TRAINING PROGRAMME ON ULITISATION OF SPSS SOFTWARE PACKAGE FOR QUALITATIVE ANALYSIS OF EDUCATIONAL DATA FOR IASE/CTE AND DIET FACULTY OF TAMILNADU AND PONDICHERRY

Academic Coordinator

DR. G. VISWANATHAPPA
Senior Lecturer in Education
Regional Institute of Education (NCERT)
MYSORE – 570 006

Dates
20th November to 1st December 2006

Regional Institute of Education (NCERT)
MYSORE – 570 006

Regional Institute of Education (NCERT), Mysore 570 006

Training Programme on the Utilisation of SPSS Software in Educational Data Analysis for the DTERT, IASE and CTE Staff of Tamilnadu

Day & Date	9.30 to 10.00	Session – 1 (10.00 to 11.30)	Session – 2 (11.45 to 1.15)	Session – 3 (2.15 to 3.45)	Session – 4 (4.00 to 5.30)
20-11-2006 Monday	Registration & Inauguration		Introduction to Educational Data-Quantitative data & Qualitative data (CSN)	Tools and Techniques for collecting Educational Data (VDB)	Review of Tools and Techniques for collecting Educational Data (VDB&GV)
21-11-2006 Tuesday	Review of Previous day Programme	Introduction to Educational Data-Levels of Measurement (GV)	Introduction to Statistical Methods used in Qualitative data (DB)	Introduction to Statistical Methods used in Quantative data (GV)	Group work on Statistical Methods (GV/DB/ACJ)
22-11-2006 Wednesday	Review of Previous day Programme	Data entry and Manipulation in MS-Excel (ACJ)	Practicum on Data Entry in MS- Excel (GV/DB/ACJ)	Data Analysis in MS-Excel	Practicum on Data Analysis in MS-Excel (ACJ & GV)
23-11-2006 Thursday	Review of Previous day Programme	Comparison of Data base software (MS-Excel & SPSS) (GV/DB/ACJ)	Working with MS-Excel (ACJ/DB)	Planning and Preparation of Data file in SPSS (GV/DB)	Practicum on Data Entry in SPSS(GV/DB)
24-11-2006 Friday	Review of Previous day Programme	Data exploring techniques (assumption; estimation and hypotheses) (GV/DB)	Practicum on creating Data File in SPSS(GV/DB/ACJ)	Data Processing – Descriptive Statistics(GV/DB)	Practicum on Descriptive Statistics(GV/DB/ACJ)

Day & Date	9.30 to 10.00	Session – 1 (10.00 to 11.30)	Session – 2 (11.45 to 1.15)	Session – 3 (2.15 to 3.45)	Session – 4 (4.00 to 5.30)
25-11-2006 Saturday	Review of Previous day Programme	Data Processing-Statistical Techniques (GV/DB/ACJ)	Practicum on Descriptive Statistics (GV/DB/ACJ)	Data Processing – Bivariate Statistics(GV/DB/ACJ)	Practicum on Bivariate Statistics(GV/DB/ACJ)
26-11-2006 Sunday	Review of Previous day Programme	Data Processing – Inferential Statistics(z-Test / t-Test) (GV/DB)	Practicum (GV/DB/ACJ)	Data Processing – Inferential Statistics(Chi-square Test) (GV/DB)	Practicum on Chi square Test(GV/DB/ACJ)
27-11-2006 Monday	Review of Previous day Programme	Data Processing – Inferential Statistics(ANOVA-and ANCOVA)-I (GV/DB)	Practicum(GV/DB/ACJ)	Data Processing – Inferential Statistics(ANOVA and ANCOVA)-II (GV/DB)	Practicum(GV/DB/ACJ)
28-11-2006 Tuesday	Review of Previous day Programme	File transformations - I(GV/DB)	Practicum on File Transformation(GV/DB/ACJ)	File transformations - II(GV/DB)	Practicum on File Transformation(GV/DB/ACJ)
29-11-2006 Wednesday	Review of Previous day Programme	Data Processing – Multivariate Statistics (GV/DB)	Practicum on Multivariate Statistics(GV/DB/ACJ)	Working with Output view(GV/DB)	Practicum(GV/DB/ACJ)
30-11-2006 Thursday	Review of Previous day Programme	Qualitative research Methods of Data Analysis (DB)	Interpretation and discussion based on Quantitative data (GV/DB/ACJ)	Creating Graphs in MS-Excel(GV/DB/ACJ)	Practicum on Graphs(GV/DB/ACJ)
01-12-2006 Friday	Review of Previous day Programme	Creating Graphs in SPSS(GV/DB)	Practicum on Graphs (GV/DB/ACJ)	Participants Feedback and Valedictory Session	Disbursement of TA/DA

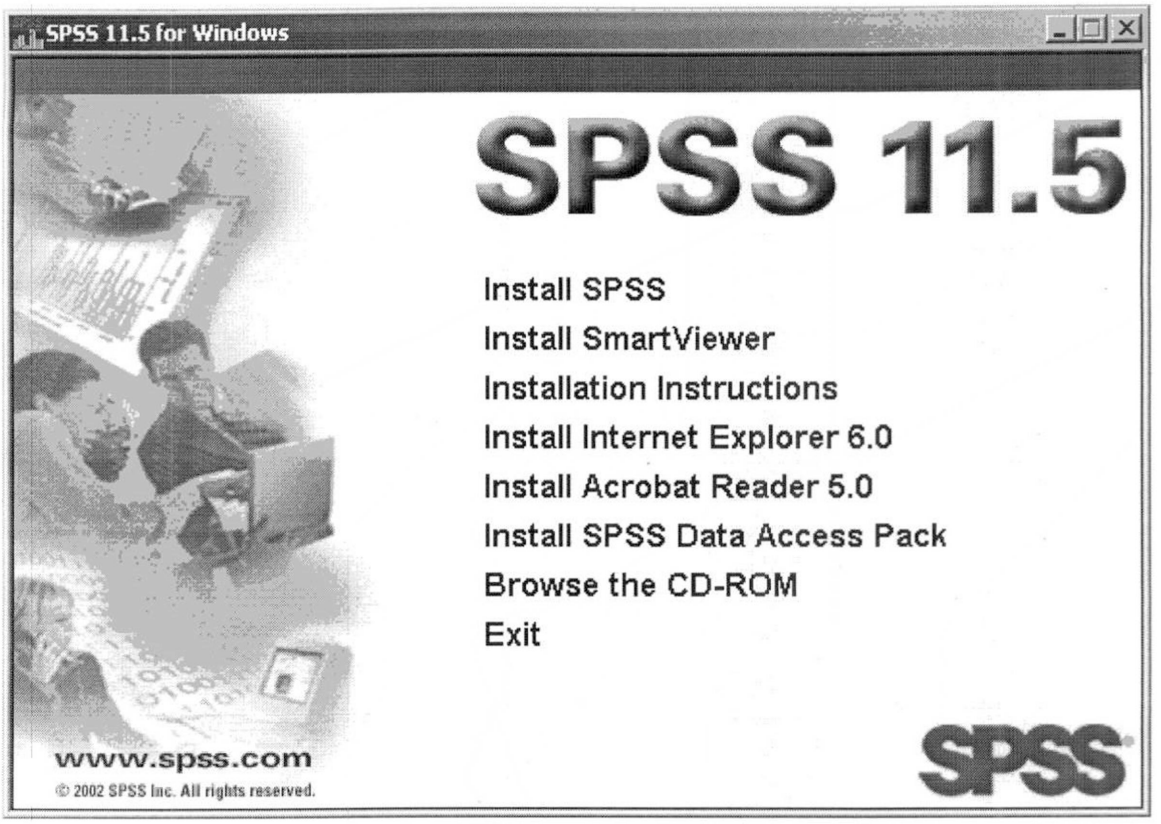
List of Resource Persons: 1. CSN: Prof. C.S. Nagaraju, Principal: GV: Dr. G. Viswanathappa; DB: Prof. D. Basavayya; ACJ: Mr. A. C. Josy

SPSS 11.5 for Windows Installation

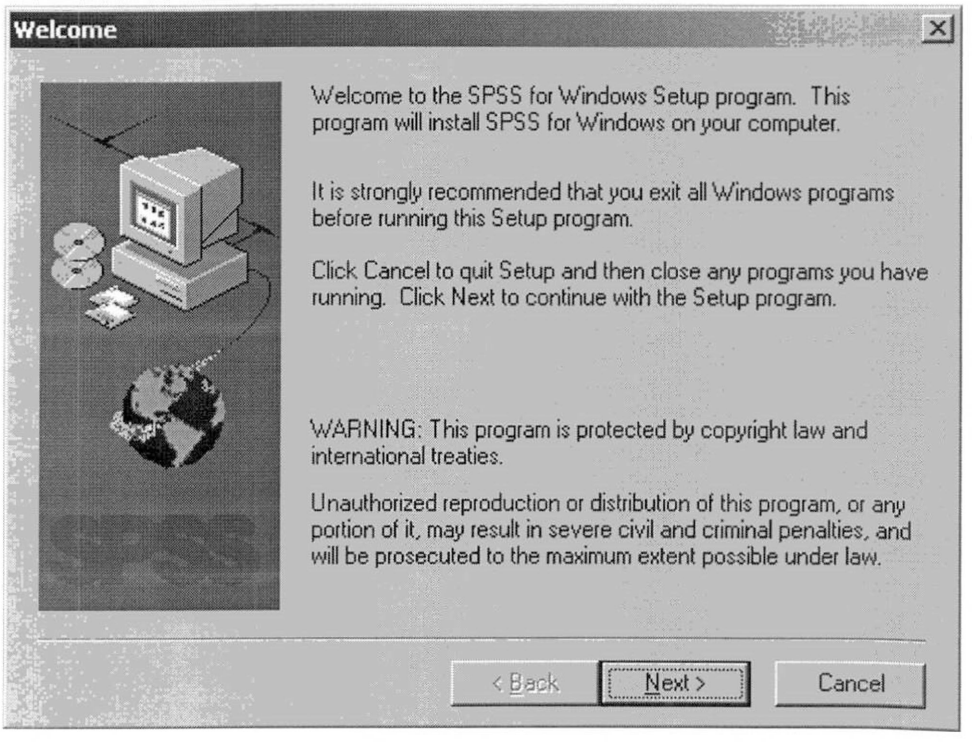
SPSS 11.5 for Windows 98/ME/XP/2000/NT

Installation instructions

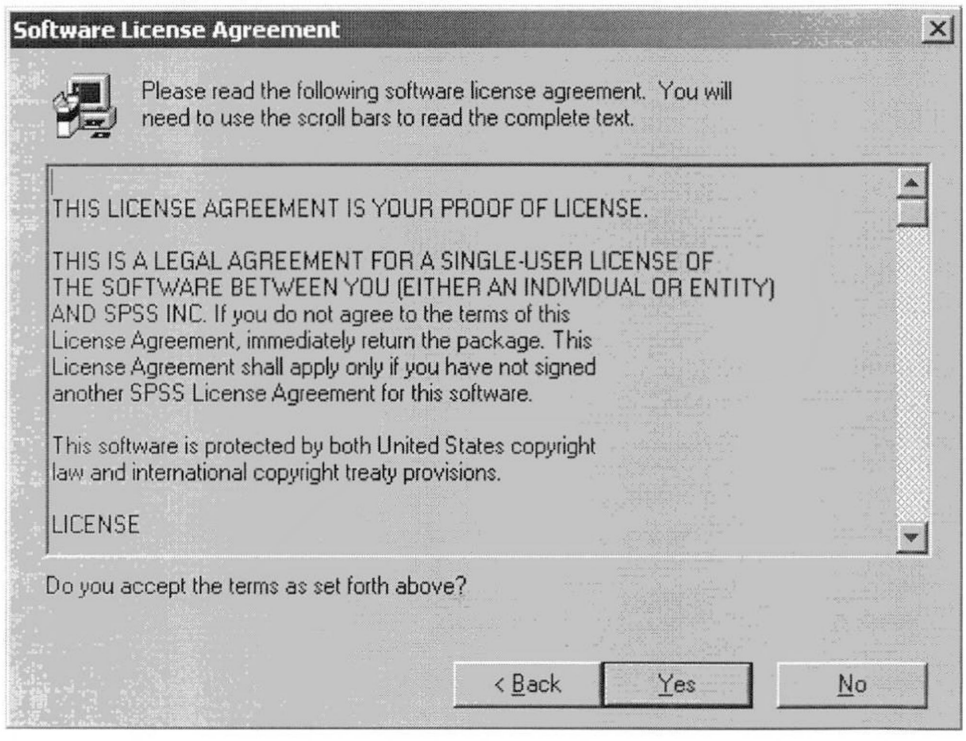
1. Make sure there is 84 megabytes of available space on the hard drive before installation of this program.
2. Make sure you have the "**serial number**" and "**license code**" for SPSS. This information would have been emailed to you. Keep a printout of this email handy so you can type in the "**serial number**" and "**license code**" when prompted during the installation.
3. Insert the "**SPSS 11.5 for Windows**" CD into the disk drive. After several seconds the install splash screen will appear. If the install wizard fails to activate then press "**Start | Run**" menu items. Navigate the Run dialog box to the CD-ROM containing the software and select the "**setup.exe**" executable file. Press the "**Open**" button to select the program and return to the run dialogue box then press the "**Ok**" button to activate the setup program.



Click on '**Install SPSS**'.

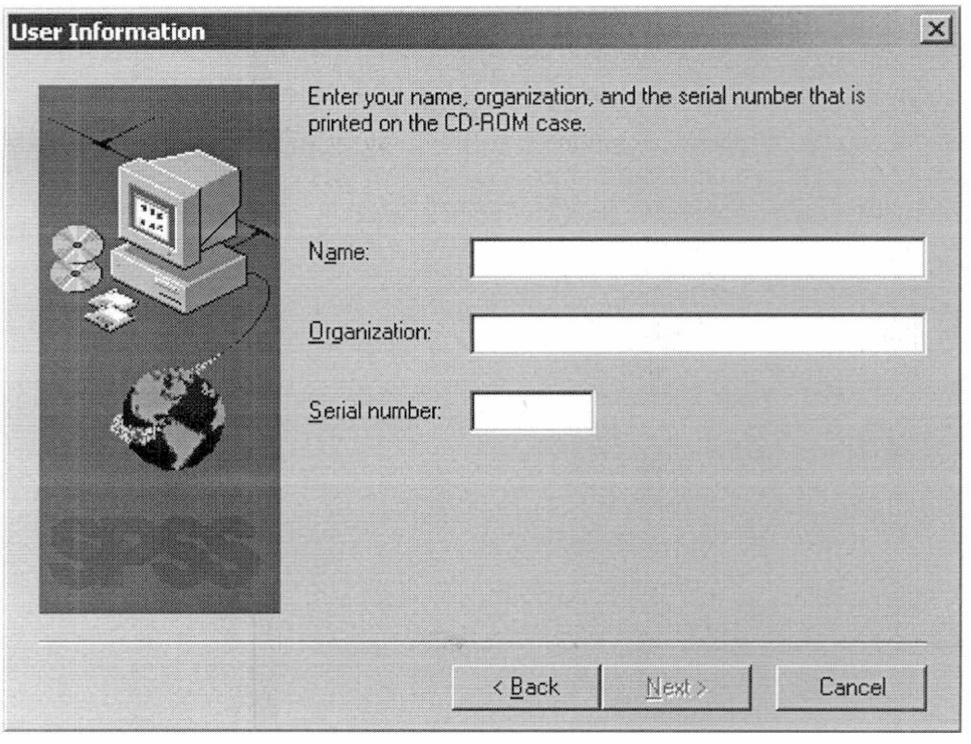


Click on '**Next**'.

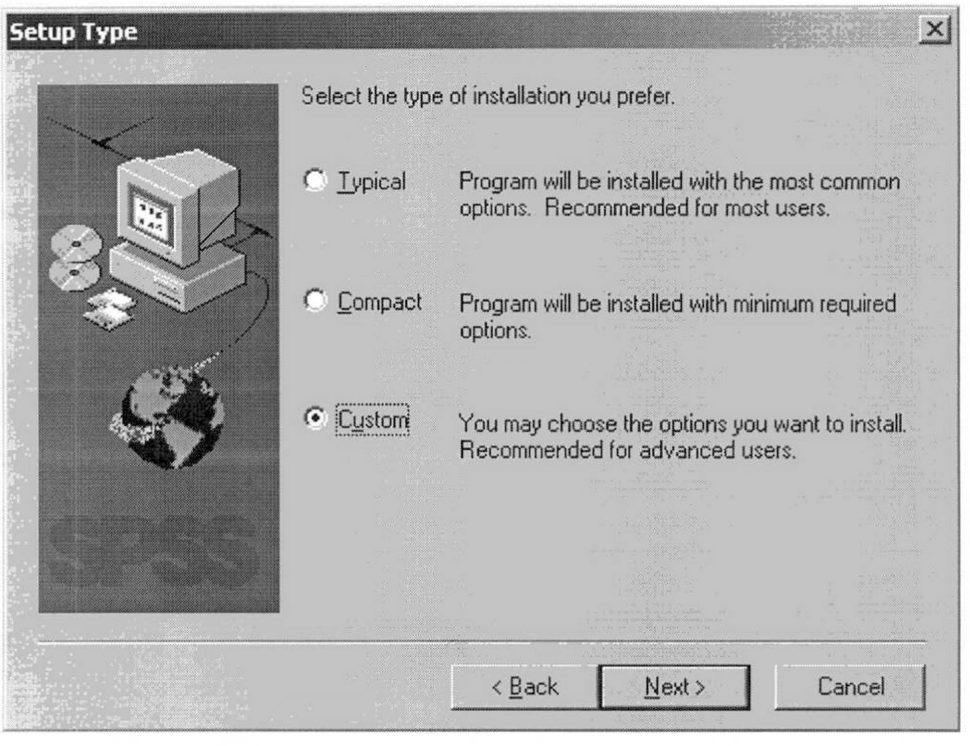


6. Click on '**Yes**'.

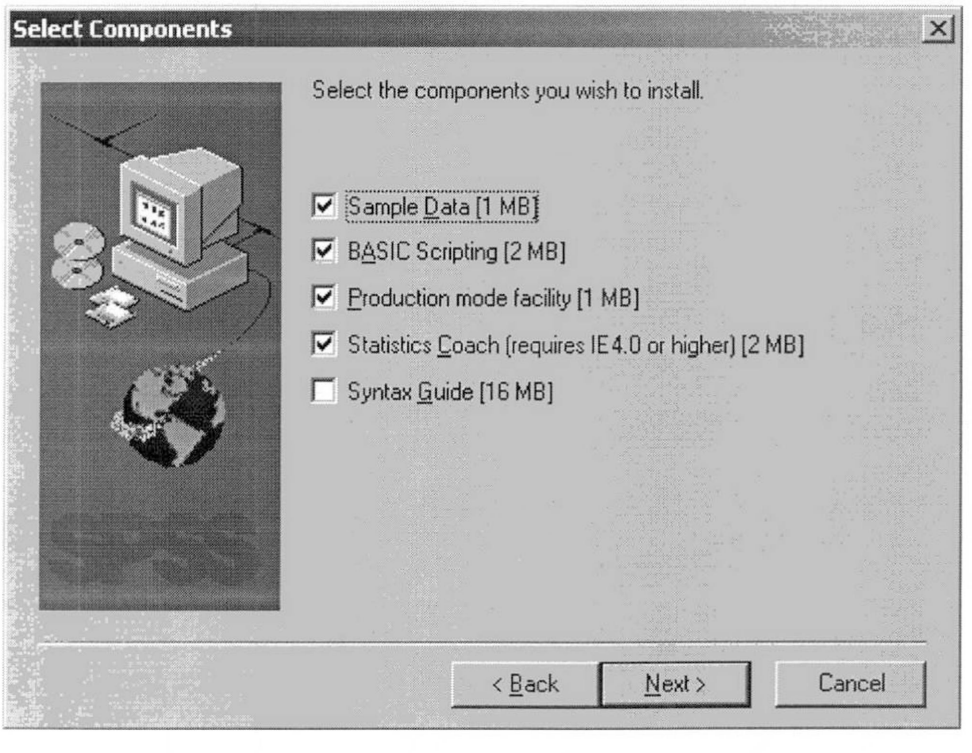
7. Click on '**Next**'.



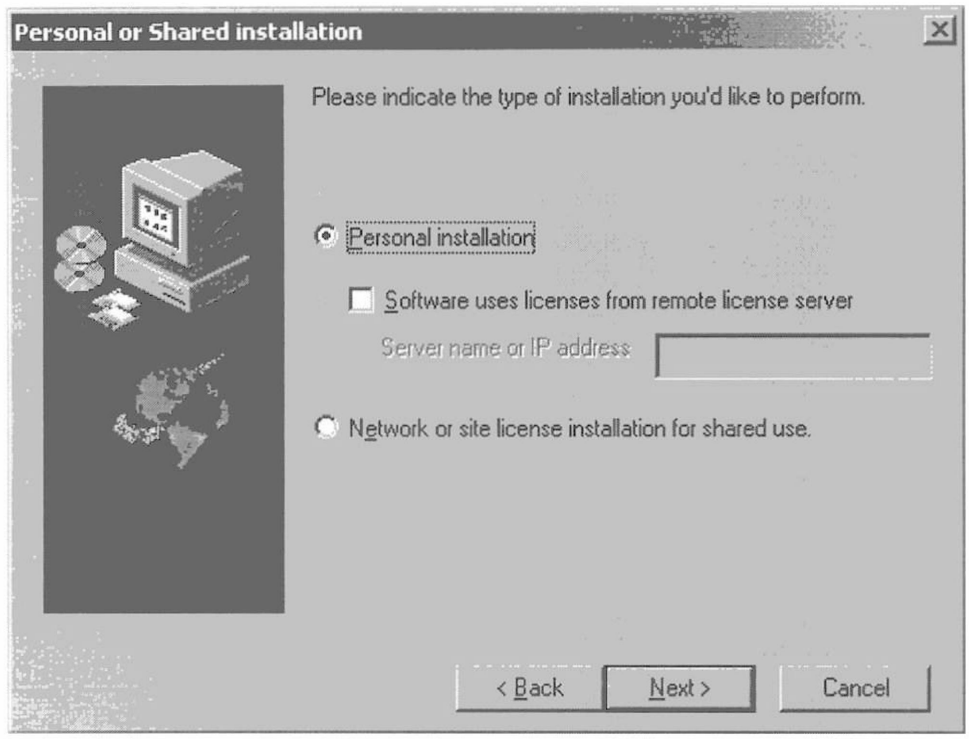
Under '**Name**', enter **your name**, under '**Organization**', enter '**University of Cincinnati**', and under '**Serial number**', enter the serial number supplied to you in the e-mail. Once this is done, click '**Next**'.



