

Introduction to SPSS

1

2nd Class
Chemistry Dept.

Instructor :- Yahiea Al-Naiemy

What is in this workshop

2

- SPSS interface: data view and variable view
- How to enter data in SPSS
- How to import external data into SPSS
- How to clean and edit data
- How to transform variables
- How to sort and select cases
- How to get descriptive statistics

SPSS interface

3

- **Data view**
 - The place to enter data
 - Columns: variables
 - Rows: records
- **Variable view**
 - The place to enter variables
 - List of all variables
 - Characteristics of all variables

Before the data entry

4

- You need a code book/scoring guide
- If you use online survey, you need something to identify your cases.
- You also can use Excel to do data entry.

Example of a code book

5

2009 Youth Risk Behavior Survey Code Book sample.pdf - Adobe Reader

File Edit View Document Tools Window Help

1 / 1 173% Find

2009 Youth Risk Behavior Survey Code Book

Q1. How old are you?

Response options: 1 = 12 years old or younger; 2 = 13 years old; 3 = 14 years old; 4 = 15 years old; 5 = 16 years old; 6 = 17 years old; 7 = 18 years old or older

Q2. What is your sex?

Response options: 1= Female; 2 = Male

Q20. During the past 12 months, did your boyfriend or girlfriend ever hit, slap, or physically hurt you on purpose?

Response options: 1 = Yes; 2 = No

Recoded variable name: Q20r

Response options: 1 = Yes; 0 = No

A code book is about how you code your variables. What are in code book?

- 1. Variable names**
- 2. Values for each response option**
- 3. How to recode variables**

Enter data in SPSS 19.0

6

Untitled2 [DataSet3] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

Visible: 0 of 0 Variables

Columns: variables

Rows: cases

Under Data View

Data View Variable View

IBM SPSS Statistics Processor is ready

Enter variables

7

1. Click this Window

2. Type variable name

3. Type: numeric or string

4. Description of variable

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Q01	Numeric	8	0	Age	None	None	8	Right	Ordinal	Input
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											
32											
33											
34											
35											
36											
37											
38											

IBM SPSS Statistics Processor is ready

1. Click Variable View
2. Type variable name under Name column (e.g. Q01).
NOTE: Variable name can be 64 bytes long, and the first character must be a letter or one of the characters @, #, or \$.
3. Type: Numeric, string, etc.
4. Label: description of variables.

Enter variables

8

Based on your code book!

The screenshot shows the IBM SPSS Statistics Data Editor window with the 'Value Labels' dialog box open. The dialog box is titled 'Value Labels' and has a 'Value Labels' tab selected. The 'Value' field is empty, and the 'Label' field is empty. The 'Add' button is visible. The background shows the Data Editor window with the 'Age' variable selected. A red cloud contains the text 'Based on your code book!' with arrows pointing to the 'Value Labels' dialog box and the 'Age' variable in the Data Editor.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns
1	Q01	Numeric	8	0	Age	None	None	8
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								

Value Labels

Value:

Label:

Add

Spelling...

Value Labels

Value:

Label:

1 = "12 years old or young"

2 = "13 years"

Add

Change

Remove

Spelling...

OK Cancel Help

Enter cases

9

The screenshot shows the IBM SPSS Statistics Data Editor window. The title bar reads "Untitled2 [DataSet3] - IBM SPSS Statistics Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations, data manipulation, and analysis. The main data grid has 20 columns and 36 rows. The first two columns are labeled "Code" and "Q01". The first five rows contain data: (1001, 1), (1002, 2), (1003, 2), (1004, 3), and (1005, 2). The remaining columns are labeled "var". The status bar at the bottom indicates "Visible: 2 of 2 Variables".

8: Visible: 2 of 2 Variables

	Code	Q01	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1001	1																	
2	1002	2																	
3	1003	2																	
4	1004	3																	
5	1005	2																	
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22																			
23																			
24																			
25																			
26																			
27																			
28																			
29																			
30																			
31																			
32																			
33																			
34																			
35																			
36																			

Under Data View

1. Two variables in the data set.
2. They are: Code and Q01.
3. Code is an ID variable, used to identify individual case (NOT people's real IDs).
4. Q01 is about participants' ages: 1 = 12 years or younger, 2 = 13 years, 3 = 14 years...

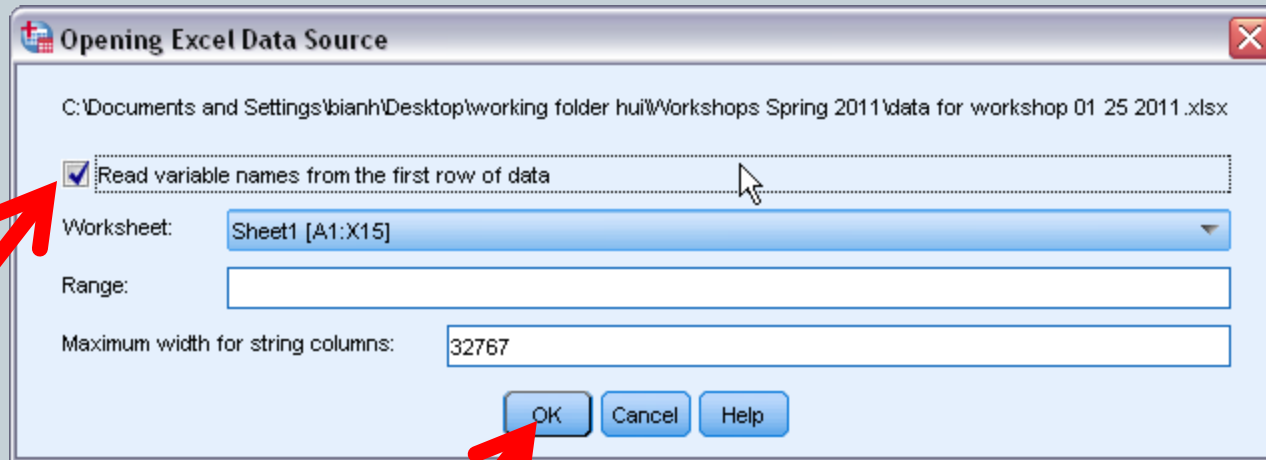
Data View Variable View

IBM SPSS Statistics Processor is ready

Import data from Excel

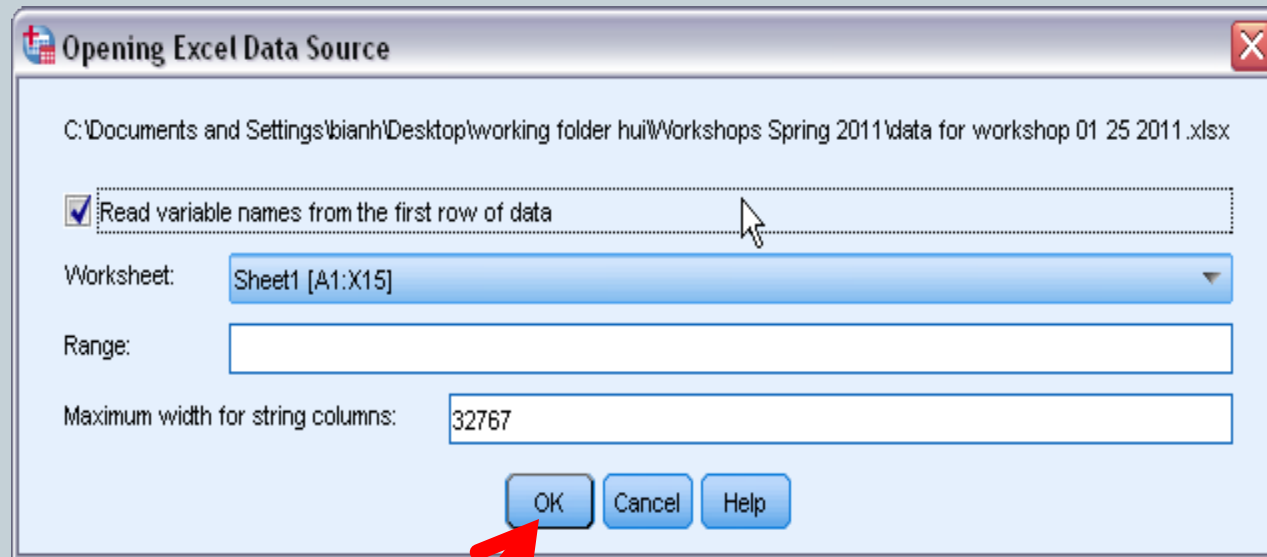
10

- Select File → Open → Data
- Choose **Excel** as file type
- Select the file you want to import
- Then click Open



Open Excel files in SPSS

11



Import data from CVS file

12

- CVS is a comma-separated values file.
- If you use Qualtrics to collect data (online survey), you will get a CVS data file.
- Select File → Open → Data
- Choose **All files** as file type
- Select the file you want to import
- Then click Open

Continue

13

Text Import Wizard - Step 1 of 6

Welcome to the text import wizard!

