Lecture 5: Data Transformation and Data in SPSS

<u> Tasks :</u>

Recode into Different Variable:

- Modify data values of academic ability into 10 categories that ranges from 1-10. [For instance, student who scored in the ranges '0-9' into 1, '10-19' into 2, and so on. The last one will assign the values '90-100' into 10.]
- Modify the academic ability of students who scored in the ranges "0-39" into fail, '40-49' into "marginal fail", '50-59' into pass, '60-69' into merit" and '70-100' into honor categories.
- 3. If it is a matter of "pass" or "fail", it can also be recoded into fail for range '0-49' and pass for range '50-100'.

Recode into Same Variable:

 Decompose the gender variable using **recode** procedure into two groups such as male and female. [Note that by selecting into Same Variable, you will overwrite the original data.]

Compute:

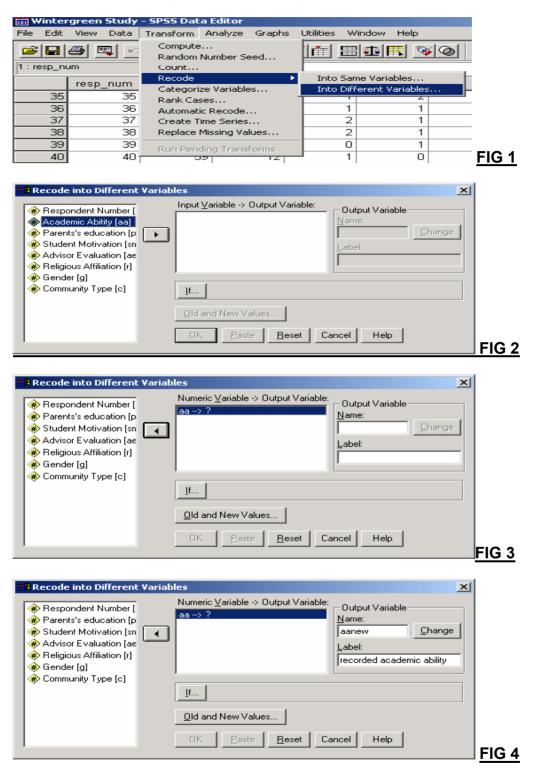
- 5. Compute a new variable by adding 1 to "**aa**" and multiply this with 0.1.
- 6. Create the natural log of the existing variables.

<u>Data:</u>

- 7. Select the students who come from a rural community.
- 8. List the first 10 students' parent education.

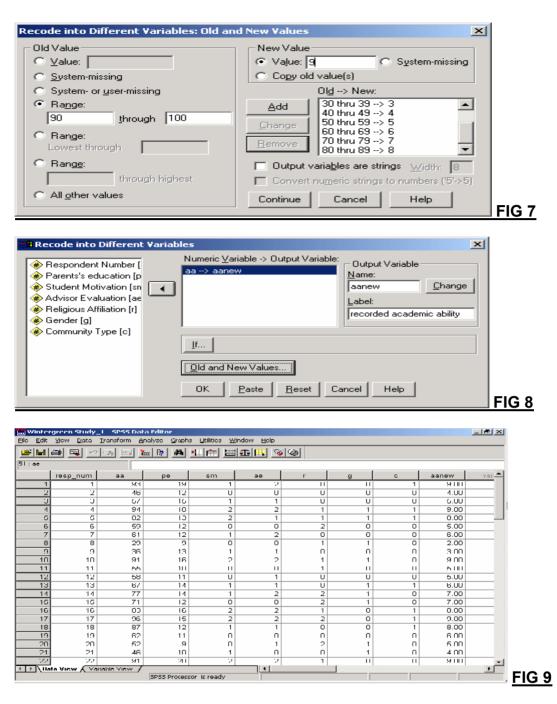
Tips For Task 1

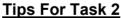
1. This task is accomplished by recoding the data.



