

Fondren Library
Data @ Rice Workshop Series

Introduction to SPSS

Overview:

The software name originally stood for Statistical Package for the Social Sciences (SPSS)

Interface

Data View

🝓 Untitled1 [l	DataSet0] - IBI	M SPSS Statistic	s Data Editor			<u>- 🗆 ×</u>
<u>F</u> ile <u>E</u> dit <u>V</u>	iew <u>D</u> ata	<u>T</u> ransform <u>A</u> n	alyze <u>G</u> raphs	<u>U</u> tilities E <u>x</u>	tensions <u>W</u> in	dow <u>H</u> elp
		🛄 🗠	2		M .	h
					Visible: 0 o	f 0 Variables
	var	var	var	var	var	var
1						
2						
3						
4						
5						-
	4		222			
Data View	ariable View					

Variable View

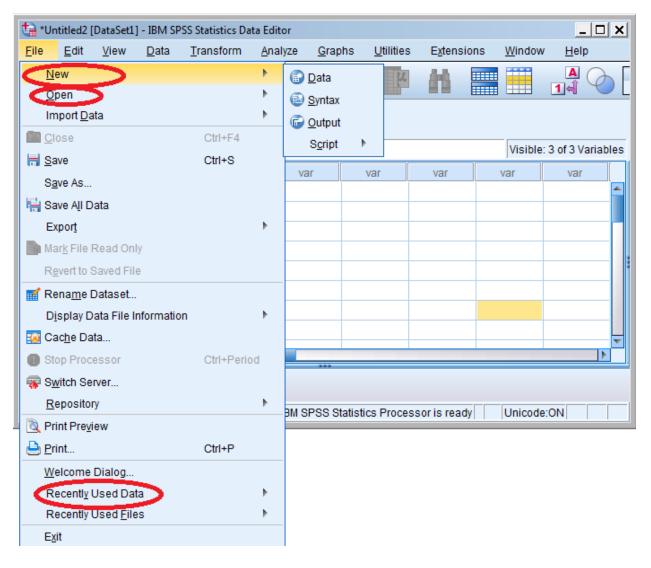
🔚 *Untitled1	[DataSet0] - IBM	SPSS Statistics Da	ata Editor		_	
<u>F</u> ile <u>E</u> dit <u>V</u>	jew <u>D</u> ata <u>T</u> ra	nsform <u>A</u> nalyze	<u>G</u> raphs <u>L</u>	<u>I</u> tilities E <u>x</u> ter	nsions <u>W</u> indow	<u>H</u> elp
🗁 🗄					F H	
	Name	Туре	Width	Decimals	Label	
1	VAR00001	String	4	0		N
2	VAR00002	Numeric	8	2		N
3						
4						
5						
6						_
	4					•
Data View	ariable View					

Data entry and change variable properties(name, type, lable, values, measures)

Useful options under the Menus

File

- > New
- > Open
- > Recently used



View

> Value lables

When it's off, it shows the numaric values. When it's on, it shows the text values input in the Variable View

the stress	titlad2 []	DataSat11 -	IDM CDCC	Statistics Dat	a Editor						
						Oranha	Litilitie	. Evtensi	iono Mindu		
<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	s E <u>x</u> tensi		ow Help	
			III 1	<u>ר מ</u>		* =	۲	H		A (\rightarrow
										\sim	
7:									Visib	le: 3 of 3 Vari	iables
		💑 name	🛷 age	8 gend	er	ar	var	var	var	var	
1		bob	4.00		1						
2		kate	5.00		2						
3											
4											
5											-
6											
7											
8											
0		4									
Data \	/iew V	ariable Vie	w			***					
					IBM S	PSS Statis	tics Proce	ssor is read	ly Unico	de:ON	

ta *Unt	titled2 [DataSet1] -	IBM SPSS S	tatistics Dat	a Editor					_ [×
<u>F</u> ile	<u>E</u> dit	<u>V</u> iew	<u>D</u> ata <u>T</u> r	ansform	<u>A</u> nalyze	<u>G</u> raphs	s <u>U</u> tilitie	s E <u>x</u> tens	ions <u>W</u> ind	ow <u>Holp</u>	
				<u>ר</u> מ						14 1	
7:									Visil	ole: 3 of 3 Vari	ables
		_{aname}	🔗 age	🗞 jende	er va	ar	var	var	var	var	[
1		bob	4.00	M	lale						
2		kate	5.00	Fem	ale						
3											
4											
5											
6											
7											
8											
0		4									▼ ▶
Data V	/iew V	ariable Vie	w			***					
Customize Toolbars IBM SPSS Statistics Processor is ready Unicode:ON											