Choose '**Custom**' and click on '**Next**'.

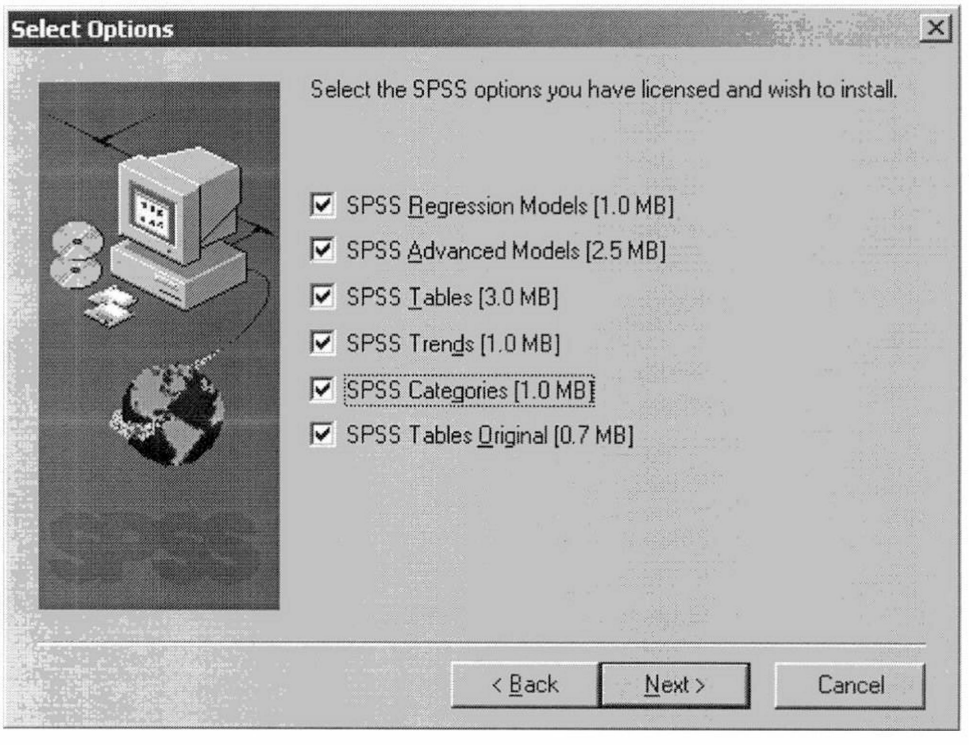


Check the '**Syntax Guide** (16 MB)' check box and click on '**Next**'.

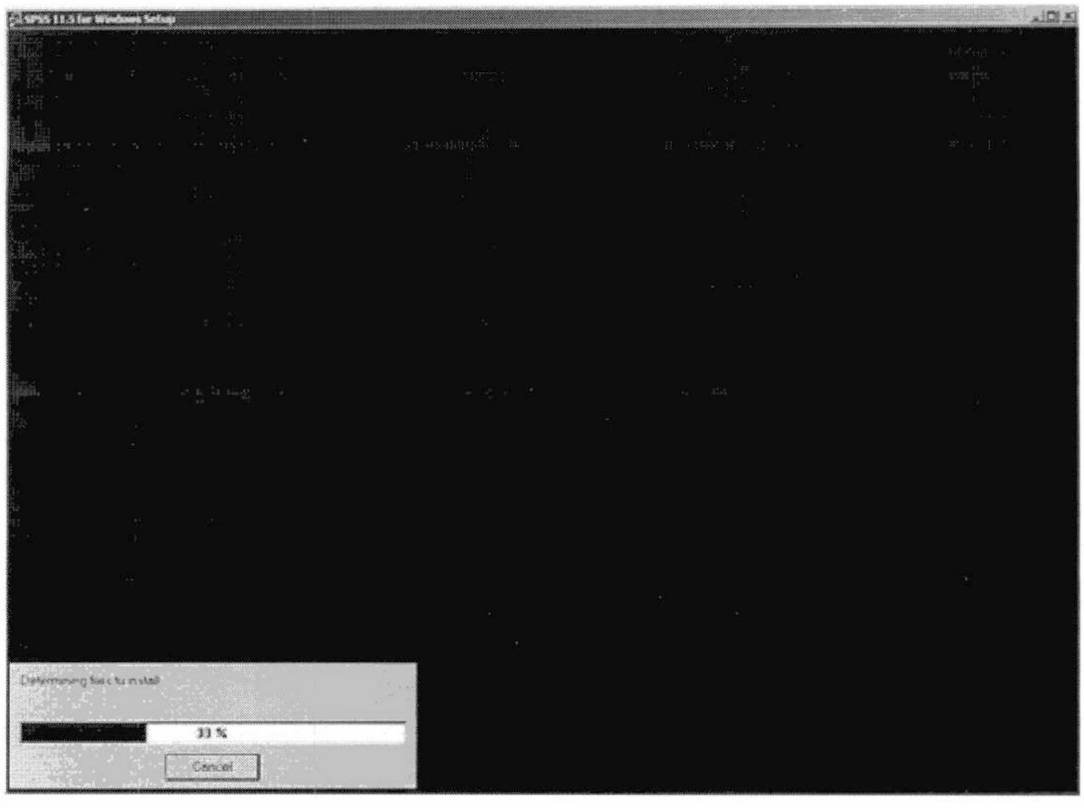


Click on '**Next**'.

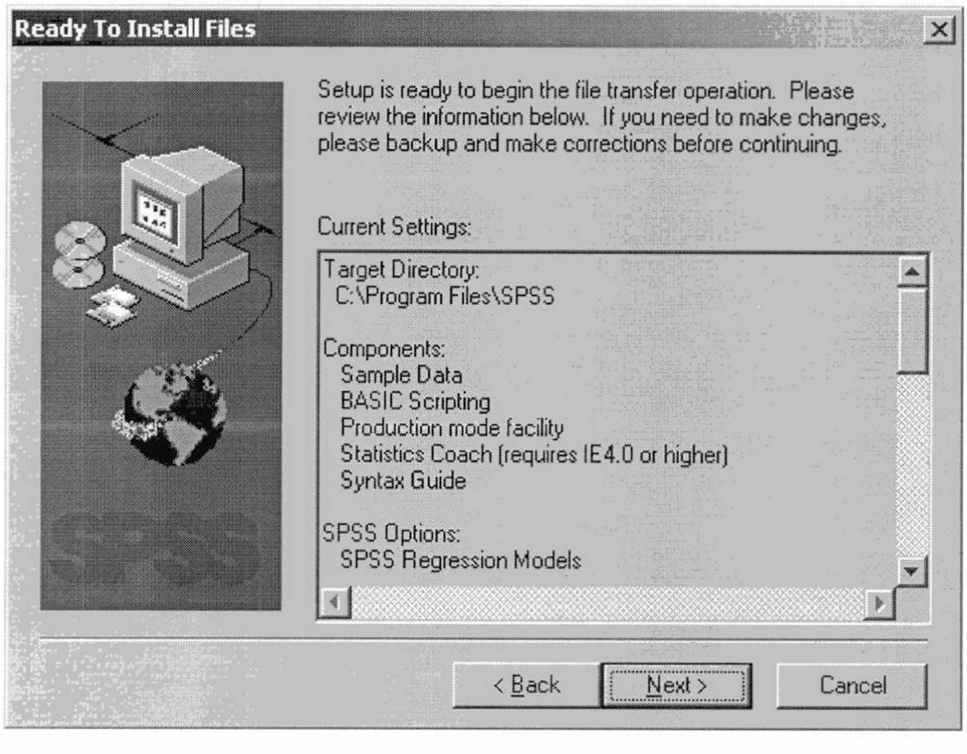
Enter the license number which has been e-mailed to you in the box above. Be sure to enter the license number without any dashes. Then click '**Next**'.



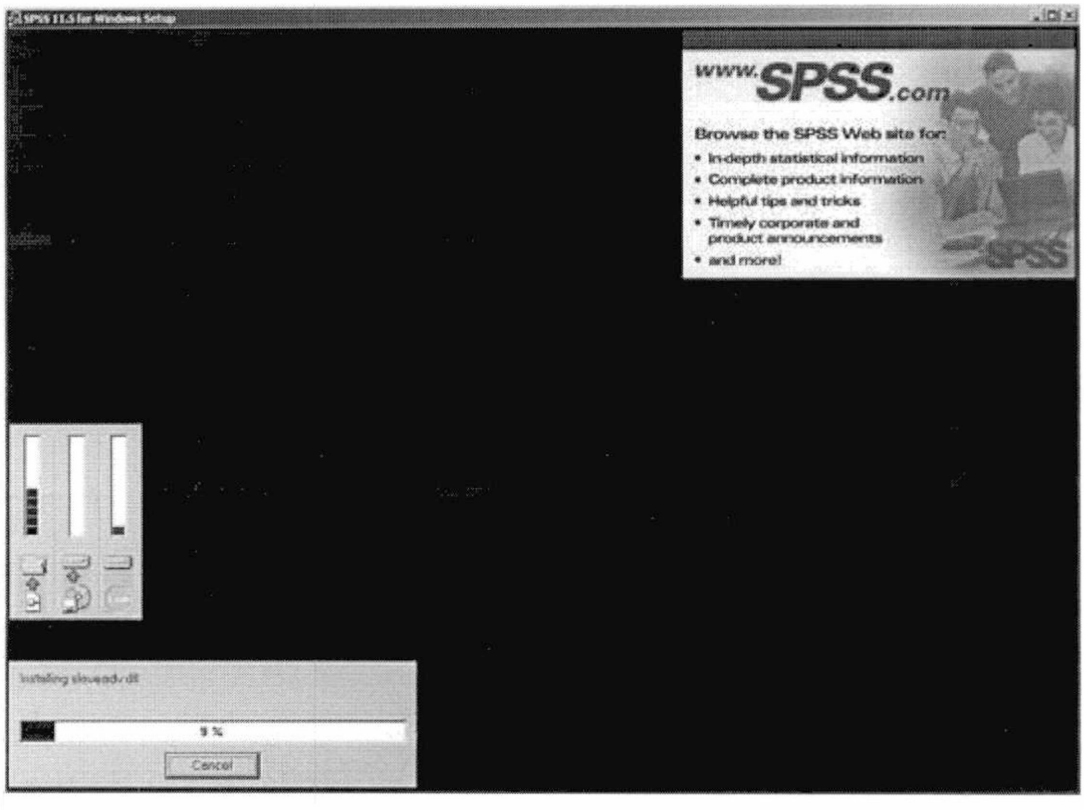
Click '**Next**'.



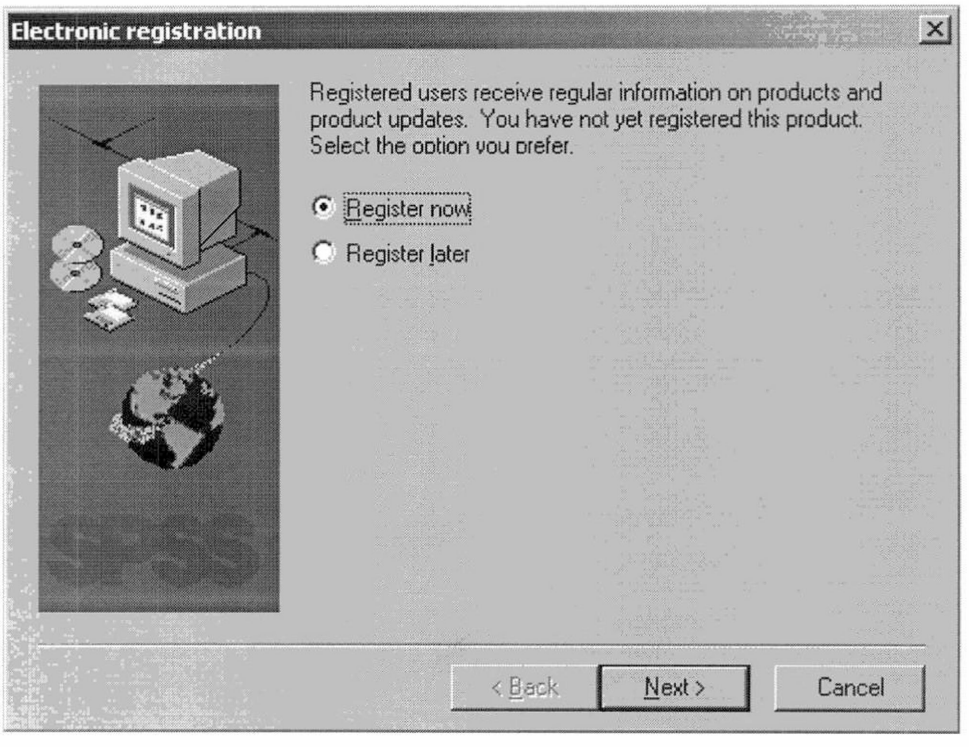
At this point the installation begins and you will see a screen like the screen above. Wait till the screen changes to the one shown below.



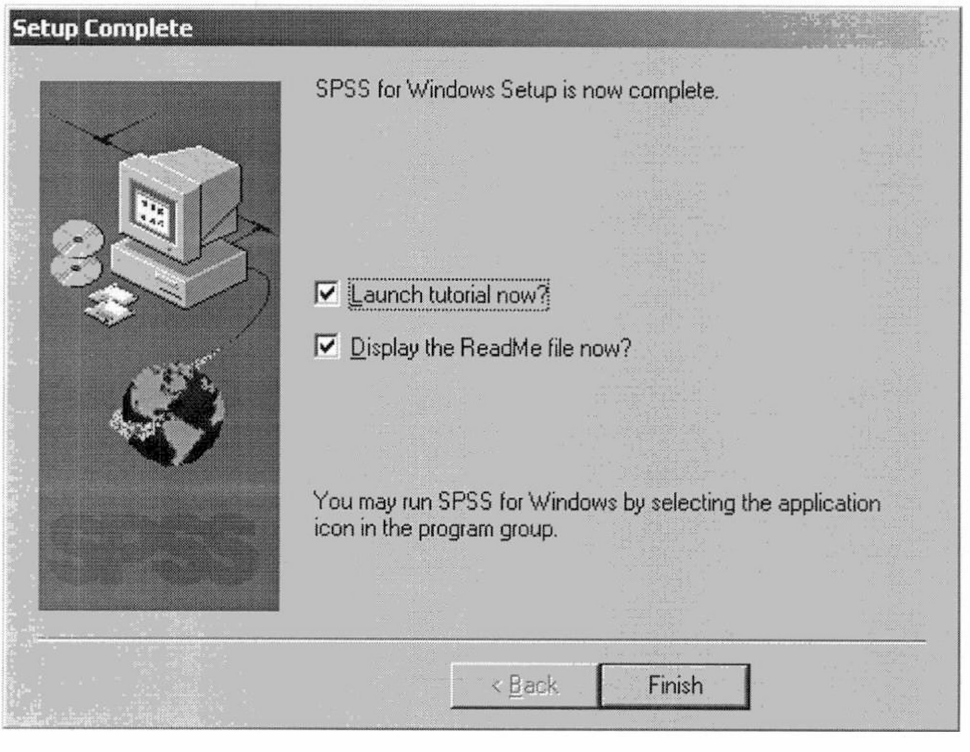
Click '**Next**'.



You will see an install screen like the one shown above. Once the installation is complete, you get a screen like the screen below.



Select the option '**Register later**' in the screen above and click '**Next**'.



Uncheck the '**Launch tutorial now?**' and the '**Display the ReadMe file now**" checkboxes and click on '**Finish**'.

SPSS 11.5 for Windows

Uninstall instructions:

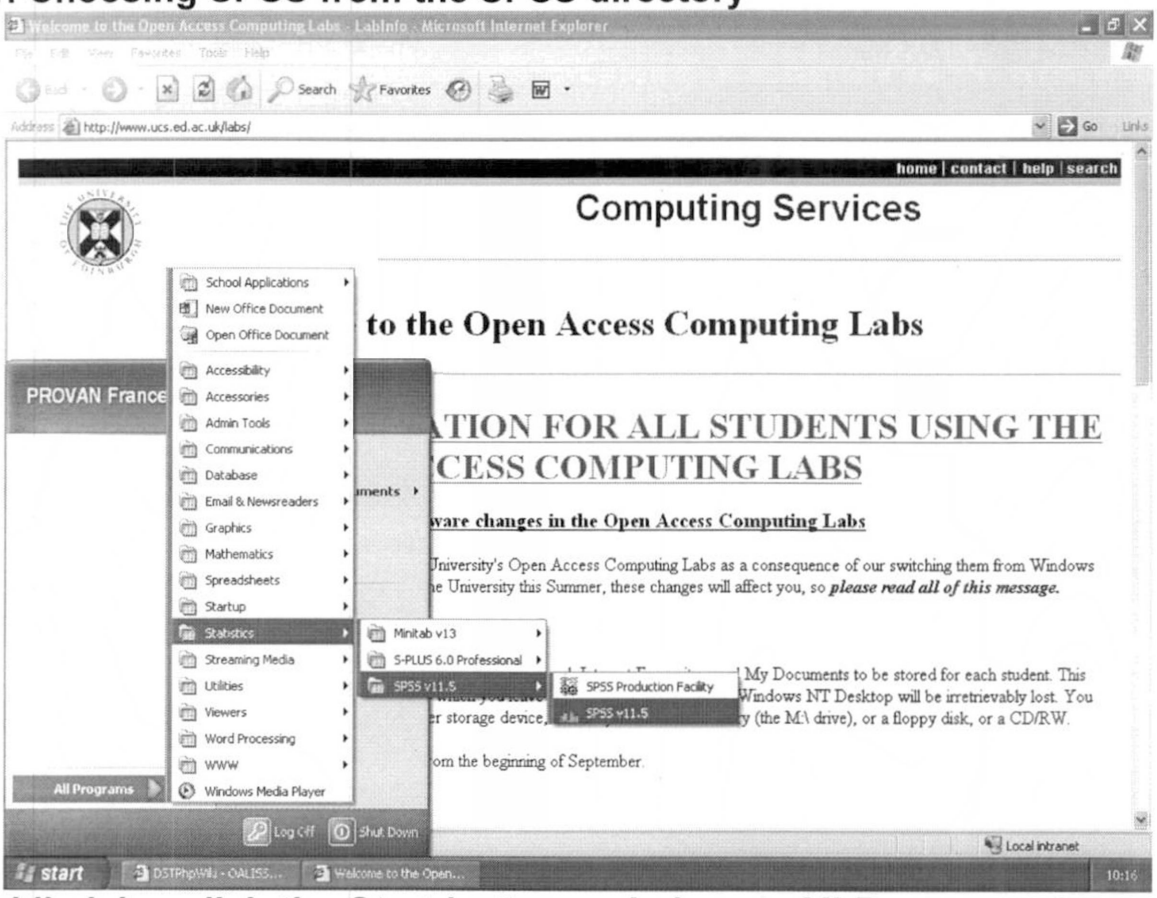
1. Select the following menu items to arrive at the dialogue box where you can uninstall the software package "**Start | Settings | Control Panel | Add/Remove Programs**".
2. Scroll down the listing and select the entry for "**SPSS 11.5 for Windows**"
    a. Click the Change/Remove button to active the "Remove programs from your Computer" wizard. The following components will be removed …
        i. Shared program files
        ii. Standard program file
        iii. Folder items
        iv. Program folders
        v. Program directories
        vi. Program registry entries
3. A progress bar indicates the progression of the removal task.
4. You will receive uninstall completed message. Some elements could not be removed. You should remove items related to the application.
5. Click on the "**Ok**" button. The entry in the "**Add/Remove Programs**" listing will disappear and disk space on the hard drive should increase by 82 megabytes.
6. Restart the computer.

Starting up SPSS

SPSS can be opened from **Start** menu. The button for the **Start** menu is usually at the bottom left hand corner of the desktop. **Pic 1** is a picture of a PC desktop

running Windows XP in the public labs. You can see from the **Start** menu, **All Programs** has been selected, from there **Statistics** has been chosen, and then from **Statistics**, **SPSS 11.5** has been chosen and then **SPSS 11.5** has been chosen from its menu.

Pic 1: Choosing SPSS from the SPSS directory



In Public labs, click the Start button and choose All Programs, from there choose Statistics, then choose SPSS 11.5 and SPSS 11.5 again to open SPSS.

