SPSS TUTORIAL



Step-by-Step Instructions for some basic SPSS operations by SEMPAXCONSULTING.COM

SPSS TUTORIAL

www.SempaxConsulting.com

Scatter Plot Regression ANOVA GLM Recoding Data

TABLE OF CONTENTS

1. Scatter Plot	4
Procedure	4
Assign Variable	4
Fit	5
Spikes	6
Titles	6
Options	7
Output	7
2. Linear Regression	8
Simple Linear Regression	
Output	10
Multiple Linear Regression	11
Output	12
3. ANOVA	14
One Way ANOVA	14
Output	15
4. General Linear Model (GLM)	17
GLM-Univariate	17
GLM Univariate-Fixed Factor(s)	17
Output	20
GLM Univariate-UNCOVA	23
Output	25
GLM Multivariate	27
Output	29
GLM Repeated Measures	31
Output	33
5. Recoding Data	35
FAED/MAED	35
Labeled	35
Recoding Data With Syntax	42
2 For any SPSS help request please e-mail me at	

1. Scatter Plot

A scatter plot may help you to understand how well linear regression fits your data. You may find that a quadratic equation would be more appropriate than a linear one.

Procedure

For example in this section we shall create a scatter plot for Employee Data.sav from SPSS data sample. Once the Employee Data.sav dataset is open, pull down the **Graphs** menu and point to **Interactive** and click on **Scatterplot** option.

Create Scatterplot dialog will appear. There are 5 tabs in Create Scatterplot dialog; assign variable, fit, spikes, title and option.

Assign Variable

On Assign variable you can select scatter plot coordinate between 2-D or 3-D, and then assign variable for each axis. If you select 2-D coordinate you must choose the variables you want on the X-axis and Y-axis, and if you select 3-D coordinate you must choose the variables you want on the X-axis and Y-axis. Drag and Drop variable name into axis field. This tutorial demonstrate sample for 2-D scatterplot, we have chosen 'previous experience' vs 'salary' from employee data. sav

Create Scatterplot
Assign Variables Fit Spikes Titles Options Case [\$case] Count [\$count] Percent [\$pet] Beginning Salay [sa Educational Level [the Employment Categor Gender [gender] Minority Classificatic Months since Hire[Legend Variables Color: Size: Panel Variables Label Cases By:
OK <u>P</u> aste <u>R</u> eset Cancel Help

Fit

Select fit method, there are 4 options; None, Regression, Mean and Smoother. Select None for this case

Create Scatterplot	
Assign Variables Fit	Spikes Titles Options
Method	
None	v
None Regression Mean Smoother	
Prediction Lines-	Individual Confidence Interval: 95.0
- Fit lines for	
<u>√</u> <u>T</u> otal	
<u>Subgroups</u>	
ОК	Paste Reset Cancel Help

Spikes

Use spikes options if you want to mark spikes data

Create Scatterplot	
Assign Variables Fit Spikes Titles Options	
Spike to: Drigin Corner Total Centroid Subgroup Centroid X1 Axis Y Axis Floor Color spikes by color legend Style spikes by style legend	
OK <u>P</u> aste <u>R</u> eset	Cancel Help

Titles

Fill on chart title, chart sub-title and caption, as you need

Create Scatterplot	\mathbf{X}
Assign Variables Fit Spikes Titles Options	
Chart <u>T</u> itle:	
Prev.Experiencs vs Salary	
Chart Sublitie:	
Caption:	
Note Because a graph's size is affected by the sizes of all of its parts, you should break long text into multiple lines by using the Enter key.	
OK <u>P</u> aste <u>R</u> eset Cancel Help	

Options

Select options as you need and then click **OK** to produce scatter plot diagram

Create Scatte	erplot		
Assign Variabl	es Fit Spikes Titles	Options	
Categorical	Drder	Ch	ectory:
Variable:	Gender	✓ C:	\Program Files\SPSS Evalu 🔽
Order by:	Values		Default> None> nalkboard assic
Of:	Beginning Salary	D De	ante
	Type: 🖉 Scale	Gr	rayscale
Sort:	ing O Descending	Ne St	eon eel
Exclude	empty categories		
- Scale Rang	e		Browse
Variable:	Count	× Ax	es
🔽 Auto			t∠, ⊻: 3.00 🗢 in.
Minimum:			t∡ X <u>1</u> : 3.00 🔷 in.
Maximum:			12. X2: 3.00 ♀ in.
ОК	Paste <u>R</u> e	eset	Cancel Help

Output

Scatter plot diagram will appear on output window



Prev.Experiences vs Salary

2. Linear Regression

This tutorial will explain two types of linear regression, there are simple linear regression and multiple linear regression.

Simple Linear Regression

Linear regression it is possible to output the regression coefficients necessary to predict one variable from the other. To do linear regression click on **Analyze => Regression => Linear**.

File Edit	View Data	Transform	Analyze Graphs Utilities Add-ons Window Help	
🗁 🔲 17 : Sal	🖹 📴 🦘 es_Num	۰ ե	Reports Descriptive Statistics Tables	
1 2 3	Sales Num 5 6 7	Revenue 12 14 16	Compare Means var var var General Linear Models • • • Mixed Models • • • Correlate • • •	
4 5 6	9	17 19 21	Regression Linear Loglinear Curve Estimation	
7	11	21 22 25	Classify Binary Logistic Data Reduction Multinomial Logistic	
9 10	13 14	27 29	Nonparametric Tests Probit	
11 12	15 16	30 31	Survival Nonlinear Multiple Response Weight Estimation Missing Value Analysis 2-Stage Least Squares	
13 14	17 18	33 35	Complex Samples Quality Control	
15 16	19 20	37 390	ROC Curve	

• Linear Regression Dialog will appear. Further, there is a need to know which variable will be used as the dependent variable and which will be used as the independent variable(s). In our current example, Revenue will be the dependent variable, and Sales Number will act as the independent variable.

Linear Regression		×
Sales Number [Sales_	Dependent: Thousand U\$ [Reven] Block 1 of 1 Previous <u>Next</u> [ndependent(s): Sales Number [Sales_Num] <u>M</u> ethod: Enter	OK Paste Aleset Cancel Help
	Selection Variable: Case Labels: WLS Weight:	
	Statistics Plots Save Optic	ons

• Click on **Statistics** button, and select **Estimates** and **Model fit** (as default)

Linear Regression: Sta	tistics	
Regression Coefficients	 ✓ Model fit R squared change Descriptives Part and partial correlations Collinearity diagnostics 	Continue Cancel Help
Durbin-Watson Darbin-Watson Casewise diagnostics Outliers outside: All cases	3 standard deviations	

- Click **Continue** button
- Click **Options** button and define confidence interval for F-test

Stepping Method Criteria Use probability of F Entry: Bemoval: .10 Use F value Entry: 3.84 Removal: 2.71	Continue Cancel Help
 ✓ Include constant in equation Missing Values ⊙ Exclude cases listwise ○ Exclude cases pairwise ○ Replace with mean 	

- Click Continue
- Click **OK** and output will appear

Output

Output for this case is:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.998(a)	.997	.996	5.106

a Predictors: (Constant), Sales Number

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	107974.94 4	1	107974.944	4140.871	.000(a)
	Residual	365.056	14	26.075		
	Total	108340.00 0	15			

a Predictors: (Constant), Sales Number

b Dependent Variable: Thousand U\$

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		В	Std. Error	Beta	В	Std. Error
1	(Constant)	34.243	3.690		9.281	.000
	Sales Number	17.821	.277	.998	64.350	.000

a Dependent Variable: Thousand U\$

Linear Regression Formula Model for this case is :

Y=34.243 + 17.821X

Multiple Linear Regression

- Click on **Analyze => Regression => Linear**.
- In this case we use **revenue** as **dependent variable**, **product price** and **sales number** as independent(s) variable.

