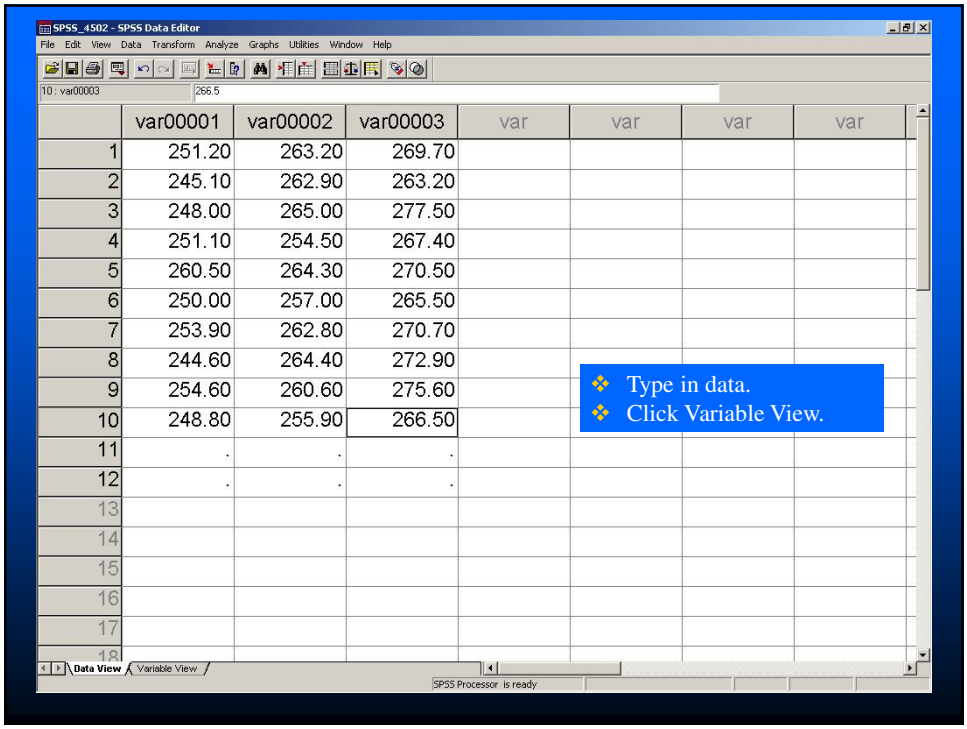
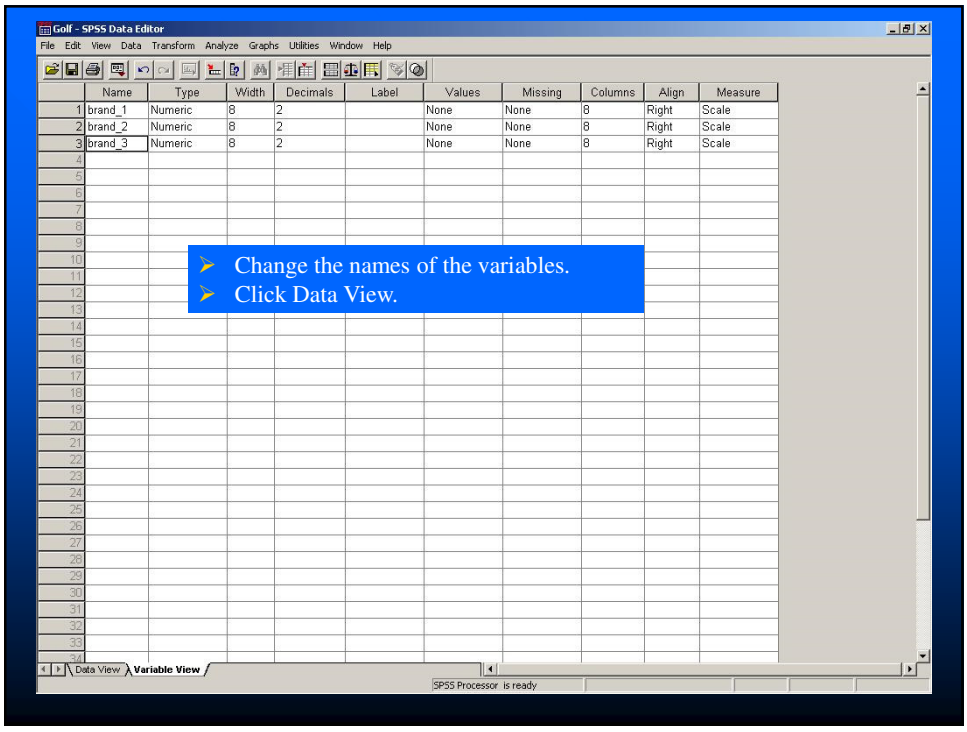
# Descriptive Statistics