This wizard will help you read data from your text file and specify information about the variables.

Does your text file match a predefined format?

☐ Yes ☒ No

Text file: C:\Documents and Settings\bianh\Desktop\working folder hui\Workshops Spring 2011\CSV data

	var1	var2	var3	var4
1				
2				
3				
4				

0 10 20 30 40 50 60

1	Q1,Q2,Q3,Q4,Q5,Q6,Q7,Q8,,
2	4,2,3,1,1,1,3,3,,
3	2,2,2,1,1,1,3,5,,
4	2,2,3,1,1,1,3,5,,
5	2,2,2,1,1,1,3,4,,

< Back Next > Finish Cancel Help

14

Text Import Wizard - Step 2 of 6

How are your variables arranged?

- ☒ **D**elimited - Variables are delimited by a specific character (i.e., comma, tab).
- ☐ **F**ixed width - Variables are aligned in fixed width columns.

Are variable names included at the top of your file?

- ☒ **Y**es
- ☐ **N**o

Text file: C:\Documents and Settings\bianh\Desktop\working folder hui\Workshops Spring 2011\CSV dat

	0	10	20	30	40	50	60
1	4	2	3	1	1	1	3
2	2	2	2	1	1	1	3
3	2	2	3	1	1	1	3
4	2	2	2	1	1	1	3
5	1	2	1	1	1	1	5

< Back Next > Finish Cancel Help

15

Text Import Wizard - Delimited Step 3 of 6

The first case of data begins on which line number?

How are your cases represented?

☒ Each line represents a case

☐ A specific number of variables represents a case:

How many cases do you want to import?

☒ All of the cases

☐ The first cases.

☐ A random percentage of the cases (approximate): %


Data preview

	0	10	20	30	40	50	60
1	4,2,3,1,1,1,3,3,						
2	2,2,2,1,1,1,3,5,						
3	2,2,3,1,1,1,3,5,						

< Back Next > Finish Cancel Help

Continue

16



Text Import Wizard - Delimited Step 4 of 6

Which delimiters appear between variables?

☐ Tab ☐ Space

☒ **Comma** ☐ Semicolon

☐ Other:

What is the text qualifier?

☒ None

☐ Single quote


☐ Double quote

☐ Other:

Data preview

Q1	Q2	Q3	Q4	Q5	Q6	Q7
1	2	1	2	1	1	1
1	2	1	1	1	1	1
1	2	1	1	1	1	1
2	2	1	1	1	1	1
4	1	2	1	1		3
1	2	1	1	1		1
3	2	5	1	1		3
2	2	2	1	1		1

< Back Next > Finish Cancel Help



Continue

17

Text Import Wizard - Step 5 of 6

Specifications for variable(s) selected in the data preview

Variable name: Original Name: Q1

Data format:

Data preview

Q1	Q2	Q3	Q4	Q5	Q6	Q7
4	2	3	1	1	1	3
2	2	2	1	1	1	3
2	2	3	1	1	1	3
2	2	2	1	1	1	3
1	2	1	1	1	1	1

< Back Next > Finish Cancel Help

Continue

18

Text Import Wizard - Step 6 of 6

You have successfully defined the format of your text file.

Would you like to save this file format for future use?

☐ Yes ☒ No

Save As...

Would you like to paste the syntax?

☐ Yes ☒ No

☒ Cache data locally

Press the Finish button to complete the text import wizard.

Data preview

Q1	Q2	Q3	Q4	Q5	Q6	Q7
4	2	3	1	1	1	3
2	2	2	1	1	1	3
2	2	3	1	1	1	3
2	2	2	1	1	1	3
1	2	1	1	1	1	1
1	2	1	2	1	1	1

< Back Next > Finish Cancel Help

Continue

19

SPSS Data Editor window showing a dataset with 37 rows and 18 columns. The columns are labeled Q1 through Q8, followed by 10 unlabeled variable columns. The data is organized into a grid with a blue header row and a blue footer row. A red starburst graphic is overlaid on the right side of the data grid, containing the text "Save this file as SPSS data".

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	var	var	var	var	var	var	var	var	var	var
1	4	Female	3	1	1	1	3	3										
2	2	Female	2	1	1	1	3	5										
3	2	Female	3	1	1	1	3	5										
4	2	Male	2	1	1	1	3	4										
5	1	Male	1	1	1	1	1	5										
6	1	Female	1	2	1	1	1	1										
7	1	Male	1	2	1	1	1	5										
8	1	Female	1	1	1	1	1	3										
9	1	Female	1	1	1	1	1	5										
10	2	Female	1	1	1	1	1	1										
11	4	Female	2	1	1	1	3	5										
12	1	Male	1	1	1	1	1	3										
13	3	Male	5	1	1	1	3	5										
14	2	Male	2	1	1	1	1	5										
15	1	Male	1	2	1	1	1	1										
16	1	Female	1	1	3	1	1	4										
17	1	Male	1	1	1	1	3	3										
18	1	Female	1	1	1	1	1	3										
19	1	Female	1	1	1	1	1	3										
20	1	Male	2	1	1	1	3	1										
21	1	Female	1	1	1	1	3	5										
22	1	Female	1	1	1	1	1	4										
23	3	Female	2	1	1	1	3	5										
24	1	Male	1	1	1	1	3	1										
25	1	Male	1	1	1	1	1	3										
26	1	Female	1	2	1	1	1	3										
27	1	Male	1	3	1	1	1	5										
28	3	Female	2	1	3	1	3	5										
29	5	Male	2	1	2	1	3	5										
30	4	Male	4	1	1	1	3	1										
31	1	Male	1	1	1	1	1	5										
32	1	Female	1	1	1	1	1	3										
33	1	Male	1	1	1	1	1	4										
34	1	Female	1	3	1	1	1	2										
35	1	Male	1	1	1	1	1	5										
36	2	Male	2	1	1	1	3	2										
37	2	Male	2	1	1	1	3	2										

Save this file as SPSS data

Clean data after import data files

20

- Key in values and labels for each variable
- Run frequency for each variable
- Check outputs to see if you have variables with wrong values.
- Check missing values and physical surveys if you use paper surveys, and make sure they are real missing.
- Sometimes, you need to recode string variables into numeric variables

Continue

(21)

Q1 How old are you

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 12 years old or younger	19	.1	.1	.1
	2 13 years old	17	.1	.1	.2
	3 14 years old	1851	11.3	11.3	11.5
	4 15 years old	4045	24.7	24.7	36.3
	5 16 years old	4234	25.8	25.9	62.2
	6 17 years old	3963	24.2	24.2	86.4
	7 18 years old or older	2215	13.5	13.6	100.0
	8	1	.0	.0	100.0
	9	1	.0	.0	100.0
	Total	16344	99.6	100.0	
Missing	System	66	.4		
Total		16410	100.0		

