
UNIT I INTRODUCTION TO SPSS, DATA MANAGEMENT AND TRANSFORMATION- CREATING VARIABLES-ENTERING DATA

- 1.1 Introduction
- 1.2 Objectives
- 1.3 Overview of Concepts in Quantitative Study
- 1.4 Introduction to SPSS
- 1.5 Creating a Data File
- 1.6 Data Management
- 1.7 Summary
- 1.8 Glossary
- 1.9 Answer to Check Your Progress
- 1.10 Reference/ Bibliography
- 1.11 Suggested Readings
- 1.12 Terminal & Model Questions

1.1 INTRODUCTION

Statistics is that branch of mathematics that converts data into some useful information. This transformation requires complex calculations which can easily be done using computers. Attempting these computations manually is sometimes a herculean task for researchers. Therefore, SPSS was created to help researchers in handling large volume of data that they collect during their research study.

In this unit you would be learning about the role of SPSS Software in transforming statistical data into some meaningful information. SPSS provides a good way to introduce you to basic statistical methods and demonstrate how these methods can help you in decision making for your research activity.

1.2 OBJECTIVES

After reading this unit you will be able to:

- ❖ Learn about the characteristics of SPSS.
- ❖ Conversant with the terminologies used in SPSS.
- ❖ Learn the procedure for entering data into SPSS

- ❖ Understand the important features of the SPSS that will help in data processing.

1.3 OVERVIEW OF CONCEPTS IN QUANTITATIVE STUDY

Research is basically a systematic enquiry or an objective process of gathering data for the purpose of making decisions. Research is a fact finding process which is undertaken through a systematic procedure which includes collection, compilation presentation and interpretation of data. The research which aims to establish some theory is known as basic research, while the research conducted for solving problems with the use of theories is termed as applied research.

There are certain concepts or terminologies associated with quantitative study which will be repetitively used throughout the module.

Population: A population consists of all the items or individuals about which you want to draw a conclusion.

Variable: It refers to properties or characteristics that can be taken in different values, quantitative and qualitative. There are two type of variable, one is independent variable and the other is dependent variable.

Construct: It is a depiction or an abstract idea which is formulated for the given research or a theory. A construct is a redefinition of the concepts as per the research undertaken.

Model: It represents some system that have been established to study some aspect of that system or the system as a whole.

Sample: It is the part of a population that is selected for analysis.

Parameter: It represents characteristics of the population.

Hypothesis: It is a proposition that can be experimentally verifiable and has a definite practical consequence.

Laws: Generalisations that are accepted as true are laws either affirmative or conditional.


Principles: Hypotheses that are acceptable becomes principles.

Research approaches are divided into two categories one is deductive and the inductive research and second one is qualitative and quantitative research.

Inferences that are drawn from a general principle to a particular conclusion constitute deductive research and inferences which are drawn from the evidence in the form of conclusions which explain evidences of facts constitute inductive research.

The quantitative approach includes generalisation based on the results of the study conducted and quantitative research aims as better understanding of the problem.

Measurement Scales: Refer to the file 'Scale' posted in LMS for knowing more about measurement scales. For further exploring about the scales, refer to the following link-

 http://en.wikipedia.org/wiki/Level_of_measurement

Coding of questionnaire items- Refer to the link provided for knowing about coding of questionnaire

http://en.wikipedia.org/wiki/Coding_%28social_sciences%29

1.4 INRODUCTION TO SPSS

SPSS is a software package used for statistical analysis. It is statistical software that accepts raw data and converts into some relevant statistics that can be used for further analysis. It is termed as Statistical Product and Service Solutions widely known as Statistical Package for the Social Sciences. It is a comprehensive tool for analyzing statistical data. SPSS is compatible in accepting data of different file format and use them to generate tabulated reports, charts, and graphs including descriptive and inferential statistics.

You can learn more about History of SPSS from the following url;
<http://en.wikipedia.org/wiki/SPSS>

1.4.1 STARTING SPSS

An SPSS session can be opened like any other window based program. To do this on most computers, you need to click Start Button, move cursor over the menu folder and then over SPSS programme icon. Just click to SPSS button and accordingly after clicking SPSS session will begin. The window desktop will look similar to that shown in screenshot.

Path: CLICK



After doing this, SPSS session would open and the view would be in the form of SPSS Statistics

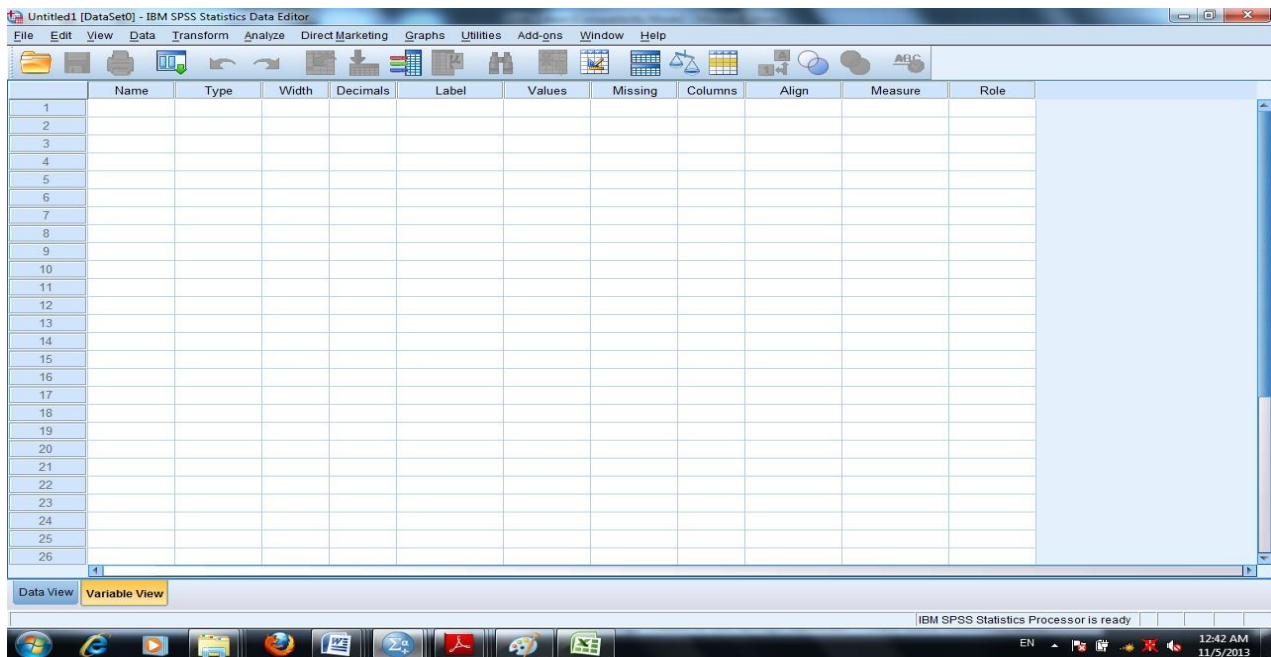
Data Editor as shown in the image below:

Depending on how your software is configured, you may get an option window with OK and Cancel buttons. If so, click the cancel buttons. In either case, an empty Data Editor Windows appears. In option window, you will also find options regarding opening an existing file or importing a data file in other formats like Excel, SQL data and the like.