Data

- > Identify duplicated case
 > Transpose
 > Merge files
 > Select cases

- ➢ Weight cases

ta *I	Jntitled2	[DataSet1] - IBM S	PSS Statistics I	Data Editor							<u>- 🗆 X</u>
<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	E <u>x</u> tensio	ons <u>W</u> indo	ow <u>H</u> elp		
			<u>3</u> Set	fine <u>V</u> ariable F Measuremen py Data Prope	t <u>L</u> evel for U rties							
5:			-	w Custom Attri							Visible: 3 of 3	Variables
		🦂 na		fine date and t				var	var	var	var	
	1	bob	🔡 De	fine <u>M</u> ultiple R	esponse S	ets						
	2	kate 👎		idation	2		•	🧟 Load F	redefined Ru	ules		
	3	том 🤇	🔡 Ide	ntify D <u>u</u> plicate	Cases	>		📓 <u>D</u> efine	Rules			
	4	jim	🏹 Ide	ntify Unusual (Cases			遲 <u>V</u> alidat	e Data			
	5	1	💦 Co	m <u>p</u> are Datase	ts		l					1
	6	1	🔊 S <u>o</u> i	t Cases								
	7	1	📷 Sor	t Varia <u>b</u> les								
	0		Tra	nspose								
		4	-	rge Files			•					
Data	a View	Variable	🐺 <u>R</u> e	structure								
Valio	dation		🔁 Agg	gregate				statistics Pr	ocessor is re	eady U	Inicode:ON	
			Ort	<u>h</u> ogonal Desi	gn		•					
			🚼 Co	py <u>D</u> ataset								
			📰 Spl	it <u>F</u> ile								
			Sel	ect Cases								
				ight Cases.								

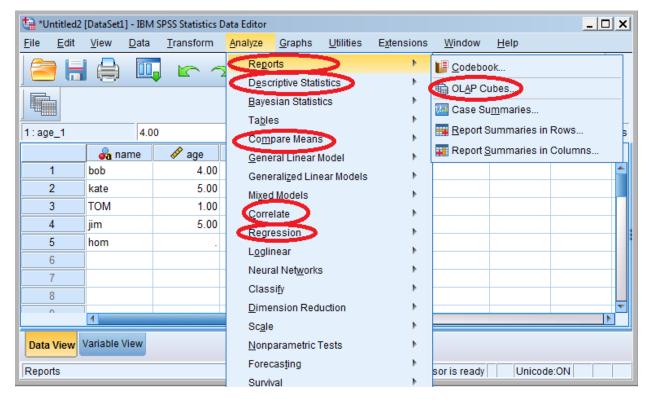
Transformation

- > Compute variable
- > Recode into different variable

ta *I	Untitled2	[DataSet	L] - IBM S	PSS Statistics [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	Exte	nsions	<u>W</u> in	dow	<u>H</u> elp		
				Compute	-				nî.					
				Shi <u>f</u> t Valu	Jes									
5:				🔤 Recode		and the second value of th						Vi	sible: 4 of 4 Va	riables
		🚑 n	ame	🔤 <u>R</u> ecode 🛐 <u>A</u> utomati		the second s	ā			var		var	var	
	1	bob		Visual <u>B</u> i	nning									
	2	kate		🔀 Opt <u>i</u> mal I	Binning									
	3	TOM		<u>P</u> repare	Data for M	odeling		•						
	4	jim		Ran <u>k</u> Ca	ses									
	5			🗎 Date and	d Time Wiza	ard								
	7			🚾 Create T	i <u>m</u> e Series	·								
	8			📲 Replace	Missing <u>V</u> a	alues								
		4		🍘 Random	Number <u>G</u>	enerators				_	_			
Data	a View	Variable	View	Run Pen	ding Trans	forms	Ctrl+(G						
Com	pute Va	riable				1	BM SPSS S	tatistic	s Proc	essor is	ready	Un	icode:ON	

Analyze

- Report-OLAP Cubes
- Descriptive Statistics
- Compare Means
- Correlate
- Regression



Graphs

- Chart BuilderGraphboard Template Chooser

t a ∗∪	ntitled2	[DataSet1] - IB	3M SPSS Statistics	Data Editor						_ [×
<u>F</u> ile	<u>E</u> dit	<u>V</u> iew <u>D</u> a	ta <u>T</u> ransform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	E <u>x</u> tensions	<u>W</u> indow	<u>H</u> elp		
						Builder) Iplate Choose		 €		
						cy Dialogs		Þ			
1:ag	e_1	4	4.00						Visibl	e: 4 of 4 Varia	ables
		🍶 name	🧳 age	💰 gender	ø	age_1	var	var	var	var	
	1	bob	4.00		0	4.00					
	2	kate	5.00		1	5.00					
	3	TOM	1.00		4	1.00					
	4	jim	5.00		1	5.00					
	5	hom			0	3.75					
	6										
	7										
	8										
	0	4									
Data View Variable View											
					IBN	I SPSS Sta	tistics Proces	sor is ready	Unicod	e:ON	