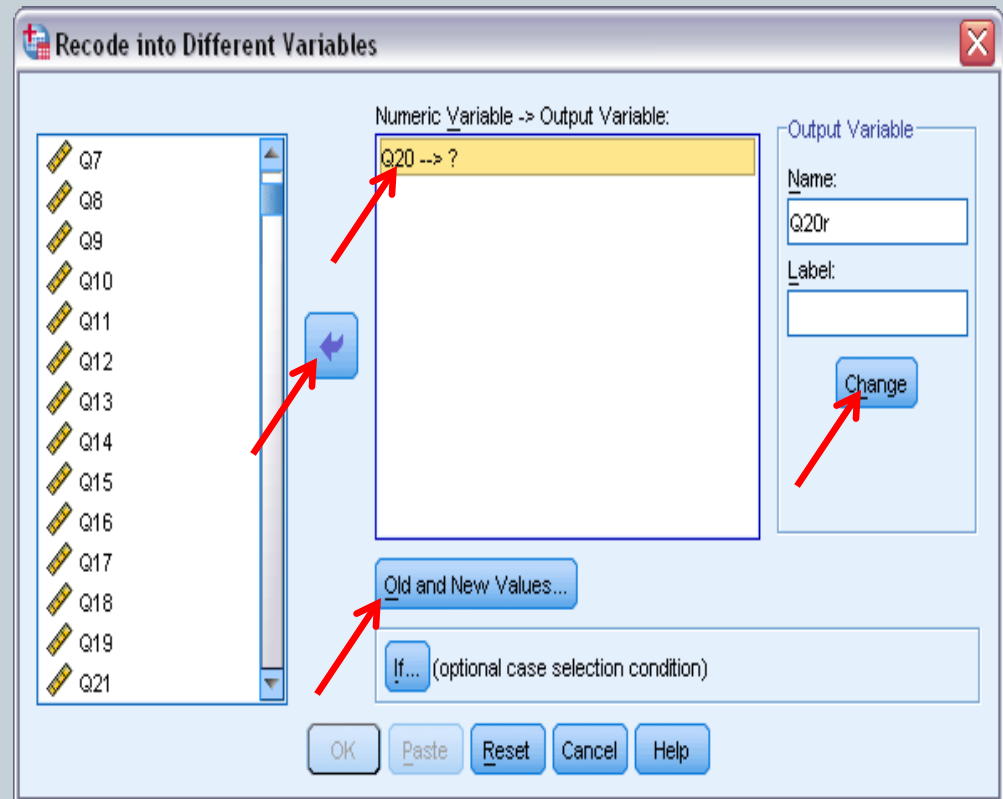
Wrong
entries

Variable transformation

22

- **Recode variables**

1. Select Transform → Recode into Different Variables
2. Select variable that you want to transform (e.g. Q20): we want 1= Yes and 0 = No
3. Click Arrow button to put your variable into the right window
4. Under Output Variable: type name for new variable and label, then click Change
5. Click Old and New Values



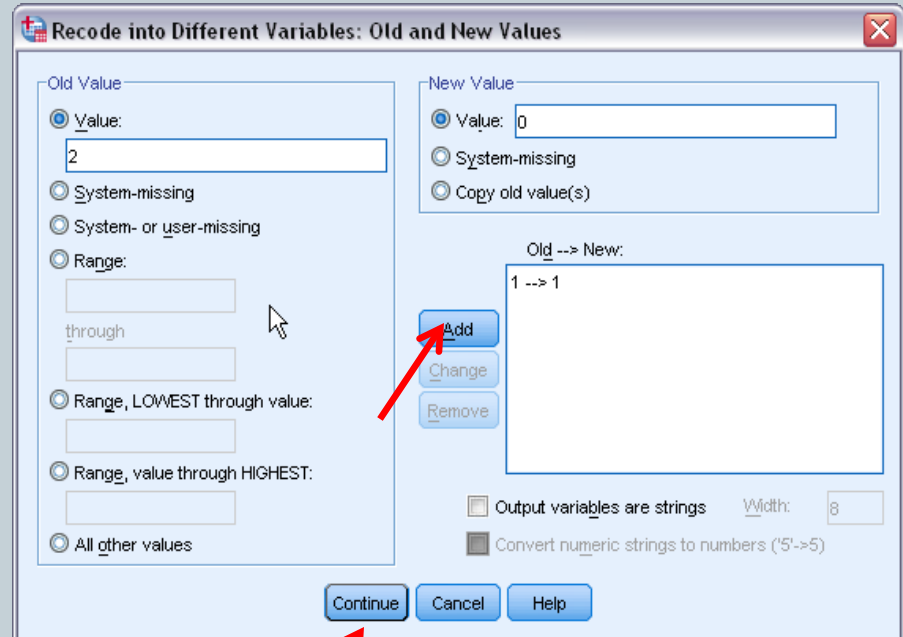
Continue

23

6. Type 1 under Old Value and 1 under New Value, click Add. Then type 2 under Old Value, and 0 under New Value, click Add.


7. Click Continue after finish all the changes.

8. Click Ok



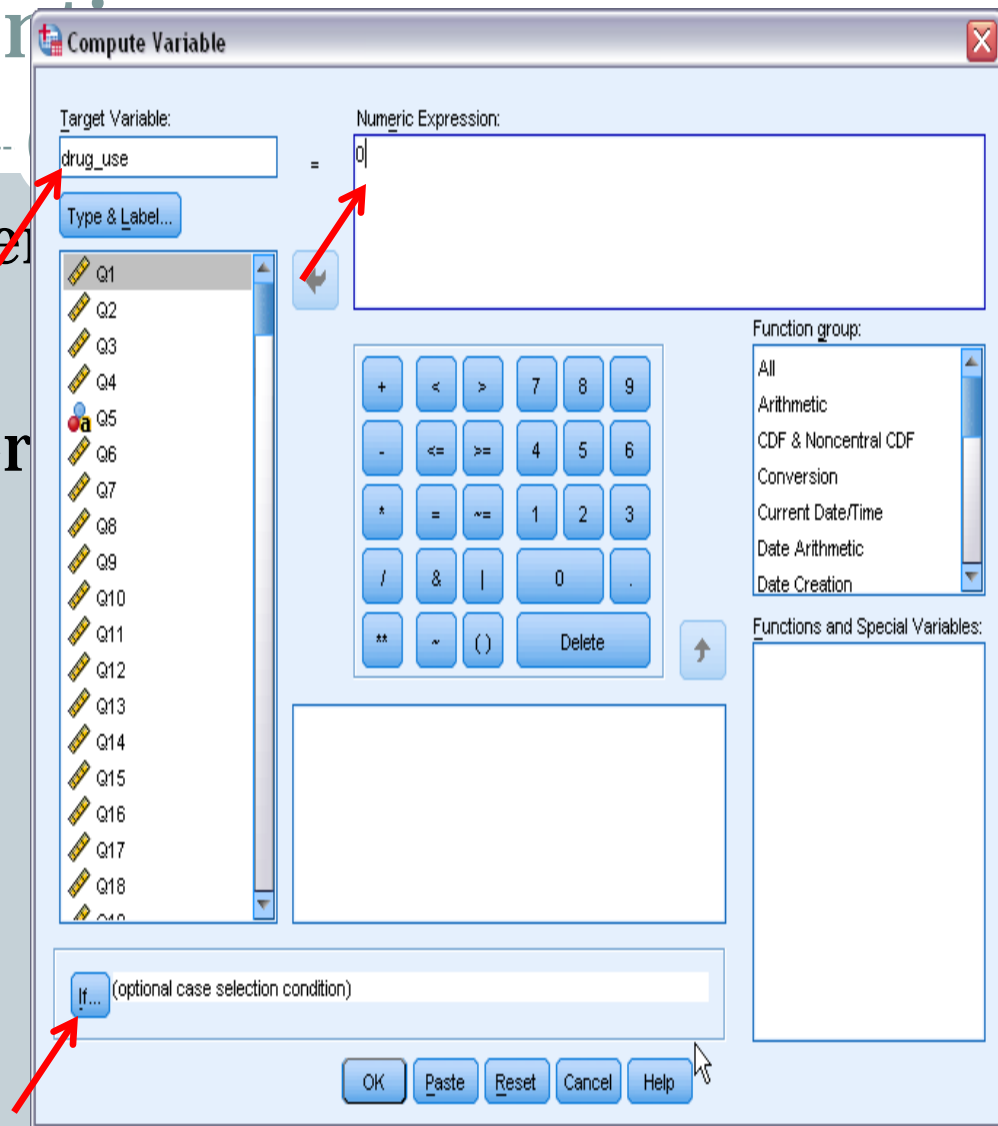
Variable transformation

24

- Compute variable (use YRBSS 2009 data)
 - Example 1. Create a new variable: drug_use (During the past 30 days, any use of cigarettes, alcohol, and marijuana is defined as use, else as non-use). There are two categories for the new variable (use vs. non-use). Coding: 1= Use and 0 = Non-use
 1. Use Q30, Q41, and Q47 from 2009 YRBSS survey
 2. Non-users means those who answered **0** days/times to all three questions.
 3. Go to Transform  Compute Variable