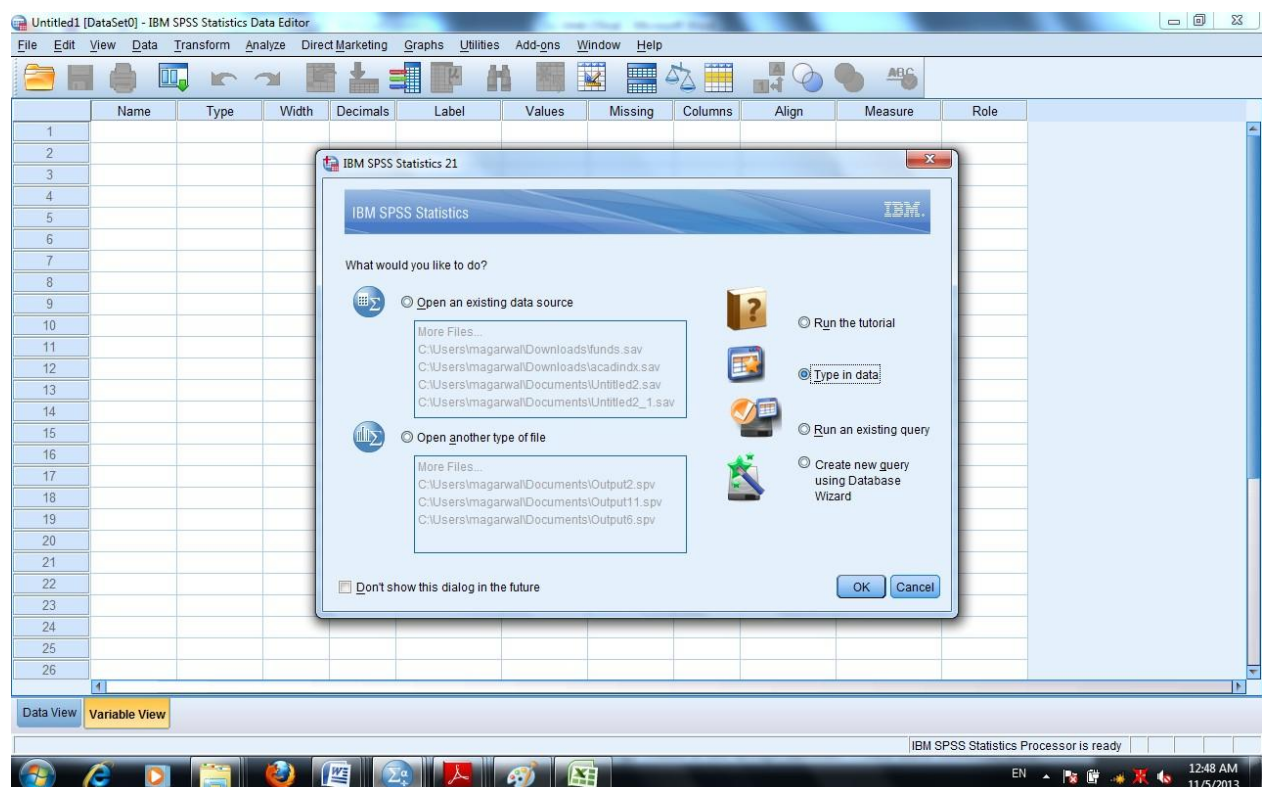
(Note: SPSS can take data from many sources. It can read data from a text file, a database, or files produced by programmes like Access or Excel).

For creating a new file; you have to click Type in data and click OK or you can click to cancel button and you will come to the Data Editor for typing a new set of data. However, for opening existing file: click or select to open an existing data source button, then select the

identified data source and then click to open. The existing file is ready to use. This is shown in the figure below –

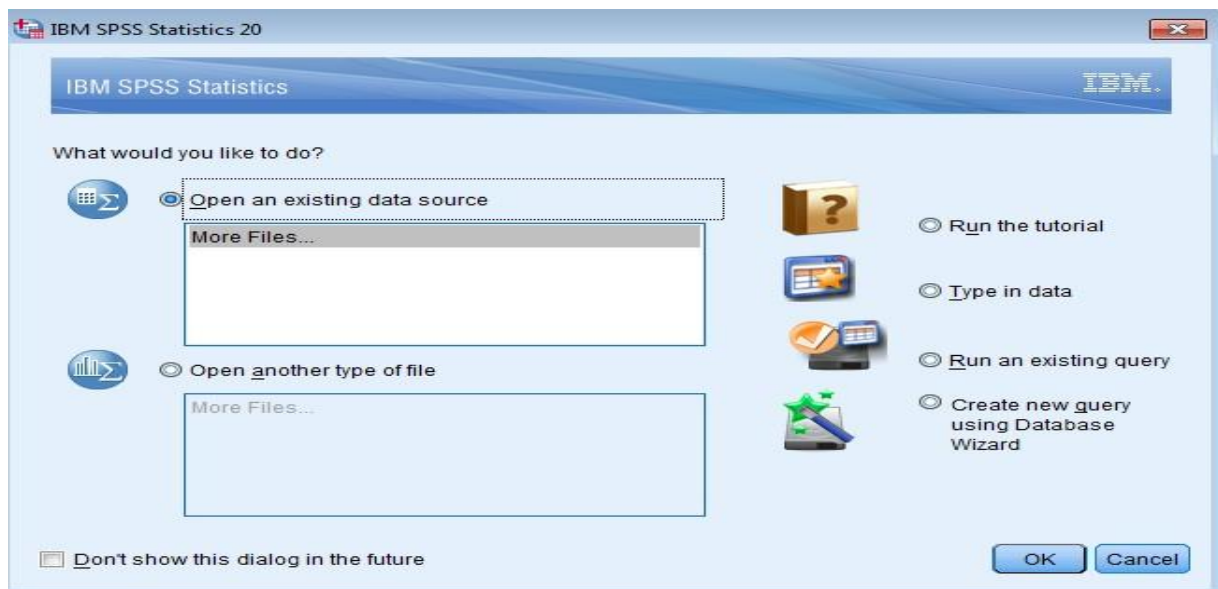


For New File Creation in SPSS follow the depicted path;





For opening existing file in SPSS follow the steps provide as under;



For typing a new set of data or for creating a new file, you have to go the mode i.e. Variable View. These two modes are indicated by the tab at the bottom of the window to the left side of the screen, one is data view and the other is variable view.

1.4.2 PLATFORMS OF SPSS

In order to work with SPSS, you should have a brief idea of various interfaces of SPSS. The main interfaces are;

