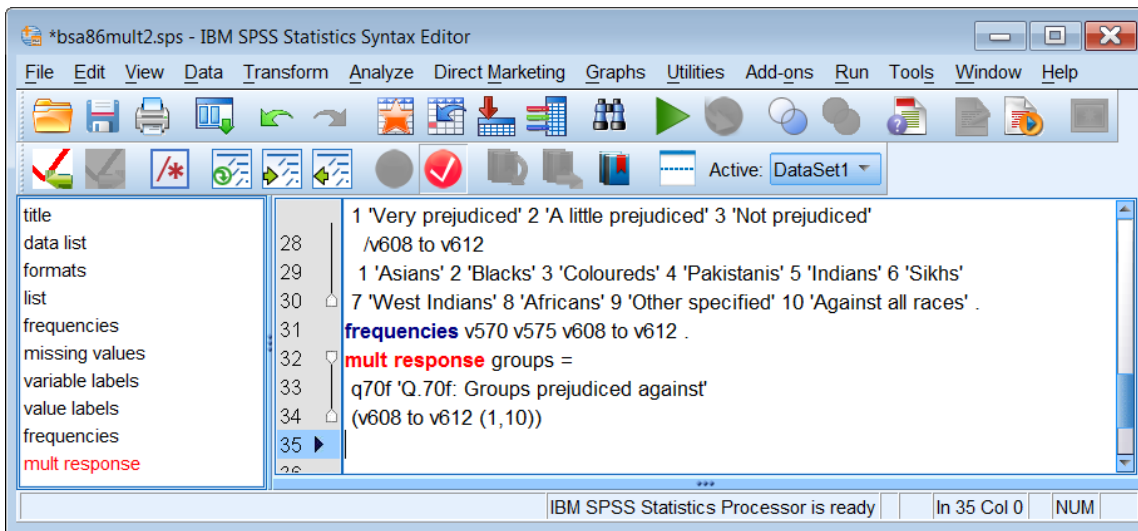
mult response is in **red** because the command is not yet complete. If you try to run the job, you will get an error message. We need to specify what analysis we want as well, in this case a group frequency count. You can't use the ordinary FREQUENCIES command, you have to specify the frequencies within the mult response command, as a sub-command preceded by a forward slash (and inset by at least one space if it's on a separate line):

      /frequencies q70f .
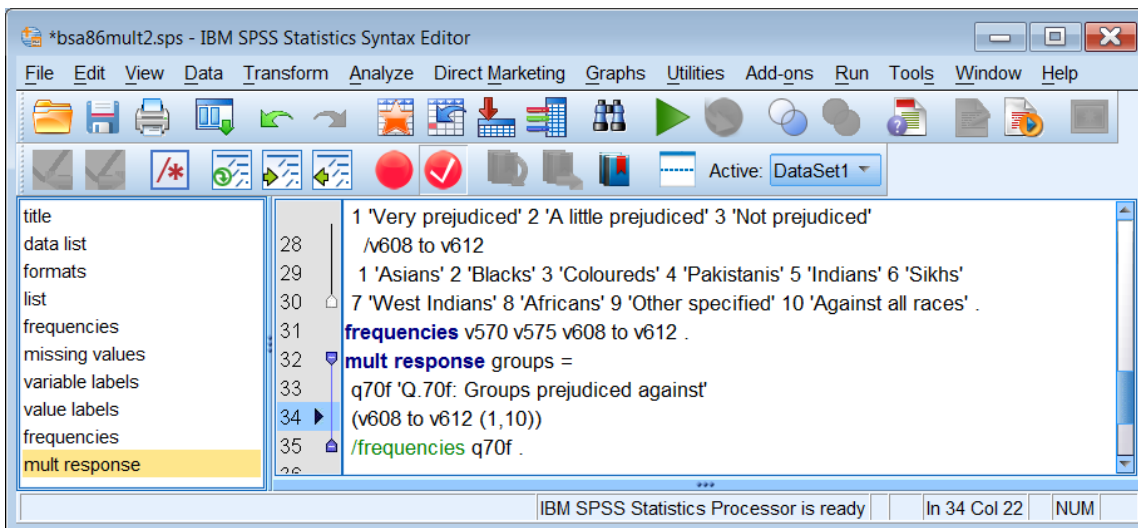
Click on

Case Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
q70f[a]	1043	33.6%	2057	66.4%	3100	100.0%

a. Group

q70f Frequencies

		Responses		Percent of Cases
		N	Percent	
Q.70f: Groups prejudiced against[a]	Asians	173	14.0%	16.6%
	Blacks	81	6.6%	7.8%
	Coloureds	53	4.3%	5.1%
	Pakistanis	129	10.5%	12.4%
	Indians	51	4.1%	4.9%
	Sikhs	4	.3%	.4%
	West Indians	98	7.9%	9.4%
	Africans	28	2.3%	2.7%
	Other specified	44	3.6%	4.2%
	Against all races	573	46.4%	54.9%
Total		1234	100.0%	118.3%

a. Group

is the grouped frequency table for the group variable Q.70f. The group variable **cannot be saved**. Group variables have to be recreated every time you need to use them, but you don't have to write out the syntax each time: you can save the specifications into another syntax file and then copy them back next time you want to use them

End of tutorial

Next tutorial: 3.3.3.3 Analysing multiple response exercise 3 - More replies than values

[In preparation]

Using question Q.125 about welfare benefits received, this exercise demonstrates how to handle a question in which the same code values 1 - 6 are repeated within a set of 12 responses to the same question (but in different fields). This involves recoding the second set of values, but needs great care whn dealing with missing values.