Linear Regression		X
 ✓ Sales Number [Sales_ ✓ Product Price - U\$ [Pr 	Dependent: Previous Product Price - U\$ [Pric Sales Number [Sales_Nt] Method: Enter	OK <u>P</u> aste <u>R</u> eset Cancel Help
	Sglection Variable: Pase Labels: WLS Weight: Statistics Plots Save Opti	ons

• Click on **Statistics** button and select Estimates, Model Fit,Colineary diagnostics and Durbin-Watson

Linear Regression: Stat	istics	X
Regression Coefficients Estimates Confidence intervals Covariance matrix	<u>M</u> odel fit R squared change <u>D</u> escriptives <u>Part</u> and partial correlations <u>Collinearity diagnostics</u>	Continue Cancel Help
Residuals Durbin-Watson Casewise diagnostics Qutiers outside: All cases	3 standard deviations	

- Click **Continue** button
- Click **Options** button and define confidence interval for F-test



- Click Continue
- Click OK and output will appear

Output

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.993(a)	.986	.983	5.758	1.910

a Predictors: (Constant), Sales Number, Product Price - U\$

b Dependent Variable: Revenue - Thousands U\$

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27662.125	2	13831.063	417.148	.000(a)
	Residual	397.875	12	33.156		
	Total	28060.000	14			

a Predictors: (Constant), Sales Number, Product Price - U\$

b Dependent Variable: Revenue - Thousands U\$

Coefficients(a)

Ur Model	nstandardized Coefficients	Standardized Coefficients	t	Sig.	Collinearity Statistics
-------------	-------------------------------	------------------------------	---	------	----------------------------

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Colline Statis	earity stics
		В	Std. Error	Beta	Tolerance	VIF	В	Std. Error
1	(Constant)	60.444	12.710		4.756	.000		
	Product Price - U\$	396	.113	143	-3.491	.004	.703	1.42 2
	Sales Number	11.313	.436	1.064	25.945	.000	.703	1.42 2

a Dependent Variable: Revenue - Thousands U\$

Formula Model for this case is:

Y=30.444+11.313X(sales number) – 0.143X(product price)

3. ANOVA

analysis of variance (ANOVA) is a collection of statitical model and their associated procedures, in which the observed variance partitioned into components due to different explanatory variables.

Example case for this section is research about relationship between course period and grade. There are 3 kinds of course; 3 month, 6 month and 9 month.

One Way ANOVA

• Click on Analyze => Compare Means => One-Way ANOVA

File Edit	View Data	Transform	Analyze Graphs Utilities Add-ons Window Help
🗁 🔒	兽 🖭 🦘	• 🔶 🐜	Reports Descriptive Statistics Tables
18 : 1 2 3 4 5 6 7 8 9 10 11 12 13	Course 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Grade 7.0 8.0 7.2 7.5 7.0 7.5 7.7 6.5 7.0 7.5 8.0 8.5 8.0	Tables Means Compare Means One-Sample T Test General Linear Model One-Samples T Test Generalized Linear Models Independent-Samples T Test Mixed Models Paired-Samples T Test Mixed Models One-Way ANOVA Correlate One-Way ANOVA Regression Independent Samples T Test Loglinear Independent Samples T Test Classify Independent Samples T Test Data Reduction Independent Samples T Test Scale Independent Samples T Test Nonparametric Tests Independent Samples T Test Multiple Response Independent Samples Missing Value Analysis Independent Samples
14 15	2	8.2 7.5	Quality Control • ROC Curve •

• One-Way ANOVA dialog will appear, select **Grade** variable as dependent list and **Course** variable as factor

One-Way ANOVA		
	D <u>ependent List:</u> ▲ Grade	OK <u>P</u> aste <u>R</u> eset Cancel
	Factor:	Help
	Contrasts Post <u>H</u> oc	tions

• Click **Option** and select **Descriptive** and **Homogenety of variance test**

One-Way ANOVA: Options	
Statistics	Continue Cancel Help
 Means plot Missing Values Exclude cases <u>a</u>nalysis by analy 	sis
O Exclude cases listwise	

- Click Continue
- Click **Post Hoc** and select **LSD**
- Click Continue
- Click **Contrast** and enter **coefficients number** 0 and click **Add**, then enter coefficient number -1 and 1.
- Click Continue
- Click OK, and output will appear

Output

Descriptives

Grade

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
3 Months	10	7.3000	.43780	.13844	6.9868	7.6132	6.50	8.00

_								
6 Months	10	8.0750	.42573	.13463	7.7704	8.3796	7.50	8.75
9 Months	10	8.8750	.44488	.14068	8.5568	9.1932	8.00	9.50
Total	30	8.0833	.77774	.14200	7.7929	8.3737	6.50	9.50

ANOVA

Grade

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.404	2	6.202	32.595	.000
Within Groups	5.138	27	.190		
Total	17.542	29			

Test of Homogeneity of Variances

Grade

01000			
Levene Statistic	df1	df2	Sig.
.006	2	27	.994

Multiple Comparisons

Dependent Variable: Grade LSD

		Mean Difference				
(I) Term	(J) Term	(I-J)	Std. Error	Sig.	95% Confide	nce Interval
		Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound
3 Months	6 Months	77500(*)	.19508	.000	-1.1753	3747
	9 Months	-1.57500(*)	.19508	.000	-1.9753	-1.1747
6 Months	3 Months	.77500(*)	.19508	.000	.3747	1.1753
	9 Months	80000(*)	.19508	.000	-1.2003	3997
9 Months	3 Months	1.57500(*)	.19508	.000	1.1747	1.9753
	6 Months	.80000(*)	.19508	.000	.3997	1.2003

* The mean difference is significant at the .05 level.

Contrast Coefficients

Contrast	Term				
	3 Months	6 Months	9 Months		
1	0	-1	1		

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Grade	Assume equal variances	1	.8000	.19508	4.101	27	.000
	Does not assume equal variances	1	.8000	.19472	4.108	17.965	.001

4. General Linear Model (GLM)

This tutorial will explain four types of GLM, there are; GLM Univariate-Fixed Factor(s), GLM Univariate-UNCOVA, GLM-Multivariate and GLM-Repeates Measures.

GLM-Univariate

GLM-Univariate analysis is regression analysis and variance one dependent variable with two or more factor variable or other variables.

GLM Univariate-Fixed Factor(s)

Example case for univariate-fixed factor is to know customer shopping trend.

• Click on Analyze => General Linear Model => Univariate

File Edit	View Da	ta Transform	Analyze Graphs Utilities Add-ons Window Help
🖻 🔲	🔒 📴	승 🄶 ዀ	Reports 💦 🌾 🐼 💿
1 : gend	er		Tables
	gender	Cust_Cat fr	Compare Means
25	0	2	General Linear Model 🔸 Univariate
26	0	1	Generalized Linear Models 🕨 Multivariate
27	1	2	Mixed Models Repeated Measures
28		1	Correlate Variance Components
20	0		Regression
29	0	2	Loglinear
30	U	3	Classify •
31	1	2	Data Reduction 🕨
32	1	3	Scale 🕨
33	0	2	Nonparametric Tests
34	0	1	Time Series
35	1	1	Survival
36	0	2	Missing Value Apalysis
37	0	3	Complex Samples
38	0	2	Quality Control
39	0	2	ROC Curve
40	0	1	1 260476

• Uniavariate dialog will appear, select **shopping value** as dependent variable, **frequency** and **customer category** variable as fixed factor(s).

