

Foreign Affairs, Trade and Development Canada



### STATISTICAL ANALYSIS OF POST EVENT COVERAGE SURVEY (PECS) DATA USING SPSS



March 2014

### PECS Survey Data Analysis using SPSS V2

1	Ореі	ning Data Files in SPSS	
	1.1	How to open SPSS	4
	1.2	Importing Data from SPSS, Excel or Stata Files	7
	1.3	Importing Data from dBase or Access Files	11
2	Data	Cleaning	
	2.1	Calculating Age from Date of Birth and Confirming Eligibility	17
	2.2	Data Cleaning - Ensuring the Data Make Sense	35
	2.3	Data Cleaning - Identifying Duplicate Surveys	39
3	Data	Analysis	
	3.1	Data Analysis - Descriptive Statistics	43
	3.2	Data Analysis: Chi Square Calculation	45
	3.3	Data Analysis: Comparison of Means	50

56

3.4 Data Analysis: Logistic Regression

# **Opening Data Files in SPSS**

#### How to open SPSS

#### If you have a SPSS icon on your desktop



Double click the IBM SPSS Statistics Icon to open SPSS.

#### If you do not have an SPSS icon on your desktop



Click on the Start menu as shown in #1. Then left click "All Programs" (#2) to display the programs on your computer.



Click the IBM Statistics folder (#1) to display its contents and then left click "IBM SPSS Statistics" (#2). This will open the SPSS program.



You are now ready to import your data and begin the survey analysis!

Data can be imported from a SPSS (\*.sav, \*.sys, \*.por), Excel (\*.xlm, \*.xlsx, \*.xlsm), dbase (\*.dbase), or STATA(\*.dta) file. **NOTE: These procedures will not work if your data file is open in another window.** 

#### Select the data file to be opened.



First left click "File" (#1). From the drop down menu select "Open" (#2). Then left click on "Data" (#3).

🍓 Open Data	x
Look in: 🌗	Data Manual docs 💦 🔯 🔝 🔠
Nigeria_c Nigeria_p Nigeria_p	dataEW.sav part1.sav part2.sav
Wigeria_p	part3.sav
File name:	3 <u>Open</u>
Files of type:	SPSS Statistics (*.sav)
<u>E</u> ncoding:	
🔲 Minimize	e string widths based on observed values
	Retrieve File From Repository

In the "Look in:" drop down menu (#1), go to the location where your data set is saved. If your data is not a SPSS file with a \*.sav extension, click the "Files of type" drop down menu and select the correct file type (#2). If you are not sure of the file type, select "All files (\*.\*) from the "Files of type" menu. Once located, single left click the desired data set and then click "Open" (#3).

#### Your data will be displayed in DataView

File Edit	taEW.sav [DataS View Data	et1] - IBM SPS	SS Statistics Data	Editor	Sranhs Utilities	Add-ons Window Heln		_ 0 ×
-							1	Visible: 189 of 189 Variables
	q0iiUniqid	q0iiiState	q0ivLGA	q0vWard	q0viVillage	q0viiitype	q0ixDateofint erview	OLDQ0XNAME(
1	1822		1 10	18	18 1		1-Feb-2012	<u>~</u>
2	1821		1 10	18	18 1		1-Feb-2012	
3	1820		1 10	18	18 1		1-Feb-2012	
4	1819		1 10	18	18 1		1-Feb-2012	
5	1818		1 10	18	18 1		1-Feb-2012	
6	1817		1 10	18	18 1		1-Feb-2012	
7	1816		1 10	18	18 1		1-Feb-2012	
8	1815		1 10	18	18 1		1-Feb-2012	
9	1814		1 10	18	18 1		1-Feb-2012	
10	1813		1 10	18	18 1		1-Feb-2012	
11	1812		1 10	18	18 1		1-Feb-2012	
12	1811		1 10	18	18 1		1-Feb-2012	
13	1808		1 10	18	18 1		1-Feb-2012	
14	1807		1 10	18	18 1		1-Feb-2012	
15	1806		1 10	18	18 1		1-Feb-2012	
16	1805		1 10	18	18 1		1-Feb-2012	
17	1804		1 10	18	18 1		1-Feb-2012	
18	1803		1 10	18	18 1		1-Feb-2012	
19	1802		1 10	18	18 1		1-Feb-2012	
20	1801		1 10	18	18 1		1-Feb-2012	
21	1829		1 10	18	18 1		1-Feb-2012	
22	1828		1 10	18	18 1		1-Feb-2012	
-Ô-	1827		1 10	18	18 1		1-Feb-2012	
			1 10	10	10 1		1 Eab 2012	
Data View	Variable View					***		

As the name suggests, Data View (#1) provides a window for viewing data, whereas Variable View (#2) shows the variable names and characteristics.

#### **Reclassify Categorical Data as Numeric**

🔒 *akwa	ibom and Ben	ue Combined fina	al data.sav [Dat	aSet12] - IBM :	SPSS Statistics [	Data Editor						-	CO.		_ 🗆 X
<u>F</u> ile E	ta *Untitled2	[DataSet16] - IB	M SPSS Statistic	s Data Editor	in preserve	-		-		23					
	<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Transform	<u>A</u> nalyze Di	irect <u>M</u> arketing	<u>G</u> raphs <u>U</u> tili	ties Add- <u>o</u> r	ns <u>W</u> indow <u>H</u>	elp						
	l 👝 🖿			-	J	A-0. 185		🔤 📥 🎬							
											Column	a Align	Measure	Role	
166		Name	Type	Width	Decimals	Label	Values	Missing	Columns		12	🚟 Right	🗞 Nominal	🔪 Input	<u> </u>
167		dovicoid	String	16	0		None	None	10		12	🚟 Right	💦 Nominal	🔪 Input	
168		subscriberid	String	15	0		None	None	15		12	Right	💦 Nominal	🔪 Input	
169	6	simserial	String	15	0		None	None	15		12	🚎 Right	💦 Nominal	🖒 Input	
170	7	phonenumber	String	12	0		None	None	12		12	Right	💫 Nominal	🔪 Input	
171	8	eng	String	11	0		None	None	11		12	Right	🗞 Nominal	🔪 Input	
172	9	code enque	String	2	0		None	None	2		12	III Right	🗞 Nominal	🔪 Input	
173	10	quartier dist	String	9	0		None	None	9	EI	12	端 Right	🗞 Nominal	🔪 Input	
174	11	quartier aire	String	24	0		None	None	24	ELe	12	Right	🗞 Nominal	🔪 Input	_
175	12	quartier	String	25	0		None	None	25	ELe	12	Right	💑 Nominal	> Input	
176	13	intrevue date	String	10	0		None	None	10	ELe	50	≣≣ Left	💦 Nominal	> Input	
177	14	no qst	String	1 🛃 Varia	able Type		No.				50	E Left	💑 Nominal	> Input	_
178	15	consent note	String							- e	23	Right	Nominal	> Input	
179	16	q1 <b>1</b>	String	2 Q N	umeric					Le	17	Right	Nominal	> Input	
180	17	q1 autre	String	2 ○ ⊆	omma			Width:	2	Le l	15	≡ Right	Nominal	> Input	-
181	18	q2	String	2 O Do	ot			Decimal Places:	0	- Le	11	≡ Right	Nominal	S Input	-
182	19	q2_autre	String	7 © <u>S</u> o	ientific notation	1		-		e	11	≔ Right	Nominal	S Input	-
183	20	q3	String	1 O D <u>a</u>	ate					Le	10	= Right	Nominal	> Input	-
184	21	q4	String	1 O Do	ollar					Le	10	I Right	Nominal	S Input	-
185	22	q5	String	1 <sup>O</sup> C <u>i</u>	istom currency					_e					-
186		4		. O St	ring										-
107	Data View	Variable View		0 R <u>e</u>	estricted Nume	ric (integer with le	ading zeros)								-
100		variable view			The Numeric to	ne honors the dir	it arouning se	atting while the Re	stricted						
				🛡	Numeric never	uses digit groupi	ng.	stang, while the ree	Stricted						
Data Vi	Data View Variable View 30 OK Connel List														
							Cancer	Telb				IDM OF	PC Statistics Broos	ager in ready	
													- 33 Statistics Proce	ssor is ready	
	e 1				🤹 💫					Σa	X		A TOTAL OF	🥠 🔺	17:10
													A COLON		05/03/2014

For all variables with categorical data, ensure that the variable type is listed as Numeric and not String in the Variable View. To change the Type of a variable, click on the cell where "String" is given and change the Variable Type to Numeric. Then click on "OK"

Data from dBase files (\*.dbase) or Microsoft Access files (\*.mdb) file can be imported into SPSS using the following steps:

#### NOTE: These procedures will not work if your data file is open in another window.



To open data from a dBase or Access file, first single left click "File" (#1), then "Open Database" (#2), and finally "New Query" (#3).



Select the appropriate data source from the list (dBASE File, Excel File or MS Access Database) and click "Next."

Datab	base Wizard	Uniters waterface (Dealers (Dealers	X					
We	Icome to	the Database Wizard!						
This You c	wizard 🛟 OD	BC Driver Login	r0					
retrie to SP	val, an Dat PSS St:	Data Source: MS Access Database						
From	n which	OK Cancel						
	Open	9 818118	22					
	Look in: 🔒 I	Data Manual docs 💦 👔 🔯 🔢 🗄						
	Adamawa Akwa Ibon Benue PE	n PECS Data - Double Entry.MDB n PECS Data - Double Data Entry.MDB iCS Data - Double Data Entry.MDB						
	File name:		Open					
	Files of type:	Access (*.mdb)	Cancel					

To locate your data file, click Browse and follow the path to the file using the "Look in:" drop down menu. Once you have located the correct folder, click the desired data file and single left click "Open".

🖬 ODBC Driver I	.ogin
Data Source:	MS Access Database
Database:	Ial docs\Adamawa PECS Data - Double Entry.MDB
	Cancel

Click "OK" to continue the database wizard.



Select the table that contains the data you will be analyzing from the "Available Tables" column on the left (#1) and click the blue arrow (#2) to move its variables to the "Retrieve Fields in this Order" column on the right. All of the variables in the database should now be listed in the "Retrieve Fields in this Order" column. Then single left click the "Finish" button (#3).