Recently used and Customize

Scrips

ta Ca	aseToVar	s.sav [DataSet10	0] - IBM SPSS Sta	tistics Data Edit	or				_	
<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> tilitie	es E <u>x</u> tension	s <u>W</u> indow	<u>H</u> elp		
			🖡 🗠 r	M 🎬	▙ᆿ▋	<u>۲</u>		(
11 : F	rench							Visi	ble: 5 of 5 Va	riable
		🍰 Name	💰 English	🔗 Math	🖋 Histroy	🔗 French	var	var	var	
	1	Tom	95			89.00				
	2	Tom		89.00	44.00					
	3	Jim	98							
	4	Jill				66.00				
	5	Jim		59.00						
	6	Jim				77.00				
	7									
	8									
	0	1			ine al anora					
Data	View	Variable View						· · · · ·		

taseToVars.sav [DataSet10] - IBM SPSS Statistics 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> tilitie	es E <u>x</u> tension	s <u>W</u> indow	<u>H</u> elp		
			¥ 🎬		4		(
1 : Name	Jill						Visi	ible: 5 of 5 Vari	iables
	🎝 Name	💑 English	🔗 Math	🖋 Histroy	🔗 French	var	var	var	
1	Jill				66.00				
2	Jim	98	59.00		77.00				
3	Tom	95	89.00	44.00	89.00				
4									
5									
6									
7									
8									
0	4								
Data View Variable View									

<u>Terms</u>

Data type

- String
- > Numeric

Level of measures

- Norminal (No ranking, "Race", "Gender")
- > Ordinal (Scale, "Strongly Disagree", "Strongly Agree")
- Scale ("Age", "Height", "Income")

Cross Tablulation

Cross-tabluations are frequency distributitons for two variables together. It gives you a basic picture of how two variables inter-relate.

Gender Gender * Happy happy Crosstabulation

Count

		Нарру	Sad	Bored	Total
Gender	Male	3	2	9	14
	Female	10	0	1	11
Total		13	2	10	25

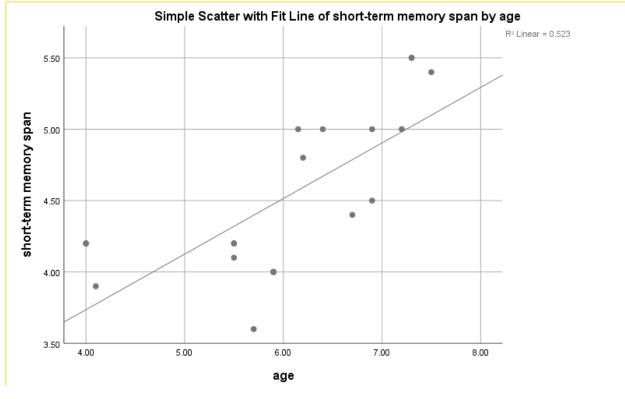
Chi square is used to test the relationship between two <u>norminal or ordinal</u> variables. (If **p-value** is less than **0.05**)

Correlation measures the strength and direction (-1,1) of association between two <u>quantitative</u> <u>variables</u>.

Correlations								
		short-term memory span	age					
short-term memory span	Pearson Correlation	1	.723**					
	Sig. (2-tailed)		.000					
	Ν	20	20					
age	Pearson Correlation	.723**	1					
	Sig. (2-tailed)	.000						
	Ν	20	20					
**. Correlation is significant at the 0.01 level (2-tailed).								

Regression predicts the value one variable(dependent) base on another variable(independent, predictor)

GGraph



Y=2.2+0.39X

Exercise

- 1. Find out how many girls vs boys in our dataset
- 2. Make a pie chart for gender distribution
- 3. What's the average age of the children?
- Recode "Age" into "Age1" (below or equal 6 code as 1, above code as 2) Recode "Read Ability" into "ReadAbility1" (below or equal to 6 code as 1, above code as 2)
- 5. Crosstab between "Age1" and "ReadAbility1"
- 6. Correlation between "Age" and "Memory Span"
- 7. Linear Regression "Age" and "Memory Span"
- 8. Plot "Age" and "Memory Span" with a trend line
- 9. OLAP Cubes Average Grad in 6 years for Black/White weighted by Cohort counts