Univariate		
🚓 Gender [gender]	Dependent Variable:	<u>M</u> odel Co <u>n</u> trasts
	Customer Category	Plo <u>t</u> s Post <u>H</u> oc
	Random Factor(s):	<u>S</u> ave Options
	Covariate(s):	
	MLS Weight:	
ок [aste <u>R</u> eset Cancel Help]

• Click **Plots** and Univariate: Profile Plots dialog will appear, enter **frequency** variable into horizontal axis and **customer category** (**Cust_Cat**) into separate lines and then click **Add**. frequency*cust_cat variable will move into Plots box.

Univariate: Profile Plot	is		
Eactors: Cust_Cat frekun	$\overline{\mathbf{b}}$	Horizontal Axis: <u>Beparate Lines:</u> Separate Plots:	Continue Cancel Help
Plo <u>t</u> s: <u>A</u> dd		<u>C</u> hange <u>R</u> emove	
frekun*Cust_Cat			

- Click Continue
- Click **Post Hoc** and Univariate:Post Hoc dialog will appear, Select Equal Variances Assumed **Turkey** and Equal Variance Not Assumed **Tamhane**

Univariate: Post I	loc Multiple Comparisons for Observed Means 🛛 🔀
Eactor(s): Cust_Cat frekun	Post Hoc Tests for:
Equal Variances As LSD Bonferroni Sidak Scheffe B.E.G.W F R.E.G.W Q Equal Variances No V Tamhane's T2	sumed S-N-K Waller-Duncan Iukey Type I/Type II Error Ratio: 100 Tukey's-b Dunnett Duncan Control Category: Last Hochberg's GT2 Gabriel 2-sided < Control > Control t Assumed Dunnett's T3 Games-Howell Dunnett's C

- Click Continue
- Click **Option** and Univariate:Option dialog will appear, move frequency*Cat_Cus from Factor(s) and Factor Interactions box into Display Means for box. Select Descriptive statistic, Estimates of effect size, Homogenety test and spread vs level plot in Display groupbox.

Univariate: Options	X
Estimated Marginal Means Factor(s) and Factor Interactions: (OVERALL) Cust_Cat frekun Cust_Cat [*] frekun	Display <u>M</u> eans for:
	Compare main effects Confidence interval adjustment: LSD (none)
Display ✓ Descriptive statistics ✓ Estimates of effect size Observed power Parameter estimates	 <u>H</u>omogeneity tests Spread vs. level plot <u>R</u>esidual plot <u>L</u>ack of fit
Contrast coefficient matrix	<u>G</u> eneral estimable function lence intervals are 95% Continue Cancel Help

- Click Continue
- Click **OK** and output will appear

Output

Between-Subjects Factors

		Value Label	Ν
Customer Category	1	individu	337
0,	2	couple	287
	3	family	176
frequency	1	once-two weeks	187
	2	once-a week	461
	3	several - a week	152

Descriptive Statistics

Dependent Variable: Shopping Value

Customer Category	frequency	Mean	Std. Deviation	Ν
individu	once-two weeks	241549.03	51076.881	86
	once-a week	267907.85	47644.510	191
	several - a week	297827.14	46527.810	60
	Total	266508.14	51569.954	337
couple	once-two weeks	298406.07	50142.015	61
	once-a week	324952.65	47765.161	165
	several - a week	342457.78	48419.239	61
	Total	323030.95	50393.788	287
family	once-two weeks	384409.50	69602.060	40
	once-a week	400183.56	78776.554	105
	several - a week	421745.30	68918.993	31
	Total	400396.35	75637.588	176
Total	once-two weeks	290654.37	77743.025	187
	once-a week	318453.06	75860.179	461
	several - a week	341010.90	69289.815	152
	Total	316241.11	76812.882	800

Tests of Between-Subjects Effects

Dependent Variable:	Shopping Value					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	22909853751 27.443(a)	8	28637317189 0.930	93.477	.000	.486
Intercept	63790494928 399.500	1	63790494928 399.500	20822.226	.000	.963
Cust_Cat	15978262869 44.965	2	79891314347 2.483	260.778	.000	.397

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
frekun	16112594323 5.640	2	80562971617. 820	26.297	.000	.062
Cust_Cat * frekun	6899347068. 037	4	1724836767.0 09	.563	.690	.003
Error	24232895140 93.421	791	3063577135.3 90			
Total	84721024114 141.800	800				
Corrected Total	47142748892 20.860	799				

a R Squared = .486 (Adjusted R Squared = .481)



Estimated Marginal Means of Shopping Value

Customer Category * frequency

Dependent Variable: Shopping Value

Customer Category	frequency	Mean	Std. Error	95% Confide	ence Interval
			Upper		
		Lower Bound	Bound	Lower Bound	Upper Bound
individu	once-two weeks	241549.033	5968.500	229833.061	253265.004
	once-a week	267907.849	4004.956	260046.250	275769.447

Customer Category	frequency	Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound	Lower Bound	Upper Bound
	several - a week	297827.138	7145.601	283800.554	311853.721
couple	once-two weeks	298406.073	7086.789	284494.936	312317.210
	once-a week	324952.647	4308.960	316494.299	333410.996
	several - a week	342457.775	7086.789	328546.639	356368.912
family	once-two weeks	384409.496	8751.539	367230.510	401588.483
	once-a week	400183.564	5401.567	389580.464	410786.665
	several - a week	421745.298	9941.080	402231.281	441259.316

Multiple Comparisons

Dependent Variable: Shopping Value Tamhane

		Mean Difference				
 (I) frequency 	(J) frequency	(I-J)	Std. Error	Sig.	95% Confide	ence Interval
		Lower	Upper	Lower		
		Bound	Bound	Bound	Upper Bound	Lower Bound
once-two weeks	once-a week	-27798.69(*)	6693.576	.000	-43861.00	-11736.37
	several - a week	-50356.53(*)	7994.172	.000	-69540.69	-31172.37
once-a week	once-two weeks	27798.69(*)	6693.576	.000	11736.37	43861.00
	several - a week	-22557.84(*)	6638.469	.002	-38504.28	-6611.40
several - a week	once-two weeks	50356.53(*)	7994.172	.000	31172.37	69540.69
	once-a week	22557.84(*)	6638.469	.002	6611.40	38504.28

Based on observed means.

* The mean difference is significant at the .05 level.

GLM Univariate-UNCOVA

Example case for this section is research about house hold income before and after participate in government program.

- Click on Analyze => General Linear Model => Univariate
- Enter result_after as dependent variable, program status variable as Fix Factor(s) and result_before as Covariate(s)

Univariate			X
	\rightarrow	Dependent Variable: Øresult after program [re <u>Eixed Factor(s):</u>	Model
	\rightarrow	💑 program status (progra	Plo <u>t</u> s Post <u>H</u> oc
	_	R <u>a</u> ndom Factor(s):	<u>S</u> ave
	\mathbf{r}		Options
		<u>C</u> ovariate(s):	
	◀	result before program [
	\rightarrow	<u>W</u> LS Weight:	
OK E	aste	<u>R</u> eset Cancel Help	

• Click **Models** and Univariate: Models dialog will appear, Select **Custom** in Specify Model. Select **program** variable move into Model box, select **result_before** and move into Model box. Select both program variable and result_before and move into Model box and then **program*result_before** variable will appear. Select **Interaction** on the Build Term(s) dropdown.