#### Your data will be displayed in the Data View window

💼 *Untitled2	[DataSet1] - IBM	A SPSS Statistic	s Data Editor					_ 0 ×
<u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Transform A	Analyze Direct	t <u>M</u> arketing G	aphs <u>U</u> tilities	Add- <u>o</u> ns <u>W</u> indow <u>H</u> elp		
			- 17	↓			ABG	
j 🔁 🗖								
								Visible: 189 of 189 Variables
	q0iiUniqid	q0iiiState	q0ivLGA	q0vWard	q0viVillage	q0viiitype	q0ixDateofint	OLDQ0XNAME(
							erview	
	4000.00	1.00	40.00	40.00	40.00.4		1510010	
1	1822.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
2	1821.00	1.00	10.00	18.00	18.00 1		1-FeD-2012	
3	1820.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
4	1019.00	1.00	10.00	10.00	10.00 1		1-Feb-2012	
 	1010.00	1.00	10.00	10.00	10.00 1		1-Feb-2012	
0	1017.00	1.00	10.00	10.00	10.00 1		1 Eab 2012	
0	1816.00	1.00	10.00	10.00	10.00 1		1-Feb-2012	
0	1814.00	1.00	10.00	10.00	18.00 1		1-Feb-2012	
3	1014.00	1.00	10.00	10.00	10.00 1		1-Feb-2012	
10	1813.00	1.00	10.00	10.00	10.00 1		1-Feb-2012	
12	1012.00	1.00	10.00	10.00	18.00 1		1 Epb 2012	
12	1909.00	1.00	10.00	10.00	18.00 1		1 Eob 2012	
14	1807.00	1.00	10.00	10.00	10.00 1		1 Eob 2012	
14	1806.00	1.00	10.00	18.00	18.00 1		1 Fab 2012	
16	1805.00	1.00	10.00	18.00	18.00 1		1 Feb 2012	
10	1804.00	1.00	10.00	18.00	18.00 1		1 Feb 2012	
18	1803.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
10	1802.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
20	1801.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
21	1829.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
22	1828.00	1.00	10.00	18.00	18 00 1		1-Feb-2012	
23	1827.00	1.00	10.00	18.00	18.00 1		1-Feb-2012	
24	1020.00	1.00	10.00	10.00	10.00 1		1 Eab 2012	
	1			_				
Data View	Variable View							
							IDM CDCC Clatining Processor	r is ready (

## **Data Cleaning**

#### Calculating Age from Date of Birth and Confirming Eligibility

Before analyzing data, it is important to ensure all respondents meet the eligibility requirements. For vitamin Asupplementation post event coverage surveys (PECS), children must have been between 6 and 59 months of age at the time of the supplementation event. During the data cleaning step, all children outside of the 6-59 month age range should be identified and excluded from the analysis.



To enter the date of the vitamin Asupplementation event, on the top menu, click "Transform" (#1) and select the Compute Variable option (#2).



When the compute variable window opens, perform the following steps to add the last date of the supplementation event to each row:

1. Enter the name of the variable you want to create for the date of the supplementation event in the Target Variable box (#1).

2. In the function group list, left click "Date Creation" (#2).

3. Double click the format you wish to use to enter your date (#3). In this example the mm/dd/yyyy (Mdy) format was selected.

4. DATE.MDY(?,?,?) automatically appears in the Numeric Expression window. Inside the parentheses, change the "?" to the month, day, and year of the last date of the supplementation event (#4). Take care not to remove the commas or parentheses.
5. Click "OK" (#5).

Char	Change the format of the new variable to date format												
192	4												T
Data View	Variable View	-									IBM SPSS Statistics I	Processor is ready	

Click the "Variable View" tab at the bottom of the data table. This will display all of the variables in your database.

le <u>E</u> dit ⊻	∕jew <u>D</u> ata <u>T</u> ra	ansform <u>A</u> nalyz	e Direct <u>M</u> a	rketing <u>G</u> rap	hs <u>U</u> tilities Add-	ons <u>Wi</u> ndow	Help			
<b>&gt;</b> H				▙ ⊒	H		<i>₽</i>			
	Name	Туре	Width	Decimals	Label	Values	Missing	Columns	Align	
185	q199Other	String	255	0		None	None	50	📰 Left	R
186	q444Dontkn	String	255	0		None	None	50	📰 Left	
187	qOxNameofi	Numeric	8	2		None	None	8	🗮 Right	R
188	q45Addition	String	255	0		None	None	50	🧮 Left	R
189	qOiMotherca	String	255	0		None	None	50	📰 Left	2
190	dateofcomp	Date	20	0		None	None	8	🗮 Right	Þ
191	ageinmonths	Numeric	8	0		None	None	13	🗮 Right	ø
192	agecat	Numeric	8	0		{1,6-11mon	None	10	🗮 Right	2
193	ageinmonth	Numeric	8	2		None	None	23	🗮 Right	Þ
194	StartDate	Numeric	11	0		None	None	8	🗮 Right	١U
195		_								$\top$

Scroll down to the last variable to find the new variable name that you entered. Then click on the cell in the "Type" column.



When the Variable Type window opens, first select "Date" (#1) and then select the date format that you prefer (#2). For example, in the example above, the "mm/dd/yyyy" format is selected. Once this is complete click "OK" (#3). Return to the Data View window and confirm that the new variable

has been entered correctly. The new variable will appear as the last column in the database.

#### Events with varying dates of distribution



If different survey locations had different dates of distribution, then use the "If" condition in the Compute Variable window (#1). The "Compute variable: if case" window appears. Select "Include case if satisfies condition" (#2).

From the list of variables in the first column, select the one that identifies the various regions. In this case, dates of distribution differed by state hence state was selected (#3). Click the blue arrow (#4) to send it to the text box.

Specify the condition to be met. e.g. state = 1. Then click continue (#5).

Enter the date of distribution for that location and select OK (#6).

Repeat this procedure for each region until the dates for all regions have been entered.

#### Calculate the age of each child during the supplementation event



Calculate the age of each child at the last vitamin Asupplementation event for children with known dates of birth using the Date and Time Wizard. First click "Transform" (#1) and then select "Date and Time Wizard" (#2) from the drop down menu.



When the wizard starts, select "Calculate with dates and times" (#1) and left click "Next" (#2).



For step 1 of 3, choose "Calculate the number of time units between two dates" (#1) and left click "Next" (#2).



Step 2 of 3 in the Date and Time Wizard is where you will select the two dates from column #1 that you will use to calculate each child's' age. First select the date of the vitamin Asupplementation event that you created in the Variables box (#1) and click on the right arrow (#2). Then select the variable name for the child's date of birth (#3) and click on the lower right arrow (#4). To have the child's age calculated in months, select "Months" from the Unit drop down menu (#5). Finally choose "Truncate to integer" and then left click "Next" (#6).

🔛 Niger Care	taker Data.sav [DataSet1] - IBM SPSS Statistics	Data Editor		ē	J XX
<u>File</u> Edit	View Data Transform Analyze Graph	s <u>U</u> tilities Add- <u>o</u> ns <u>W</u> indow <u>H</u> elp			
🔁 🖿					
				Visible: 61 of 61	1 Variables
	Codeident Codeenque	Date and Time Wizard - Step 3 of 3	Q13	Q14	Q15
1	9 1	Calculation: Datenaiss - date2	2	1	-
2	12 1		2	1	
3	12 3		2	2	
4	21 4		2	1	
5	30 4		1	1	
6	9 4		2	1	
7	11 4		2	1	
8	14 17	Result Variable: Units:	2	2	
9	27 4	Age_Months World's	1	2	
10	22 3	11 12 1 X Variable Label:	1	1	
11	13 3	Age in months when date of birth is known	2	1	
12	4 3	8 4 Execution	2	2	
13	2 3	© Create the variable now O Paste the syntax into the syntax window	2	2	
14	10 3		1	1	
15	12 3		2	2	
16	14 3		1	1	
17	14 5		1	2	
18	8 5			· · · · · · · · · · · · · · · · · · ·	[]
19	28 6		2	1	
20	10 5		1	1	
21	3 5	< Back 4 Finish Cancel Help	1	2	
22	5 5			2	
23	28 6	3 5 3 Z 5 1	1	2	<b></b>
Data View	Variable View				
Date and Tim	ne Wizard	IBI	M SPSS Statistics Process	or is ready	
	🗃 🚺 💁 😱 🕻			▲ 📴 🔐 👫 🕺	:39 AM 25/2014

In the final step, Step 3 of 3, assign a new variable name for the age of the child at the time of the supplementation event (#1). Then add a description of the variable in the "Variable Label" field (#2). Ensure that the "Create the variable now" option is selected in the Execution box (#3). Then click on "Finish" (#4).