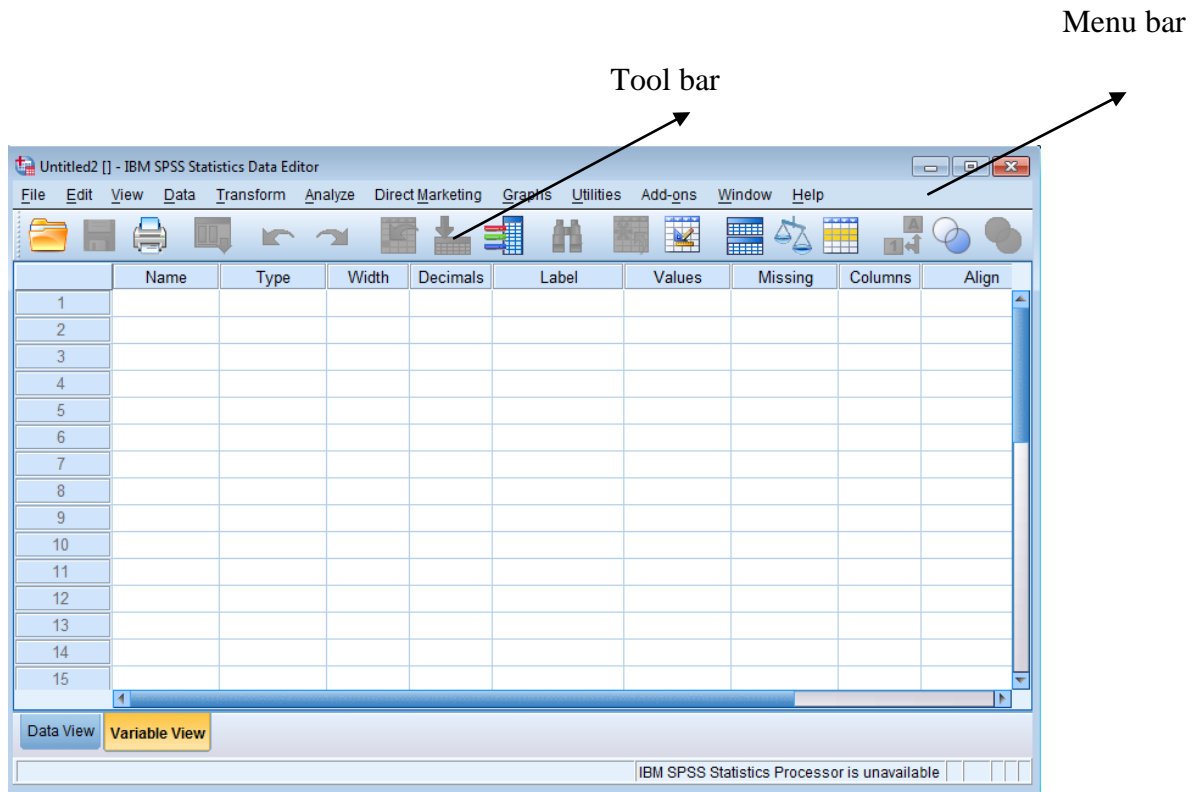
1.4.2.1 Data Editor

1.4.2.2 The Output Window- The output

1.4.2.3 Syntax Editor









1.4.2.4 Script Editor









1.4.2.1 DATA EDITOR: TOOL BAR AND MENU BAR



In data editor the menu bar and the tool bar are located at the top of the screen. Tool bar provides the different aid tools available in SPSS. Some icons are bright and clear and others are greyed. Greyed icons are those that are only available after data is added in the SPSS. It is below the menu bar which provides a quick and easy access to regularly used features. The tool bar shows different tools with different editors. A brief description of tool bar is presented as under-



	Print a particular data file		Give different weights to cases
	Open recently used commands		Case selection on the basis of some criteria
	Redo or undo any action		Depicts value label
	Go to particular case		Only variables in the selected sets will appear in the Data Editor and in the dialogs

	Open a particular data set	 	Insert variable/case
	Save a particular data file		Split file as per groups
	the variable. Missing , label and value label etc		
	Search particular data		

Menu Bar : The menu bar contains the following Menu items:

Items	Description
File	It helps in creating new data file, opening existing one and save and print data file.
Edit	Allow editing functions like copying, cut, paste.
View	Shows menu editor for change in fonts, grid lines , value labels etc.
Data	Details about entering data, define variable properties. Helps in identifying duplicate cases, sorting cases, transposing, restructuring and aggregating data.
Transform	Transforms data compatible for analysis using recode, replace missing value etc.
Analyse	All inferential statistics are available. It contains statistical tools and techniques
Graphs	Builds different types of charts, graphs etc. Using chart builder
Utilities	Command which are used for more complex statistical computations
Add-ons	List extra features available for advanced level
Window	Allows you to arrange, select and control the attributes of windows

Help

Supports in gaining insights about the procedures of SPSS. Contains tutorials, coach etc.

1.4.2.2 The Output Window

The output window assimilates the results of the work done in SPSS. This means that it identifies the results of previously conducted analysis. The picture below summarises the key elements of the output window and presents a detailed tool bar on the output window.

You will find several options in the tool bar options which are similar to the Data Editor window, however, the output window works with the output and the Data Editor works on the data with the options available in the tool bar.

>Warning # 67. Command name: GET FILE
>The document is already in use by another user or process. If you make changes to the document they may overwrite changes made by others or your changes may be overwritten by others.
>File opened C:\Users\magarwal\Documents\Untitled2.sav
DATASET NAME DataSet1 WINDOW=FRONT.
FREQUENCIES VARIABLES=Jobsatisfaction
/ORDER=ANALYSIS.

Frequencies

[DataSet1] C:\Users\magarwal\Documents\Untitled2.sav

Statistics

Jobsatisfaction

N	Valid	14
	Missing	0

Jobsatisfaction

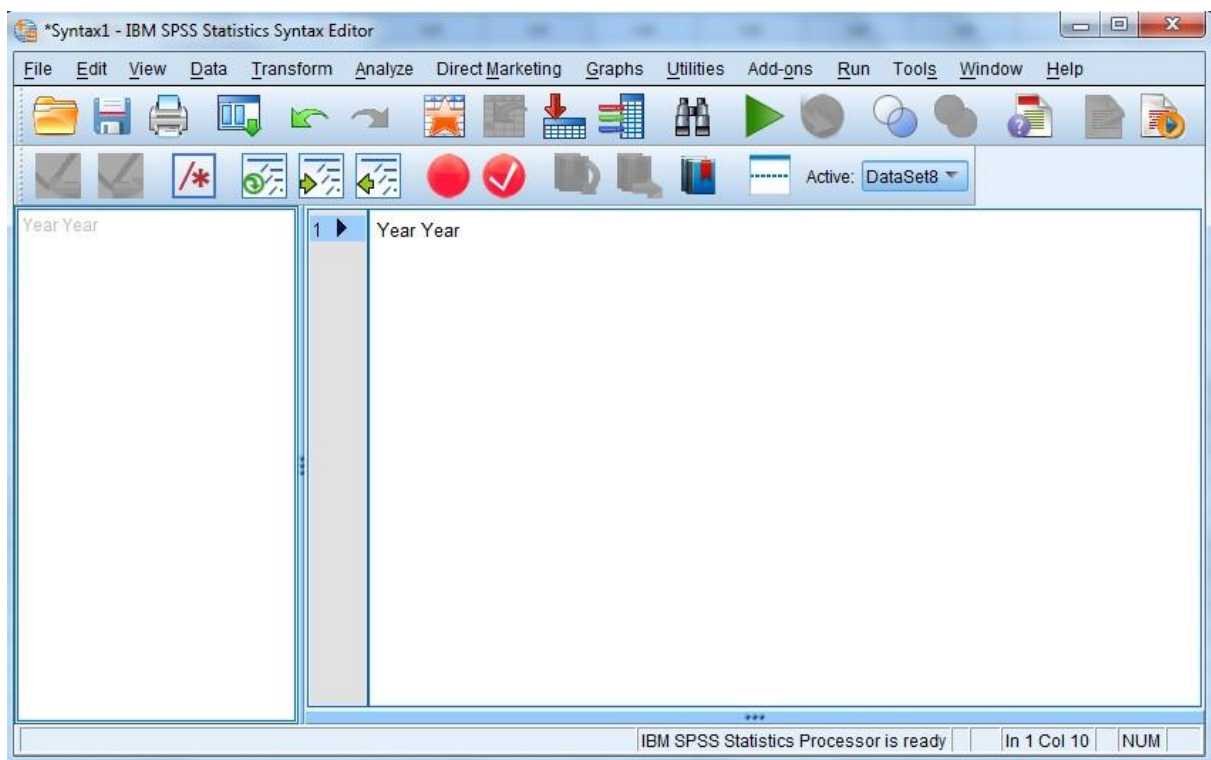
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Not Satisfied	4	28.6	28.6	28.6
Moderately Satisfied	4	28.6	28.6	57.1
Satisfied	2	14.3	14.3	71.4
Highly satisfied	4	28.6	28.6	100.0
Total	14	100.0	100.0	

In the output window, the right side is the output from the SPSS procedures that were run and on the left side is the outline of that output. The SPSS output is actually composed of series

of output objects that can be titles like frequencies, descriptive, cross tables, table of charts and numbers etc. Each of these objects is listed in the outline view. The outline view makes navigating the output easier. If you want to move to other output, you merely need to click on the word in the outline view and the output selected will emerge in the output windows. Similarly, if you want to delete, simply click on the word in the outline view and then click on the word in the outline view and then click Delete. If you want to move some output from one section to another you can select an output object and select Edit and then click Cut. Further, choose Edit and click Paste After. If you want to see notes then just double click on the closed book icon to the left of the notes title. The closed book icon will then become an open book icon and the notes will materialise in the window to the right. Double click again and the notes will disappear and the book will close.