Univariate: Model Specify Model	⊙ <u>C</u> ustom	
Eactors & Covariates: program(F) result_before(C)	Build Term(s) Interaction	
Sum of sguares: Ty	ce III 💌 🔽 Include intercept in mode	el Cancel Help

- Click Continue
- Click **Options** and select Estimates of effect size.

Univariate: Options	
Estimated Marginal Means Eactor(s) and Factor Interactions: (OVERALL) program	Display Means for:
Display Descriptive statistics Descriptive statistics Descriptive statistics Descriptive statistics Descriptive statistics Descriptive power Parameter estimates Contrast coefficient matrix Significance level: .05 C	<u>Homogeneity tests</u> Spread vs. level plot <u>R</u> esidual plot <u>L</u> ack of fit <u>G</u> eneral estimable function
	Continue Cancel Help

- Click Continue
- Click **OK** and output will appear.
- The next step is covarian analysis. Open Univariate dialog again click **Model** and select **Full Factorial** in Specify model.

Univariate: Model		×
Specify Model Full factorial	Custom	
Eactors & Covariates: program(F) result_before(C)	Build Term(s) program result_before program*result_before Interaction Interaction	
Sum of sguares:	Type III	Help

- Click Continue
- Click **Option** and select Descriptive statistics, Estimates of effect size, Homogenety test, and Parameter Estimates.

Univariate: Options	X
Estimated Marginal Means Eactor(s) and Factor Interactions: (OVERALL) program	Display Means for:
Display ✓ Descriptive statistics ✓ Estimates of effect size □ Observed power ✓ Parameter estimates □ Contrast coefficient matrix	Homogeneity tests Spread vs. level plot <u>R</u> esidual plot Lack of fit <u>G</u> eneral estimable function
Significance le <u>v</u> el: .05 C	onfidence intervals are 95% Continue Cancel Help

- Click Continue
- Click **OK** and output will appear.

Output

Between-Subjects Factors

		Value Label	Ν
program status	0	not participate	293
	1	participate	307

Tests of Between-Subjects Effects

Dependent Variable: result after program

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	16557688.72 9(a)	3	5519229.576	220.962	.000	.527
Intercept	291519.199	1	291519.199	11.671	.001	.019
program	114931.433	1	114931.433	4.601	.032	.008
result_before	9598160.112	1	9598160.112	384.262	.000	.392
program * result_before	11836.725	1	11836.725	.474	.491	.001
Error	14886973.77 1	596	24978.144			
Total	453042500.0 00	600				
Corrected Total	31444662.50 0	599				

a R Squared = .527 (Adjusted R Squared = .524)

Descriptive Statistics

Dependent Variable: result after program

program status	Mean	Std. Deviation	Ν
not participate	728.84	194.917	293
participate	942.67	210.011	307
Total	838.25	229.118	600

Levene's Test of Equality of Error Variances(a)

Dependent Variable: result after program

F	df1	df2	Sig.	
.605	1	598	.437	

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a Design: Intercept+result_before+program

Tests of Between-Subjects Effects

Dependent Variable: result after program

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	16545852.00 4(a)	2	8272926.002	331.499	.000	.526
Intercept	285238.337	1	285238.337	11.430	.001	.019
result_before	9691004.737	1	9691004.737	388.322	.000	.394
program	6459841.024	1	6459841.024	258.848	.000	.302
Error	14898810.49 6	597	24956.131			
Total	453042500.0 00	600				
Corrected Total	31444662.50 0	599				

a R Squared = .526 (Adjusted R Squared = .525)

Parameter Estimates

Dependent Variable: result after program

							Partial Eta
Parameter	В	Std. Error	t	Sig.	95% Confide	95% Confidence Interval	
	Lower	Upper	Lower	Upper		Upper	
	Bound	Bound	Bound	Bound	Lower Bound	Bound	Lower Bound
Intercept	227.856	37.378	6.096	.000	154.448	301.264	.059
result_before	1.587	.081	19.706	.000	1.429	1.745	.394
[program=0]	-207.641	12.906	-16.089	.000	-232.987	-182.294	.302
[program=1]	0(a)				-	•	

a This parameter is set to zero because it is redundant.

GLM Multivariate

Example case for this section is research about the impact of gender factor to expense for life style.

• Click on Analyze => General Linear Model => Multivariate

File Edit	View Data	Transform	Analyze Graphs Utilities Add-ons Window Help
🗁	🔒 📴 🦘	+ 🐱	Reports Descriptive Statistics
1 : gend	er		Tables 🕨
	gender	food	Compare Means Var Var Var Var
1	1	40	General Linear Model Univariate
2	0	55	Generalized Linear Models Multivariate
3	0	40	Mixed Models Repeated Measures
4	1	50	Variance Components
5	1	65	
6	0	35	Classify •
7	0	45	Data Reduction 🕨
8	0	50	Scale 🕨
9	1	35	Nonparametric Tests
10	1	50	Time Series
11	1	40	Survival
12	1	45	Milliple Response
13	0	40	Complex Samples
14	0	45	Quality Control
15	0	35	ROC Curve
16	0	50	0 450

- Multivariate dialog box will appear, select **food cost** variable and **lifestyle** variable and move into Dependent Variable
- Move Gender variable into Fix Factor(s)

Multivariate		X
	Dependent Variables:	Model Contrasts Plots
	Eixed Factor(s):	Post <u>H</u> oc Save
	Covariate(s):	
	LS Weight:	
ок [aste <u>R</u> eset Cancel Help	

• Click **option** and then select Descriptive statistics, Estimates of effect size, and Parameter estimates

Multivariate: Options	
Estimated Marginal Means Eactor(s) and Factor Interactions: (OVERALL) gender	Display Means for:
Display ✓ Descriptive statistics ✓ Estimates of effect size Observed power ✓ Parameter estimates SSCP matrices Residual SSCP matrix	 Transformation matrix Homogeneity tests Spread vs. level plots Besidual plots Lack of fit test General estimable function
Significance le <u>v</u> el: .05 Co	onfidence intervals are 95% Continue Cancel Help

- Click Continue
- Click **OK** and output will appear

Output

Between-Subjects Factors

		Value Label	Ν
gender	0	male	189
	1	female	211

Descriptive Statistics

	gender	Mean	Std. Deviation	Ν
food cost	male	445.24	82.375	189
	female	451.18	79.947	211
	Total	448.38	81.056	400
lifestyle cost	male	723.54	195.055	189
	female	945.26	209.481	211
	Total	840.50	230.880	400

Multivariate Tests(b)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.969	6237.900(a)	2.000	397.000	.000	.969
	Wilks' Lambda	.031	6237.900(a)	2.000	397.000	.000	.969
	Hotelling's Trace	31.425	6237.900(a)	2.000	397.000	.000	.969
	Roy's Largest Root	31.425	6237.900(a)	2.000	397.000	.000	.969
gender	Pillai's Trace	.308	88.173(a)	2.000	397.000	.000	.308
	Wilks' Lambda	.692	88.173(a)	2.000	397.000	.000	.308
	Hotelling's Trace	.444	88.173(a)	2.000	397.000	.000	.308
	Roy's Largest Root	.444	88.173(a)	2.000	397.000	.000	.308

a Exact statistic

b Design: Intercept+gender

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	food cost	3525.673(a)	1	3525.673	.536	.465	.001
	lifestyle cost	4900914.469(b)	1	4900914.469	119.169	.000	.230
Intercept	food cost	80114325.673	1	80114325.673	12179.717	.000	.968
	lifestyle cost	277648789.469	1	277648789.46 9	6751.241	.000	.944
gender	food cost	3525.673	1	3525.673	.536	.465	.001
	lifestyle cost	4900914.469	1	4900914.469	119.169	.000	.230
Error	food cost	2617918.077	398	6577.684			