Recode into Different Variables	×
 Respondent Number [Parents's education [p Student Motivation [sn Advisor Evaluation [ae Religious Affiliation [r] Gender [g] Community Type [c] 	Values
	Paste <u>R</u> eset Cancel Help <u>FIG 5</u>
Recode into Different Variables: Old an	d New Values
- Old Value	- New Value
	Value: O System-missing
C System-missing	C Copy old value(s)
C System- or user-missing	Old> New:
Range: through Range: Lowest through	Add Change Remove
C Range:	Output variables are strings Width: 8
through highest	Convert numeric strings to numbers ('5'->5)
C All other values	Continue Cancel Help FIG 6

- Select old value **Range** and enter "**0**" through "**9**". Then enter "1" for the **New Value,** and click the **Add** button.
- Return to the Range box and enter "10" through "19" for the range, return to the New Value box, enter a "2" for the value, and click the Add button. Continue for the remaining groups of ranges (the last one will assign the values "90" through "100" to group number "10") (see Fig 7).
- If you make a mistake along the way, you can correct it by using the **Change** and **Remove** buttons.
- Once you have completed defining groups, click the **Continue** button.
- You will return to the previous dialog box (see Fig 8). If you now click the **OK** button, the new variable will be created and it will be added into the dataset as "**aanew**" (see Fig 9).





From Transform>Record>Into Different Variables, the range of numbers can be prepared by clicking on the forth radio button and entering the values 0 through 39, then click on the check box output variables are strings and type fail into the new value box and click on add button (before start to apply the new process, you need to remove the previous numbers or interval by using the remove button). This procedure should be repeated for the other intervals (i.e. "honor", 70-100) (see fig 10). In this step, you can use variable aanew to change it as aanew1 if you want it to be added into the database under a new variable name. In the end, you can return to the previous dialog box (see fig 11). In order to see the final outcome, refer to fig 12.

Recode into Different Variat	oles: Old and	New Values		×	1
- Old Value		- New Value			
		Value: ho	nor o s	ystem-missing	
C System-missing		C Copy old v			
C System- or user-missing			01 <u>d</u> > New:		
• Range:			0 thru 39> 'fail'		
70 <u>through</u> 10	00		40 thru 49> 'ma 50 thru 59> 'pa:		
C Range: Lowest through			60 thru 69> 'me		
C Range:			ria <u>b</u> les are strings	1. Gebber 🛛	
through high	nest		umeric strings to n		
C All <u>o</u> ther values		Continue	Cancel	Help	
					<u> </u>
					-
Recode into Different Varial	bles			×	1
Respondent Number [able -> Output Varia	able: 🗖 Output Varia	able	
Parents's education [p	aa> aanew	1	Name:		
🛞 Student Motivation [sn			aanew1	<u>C</u> hange	
Advisor Evaluation [ae			Label:		
Religious Affiliation [r]			recorded aa		
() Gender [g]	J		1		
Community Type [c] Community Type [c]					
recorded academic at	<u>l</u> f				
	(<u>O</u> ld and New	w Values			
				-1	
I		Paste <u>R</u> eset	Cancel Help	<u>'</u>	
					_ <u>FIG 11</u>
					-
Wintergreen Study_1 File Edit View Data Tr.		a Editor nalyze Graph	is Utilities W	/indow <u>H</u> elp	
		- <u>?</u> 44			>
51 : ae		_			т
1 0	g O	<u>с</u> 1	aanew	aanew1 honor	Ļ
2 0	0	0		marginal	+
3 0	0	0	5.00		†
4 1	1	1	9.00	honor	Ť
5 1	1	1	8.00	honor	Ť
6 2	0	0	5.00	pass	Ť
			C 00		+
7 0	0	0	6.00	merit	
7 0 8 1	0 1	0	2.00		+

Tips For Task 3

In the second step the first range of numbers **lowest thru 49** is prepared by clicking on the fifth radio button and entering the value 49 in the box, then click on the check box **output variables are strings** and type **less** into the **new value** box and click on **add** button (this time you need to use another variable i.e. aanew2 (see Fig 13).

The second range of numbers 50 thru highest is prepared in a similar manner after clicking on the sixth radio button, entering the value 50 in the relevant box, and the more in the new value box (see also Fig 13).

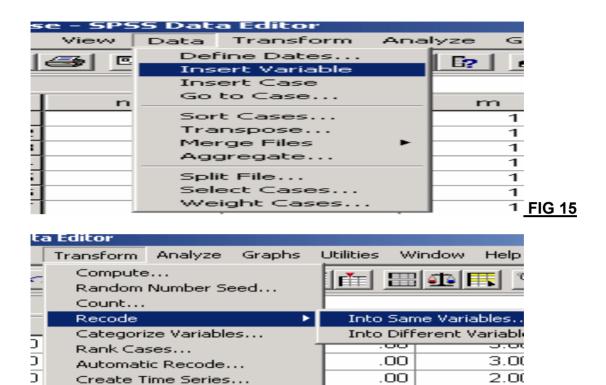
Recode into Different ¥ariables: Old a	and New Values
Old Value C ⊻alue: C System-missing	New Value System-missing Copy old value(s)
C System- or <u>u</u> ser-missing	0l <u>d</u> > New:
Range: through Range: Lowest through	Add Lowest thru 49> 'less' Change Bernove
Range through highest All other values	Output variables are strings Width: 8 Convert numeric strings to numbers ('5'>5) Continue Cancel Help
	FI