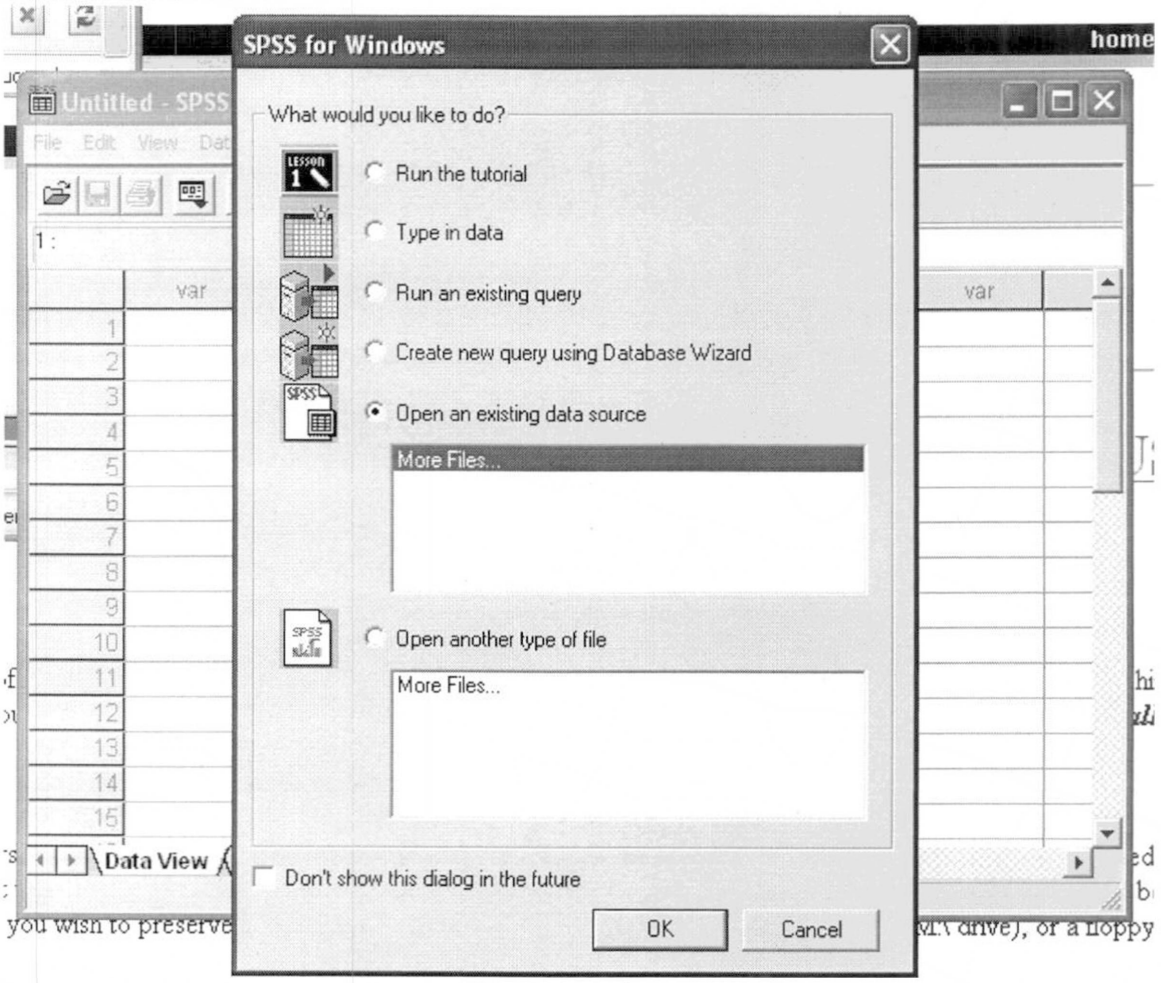
The exact sequence might vary depending on how the computer has been setup but SPSS should be available through the **All Programs** or **Programs** menu if it is there. If you can't find it, search for the SPSS directory using the **Start** menu choose

Search or **Find**, then select **All files or folders** or **Files!or!Folders** and search the computer for the word "spss".

SPSS for Windows dialog

Starting up SPSS will open the **SPSS for Windows** dialog box shown in **Pic 2**.

Pic 2: First view of SPSS



SPSS should look like this in the Public labs and Computing Services training suite and when it is first installed (unless SPSS has been customised on the computer you are using). There are several

choices of what to do next: -

- **Run the Tutorial** runs a slide show tutorial to show SPSS's main features.
- **Type in data** presents an empty untitled data window SPSS Data Editor.
- **Run an existing query** will run a query to bring data in from a database.
- **Create new query using Database Capture Wizard** takes you through creating your own query to extract data from a database or Excel.

- **Open an existing data source** to select an SPSS or other type of data file to analyse.
- **Open another type of file** to open any other sort of file.
- **Don't show this dialog box in the future** will stop the dialog box appearing in the future.

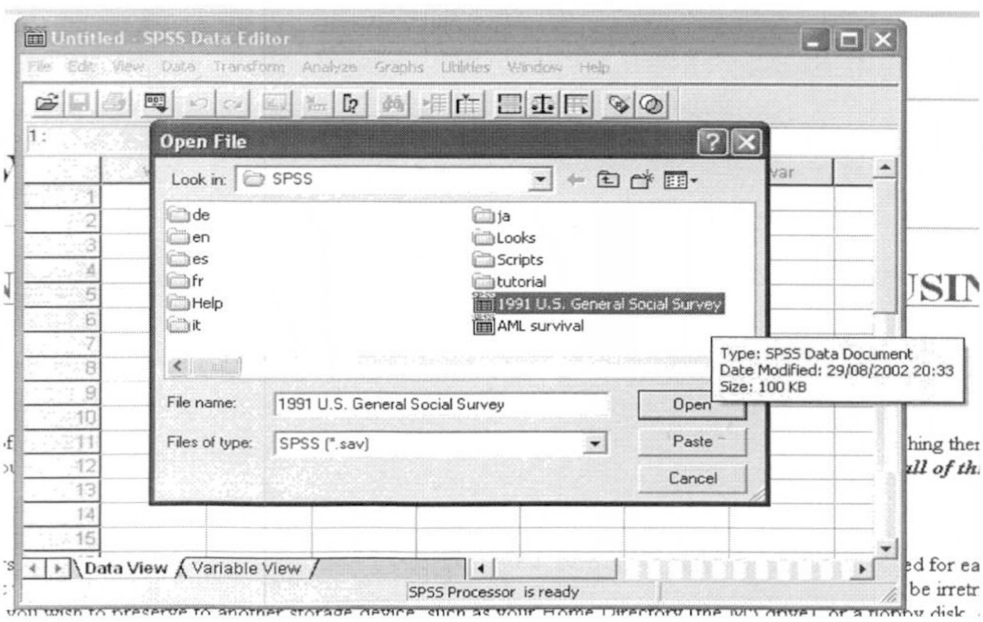
Click OK to choose Open an existing data source.

If you don't see the dialog box shown in Pic 2 then use the **Open File** button (first left on the tool bar, see Pic 9) or choose **File/Open/Data** from the menus to activate the **Open File** dialog box.

Starting up SPSS without the dialog box should open the SPSS Data Editor window, as in Pic 4 but without any data in it.

We will use a data set called **1991 U.S. General Social Survey.sav** that comes with SPSS and in general there should be a copy of it in the folder where SPSS is installed. In the labs it is drive T, in the Skills centre there is a copy in the spsswork directory as well. On your own computer it is likely to be on the C or D drive in the **Program Files** or **Programs** folder. You might have to scroll to the right past lots of folders before you get to the data files.

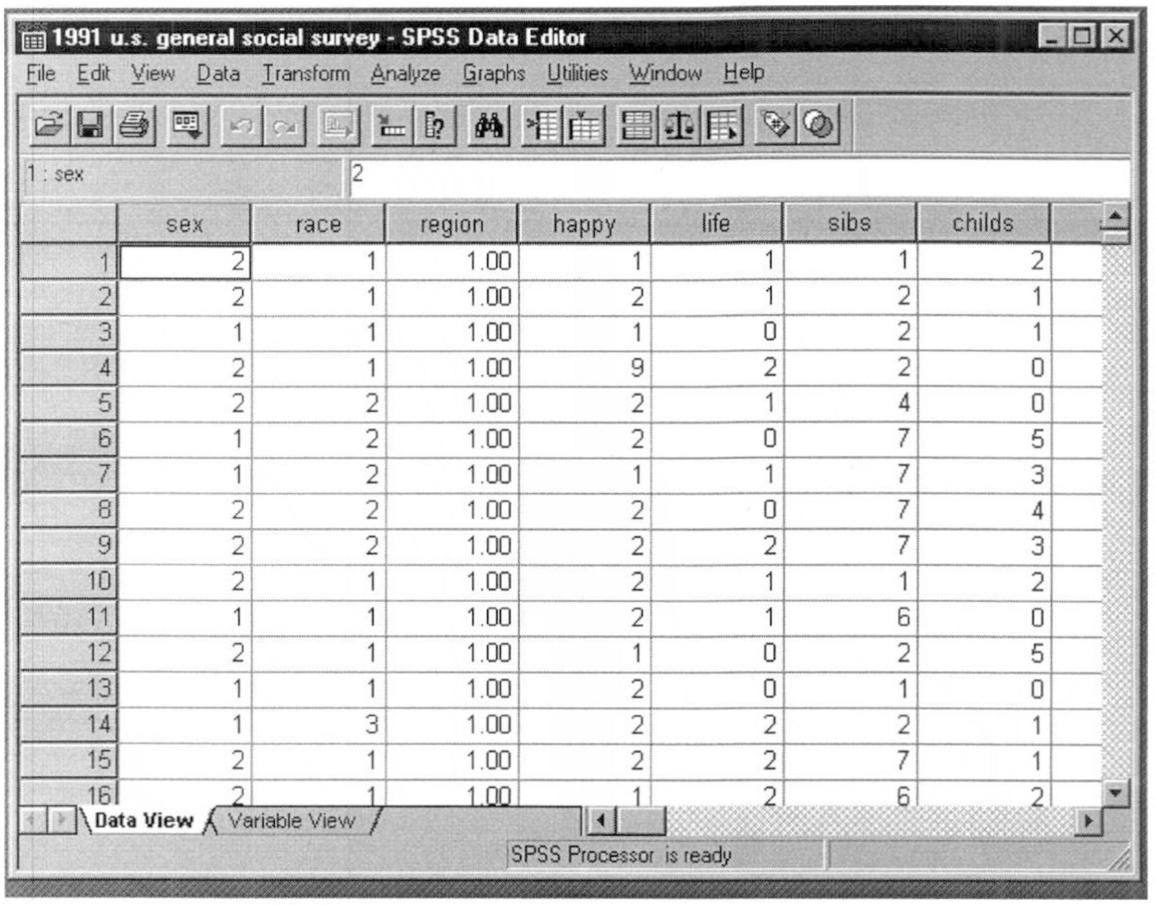
Pic 3: Open File dialog box showing data (.sav) files in the Spss11 folder



Select the file 1991 U.S. General Social Survey.sav and click Open to read the data file into the DataEditor window.

Data in the SPSS Data Editor

Pic 4: Survey data without value labels



	sex	race	region	happy	life	sibs	childs
1	2	1	1.00	1	1	1	2
2	2	1	1.00	2	1	2	1
3	1	1	1.00	1	0	2	1
4	2	1	1.00	9	2	2	0
5	2	2	1.00	2	1	4	0
6	1	2	1.00	2	0	7	5
7	1	2	1.00	1	1	7	3
8	2	2	1.00	2	0	7	4
9	2	2	1.00	2	2	7	3
10	2	1	1.00	2	1	1	2
11	1	1	1.00	2	1	6	0
12	2	1	1.00	1	0	2	5
13	1	1	1.00	2	0	1	0
14	1	3	1.00	2	2	2	1
15	2	1	1.00	2	2	7	1
16	2	1	1.00	1	2	6	2

Now you should be able to see the Data Editor window dominated by the grid used to display and change existing data or enter new data. Use the vertical and horizontal scroll bars to see more of the data, there are over 1500 rows representing the opinions of people from all over the states. Above the grid and below the title bar are the menu headings as seen in Pic 5.

Pic 5: SPSS Data Editor detail, Menu Headings

File Edit View Data Transform Analyze Graphs Utilities Window Help

Click each heading to look at the sort of items on each menu.

Most menu items are greyed at and cannot be selected until there is some data in the Data Editor.

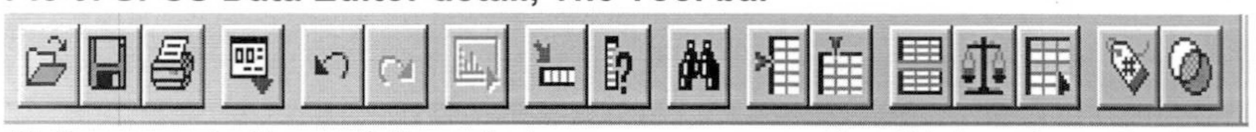
The **Help** menu contains lots of useful information about using SPSS including what all the menus and windows are for. Try **Help/Topics** and look at the various topics, including **SPSS at a glance**.

Help/Tutorial guides you through the basics of running SPSS – use it to supplement these notes.

Just below the menu headings is the **Tool bar**, its buttons are shortcuts for many common tasks in

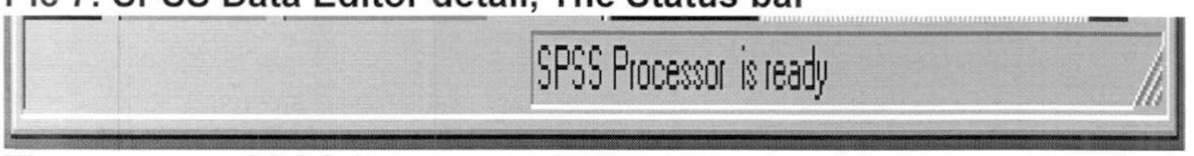
SPSS rather than using the menus.

Pic 6: SPSS Data Editor detail, The Tool bar



At the very bottom of the window you can see the **Status bar**. It shows the status of SPSS at the time, including the progress of any procedure and any special transformations in place (i.e. filters, splits or weights).

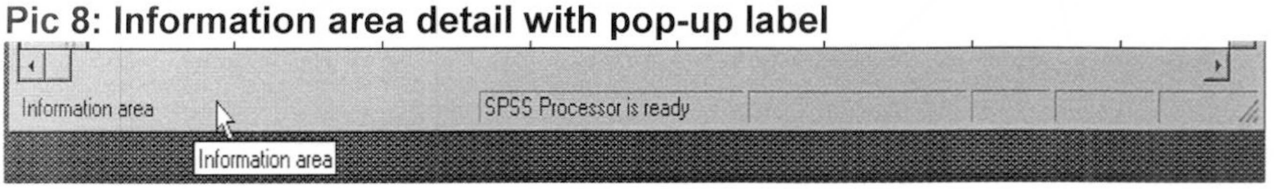
Pic 7: SPSS Data Editor detail, The Status bar



The message **SPSS Processor is ready** is showing in our example, indicating that SPSS is waiting for something to do. The first space on the status bar is known as the information area. It provides a

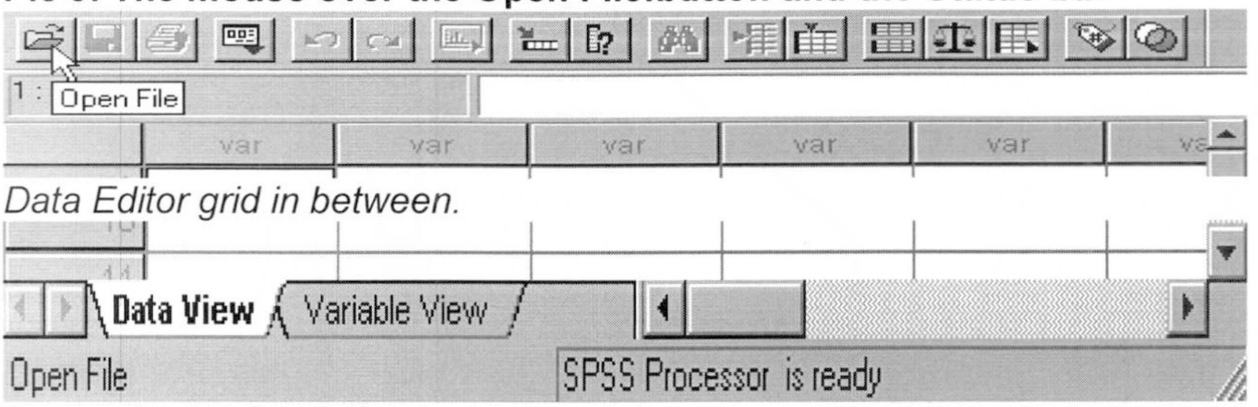
phrase and sometimes a pop up label to describe whatever the mouse is pointing to.

Pic 8: Information area detail with pop-up label



Place the mouse pointer over each button in the tool bar and watch the information area on the status bar at the bottom of the window as it changes.

Pic 9: The mouse over the Open File!button and the Status bar



Data Editor grid in between.

The contents of both the tool bar and the menus will change depending on what type of SPSS window you are looking at. Toolbars can be changed or new ones can be created in the **Show**

Toolbars dialog box, opened by choosing **View/Toolbars**. Menus can be customised using the **Menu Editor** opened by choosing **Utilities/Menu Editor**.

Data can either be typed directly into the **Data Editor**'s **Data View** or an existing data file can be read in. The Data Editor looks like a spreadsheet, but has a much more rigid data structure.

SPSS phrase book

CASE: Each row of the data editor represents a **case**. Cases are the units or individuals that make up the study, e.g. survey responses, products in an inventory, schools in a district, countries, plots in an agricultural experiment, etc.

VARIABLE: Each column of the data editor represents a **variable**. A variable is a particular type of measurement for each case, e.g. it could be someone's height, a code representing their social class or a recording of their opinion on some subject.

VARIABLE NAME: Above the data in the Data View, each variable has a **variable name.** In Pic 4, you can see the variable names sex, race, region,

happy, life and sibs above the data. Variable names can be up to eight characters in length and must start with a letter. There must be no spaces, full stops, apostrophes or other punctuation marks in the variable name.

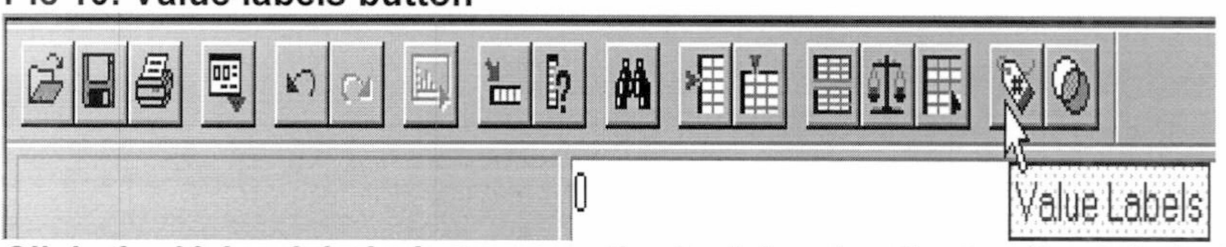
VARIABLE LABEL: Variables can also have a **variable label.** It is a longer bit of text to describe the variable. SPSS uses the variable label in the output when the variable is used and you can see what the label is when you point the mouse at the head of the variable as in Pic 13.

VARIABLE VALUE: Each cell in the Data View is known as a **variable value**. It represents the variable measurement for that particular case or row. These variable values can be numbers, numbers representing codes (nominal/ordinal) or text. SPSS

assumes that the data is numeric unless it is told otherwise.

VALUE LABEL: Sometimes a piece of text called a **value label** is associated with each variable value. You should see the data with the value labels switched off. They can be displayed by using the **Value labels** button situated second from the right on the tool bar.

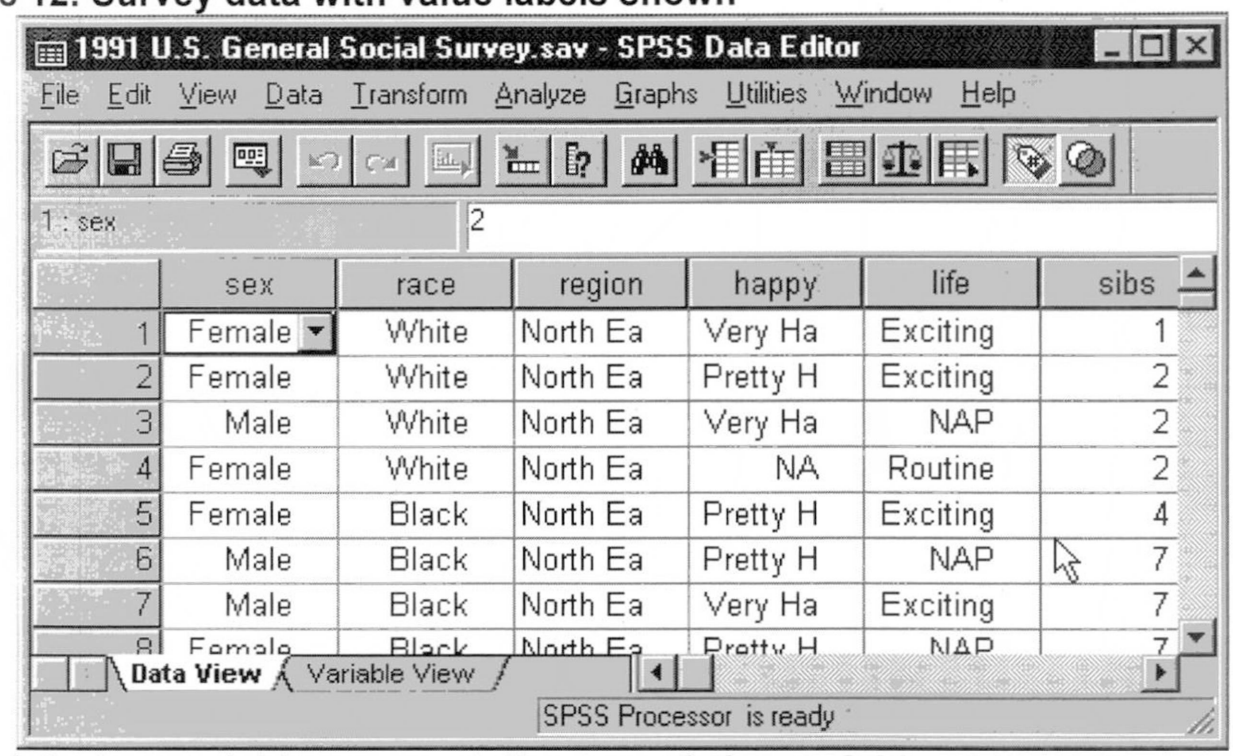
Pic 10: Value labels button



Click the Value labels button on the tool bar to display the labels if you cannot see them.