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
	lifestyle cost	16367985.531	398	41125.592			
Total	food cost	83037500.000	400				
	lifestyle cost	303845000.000	400				
Corrected Total	food cost	2621443.750	399				
	lifestyle cost	21268900.000	399				

a R Squared = .001 (Adjusted R Squared = -.001) b R Squared = .230 (Adjusted R Squared = .228)

Parameter Estimates

								Partial
						95% Cor	nfidence	Eta
Dependent Variable	Parameter	В	Std. Error	t	Sig.	Inter	val	Squared
		Lower	Upper	Lower	Upper	Lower	Upper	Lower
		Bound	Bound	Bound	Bound	Bound	Bound	Bound
food cost	Intercept	451.185	5.583	80.809	.000	440.208	462.161	.943
	[gender=0]	-5.947	8.123	732	.465	-21.915	10.022	.001
	[gender=1]	0(a)						
lifestyle cost	Intercept	945.261	13.961	67.707	.000	917.814	972.707	.920
	[gender=0]	-221.716	20.310	-10.916	.000	-261.644	- 181.787	.230
	[gender=1]	0(a)	-					-

a This parameter is set to zero because it is redundant.

GLM Repeated Measures

Example case for this section is research about performance of 4 weeks diet program between male and female.

- File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help Reports ۲ ¥ 🗁 🔲 🚔 -- 🙆 🌒 ШŤ. - **to** - **cb** -• ۲ Descriptive Statistics 1 : Gender Tables ٠ Compare Means ۲ Weinht3 | Weinht4 | Gender Weight0 var General Linear Model Univariate... 90.0 1 1 Generalized Linear Models ۲ Multivariate... 85.7 2 1 Repeated Measures... Mixed Models ۲ 3 79.0 1 Correlate ۲ Variance Components... 4 1 80.0 Regression ۲ 5 1 83.5 75.50 72.00 Loglinear ۲ 6 1 87.0 78.50 75.00 Classify ۲ 7 1 89.0 Data Reduction 80.75 77.00 ۲ 8 1 86.0 Scale ۲ 77.00 74.50 Nonparametric Tests ۲ 9 1 83.0 75.00 72.50 Time Series ۲ 10 1 81.5 73.00 72.00 Survival ۲ 11 2 81.2 75.00 73.00 Multiple Response ۲ 2 12 75.0 68.00 66.00 Missing Value Analysis... 2 13 71.0 65.50 62.00 Complex Samples ۲ 2 14 77.5 71.50 69.00 Quality Control ۲ 15 2 78.0 ROC Curve... 72.50 71.00 2 16 76.50 68.00 75.00 72.75 70.00 2 17 75.00 72.00 66.00 69.00 69.00
- Click on Analyze => General Linear Model => Repeated Measures

• Repeated Measures Define dialog will appear, write **weight** in the Within-Subject Factor Name and enter **5** on Number of Levels. Click **Add** and weight5 will move into box

Within-Subject Factor Name:	
Number of Levels:	De <u>f</u> ine <u>R</u> eset Cancel
Add weight(5) Change Remove	Help
Add Change Bemove	

• Click **Define** and Repeated Measures dialog will appear. Enter dependent variable from **weight0**, **weight1**, **weight2**, **weight3** and **weight4** in Within-Subjects Variables (weight) and **gender** variable in Between-Subjects Factor(s) box.

Repeated Measures	×
Within-Subjects Variables (weight):	OK Paste
Weight0(1) Weight2(2) Weight2(3) Weight3(4) Weight4(5)	<u>R</u> eset Cancel Help
Between-Subjects Factor(s):	
Model Contrasts Plois Post Hoc Save Options	

- Click **Option**, select Descriptive statistics, Estimates of effect size, and Parameter estimates
- Click Continue
- Click **OK**, and Output will appears

Output

Within-Subjects Factors

Measure: MEASURE_1

weight	Dependent Variable
1	Weight0
2	Weight1
3	Weight2
4	Weight3
5	Weight4

Between-Subjects Factors

		Value Label	Ν
Gender	1	male	10
	2	female	10

Descriptive Statistics

	Gender	Mean	Std. Deviation	Ν
Weight before program	male	84.4750	3.70144	10
	female	75.5000	3.12027	10
	Total	79.9875	5.68324	20
Weight Weeks1	male	82.0500	3.37021	10
	female	73.5000	3.30824	10
	Total	77.7750	5.45912	20
Weight Weeks2	male	78.9250	3.21898	10
	female	71.4250	3.26609	10
	Total	75.1750	4.97633	20
Weight Weeks3	male	77.0250	4.77617	10
	female	70.4000	4.03320	10
	Total	73.7125	5.48279	20
Weight Weeks4	male	74.5000	4.99444	10
	female	68.1250	4.21843	10
	Total	71.3125	5.56237	20

Multivariate Tests(b)

Effect		Value	F	Hypothesi s df	Error df	Sig.	Partial Eta Squared
weight	Pillai's Trace	.981	193.405(a)	4.000	15.000	.000	.981

	Wilks' Lambda	.019	193.405(a)	4.000	15.000	.000	.981
	Hotelling's Trace	51.575	193.405(a)	4.000	15.000	.000	.981
	Roy's Largest Root	51.575	193.405(a)	4.000	15.000	.000	.981
weight * Gender	Pillai's Trace	.569	4.960(a)	4.000	15.000	.009	.569
	Wilks' Lambda	.431	4.960(a)	4.000	15.000	.009	.569
	Hotelling's Trace	1.323	4.960(a)	4.000	15.000	.009	.569
	Roy's Largest Root	1.323	4.960(a)	4.000	15.000	.009	.569

a Exact statistic

b Design: Intercept+Gender Within Subjects Design: weight

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
weight	Sphericity Assumed	922.129	4	230.532	73.811	.000	.804
	Greenhouse-Geisser	922.129	1.118	824.571	73.811	.000	.804
	Huynh-Feldt	922.129	1.206	764.356	73.811	.000	.804
	Lower-bound	922.129	1.000	922.129	73.811	.000	.804
weight * Gender	Sphericity Assumed	26.271	4	6.568	2.103	.089	.105
	Greenhouse-Geisser	26.271	1.118	23.492	2.103	.162	.105
	Huynh-Feldt	26.271	1.206	21.776	2.103	.159	.105
	Lower-bound	26.271	1.000	26.271	2.103	.164	.105
Error(weight)	Sphericity Assumed	224.875	72	3.123			
	Greenhouse-Geisser	224.875	20.130	11.171			
	Huynh-Feldt	224.875	21.715	10.356			
	Lower-bound	224.875	18.000	12.493			

Tests of Between-Subjects Effects

Measure: MEASURE_1 Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	571422.606	1	571422.606	9246.269	.000	.998
Gender	1445.901	1	1445.901	23.396	.000	.565
Error	1112.406	18	61.800			

5. Recoding Data

You can recode data into either the same variable or into a new one by going to **Transform > Recode.** This tool is especially useful for creating dummy variables, changing values from letters to numbers, increasing or decreasing the number of possible values, or for creating specialized filters that let you have fine-tuned control over which cases to exclude.