1.4.2.2 THE SYNTAX EDITOR

It records the sequence of the operations that are “pasted”. This window is activated when you click on the paste function. The editor of the Syntax Command language is capable of recognising various language parts and highlighting them for you.



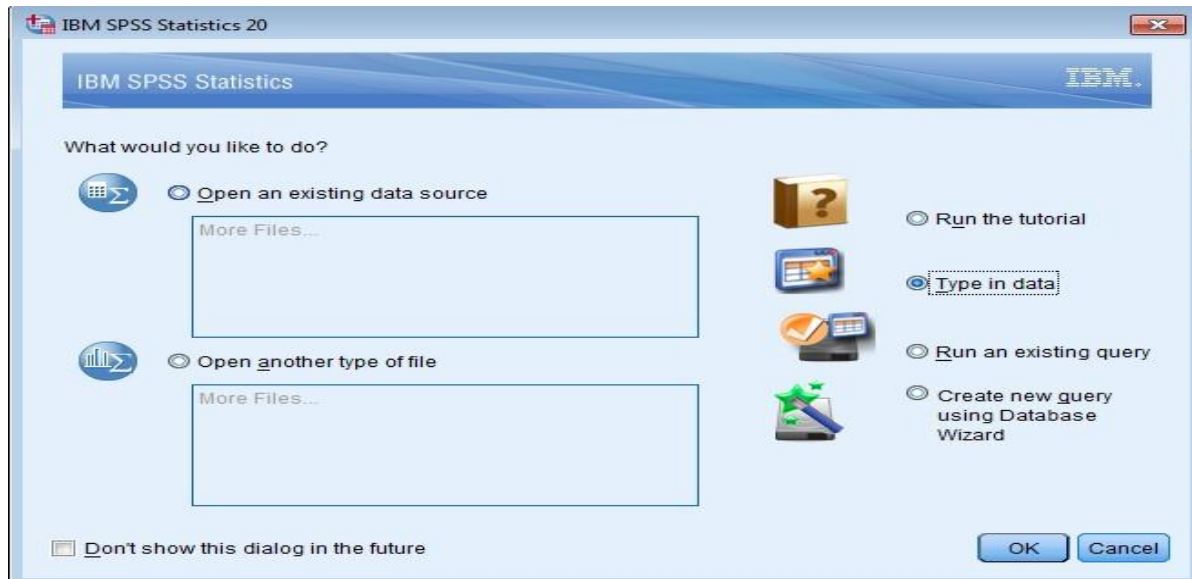
1.4.2.3 THE SCRIPT EDITOR

It is used to determine some fundamental defaults about scripts. This setting determines which script editor is launched when new scripts are created. The default script language is Basic. No other choice is available unless you have the Python add-on installed.

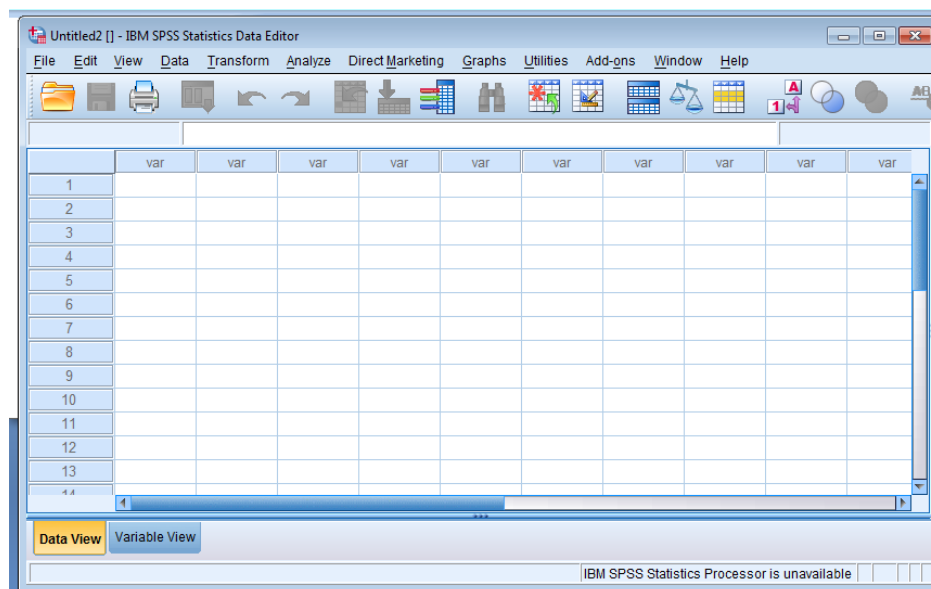
1.5 CREATING A DATA FILE

The first step in using SPSS is to enter your data and create a SPSS data file . SPSS can read data files created by other programmes as MS. Excel, SAS STATA, MS Access, ASCII files and .sav files created in earlier versions of SPSS.

In order to create a new data file, you have to click Type in Data in the window shown below and then click OK.



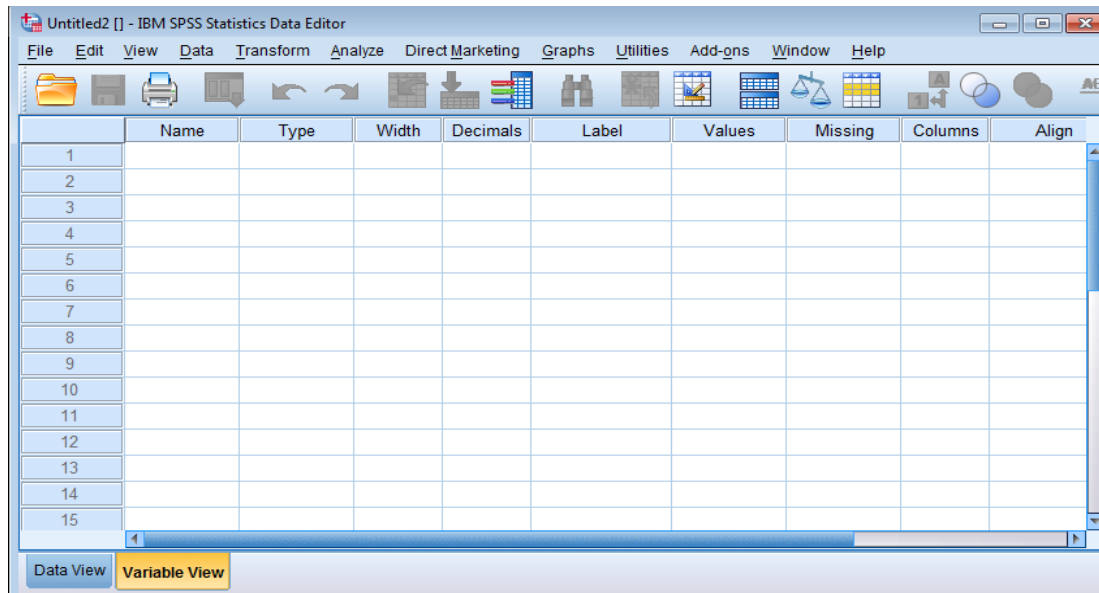
This will open data editor window shown below.