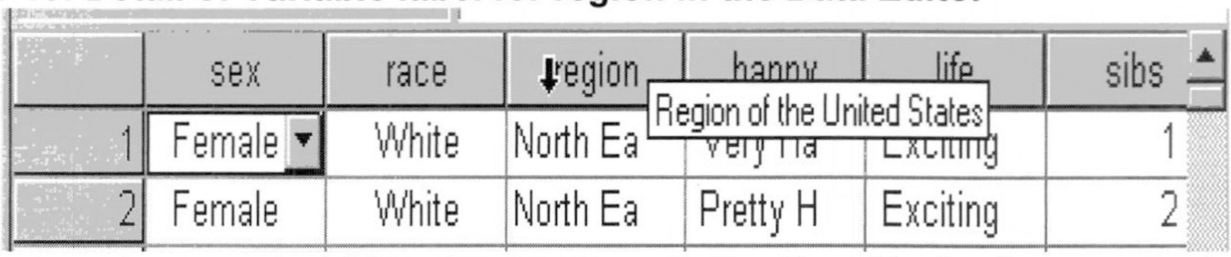
Click on Value Labels button again so you can see the labels.

Pic 12: Survey data with value labels shown



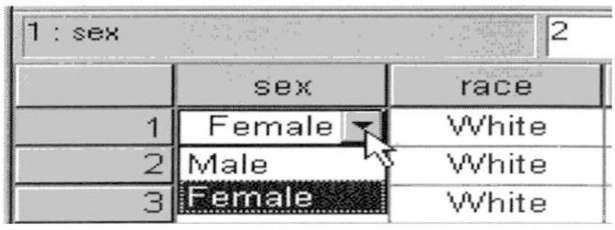
Point at a variable name at the top of the variable column to see its variable label, as shown below in Pic 13.

Pic 13: Detail of Variable label for region in the Data Editor



You can change a variable value just by selecting it and typing in a new code. If **Value Labels** are switched on, you can also change it by choosing another label from the list of value labels or more tediously by typing in the whole value label exactly as it appears in the list.

Pic 14: The list of value labels for the sex variable



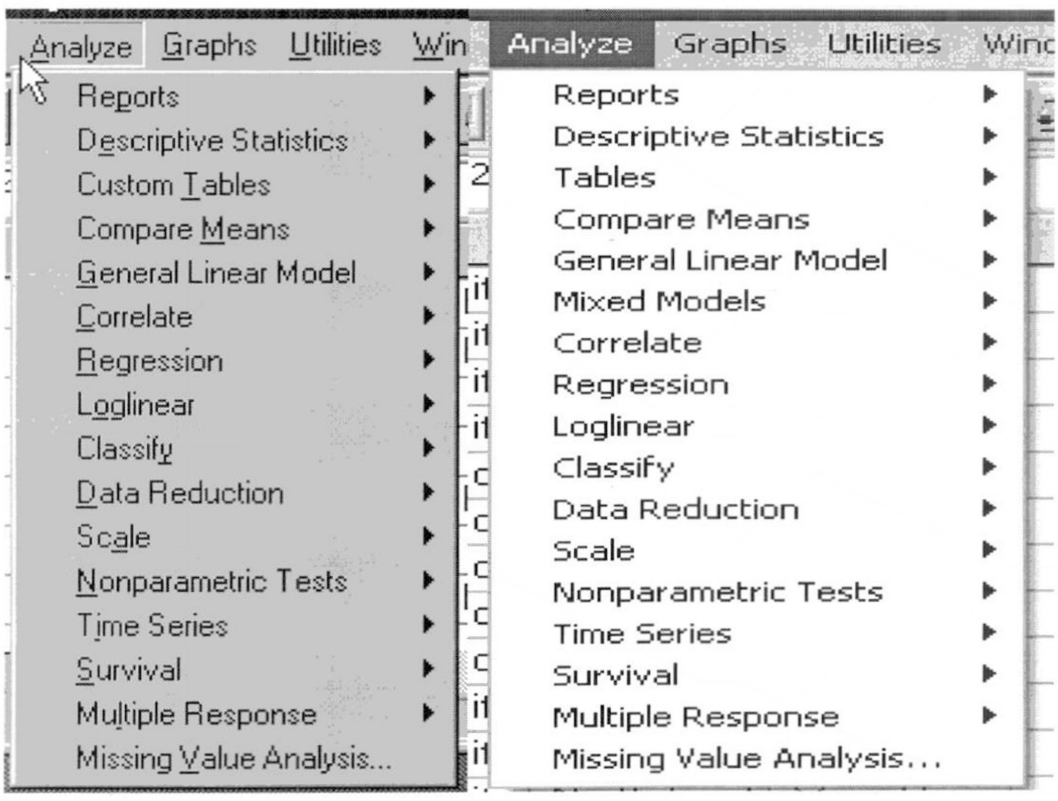
Select a cell in the Data Editor and click on the small arrow beside it to see the list of labels defined for that variable.

If you change the data remember to return the data to its original state before going on or use the **File /Open** menu to open the original system file, clicking on **No** when it asks you to save changes to the data.

Running an analysis

Now to look at how you would run an analysis in SPSS. The **Analyze** menu contains the statistical analyses available. We will run through one procedure that deals with variables individually (univariate analysis) and another that deals for a tabulation using two variables together. From there you should be able to apply the same principles to other analyses.

Pic 15 & Pic 15a: Different sorts of statistical analysis available for SPSS 10 and SPSS 11.5



Choose Analyze from the menus to see what is available.

An arrow beside a menu item indicates that it has a sub menu. Notice all the items in the **Analyze** menu have submenus. The list of menu items shown here is a fairly standard SPSS installation.

Move the mouse over each arrow to see a list of the procedures available in each sub menu. Other items may appear on this menu depending on what modules are installed and whether the menus have been customised.

From the Analyze menu, choose Descriptive Statistics to see a list of the descriptive procedures.

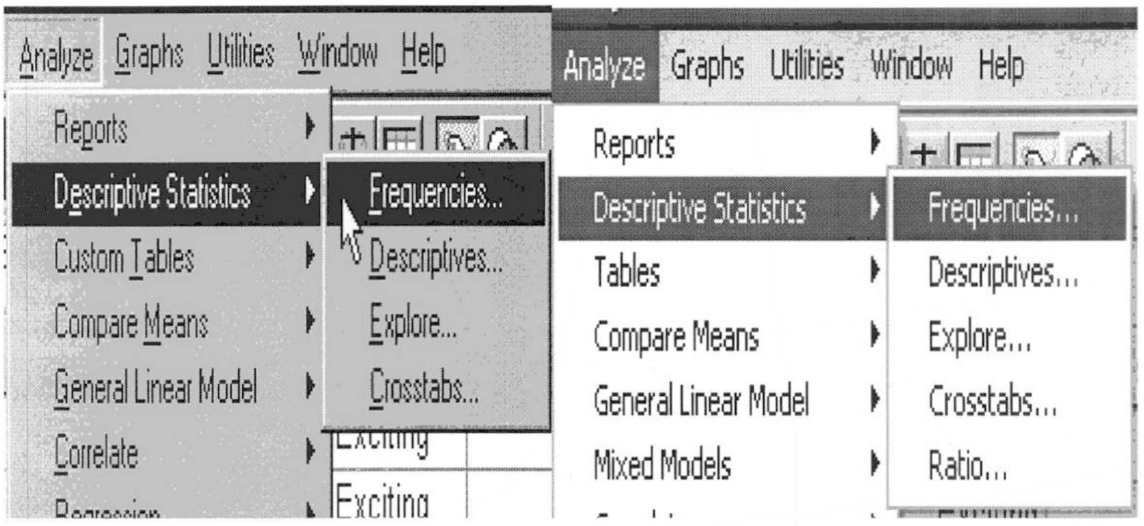
These notes are intended to show you how to use SPSS, not to teach you statistics.

There is a **Statistics Coach** in the **Help** menu to help you choose the correct analysis for your data. The coach is not very detailed so get take advantage on any statistics courses your department may run or get a good introductory statistics book.

Frequencies

The **Frequencies** procedure generates summary tables of frequencies of each value for a variable. It can also be used to calculate summary statistics and charts for each variable individually as well.

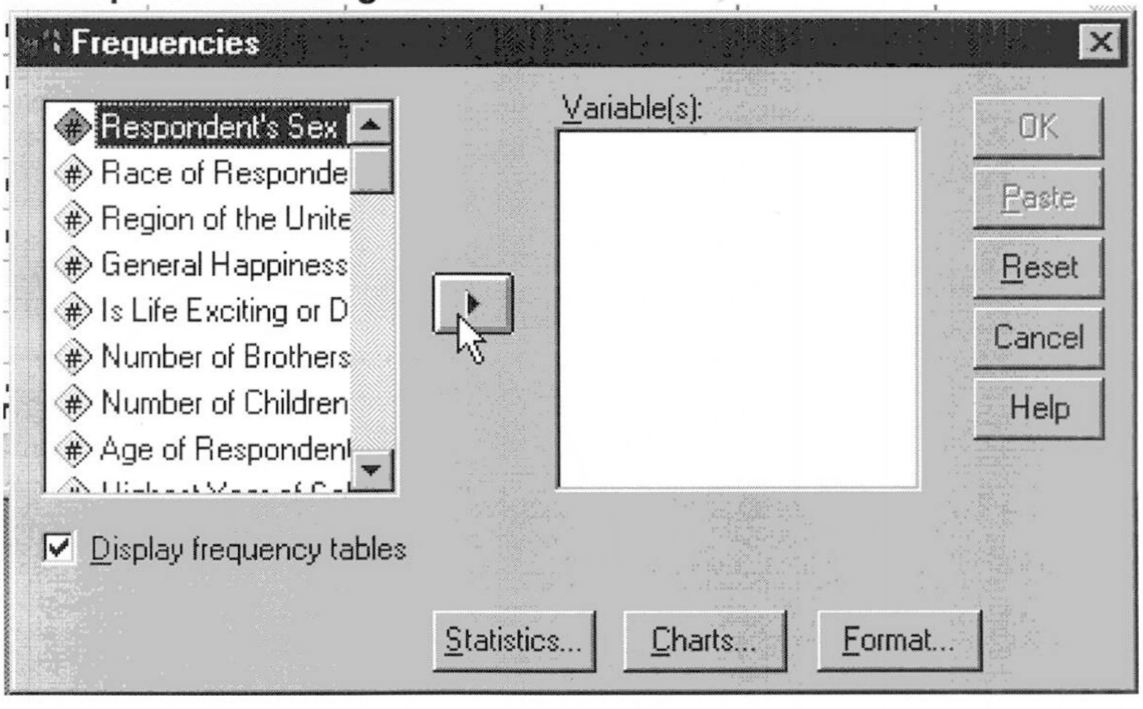
Pic 16 & Pic 16a: Items from the Analyze/ Descriptive Statistics submenu for SPSS 10 and SPSS11.5



Now choose Analyze/Descriptive Statistics/Frequencies.

The **Frequencies** dialog box consists of two boxes of variable lists, below them are buttons to customise the procedure results and to the right of the lists are the five standard dialog buttons. To produce frequencies tables for a variable simply select it from the left hand list of all the variables in the data set and click on the arrow button to put it in the **Variable(s):** box.

Pic 17: Frequencies dialog box



The three buttons along the bottom of the dialog box are for making additions or changes to the default output. **Statistics...** is used to add summary statistics for each variable in the **Variable(s)** list. **Charts...** is used to produce a graph for each variable in the **Variable(s)** list. You can choose between a bar chart, pie chart or histogram. **Format...** is used to change the format of the standard frequency table and the order objects are displayed in.

The five buttons down the right hand side of the box are common to most dialog boxes.

Clicking **OK** tells SPSS that the dialog box is complete and is to be acted on.

Paste will paste the underlying SPSS command into a syntax window - the SPSS commands

course deals with this in more detail.

Reset can be used to clear the dialog box of any selections made so you can start specifying

the procedure from scratch.

Cancel is used to exit from the dialog box without running the procedure.

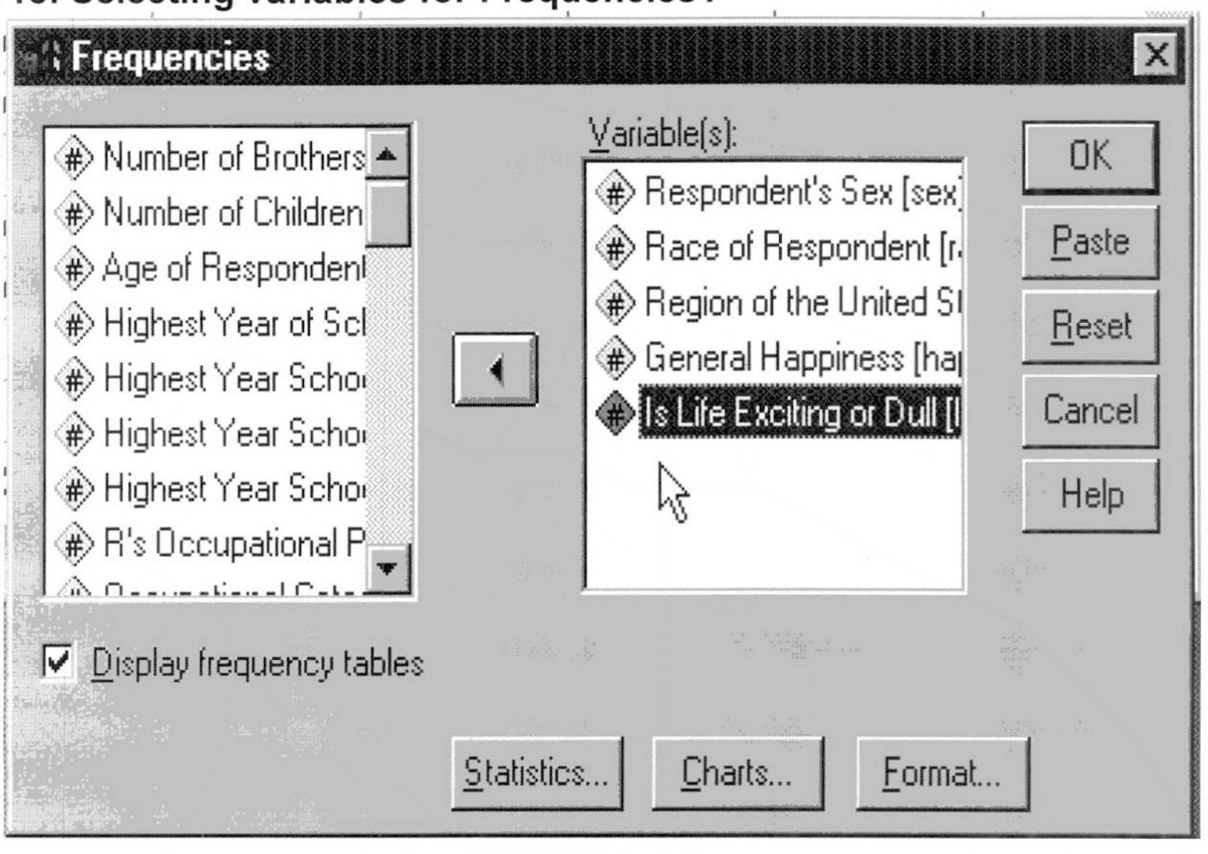
Help has information about the procedure and how to use the dialog box.

Above in Pic 17 the first variable in the list, Respondent's Sex (gender), has been selected and is

about to be put in the **Variable(s):** box using the arrow button.

Put a few categorical variables in the Variable(s) list, e.g. variables sex, race, region, happy and life were chosen in Pic 18.

Pic 18: Selecting variables for Frequencies1



Double clicking on the variable will also transfer it into the **Variable(s):** box. Once there, if you decide not to use a variable select it in the **Variable(s):** list and move it back using the arrow button which should be pointing back to the original list or double click on it again.

Notice that the **OK** button (and **Paste** button) is active now in Pic 18 – as soon as SPSS has enough information to run a procedure it will activate the **OK** button. By comparison the OK button in Pic 17 is inactive and looks faded or "greyed out". You could run the **Frequencies** procedure just now if you wanted to2. However we will look at the procedure in a bit more detail before we run it, just

to heighten the sense of anticipation!

1 Each variable has a small icon beside it, showing **#** for a numeric variable (i.e. values are numbers). A string or text variable will show **A<** or **A>** for ones over 8 characters wide.

2 If you do run Frequencies at the moment, then choose **Analyze/ Descriptive Statistics/ Frequencies** from the menus to continue with the next section. Page 19 will tell you about the output.

More information on procedures

Most Windows applications have a Help facility where you can get more information. It is usually accessible from the **Help** menu or by a **Help** button in SPSS dialog boxes. The **Help** button is "context sensitive", i.e. in the **Frequencies** dialog box it points directly to the information about **Frequencies** in the Help system. Any time you lose track of what you are doing in a dialog box, use the Help button or the Help menu to find out more about where you are.

Click Help button and SPSS will tell you more about Frequencies.

Pic 19: Frequencies information in the Help window in SPSS 11.5

Click Show me to open a tutorial about Frequencies or click a Related Topics link to find out more. In SPSS 10 the See Also button in the Help window will show you related topics.

Once you've finished reading, close the Help (and Tutorial) window and get back to SPSS and the Frequencies dialog box.

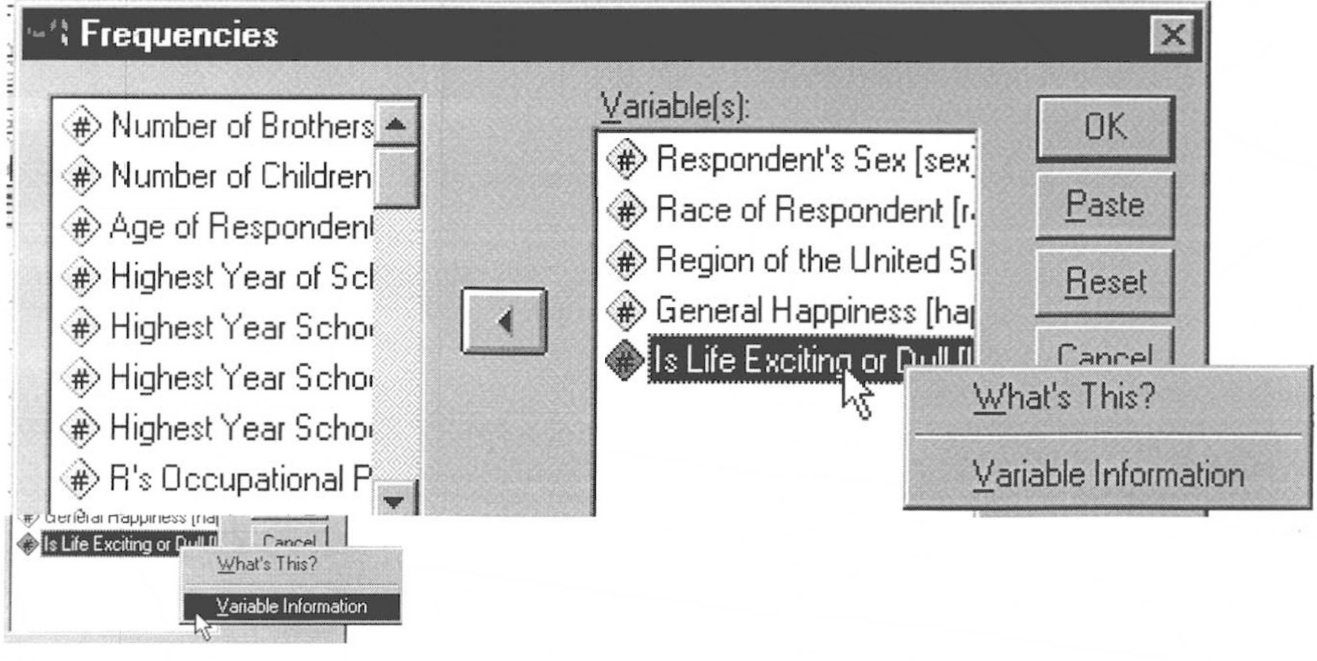
Context Sensitive Menus

A "context sensitive" menu can be activated for an object in Windows by selecting it and clicking on it with the right hand mouse button. The contents of menu will depend on the thing you selected, showing what you can do in that context. Normally one of the items is called "**What's This?**" or "**Properties**" and it will tell you about the selected object.