Now, you will return to the previous dialog box and if you click on the **OK** button, the new variable will be created and added into the data set as **aanew2** (see Fig 14).

📰 Winter	rgreen Study	_1 - SPSS Dat	ta Editor			
<u>File E</u> dit	⊻iew <u>D</u> ata	<u>T</u> ransform <u>A</u>	nalyze <u>G</u> raph	s <u>U</u> tilities <u>W</u>	<u>/indow H</u> elp	
2	a 🖳 🗠		- I? M	× 🛍 🗄	1 1 1	0
51 : ae						
	r	g	С	aanew	aanew1	aanew2
1	0	0	1	9.00	honor	more
2	0	0	0	4.00	marginal	less
3	0	0	0	5.00	pass	more
4	1	1	1	9.00	honor	more
5	1	1	1	8.00	honor	more
6	2	0	0	5.00	pass	more
7	0	0	0	6.00	merit	more
8	1	1	0	2.00	fail	less
9	0	0	0	3.00	fail	less
40			-	0.00		1

Tips For Task 4

Create two new variables as *m* (male) and *f* (female) by using *Data* > *Insert Variable* (see Fig 15). (These variables can be generated conducting by copy and paste from the original gender variable.) Then, decompose the gender variable using *Recode* procedure into two groups such as male and female. Now, click on **Transform> Recode** and select **Into Same Variables** (see Fig 16).



 Replace Missing Values...
 .00
 2.00

 Run Pending Transforms
 .00
 4.00

 The resulting dialog box, shown in Fig 17, contains a list of your variables on

the left. Put the relevant one (i.e. gender) into the right box and then click on old and new values.

Recode into Same Variables	×
Image: Numeric ⊻ariables: Image: Numeric ⊻ariables: Image: Imag	OK <u>P</u> aste <u>R</u> eset Cancel Help
recoded level [ibne	
Approximately 50 %	

Change gender as male and female. For female just change the variable name as female. For male click on the **Old and New Values** to open **Recode into Same Variables: Old and New Values** sub-dialogue box.

- In the Old Value box, type the old value (i.e. 0). In the New Value box, type the new value (i.e. 2). Then, Click on the Add Command pushbutton.
- Then, in the Old Value box, type the old value (i.e. 1). In the New Value box, type the new value (i.e. 0). Then, Click on the Add Command pushbutton.
- Two variables are formed as in figure 18.

m	f	
1.00	.00	
1.00	.00	1
1.00	.00	
1.00	.00	
1.00	.00	
1.00	.00	
1.00	.00	
.UU.	2.00	1
1.00	.00	1
1 00	00	
1.00	.00	
1.00	.00	
.00	2.00	
1.00	.00	<u>FIG 18</u>

Tips for Task 5

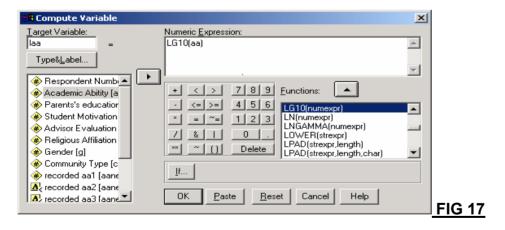
From the **Transform** pull-down menu, select **Compute.** To create the new computed variable, enter "aa**comp**" in the **Target Variable** box. Then, in the **Numeric Expression** box, enter **(aa+1)*0.1** and click the **OK** button to compute the new variable (see Fig 15 and 16).