SPSS Considers each row as respondents or cases. Columns are used for entering variables information.

1.5.1 DEFINING VARIABLES

In order to define variables for entering data set you have to select variable view button from bottom left corner of data editor window.



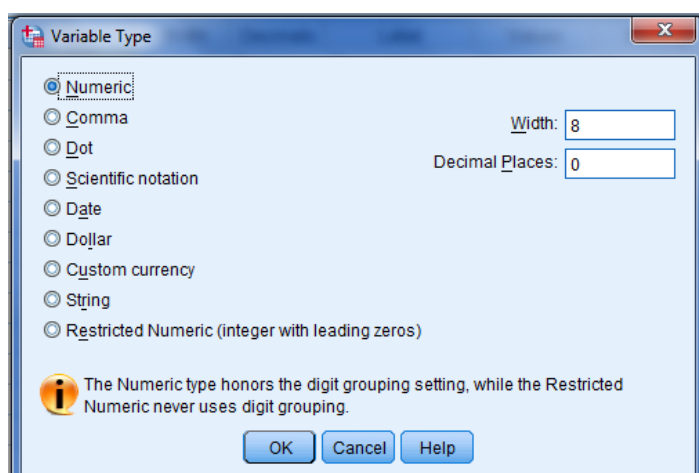
In this view you have to define characteristics of variables. The following variables characteristics would appear in the variable view-

1. Name- You should specify the name of the variable here. The first character needs to be alphabetic; however, the remaining can be numeric or alphabetic, symbol etc. While naming the variable the following have to be kept in mind:

- There should not be any space while naming.
- It should not end with full stop or underscore.
- Duplication has to be avoided.
- Variable names are not case sensitive to upper or lower case.
- It should not include words like all, with, not, get, or etc.
- Special characters like *,", 1 etc. should not be used.
- Maximum name can be written up to 64 characters or bytes.

Note- If you want to export data to another application, then you should ensure the names you use are in a form acceptable to that application.

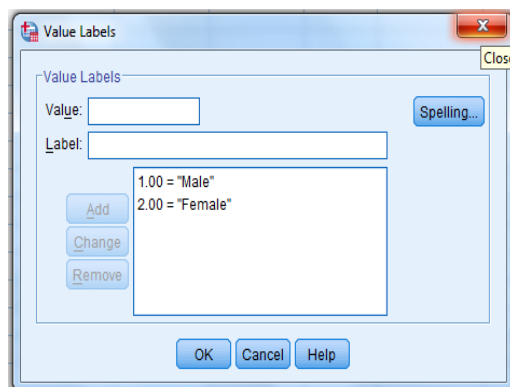
2. Type- A gray box would be displayed when you click to TYPE. Here you have to specify the nature of the data. Most of the data you will enter into SPSS would be numeric.



Options	Details	Width	Decimal places
Numeric	Standard numbers are recognizable form.	The total number of all characters. Also include any positive or negative signs.	It denotes the total number of digits displayed after the decimal point.
Comma	Explains numeric values with commas inserted between three digit groups.	Denotes total width including commas and decimal point.	Specifies number of digits after decimal.
Dot	Used to group the digits into three and comma are used for decimal point.		
Scientific Notation	A numeric variable that always includes the E to designate the power of ten exponents. The base, the part of the number to the left of E. Example- 6.346E3	Determines the number of characters used to display the value.	It denotes the total number of digits displayed after the decimal point.
Date	It includes the year, month, hour, minute, day etc.	--	--

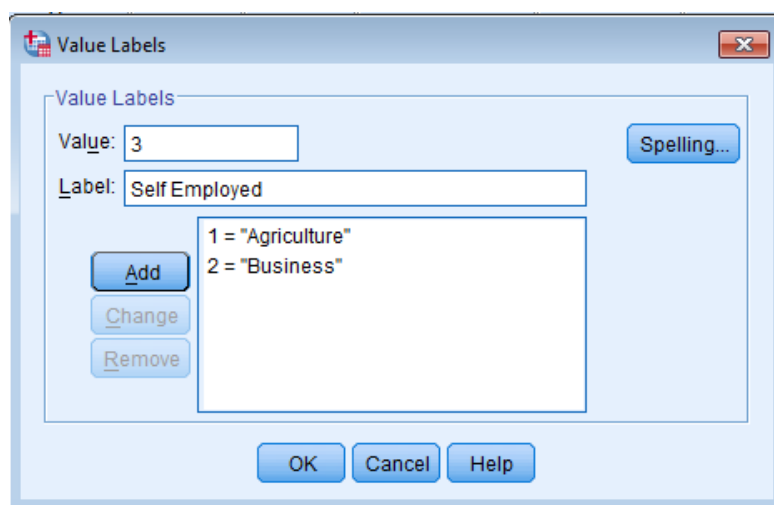
Dollar	Dollar values are always displayed with a leading dollar sign and a period for a decimal point and, for large values will include commas to collect the digits.	Helps in vertical alignment in the output.	
Custom Currency	The five custom formats for currency are named CCA,CCB, CCC,CCD and CCE. It can be changed by choosing edit options and then currency tab.		
String	It is freeform nonnumeric item. It will not be used for calculations.	Length would be specified in characters.	

3. Label- It helps you in explaining variable whose meaning is not clear from the variable name. For example, Occu in variable name can be explained as Occupation, gender can be explained with longer label as Men and Women or Boys and Girls.

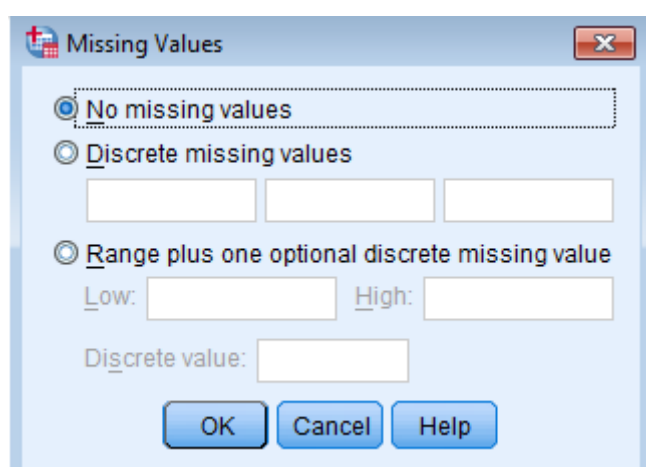


4. Values- The values column allow you to identify levels of a variable. Entering value label is very important for interpreting output. It allows up to 60 characters for each value label. A click on any cell in the values column will lead to the small blue or greyed box. Type the value code in value box, type value label in label box and then add. In order to change particular values you can use change button. Further, remove is used for removing particular entries and spelling is used for checking spelling.

For Example, for a variable name Occu you could have the value 1 assigned the label Agriculture, 2 assigned for Business and 3 for Self Employed and 4 for Others.



5. Missing- Its purpose is to designate different types of missing values in your data. It is generally found in questionnaires that some of the questions are not answered by the respondents. In such cases, some missing values are there that need to be dealt with by the researcher.



Missing value codes must be of the same data type. Generally by convention, the missing value is taken as 9. However, you can give different types of missing values in data. For example, those who refused to answer an occupation question might be coded as 8 and those who were of a different occupation other than listed might be coded as 9. You can specify up to three specific values (called discrete values) to represent missing data or you can specify a range of numbers along with one discrete value, all to be considered missing.