Information about variables

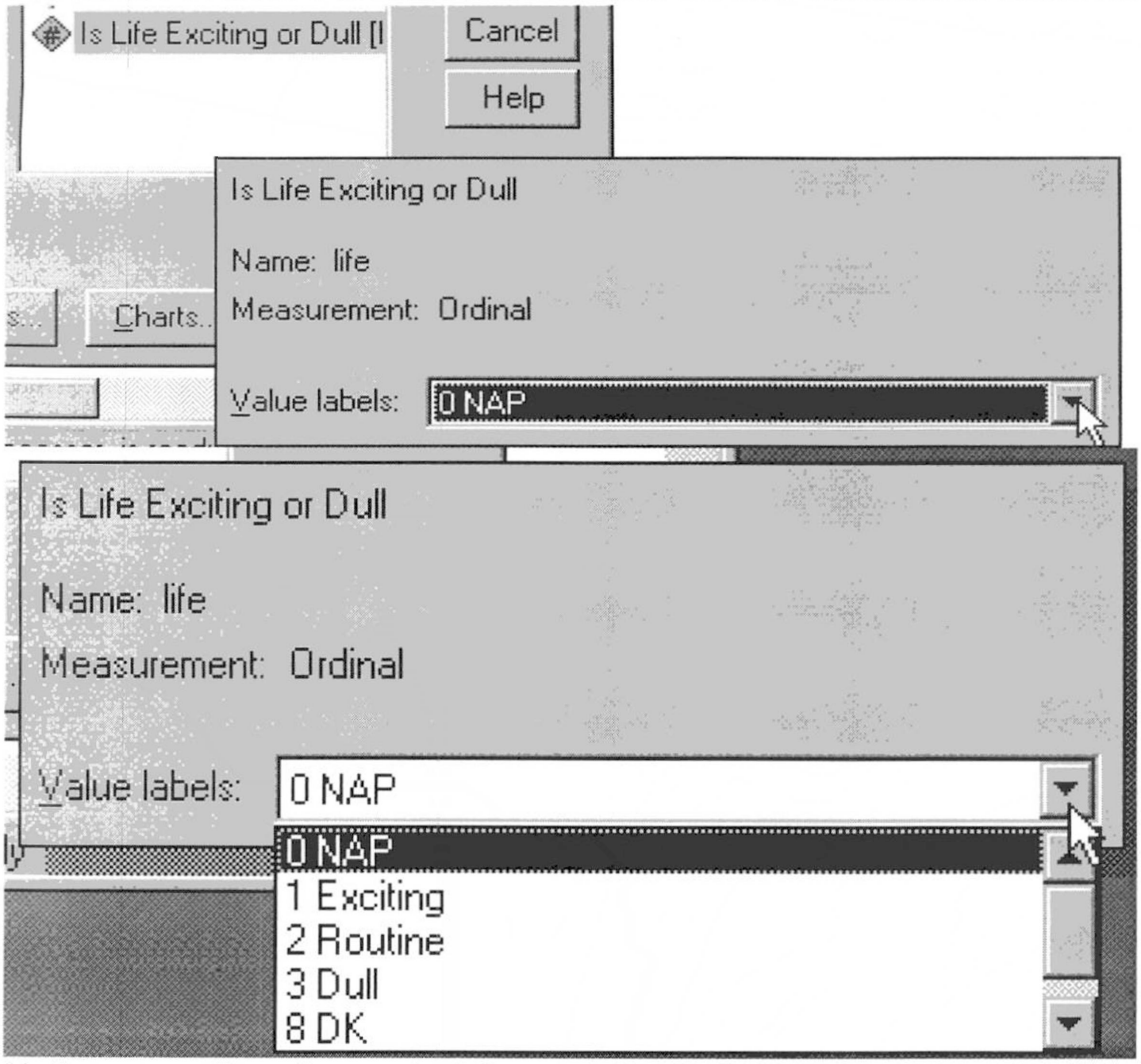
A variable information box is available via a context sensitive pop-up menu from SPSS dialog boxes. To do the following make sure the **Frequencies** dialog box is open.

Pic 20 & Pic 20a: Pop-up menu & choosing Variable Information



Click on one or two variables in either list using the right mouse button to display a pop-up menu similar to Pic 20. Choose Variable Information from the menu to see information about it.

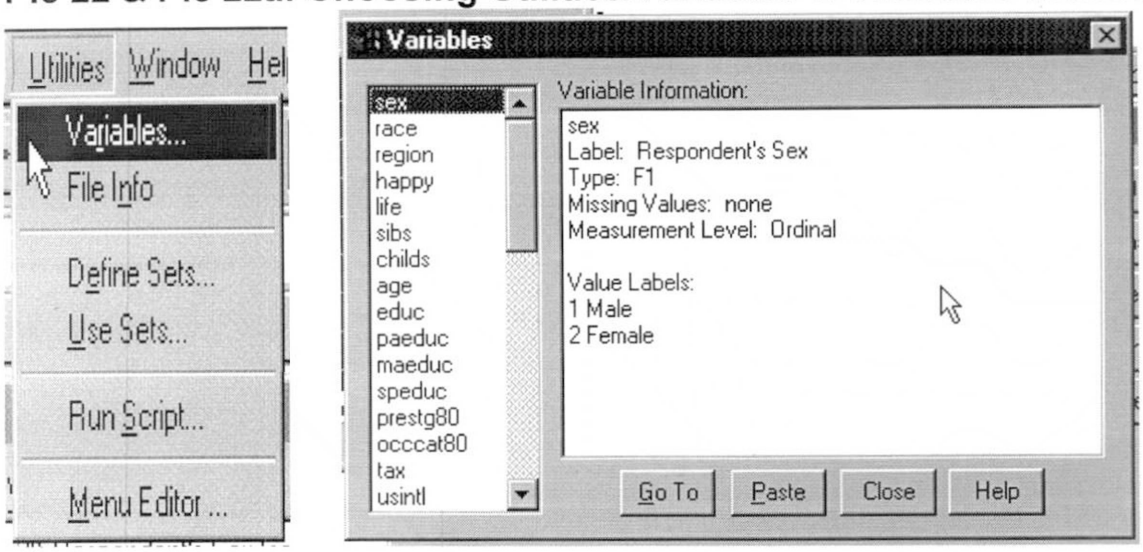
Pic 21 & Pic 21a: Variable information for life variable & value labels



Notice as well as the rest of the variable definition there is a drop-down menu of defined value labels. If you click (left mouse click this time) on the dialog box in the background then the variable information will disappear. The other way you can get information about the variables is by selecting **Variables...** from the **Utilities** menu to open the **Variables** window.

Choose Utilities/Variables... to see the Variables window.

Pic 22 & Pic 22a: Choosing Utilities/Variables & Variables window



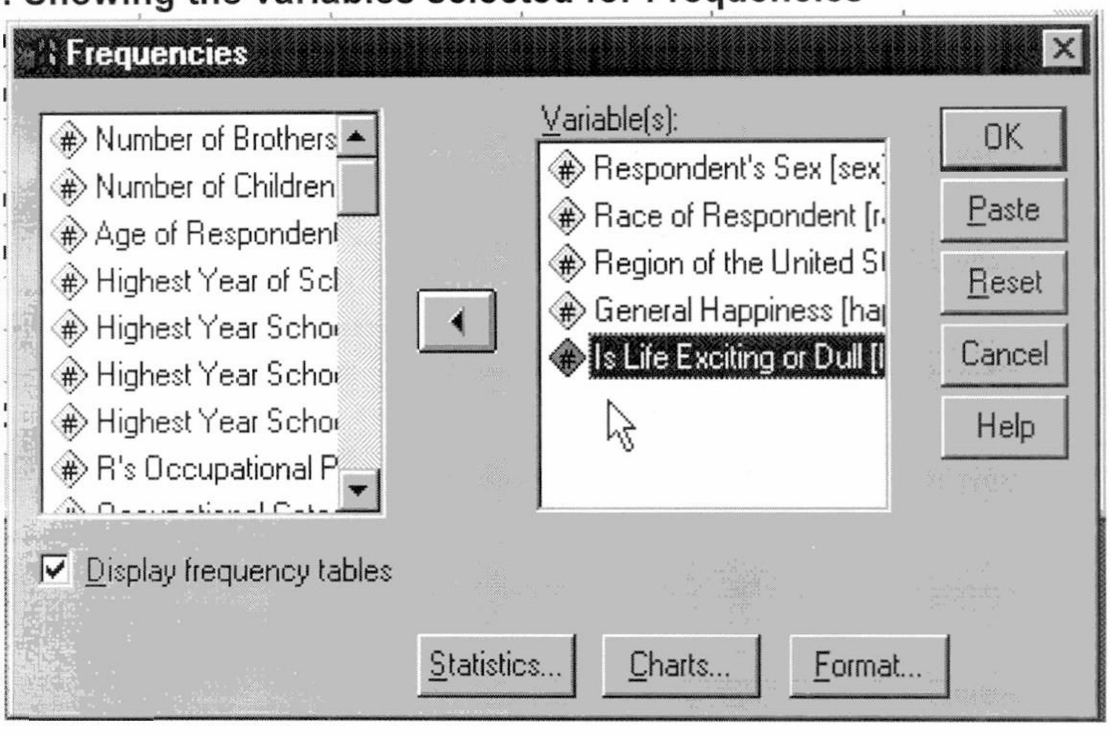
The **Variables** window contains the same sort of information in Pic 21, but in a window rather than a pop-up menu. The window can remain open in the background while you work elsewhere in SPSS. It can also be opened at practically anytime for reference in SPSS.

Click the Close button to close the Variables window, so the Frequencies dialog box is at the front.

We will now look at some ways we can customise the Frequencies output using the **Charts...** and **Format...** buttons. Most of the statistics available via **Statistics...**button are more suitable for scale

variables than the categorical ones we have chosen – so we will not bother with them at the moment. If you've lost the **Frequencies** dialog box, choose **Analyze/ Descriptive Statistics/Frequencies** from the menus again and SPSS will bring it back to the front. Just in case you have forgotten what it looks like, Pic 23 (a copy of Pic 18) shows the dialog box before we got diverted.

Pic 23: Showing the variables selected for Frequencies

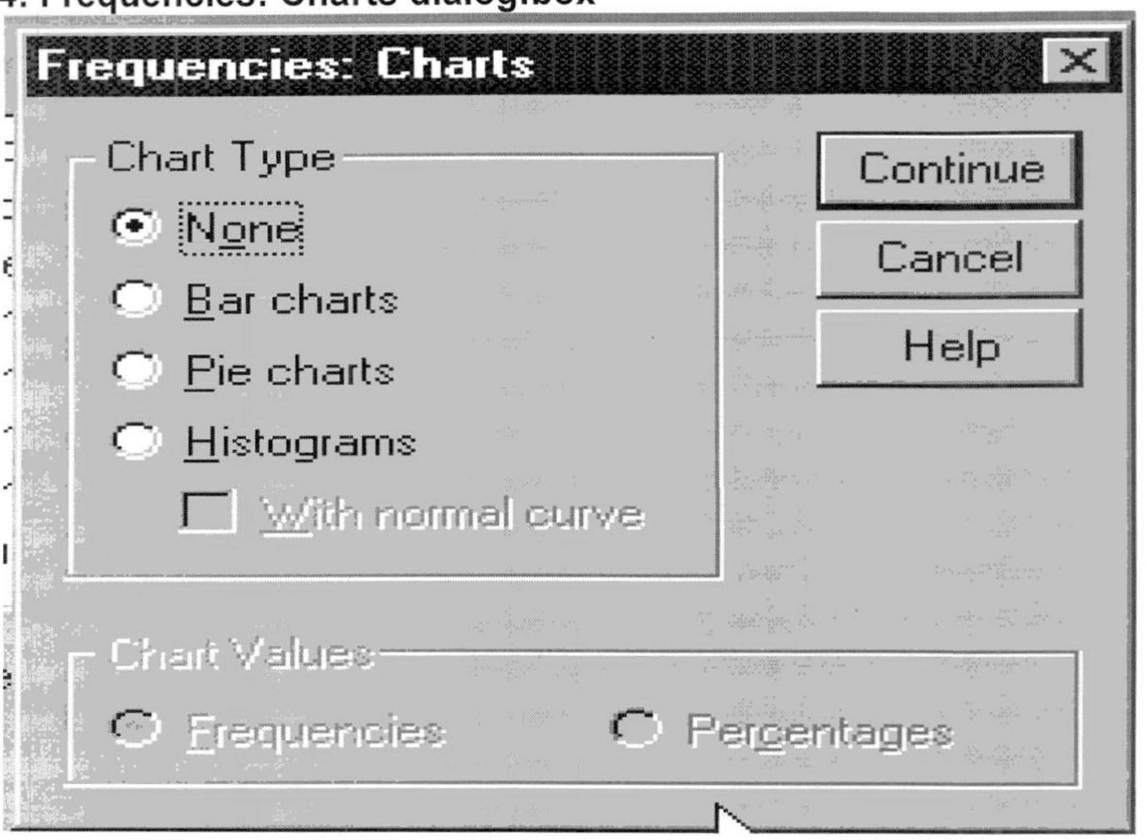


Frequencies: Charts, Specifying a bar chart

The **Charts...** button on the **Frequencies** dialog box opens the **Frequencies: Charts** dialog box where you can choose to produce graphical equivalents of the frequency tables. There are three sorts of chart available. **Bar charts** and **Pie Charts** are best for variables with few distinct values, i.e. categorical (nominal or ordinal) variables. **Histograms** are best for variables with lots of values measured on a scale, i.e. quantitative variables.

Click on the Charts... button to bring up the Frequencies: Charts box.

Pic 24: Frequencies: Charts dialog!box



Click on Bar Charts to produce bar charts with the tables.
Then click Continue to go back to the main Frequencies dialog box.
Make sure you selected **Bar Charts** and then clicked **Continue**. If you click **Cancel** or otherwise close the dialog box SPSS will close the **Frequencies:!Chart** box without saving your selection and no chart will be specified.

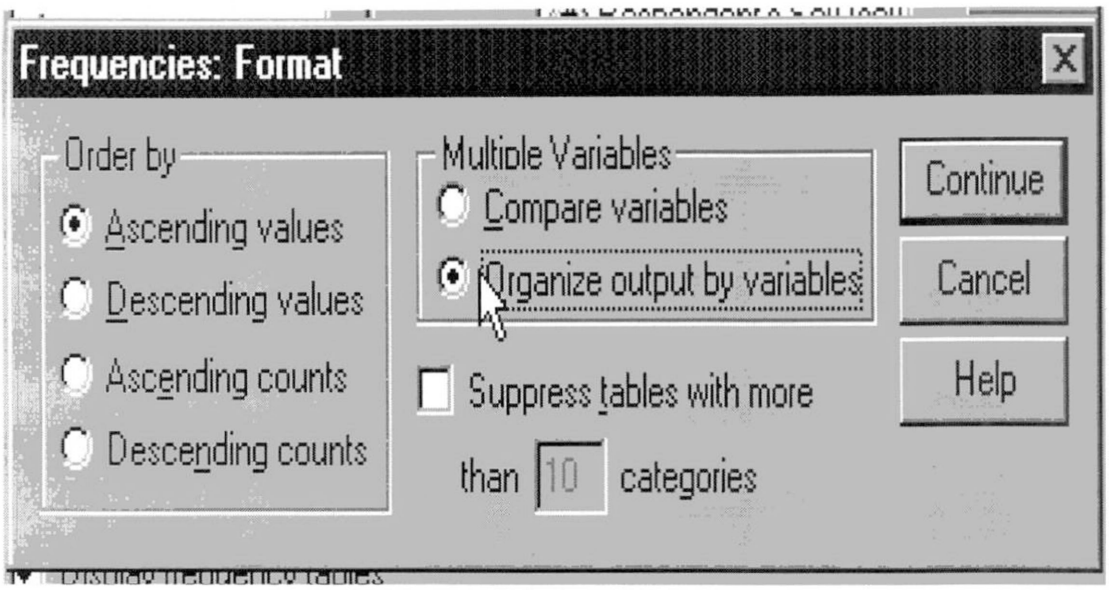
Frequencies Format

The **Format...** button opens the **Frequencies: Format** dialog box. It contains options to change the order of the Frequencies output or even suppress the output under certain conditions. We will change the setting and use **Organize output by variables** so we can see the frequency table for a variable beside its pie chart.

Click the Format... button and choose Organize output by variables button.

Click Continue to get back to the Frequencies dialog box to run it.

Pic 25: Selecting Organize output by variables



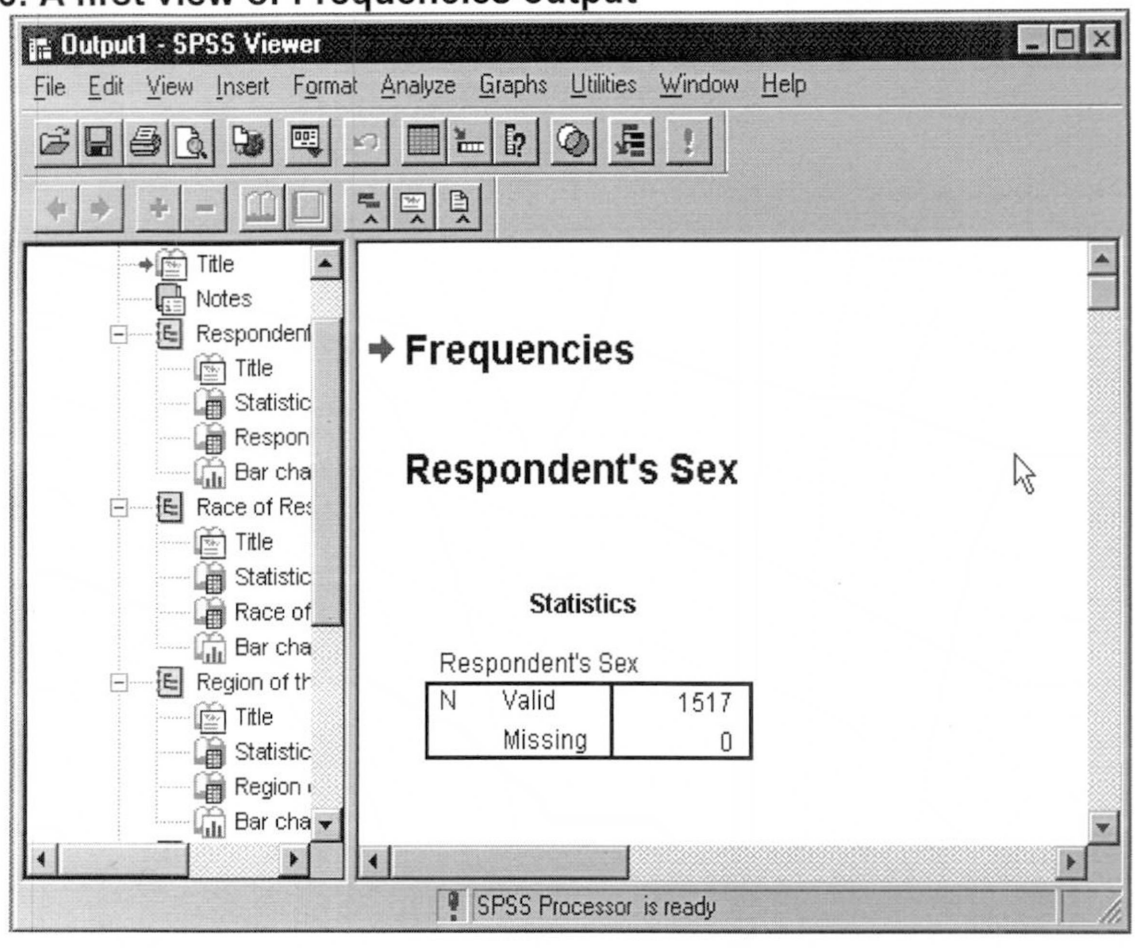
Click the OK button in the Frequencies dialog box to run it.

Again make sure you click **Continue** after making your selection to save your selection.

The output viewer

As soon as you click **OK** you should see the SPSS Output viewer come to the front. And depending on how fast the computer is, you may see the status bar at the bottom change to **Running FREQUENCIES....** before the results are written to the Output window. It should change back to **SPSS Processor is ready** as illustrated below when the procedure has finished.

Pic 26: A first view of Frequencies output



The **SPSS Output Viewer** will usually come to the front as soon as a statistical procedure is run. The Output window is split vertically into two panes. Above the split window, the menus have changed and different buttons are active on the tool bar to reflect the fact they are for use with different sort of window. The two "panes" of the viewer each have separate scroll bars which means you can scroll down the results or you can scroll through the outline or click on an icon to jump to that bit of the results.

The **left** side of the output window is the viewer outline, consisting of small icons representing the different parts of the SPSS output.

The **right** hand side contains the output itself.

In Pic 26 you can see the start of frequency output with the default summary statistics table for the first variable from our list, **Respondent's Sex**. **N Valid** is

the number of cases with valid values, i.e. there are 1517 where gender has been recorded and **N Missing** is the number of missing cases, where there is no valid value, none in this example. If we had used the **Statistics...** button to add more summary statistics they would appear in this table.

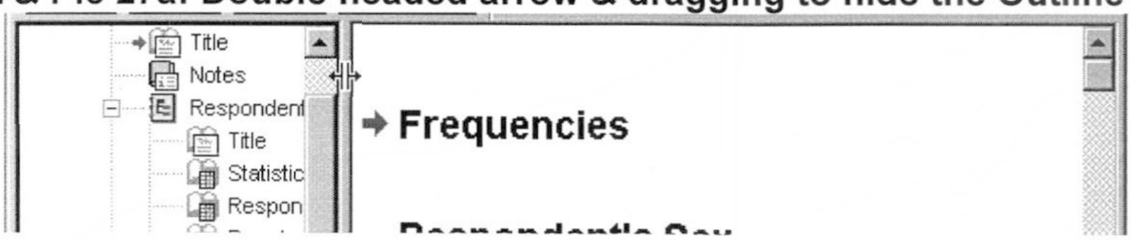
Hiding the outline

The outline side of the output window can be easily hidden to give you a more space on the screen

to inspect the results.

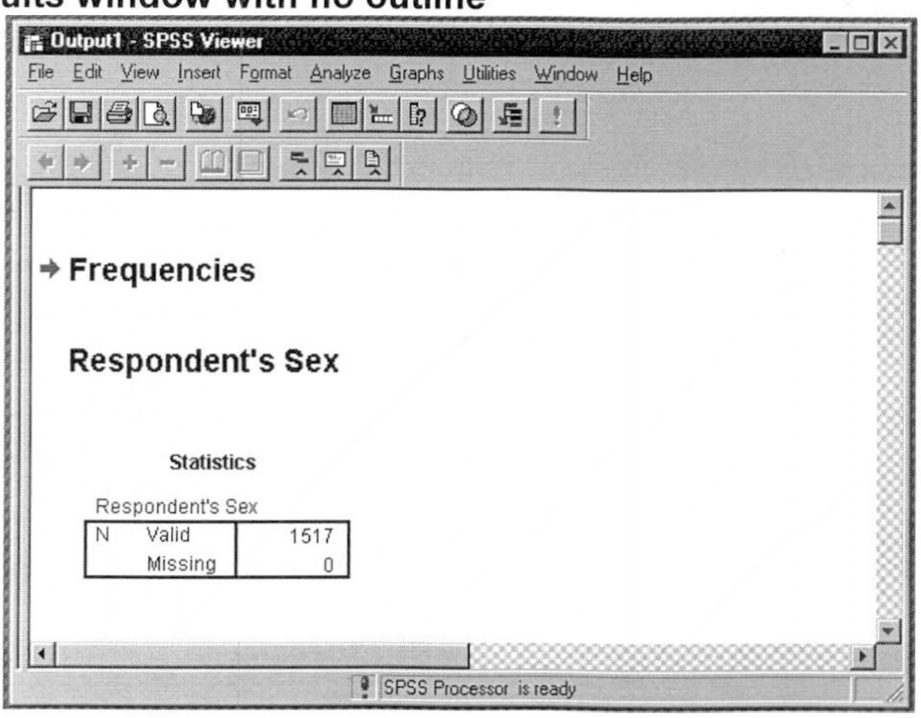
Place the mouse pointer over the divide between the two parts of the window. You should see the pointer change to a double-headed arrow.