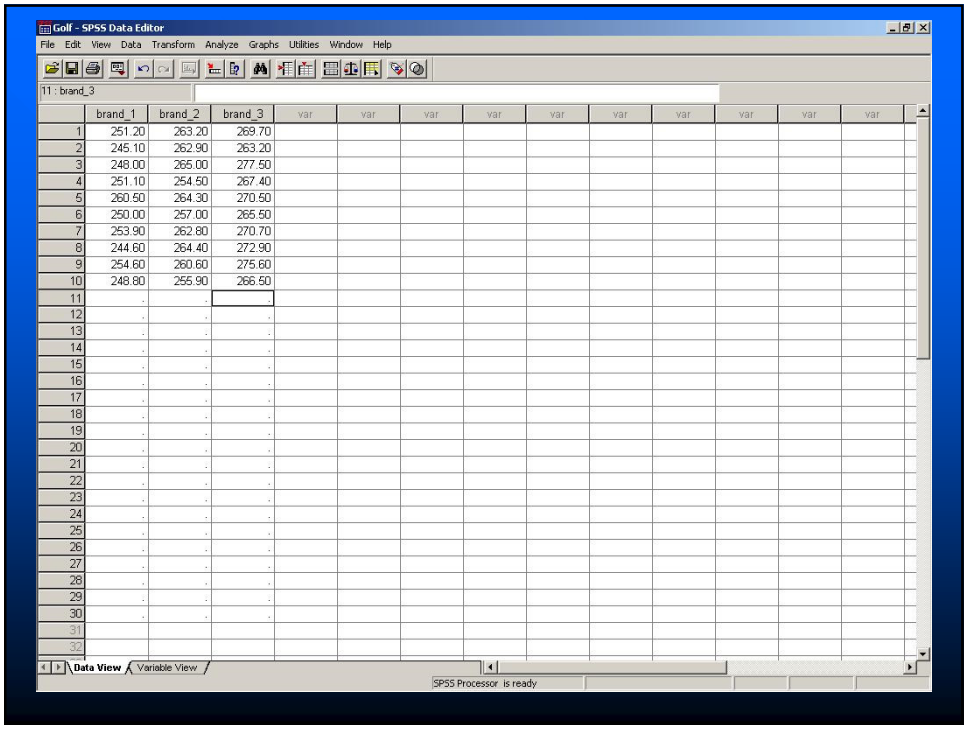
## GOLF DATA

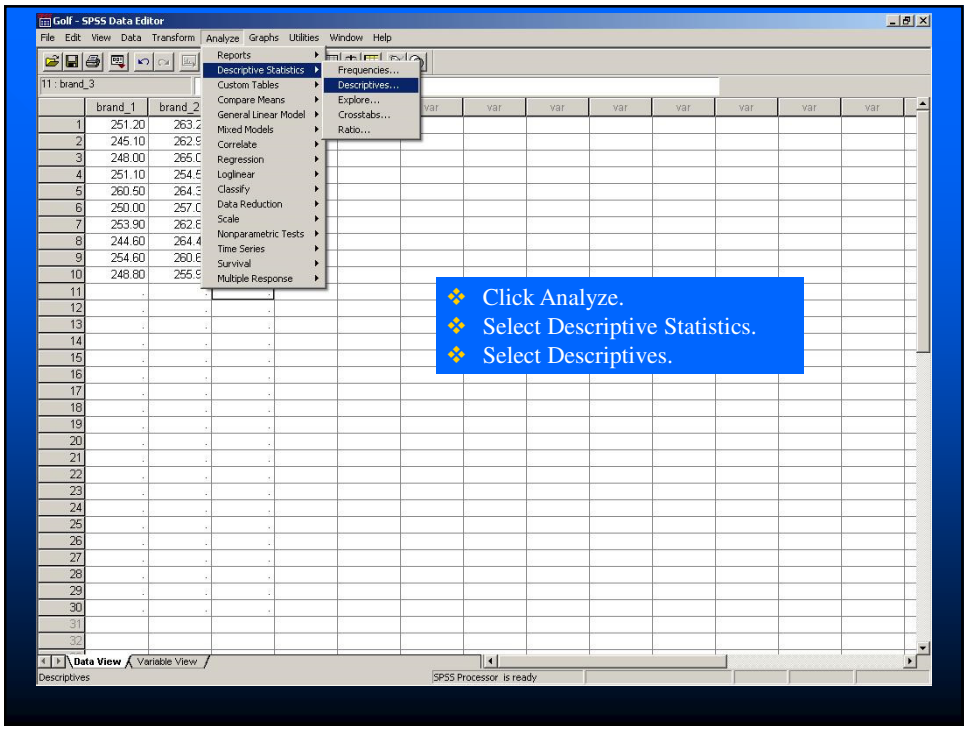
To compare the average distances traveled by three brand's of golf balls, ten balls were randomly selected from each of the three brands and the distance traveled by each ball was recorded.

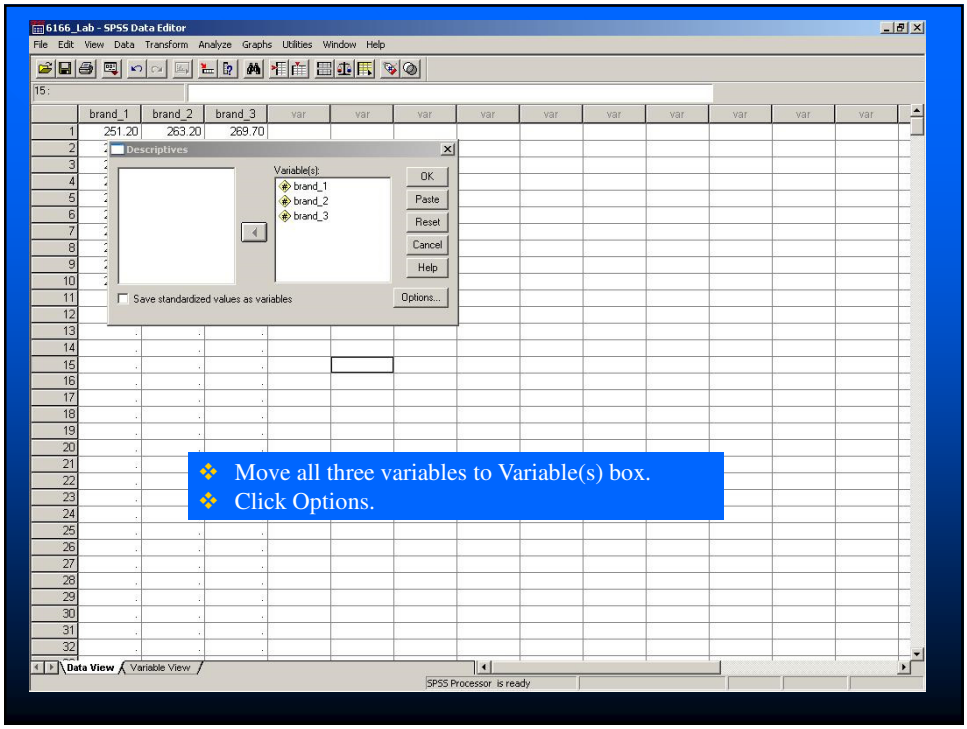
Brand A: 251.2, 245.1, 248, 251.1, 260.5, 250, 253.9, 244.6, 254.6, 248.8
Brand B: 263.2, 262.9, 265, 254.5, 264.3, 257, 262.8, 264.4, 260.6, 255.9
Brand C: 269.7, 263.2, 277.5, 267.4, 270.5, 265.5, 270.7, 272.9, 275.6, 266.5

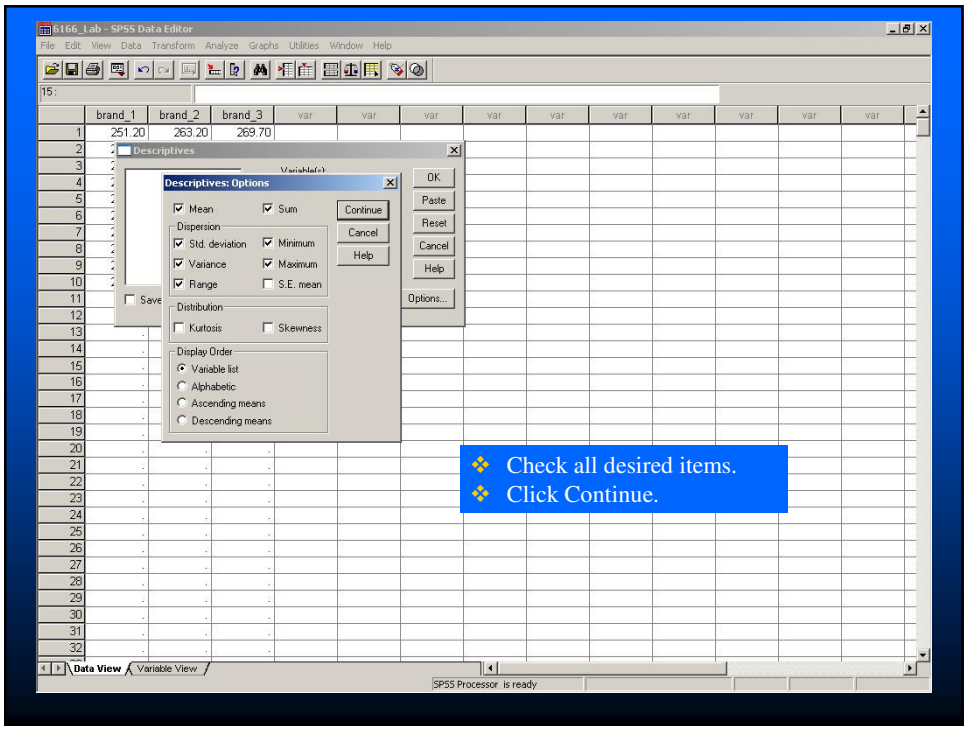

SPSS_4502 - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
10 : var00003
266.5
var00001 var00002 var00003 var var var var
1 251.20 263.20 269.70
2 245.10 262.90 263.20
3 248.00 265.00 277.50
4 251.10 254.50 267.40
5 260.50 264.30 270.50
6 250.00 257.00 265.50
7 253.90 262.80 270.70
8 244.60 264.40 272.90
9 254.60 260.60 275.60
10 248.80 255.90 266.50
11 . . .
12 . . .
13
14
15
16
17
Data View Variable View
SPSS Processor is ready
Type in data.
Click Variable View.