#### Recode missing data as 0

🔓 *Niger Caretaker Data.sav [DataSet1] - IBM SPSS Statistics Data Editor														
<u>File</u> Edit	<u>V</u> iew <u>D</u> at	a <u>T</u> ransforr	n <u>A</u> nal	yze <u>G</u> raphs <u>U</u> til	lities Add- <u>o</u> ns <u>W</u> ir	ndow <u>H</u> elp								
😑 H		🛄 🗠		- E to :	1 H	2	- 4 <u>3</u>		AB6					
1: Age_DOB		0										٧	isible: 64 of 64	Variables
	28	codes	qu	dateenquete	date2	AGE MOIS	Datenaiss	Ageenfant	Campaign date	Age DOB	VAR00003	VAR00002	VAR00001	
5239			. 8	23-Jan-2012	16-Oct-1582	33	02-Apr-2009	33	14-May-2012	37				<u></u>
5240			5				01-Apr-2009	33	14-May-2012	37				
5241			8	23-Jan-2012	16-Oct-1582	33	30-Mar-2009	33	14-May-2012	37				
5242			6	23-Jan-2012	16-Oct-1582	33	24-Mar-2009	33	14-May-2012	38				
5243			1	23-Jan-2012	16-Oct-1582	34	21-Mar-2009	34	14-May-2012	38				
5244			3	23-Jan-2012	16-Oct-1582	34	18-Mar-2009	34	14-May-2012	38	-			
5245			5	23-Jan-2012	16-Oct-1582	34	18-Mar-2009	34	14-May-2012	38				
5246			3	23-Jan-2012	16-Oct-1582	34	16-Mar-2009	34	14-May-2012	38				
5247			1	23- Jan-2012	16-Oct-1582	3/	15-Mar-2009	34	14-May-2012	38				
5248			Eine Fin	d and Replace - Dat	a View			38	14-May-2012					
5249								38	14-May-2012					
5250			Fin	d Replace				38	14-May-2012					
5251			Colu	imn: Age DOB	1			38	14-May-2012					
5252			E and					38	14-May-2012					
5253			Fina				· · · ·	38	14-May-2012					
5254	2		Repl	ace with: 0 4			-	38	14-May-2012					
5255	2							38	14-May-2012					
5256	2			Mat <u>c</u> h case				38	14-May-2012					
5257			Sho	ow Options >>				38	14-May-2012					
5258								38	14-May-2012					
5259				Find Next Re	place Replace A	U 5 Close	Help	38	14-May-2012					
5260								38	14-May-2012					
5261	5		2	23-Jan-2012				38	14-May-2012					-
	4											and the second se		•
Data View	Variable Vie	ew												
										IDM ODOO	Statiation Pro-			
	-								-	IBM SPSS	Statistics Proc	essor is ready		55 ANA
		<u> </u>										🔺 🏴 atl	I 📴 😽 8∷ 2/2	25/2014

Select the column in which age was calculated from date of birth (#1). Click on the Find icon (#2) to open the Find and Replace window. In the find section (#3) type . (dot) as shown in the photo above. In the "Replace With" section (#4), type 0 (zero) and click "Replace All" (#5). Using this procedure, all cases where the child's date of birth is not known will be replaced a with 0 (zero).

#### Create a new variable containing the age of children whose date of birth is not known

🔒 Niger Caretaker Data.sav [DataSet1] - IBM SPSS Statistics Data Editor												
<u>File</u> Edit	View <u>E</u>	Transform Anal	lyze <u>G</u> raphs <u>U</u>	tilities Add- <u>o</u>	ns <u>W</u> indow <u>H</u> e	lp						
e 🖛		📄 Compute Va	riable 2				5 🎹 🔺 (	ARC ARC				
		Count Value	s within Cases				′≙ 🎹 🖪 🕯 `		l internet in the second se			
		Shift Values.									Visible: 6	1 of 61 Variables
	Codeide	Recode into	Same Variables		Q3	Q5	cat.Q5	Q8	Q11	Q13	Q14	Q15
1		Recode into	Different Variables		23	3	1 1	2	1	2	1	
2		Martamatia D	Dillerent variables		23	3	1 1	2	1	2	1	
3	Automatic Recode 14 3 2 3 1 2 2											
4		Visual Binnir	1g		:	}	3 2	2	3	2	1	
5		🕌 Ran <u>k</u> Cases			2	)	2 1	1		1	1	
6		🗎 Date and Tin	ne Wizard		2	)	3 2	3	1	2	1	
7		🚾 Create Ti <u>m</u> e	Series		20	6	3 2		1	2	1	
8		Replace Mis	sing Values		31	)	3 2	1	1	2	2	
9		Random Nu	mber Generators		24		3 2	1	1	1	2	
10	Pun Panding Transforms Ctrl+C						3 2	3	1	1	1	
11		Run Pending		CIII+G	2		1 1	1	3	2	1	
12	-	4	3	2	2	3	3 2	3	1	2	2	
13		2	3	2	2		3 2	1	5	2	2	
14		10	3	2	2		1 1	1	1	1	1	
15		12	3	2	2	2	3 2	2	1	2	2	
16		14	3	2	2.	5	3 2	2	1	1	1	
1/		14	5	3			1 1	3	1	1	2	
18		8	5	3		5	3 2	3			· · ·	
19		28	6	3	1:	,	3 2	2	1	2	1	
20		10	5	3	2.		ວ 2 ວ ວ	2	1	1	1	
21	1	5	5	2		) )	2		1	1	2	
22		28	6	э э			2 2		4	1	2	
23	4	20	0	3		,	2	5	'	1	2	▼ 
Data View	Variable View						***		10.0	0000 Obtinting C		
Compute Va	riable								IBM	SPSS Statistics Proce	ssor is ready	0.25 AM
					Σ¢						- 🖻 al 🛱 🍀	8:25 AM 2/25/2014

Select Transform (#1) and then Compute Variable (#2).

ta *caretakerkenya final-SPSS-with regrouped other(category).sav [DataSet1] - IBM SPSS Statistics Dat	ata Editor	o X
File Compute Variable		
*caretakerkenya final-SPSS-with regrouped other(category).sav [DataSet1] - IBM SPSS Statistics Data File          Target Variable         1         Age2         1         Age2         2         4         5         6         0.LDQUIENUME         6         0.LDQUIVENUME         6         0.LDQUIVENUME         6         0.LDQUIVENUME         6         0.LDQUIVENUME         6         0.LDQUIVENUME         6         9         12         6         9         12         13         14         15         16         17         18         19         11         12         13         14         15         16         17         18         19         10         11         12         13         14         15         16         17 <td>te Editor  Editor Editor  Editor  Editor  Editor  Editor  Edit</td> <td>08 Variables var</td>	te Editor  Editor Editor  Editor  Editor  Editor  Editor  Edit	08 Variables var
	Age_Group Age_Gr	
22 8 8.00 0 1.00	Continue Cancel Help	
Data View Variable View		
	IBM SPSS Statistics Processor is ready	2.42.014
		2/25/2014

In the Target Variable box (#1), enter the name you want to give the new variable. In the Type and Label column, select the variable for age in months (#2) given for children whose date of birth is not known. Click the arrow (#3) so that the variable moves to the Numeric Expression box. Click on "If" (#4) and select "Include if case satifies condition" (#5). Select the variable that you created for age in months calculated from date of birth (#6) and click on the arrow (#7). Then type "=0" in the text box after the variable name (#8) and press continue (#9) and then OK (#10).

🄄 *Niger Car	etaker Data.sav	[DataSet1] - IBM	M SPSS Statisti	cs Data Editor								
<u>File</u> Edit	<u>V</u> iew <u>D</u> ata	Transform	Analyze <u>G</u> raj	phs <u>U</u> tilities	Add- <u>o</u> ns <u>W</u> in	dow <u>H</u> elp						
📄 🔓			<b>- 1</b>	<b>*</b> =	P M		- S	1	A 🕗 🌒	ABS		
	Name	Туре	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role	
36	Q237	Numeric	12	0		None	None	12	端 Right	\delta Nominal	💊 Input	A
37	Q238	Numeric	12	0		None	None	12	端 Right	💑 Nominal	💊 Input	
38	Q239	Numeric	12	0		None	None	12	疆 Right	💑 Nominal	🔪 Input	
39	Q2310	Numeric	12	0		None	None	12	遭 Right	💑 Nominal	🔪 Input	
40	Q241	Numeric	12	0		None	None	12	端 Right	\delta Nominal	💊 Input	
41	Q242	Numeric	12	0		None	None	12	端 Right	\delta Nominal	🦒 Input	
42	Q243	Numeric	12	0		None	None	12	🚟 Right	💑 Nominal	💊 Input	
43	Q244	Numeric	12	0		None	None	12	遭 Right	💑 Nominal	🔪 Input	
44	Q245	Numeric	12	0		None	None	12	遭 Right	💑 Nominal	🔪 Input	
45	Q248	Numeric	12	0		None	None	12	疆 Right	\delta Nominal	🔪 Input	
46	Q249	Numeric	12	0		None	None	12	疆 Right	🙈 Nominal	🦒 Input	
47	Q25	Numeric	12	0		None	None	12	🚟 Right	💑 Nominal	🔪 Input	
48	Q26	Numeric	12	0		None	None	12	遭 Right	💑 Nominal	🔪 Input	
49	Q27	Numeric	12	0		None	None	12	Right	🙈 Nominal	🦒 Input	
50	Q28	Numeric	12	0		None	None	12	疆 Right	🙈 Nominal	🦒 Input	
51	codesup	Numeric	12	0		None	None	12	疆 Right	🙈 Nominal	🦒 Input	
52	dateenquete	Date	11	0		None	None	11	🖷 Right	🛷 Scale	🦒 Input	
53	date2	Date	11	0		None	None	12	🖷 Right	🛷 Scale	🦒 Input	
54	AGE_MOIS	Numeric	12	0		None	None	12	Right	🛷 Scale	🦒 Input	
55	Datenaiss	Date	11	0		None	None	11	疆 Right	🛷 Scale	🦒 Input	
56	Ageenfant	Numeric	12	0		None	None	12	疆 Right	🛷 Scale	🦒 Input	
57	Campaign	Date	11	0		None	None	15	🖷 Right	🛷 Scale	🦒 Input	
58	Age_DOB	Numeric	8	0	-	None	None	8	🖷 Right	🛷 Scale	🦒 Input	
59	Age2	Numeric	8	0 🚔	2	None	None	13	疆 Right	💑 Nominal	🦒 Input	
c0	Ago montho	Numorio	0	n		None	Mono	10	I Diaht	A Saala	N. Innut	
Data View	Variable View	4										
											1011.01	
	_										IBM SI	PSS Statistics Processor is ready
				Σ	29 📉							▲ 🔤 🔐 👘 🏶 8:49 AM

Click on the variable view tab and scroll to the bottom to view the new variable that was created. Enter 0 in the Decimals column for the new variable.