Pic 27& Pic 27a: Double headed arrow & dragging to hide the Outline



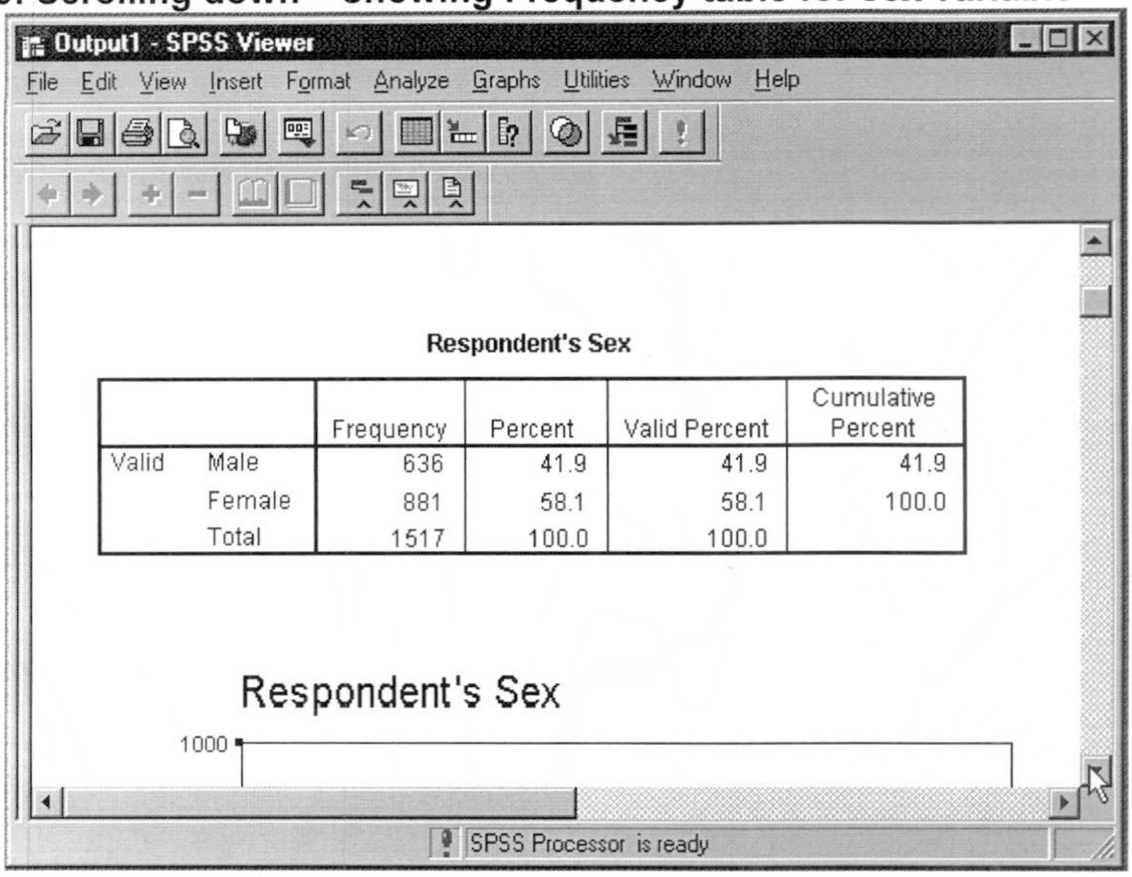
Click and drag the division over to the left most point and the outline will be hidden.

Pic 28: Results window with no outline



Use the scroll bars to inspect the rest of the output.

Pic 29: Scrolling down – showing Frequency table for sex variable



The outline can be seen again by clicking on the left hand side of the window and dragging the division bar towards the middle of the window. (mouse should be the double headed arrow again)

This method can also be used to change the relative size of the panes by dragging the bar between them but just not all the way across.

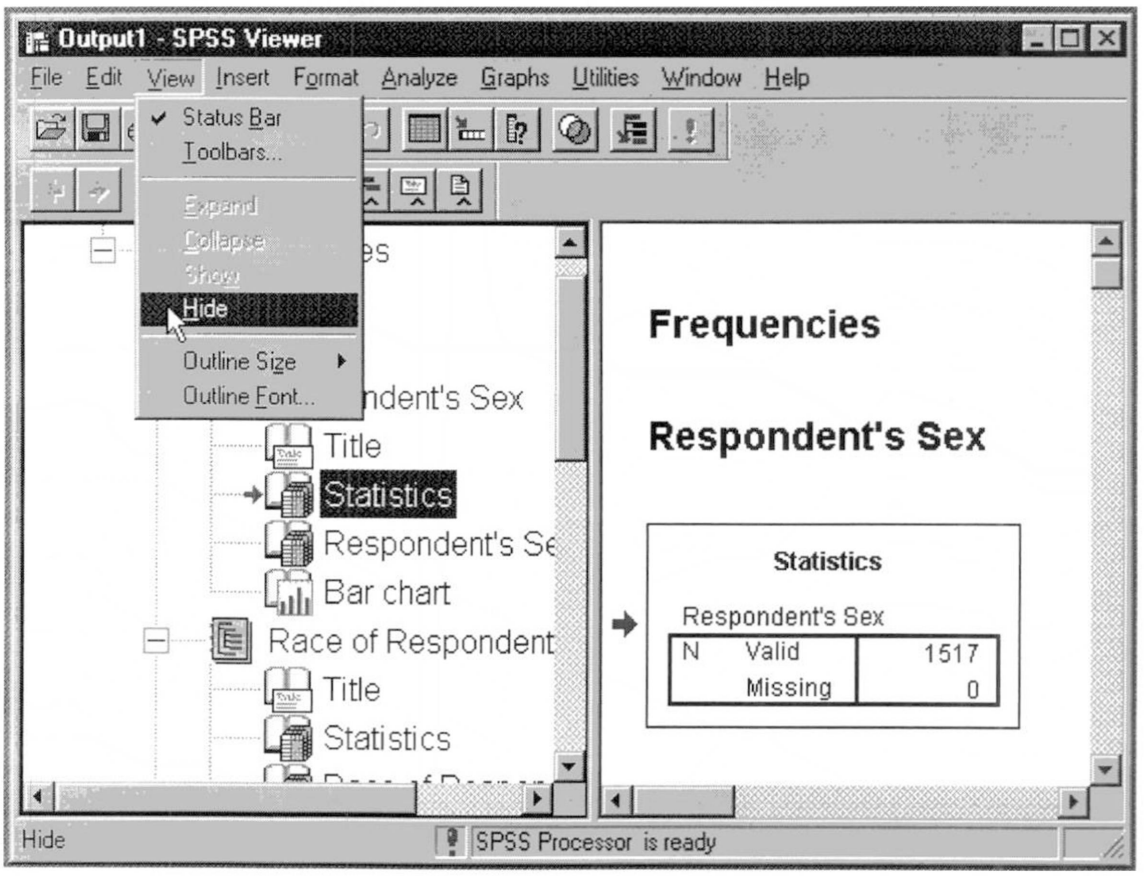
Restore the outline by clicking on the left hand side of the window and dragging the division bar into the middle again.

Using the outline

The outline can be used to organise the results you want to see to jump to a particular section of output. It is possible to hide items in the results by selecting it in the outline and "hiding" them.

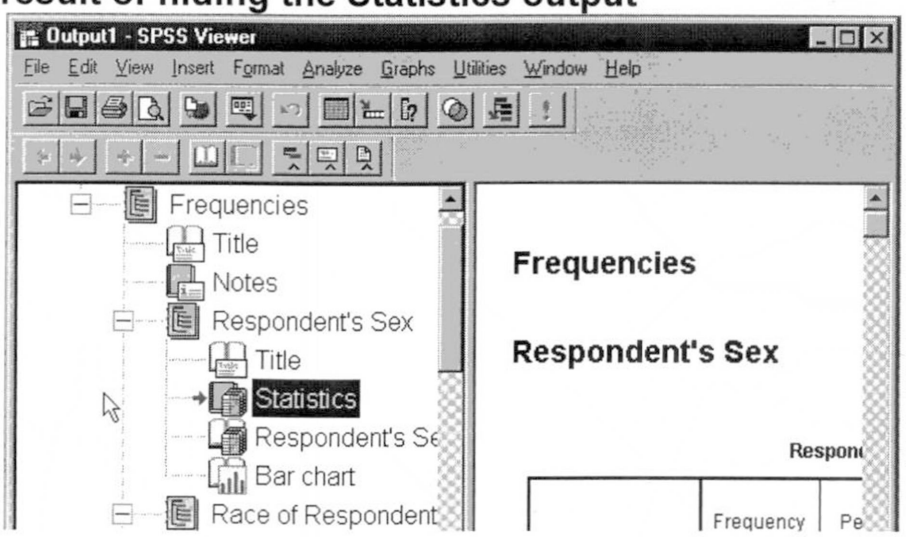
Hide a selected icon by choosing **View/Hide** from the menus or double clicking on the outline icon. Click once to go to the output the icon represents and double click to hide the output.

Pic 30: Using the View menu to hide some output

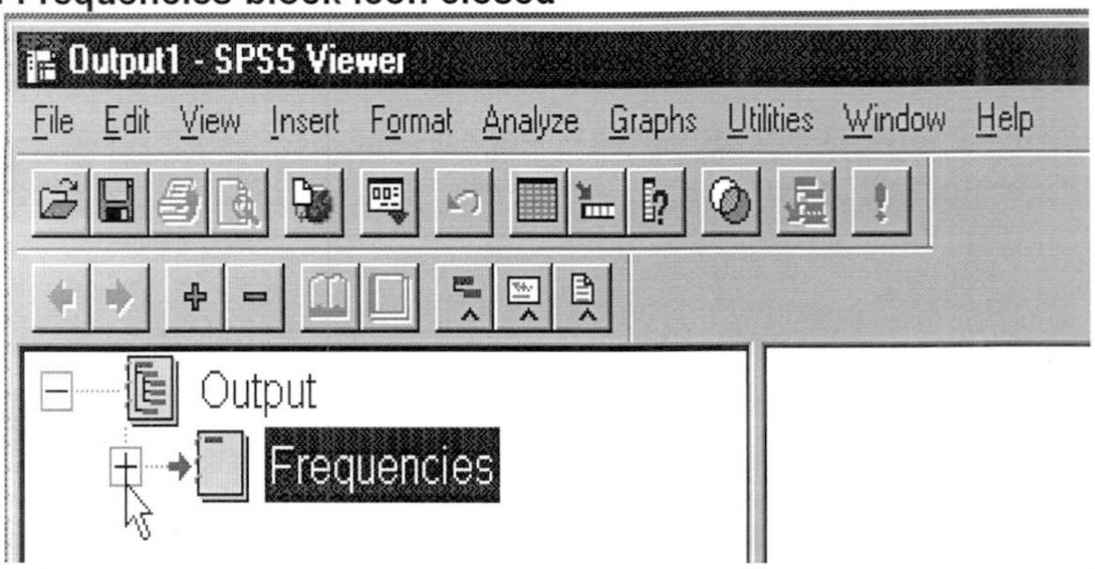


Select an item from the outline and choose Hide from the View menu.

Pic 31: The result of hiding the Statistics output



Pic 32: Frequencies block icon closed



You can see in Pic 32 what the output icon representing an output block looks like when it's open, labelled **Output**, and when it is closed, the one labelled **Frequencies**. Also the small minus, -, and plus, +, buttons can be pushed to close or open an output block respectively.

About the Frequencies output

There are three different parts in Frequencies output, the summary statistics table, the frequency table itself and a chart if one has been specified. The frequency table list the number and percentage of cases with each variable value, e.g. in Pic 33, the number in the sample who consider life dull was 41 and they were 2.7% of the whole sample and 4.2% of the people asked.

Pic 33: Frequency table for the life variable

Is Life Exciting or Dull

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Exciting	434	28.6	44.3	44.3
	Routine	505	33.3	51.5	95.8
	Dull	41	2.7	4.2	100.0
	Total	980	64.6	100.0	
Missing	NAP	524	34.5		
	DK	8	.5		
	NA	5	.3		
	Total	537	35.4		
Total		1517	100.0		

Look at the frequency tables for other variables in the Frequencies output and write down the number and percentage in the sample who:

1. lived in the North East region

2. were very happy at the time of the survey.

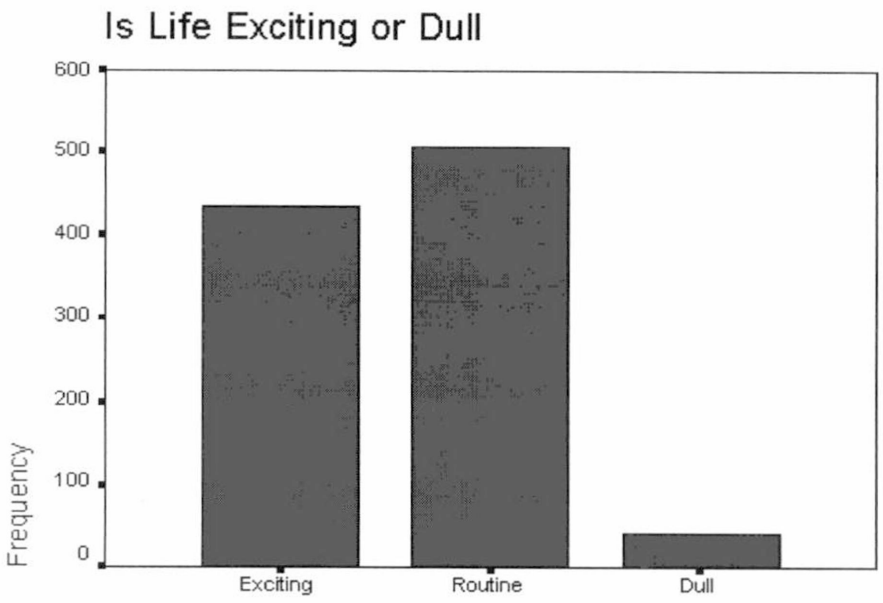
The charts are visual illustrations of the frequency table. In Pic 34 you can see the proportion of people who find life dull from the bar chart in relation to those who found life exciting and routine.

Pic 34: Barchart for the life variable

Notice that there are no bars representing missing values, the charts created with the Frequencies procedure only has the valid values in it.

ⓘ Choosing **Graphs/Bar...** from the menus will permit you to create a chart showing the

relative number of missing values. Look out for the chart tour coming soon which will have more on how to do this. Notice that there are no bars representing missing values, the charts created with the Frequencies procedure only has the valid values in it.



Is Life Exciting or Dull

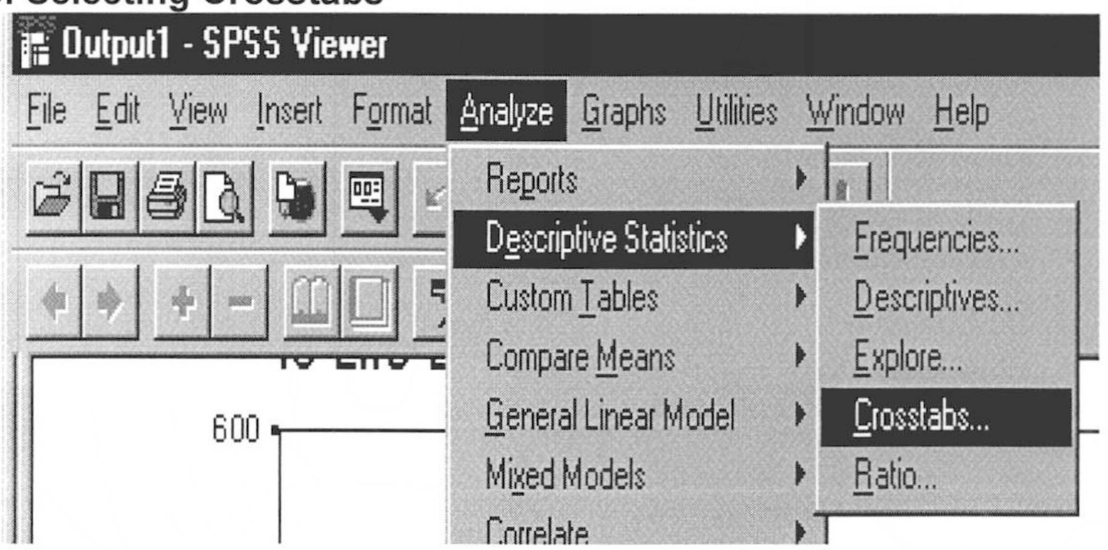
Notice that there are no bars representing missing values, the charts created with the Frequencies procedure only has the valid values in it.

Choosing **Graphs/Bar...** from the menus will permit you to create a chart showing the relative number of missing values. Look out for the chart tour coming soon which will have more on how to do this. **Crosstabs**

Frequencies deals with variables separately but **Crosstabs** displays two or more categorical variables in association with each other in a table. In fact it's called a crosstabulation and it's in the same menu as Frequencies, so choose **Analyze/ Descriptive Statistics/ Crosstabs....** will open the Crosstabs dialog box.

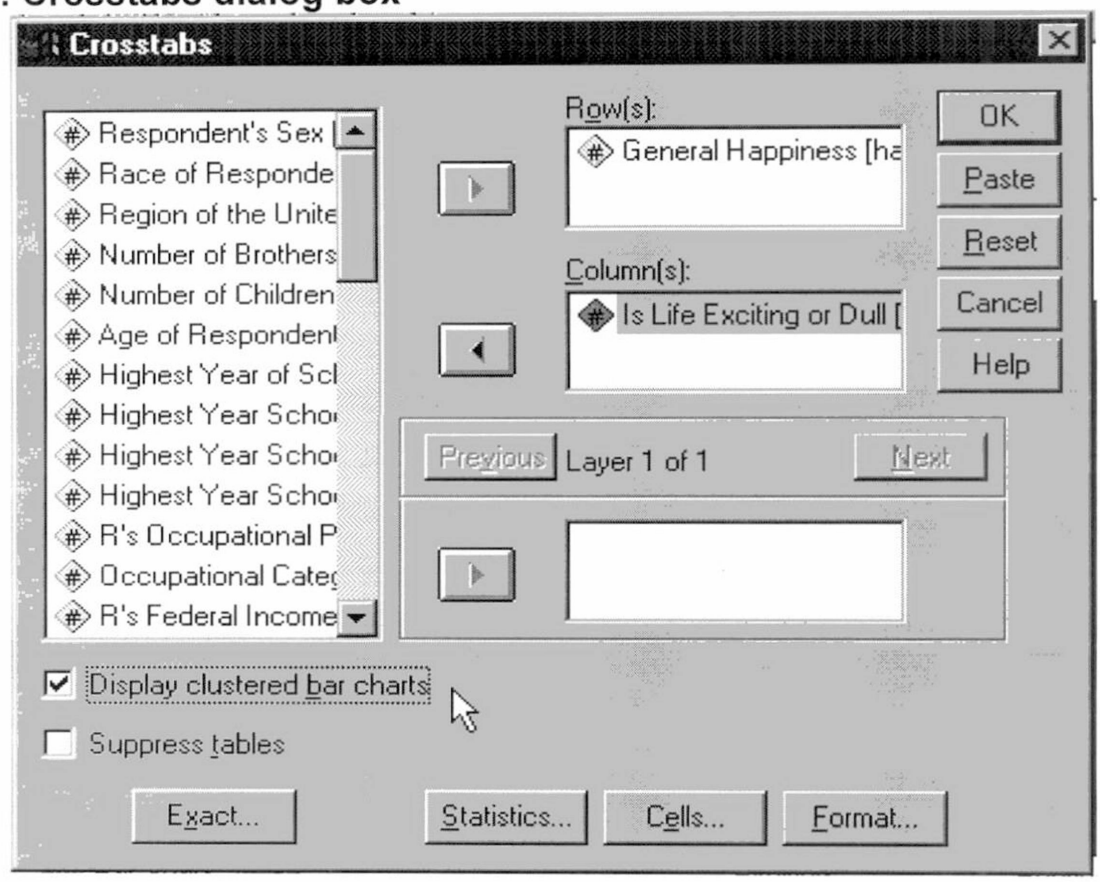
Choose Analyze/ Descriptive Statistics/ Crosstabs....

Pic 35: Selecting Crosstabs



Select the General Happiness (happy) variable and put it in Row(s): Select the Is Life Exciting or Dull (life) variable put it in Column(s): Click the Display clustered bar charts option, as shown above. Click OK to run the Crosstabs procedure.

Pic 36: Crosstabs dialog box



When it is finished you should be looking at the Output Viewer again. You can see in the outline that there are several blocks to the Crosstabs output including the table below in Pic 37 and the bar chart shown in Pic 38.

Pic 37: Crosstabs table

Each valid value of the row variable will define a row and each valid value of the column variable will define a column of the table.

Pic 38: Crosstabs Clustered bar chart

Tables and charts can be customised to suit your own taste. Look out for the table tour and the chart tour coming soon!