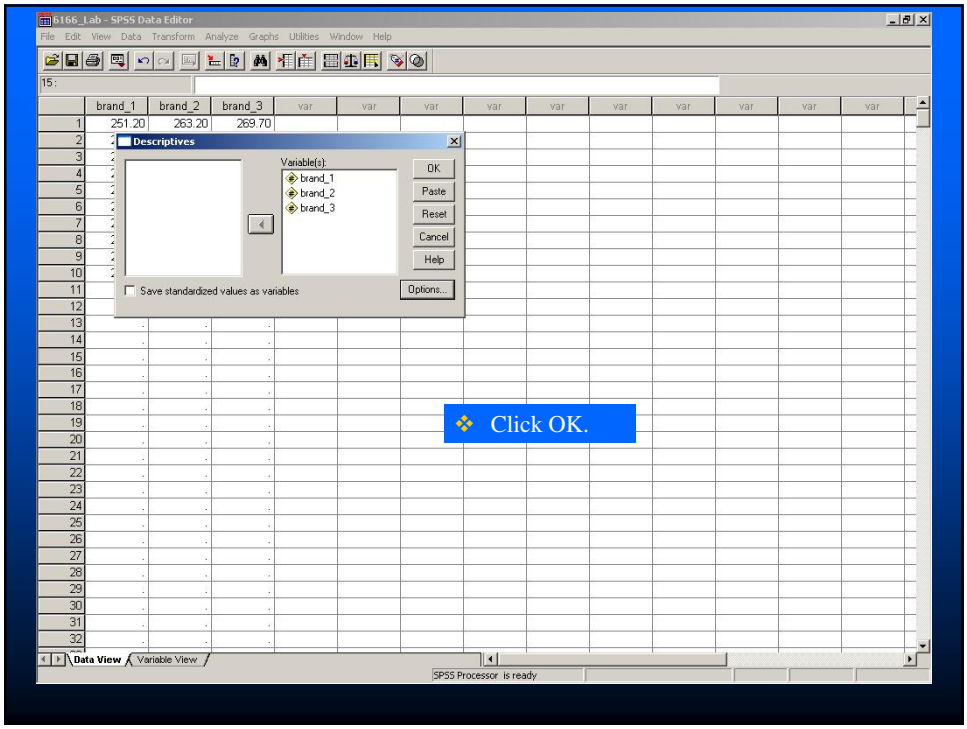

Golf - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
Name Type Width Decimals Label Values Missing Columns Align Measure
1 brand_1 Numeric 8 2 None None 8 Right Scale
2 brand_2 Numeric 8 2 None None 8 Right Scale
3 brand_3 Numeric 8 2 None None 8 Right Scale
Data View Variable View
SPSS Processor is ready
Change the names of the variables.
Click Data View.


Golf - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
11 : brand_3
brand_1 brand_2 brand_3 var
1 251.20 263.20 269.70
2 245.10 262.90 263.20
3 248.00 265.00 277.50
4 251.10 254.50 267.40
5 260.50 264.30 270.50
6 250.00 257.00 265.50
7 253.90 262.80 270.70
8 244.60 264.40 272.90
9 254.60 260.60 275.60
10 248.80 255.90 266.50
Data View Variable View
SPSS Processor is ready


Golf - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
Reports
Descriptive Statistics
Custom Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Frequencies...
Descriptives...
Explore...
Crosstabs...
Ratio...
Click Analyze.
Select Descriptive Statistics.
Select Descriptives.
Data View Variable View
Descriptives
SPSS Processor is ready


6166_Lab - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
brand_1 brand_2 brand_3
251.20 263.20 269.70
Descriptives
Variable(s):
brand_1
brand_2
brand_3
OK
Paste
Reset
Cancel
Help
Save standardized values as variables
Options...
Move all three variables to Variable(s) box.
Click Options.
Data View
Variable View
SPSS Processor is ready

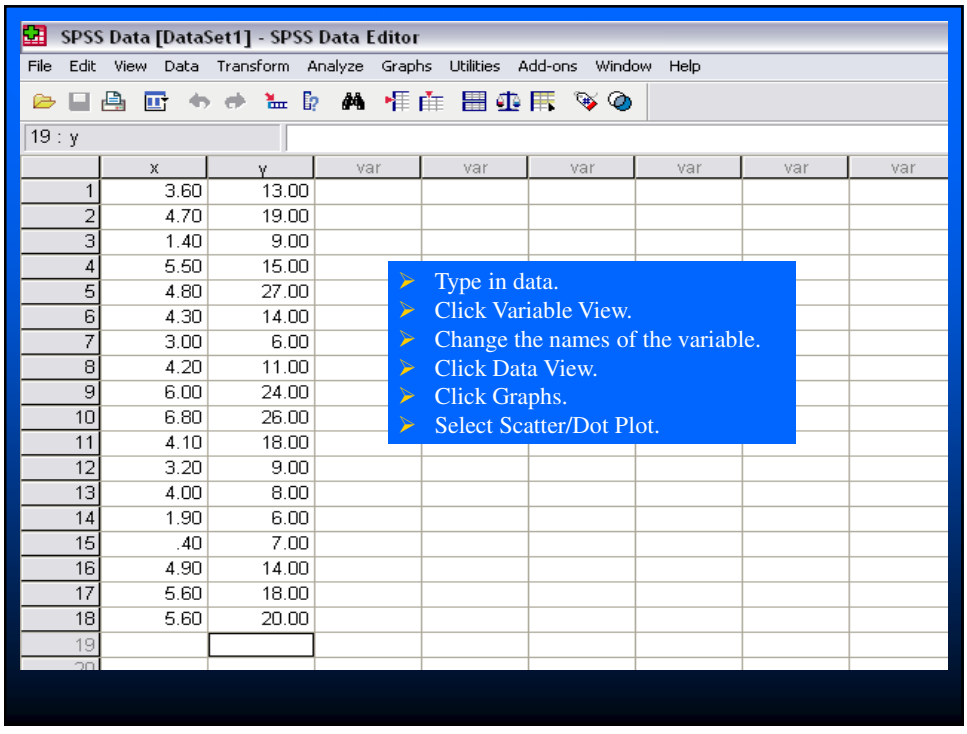

6166_Lab - SPSS Data Editor
Descriptives: Options
Mean
Sum
Continue
Cancel
Help
Dispersion
Std. deviation
Minimum
Variance
Maximum
Range
S.E. mean
Distribution
Kurtosis
Skewness
Display Order
Variable list
Alphabetic
Ascending means
Descending means
Check all desired items.
Click Continue.