👯 Compute Variable	×	1
Target Variable: aacomp Type&Label * Respondent Numb Academic Abitity [a * Academic Abitity [a * Academic Abitity [a * Advisor Evaluation * Advisor Evaluation * Religious Affiliation * Gender [g] * Community Type [c * recorded academic * recorded academic	Numeric Expression: (aa+1) * 0.1 + < > 7 8 9 - < = > = 4 5 6 ABS(numexpr) ANY(test, value, value,) ABSIN(numexpr) ABTIN(numexpr) ABTAN(numexpr) CDFNORM[zvalue) [f]	
🚯 recorded aa1 faane 🔼	OK Paste Reset Cancel Help	FIG 1

📰 Winter	rgreen Study	_1 - SPSS Dal	ta Editor		
<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	<u>T</u> ransform <u>A</u>	nalyze <u>G</u> raph	ns <u>U</u> tilities <u>V</u>	<u>V</u> indow <u>H</u> elp
2	a 🔍 🗠		- <u> </u>		14
51 : ae					
	С	aanew	aanew1	aanew2	aacomp
1	1	9.00	honor	more	9.40
2	0	4.00	marginal	less	4.70
3	0	5.00	pass	more	5.80
4	1	9.00	honor	more	9.50
5	1	8.00	honor	more	8.30
6	0	5.00	pass	more	6.00
7	0	6.00	merit	more	6.20
8	0	2.00	fail	less	3.00
9	i n	3.00	fail	less	3.70

Tips for Task 6

To create the natural log of the existing variables, go to **the transform dropdown menu** and choose **compute** option. Select the relevant function (**LG10**) from the **function box** and click function button to get it into the upper screen, and then choose the relevant variable (**academic ability**) with clicking the nearest button (see fig 17). To get the final product, look at the **data view** section. Apply the same process for **the other variables** and make sure the log version of the relevant variables exist in data view section.



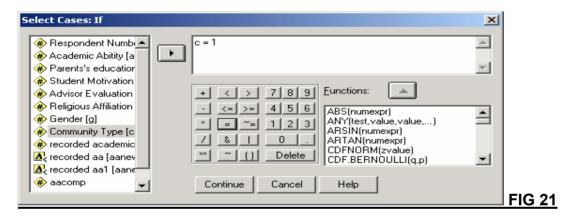
Compute Variable: Type and Label	×	
Label Iog of academic ability	Continue	
C Use expression as label	Cancel	
Type Mumeric	Help	
C <u>S</u> tring <u>W</u> idth: 8		FIG 1

Tips for Task 7

First, from the **Data** pull-down menu choose **Select** Case (see fig 19). Then select **If condition is satisfied** and click the **If** button (see Fig 20). Click on the variable "c" from the list on the left, then type **c=1**. Then click the **Continue** button, and finally click the **OK** button to select the cases (see Fig 21 and 22). When you select cases, you will notice that in the Data editor window **a slash mark** / appears through the record number of those cases that were not selected. Here you will also notice that a new column labeled **filter _\$,** and containing **1 and 0** has appeared. These two values represent the selected and unselected cases, respectively (see Fig 23). If you wish to include all the cases in later analysis, simply return to the **Data** pull-down menu and choose **Select Cases** dialog box and select **All cases**.

📰 Winter	green	Study	_1 - SPSS D	ata	Edita)r		
File Edit	View	Data	Transform	Ana	lyze	Graph	s Utilities	- V
훋 日 51 : ae	e 1	Ins	ine Dates ert Variable ert Cases		[?	#\$		
	resp	Go	to Case		ре	е	sm	
1		Sor	t Cases			19		1
2		Tra	nspose			12		0
3			structure			15		1
4			rge Files			18		2
5		- Agg	pregate			13		2
6			t File			12		0
7			ect Cases			12		1
8		We	ight Cases			9		0 510
0		0				10		<u> </u>

Select Cases	
Respondent Number [Academic Abitity [aa] Parents's education [p Student Motivation [sr Advisor Evaluation [ae Religious Affiliation [r] Gender [g] Community Type [c] Frecorded academic ab Pased on time or case range Based on time or case range Unselected Cases Are Filtered Dejeted	
Current Status: Do not hiter cases	
OK <u>Paste</u> Cancel Help FIG	20