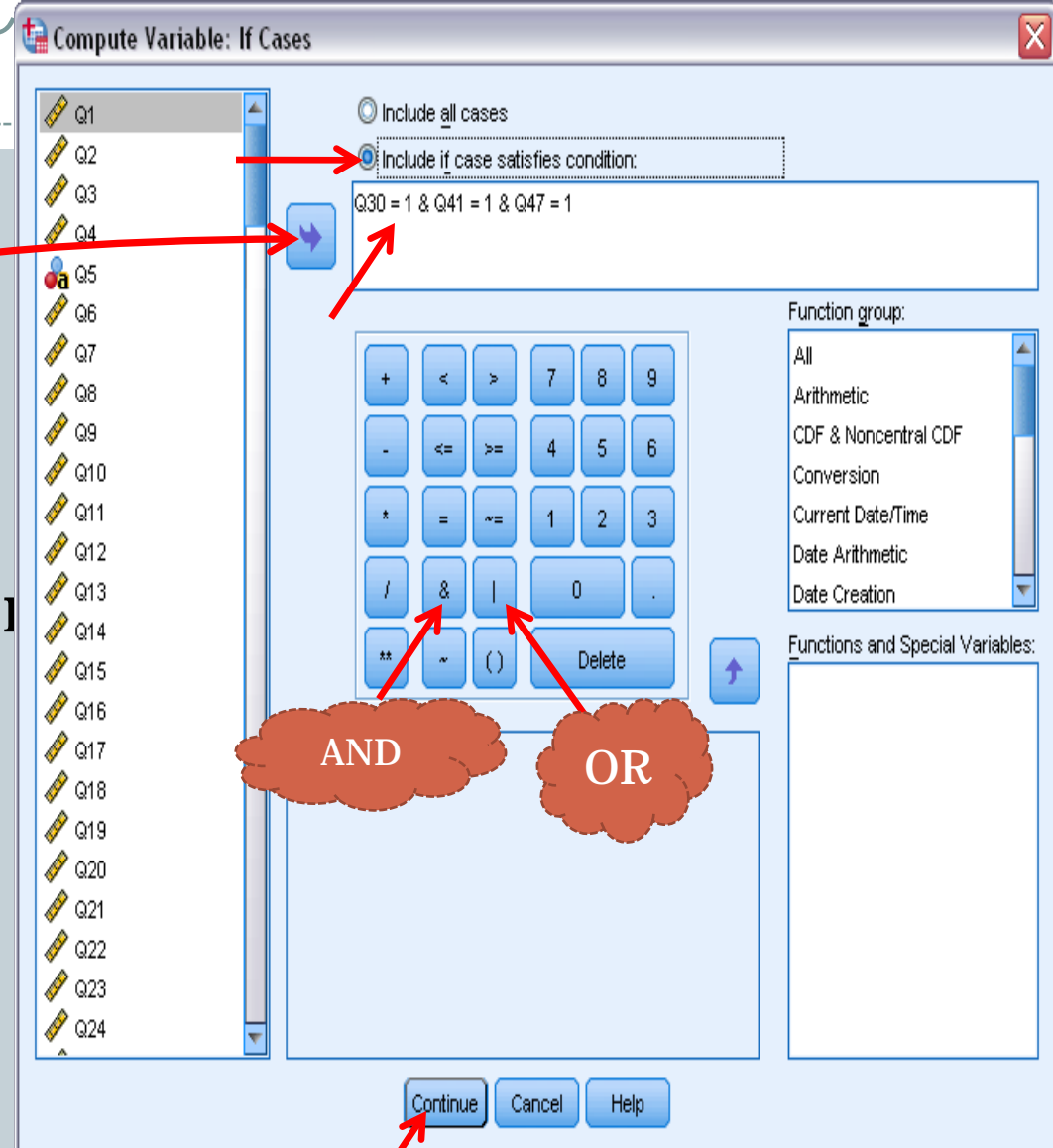
Continue

4. Type “drug_use” under Target Variable
5. Type “0” under Numeric Expression. 0 means Non-use
6. Click If button.

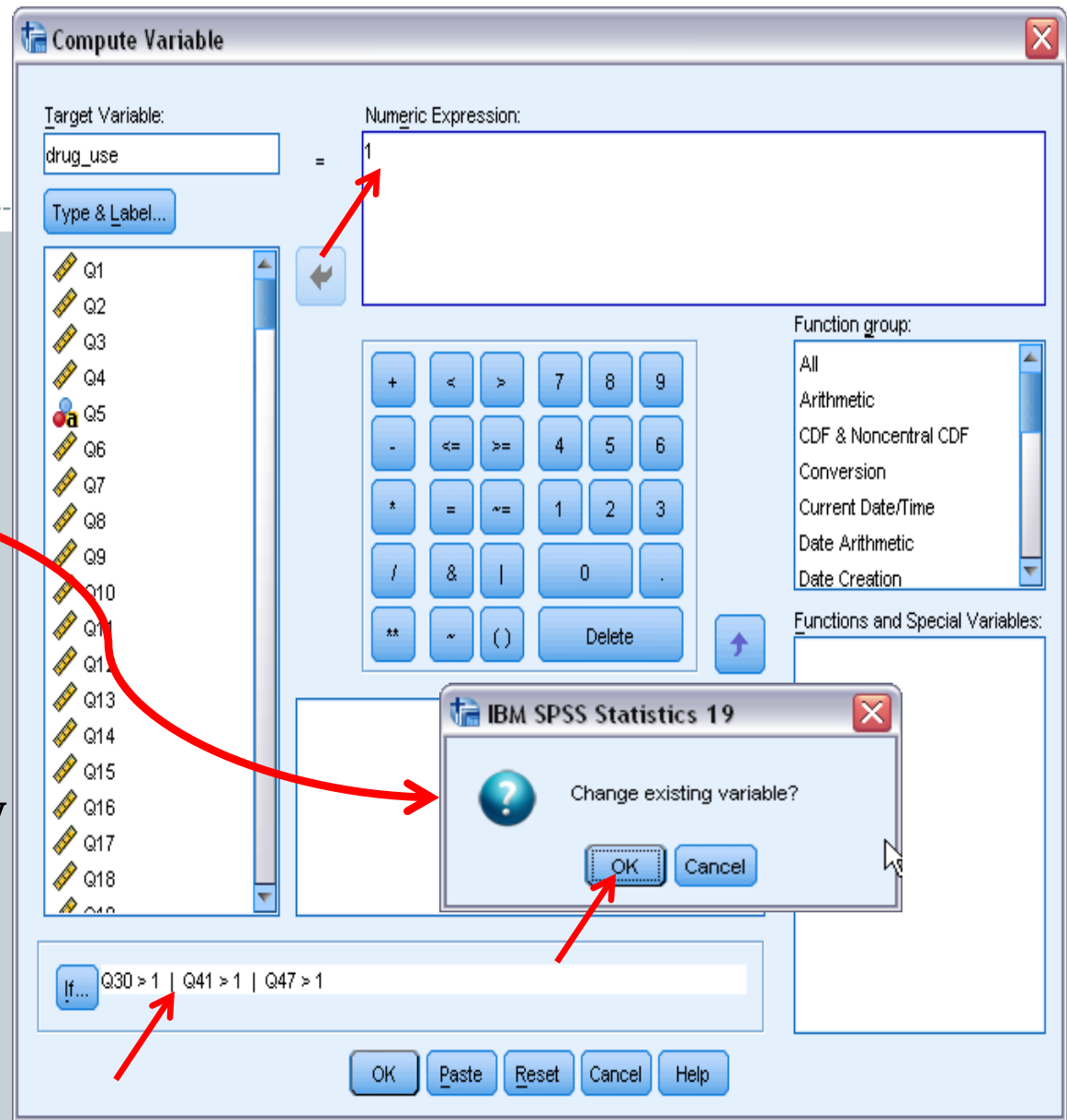


7. With help of that
Arrow button, type
 $Q30 = 1 \& Q41 = 1 \& Q47 = 1$
then click Continue
8. Do the same thing for
Use, but the numeric
expression is different:
 $Q30 > 1 \mid Q41 > 1 \mid Q47 > 1$

Continue



9. Click OK
10. After click OK, a small window asks if you want to change existing variable because drug_use was already created when you first define non-use.
11. Click ok.