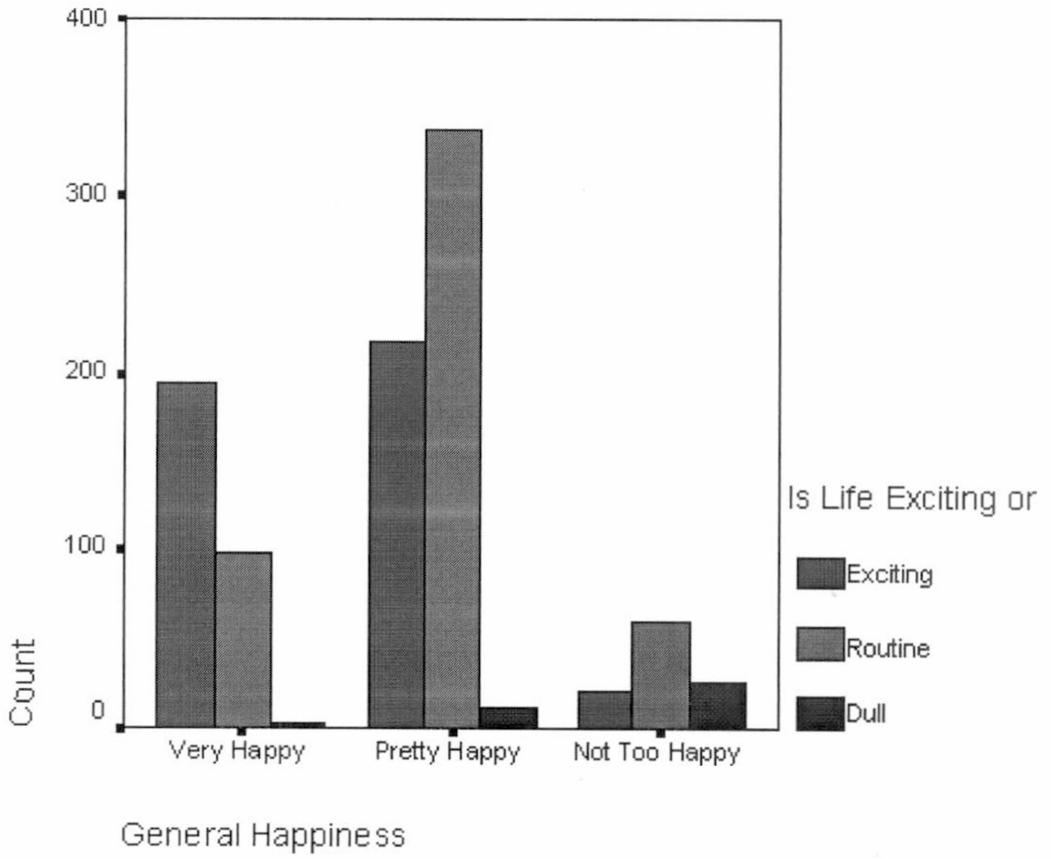
General Happiness * Is Life Exciting or Dull Crosstabulation

Count

		Is Life Exciting or Dull			
		Exciting	Routine	Dull	Total
General Happiness	Very Happy	195	98	2	295
	Pretty Happy	218	338	12	568
	Not Too Happy	21	61	26	108
Total		434	497	40	971

Each valid value of the row variable will define a row and each valid value of the column variable will define a column of the table.

Pic 38: Crosstabs Clustered bar chart

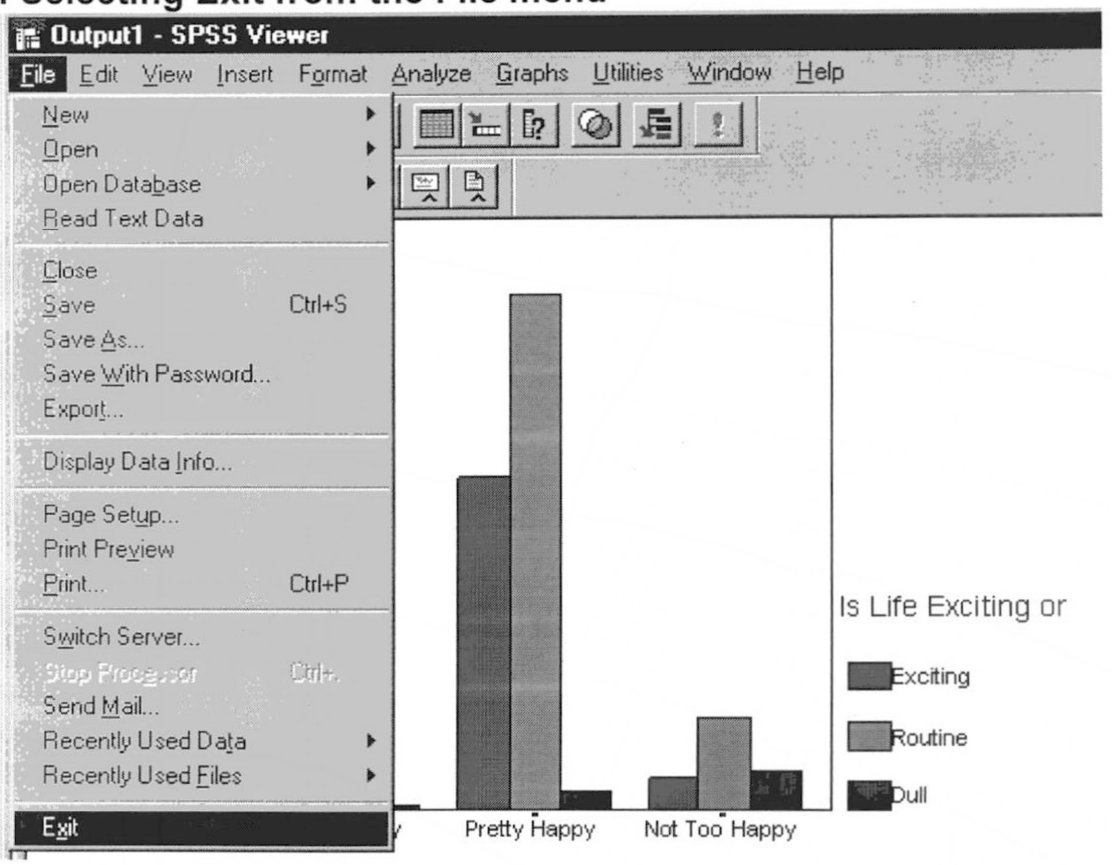


Tables and charts can be customised to suit your own taste. Look out for the table tour and the chart tour coming soon!

Leaving SPSS

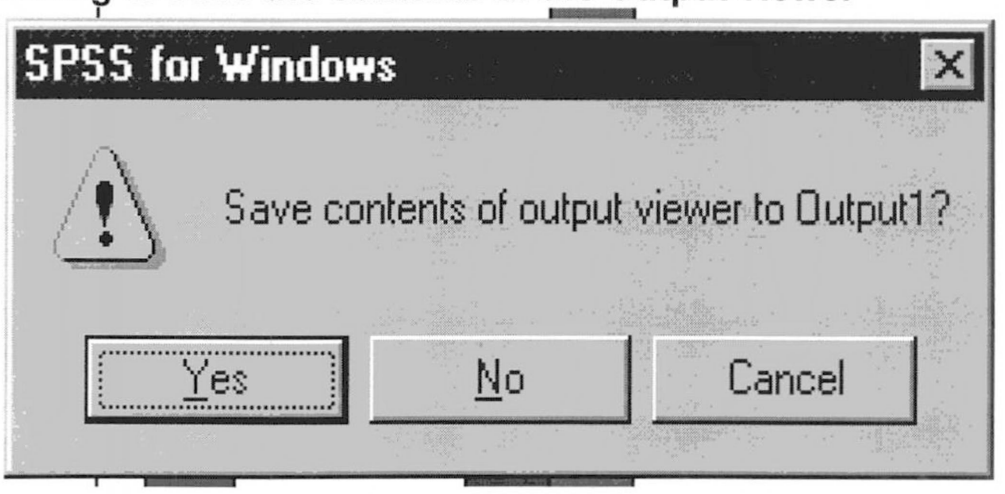
To finish the SPSS session, select File /Exi t from the menus.

Pic 39: Selecting Exit from the File menu



Before SPSS shuts down you will be asked about saving the contents of each window in turn, in this case the output viewer window and the data window.

Pic 40: Warning to save the contents of the Output viewer

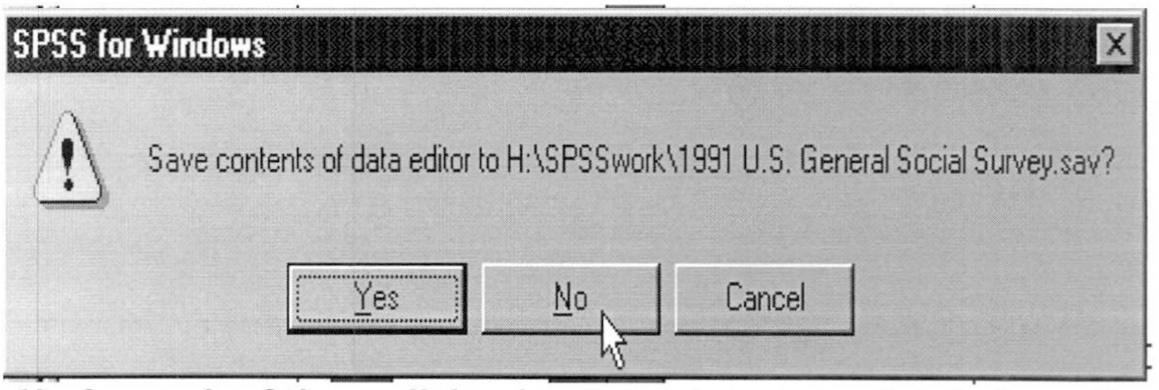


For example in Pic 40, SPSS is asking whether to save the contents of the output viewer. The three buttons have similar functions whatever the type of window that is about to be closed. **Yes** saves the contents to a file, either to a

named file or a dialog box is opened to name a file where that window's contents are to be stored. **No** closes the window and none of the changes you have made during the session will be saved (unless you've used the save option in the file menu).

Cancel stops SPSS closing and leave the current window open.

Pic 41: Warning to save the contents of the data file



Click No for each of these dialog boxes, unless you wish to keep the data file or the output you made.

It is not essential to keep the contents of any of the windows created during this session – unless you want to keep a souvenir of the tour.

Each saved window can be opened up in SPSS, using the **File /Open** from the menus. Shortcuts for the most recently used files are available in the File menu under **File/Recently Used Data** and **File/Recently Used Files**. Choosing the shortcut from the menu will open the file in SPSS as long as it hasn't been deleted in the mean time. SPSS will associate a default file extension with different types of file. You can see these extensions in the drop down menu, **Files of type:** in **Open File** dialog box. The two most common extensions are .SAV and .SPO. **SPSS data windows** are saved to files with .**SAV** extensions and **Output windows** are given .**SPO** extensions.

References:

1. http://www.ucs.ed.ac.uk/fmd/central_labs.html
2. http://www.ucs.ed.ac.uk/ucsinfo/deptalabs.html
3. http://www.ucs.ed.ac.uk/
4. http://www.ucs.ed.ac.uk/usd

QUALITATIVE VERSUS QUANTITATIVE RESEARCH METHODS IN EDUCATION

I. Introduction

There are almost endless number of meanings attached to word 'education' and all of them can be considered accurate within the context intended by the user. In the present context, it refers to the aggregate of all deliberate attempts to bring about development or positive change in human behaviour. Operationally speaking, education is considered as a process, a science and an interaction of teaching and learning. Likewise, the word 'research' has varying meaning from person to person. Essentially, if in systematic investigation of a problem or series of problems or addition to knowledge. The technique of investigation varies from problem to problem and it depends on the purpose of research study. These two words Education and Research have great relevance in the context of programmes of NCERT.

II. Types of Research

Educational research can be broadly classified into number of categories.

1. Basic and Applied Research
2. Basic and Action Research and
3. Analytical and Descriptive Research;
4. Experimental and Non-Experimental Research, etc.

The distinction between these categories as well known needs no elaboration. For any educational practitioner to become educational researcher, three qualifications are essential.

i) Knowledge of various methods of research,

ii) Communicative fluency in the language of research,

iii) An understanding of theory behind application of processes. For our present discussion first one is most significant. Research methods can be classified into two categories :
   - Qualitative Research methods
   - Quantitative Research methods

III. Qualitative versus Quantitative Research

Qualitative research is a growing enterprise worldwide. Qualitative research methods are well developed in sociology and anthropology of education rather than in current issues in education. These are considered as non-experimental studies, hence beginners in this are uncertain of scientific value of such studies. Qualitative research can be both understandable and rigorous. It examines people's words, actions in narrative or descriptive ways and analyses patterns emerging out of data. Data Analysis is the heart of qualitative research. The outcome of such researches are not sweeping generalisations but are contextual findings. The area is developing fast.

On the other hand, quantitative research is based on the observation that are converted into discrete units that can be compared with other units using statistical techniques. Statistical analysis is an important part of quantitative research. Quantitative approach looks past words, actions, patterns but records their mathematical significance. It is based on some valid assumption. Outcome of quantitative research can be generalised. A quantitative researcher makes a guess or forms a hypothesis which is used to test the data.

It is relevant to quote Lord Kelvin – I often say that when you measure what you are speaking about and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of a science, whatever the matter may be.

Above two types of approaches can be further distinguished by their philosophic underpinning. Qualitative research is based on phenomenological position while quantitative research is based on positivist position. These will be further discussed during deliberation.

IV. Characteristics of Qualitative Research

Eight important characters of qualitative research summarised in the figure can be further dealt with to refine research skills.

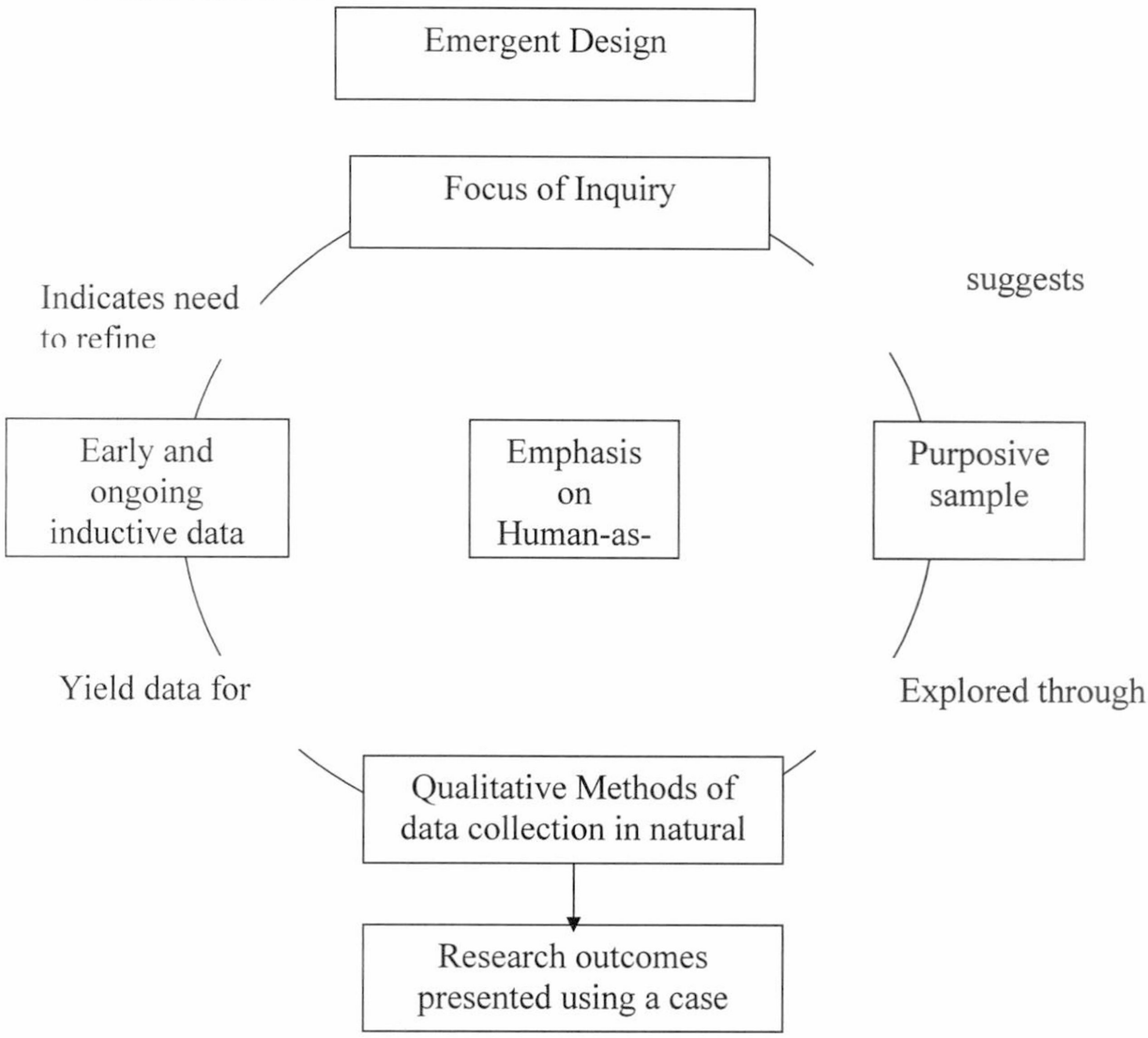


Figure – Characteristics of qualitative research (*Adapted from Lincoln and Guha (1985)).

V. Further Reading

1. Lincoln Y.S. and Guha E.G. (1985) Naturalistic Inquiry Beverly Hills C.A. Sage.
2. Berg B (1989), Qualitative Research Methods for Social Sciences, Boston M A Allyn and Bacon.
3. Pamela M and Richard M (1994) Beginning Qualitative Research – A Philosophical and Practical Guide, The Falmer Press, Washington DC.
4. Bogdem R and Taylor S. (1975). Introduction is qualitative research methods – A Phenomological Approach to Social Science, New York Wileg.

ONE-WAY ANALYSIS OF VARIANCE: PROCEDURE ONEWAY

Examination of Data

Before any statistical analysis, one should examine the distribution of data values to make sure that there is nothing unusual i.e. no outlying or extreme values exist in the data. Use EXAMINE command to make boxplot for each group.

Examination of Sample Means and Confidence Intervals

The sample means for a group provides the single best guess for the unknown population value (parameter), the mean. However, the value of sample mean may not be exactly equal to population value, but it is not too different.

Based on sample mean, one can calculate range of values, called a confidence interval, which may include the population value. The 95% confidence interval means, that if a experiment is repeated 100 times under the same condition with the same sample sizes in each group, 95% of these intervals would contain the unknown population parameter value.

Testing the Null Hypothesis

The statistical technique, which test the null hypothesis, that several population means are equal, is called analysis of variance (ANOVA). This examines the variability of the observations within each group as well as between groups.

SPSS/PC+ contains two different analysis of variance procedures, namely, ONEWAY and ANOVA. ONEWAY is used when one variable is classified into independent different groups. ANOVA is used when two or more variables are classified into groups.

Assumptions

Each of the group is an independent random sample from the normal populations.

--- In the populations, the variances of the groups are equal.

One can use procedure EXAMINE to make stem and leaf plots or histograms for

each group and calculate variances. Statistical tests can be used to check the assumptions of normality and equal variances.

Levene Test

ONEWAY procedure provides Levene Test to test null hypothesis that the groups come from population with the same variance.

The command to use Levene test is:

```
ONEWAY
/ VARIABLES = SCORE BY SEX ( 1,2 )
/ STATISTICS  = 3.
```

(Details of the procedure will be explained in subsequent sections).

Procedure ONEWAY:

ONEWAY procedure produces a one-way analysis of variance. Although ANOVA (to be discussed in subsequent lectures) can also produce one-way analysis of variance. ONEWAY allows to test for nonlinear trends, specify contrasts and use multiple comparison tests.

The basic specification is a dependent (criterion) variable list and an independent (classificatory) variable with its range of values.

How to Specify the Variables: Subcommand VARIABLES

The VARIABLES subcommand is the only required specification for ONEWAY and consists of the name of at least one dependent variable, the keyword BY, and the name of the independent variable followed by its minimum (the lowest) and maximum (the highest) values enclosed in parenthesis and separated by a Comma (,). Cases with values outside this range are excluded from the analysis.

```
ONEWAY
/ VARIABLES = SCORE BY SEX ( 1, 2 ) .
```

Important Note: The above command can only be issued if the System Active File U *.SYS" exists during the session, when you are working. It can be made active only after issuing the either DATA LIST or GET command.

The keyword (subcommand) VARIABLES can be omitted by issuing the following command.

ONEWAY SCORE BY SEX (1,2).

----- Only one independent variable, consisting of two or more groups, can be used in an analysis list.

More than one dependent variable can be specified.

For more than one dependent variables, a separate one-way analysis of variance is produced for each dependent variable.

How to specify Multiple Comparison Tests: Subcommand RANGES

The subcommand RANGES produces any of seven multiple comparison tests available in ONEWAY. One can specify multiple RANGES subcommands. In some tests, a number in parenthesis (p) indicates the significance level. If a significance level is not specified, ONEWAY uses 0.05 significance level.

DUNCAN (p) - Duncan's multiple-range test at p ($0 < p < 1$)

SNK - Student-Newman-Keul's test at p = 0.05

STUKEY - Tukey's alternate test at p = 0.05

TUKEY - Honestly significant different at p = 0.05

SCHEFFE(P) - Scheffe test at p ($0 < p < 1$)

ONEWAY SCORE BY SCHOOL (1,4)

/ RANGES = SCHEFFE (0.05).

Multiple comparisons for all groups (in this case four groups) are produced whenever RANGES is used. (Please refer manual for details).

How to partition Sum of squares for Linear and Non-Linear Trends: Subcommand POLYNOMIAL

The POLYNOMIAL subcommand partitions the between-groups sums of squares into linear, quadratic, cubic and higher order trend components. Its specification consists of a single number which indicates the degree of the highest-order polynomial to be used.

ONEWAY SCORE BY SCHOOL (1,4)

/ POLYNOMIAL = 2.

The above command specifies order 2 (quadratic) as the highest order polynomial.

Some of the restrictions are:

- The number specified must be positive integer and should be less than or equal to 5.
- The number giving order of polynomial, should be less than number of groups.
- The subcommand POLYNOMIAL follows the VARIABLES subcommand and can be used only once.

How to specify the contrasts: Subcommand CONTRASTS

The subcommand CONTRASTS specifies a prior contrasts to be tested by the t statistics. The sequential order of the coefficients is important since it corresponds to ascending order of the category values of the independent variable.