#### Recode missing data as 0

🛤 *Niger Car	etaker Data.sav (	DataSet11	- IBM SPSS Stat	tistics Data Editor										_ 0	23
File Edit	View Data	Transform	Analyze (	Graphs Utilities	Add-ons Windo	w Help								<u> </u>	
e 1				ă 🛓 🗐	P A (			\$ }	 1 ⇔์						
1 : Age2	0													Visible: 61 of 61 Vari	iable
2.00	Q27	_	Q28	codesup	dateenquete	d	ate2	AGE_M	IOIS	Datenaiss	 Ageenfant	Campaign_date	Age_DOB	1 Age2	
2791	1	2	5		2 23-Jan-20	12					24	14-May-2012	0	<u> </u>	24
2792		1			2 23-Jan-20	12					24	14-May-2012	0		24
2793		1			2 23-Jan-20	12					24	14-May-2012	0		24
2794		1	2		2 23-Jan-20	12					24	14-May-2012	0		24
2795	]	1	2		2 23-Jan-20	12					24	14-May-2012	0	:	24
2796	]	1	2		2 23-Jan-20	12					24	14-May-2012	0	:	24
2797		2	2		2 23-Jan-20	12					24	14-May-2012	0		24
2798	]	1	2		2 23-Jan-20	12					24	14-May-2012	0		24
2799	]	2 (	Eind and I	Replace - Data View		_		×	) .		24	14-May-2012	0	:	24
2800	]	1		tepiste bata nen	<i></i>						24	14-May-2012	0	:	24
2801	]	1	Find Do	place							24	14-May-2012	0		24
2802		1	Find	place							24	14-May-2012	0		24
2803	1	2	Column:	Age2							24	14-May-2012	0	:	24
2804	]	1	Find:					-			24	14-May-2012	0	:	24
2805	1	1									24	14-May-2012	0		24
2806	1		Replace wit	th: 0 4				~			24	14-May-2012	0		24
2807		1	Match (								24	14-May-2012	0		24
2808		1	imateri e								24	14-May-2012	0		24
2809		1	S <u>h</u> ow Opti	ons >>							24	14-May-2012	0		24
2810		1				-					24	14-May-2012	0		24
2811		1	Eind Ne	ext <u>R</u> eplace	Replace All	5 Close		Help			24	14-May-2012	0		24
2812		1			-				J .		24	14-May-2012	0		24
2813		1			2						24	14-Mav-2012	0		24
Data View	Variable View											IBM SPSS Statistics Pro	cessor is read	y	
		S		2							~		▲ P*	ıl 📴 😽 8:58 Al	M 014

Select the column for the new variable for age in months (#1). Click on the Find icon (#2) to open the Find and Replace window. In the find section (#3) type . (dot) as shown in the photo above. In the Replace With section (#4), type 0 (zero) and click Replace All (#5). Using this procedure, all cases where the child's date of birth is known will be replaced a with 0 (zero).



In order to analyze coverage by age, a single variable needs to be created that contains the age of the child in months. Click on Transform (#1) and then select Compute Variable (#2).

tare *care	takerkenya final-SPSS-with regr	ouped other(ca	tegory).sav [DataSet1] -	IBM SPSS Statistics Da	ta Editor	-									\$
Eile	Compute Variable		pages passes of		-	×									
	Target Variable:	Num	eric Expression:					ABC							
	/ige_inionalio_/iii	] = [=											Visible: 208	of 208 Variab	les
1 2 3 4 5 6 7 8 9	Type & Label	5 ·	                         	8 9 5 6 2 3 0 . Delete	Function group: All Arithmetic CDF & Noncentral CDF Conversion Current Date/Time Date Arithmetic Date Arithmetic Date Creation	ables:		ar	var	var	var	var	var	var	
10 11 12 14 14 15 16 17	a q91Anyadditional       a q3813Communit       d astaidfirence       A GE_IN_MONTHS       A AGE_IN_MONTH       A GE_Group       Campaign_End       It       (ptional case select)	DATE.DMY( correspondi arguments i 31, month b greater than date format	day,month,year). Nume ing to the indicated day, must resolve to integers etween 1 and 13, and y 1582. To display the re to the result variable.	ric. Returns a date val month, and year. The s, with day between 1 year a four-digit integer ssuit as a date, assign	Date.DMdy Date.Mdy Date.Moyr Date.Qyr Date.Wkyr Date.Yrday										
18															
19 20			OK Paste Res	et Cancel Help											
21	0	8.00	0	1.00	12-Dec-2012										
22	8	8.00	0	1.00	12-Dec-2012										÷.
	4		-				_						-		
Data V	ata View Variable View														
									~	part of		<b>▲</b>	• al () 🕈	3:50 PM 2/25/2014	

In the Target Variable box (#1), enter the name of the final variable you want to create for age in months. Select the first variable for age in months that you created (for children with a known date of birth) from the Type and Label column (#2) and move it using the arrow (#3) to the Numeric Expression section (#4). Click on the addition sign (#5), below the Numeric Expression box and go back to the Type & Label column to select the second variable for age in months that you created (for children with an unknown date of birth) and send it to #4. The numeric expression box should now contain an equation for the sum of two ages that were calculated. Click OK (#6) to create a variable for age in months for all children.

#### Exclude surveys of children who were not 6-59 months at the time of the supplement event

<u>D</u> ata	form	<u>A</u> nalyze	Direct <u>M</u> arketing
😡 C	Define <u>V</u> ariable	Propertie	S
<b>%</b> s	Set Measurem	ent <u>L</u> evel fo	or Unknown
	Copy Data Prop	perties	
	New Custom A	ttri <u>b</u> ute	
8	) <u>e</u> fine Dates		
<b>##</b> C	Define <u>M</u> ultiple	Response	e Sets
٧	/alidation		4
	dentify D <u>u</u> plica	te Cases	
<u>s</u> 10	dentify Unusua	al Cases	
🔣 🔁	Com <u>p</u> are Data	sets	
🔊 S	3 <u>o</u> rt Cases		
<b>36</b> S	Sort Varia <u>b</u> les.		
T 🔝	ra <u>n</u> spose		
N	llerge Files		۲.
F	Restructure		
<b>a</b>	ggregate		
C	Ort <u>h</u> ogonal De	sign	۲.
🔣 🔁	Copy <u>D</u> ataset		
== s	Split <u>F</u> ile		
<u> </u>	elect Cases	2	
4 <b>1</b> 2 V	Veight Cases.	~	

First, go to the Data menu (#1) and click "Select Cases" (#2).

Select Cases
Image: Construction of the product
OK Paste Reset Cancel Help

When the Select Cases window opens, select the option "If condition is satisfied" (#1) and left click the blue "If..." button below (#2).

Select Cases: If	Age >= 6 & Age <= 59	<u></u>
aq q43990ther         aq q4399aSpecifyot         aq q44Howoftensho         aq q442Every6mont         aq q442Every6mont         aq q442Every6mont         aq q44990ther         aq q44990ther         aq q44990ther         aq q44990ther         aq q44990ther         aq q44990ther         aq q412Know2b         aq q412Know2c         aq q413Know2c         aq q414Know2d         aq q414Know2d         aq q414Know2d         aq q45Additionalinf         aq 00Nothercaretak         Ag q45Additionalinf         Ag q0iMothercaretak         Ag ag at Malezi Bor	+ < > 7 8 9 - <= >= 4 5 6 * = ~= 1 2 3 / &   0 . ** ~ () Delete	All Arithmetic CDF & Noncentral CDF Conversion Current Date/Time Date Arithmetic Date Creation
	4 Continue Cancel Help	

The Select Cases If window will appear where you can specify what cases you want to select. In this case, we want to only include children who were between 6 and 59 months of age at the time of the supplementation event. First select the variable that corresponds to the child's age at the

time of the supplementation event (#1). This is the variable we created in previous step. Then click the right arrow (#2). Enter the conditions for eligibility as shown above, for example, if age >=6 and age <=59.

\*\*Note: Using the survey end date to calculate age at the time of the supplementation event ensures that all children who were six months at the time of the campaign will be included in the analysis. However, it is possible that children at the upper end of the age range could be excluded. It is important to check that children listed as having 60 months were not mistakenly excluded.

<u>D</u> ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Ado	d- <u>o</u> ns	<u>W</u> indow	<u>H</u> elp					
<u>}</u> []		Reg	orts			pe	44	*				_	
	•	D <u>e</u> s	scriptive Sta	atistics	9	123	Frequencie	es				1	
		Ta <u>b</u>	les		•	Pa	Descriptive	S					
umber	qiiiEnum	Cor	<u>m</u> pare Mea	ns	•		Explore		cation	qvii	Village		¢
		<u>G</u> er	neral Linea	r Model	•		Crosstaba						i
2		Ger	nerali <u>z</u> ed Li	inear Mod	lels				1			4	
		Mi <u>x</u> e	ed Models		•		TURF Anal	ysis				1	
7		<u>C</u> or	relate			1/2	<u>R</u> atio		1			1	
10		<u>R</u> eg	gression		•	<b>*</b>	P-P Plots		1			1	
13		L <u>o</u> g	linear		•	-	<u>Q</u> -Q Plots		1			1	
18		Cla	ssify		•		1					1	
22		Dim	nension Re	duction	•		1		. 1			1	
29		Sc <u>a</u>	le		•		1		1			1	
9		Nor	nparametri	c Tests	•		1		1			1	
17		For	ecas <u>t</u> ing				1		1			1	
19		<u>S</u> un	vival		•		1		1			1	
21		Mult	tiple Respo	onse	•		1		1			1	
24		🐺 Sim	ulation				1		1			1	
28		Qua	ality Control	I			1		1			1	•
31			C Curve				2		2			2	•
39		1	<u> </u>	2			2		2			2	
41		1		2			2		2			2	•
46		4		2	2		2		2			2	ſ

**Data Validation - Cross Variables** 

It is important to check that your data makes sense before embarking on analysis. To begin with, it is advisable to check how your sample size is distributed among the smallest units (in this case clusters). Go to Analyze (#1), then click on Descriptive Statistics (#2) and finally Crosstabs (#3).