Continue

28

- **Compute variables**

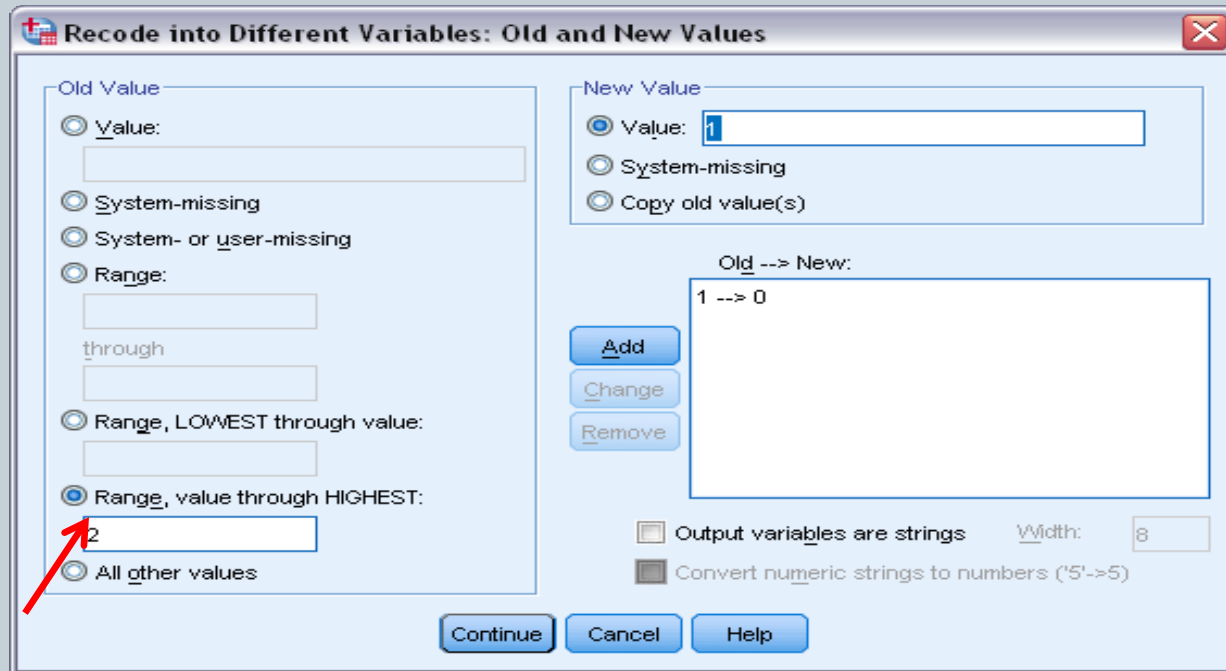
- **Example 2. Create a new variable drug_N that assesses total number of drugs that adolescents used during the last 30 days.**
 1. Use Q30 (cigarettes), 41 (alcohol), 47 (marijuana), and 50 (cocaine). The number of drugs used should be between 0 and 4.
 2. First, recode all four variables into two categories: 0 = non-use (0 days), 1 = use (at least 1 day/time)
 3. Four variables have 6 or 7 categories

Continue

29

4. Recode four variables: 1 (old) = 0 (new), 2-6/7 (old) = 1 (New).

5. Then select Transform → Compute Variable

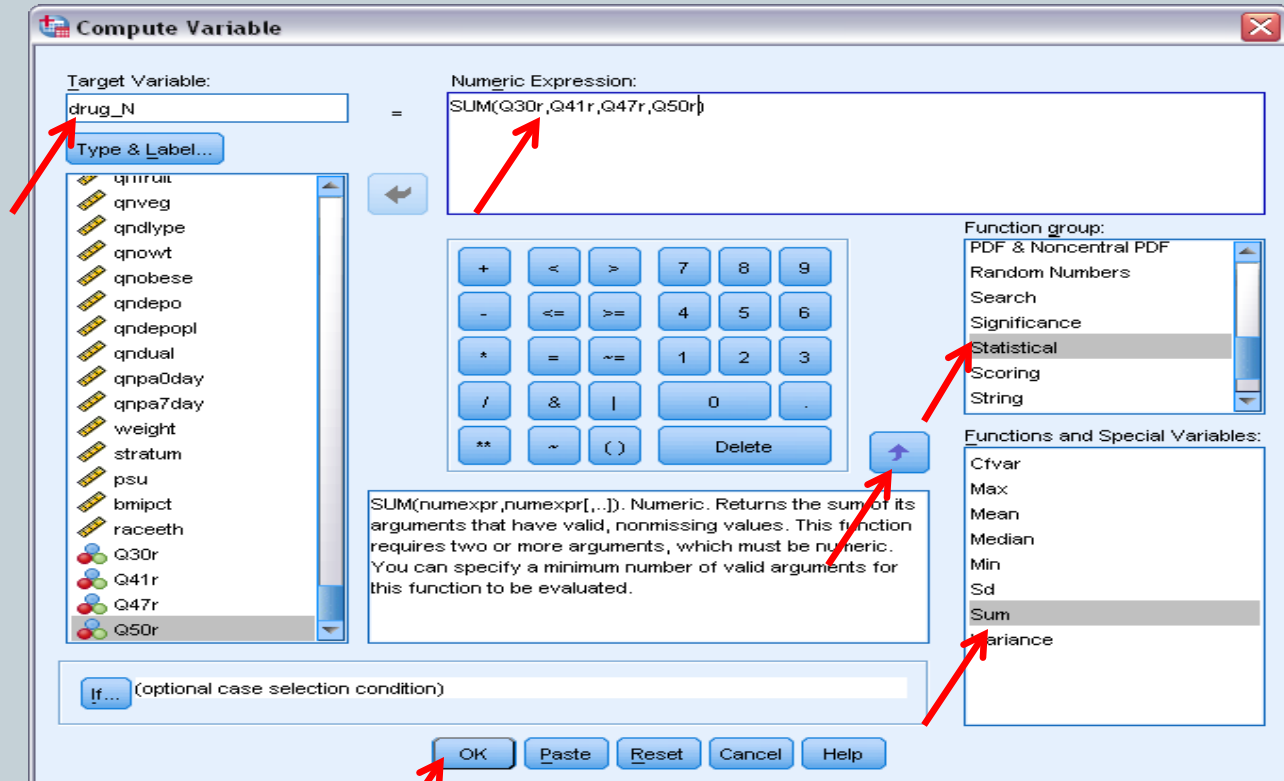


The image shows the 'Recode into Different Variables: Old and New Values' dialog box in SPSS. The 'Old Value' section on the left has the 'Range, value through HIGHEST:' option selected, with the value '2' entered in the adjacent text box. A red arrow points to this '2'. The 'New Value' section on the right has the 'Value:' option selected, with the value '1' entered in the adjacent text box. Below these sections is a list of mappings under the heading 'Old --> New:', which currently contains '1 --> 0'. To the right of this list are buttons for 'Add', 'Change', and 'Remove'. At the bottom of the dialog are 'Continue', 'Cancel', and 'Help' buttons. There are also checkboxes for 'Output variables are strings' (unchecked) and 'Convert numeric strings to numbers ('5' -> 5)' (checked), along with a 'Width' field set to '8'.