SPSS allows us to recode variables and then use the recoded variables in statistical analyses.

The values in variables **FAED** (father's education) and **MAED** (mother's education) range from 2 to 10 indicating 9 levels of education as:

	Labeled
FAED/MAED	
2	Less than high school
3	High school graduate
4	Less than 2 years' vocational education
5	More than 2 years' vocational education
6	Less than 2 years' college education
7	More than 2 years' college education
8	College graduate
9	Master's degree
10	MD/PhD degree

Now we want to regroup (recode) the nine levels into four levels as:

FAED/MAED	FAEDNEW / MAEDNEW	Labeled
2	1	Less than high school
3	2	High school graduate
4,5,6,7	3	Some post-secondary education
8,9,10	4	College graduate & beyond

To recode the variables, please follow the steps:

• You will see the data in the SPSS Data Editor window:

📰 Unti	tled - SF	PSS Da	ta Editor					- 🗆 ×		
<u>F</u> ile <u>E</u> o	lit ⊻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>	indow <u>H</u> elp				
B B B C C C C C C C C C C										
6: alg 1										
	i	d	visual	mosaic	grades	mathgr	alg	alg2 📤		
	1	1	8.750	31.000	4	0	0			
	2	2	1.000	24.000	8	1	1			
	3	3	6.000	5.000	7	1	1			
	4	4	3.500	23.000	6	0	0			
	5	5	12.250	22.000	7	1	1			
	6	6	11.000	35.500	8	1	1			
	7	7	3.500	26.000	6	1	0			
	8	8	9.750	30.500	4	0	0			
	9	9	4.750	56.000	5	0	0			
1	0	10	4.750	25.000	6	1	0			
1	1	11	1.000	22.000	3	0	1			
1	2	12	2.250	17.500	3	0	0			
1	3	13	7.250	28.500	5	0	1	-		
	Data Vie	N K Va	riable View 🖊					▶		
				SPSS Proce	ssor is ready					

- Before you recode the data, you should make a copy of original data. Make sure you save the new file into the same place as the original file.
- Recode the **FAED** (father's education) variable into a new variable
- From **Transform** menu, choose **Recode**, then **Into Different Variable**.

🛗 Untitled - SPSS Data 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> raphs	<u>U</u> tilities	<u>W</u> indo	w <u>H</u> elp			
🖻 🖬	a 🔍 🖻	<u>C</u> ompute Random N Count	Iumber <u>S</u> eed	Ť		1	0		
pr. raca		Becode			nto Same	Variables		In	
	alg	Categorize	Variables		nto Differ	ent Variable:	s ach	faed	
1	0	Rank Cas	es				9.000 [10	
2] 1	 <u>A</u> utomatic	Recode		1	0	22.333	7	
3	1	Create Ti <u>n</u>	je Series		0	0	9.000	4	
4	0	Replace M	fissing <u>V</u> alues…		0	0	6.333	3	
5	1	Bun Pend	ing Transforms		1	0	22.333	8	
6	1		Ing Transiening		1	1	23.667	5	
7	0	0	0		0	0	7.667	2	
8	0	0	0		0	0	8.000	7	
9	0	0	0		0	0	10.333	7	
10	0	0	0		0	0	7.667	2	
11	1	0	0		0	0	5.000	3	
12	0	0	0		0	0	-1.667	3	
13	1	0	0		0	0	18.333	2	-
▲ ▶ \ Da	ta View 🖌 🗸	ariable View 🏼 /		•				•	
Recode In	to Same Variab	les	SPSS Proces	sor is rea	dy				

• In the "Recode into Different Variable" dialogue box, you will see a list of variables in the box on the left. Click on "**faed**", and then click on the arrow button. You will see "**faed**" appears in the right box.

🚮 Recode into Different Variables	X	<mark>%</mark> Recode into Di	fferent Variables		X
Input ⊻ariable → Output Variable Imput ½ariable → Output ½ariable Impu	Cancel Help	 id visual mosaic grades mathgr alg alg2 geo lrig calc mathach maed 	Numeri action of the second o	Ind New Values	Output Variable Name: [raednew Change] Labet Father's education

- Type "**faednew**" in the Output/Variable-Name box as the name of the new variable. Type "Father's education" as the label for the new variable. Click on the **Change** button (see above).
- Click on the button "Old and New Values", you will see the Old and New Values dialogue box. Under Old Value section, type 2 in the Value box, and type 1 into the Value box under New Value section this will recode the old value 2 into a new value 1 (as shown in the tables at the beginning of this module). Click on Add button.

Recode into Same Variables: Old and New Values	Recode into Same Variables	Recode into Same Variables: Old and New Values		
Old Value New Value © Value: 2 © System-missing © Value: © System-or user-missing Add © Range: Ehenge Lowest through Eemove © Range: Hrough kighest	Old Value Yalue: Yalue: System-missing System-or user-missing System-or user-missing Range: Hrough Range: Lowest through Range: Hrough high	New Value Value System-missing Old -> New: Add Old -> New: Add Change Bemove		
C All other values Continue Cancel	Help C All other values	Continue Cancel Help		

- Type 3 in Old value box, and 2 in New value box. Click on Add button. The recoding shows in the Old → New box.
- Check the **Range** radio button, then type 4 in the first box, and type 7 in the box after the word "through". Then type 3 in the New value box. Click on **Add** button. You will see:

Recode into Same Variables: Old and New Values						
Old Value Value: System-missing System- or user-missing Range: 8 through Lowest through Range: Lowest through	New Value Image: Value: 4 Old> New Add 2> 1 3> 2 Change Hemove	C System-missing r				
C All <u>o</u> ther values	Continue Cancel	Help				

• Type 8 and 10 in the range boxes, and 4 in the New value box. Click on **Add**, you have recoded the nine old values into four new values:

Recode into Same Variables: Old and Ne	w Values	×
Old Value Value: System-missing System-or user-missing Range: through Range: Lowest through Range: through Hange:	New Value Image: New Value: Image: New Value: Add Add Image: Strength of the strengt of the strength of the strength of the strengeh of the streng	© System-missing New: 7 → 3 10 → 4
C All <u>o</u> ther values	Continue Can	cel Help

- Click on Continue button, you will be back to the Recode into Different Variable dialogue box. Now you will recode another variable **maed**—mother's education.
- Recode the MAED (mother's education) variable
- In the Recode into Different Variable box, from the variable list, click on **maed**, and click on the arrow button to add the variable **maed** into the right box, it should be under **faed** variable.