#### **Crosstab Variables**

taretakerk *	enya final-SPSS-with regrouped	other(category).	.sav [DataSet1] - IBM SPSS Statistics Data Editor	
<u>File</u> Edit	View Data Transform An	alyze <u>G</u> raphs	s <u>U</u> tilities Add- <u>o</u> ns <u>W</u> indow <u>H</u> elp	
		- 🎬 🗸		ABG
<u> </u>				
				Visible: 208 of 208 Variables
	qiDate	qiiSurveynum	OLDQIIIENUMERATOR	OLDQIVPROVINCE
		ber	f Crosstahr	
1	14 Jun 2012 00:00:00	20.00		
2	14-Jun-2012 00:00:00	29.00	R <u>o</u> w(s):	Statistics
3	13-Jun-2012 00:00:00	25.00	ajDate	Calls
4	11-Jun-2012 00:00:00	23.00		Somet
5	9-Jun-2012 00:00:00	28.00	OLDQIVPROVINCE Column(s):	
6	12-Jun-2012 00:00:00	6.00	a OLDQVDISTRICT	Crosstabs: Cell Display
7	21-Jun-2012 00:00:00	22.00		
8	19-Jun-2012 00:00:00	10.00	al aver 1 of 1	Counts z-test
9	15-Jun-2012 00:00:00	17.00	a 2Doyouhaveachildhe	Observed Compare column proportions
10	20-Jun-2012 00:00:00	8.00	a q2alfthecaretakerhas Previous Net	t Expected Adjust p-values (Bonferroni method)
11	11-Jun-2012 00:00:00	27.00	Q2blfthechilddoesnoth	Hide small counts
12	13-Jun-2012 00:00:00	19.00	a daverifvusinghealthc	Less than 5
13	12-Jun-2012 00:00:00	15.00	a g3bIndicatedateofme	Percentages Residuals
14	27-Jun-2012 00:00:00	21.00		Bow A Unstandardized
15	20-Jun-2012 00:00:00	25.00	Display clustered bar charts	Column Standardized
16	20-Jun-2012 00:00:00	30.00	Suppress tables	Total Adjusted standardized
17	15-Jun-2012 00:00:00	23.00		
18	20-Jun-2012 00:00:00	2.00	OK Paste Reset Cancel Help	Noninteger Weights
19	12-Jun-2012 00:00:00	30.00	<u></u>	Round cell counts     O Round case weights
20	18-Jun-2012 00:00:00	22.00		O Truncate cell counts      O Truncate case weights
21	14-Jun-2012 00:00:00	11.00		○ No adjustments
- 00	14 Jun 2012 00-00-00	26.00		
			***	
Data View	variable view			
				IBM SPSS Statistics Processor is ready
				▲ 📑 🖉 🕂 🗇 🐲 4:07 PM
				2/25/2014

Crosstabs allows you to view the responses to two questions side by side to make sure the responses make sense. For example, we can view the Area Type selected by each respondent broken down by cluster. In most cases, a rural village should not have been marked as urban and vice versa. In cases where both area types are selected for a cluster, the reason for the discrepancy should be investigated. To do the crosstabs calculation, select one variable (e.g. cluster) and move it to the Rows box (#1), then select the other variable (e.g. Area type) move it to the Columns box (#2). Click on Cells (#3) to open the Cell Display window. Here click on "Row" to display the percent who gave each possible response. Click on Continue (#5) and finally OK (#6) to run the analysis.

			2	Total	
		)	1	2	
	Count	2 1	27	0	28
0	% within <u>qviiiAreatype</u>	100.0%	3.8%	0.0%	3.1%
-	Count	0	2	0	2
(	% within <u>qviiiAreatype</u>	0.0%	0.3%	0.0%	0.2%
45	Count	0	0	30	30
15	% within <u>gviiiAreatype</u>	0.0%	0.0%	16.4%	3.3%
10	Count	0	31	30	31
10	% within <u>gviiiAreatype</u>	0.0%	0.1%	16.4%	3.4%
17	Count	0	29	0	29
17	% within <u>gviiiAreatype</u>	0.0%	4.1%	0.0%	3.2%
10	Count	0	27	4 3	30
18	% within <u>gviiiAreatype</u>	0.0%	3.8%	1.6%	3.3%
00	1	715	183	899	
23	100.0%	100.0%	100.0%	100.0%	
Total					
Total					

#### qviiVillage \* qviiiAreatype Crosstabulation

In this output you will identify a number of mistakes. Column #1 shows missing information on area type from cluster 6 (#2). #3 and #4 indicate clusters in which area type responses are inconsistent. These entries should be given a second look to see if revisions are needed.

#### Validation Check - Incorrect Responses



In some cases, incorrect entries can be determined by running a simple frequency analysis. If you want to investigate if out of range responses appeared in any question, it is recommended to run descriptive statistics. To do this click on Analyze (#1) and then Descriptive Statistics (#2). Select Frequencies (#3) and a Frequency window appears. Select all the variables to investigate and click on the arrow to move them to the Variable(s) box (#4). Click OK (#5) and go to output window to see the frequency of responses to each variable (#6). In the example above, the question asks the gender of the child, however there are five different options given whereas only two options (male and female) were expected. In this case you should go back to the original entry and try to determine the correct answer.

Image: Construction of the properties	<u>File E</u>	dat ⊻i	iew	Data	(1) tm	Analyze	Direct Marketing	Graphs	Utilities Add	i- <u>o</u> ns <u>Wi</u> ndow	Help		
t : q0/SeriaNumber       Copy Data Properties       merato       qi/Province       qv/District       qv/		H	e	😡 D 🏹 S	et Measureme	Properties	s or Unknown		AA 🐮	<u> </u>	42 🔛	1 A (	
Image: Construction parameter         mercato         qr/District         qr/District         qr/Sublocation         qr/Willage         qr/Willage <th>1 : q0iSe</th> <th>riaNum</th> <th>ber</th> <th></th> <th>oov Data Pros</th> <th>verties .</th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>	1 : q0iSe	riaNum	ber		oov Data Pros	verties .							
Image: Construction Analysis         Image: Construction Analysis           Image: Construction Analy			0		eur Custom &	theira da		merato	qivProvince	qvDistrict	qviSublocatio	qviiVillage	qviiiAreaty
Image Set and S					efine Deles	angere		r			n		
1       12       aiii Centre graphe response deal       1.00			10.1		efine Midinia	Remonee	Sate	1.00	4.00	4.00	1.000	4.00	
2       124       Vapadol       1.00	1	_	12-J		childeline.	response	50t5	1.00	1.00	1.00	1.00	1.00	1
3       12.5       100       1.00 <td< th=""><th>2</th><th></th><th>12-5</th><th></th><th></th><th></th><th>6</th><th>1.00</th><th>1.00</th><th>1.00</th><th>1.00</th><th>1.00</th><th>1</th></td<>	2		12-5				6	1.00	1.00	1.00	1.00	1.00	1
4         12-1         1.00         1.	3	_	12-J		енину одресе	te cases		1.00	1.00	1.00	1.00	1.00	1
5       124       S ort Variagles       1.00<	4		12-J	R-4 10	lentify Unusua	al Cases		1.00	1.00	1.00	1.00	1.00	1
b         124         Soft Variables         1.00	5		12-J	₿ S	ort Cases			1.00	1.00	1.00	1.00	1.00	1
7       12-5       Image Files       1.00	6	_	12-J	S 🔤	ort Variables.			1.00	1.00	1.00	1.00	1.00	1
8         12-4         Merge Files         1.00	/		12-5	T 🔝	ranspose			1.00	1.00	1.00	1.00	1.00	1
9       12-       Image: Restructure       2.00       1.00	8	_	12-J	M	lerge Files		*	1.00	1.00	1.00	1.00	1.00	1
10       12-4       2.00       1.00	9		12-J	🗰 B	estructure			2.00	1.00	1.00	1.00	1.00	1
11         12-s         Orthogonal Design         200         1.00	10	_	12-J	a 🔤	ggregate			2.00	1.00	1.00	1.00	1.00	1
12       12       12       12       1.00       1.0	11		12-J	0	vthogonal Des	ign -	*	2.00	1.00	1.00	1.00	1.00	1
13         12- 3         Spit Fie         2.00         1.00	12		12-J	🐂 c	opy <u>D</u> ataset			2.00	1.00	1.00	1.00	1.00	1
14       12-3       2.00       1.00       <	13		12-J	= s	plt Fie			2.00	1.00	1.00	1.00	1.00	1
15       12-1       100       1.00 <t< th=""><th>14</th><th></th><th>12-J</th><th>🔳 s</th><th>elect Cases</th><th></th><th></th><th>2.00</th><th>1.00</th><th>1.00</th><th>1.00</th><th>1.00</th><th>1</th></t<>	14		12-J	🔳 s	elect Cases			2.00	1.00	1.00	1.00	1.00	1
16         12-0         1.00         1	15		12-J	ath w	wint Cases			2.00	1.00	1.00	1.00	1.00	1
17         12-Jun-201         17.00         17.00         3.00         1.00	16		12-J	-1	ogni ouses.			2.00	1.00	1.00	1.00	1.00	1
18         12-Jun-201         18.00         18.00         3.00         1.00	17		12-J	un-201.	1	7.00	17.00	3.00	1.00	1.00	1.00	1.00	1
19         12-Jun-201         19.00         19.00         3.00         1.00	18	_	12-J	un-201.	1	8.00	18.00	3.00	1.00	1.00	1.00	1.00	1
20         12-Jun-201         20.00         20.00         3.00         1.00         1.00         1.00         1.00         1           21         12-Jun-201         21.00         21.00         3.00         1.00         1.00         1.00         1         <	19		12-J	un-201.	1	9.00	19.00	3.00	1.00	1.00	1.00	1.00	1
21         12-Jun-201         21.00         21.00         3.00         1.00         1.00         1.00         1.00         1           22         12-Jun-201         22.00         22.00         3.00         1.00         1.00         1.00         1         1.00         1           23         12-Jun-201         23.00         23.00         3.00         1.00         1.00         1.00         1<	20		12-J	un-201	2	0.00	20.00	3.00	1.00	1.00	1.00	1.00	1
22 12-Jun-201 22.00 22.00 3.00 1.00 1.00 1.00 1.00 1 23 12-Jun-201 23.00 23.00 3.00 1.00 1.00 1.00 1.00 1.00 1	21		12-J	un-201	2	1.00	21.00	3.00	1.00	1.00	1.00	1.00	1
23 12 Jun 201 23 00 23 00 3 00 1 00 1 00 1 00 1 0	22		12-J	un-201	2	2.00	22.00	3.00	1.00	1.00	1.00	1.00	1
20 1200 2000 100 100 100 100 100 100 1	23		12-J	un-201	2	3.00	23.00	3.00	1.00	1.00	1.00	1.00	1

Each survey should have a unique identifier (survey number) in order to identify duplicates. If there are two surveys with the same number, then the reason for this needs to be addressed. Perhaps the same data was entered twice, or there was a data entry error and the wrong survey number was entered. To search for duplicates, go to Data (#1) and select "Identify Duplicate Cases" (#2).