👷 🖪 Select Cases		×
 Respondent Number [r Academic Abitity [aa] Parents's education [p Student Motivation [sr Advisor E valuation [ac Religious Affiliation [r] Gender [g] Community Type [c] recorded academic ab aacomp 	Select All cases If gondition is satisfied I	
Current Status: Do not filter ca	ises	_
	OK <u>Paste R</u> eset Cancel Help	

	rgreen Study Mew Data			is Utilities Wi	ndow Help					<u>– 7 ×</u>	1
						6					9
1.az											Office
	aanow	aanew1	aanew2	aacomp	filter_\$	var	var	var	var	var 🔺	8
1	9.00	honor	more	9.40	1		1				
	1 Z I I I	marginal	less	4 711							
	5.00	paco	more	5.80	0						⊇
4		honor	more	9.60	1						1
6	8.00	honor	more	8.30	1				1		2
6	5.00	pass	more	6.00	0				1		۵¥
7	6.00	merit	more	6.20	Π						
<u>8</u>	2.00	fail	less	3.00	0						=
	3 00	fail	less	3 70	n						#
~~t0	9.00	honor	more	9.20	0						0.020
11	1 s.uu	pass	more	5.60	U						13
~~ t2		расс	more	5.90	0						C
10	6.00	ment	more	6.00	1						3
14	7.00	honor	more	7.80	0						26
15	7.00	hunor	more	7.20	0						
16	800	hanor	more	8 40	1						e
17	9.00	honor	more	9.70	1						
18	8111	hanor	more	8 80	1						
19	6.00	merit	more	6.30	0						
20 بر	1 S.UU	pass	more	5.00	U						3
21	1 ∠.00	marginal	leco	4.70	0						S
	9.00	honor	more	9.20	0					-	0 S
▶\D:	ata View 🆯 🗸	ariable View /	,		_]				Ŀ	Microsoft
			SPSS Froces	sor is ready	J			Hiter On	J		14
A Start	📶 🈂 🖻	🗊 🛛 छjwcc	k2 Microsoft V	vord 📴 Se	ossion6ma Mic	rosoft W	🛗 Wintergr	een Shidy_1.		🍕 🛃 🛛 16:	27

Tips for Task 8

From the Analyze pull-down menu, select **Reports** and then choose **Case Summaries** (see fig 24). Highlight the "pe" variable from the list on the left and click the button with the right-hand arrow. Make sure the **Display cases** option is selected (it will have a check mark in the box next to the option), and in the **Limit cases to first**, enter 10 (see Fig 25). Click the **OK** button to run this procedure and get the final outcome as in Fig 26. If you want to copy and paste the outcome, you can able to do that when summary output window appears (see Fig 27).

You can also save your output file as in Fig 28.

誧 Wintergreen Study_1 - SPSS D	ata Editor	
File Edit View Data Transform	Analyze Graphs Utilities	Window Help
	Reports 🕨 🕨	OLAP Cubes
51 : ae	Descriptive Statistics Compare Means	Case Summaries
1	General Linear Model	Report Summaries in Rows
resp_numaa	Correlate	Report Summaries in Columns.
1 1 9	Regression •	1 2 0
2 2 4	Classify •	0 0 0
3 3 5	Data Reduction	1 1 0
4 4 9	Scale 🕨 🕨	2 2 1
5 5 8	Nonparametric Tests 🔸	2 1 1
6 6	Multiple Response 🔹 🕨	0 0 2

🐺 Summarize Cases	×	
 Respondent Number [rs Academic Abitity [aa] Student Motivation [sm] Advisor Evaluation [ra] Religious Affiliation [r] Gender [g] Community Type [c] recorded aa [aanew1] recorded aa [aanew2] acomp c = 1 (FILTER) [filter_\$] 	∩K <u>P</u> aste <u>H</u> eset Cancel Helµ	
Display cases		
Limit cases to <u>first</u> 10		
Show only valid cases		
Show case numbers Statistics Options		FIG

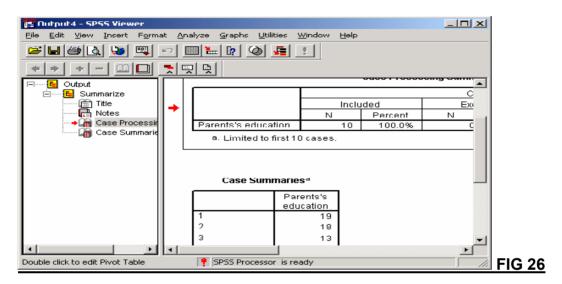


FIG 27

Save As				<u>?</u> ×	
Save jn: 🤂	Data04	- 🗲 🔁	📸 🎫 •		
File <u>n</u> ame:	Output1		Sav		
Save as <u>t</u> ype:	Viewer Files (*.spo)	_	Canc		FIG 28

• If you want to find out **first 15 students' religious affiliation**, you can apply the same procedure. The final outcome as follow: (see fig 29).

<u>FIG 29</u>