6. Columns- Columns is where you specify the width of the column you will use to enter the data. It explains the column width that is how much space needs to be provided to the data and labels.

7. Align- The Align column provides a drop-down menu that allows you to align the data in

each cell as right, left or centre. Aligning to the left means inserting all blanks on the right; aligning right inserts the entire extra spaces on the left; centering the data splits the spaces evenly on each side.

8. Measure- The level of measurement should be correctly defined otherwise the analysis or results would not come accurately. It depends on the qualitative or quantitative nature of the variable. It allows you to select three options based on the nature of your data into scale, ordinal and nominal.

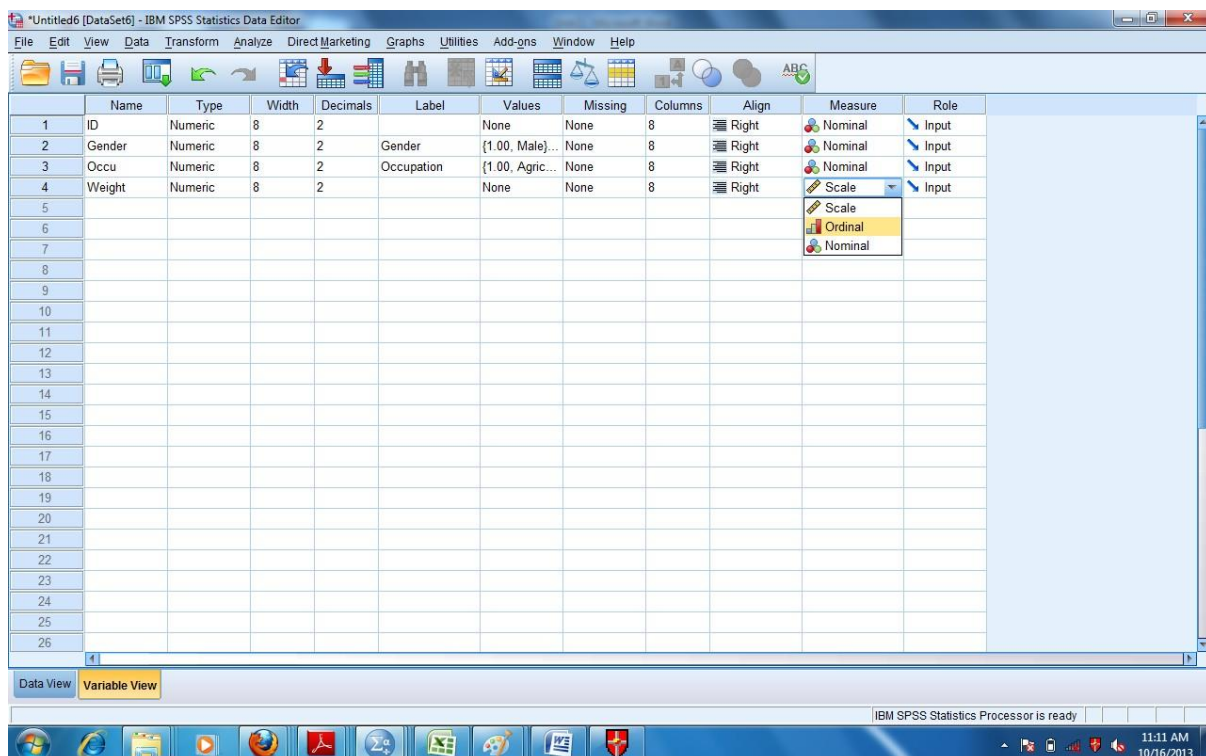
Scale- A number that specifies a magnitude. Scale is default for numeric variables. It can be distance, weight, age or count of something. Its technical name is cardinal.

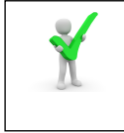
Ordinal- These numbers denotes the position or order of something in a list.

Nominal- It is used for identification but have no intrinsic order i.e lesser or greater. It specifies categories or type of things.

9. Role-It explains the variables according to their role and contribution in analysis as used for input or output or both. The other options are like split, partition, target etc. will require understanding of concepts of split file and regression analysis.

Note: In all analysis SPSS handles both ordinal and scale variables identically.





Check Your Progress-A

Q1. What is the function of data editor in SPSS?

Q2. Explain the procedure for adding variables to the data file.

Q3. How you can value labels in SPSS?

Q5. Which of the following command helps in switching over from one window to another or for resizing windows?

- A. Windows
- B. Transform
- C. Help
- D. View

Q7. Which of the following records the sequence of the operations that are “pasted”?.?

- A. Script Editor
- B. Syntax Editor
- C. Data Editor

Q8. Fill in the blanks with appropriate word or words.

- a) The _____ assimilates the results of the work done in SPSS.
- b) SPSS can read data files created by other programmes as files created in earlier versions of SPSS.
- c) SPSS Statistics was released in its first version in year.....
- d) SPSS was developed by

1.6 DATA MANAGEMENT

Data plays a significant role in any research. Therefore for you as researchers it is necessary to manage the data in SPSS so that the collected data can facilitate statistical analysis. This part will help you in entering data, storing data, sorting data, data aggregation etc. and other related procedures.

The following are its various facets:

1.6.1 ENTERING DATA

After creating the variables data would be entered in the Data View. You can enter data by variable or by case. You may enter the data in the data editor as you do it in Ms Excel and enter the values accordingly. First click on the first empty cell under the first variable, type the value and then press the Down Arrow Key, then type the next number. Instead of Down Arrow Key you can also use Enter Key and while entering data by case you may use Right-Arrow Key or TAB key.

	ID	Gender	Occu	Weight	var	var	var	var	var	var	var	var	var	var
1	1.00	1.00	1.00	70.00										
2	2.00	2.00	3.00	60.00										
3	3.00	1.00	2.00	68.00										
4	4.00	2.00	4.00	55.00										
5	5.00	2.00	2.00	67.00										
6	6.00	1.00	1.00	59.00										
7	7.00	1.00	3.00	60.00										
8	8.00	1.00	2.00	77.00										
9	9.00	1.00	1.00	69.00										
10	10.00	2.00	2.00	50.00										
11	11.00	1.00	2.00	68.00										
12	12.00	1.00	3.00	69.00										
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														
24														

1.6.2 EDITING DATA

For editing any value while entering data you simply have to click on the cell of interest, type the new value and then hit enter. When a file has many variables and if you want one or two column always visible then you have to place the cursor at the bottom right of the screen, look for a small area to the right of the vertical scroll bar and drag the line to the column or variable you want to refer while entering data.

	id	Gender	reading	writing	acadindx	var	var	var	var	var	var	var	var	var
1	1.00	1.00	34.00	44.00	147.00									
2	2.00	1.00	39.00	41.00	143.00									
3	3.00	2.00	63.00	65.00	179.00									
4	4.00	1.00	44.00	50.00	148.00									
5	5.00	2.00	47.00	40.00	156.00									
6	6.00	1.00	47.00	41.00	154.00									
7	7.00	2.00	57.00	54.00	174.00									
8	8.00	1.00	39.00	44.00	164.00									
9	9.00	2.00	48.00	49.00	164.00									
10	10.00	1.00	47.00	54.00	170.00									
11	11.00	2.00	34.00	46.00	152.00									
12	12.00	2.00	37.00	44.00	152.00									
13	13.00	1.00	47.00	46.00	154.00									
14	14.00	2.00	47.00	41.00	164.00									
15	15.00	2.00	39.00	39.00	138.00									
16	16.00	2.00	47.00	31.00	148.00									
17	17.00	1.00	47.00	57.00	160.00									
18	18.00	2.00	50.00	33.00	161.00									
19	19.00	1.00	28.00	46.00	155.00									
20	20.00	2.00	60.00	52.00	186.00									
21	21.00	2.00	44.00	44.00	179.00									
22	22.00	2.00	42.00	39.00	163.00									
23	23.00	1.00	65.00	65.00	190.00									
24	24.00	2.00	52.00	62.00	181.00									
25	25.00	1.00	47.00	44.00	152.00									

1.6.3 MANAGING DATA FILES AND DATA MANOEUVRING

In this part of unit you would be grounded with the tools and techniques for data management. Since for analysing the data from different perspectives or dimensions you need to derive required data from the original data file. Once all the data is entered you can save the data file by clicking File Menu and Save and then you have to enter the name of the file and choose the location and then press save. The file will save with an extension .sav.