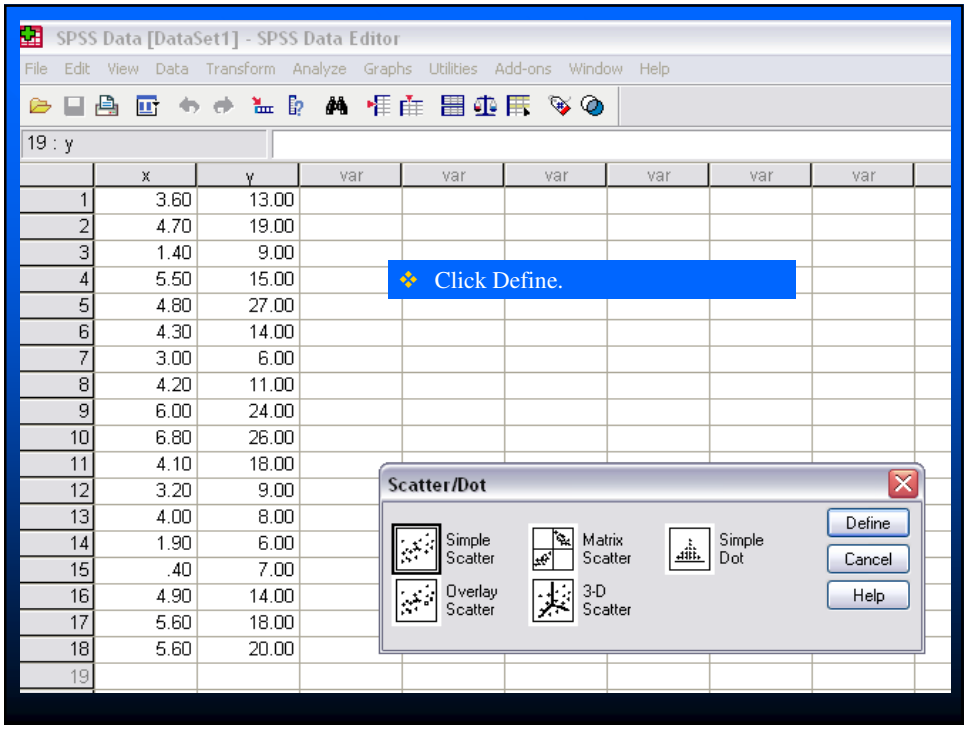


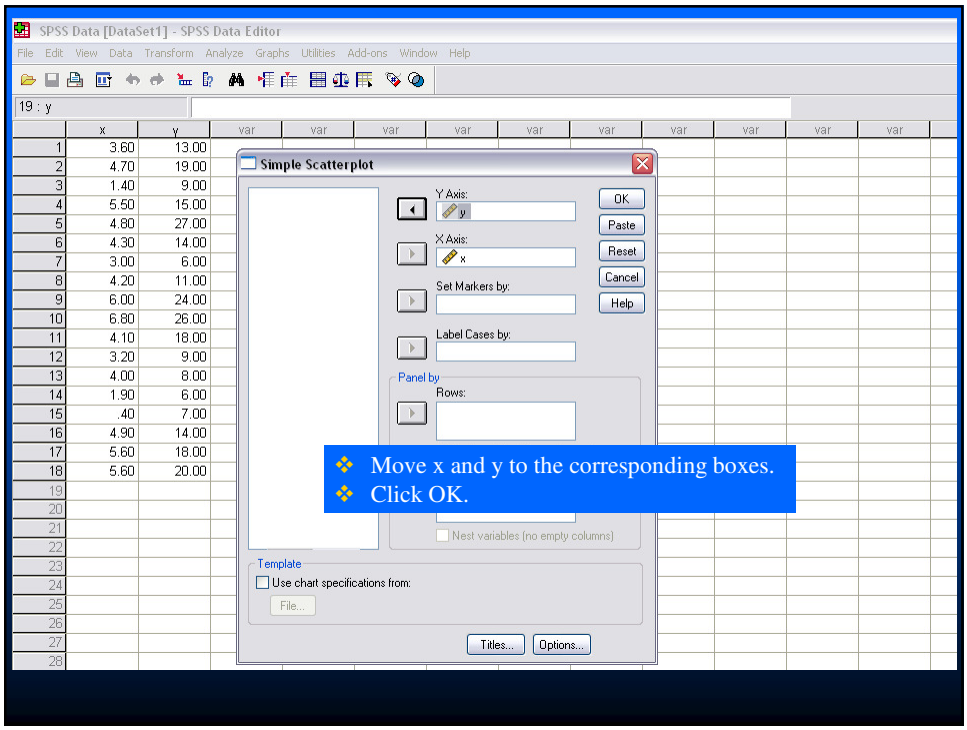

## Scatter Plot

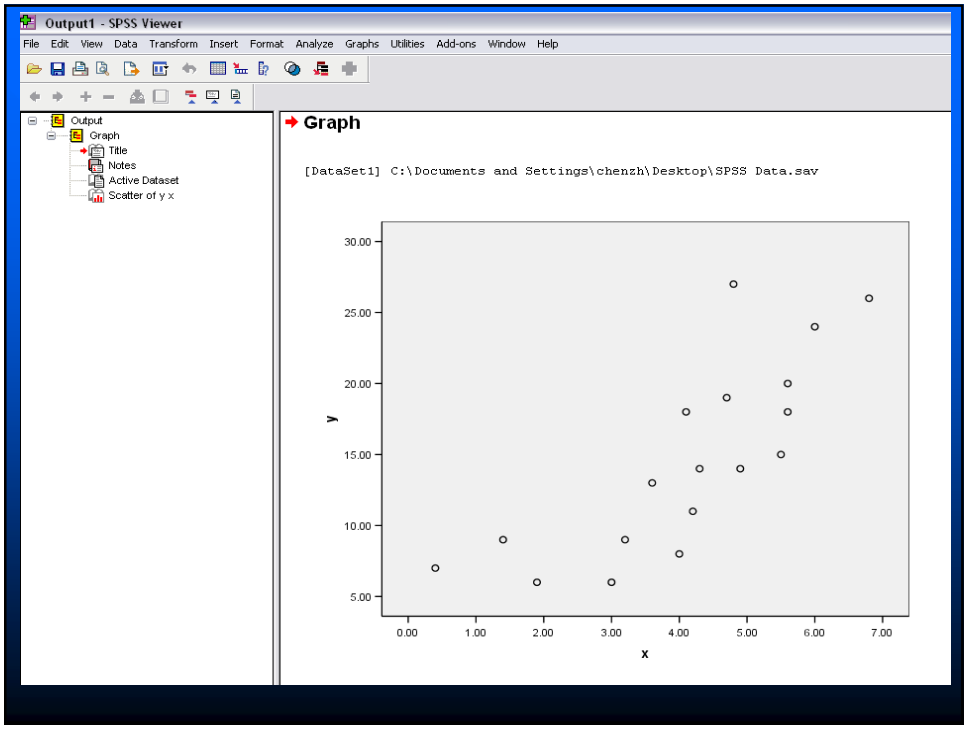
Construct a scatter plot for the following data set:

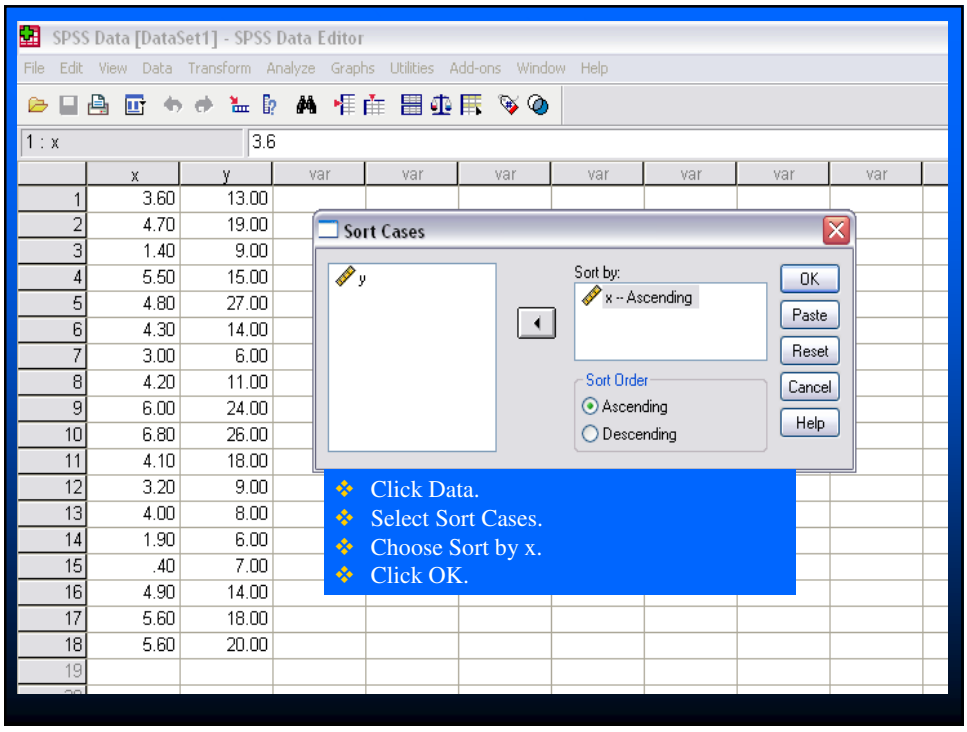
(3.6, 13), (4.7, 19), (1.4, 9), (5.5, 15), (4.8, 27), (4.3, 14), (3.0, 6), (4.2, 11), (6.0, 24), (6.8, 26), (4.1, 18), (3.2, 9), (4.0, 8), (1.9, 6), (0.4, 7), (4.9, 14), (5.6, 18), (5.6, 20) .