The following command tests difference between means of two groups

```
ONEWAY SCORE BY SCHOOL (1,4)
/ CONTRAST = 1  -1  0   0
/ CONTRAST = 1   0 -1   0
/ CONTRAST =-1   0  1   0
/ CONTRAST = 0   0  1  -1
```

Subcommand STATISTICS

Three optional sets of statistics or keyword ALL on the subcommand STATISTICS can be specified.

Statistic 1 - Group descriptive statistics - Count, mean, standard deviation, standard error, minimum, maximum and 95% confidence interval

Statistic 2 - Fixed and random effects statistics

Statistic 3 - Homogeneity of variance tests- the Levene test

ALL - All statistics

The following command

```
ONEWAY
/ VARIABLES - SCORE BY SEX (1,2)
/ STATISTICS = 1 3.
```

provides descriptive statistics and Levene test for homegeneity of variances.

ANALYSIS OF VARIANCE: PROCEDURE ANOVA

Achievement scores can be suhdivided into two categories of Gender (Male, Female) and four categories of Management (Govt, LB, PA, PUA). Here Achievement score is the criterion (dependent) variable, whereas the Gender and Management are the two independent variables.

Here three hypothesis are of interest

i) Does Gender affect the achievement scores

ii) Does Management affect the achievement scores

iii) Is there any interaction between the effects of Gender and Management

The statistical technique used to evaluate these hypothesis is an extension of the oneway analysis of Variance.

The same assumptions are needed i.e. the observations should be independently selected from normal populations with equal variances.

Total Sum of Squares = Gender Sum of Squares
+ Management Sum of Squares
+ (Gender X Management) Sum of Squares
+ Residual Sum of Squares

The sum of squares for each independent variable are termed as Main Effect sum of squares. The sum of squares due to Gender and Management are termed as Interaction sum of squares. The explained sum of squares is the total sum of squares for main effects and interaction.

Procedure ANOVA

ANOVA performs analysis of variance for factorial designs. By default, it provides full factorial model for five or less than five factors. It tests the hypothesis that the group means of criterion (dependent) variable are equal. One or more categorical variables, also known as independent variables, define the grounds and are termed as factors. ANOVA also allows you to include Continuous explanatory variables, termed as Covariates.

How to specify the Variables: Subcommand VARIABLES

The only required specification for ANOVA is the VARIABLES subcommand with

--- A list of one or more dependent variables

--- the keyword BY and

--- One to five factors (independent variables) followed by their minimum and maximum values enclosed in parenthesis and separated by a comma or space.

ANOVA VARIABLES = SCORE BY SCHOOL (1,4) SEX (1,2).

The actual keyword VARIABLES can be omitted by issuing following command.

ANOVA SCORE BY SCHOOL (1, 4) SEX (1 , 2) .

Caution: The values of factors are arbitrary from statistical point of view. However, they are not arbitrary from computational point of view as they define the dimension of the table of means and variances. If the factor values are not consecutive, they should be made consecutive using RECODE or AUTORECODE command before running ANOVA .

Important Note: The above command can only be issued if the System Active File "*.SYS" exists during the session, when you are working. It can be made active only after issuing the either DATA LIST or GET command

If two or more dependent variables are specified, separate analysis of variance are produced for each variable. (They are not analysed jointly).

ANOVA SCORE1 SCORE2 BY SCHOOL(1,4)
SEX (1,2).

If two or more factor variables have the same value range, the range can be listed following the last factor.

ANOVA SCORE BY PRIMARY SEX (1 , 2) .

Here factor PRIMARY has two Values i.e. Schools with and without Primary sections.

How to specify more than one design

More than one design can be specified on an ANOVA hy using multiple analysis lists separated by slashes.

```
ANOVA SCORE1 BY PRIMARY SEX(I,2)
        / SCORE2 BY SCHOOL (1,4) SEX ( 1 , 2 ) .
```

How to specify the Covariates

Covariates can be specified after the factor list (independent variables) following keyword WITH

```
ANOVA SCORE BY SCHOOL (I, 4) SEX (1,2)
        WITH AGE.
```

MEASURING LINEAR ASSOCIATION

1. Introduction

Quite often, researcher is interested in cause and effect relationships between two or more continuous variables. The linear relationship is mostly preferred because it is theoretically very well defined. In this endeavour, values of two variables are plotted together to have an idea of relationship between them. Plotting the values of two variables .is known as bivariate scatterplot. A bivariate scatterplot using a third variable is also required to be displayed in one plot frame. This scatterplot is known as controlled scatterplot. Besides visual display of relationship, the linear relationship is expressed in terms of coefficient for continuous variables is the Pearson product moment correlation coefficient. This unit, therefore, aims at developing sufficient skills for drawing the above mentioned coefficient.

It is presumed that all the execution options preferred through 'Set' command are known to the trainees. These are neither deliberated in this unit nor these are part of the illustrations for .the sake of brevity.

2. Training Outcome

After completion of the unit, it would be possible to execute the following activities at the SPSS/PC+ platform.

i. Draw bivariate scatterplot for single and multiple cases when all the display parameters have default values.
ii. Draw several bivariate scatterplots by using single subcommand.
iii. Draw overlay scatterplot by using default parameters.
iv. Draw bivariate and controlled scatterplot by selecting values of display parameters such as plot titles, scaling and labeling of plot axes.
v. Basic specification of the subcommand for computation of correlation, coefficient.
vi. Computation of correlation coefficient using different options, i.e., treatment of missing values, display of results in square matrix, counts and probabilities, univariate statistics and cross product deviations and covariances.

3. Data Set Description

Illustrations in this unit use a data set stored in he form of SPSS/PC+ system file named as 'pupil. sys' . This file contains students achievement scores and data on their socio-economic background. The variables used in the illustrations are defined as follows:

	Variable description	Variable label	Variable value code
(i)	Mathematics Score	Math	0-40
(ii)	Language Score	Lang	0-80
(iii)	Student gender	Sex	Girls-1, Boys-0
(iv)	Student's social class	SC/ST	SC or ST=1 Non-SC/ST=0
(v)	Father's Education	daded	Transformed to ratio data type and centered
(vi)	Mother's Education	mumed	Transformed to ratio data type and centered
(vii)	Father's Occupation	dadocc	Transformed to ratio data type and centered
(viii)	Ever repeated a class	repeat	Repeated in any class=1 else=0
(ix)	Socio-economic status	ses	Average of daded, mumed and dadocc variable

For executing the illustrations please ensure that this data set is copied into the active file.

4. Scatterplot Under Default Options

In order to draw the bivariate scatterplot for math and daded variable in default mode, execute the following commands.

```
GET             FILE='pupil.sys'
PLOT/PLOT    math    WITH
daded.
```

The first 'PLOT' subcommand is generally used to select various parameters of the plot whereas the second 'PLOT' subcommand is used to specify the variables. The scatterplot, as illustrated above, can also be obtained under default conditions by introducing a blank space between the two 'PLOT' subcommands.

You must have noticed in the above illustration that different symbols have been used to display two or more cases with similar values on the same point in the scatterplot. The following symbols are used to represent multiple cases.

Frequency		Symbol	
From	To	From	To
1	9	1	9
10	35	A	Z

The above scatterplot symbols 'D' and '5' indicate 13 and 5 cases respectively.

The following command provides four bivariate scatterplots, i.e. math with dadocc, math with daded, lang with dadocc and lang with daded.

```
GET FILE='pupil.sys'.
PLOT/PLOT math lang WITH dadocc daded.
```

Variables can be taken into pairs. For example, to draw two scatterplots, math with dadocc and math with daded, we use the following command:

```
GET FILE='pupil.sys'
PLOT/PLOT math WITH dadocc daded (Pair).
```

Controlled scatterplot is obtained by including the third variable 'girl' as it is shown below:

```
GET FILE='pupil.sys'.
PLOT/PLOT math WITH daded by girl.
```

Multiple plots can be obtained by using a semicolon (;), a separator, between two plot specification.

GET FILE='pupil.sys'.

PLOT/PLOT math WITH dadocc; math WITH mumed.

5. Scatterplot Under Different Options

(a) The type of plot is specified by the subcommand 'format' having the following values:

(i)	Overlay plot	:	Several variables use the same type of measurements or when the same variable is measured at different times.
(ii)	Regression	:	Regression of y-axis variable on x-axis variable. The intercept is marked by R.
(iii)	Contour (n)	:	Controlled plot having n layers (min. 10 and max. 36) is obtained for a continuous controlled variable.

Example

```
GET FILE='pupil.sys'.
PLOT/FORMAT=OVERLAY/PLOT daded mumed WITH lang.
PLOT/FORMAT = REGRESSION/PLOT math WITH lang.
```

(b) 'SYMBOLS' and 'CUTPOINT' subcommands are used for selection of symbols and fixing the interval of the frequency at a point in the scatterplot. The default values for symbols is as follows

(i)	Alphanumeric	:	the 36 symbols listed in the earlier symbol list given in section 4.
(ii)	Numeric	:	Symbols 1 to 9 are used to indicate 1 to 9 frequencies and '*, for more -than 9.

Besides, the above options you can choose specific hexadecimal symbols by enclosing them within two apostrophe (") without a blank or comma(,) between two symbols.

Examples:

GETFILE='pupil.sys'.

(i) PLOT/FORMAT=OVERLAY/SYMBOLS=ALPHANUMERIC.
/PLOT= daded mumed WITH lang.

GET FILE='pupil.sys'.

(ii) PLOT/FORMAT=CONTOUR/SYMBOLS= '.-=.+%#o&x'
/PLOT= math WITH lang by mumed.

GETFILE='pupil.sys'.

(iii) PLOT/FORMAT=OVERLAY/SYMBOLS=NUMERIC.
/PLOT= daded mumed WITH lang.

(c) 'CUTPOINT' subcommand is used to alter the width of the frequency or category represented by plot symbols. The following keywords are specified with the CUTPOINT subcommand:

(i) EVERY (n) : Frequency interval of width n

(ii) [value list] : Each value defines a cutpoint. Not more than thirty five points can be defined. The number is one less than the number of cut point.

Examples:

(i)
GET FILE='pupil.sys'.
PLOT/CUTPOINT=EVERY(5)/PLOT math WITH lang.

(ii)
GET FILE='pupil.sys'.
PLOT/CUTPOINT=(5, 10, 15)/PLOT math WITH lang.

(d) Scaling and Labeling of Axes

The 'VERTICAL' and 'HORIZONTAL' subcommands are used to control the scaling and labeling of the vertical and horizontal axes respectively. These subcommands can be used once before each subcommands and apply to all plots specified in the next PLOT subcommand. The following specifications are available.

(i)	'label'	:	Label for the axis can have not more than 40 characters and truncated if the axis size is shorter than label.
(ii)	MIN (n)	:	Minimum axis value in integer.
(iii)	MAX (n)		Maximum axis value in integer.
(iv)	UNIFORM	:	The Uniform axis scale based on minimum and maximum observed values.
(v)	REFERENCE (values)	:	Draws reference lines at the values specified.
(vi)	STANDARDIZE	:	Plot standardized variables when they are on different scales.

Examples:

GET FILE= 'pupil.sys'.

(i) PLOT/VERTICAL='ACHIEVEMENT SCORES IN MATHS' MIN(O) MAX(40)/ HORIZONTAL='ACHIEVEMENT SCORES IN LANGUAGE' MIN(O) MAX(80) REFERENCE(40)/PLOT maths WITH lang.

GET FILE='pupil.sys'.

(ii) PLOT/VERTICAL=UNIFORM STANDARDISE /HORIZONTAL=UNIFORM/PLOT=math WITH lang.

(e) Setting the Plot Size

The subcommand 'VSIZE' and 'HSIZE' control height and width of the plot respectively. To control the axis values displayed and the interval between them, VSIZE and HSIZE are used in conjunction with 'MIN' and 'MAX' subcommands of 'HORIZONTAL' and 'VERTICAL' with the help of (MAXMIN)*S/(INTERVAL)= VSIZE OR HSIZE, if (MAX-MIN)/(Interval) is non-integer, adjust value of MAX OR MIN accordingly.

Page length and width defined through 'VSIZE' AND 'HSIZE' are to be less than the respective default values 16 and 38 if these are controlled through the SET COMMAND at LENGTH=79 and WIDTH=24.

Example:

```
GET FILE= 'pupil.sys'.
PLOT/VSIZE=10/HSIZE=20/PLOT math WITH lang.
```

(f) Specifying the Plot Titles

The 'TITLE' subcommand is used to specify the title of the plot. It can contain not more than 60 characters.

Example:

```
GET FILE='pupil.sys'
PLOT/SYMBOLS=
'ML'/FORMAT=OVERLAY/
TITLE='ACHIEVEMENT SCORES IN MATHEMATICS AND
LANGUAGE' /PLOT Math WITH lang WITH mumed
```

(g) Treatment to Missing Values

The 'MISSING' subcommand, used for the treatment of the missing values, It has the following options.

PLOTWISE Excludes cases having a missing values for any variable within a single plot.

LISTWISE Cases with missing values for any variable named on the 'PLOT' Subcommand are excluded from all plots.

INCLUDE Only cases with system missing values are excluded. It can used with either PLOTWISE or LISTWISE.

6. Measure of strength of Linear Association Between Two Variables

Scatterplot visually reveals varieties of associations between two variables. Most commonly such relationships encountered are (i) no discernible relationship between two variables; (ii) values of variable y increases rapidly for higher values of Xi (iii) small and large value of X are associated with large values of Yi and (iv) values of y and x variable more or less cluster around a straight line. The first step in studying association between two variables is to have a scatterplot. The next step often useful is to quantity the strength of the association by computing the commonly used Pearson correlation coefficient. It is defined as cov(x,y) / {std.dev. (x) *std.dev. (y)}. Its largest possible absolute value 1 indicates exact linear relationship between the two variables. The lowest absolute value 0 indicates no linear relationship. We use 'CORRELATIONS' command to compute Pearson correlation coefficient between two non-string variables. In the case long or short 'string' variables 'CORRELATIONS' command is not executed.

Example 1:

Compute correlation coefficient among math, lang, daded and mumed and save the results in square matrix in the file 'corr.ex l' under default conditions.

```
SET  RESULTS='corr.ex1'.  GET
FILE='pupil.sys'.
CORRELATIONS VARIABLES=math lang daded mumed.
* FINISH
```

Please observe:

i. The correlation of a variable with itself is displayed as 1.

ii. As asterisk (*) indicates that a coefficient has a two tailed probability of less than 0.01 and two asterisks (**) indicates a probability of less than 0.001.

iii. Cases that have missing values for any variable in the matrix are excluded.

Example 2:

Compute correlation of math with dadocc, daded and mumed, and of lang with dadocc, daded and mumed. Save results in rectangular matrix wherein math and lang are rows and dadocc, daded and mumed are the columns. We use the following commands.

```
SET RESULTS='corr.ex2'.
GET FILE='pupil.sys'.
CORRELATIONS VARIABLES=math lang dadocc daded mumed.
* FINISH
```

Please try the above example by omitting the 'VARIABLES' subcommand.

The command 'CORRELATIONS' has five options under subcommand 'OPTION' for handling of missing values (inclusion of user missing values and exclusion of pairwise missing values) computation of onetailed probability, writing of correlation matrix and count, and display of count and probability. The most commonly used option, i.e. onetailed probability is given in the following example.

Example 3:

Compute correlation of math with mumed and daded by computing onetailed probability.

```
SET
RESULTS='corr.ex3'.
GET FILE='pupil.sys' .
CORRELATIONS VARIABLES=math WITH mumed and daded/OPTION 3.
```

You can have univariate statistics, and cross products deviations and covariances by using the subcommand 'STATISTICS' by assigning values 1 and 2 respectively. Example 4 and 5 illustrates these.

Example 4:

Compute correlation for math, lang and mumed. Also compute univariate statistics (cases, mean and std.dev.) for all variables.

```
SET RESULTS='corr.ex4'.
GET FILE='pupil.sys'.
CORRELATIONS VARIABLES=math lang mumed/STATISTICS=l.
```

Example 5:

Compute correlation for math, lang and mumed and also provide cross-product deviations and covariance.

```
SET RESULTS='corr.ex5'. GET
FILE='pupil.sys'.
CORRELATIONS VARIABLES=math lang mumed/STATISTICS=2.
```

5. *LIST OF PARTICIPANTS:*

1. Dr. G. Devaraj
   Lecturer in Education (SG)
   Meston College of Education,
   ROYAPETTAH,
   CHENNAI – 600 014
   CELL: 94441 36571
   jeril_litani@yahoo.co.in

2. Mr.M.Kandasamy,
   Lecturer in History (S.G.),
   Govt. College of Education,
   KOMARAPALAYAM – 638 183
   CELL: 94437 69937
   Res: 04283 254368

3. Dr. K.Mohanasundaram,
   Reader in Physical Science Education,
   Govt. College of Education,
   ORTHANAD – 614 625
   THANJAUVR Dist.
   CELL: 94436 17730
   dr_mohanasundaram@yahoo.com

4. Mr.G.Subramonian
   Lecturer (SS) in Commerce
   Sri RamaKrishna Mission Vidyalaya
   College of Education,
   COIMBATORE – 641 020
   CELL: 98947 43939
   gsubramoni2005@rediffmail.com

5. Dr. S.Thangarajathi,
   Lecturer in Education,
   Annammal College of Education for Women,Thiruchendur Road,
   TUTICORIN – 628 003
   CELL: 98421 77401
   thangam_rajathi@yahoo.co.in

6. Prof.R.Charles Williams
   Senior Lecturer in Education,
   Govt. College of Education,
   PUDUKKOTTAI – 622 001
   CELL: 94433 61201
   vaathyar@gmail.com
7. Dr.S.Malthi
   S.S. Lecturer in Education,
   N.K.T. National College of Education, Triplicane,
   CHENNAI 600 005
   CELL: 98411 72972
   malasaravana@yahoo.com
8. Mrs. K.Mythili,
   S.S. Lecturer in Education,
   Govt. College of Education,
   VELLORE – 632 006
   Phone No. 0416 – 2246262

9. Sr.S.Arockia Vinotha
   St.Justin's College of Education
   161, Kamarajar Salai,
   MADURAI – 9
   Phone: 0452 – 2311012
   sr_vinotha@yahoo.com

10. Dr.Mrs Kiruba Charles,
    Reader and HoD in Education,
    Lady Willingdon Institute of Advanced study in Education,
    CHENNAI – 600 005.
    CELL: 98400 61384
    Phone: 044 – 24641462
    kirubacharles@hotmail.com

11. Dr.P.Jayaraman
    Senior Lecturer and
    HoD E.T.Branch,
    D.I.E.T., PERUNDURAI
    ERODE – 638 052
    CELL: 99420 45798
    jayaram_phy_diet@yahoo.co.in

12.Mr.P.Govindaprakash,
Lecturer in E.T. Branch,
D.I.E.T.,
UTHAMACHOLAPURAM,
SALEM – 636 010.
CELL: 98429 06860
Phone: 0427 – 2273751(O)
seppgpsmt@yahoo.com

13.Mrs.P.Jayanthi,
Lecturer in E.T. Branch,
D.I.E.T.,
RANIPETTAI,
Vellore
CELL: 94448 65612
Phone: 95416 – 2243878
jayanthipnathan@yahoo.com

14.Mrs.T.Jeeva,
Lecturer in E.T. Branch,
D.I.E.T., ODDANCHATRAM
Dindigul – 624 619
Phone: 04553 – 240793
jeenivodc@yahoo.co.in

15.Mrs. H.S.Sugatha Jeraldin,
Lecturer in Physics,
D.I.E.T.,T.KALLUPATTI,
Madurai –
CELL: 98420 73512
Phone: 04549 – 270475
sugatha_sujeer@yahoo.co.in

16.Mr.M.Anandaraj,
Lecturer in E.T. Branch,
D.I.E.T., UTHAMAPALAYAM,
THANI –
Phone: 04554 – 267703

17.Mr.S.Natarajan,
Lecturer in E.T. Branch,
D.I.E.T., KEELAPALUVUR,
Perambalur – 621 707
CELL: 94434 04051
natbuma_nivesineha@yahoo.co.in

18.Mr. S.Muthu Rengan,
Lecturer in E.T. Branch,
D.I.E.T., KILPENNATHUR,
Tiruvannamalai – 604 601
Phone: 04175 - 242308
smuthurengan@gmail.com

19.Mrs.M.Ansa,
Lecturer in Mathematics,
D.I.E.T., Lawspet,
PONDICHERRY – 605 008
CELL: 94430 74111
Phone: 0413 – 2251243(O) & 2290517(R), divi_rose@yahoo.com

20.Mrs.S.Nagalakshmi,
Lecturer in Mathematics, D.I.E.T., ADUTHURAI, Thanjayur
CELL: 93451 84506
Phone: 0435 – 2472476 (O)
04372 – 234284 (R)
nagai_velan@yahoo.co.in

21.Mrs.S.A.Annie Isabella,
Lecturer in Education (S.G.),
Lady Willingdon Institute of Advanced study in Education,
CHENNAI – 600 005.
CELL:98400 77035
Phone: 044 – 22771895
sa_annie2002@yahoo.com

LIST OF RESOURSE PERSONS:

1. Dr. G. Viswanathappa
   Senior Lecturer in Education
   Regional Institute of Education,
   MYSORE – 570 006
   CELL: 0 94482 39716
   Phone: 0821 – 2410995 (R) & - 2512153
   E-Mail Id: gvriem@gmail.com or gvriem@yahoo.co.in

2. Dr. D. Basavayya
   Professor of Mathematics (Retired)
   Regional Institute of Education,
   MYSORE – 570 006
   CELL: 0 9448978110
   Phone: 0821 – 2540410 (R)
   E-Mail Id: basavayya_d@yahoo.com

3. Mr. A. C. Josy
   Selection Grade Lecturer in Education (Retired)
   Regional Institute of Education,
   MYSORE – 570 006

4. Dr. Vasudev
   Director
   Centre for Information Science and Technology (CIST)
   University of Mysore
   MYSORE – 570 006

5. Dr. S. Ravi
   Reader in Statistics
   PG Department of Studies in Statistics
   University of Mysore
   Manasagangothri
   MYSORE – 570 006