Manipulation or Manoeuvring of the Data- To enter SPSS a click on Start in the task bar activates the start menu. After clicking the SPSS programme you could create and name a data file or edit an already existing file. The data can be manoeuvred using the tools mentioned from 1.6.4 to 1.6. 14.

1.6.4 SPLITTING FILES

Split file command is useful for performing same statistical analysis on separate groups based on the values of one or more grouping variables. For example if data is collected on several characteristics of households and the so collected data is grouped by gender, ethnicity, income and category, then Spilt File Command can be used. Split file command has two options Compare options and organise output by groups. You can choose the options as per the requirement of research questions. For example, if you select Gender as grouping variable, you can find descriptive statistics like mean, median, standard deviation separately for Males and Females.

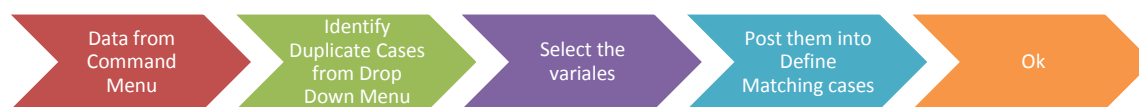
For splitting the file you have to first click Data from the Command menu and the select split file, you will get a dialogue box after this you will select Organise output by groups and move gender from the variable list box to the Groups Based on list box, then click OK. The file will be sorted on the basis of Gender Value in ascending order. To find out the descriptive for age, income, occupation or any other variable separately for males or females, you just have to use single command of descriptive statistics. This command executes same statistical tests to be performed on different categories of respondents in the same dataset.

Further, when you will execute the command you have to ensure that you select the “Sort the file by grouping variables”. Further if you have already sorted the file on the basis of gender then you click to “File is Already Sorted”.

- By selecting **Compare groups**, results for both males and females will appear in the same output table.
- By selecting **Organize output by groups** the group’s results will be displayed on two separate tables denoting one for each group.

1.6.5 IDENTIFYING DUPLICATE CASES

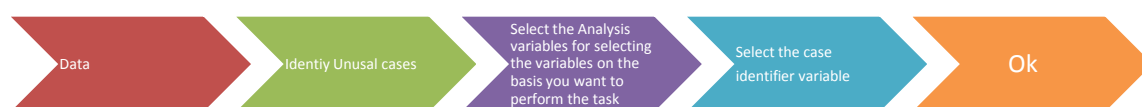
This is used to track errors in data entry “Identify Duplicate Command” helps us in detecting existence of duplicate entries while adding data into SPSS. This is also termed as Data Cleaning. For doing this go through the following path:



The output would be displayed in the output viewer window. First, you will have the title Frequencies followed by the two tables- Statistics and Indicator of Each Last Matching Case as Primary. Statistics table will show the number of cases taken into consideration while finding duplicate cases. Indicator of Each Last Matching Case as Primary table shows the number of duplicate entries in the entire dataset. The cases so detected are placed at the top of the data file. Further, the last column PrimaryLast is created on execution of Identify Duplicate Case as by default Last case in each group in Primary. There are two values as 0 and 1. 1 will proceed after 0 that means the 1 means the primary or unique case and 0 indicates a duplicate entry of the same case.

1.6.6 IDENTIFYING UNUSUAL CASES

This function will help you in identifying unusual behaviour of samples which deviate from the group norms to which they belong to. You can track those using the following path:



After this, anomaly so detected will be displayed as the title of the output under the main title Detect Anomaly. In this three tables would be displayed. First Anomaly index is calculated. Next, Anomaly Case Reason explains the reason for anomaly, impact of the variable, the value of that variable and variable norm as per the cluster behaviour. Further, if you take the reciprocal of total number of variables used for anomaly detection and if the variable impact value is greater than the reciprocal of number of variables then the impact of this variable is said to be large. Variable norm is the average value of that variable for the cluster in which anomalous case lie.

1.6.7 SELECTING CERTAIN CASES AND INSERTING CASES

Selecting a portion of the data for analysis is frequently used by researchers. The usage of Select cases helps in selecting a subset of data for analysis. It can also perform analysis on Random Sample of Cases. For example, if you want to perform what the mean total score is for females, then you needs to follow the path mentioned as under-



In If then click gender mention =1 and then press continue and O.K

Further, if you want to cross check identified cases that are selected then it is easily recognisable by the diagonal line placed through the row number for the non selected cases. In this example, the lines through case number represent males in the sample. Further, SPSS also creates a new variable named filter_\$ which codes selected cases as 1 and non selected cases as 0. Moreover, if you will select ALL cases option then the Select Case will automatically get turn off.

If you need to insert case or variables then you can use Insert cases or Insert Variables from Edit menu in the tool bar. You will then have to click to any cell in the case below the position where you wish to insert the new case and click insert cases from edit menu. On the other hand if you want to insert any cell in variable to the right of the position where you want to insert the new variable and click insert variable from edit menu.

1.6.8 DATA AGGREGATION

If you want to aggregate the data collected cumulatively like you might have collected data on

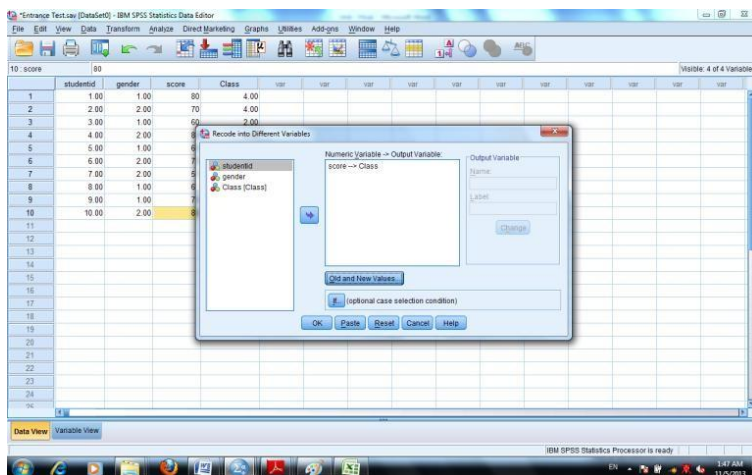
entrance test marks attained in English language Comprehension, Reasoning, Mathematics and General Knowledge and you want to know their mean score or knowing the highest level of total marks attained then it is possible in SPSS, using aggregate command. This command performs arithmetic or statistical operations on similar categories of respondents. Functions like sum, maximum, mean etc. can be calculated for cases or respondents with the identical attributes.

Path:



In Break Variable you have to specify the level variable (Marks on different parameters) on which aggregation is to be performed. Then in “Select The Variable” you have to select Score in Entrance Examination and transfer it into the Summaries of Variables list box. Then you have to click Function button then another dialogue box will open in which you can select mean, maximum, that you want to perform on the variable and click continue. In Save Frame you have to select Create a new dataset containing only the Aggregated variables and give some name to the dataset. In options, for very large datasets, select Sort File before aggregating and then OK. A new file will be formed on the applying the aggregate command.