Awindow for identifying duplicate cases will appear. In the first column, select the variable name for the unique identifier (#1), and click the blue arrow (#2) to move the variable to the "Define matching cases by" box (#3). By default, in the Variables to create box, "Last case in each groups is primary" (#4) will be selected. This means that if a duplicate is found, the first entry will be treated as a duplicate and the second/last entry will be treated as a duplicate. #5 indicates that a new variable will be created and the variable name will be PimaryLast. Click OK to view if there are any duplicates.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Duplicate Case		.1	.1	.1
1 rimary Case	899	99.9	99.9	100.0
Total	900	100.0	100.0	

Indicator of each last matching case as Primary

In the output window, you will have a table showing the number of duplicates identified (#1). In this case, there is one duplicate (#2).

#### **Deleting the duplicates**

<u>F</u> ile	Edit	<u>V</u> iew	Data Transform	<u>A</u> nalyze <u>G</u>	raphs <u>U</u>	tilities Add- <u>o</u>	ns <u>W</u> indow	<u>H</u> elp				
6			🛄 🗠 /	<b>- 1</b>	<b>*</b> =	<u>A</u> A	*, 🔽	<b>-</b> 42	1		ABC	
20 : IN	TV mnt	h	9									
, 	-	ter_\$	BirtDate	MNCHDate	Age@M NCH	Age_Old2	IntvDATE	Age@I	AgeMonts	Try5	PrimaryLast	var
1	1		1 27.05.2011	16.07.2012	13	12	08.09.2012	15	13	05/17/2023	0	
2	2		1 16.02.2011	16.07.2012	17	12	08.09.2012	18	17	05/17/2023	1	
3	3		1 14.01.2011	16.07.2012	18	12	08.09.2012	19	18	05/17/2022	1	
4	1		1	16.07.2012	-	8	08.09.2012		8	05/17/2022	1	
5	5		1 04.01.2010	16.07.2012	30	24	08.09.2012	32	30	05/17/2022	1	
6	6		1 19.09.2009	16.07.2012	33	36	08.09.2012	35	33	05/17/2022	1	
1	7		1 15.09.2009	16.07.2012	34	36	08.09.2012	35	34	05/17/2022	1	
8	3		1 07.01.2012	16.07.2012	6	8	08.09.2012	8	6	05/17/2022	1	
9	)		1 22.10.2010	16.07.2012	20	24	08.09.2012	22	20	05/17/2022	1	
1	0		1 06.05.2011	16.07.2012	14	12	08.09.2012	16	14	05/17/2022	1	
1	1		1 12.08.2008	16.07.2012	47	48	08.09.2012	48	47	05/17/2022	1	
1	2		1 23.11.2011	16.07.2012	7	12	08.09.2012	9	7	05/17/2022	1	
1	3		1.	16.07.2012		9	08.09.2012		9	05/17/2022	1	
1	4		1 26.04.2011	16.07.2012	14	12	08.09.2012	16	14	05/17/2022	1	
1	5		1.	16.07.2012		28	08.09.2012		28	05/17/2022	1	
1	6		1 30.03.2008	16.07.2012	51	48	08.09.2012	53	51	05/17/2022	1	
1	7		1 .	16.07.2012	-	48	08.09.2012	-	48	05/17/2022	1	
1	8		1 20.09.2011	16.07.2012	9	12	08.09.2012	11	9	05/17/2022	1	
1	9		1.	16.07.2012		45	08.09.2012		45	05/17/2022	1	
2	0		1 .	16.07.2012	-	38	08.09.2012	-	38	05/17/2022	1	
2	1		1.	16.07.2012		16	08.09.2012		16	05/17/2022	1	
2	2		1 17.08.2011	16.07.2012	10	12	08.09.2012	12	10	05/17/2022	1	
2	3		1.	16.07.2012	-	36	08.09.2012		36	05/17/2022	1	
2	4		1 25.03.2011	16.07.2012	15	12	08.09.2012	17	15	05/17/2022	1	

Once you know how many duplicates you have, it is easy to locate them using the new variable that was created (PrimaryLast) (#1). By default, the duplicates will be listed first. It is recommended to go back to the surveys to confirm if two surveys were given the same number by mistake or if one record was entered twice. Delete one record if it is a duplicate.

### **Data Analysis**



To view frequencies and percentages for each variable, on the menu bar select "Analyze"(#1). From the drop down menu select "Descriptive Statistics" (#2) then "Frequencies" (#3).



When the Frequencies window opens, select all of the variables you would like to calculate frequencies for from the list on the left (#1). To select multiple variables at one time press and hold "CTRL" while selecting each. Once selected, click the blue arrow between the two lists (#2). All of the variables selected will then appear in the list on the right (#3). Ensure the "Display frequency tables" box is checked and click "OK" (#4).

请 *Output2 [Document2] - IBM SPS	SS Stat	istics Vie	wer													x
<u>File Edit View Data Trans</u>	sform	Insert	F <u>o</u> rmat	<u>A</u> nalyze D	irect <u>M</u> arke	ting <u>G</u> raphs	Utilities Add-or	is <u>W</u> indow	<u>H</u> elp							
😑 H 🖨 🙇 👌	😑 H 🖨 🙋 🕗 📖 🗠 🛥 🧝 🔛 🏪 === 📀 🕒 🦫 🔛															
<b>* * * -</b>																
Image: a constraint of the second	a29mhhd*q4cvaever												4			
Crosstab	q4cvaever															
a30Highestschoolinglevelred	Cumulative															
Title				Frequency	Percent	Valid Percent	Percent									
Clossiab Chi-Square Tests		Valid		75	9.1	9.1	9.1									
a31Typeofemploymentthatyo			No	37	4.5	4.5	13.6									
Title			Yes	711	86.4	86.4	100.0									
🔓 Crosstab			Total	823	100.0	100.0										
Chi-Square Tests						•										
g33mrelig * q4cvaever																
Title					q0viiity	pe										
Clossiab Chi-Square Tests							Cumulativ	е								
				Frequen	cy Perce	nt Valid Perce	ent Percent									
Frequencies		Valid	Not Rura	al 28	51 3	0.5 30	).5 3	0.5								
🖄 Title	1		Rural	57	72 6	9.5 69	9.5 10	0.0								
R Notes			Total	82	23 10	0.0 100	0.0									
Active Dataset																
General Statistics																
Frequency Table					q1Sexofthe	child										
Title							Cumulative	7								
Q4cvaever				Frequency	Percen	Valid Percent	t Percent									
a1Sexoffbechild		Valid		6		7 .7	7	7								
a25care			Male	405	49.	2 49.2	2 49.	9								
a q26mrel			Female	412	50.	1 50.1	100.	5								
q27Maritalstatus			Total	823	100	100.0										
🛁 🛱 q28mage																
- 🛱 q29mhhd																
					q25care											
q31Typeofemploymentth							Cumulative	1								
- La q33mrelig				Frequency	Percent	Valid Percent	Percent									_
							1	1				IDU ODOO CU		- to		
												IBM SPSS Statis	tics Processo	r is ready	H: 22, W: 9	95 pt.

After selecting "OK" the frequency tables can be viewed in the Output window.

In this section we will examine differences between participants whose child received vitamin A and those who did not. The chi-square test and respective p-value will determine whether there are significant differences between Vitamin Asupplemented and non vitamin Asupplemented individuals by given variables.



Left click "Analyze" (#1) at the top of the screen. From the drop down menu select "Descriptive Statistics" (#2) then select "Crosstabs" (#3)

Crosstabs	an and a	23		1-1 CD-2V 12
	Row(s):	Exact	Crosstabs: Statistics	
	a q0viiitype	Statistics	Chi-square 5	Co <u>r</u> relations
al a	<u>₽ a25coro</u>	Cells	Nominal	Ordinal
	<u>C</u> olumn(s):	Format	Contingency coefficient	t 📃 <u>G</u> amma
a quvivillage	🖧 q4cvaever	Bootstran	Phi and Cramer's V	Somers' d
		Doorongp	🔲 Lambda	🔲 Kendall's tau- <u>b</u>
aq0xiNoofchildrenbtwn	r 1 of 1		Uncertainty coefficient	🔲 Kendall's tau- <u>c</u>
a q1bHavebirthcertificat	evious Next		-Nominal by Interval	- Kanna
a q2adobeling			Fta	Risk
a q2cverifybycard				McNemar
a3byal2				
			Cochr <u>a</u> n's and Mantel-H	aenszel statistics
Di:	sp <u>l</u> ay layer variables in table layers		lest common odds ratio	o equals: 1
Display clustered <u>b</u> ar charts			6 Continue Cano	cel Help
Suppress tables				
OK Paste	Reset Cancel Help			3-Feb-2012
			J	28-Jan-201

When the Crosstabs window appears, select from the list on the left (#1) all of the variables you would like to examine differences for between supplemented and non-supplemented children. Once the variables are selected, click the blue arrow (#2) to add them to the Row(s) box. For the box labeled "Column(s), select the variable that corresponds to whether vitamin Awas recieved

and click the blue arrow (#3) to the right of the box. Then select the "Statistics" button (#4). In the Crosstabs:Statistics window select Chi-square by clicking the check box to its left (#5) and then press "Continue" (#6).

houseare					
🔄 Crosstabs		×	7	7	1
	Devide		28	28	1
	R <u>o</u> w(s):	E <u>x</u> act	10	10	1
- q201 specifyother		Statistics	- 20		
a d291 Specify other		Cells	<b>t</b> a (	xact Tests	
a g30Whatdoyouusefor	Column(s):				-
- 🚜 q309aSpecifyOther		Format		Asymptotic only	
- 💫 q31Doyouownaworki	➡	Bootstrap. 2	90	Monte Carlo	
💫 q32lsthereaworkingr				<u>C</u> onfidence level:	99 %
🔄 💑 q33lsthereaworkingte	Layer 1 of 1			Number of samples	s: 10000
da q34Wheredoyouseek	Previous		0	Exact	
a q341GovernmentHeal				Time limit per te	est: 5 minutes
a 242 covernment Heal			_		
a d344Privateheathdis	➡		EX8 wh	ict method will be used en computational limits	a instead of Monte Carlo
a q345Privateheatthcen			Eor		da, call counta ara alwana
			rou	nded or truncated in c	omputing the test statistics.
Display alustared has aborte	Display layer variables in table layers				
Display clustered bar charts				Continue Ca	ancel Help
- Subbress Tables					
ОК	Paste Reset Cancel Help		20	20	
			18	18	1

Left click the "Exact" button (#1) and then select Monte Carlo (#2) and press "Continue" (#3). This will give the results of the Fisher's exact test in the output.