Continue

30

6. Type drug_N under Target Variable
7. Numeric Expression: SUM (Q30r,Q41r,Q47r,Q50r)
8. Click OK



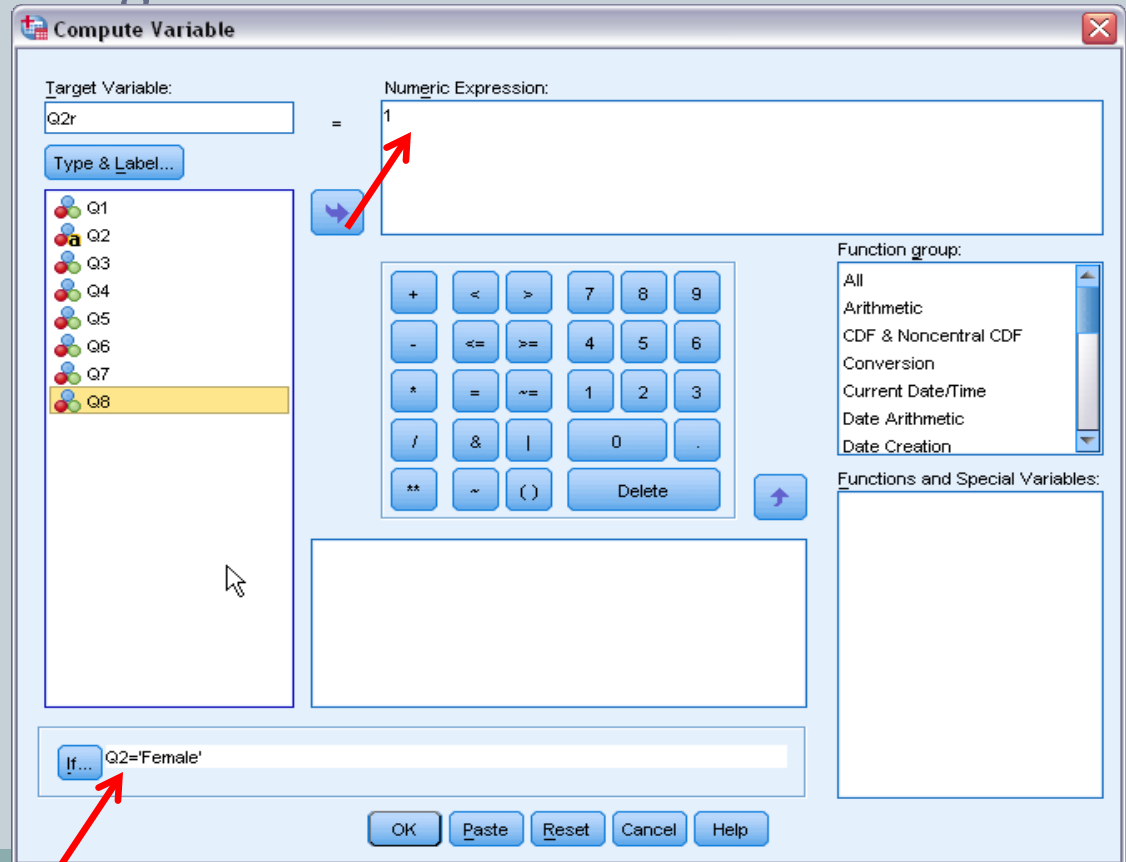
Continue

31

- **Compute variables**

- **Example 3: Convert string variable into numeric variable**

1. Enter 1 at Numeric Expression.
2. Click If button and type Q2 = 'Female'
3. Then click Ok.
4. Enter 2 at Numeric Expression.
5. Click If button and type Q2 = 'Male'
6. Then click Ok



Sort and select cases

32

- Sort cases by variables: Data → Sort Cases
- You can use Sort Cases to find missing.

The screenshot displays the IBM SPSS Statistics Data Editor window. The dataset is named 'yrbs09.sav [DataSet1]'. The variable list at the top includes Q1 through Q28. The data grid shows 37 rows of data. A red arrow points to the 'Data' menu in the top bar. Another red arrow points to the 'Sort Cases' dialog box, which is open. In the 'Sort by' list, 'Q1 (A)' is selected. The 'Sort Order' is set to 'Ascending'. The 'OK' button is highlighted.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28
1	.	2	2	5	1	1	1	1	1	1	2	1	1	1	2	2	1	2	2	2	1	1	1
2	.	2	2	2	5	2	3	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
3	.	2	4	1	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
4	2	4	5	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
5	.	2	1
6	.	1	4	2	C	.	.	1	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
7	.	2	4	1	.	.	.	1	5	2	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
8	2	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
9	.	2	1	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
10	2	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
11	2	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
12	.	1	4	2	.	.	.	2	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
13	.	1	3	2	E	.	.	6	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
14	2	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
15	.	2	1	3	1	1	1	5	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
16	1	2	3	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
17	.	1	1	2	2	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
18	1	4	3	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
19	.	2	2	1	1	2	2	1	2	2	2	1	1	1
20	1	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
21	2	3	5	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
22	1	3	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
23	2	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
24	.	1	1	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
25	2	2	5	5	5	5	5	1	1	1	1	1	2	2	1	2	2	2	1	1	1
26	.	1	2	2	C	.	.	1	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
27	.	1	3	2	C	.	.	2	5	5	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
28	2	5	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
29	.	1	4	5	2	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
30	.	1	2	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
31	.	2	2	5	1	1	3	1	1	2	1	1	1	1	2	2	1	2	2	2	1	1	1
32	2	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
33	1	3	4	3	4	3	4	2	4	.	.	2	2	1	2	2	2	1	1	1
34	.	1	2	4	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
35	.	2	4	.	C	.	.	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	1	1	1
36	.	1	1	5	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1
37	4	5	4	4	4	4	4	4	4	4	4	4	2	2	1	2	2	2	1	1	1

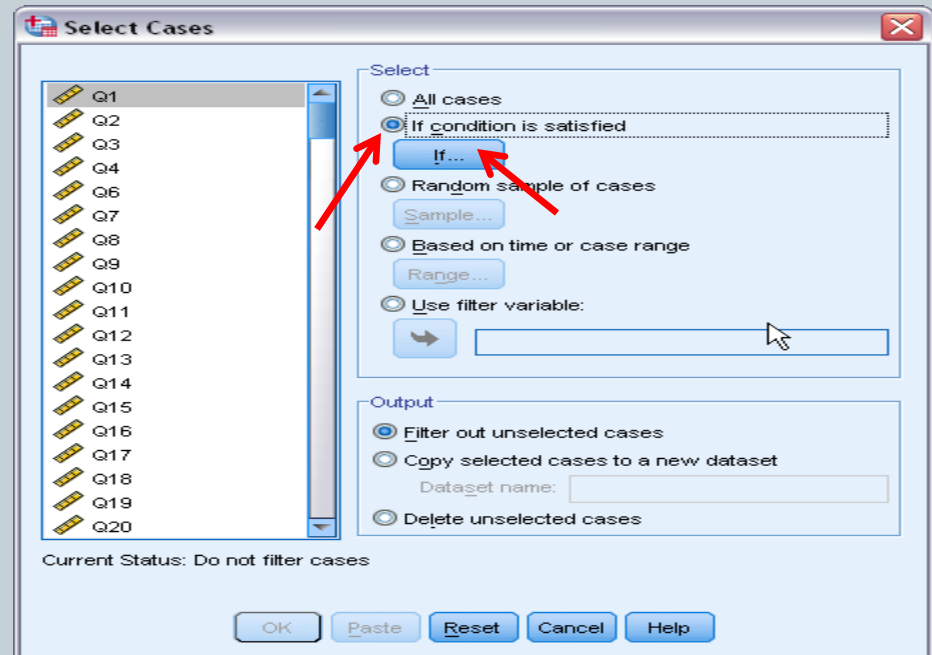
Sort and select cases

33

- **Select cases**

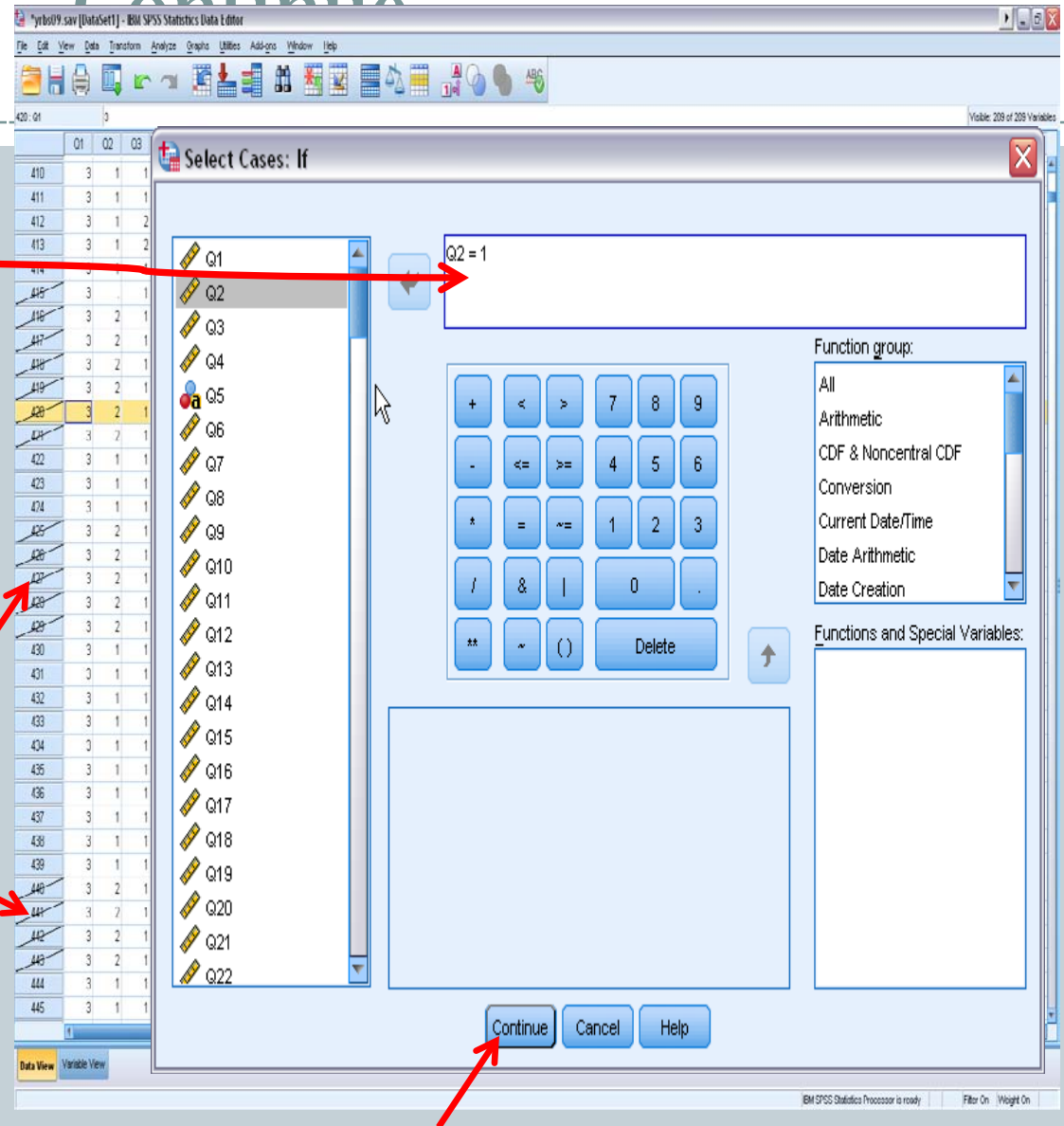
- Example 1. Select Females for analysis.