1.6.9 RECODING VARIABLES



This procedure also creates new variables by dividing a pre existing variable into categories and coding each category differently. Recode sub menu of Transform menu have two options i.e recode into same variables or recode into different variables, when you choose same variables then the original values would get replaced with the new recoded values. In second option both the values old and new would be available. For example if the Head of the Department wants to divide the scores attained by the candidates in entrance test into different grades. In this, coding is based on the percent variable and now you want to recode the score variable in class of 40 to 50=1, 50 to 60= 2, 60 to70=3, 70-80=4, 80-90 =5 and 90 to 100= 6 . Now in this case Click transform go to Recode into different variables. A dialogue box would appear as shown below:

Click on the variable which you want to recode and take it to Numeric Variable-> Output variable and mention new variable name and then click Change. Your next step would be clicking Old and new values button. Click Range Lowest through Value button and mention the number you want to depict in new value box . Repeat this step for all the values and then click continue and O.K to execute the command.

The image shows two screenshots from the SPSS software. The top screenshot is the 'Recode into Different Variables: Old and New Values' dialog box. The 'Old Value' section has 'Range' selected, with '90' in the 'through' box and '100' in the 'value through HIGHEST' box. The 'New Value' section has 'Value' selected with '6' in the input box. The 'Old --> New' list contains: '40 thru 50 --> 1', '50 thru 60 --> 2', '70 thru 80 --> 4', '60 thru 70 --> 3', and '80 thru 90 --> 5'. The 'Output variables are strings' checkbox is checked with a width of 8. The 'Convert numeric strings to numbers' checkbox is unchecked. The bottom screenshot shows the SPSS Data Editor window with the 'score' variable selected. The data table is as follows:

10 : score	studentid	gender	score	Class	var	var	var	var	var	var	var	var	var	var	var
1	1.00	1.00	80	4.00											
2	2.00	2.00	70	4.00											
3	3.00	1.00	60	2.00											
4	4.00	2.00	80	4.00											
5	5.00	1.00	60	2.00											
6	6.00	2.00	70	4.00											
7	7.00	2.00	50	1.00											
8	8.00	1.00	60	2.00											
9	9.00	1.00	70	4.00											
10	10.00	2.00	80	4.00											
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															

A new variable with class will be created in the data view with the codes given.

1.6.10 INSERTING NEW VARIABLES

If you want to insert cases or variables then you can use insert cases or insert variables from edit menu in the tool bar. You have to left click your mouse below the row where you want to insert case then go to the Data drop down menu and select Insert Case or you can select any cell in the variable to the right of the position you want to insert the new variable and click insert variable from edit menu.

Note: Here case denotes a row.

1.6.11 DELETING VARIABLES AND CASES

If you want to delete cases or variables then you have to pick clear from the Edit Menu in the tool bar. You have to click on the case number on the left side of row or select any cell in the row in which you want to delete and then click clear from the edit menu. Further, if you want to delete a variable then select the variable name that is the column in which you want to delete and then click clear from the edit menu.

1.6.12 SORTING CASES

If you want to sort cases then you have to select a variable of interest and specify which case you want to sort in ascending or descending in the Sort Cases Window which would be located in Data Menu

1.6.13 MERGING FILES

The task of data entry is carried on the basis of number of questionnaires. If there are large number of questionnaires or if modules in questionnaire are divided into different data operators then the segregated data will have to be merged to get the full data set. However, while merging you have to ensure that format files are in same data editor and identical formats are created for the each variable. You need to ensure that matching variables have identical names and the cases or variables are in the same order.

1.6.14 MERGING FILES BY CASES

It will be used if different operators are given different sets of full questionnaires for data entry. Here each respondent is a case. For the same you have to first access the file in which you want to merge and then you have to click Data then Merge Files and then Add Cases.



A dialogue box would appear on the screen. Variables in the New Active Dataset list box would be the list of variables in the active dataset as well as new dataset. If there are some variables in the external file which are not in the active dataset then those variables would be listed in the Unpaired Variables. After this you have to click the OK button to execute the command.

1.6.15 MERGING FILES BY VARIABLES

It is used when there are different data entry operators enter data of different modules of questionnaire of the same set of respondents but for different questions. For this you have to add variables to the Master file for preparing complete data set. The procedure would be that first you have to access the file in which you want to merge and then you have to click Data then Merge Files and then Add Variables.

Path:



A dialogue box would appear on the screen as *Add variables from* With the name and source of your external file included in the title. The matching variables will be listed in the Excluded Variables Window and each will be followed by a “+” sign in parentheses. Further, the variables that are in the original file that are not present in the external file are listed in the New Active Dataset Window and each one will be followed by an asterisk “*” in parentheses. It is required that the matching variables shall be in the same order identically for both the file. You have to then click to Match Cases on Key Variables in sorted files. Select the matching variable in the box to the left click the arrow button and then, click OK.

1.7 SUMMARY

This unit has introduced us with the conceptual understanding of the various characteristics of SPSS software. We learnt that the basic operations of SPSS using tool and menu bar. This unit also explained how to create, edit and format a data file. We also learnt how to save and edit output computed by SPSS. In the unit, we also looked procedures for entering data, storing data, retrieving stored data, sorting data and other related procedures. Now, in the next unit, we will be describing data numerically using SPSS.



1.8 GLOSSARY

Menus: It appears on the top of the window and just below the title bar.

Variable: It is a characteristics of an item and is represented as column in SPSS.

Variable View The screen within the SPSS Data Editor where the characteristics of variables are assigned.

Cases: A row in the Data Editor file indicates the data collected from a single participant.

Label: It helps you in explaining variable whose meaning is not clear from the variable name.



1.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress –A

5. a. Windows
6. b. Syntax Editor
7.
 - A. Output window
 - B. MS. EXCEL, SAS STATA, MS ACCESS, ASCII FILES AND .SAV
 - C. 1968
 - D. Norman H. Nie, Dale H. Bent, And C. Hadlai Hull.



1.10 REFERENCES

1. Gupta S.L. and Gupta Hitesh, SPSS 17.0 for Researchers, International Book House Pvt. Ltd., New Delhi
2. Pandya Kiran, Bulsari Smruti, Sinha Sanjay(2012) , SPSS in Simple Steps, Dreamtech press, New Delhi
3. Hooda R. P(2000), Statistics for Business and Economics, Macmillan India Ltd.
4. George Darren and Mallery Paul (2011), SPSS for windows Step by Step, Dorling Kindersley Publishing



1.11 SUGGESTED READINGS

1. Gupta S.L. and Gupta Hitesh, SPSS 17.0 for Reserachers, International Book House Pvt. Ltd., New Delhi
2. Pandya Kiran, Bulsari Smruti, Sinha Sanjay(2012) , SPSS in Simple Steps, Dreamtech press, New Delhi
3. Hooda R. P(2000), Statistics for Business and Economics, Macmillan India Ltd.
4. George Darren and Mallery Paul (2011), SPSS for windows Step by Step, Dorling Kindersley Publishing
5. *IBM* reference guide posted on the SPSS website.
6. Landau Sabine and Everitt Brian S „A Hand book on statistical analysis using SPSS“, free downloadable



1.12 TERMINAL QUESTIONS

- Q1. Explain in detail various command menu options available in SPSS?
- Q2. List steps for entering data in SPSS.
- Q3. Write a brief note on the Data Aggregation and Data Manoeuvring.
- Q4. Go to funds. sav , do the following exercises;
- a) Enter the hypothetical figure of returns obtained from the funds invested by the individuals
 - b) Search for the individuals which are from science stream.
 - c) Print out age, stream and returns obtained., fit it into one page.