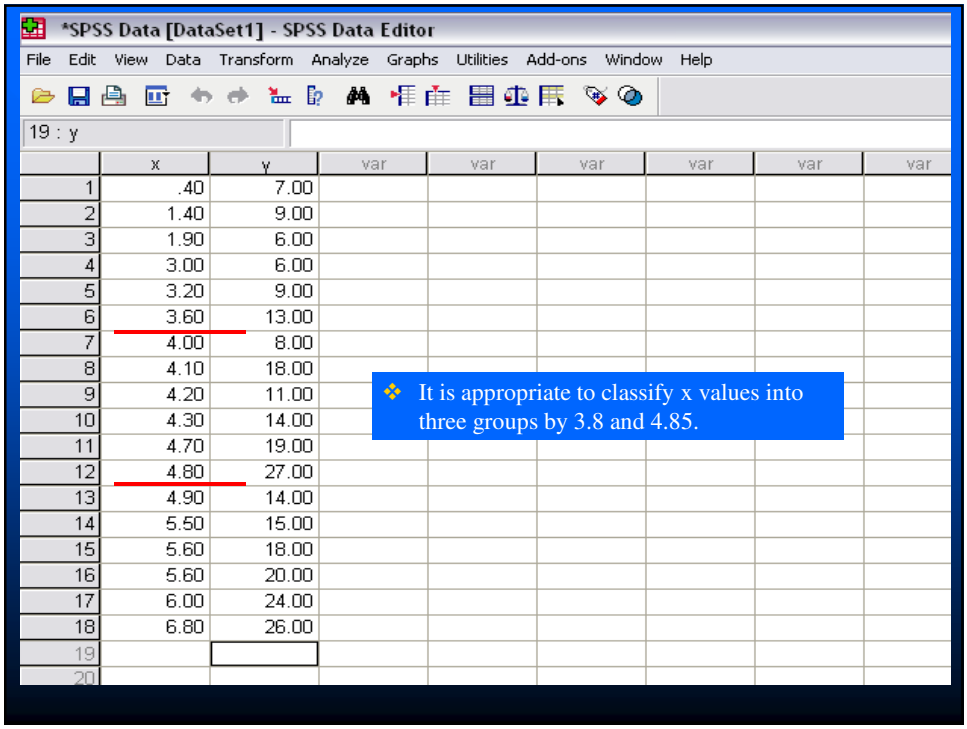


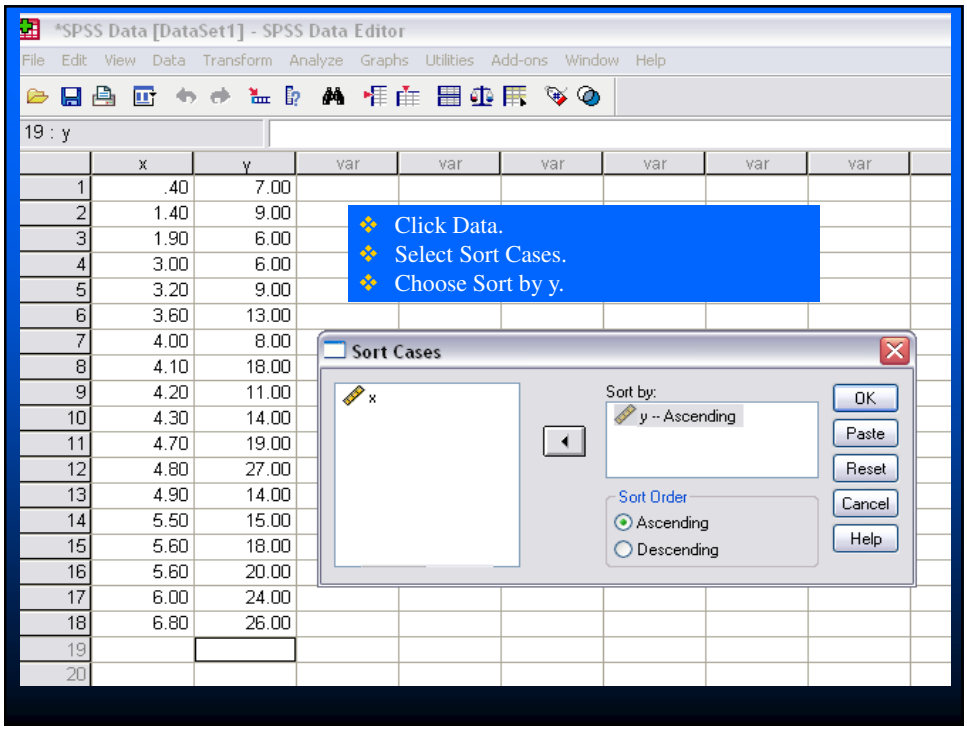

SPSS Data [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
19 : y
Click Define.
Scatter/Dot
Simple Scatter
Matrix Scatter
Simple Dot
Overlay Scatter
3-D Scatter
Define
Cancel
Help


SPSS Data [DataSet1] - SPSS Data Editor
Simple Scatterplot
Y Axis:
X Axis:
Set Markers by:
Label Cases by:
Panel by
Rows:
OK
Paste
Reset
Cancel
Help
Move x and y to the corresponding boxes.
Click OK.
Nest variables (no empty columns)
Template
Use chart specifications from:
File...
Titles...
Options...


Output1 - SPSS Viewer
File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help
Output
Graph
Title
Notes
Active Dataset
Scatter of y x
Graph
[DataSet1] C:\Documents and Settings\chenzh\Desktop\SPSS Data.sav
30.00
25.00
20.00
15.00
10.00
5.00
y
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
x


SPSS Data [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
1 : x
3.6
x y var var var var var var var
1 3.60 13.00
2 4.70 19.00
3 1.40 9.00
4 5.50 15.00
5 4.80 27.00
6 4.30 14.00
7 3.00 6.00
8 4.20 11.00
9 6.00 24.00
10 6.80 26.00
11 4.10 18.00
12 3.20 9.00
13 4.00 8.00
14 1.90 6.00
15 .40 7.00
16 4.90 14.00
17 5.60 18.00
18 5.60 20.00
19
Sort Cases
y
Sort by:
x -- Ascending
OK
Paste
Reset
Cancel
Help
Sort Order
Ascending
Descending
Click Data.
Select Sort Cases.
Choose Sort by x.
Click OK.


*SPSS Data [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
19 : y
x y var var var var var var
1 .40 7.00
2 1.40 9.00
3 1.90 6.00
4 3.00 6.00
5 3.20 9.00
6 3.60 13.00
7 4.00 8.00
8 4.10 18.00
9 4.20 11.00
10 4.30 14.00
11 4.70 19.00
12 4.80 27.00
13 4.90 14.00
14 5.50 15.00
15 5.60 18.00
16 5.60 20.00
17 6.00 24.00
18 6.80 26.00
19
20
It is appropriate to classify x values into three groups by 3.8 and 4.85.