Crosstabe				
Crosstabs				
🔍 ud Caucettia anticitat	Row(s):	Exact		
a di Sexonnechia	Qa dovadoses	Statistics		
a glindicatedateoncard		Cells	🗎 Crosstabs: Cell D	
💑 q2adobeling	Column(s):		-Courto	tost
🚜 q2cverifybycard		Format	Counts	2-1651
🚜 q3meas	➡	Bootstrap	Observed	Compare column proportions
🚜 q3bval2			Expected	Adjust p-values (Bonferroni method)
💰 q3cmdaymmomyr	Layer 1 of 1		📃 Hide small counts	
💑 q3dn meas			Less than 5	
💑 q3d1nmeastxt	Previous <u>N</u> ext		$\sim$ 2	
🍓 q6ctaken			-Percentages	Residuals
deactakentxt	→		Row	🔲 Unstandardized
a q7cloc			Column	Standardized
💑 d\acioctxt 🗾			🔲 Total	Adjusted standardized
	🔲 Display layer variables in table layers			
📃 Display clustered bar charts			-Noninteger Weights-	
Suppress tables			Round cell counts	◎ Round case weights
	Pasta Reset Cancel Holn		O Truncate cell coun	ts 🔘 Truncate case weights
			O No adjustments	-
0				
0			3 Cont	inue Cancel Help
0				
0				

To compare the various responses it is important to have their percentages and frequencies. To view the frequencies and percentages for each response, click on "Cells" (#1) and then in the Crosstabs: Cell Display window, select "Row" (#2). Click continue (#3). Once this has been

completed, select "OK" (#4) to run the analysis.

👘 *Output2 [Document2] - IBM SPSS	Statistics Viewer										
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>D</u> ata <u>T</u> ransfo	orm <u>I</u> nsert F <u>o</u> rmat <u>A</u> nalyze	Direct Mark	eting <u>G</u> rap	phs <u>U</u> tilitie	es Add- <u>o</u> n:	s <u>W</u> indow	<u>H</u> elp				
😑 🗄 🖨 🗟 🦉	) 🛄 🖛 🛥 🍹	<b>1</b>	▙╡	6	•	•		è 🔳			
Title  Crosstab  Crosstab	CROSSTABS /TABLES=q0viiitype ypeofemploymentthaty /FORMAT=AVALUE TAB /STATISTICS=CHISQ /CELLS=COLUMN /COUNT ROUND CELL. Crosstabs	qlSexof ouhave q LES	thechild 33mrelig	q25care BY q4cva	q26mrel aever	q27Marit	alstatus	q28mage q29mhhd q30Highestschoolinglevelreceive q31T			
— 🚔 Title	[DataSat11 C. \Ueare)	Dabblas\	Deakton	Data Mani	al doce	AdamawaF	inal ear				
Crosstab	[bacaseci] C. (bseis)	reppies(	Desktop(I	Jaca Main	ar uocs (	Audillawar	11141.54V				
a28mage * d4cyaever		_		_							
Title		Case	Processing	g Summary							
- Grosstab				Cas	ses						
Chi-Square Tests		Va	lid	Miss	sing	To	tal				
e q29mhhd * q4cvaever		N	Percent	N	Percent	Ν	Percent				
- Title	q0viiitype * q4cvaever	823	100.0%	0	0.0%	823	100.0%				
Crosstab	g1 Sexofthechild *										
Chi-Square Tests	q4cvaever	823	100.0%	0	0.0%	823	100.0%				
	q25care * q4cvaever	823	100.0%	0	0.0%	823	100.0%				
→ Crosstah	q26mrel * q4cvaever	823	100.0%	0	0.0%	823	100.0%				
Chi-Square Tests	q27Maritalstatus *										
q31Typeofemploymentthatyo	q4cvaever	823	100.0%	0	0.0%	823	100.0%				
🖆 Title	q28mage * q4cvaever	784	95.3%	39	4.7%	823	100.0%				
👘 Crosstab	q29mhhd * q4cvaever	823	100.0%	0	0.0%	823	100.0%				
Galactic Chi-Square Tests	q30Highestschoolingleve Ireceive * q4cvaever	823	100.0%	O	0.0%	823	100.0%				
Title	q31Typeofemploymenttha tyouhave * q4cvaever	823	100.0%	0	0.0%	823	100.0%				
Cni-Square Tests	q33mrelig * q4cvaever	823	100.0%	0	0.0%	823	100.0%				
								IBM SPSS Statistics Processor is ready H: 22, W: 995 pt.			

When the crosstabs procedure is complete, the results will be available to view in the output window. Ap-value less than or equal to 0.05 can be interpreted as a significant difference while p-values higher than 0.05 can be considered not significant.

#### **Chi-Square Output**

				q9Didyourchildre caps	eceiveavitaminA sule	
				0	1	Total
qviiiAreatype	1	Count		465	226	691
	% within qviiiAreatype			67.3%	32.7%	100.0%
	2	Count		126	53	179
		% within qvi	iiAreatype	70.4%	29.6%	100.0%
Total Count				591	279	870
		% within qvi	iiAreatype	67.9%	32.1%	100.0%
			Chi-Squ	are Tests <sup>c</sup>		
		1	Chi-Squ	are Tests <sup>c</sup> Asymp. Sig.	Exact Sig. (2-	Exact Sig.
		Value	Chi-Squ	<b>are Tests<sup>c</sup></b> Asymp. Sig. (2-sided)	Exact Sig. (2- sided)	Exact Sig. sided)
Pearson Chi-S	Square	Value .626ª	Chi-Squ df	are Tests <sup>c</sup> Asymp. Sig. (2-sided) .429	Exact Sig. (2- sided) .473	Exact Sig. sided)
Pearson Chi-S Continuity Cor	Square	Value .626 <sup>a</sup> .492	Chi-Squ df 1	are Tests <sup>c</sup> Asymp. Sig. (2-sided) .429 .483	Exact Sig. (2- sided) .473	Exact Sig. sided)
Pearson Chi-S Continuity Cor Likelihood Rat	Square rection <sup>e</sup> tio	Value .626ª .492 .633	Chi-Squ df 1 1	are Tests <sup>c</sup> Asymp. Sig. (2-sided) .429 .483 .426	Exact Sig. (2- sided) .473 .473	Exact Sig. sided)
Pearson Chi-S Continuity Cor Likelihood Rat Fisher's Exact	Square rection <sup>e</sup> tio Test	Value .626 <sup>a</sup> .492 .633	Chi-Squ df 1 1	are Tests <sup>c</sup> Asymp. Sig. (2-sided) .429 .483 .426	Exact Sig. (2- sided) .473 .473 .473	Exact Sig. sided)

c. For 2x2 crosstabulation, exact results are provided instead of Monte Carlo results.

Two tables will appear in the output window. Table 1 shows how the responses were distributed between those who were supplemented and those who were not. Table 2 shows the association between vitamin Asupplementation and each variable (In this case, area type). Below table two, you will see some notes about the data. If the first point (#3) indicates that there are 0 cells (0.0%) which have expected counts less than 5, then it is ok to use Pearson Chi-Square to determine the significance of the two variables. The significance level is indicated in the column titled Asymp.sig. In this example, since the Asymp. Sig value is more than 0.05, so there is no significant difference between groups.

Chi-Square Tests										
	-sided)									
			Asymp. Sig.		99% Confide	dence Interval				
	Value	df	(2-sided)	Sig.	Lower Bound	Upper Bound				
Pearson Chi-Square	5.961 <sup>a</sup>	2	.051	.032 <sup>b</sup>	.028	.037				
2 kelihood Ratio	6 0 9 4	2	047	035 <sup>b</sup>	.030	.039				
Fisher's Exact Test	5.793			.032 <sup>b</sup>	.028	.037				
N of Valid Cases	870									
1 (a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is .32.										
b. Based on 10000 sampled tables with starting seed 1487459085.										

If the notes below the chi-square table indicate that some cells have expected count of less than five (#1), then Fisher's exact (#2) should be used as the test statistic for these variables.

Acomparison of means can be done to determine if there are differences in numeric values, such as cost to attend the event, or distance to travel to the event. In order to determine which test to use, it first needs to be determined if the data are normally distributed.

#### 1 ect Marketing Utilities Graphs Add-ons Windor Analyze Reports Þ A-A Descriptive Statistics 2 Þ 123 Frequencies... Tables 🔚 Descriptives... Compare Means 21 Otł A Explore 3 General Linear Model 🌞 Crosstabs... Generalized Linear Models 🗾 Ratio... Mixed Models 🙍 P-P Plots... Correlate 🛃 Q-Q Plots... Regression Loglinear 0 Neural Networks Þ 0 Classify 0 **Dimension Reduction** Þ Scale 0 Nonparametric Tests Þ 0 Forecasting Þ 0 Survival 0 Multiple Response Þ 0 💕 Missing Value Analysis... 1 Multiple Imputation Þ 0 Complex Samples b 0 Quality Control Þ 0 🖉 ROC Cur<u>v</u>e...

Testing if data are normally distributed

Click on Analyze (#1), then select Descriptive Statistics (#2). Finally select Explore (#3).