🙀 Recode into Different Variables 🛛 🔀	Recode into Different Variables			
 Numeric Variable -> Output Variable Output Variable Output Variable Output Variable Itaed -> faednew Itaed -> faednew Itaet Itaet	 id visual mosaic grades mad →> maednew Labet Mother's education if if	0		

- Make sure the variable **maed** is highlighted, type **maednew** in the Output/Variable-Name box as the name of the new variable. Type "Mother's education" as the label for the new variable. Click on the **Change** button (see above).
- Click on **Old and New Values** button, you will see the previous recode settings:

Recode into Same Variables: Old and N	lew Values
Old Value ⊙ ⊻alue: ○ System-missing	New Value O Value: O System-missing
C System- or user-missing C Range: through C Range:	Add 2 -> 1 2 -> 1 3 -> 2 Change 4 thru 7 -> 3 8 thru 10 -> 4
Lowest through C Range: through highest C All other values	Continue Cancel Help

- We will use the same recode settings. So we do not need to change. Simply click on Continue. (If you need to change the settings, click on each of the recode settings, then click on Remove. You can add new transform settings).
- Now, you are back to the original dialogue box, click on **OK**. You will see the two new variables **faednew** and **maednew**.

	e into Diffe	erent Varia	bles					
 id visual mosai grade: grade: mathg alg 	c c		Numeric <u>V</u> faed> fa maed> r	ariable -> 0 ednew naednew	utput Variat	Die: Outpu Name: Maedr Label: Mothe	it Variable new r's educat	<u>C</u> hange ion
alg2								
aeo			IF 1					
a) tria			<u></u>					
ny uig Nasla					- 1			
€ calc			<u>O</u> ld and I	New Values	\$			
matha	ich	_						
🖹 sex 🗌			OK	Paste	Reset	Cancel	Help	
× i		T						
hsb50(File <u>E</u> dit	J R-b - SPSS I ⊻iew Data	Data Editor Iransform Ar	ialyze <u>G</u> raphs	Utilities Wi	ndow <u>H</u> elp			
immed in hsb500 File Edit	D R-b - SPSS View Data	Data Editor Iransform Ar	alyze <u>G</u> raphs	Ltilities Wi	ndow Help	0		
im hsb500 File Edit	DR-b - SPSS View Data	Data Editor Iransform Ar C I I I I I I I I I I I I I I I I I I I	alyze <u>G</u> raphs		ndow Help		USP	
hsb500 File Edit	D R-b - SPSS I View Data Data V M mathach	Data Editor Iransform Ar 2 faed	ialyze <u>G</u> raphs () () () () () () () ()	Utilities Wi	ndow Help	maednew 4.00	var	var _
 immul i i i i i i i i i i i i i i i i i i i	DR-b - SPSS ⊻iew Data	Data Editor Iransform Ar 2 faed 10 7	alyze <u>G</u> raphs maed 10 3	Utilities Wi	ndow Help	maednew 4.00 2.00	var	var _
hsb500 File Edit 4 : faednev 1 2 3	JR-b - SPSS I ⊻iew Data	Data Editor Iransform Ar 2 faed 10 7 4	alyze Graphs maed 10 3 2	Luilities Wi	ndow <u>Help</u>	maednew 4.00 2.00 1.00	var	var
hsb50(File Edit 2 4 : faednev 1 2 3 4	DR-b - SPSS I ⊻iew Data mathach 9.000 22.333 9.000 6.333	Data Editor Iransform Ar 2 faed 10 7 4 3	alyze <u>G</u> raphs maed 10 3 2 2	Luilities Wi	ndow Help	maednew 4.00 2.00 1.00 1.00	Var	var _
hsb500 File Edit	DR-b - SPSS View Data	Data Editor Iransform Ar 2 faed 10 7 4 3 8	maed 10 3 2 7		ndow Help faednew 4.00 3.00 2.00 4.00	maednew 4.00 2.00 1.00 1.00 3.00	var	var
hsb500 File Edit 4 : faednev 1 2 3 4 5 6	R-b - SPSS View Data	Data Editor Iransform Ar 2 faed 10 7 4 3 8 5	alyze <u>G</u> raphs maed 10 3 2 2 7 7	Utilities Wr Sex 0 0 0 1 1	ndow Help faednew 4.00 3.00 2.00 4.00 3.00 2.00	maednew 4.00 2.00 1.00 3.00 3.00 3.00	var	var -
hsb500 File Edit 2 4 : faednev 1 2 3 4 4 5 6 7	Image: Non-Series Series Series	Data Editor Tarastom År 2 faed 10 7 4 3 8 5 5 2	alyze <u>G</u> raphs maed 10 3 2 7 7 2	Litities Wi Sex 0 0 0 0 0 0 0 0 0 0 0 0 0	ndow Help faednew 4.00 3.00 2.00 4.00 3.00 1.00	maednew 4.00 2.00 1.00 3.00 3.00 3.00 1.00	var	var
hsb500 File Edit 2 4 : faednev 4 : faednev 3 4 5 6 7 8	JR-b - SPSS ¥ew Data ₩ mathach 9,000 22,333 9,000 6,333 22,667 7,667 8,000	Data Editor Iransform Ar 2 faed 10 7 4 3 8 5 2 2 7	walyze <u>G</u> raphs maed 10 3 2 2 7 7 2 2 2	Litities Wi Sex 0 0 0 0 0 0 0 0 0 0 0 0 0 0	adow <u>Help</u>	maednew 4.00 2.00 1.00 1.00 3.00 3.00 1.00 1.00	var	var
hsb50(Eile Edit 2 4 : faednew 4 : faednew 4 : faednew 5 5 6 7 7 8 9	JR-b - SPSS Yew Data Yew Data mathach 9.000 22.333 9.000 6.333 22.333 23.667 7.667 8.000 10.333	Data Editor Irarstorm Ar 2 faed 10 7 4 4 3 8 5 2 7 7 7	maed 10 3 2 7 7 7 2 2 2 2	Litities ₩i Sex 0	taednew faednew 4.00 3.00 2.00 4.00 3.00 1.00 3.00 3.00 3.00	maednew 4.00 2.00 1.00 3.00 3.00 1.00 1.00 1.00 1.00	Var	var
hsb50t Eile Edt Image: second	DB-b - SPSS ¥ew Data mathach 9,000 22,333 9,000 6,333 22,333 23,667 7,667 8,000 10,333 7,667	Lanstom Ar Iranstom Ar P Faed 10 10 10 10 10 10 10 10 10 10	walyze @raphs = 12 #4 10 3 2 7 7 7 2 2 2 2 2 2 2 2 2	Litities Wr Sex 0 0 0 0 0 0 0 0 0 0 0 0 0	ndow Help faednew 4.00 3.00 2.00 4.00 3.00 3.00 3.00 3.00 1.00 3.00 1.00	maednew 4.00 2.00 1.00 3.00 3.00 1.00 1.00 1.00 1.00	Var	Var
hsb50t Eile Edit Image: Constraint of the second se	BR-b - SPSS View Data	Data Editor Ternstom År 2 faed 10 7 4 4 3 8 5 2 7 7 7 2 2 3	xalyze Graphs T D2 040 1 maed 10 3 2 2 7 7 2 2 2 2 3	Ltitries Wr Sex 0 0 0 0 0 0 0 0 0 0 0 0 0	radow Help faednew 4.00 3.00 2.00 4.00 3.00 1.00 3.00 1.00 3.00 2.00	maednew 4.00 2.00 1.00 3.00 1.00 1.00 1.00 1.00 0.00 2.00	var	Var
■ hsb50 File Edit ■ 1 4 : faednew 4 : faednew 4 : faednew 5 6 7 8 9 10 11 12 2	PR-b - SPSS View Data → 1 → 1 → 1 → 1 → 1 → 1 → 1 → 1	Data Editor Iranstom År P P P P faed 10 7 4 3 8 5 2 7 7 7 7 7 7 2 3 3 3	xalyze Graphs	Litities W/ Sex 0 0 0 0 0 0 0 0 0 0 0 0 0 0	ndow Help faednew 4.00 3.00 2.00 4.00 3.00 3.00 1.00 3.00 1.00 2.00	maednew 4.00 2.00 1.00 3.00 3.00 3.00 1.00 1.00 1.00 1	var	var
■ hsb50(File Edit ■ 1 4 : faednew 4 : faednew 4 : faednew 4 : faednew 9 10 11 12 13 13	Br.b SPSS View Data Imathach 9.000 22.333 9.000 6.333 23.667 7.667 7.667 10.333 7.667 5.000 -1.667 1.8333 -1.667	Data Editor Tarastom ér 2 faed 10 7 4 3 8 5 2 2 7 7 7 7 7 7 2 3 3 3 2 2	Image Graphs maed 10 3 2 2 2 7 2 2 2 2 3 3 3	Lutities W/ E E E E E E E E E E E E E E E E E E E	ndow Help faednew 4.00 3.00 2.00 4.00 3.00 1.00 3.00 1.00 2.00 1.00 2.00 1.00 2.00 1.00 2.00 1.00	maednew 4.00 2.00 1.00 1.00 1.00 1.00 1.00 1.00 2.00 2	var	var 2

• For the value of these two variables, we do not need decimals. To change the decimals, look at the bottom of the Data Editor window, you can see two tabs (Data View – which is the current window, and **Variable View**). Click on the **Variable View** tab.