*SPSS Data [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
19 : y
x y var var var var var var
1 .40 7.00
2 1.40 9.00
3 1.90 6.00
4 3.00 6.00
5 3.20 9.00
6 3.60 13.00
7 4.00 8.00
8 4.10 18.00
9 4.20 11.00
10 4.30 14.00
11 4.70 19.00
12 4.80 27.00
13 4.90 14.00
14 5.50 15.00
15 5.60 18.00
16 5.60 20.00
17 6.00 24.00
18 6.80 26.00
19
20
Click Data.
Select Sort Cases.
Choose Sort by y.
Sort Cases
x
Sort by:
y -- Ascending
OK
Paste
Reset
Cancel
Help
Sort Order
Ascending
Descending

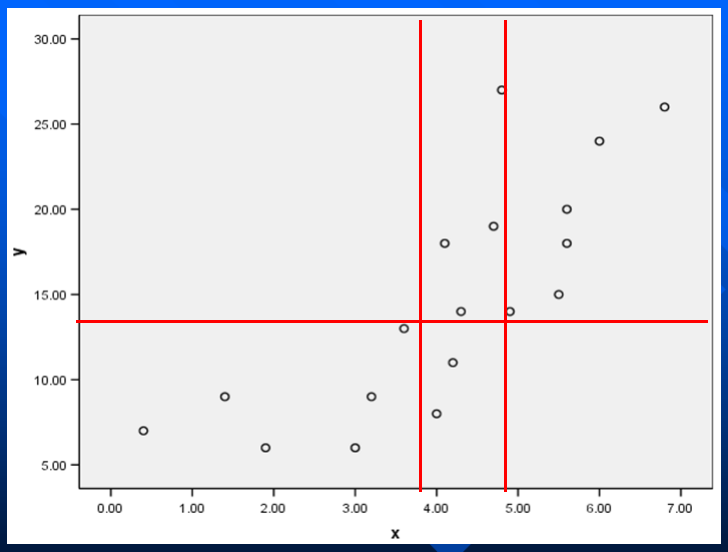
*SPSS Data [DataSet1] - SPSS Data Editor

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

19 : y

| | x | y | var | var | var | var | var | var |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.90 | 6.00 | | | | | | |
| 2 | 3.00 | 6.00 | | | | | | |
| 3 | .40 | 7.00 | | | | | | |
| 4 | 4.00 | 8.00 | | | | | | |
| 5 | 1.40 | 9.00 | | | | | | |
| 6 | 3.20 | 9.00 | | | | | | |
| 7 | 4.20 | 11.00 | | | | | | |
| 8 | 3.60 | 13.00 | | | | | | |
| 9 | 4.30 | 14.00 | | | | | | |
| 10 | 4.90 | 14.00 | | | | | | |
| 11 | 5.50 | 15.00 | | | | | | |
| 12 | 4.10 | 18.00 | | | | | | |
| 13 | 5.60 | 18.00 | | | | | | |
| 14 | 4.70 | 19.00 | | | | | | |
| 15 | 5.60 | 20.00 | | | | | | |
| 16 | 6.00 | 24.00 | | | | | | |
| 17 | 6.80 | 26.00 | | | | | | |
| 18 | 4.80 | 27.00 | | | | | | |
| 19 | | | | | | | | |
| 20 | | | | | | | | |

- It is appropriate to classify y values into two groups by 13.5 or 14.5.



Classify 18 data points into 6 cells:

(3.6, 13), (4.7, 19), (1.4, 9), (5.5, 15), (4.8, 27), (4.3, 14), (3.0, 6), (4.2, 11), (6.0, 24), (6.8, 26), (4.1, 18), (3.2, 9), (4.0, 8), (1.9, 6), (0.4, 7), (4.9, 14), (5.6, 18), (5.6, 20) .

| | $x \le 3.8$ | $3.8 < x \le 4.85$ | $x > 4.85$ | |
|---|---|---|---|---|
| $y \le 13.5$ | 6 | 2 | 0 | 8 |
| $y > 13.5$ | 0 | 4 | 6 | 10 |
| | 6 | 6 | 6 | 18 |

## Chi-square Test for Independence

## CHI-SQUARE TEST EXAMPLE

A study to decide whether an association exists between family income and education attainment of householder yielded the following data:

| | No HS Degree | HS Degree | College Degree |
|---|---|---|---|
| Under 25000 | 65 | 63 | 14 |
| [ 25000 , 50000 ) | 35 | 84 | 38 |
| [ 50000 , 75000 ) | 11 | 44 | 43 |
| 75000 or more | 4 | 23 | 72 |

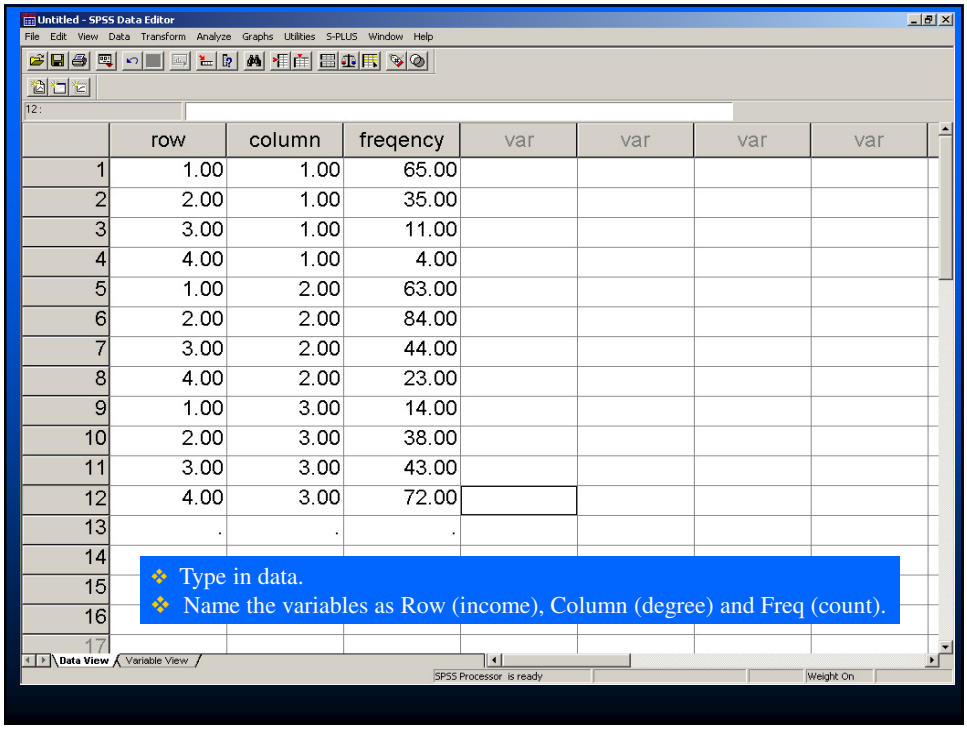
## CHI-SQUARE TEST EXAMPLE

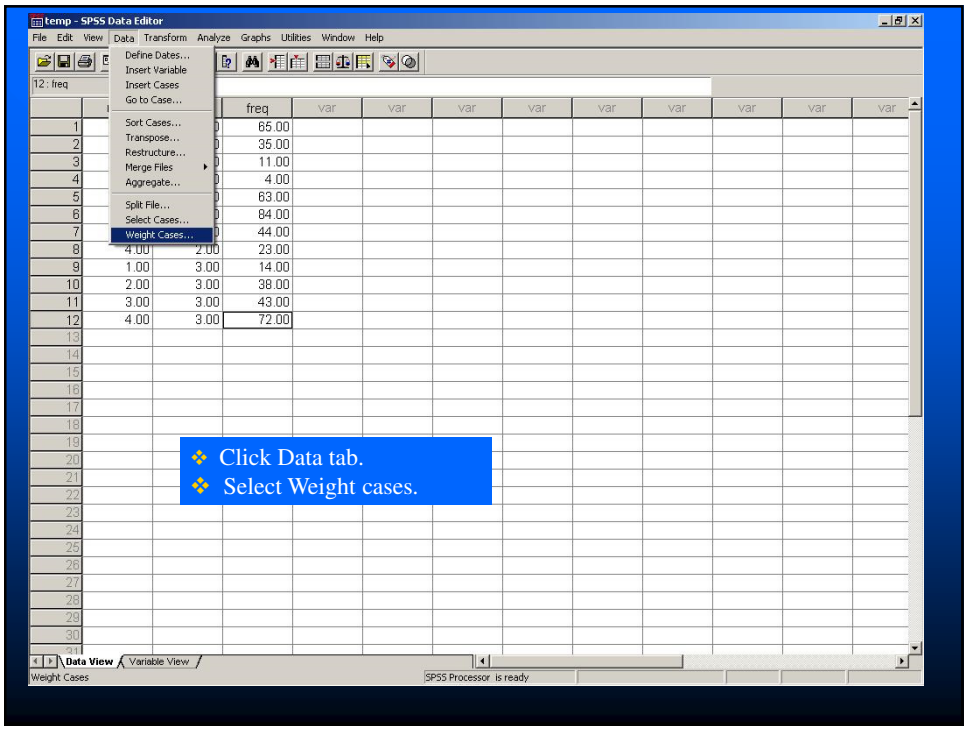
Question:

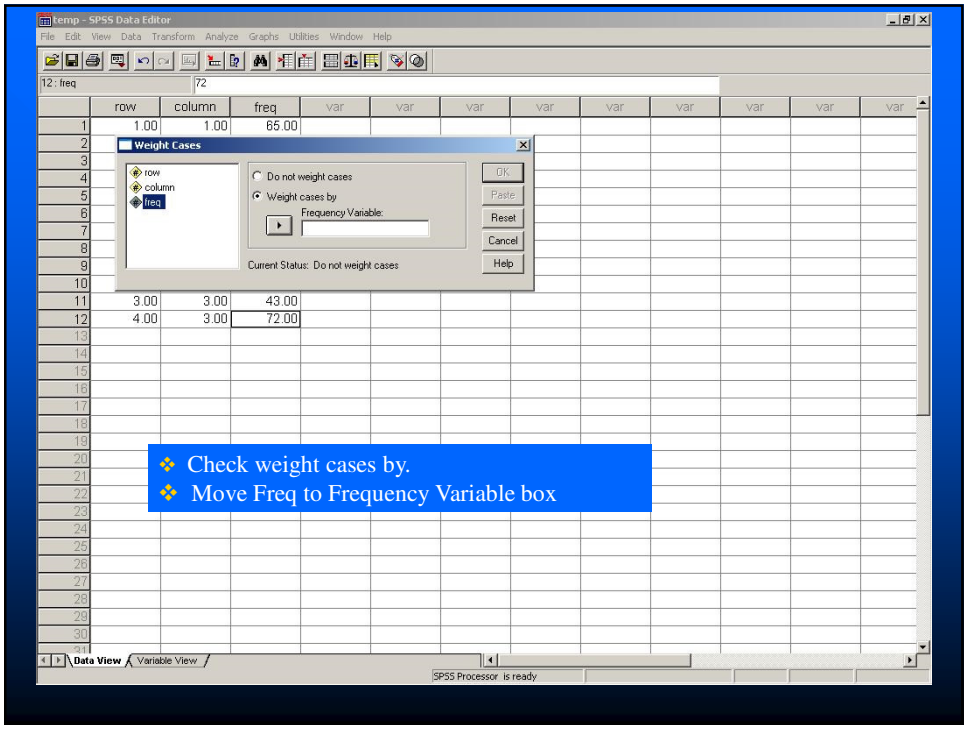
Do the data provide sufficient evidence to conclude that an association exists between family income and education attainment of householder?

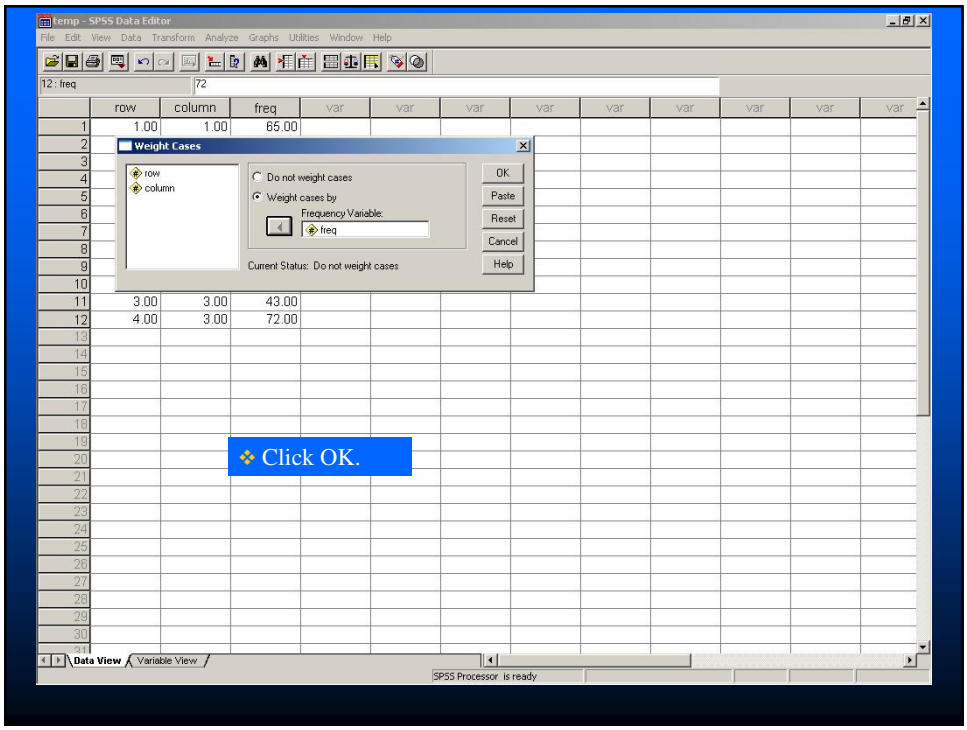
$H_0$: There is no association between family income and education attainment of householder.

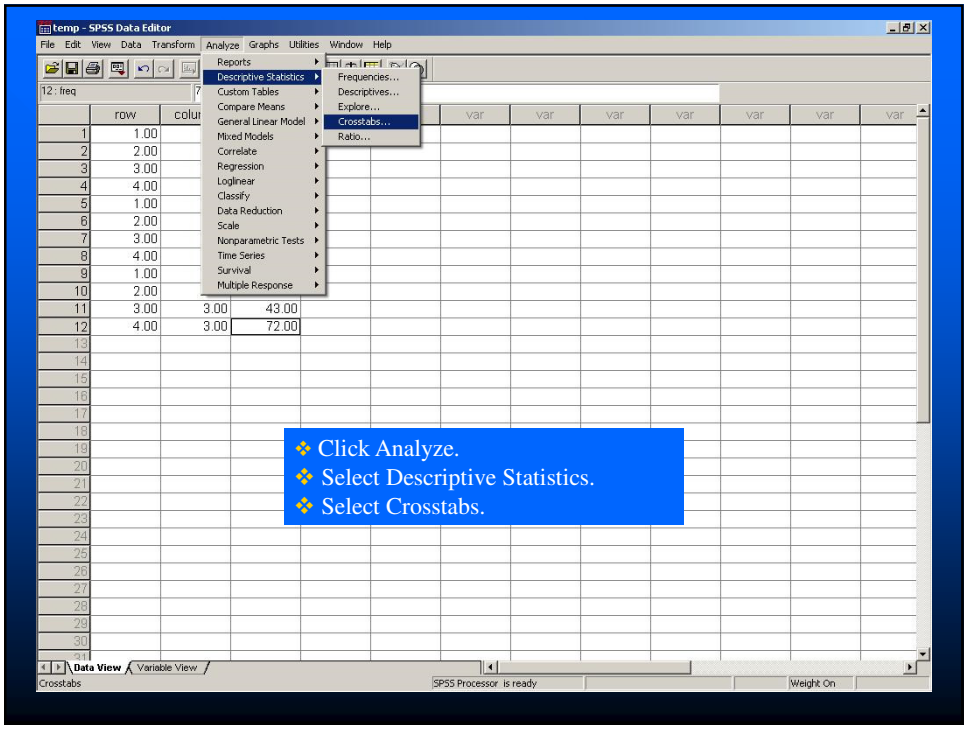
$H_a$: There exists association between family income and education attainment of householder.