1. Go to Data —→ Select Cases
2. Under Select: check the second one
3. Click If button



4. Q2 (gender) = 1,
1 means Female
5. Click Continue
6. Click Ok

Unselecte
d cases :
 $Q2 = 2$



Sort and select cases

35

7. You will see a new variable: filter_\$ (Variable view)

The screenshot shows the IBM SPSS Statistics Data Editor window with the Variable View tab selected. The list of variables includes 'filter_\$' at row 209, which is highlighted. A red arrow points to this variable. Another red arrow points to the 'Variable View' tab at the bottom left of the window.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
188	QIN97	Numeric	1	0	Get 8+ hour...	None	-179769313...	6	Right	Scale	Input
189	QIN98	Numeric	1	0	Grades mo...	None	-179769313...	6	Right	Scale	Input
190	site	String	3	0	Site Code	None	-179769313...	6	Left	Nominal	Input
191	qnfricg	Numeric	1	0	Smoked on...	None	-179769313...	9	Right	Scale	Input
192	qnanytob	Numeric	1	0	Used any t...	None	-179769313...	10	Right	Scale	Input
193	qnfrvg	Numeric	1	0	Ate 5+ fruit...	None	-179769313...	8	Right	Scale	Input
194	qnfruit	Numeric	1	0	Ate 2+ fruit...	None	-179769313...	9	Right	Scale	Input
195	qnveg	Numeric	1	0	Ate 3+ vege...	None	-179769313...	7	Right	Scale	Input
196	qndtype	Numeric	1	0	Attended P...	None	-179769313...	9	Right	Scale	Input
197	qnowt	Numeric	1	0	Overweight	None	-179769313...	7	Right	Scale	Input
198	qnobese	Numeric	1	0	Obese	None	-179769313...	9	Right	Scale	Input
199	qndepo	Numeric	1	0	Of current s...	None	-179769313...	8	Right	Scale	Input
200	qndepopl	Numeric	1	0	Of current s...	None	-179769313...	10	Right	Scale	Input
201	qndual	Numeric	1	0	Of current s...	None	-179769313...	8	Right	Scale	Input
202	qnpa0day	Numeric	1	0	Active for 6...	None	-179769313...	10	Right	Scale	Input
203	qnpa7day	Numeric	1	0	Active for 6...	None	-179769313...	10	Right	Scale	Input
204	weight	Numeric	10	0	Weight	None	-179769313...	12	Right	Scale	Input
205	stratum	Numeric	3	0	Stratum	None	-179769313...	9	Right	Scale	Input
206	psu	Numeric	7	0	Primary Sa...	None	-179769313...	9	Right	Scale	Input
207	bmipct	Numeric	5	0	Body Mass...	None	-179769313...	8	Right	Scale	Input
208	raceeth	Numeric	5	0	Race/Ethni...	{1, Am Indi...	-179769313...	9	Right	Scale	Input
209	filter_\$	Numeric	1	0	Q2 = 1 (FIL...	{0, Not Sele...	None	10	Right	Nominal	Input
210											
211											
212											
213											
214											
215											
216											
217											
218											
219											
220											
221											
222											
223											
224											
225											

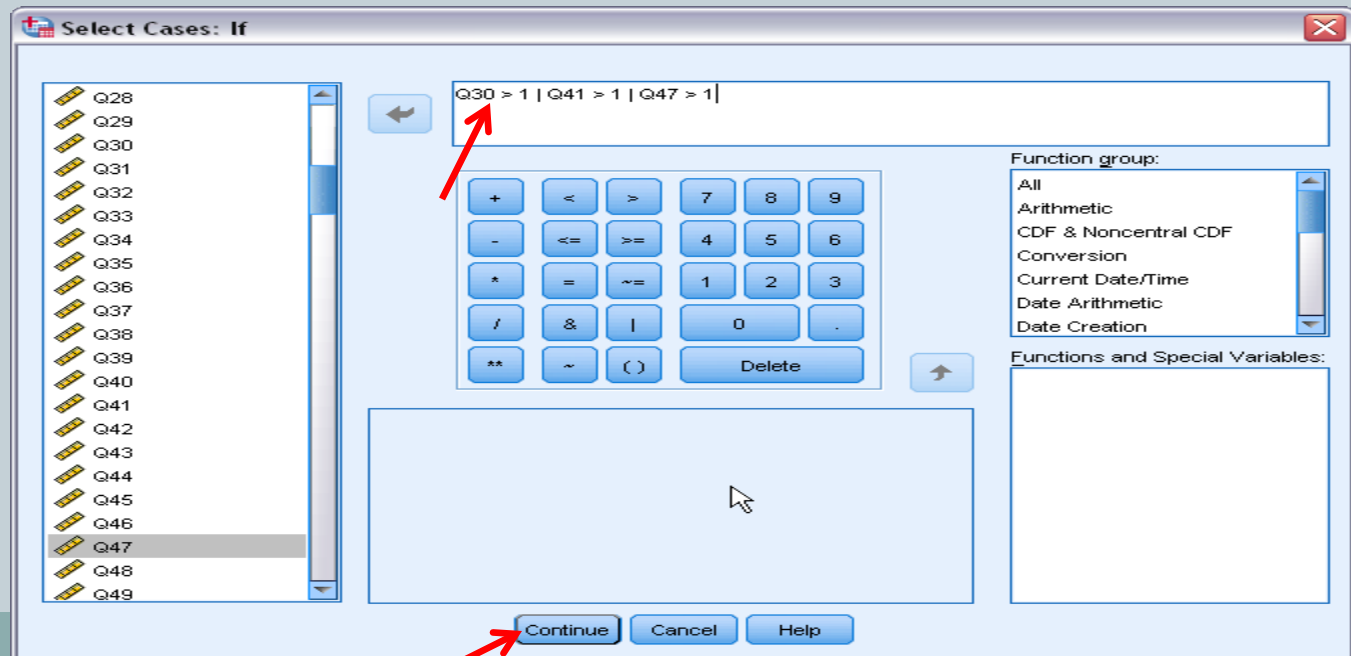
Sort and select cases

36

- **Select cases**

- Example 2. Select cases who used any of cigarettes, alcohol, and marijuana during the last 30 days.