I	10	U	U	7.667			
I	11	0	0	5.000	2	2	
I	12	0	0	-1.667	2	2	
I	13	0	0	18.333	1	2	
l	▲ ▶ \ Da	ta View 🖉 🗸	riable View 🆯		•		
ļ				SPSS Proces	ssor is ready		

• You will see the window changes to the Variable View mode:

💼 h	sb50	OR-6 - SPS	6 Data Editor					- 🗆 ×
<u>F</u> ile	<u>E</u> dit	<u>V</u> iew <u>D</u> ata	a <u>T</u> ransform <u>A</u> n	alyze <u>G</u> raphs <u>L</u>	<u>I</u> tilities <u>W</u> indow <u>H</u> e	elp		
2		a 💷 🖻	• ~ 🗐 🗄	- 1 M H		<u>v</u>		
		Name	Туре	Width	Decimals	Label	Values	<u>∧</u> ≜
	4	grades	Numeric	2	0		None	None
	- 5	mathgr	Numeric	2	0		None	None
	6	alg	Numeric	2	0		None	None
	- 7	alg2	Numeric	2	0		None	None
	8	geo	Numeric	2	0		None	None
	9	trig	Numeric	2	0		None	None
	10	calc	Numeric	2	0		None	None
	11	mathach	Numeric	8	3		None	None
	12	faed	Numeric	2	0		{2, less than H	None
	13	maed	Numeric	2	0		{2, less than H	None
	14	sex	Numeric	2	0		None	None
	15	faednew	Numeric	8	2 🚔	Father's educa	None	None
	16	maednew	Numeric	8	2	Mother's educ	None	None
	- 17							
< >		ata View λ ν	ariable View /					Ŀ
				5P55 Processor	is ready			11.

• Click on the decimal cell of the **faednew** variable, the two arrows appear for you to change the decimals. Click on the down arrow to change the decimal number to 0. Repeat this step to change the decimals for the **maednew** variable.

13	maed	Numeric	2	0		
14	sex	Numeric	2	0		
15	faednew	Numeric	8	2 🚔	Father's educa	
16	maednew	Numeric	8	2	Mother's educ	
- 17						
	ata view X v ai	riable View /				
e coun	ter area	9	PSS Processor is i	ready		

- Save the changes, to save the changes, from File menu, choose Save to save.
- Label the new values, click in the cell that crossing the **Values** column and the 12th row (**faednew** variable), you will see a small gray box.

🔠 Untitled - SPSS Data Editor								_ 🗆 >	
<u>F</u> ile	<u>E</u> dit	<u>View Data Iransform Analyze Graphs Utilities Window Help</u>							
2									
		Name	Туре	Width	Decimals	Label	Values		
	10	calc	Numeric	4	0		None	None	
	11	mathach	Numeric	8	3		None	None	
	12	faednew	Numeric	4	0		None …	None	
	13	maednew	Numeric	4	0		None	None	
	14	sex	Numeric	3	1		None	None	
	-15								

• Double click on the small gray box, you will see the Value Labels dialogue box as the following. Type 1 in the Value box, and "Less than High School" in the Value Label box. Then click on **Add** button.

Value Labels	? ×
Value Labels Value: 1 Value Label: Less than High School	OK Cancel Help
Add Change Hemove	

• Type 2 in Value box, and "High School Graduate" in the Value Label box. Click on Add. You should have:

Value Labels	? ×
Value Labels Value: Value Label: Add 1 = "Less than High School" 2 = "High School Graduate" Remove	OK Cancel Help

• Repeat Step above. Make the value label "Some Post-Secondary Education" for 3, and "College Graduate & Beyond" for 4. Then click on **OK**. You should see:

	Name	Туре	Width	Decimals	Label	Values
10	calc	Numeric	4	0		None
11	mathach	Numeric	8	3		None
12	faednew	Numeric	4	0		{1, Less th 🔤
13	maednew	Numeric	4	0		None
14	sex	Numeric	3	1		None
15						

- Repeat Step above to change the value labels for the variable **maednew**. You can repeat to add value labels to each variable.
- We can add the labels for the variables with a clear description of the variable when sometimes the meaning is not clear from the variable name itself (e.g., "mathach" we can add a label "Math Achievement" as the label). To do this:
- In the Variable View window, click in the Label column of the **mathach** variable, and type "Math Achievement" in the crossing cell.

8	geo	Numeric	2	0		None
9	trig	Numeric	2	0		None
10	calc	Numeric	2	0		None
11	mathach	Numeric	8	3	Math Achievement	None
12	faed	Numeric	2	0		{2, les
13	maed	Numeric	2	0		{2, les

- You can repeat Step above to add variable label to each variable.
- Save the changes. Make sure you save the data as SPSS (*.sav) file. Click on the Data View tab to switch to the data. Now, you are ready to use this new set of data with recoded values in **faednew** and **maednew** variable.

Recoding Data With Syntax

It is possible to use syntax when recoding variables. For example, if I had a variable that included the following values:

Redbird Bluebird

Yellowbird Elm Butterfly

and I wanted to recode any values that included 'bird' into a new value 'bird'.

To solve to problem the following syntax is an option:

DATA LIST LIST /var1(A15). BEGIN DATA Bluebird Redbird Yellowbird Butterfly Elm END DATA. STRING newVar(A15).

DO IF INDEX(UPCASE(var1), \hat{a} € BIRD \hat{a} €)>0. - COMPUTE newVar= \hat{a} € BIRD \hat{a} € . END IF. EXECUTE.

Example number two is we want to recode the above variables into variables having the same name but with the last letter being replaced by x.

DATA LIST FREE /abc, sal, age, sex1, school,v1234567. BEGIN DATA 85 95 5 87 100 1 END DATA. LIST. SAVE OUTFILE='c:\temp\mydata.sav'. * suppose we want to recode the above variables into variables having the same name but with the las letter being replaced by x. FLIP. STRING newname(A8). COMPUTE newname=CONCAT(SUBSTR(case_lbl,1,LENGTH(RTRIM(case_lbl))-

```
1),"X").
WRITE OUTFILE 'c:\temp\temp.sps'
```

/"RECODE "case_lbl" (1 THRU 87.9=1) (89 THRU 98.1=1) (ELSE=COPY) INTO "newname"."/"FREQ "newname".".

GET FILE='c:\temp\mydata.sav'. INCLUDE 'c:\temp\temp.sps'.