°.	Explore		36	3
	Image: content of the second seco	Statistics Plo <u>t</u> s Options Bootstrap	Explore: Plots  Explore: Plots  Explore: Plots  Explore  Explore	r Pescriptive IStem-and-leaf Histogram Sts vene Test Natural log
			20	5

In the Explore window, select the continuous variable you want to test for normality from first column on the left (#1). Click the blue arrow (#2) to move it to the Dependent List box (#3). Select Plots (#4), to go to the plots window. Ensure that "Normality plots with tests" is selected. Click "Continue" (#5) and finally click "OK" (#6).

Output window		

#### Tests of Normality

	Kolm	nogorov-Smi		Shapiro-Wilk		
	Statistic	Statistic df Sig.			df	Sig. 1
q24Whatisyourage	.148	861	.000	.86	4 861	.000

a. Lilliefors Significance Correction

The Shapiro-Wilk test should be used to determine if the data are normally distributed. If the Sig. value (p-value) (#1) is < 0.05 the data are not normally distributed, while if the Sig. value is > 0.05 the data are normally distributed. Since we will be comparing two groups (supplemented vs. not supplemented) the two-sample t-test should be used for if the data are normally distributed (Sig.  $\geq$  0.05) and the Mann -Whitney test should be used if data are not normally distributed (Sig. < 0.05).

#### Wilcoxon-Man-Whitney Test



To perform the Wilcoxon-Mann-Whitney test, first click "Analyze" (#1) and select "Nonparametric Test" (#2) from the drop down menu. From the available nonparametric test options, select "Legacy Dialogs" (#3) and choose "2 Independent Samples..." (#4) as the final option.



When the Two Independent Samples Test window appears, ensure the Mann-Whitney U box in the Test Type area is checked (#1). Select from the list on the left (#2) the test variable(s) you want to analyze (example: cost of transport) and use the blue arrow next to the test variable list to add them to the list (#3). In the grouping variable box, add the variable that corresponds to your dependent variable (e.g. whether or not the child received vitamin A) (#4). Next select "Define Groups..." (#5). In the Two Independent Samples box, specify the values for the two groups you would like to compare (for example 0 and 1 for those not supplemented and those who were supplemented) (#6). Please note that when defining groups, only two groups are allowed and the values should be numeric. Once this is done select Continue and then "OK" (#7)

#### Mann-Whitney Test

Ranks								
	q4cvaever	Ν	Mean Rank	Sum of Ranks				
q20mmin	0	163	85.29	13903.00				
	1	7	90.29	632.00				
	Total	170						
q22mcost	0	146	76.79	11211.50				
	1	7	81.36	569.50				
	Total	153						

#### Test Statistics<sup>a</sup>

	q20mmin	q22mcost				
Mann-Whitney U	537.000	480.500				
Wilcoxon W	13903.000	11211.500				
Ζ	268	269				
Asymp. Sig. (2-tailed)	.789	.788				

The final data will be displayed in the output window. To determine if there is a significant difference between groups, take note of each Asymp Sig. (p-value).

#### Two-sample t-test (For normally distributed data)



To conduct the two-sample t-test, click "Analyze" on the drop down menu (#1) and select "Compare Means" (#2). From the extended menu options, select "Independent-Samples T Test..." (#3).



When the menu opens, select from the list on the left the test variables you want to analyze (#1) and use the blue arrow next to the test variable list to add them to the list (#2). In the grouping variable box, add the variable that corresponds to whether or not the child received vitamin A(#3). Next select "Define Groups..." (#4) and specify the two groups you would like to compare (example 0 and 1 for those not supplemented and those who were supplemented) (#5). Once this is done select "Continue" and finally "OK" (#6).

#### Output

Group Statistics								
g4cvaever N Mean Std. Deviation Mear								
q20mmin	0	163	20.00	19.738	1.546			
	1	7	20.00	15.275	5.774			
q22mcost	0	146	132.95	209.654	17.351			
	1	7	100.00	58.595	22.147			

	Independent Samples Test									
Levene's Test for Equality of Variances				t-test for Equality of Means						
							Mean	Std. Error	95% Confidenc Differ	e Interval of the ence
		F	Sig.	t	df	Sig. (2-tailed)	Difference	Difference	Lower	Upper
q20mmin	Equal variances assumed	.046	.831	.000	168	1.000	.000	7.564	-14.933	14.93
	Equal variances not assumed		$\bigcirc$	.000	6.890	1.000	.000	5.977	-14.179	14.17
q22mcost	Equal variances assumed	.711	.401	.414	151	.680	32.945	79.619	-124.367	190.25
	Equal variances not assumed			1.171	15.386	.259	32.945	28.134	-26.891	92.78

Once the t-test test is run, the final results will appear in the output window. Look at the "Equal Variances Assumed" row to determine if there are differences between groups. Asignificant value can be determined by looking in the Sig. column and observing if the Sig. value is < 0.05 (#1). In the example, there is no significant difference in time to get to a health facility between caretakers who attended the event and caretakers who didn't attend the event since the Sig. value is  $\geq 0.05$ .

Bivariate logistic regression is performed to identify significant relationships between chosen variables and a dependent variable (e.g. vitamin Asupplementation). In this analysis we will look at odds ratios and p-values to identify characteristics of participants who had significantly higher odds of receiving vitamin Asupplementation.



First click "Analyze" (#1) from the menu on the top of the screen. Select "Regression" (#2) from the drop down menu and finally "Binary Logisitc..." (#3).

Logistic Regression		×
Image: second secon	Dependent: A q4cvaever Block 1 of 1 Previous Covariates: q1Sexofthechild(Cat) 3 26mrel(Cat) q27Maritalstatus(Cat) q30Highestschoolingleve q31Typeofemploymentth Method: Forward: Wald Selection Variable: Paste Reset Cancel Help	Categorical Save Options

When the Logistic Regression menu appears, select the variable that corresponds to whether or not vitamin Awas received as the dependent variable. This is done by selecting the variable from the list on the left (#1) and clicking on the blue arrow button adjacent to the box labeled

"Dependent:" (#2). In the larger Covariates box select and add all of the variables that you want to conduct odds ratios for (#3). Change the method to Forward: Wald (#4). Forward Wald is a method that enables you to look at only the significant variables and their odds.

#### **Reference groups**



In the Logistic Regression window, click "Categorical" (#1). Anew window called Define Categorical Variables appears. There are two columns in this case, one being Covariates (#2) and the other being Categorical Covariates. Ensure that all categorical variables except for continuous variables (e.g. age, distance) are in the covariates box on the right side. Variables can be moved to the Categorical Covariates box using the blue arrow (#3). Below the Categorical Covariates box, there is a section called Change Contrast (#4). Change the reference category to First and press "Change." By doing this, the odds ratios will be calculated in reference to the first response to each question. (Last is not used for PEC surveys since in many cases, the last response is "Other", which makes it difficult to interpret the results.). Click "Continue" (#5) to go back to logistic regression window. \*Note: Ensure that the categorical variables that you select have their Variable Type listed as numeric in the Variable View tab, as described in Section 1.2.

#### **Calculate 95% Confidence Intervals**

-	Logistic Regression	8	
	a q0viiitype     q0ixDateofintervi     Q0ixDateofintervi     q0xiNoofchildren     q1Sexofthechild     q1bHavebirthcerti     q2lndicatedateon     q2adobeling     q2cverifybycard     q3meas     q3bval2     q3cmdaymmomyr     q3dnmeas     q3dnmeas     q5vadoses     q6ctaken	Dependent: Categorical Save Block 1 of 1 Covariates: Q1Sexothechild(Cat) Q26mrel(Cat) Q30Highestschoolingleve Q31Typeofemploymentth Method: Forward: Wald Selection Variable: Rule	Statistics and Plots         Classification plots         Losmer-Lemeshow goodness-of-fit         Iteration history         Casewise listing of residuals         Cl for exp(B):         95         %         Outliers outside         2         std. dev.         3         Classification cutoff:         0.5         Entry:       0.05         Removal:       0.10         4       Maximum Iterations:         20       Conserve memory for complex analyses or large datasets
	3118 3119	2 24 31 2 24 31	Continue     Cancel Help

In the Logistic Regression window, click "Options" (#1). In the first section Statistics and Plots, check Hosmer-Lemeshow goodness of fit (#2) and CI for exp(B) (#3). In the third section called Probability for Stepwise (#4) you will define your required entry and removal criteria for independent variables. By default, the entry level is 0.05 while the removal level is 0.10. This means any independent variable whose significance level is more than 0.10 will be excluded from the model. Click "Continue" (#5) to go back to the Logistic Regression window. Flnally select "OK" (#6) and the logistic regression analysis results will appear in the output window.

#### Interpreting the output

÷							
		В	S.E.	Sig.	Exp(B)	95% C I for EXP(B)	
					3	Lower	Upper
	q35DoesyourchildattendtheEC	542	.179	.002	.581	.409	.826
Step 1ª	DEarly(1)						
	Constant	624	.082	.000	.536		
	q346Districthospital(1)	518	.180	.004	.596	.419	.848
Step 2 <sup>b</sup>	q35DoesyourchildattendtheEC	543	. 180	.003	.581	.409	.826
	DEarly(1)						
	Constant	673	.113	.000	.510		
	q346Districthospital(1)	545	.182	.003	.580	.406	.828
	q35DoesyourchildattendtheEC	447	.184	.015	.640	.446	.917
	DEarly(1)						
•	q361Routinevisitscheckupsgro	.367	.157	.019	1.443	1.062	1.962
	wth(1)						
Step 6	q3702Postersomewhereothert	-1.293	.626	.039	.274	.080	.936
	hanhealth(1)						
	q3713Religiousleaders(1)	688	.341	.043	.503	.258	.980
	q3714Communityleaders(1)	.459	.207	.027	1.583	1.055	2.374
	Constant	649	.114	.000	.523		

#### Variables in the Equation

As described above, using Forward Wald ensures that only variables that are significantly associated with the dependent variable are included in the model. It starts by inserting one variable into the model (Step 1) and determining if it is significantly associated with the dependent variable. If so, a second variable is added to the model and the cumulative effect is looked at. There will be as many steps as there are significant associated with the dependent variable) in the final model. These are the odds ratios and p-values that should be reported. To determine a variable's significance level, look at the significance level (#2, Sig. column). Then, proceed to look at the Exp(B) column (#3) which in this case gives the odds of a child being supplemented when compared to the reference group (not being supplemented).