1. Data —→ Select Cases
2. Click If button
3. Type $Q30 > 1 \mid Q41 > 1 \mid Q47 > 1$, click Continue



Basic statistical analysis

37

- **Descriptive statistics**

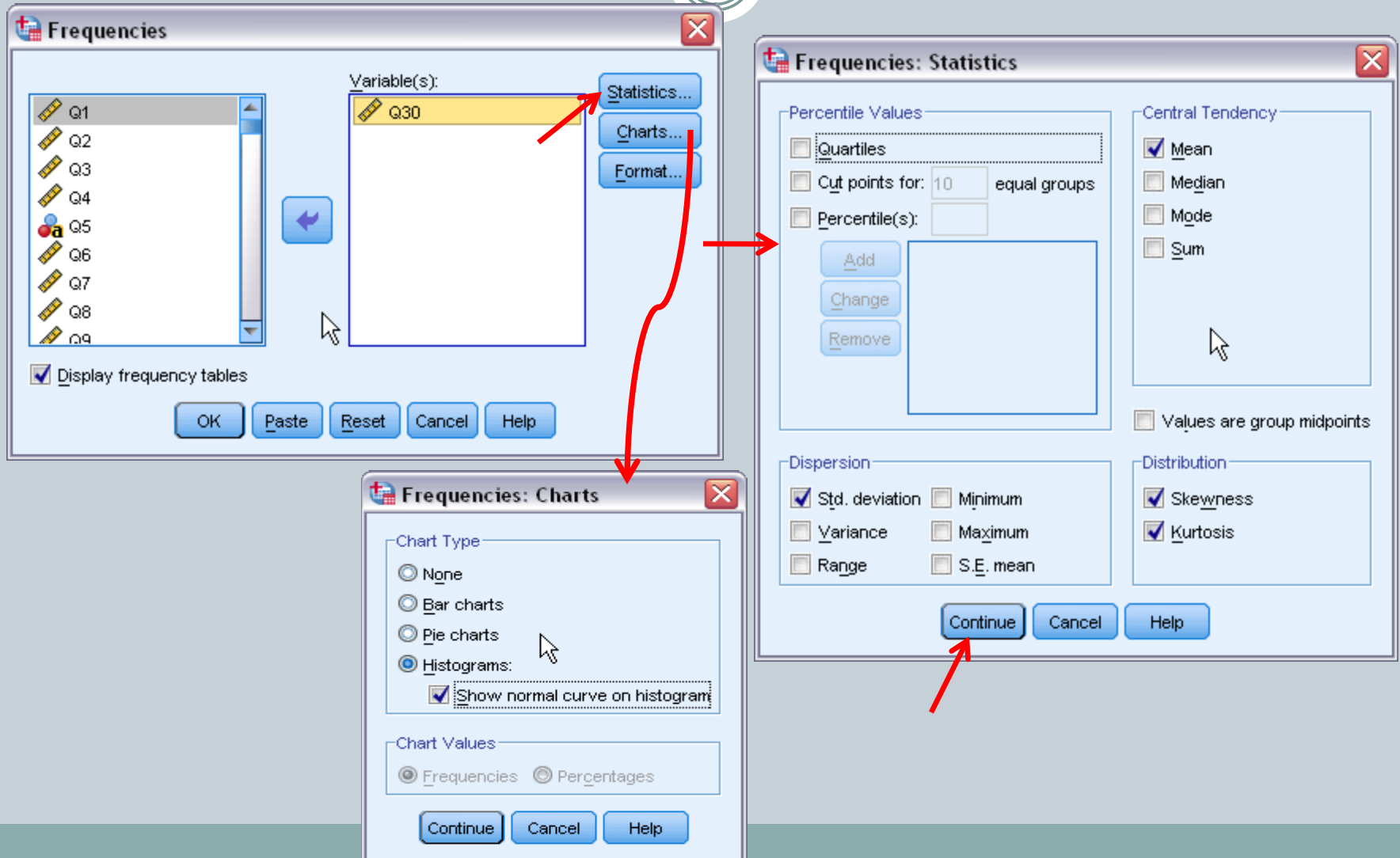
- Purposes:

1. Find wrong entries
2. Have basic knowledge about the sample and targeted variables in a study
3. Summarize data

Analyze → Descriptive statistics → Frequency

Continue

38



Frequency table

39

IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

it
.log
Frequencies
Title
Notes
Active Dataset
Statistics
Q30 How many days smoked 30 days

4 to 9 days	282	1.7	1.8	90.4
5 to 19 days	360	2.2	2.3	92.7
6 to 29 days	310	1.9	2.0	94.7
7 All 30 days	845	5.1	5.3	100.0
Total	15846	96.6	100.0	
Missing System	564	3.4		
Total	16410	100.0		

FREQUENCIES VARIABLES=Q30
/STATISTICS=STDDEV MEAN SKEWNESS SESKEW KURTOSIS SEKURT
/ORDER=ANALYSIS.

Frequencies

[DataSet1] C:\Documents and Settings\bianh\Desktop\working folder hui\Workshops Spring 2011\yrbs09.sav

Statistics

Q30 How many days smoked 30 days

N	Valid	15846
	Missing	564
Mean		1.67
Std. Deviation		1.627
Skewness		2.449
Std. Error of Skewness		.019
Kurtosis		4.630
Std. Error of Kurtosis		.039

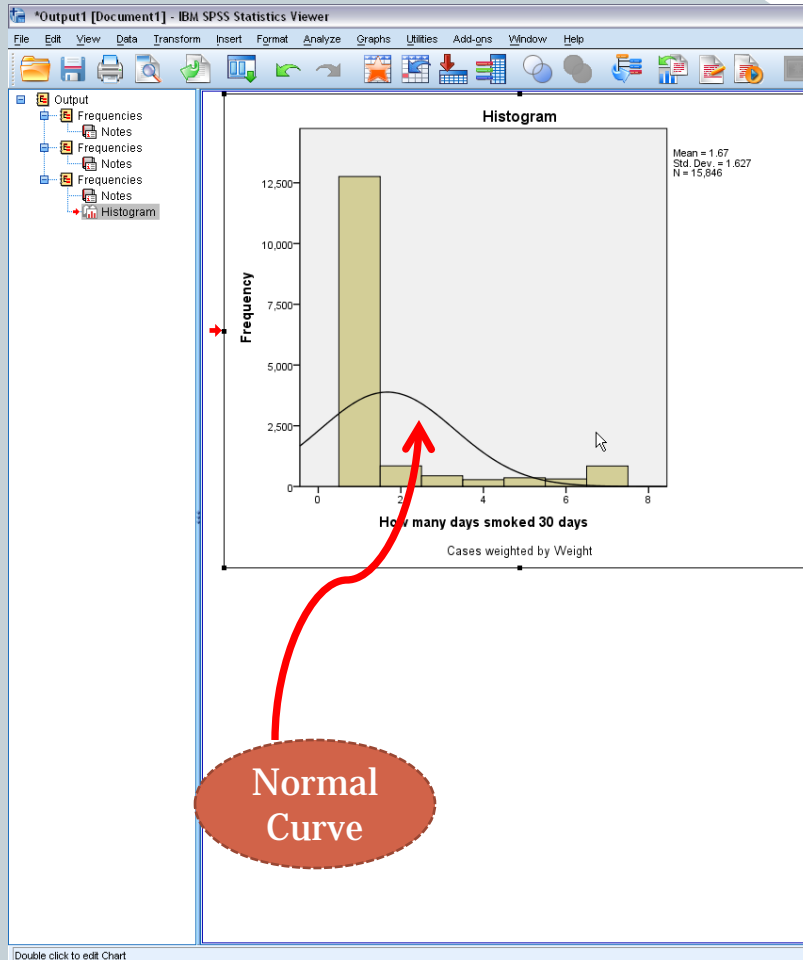
Q30 How many days smoked 30 days

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 0 days	12760	77.8	80.5	80.5
2 1 or 2 days	847	5.2	5.3	85.9
3 3 to 5 days	442	2.7	2.8	88.7
4 6 to 9 days	282	1.7	1.8	90.4
5 10 to 19 days	360	2.2	2.3	92.7
6 20 to 29 days	310	1.9	2.0	94.7
7 All 30 days	845	5.1	5.3	100.0
Total	15846	96.6	100.0	
Missing System	564	3.4		
Total	16410	100.0		

Double click to edit Pivot Table

IBM SPSS Statistics Processor is ready

H: 2.86, W: 5.22 in



1. **Skewness**: a measure of the asymmetry of a distribution. The normal distribution is symmetric and has a skewness value of zero.

Positive skewness: a long right tail.

Negative skewness: a long left tail.

Departure from symmetry : a skewness value more than twice its standard error.

2. **Kurtosis**: A measure of the extent to which observations cluster around

a central point. For a normal distribution, the value of the kurtosis

statistic is zero. *Leptokurtic* data values are more peaked, whereas *platykurtic* data values are flatter and more dispersed along the X axis.