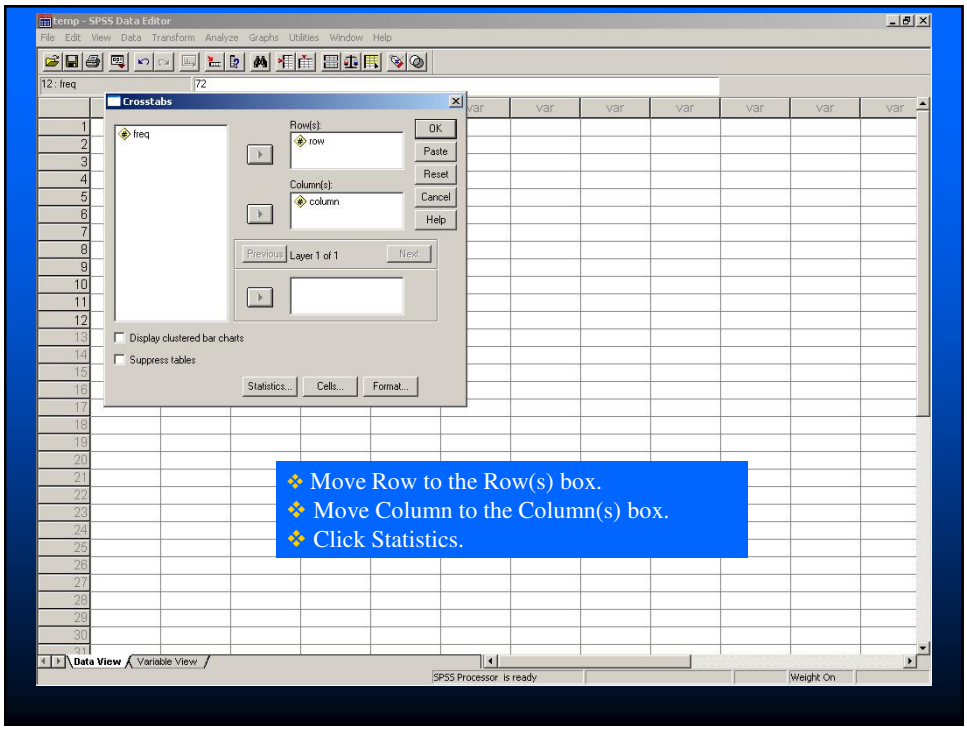

Untitled - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities S-PLUS Window Help
row column freqency var var var var
1 1.00 1.00 65.00
2 2.00 1.00 35.00
3 3.00 1.00 11.00
4 4.00 1.00 4.00
5 1.00 2.00 63.00
6 2.00 2.00 84.00
7 3.00 2.00 44.00
8 4.00 2.00 23.00
9 1.00 3.00 14.00
10 2.00 3.00 38.00
11 3.00 3.00 43.00
12 4.00 3.00 72.00
13 . . .
14
15
16
17
Data View Variable View
SPSS Processor is ready
Weight On
Type in data.
Name the variables as Row (income), Column (degree) and Freq (count).

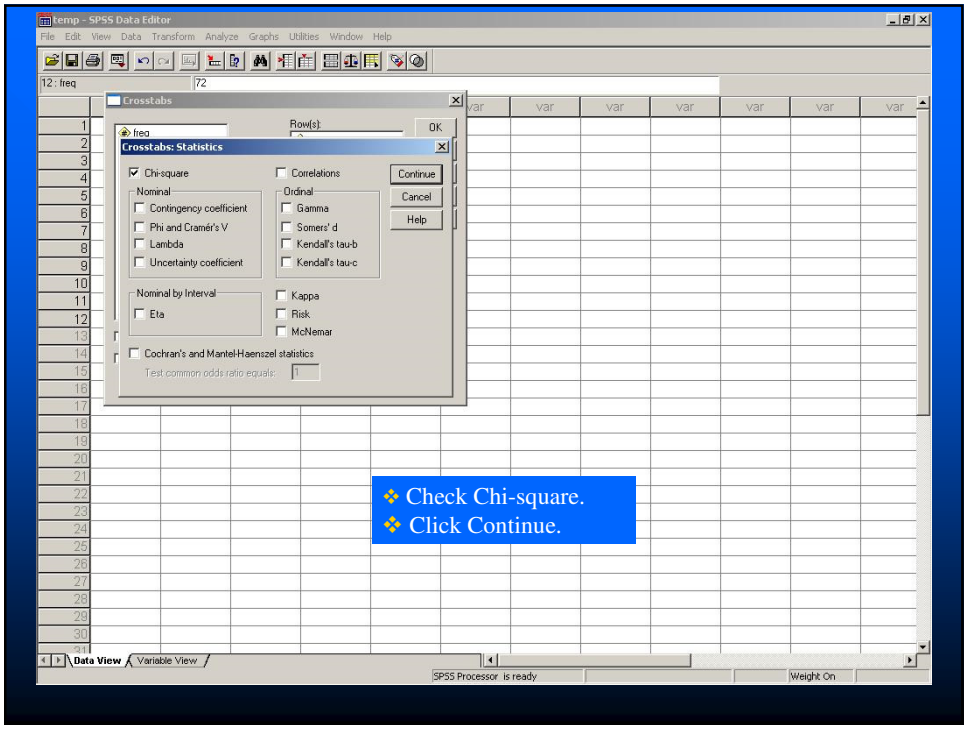

temp - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
Define Dates...
Insert Variable
Insert Cases
Go to Case...
Sort Cases...
Transpose...
Restructure...
Merge Files
Aggregate...
Split File...
Select Cases...
Weight Cases...
freq var var var var var var var var var
65.00
35.00
11.00
4.00
63.00
84.00
44.00
4.00 2.00 23.00
1.00 3.00 14.00
2.00 3.00 38.00
3.00 3.00 43.00
4.00 3.00 72.00
Data View Variable View
Weight Cases
SPSS Processor is ready
Click Data tab.
Select Weight cases.

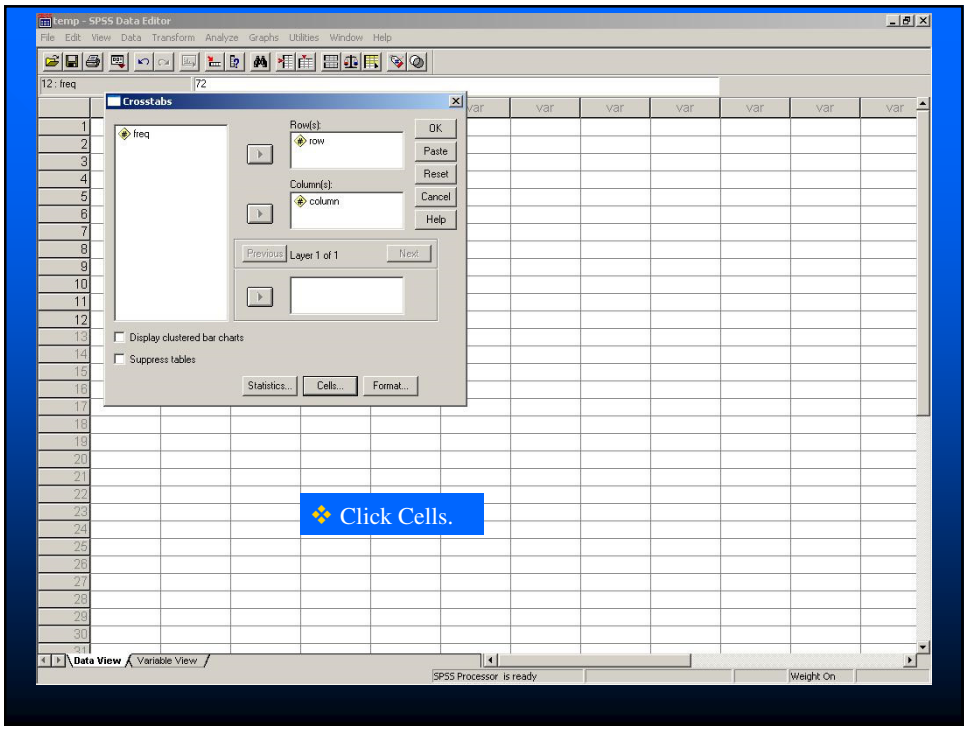

temp - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
12 : freq
72
row column freq var var var var var var var var var
1.00 1.00 65.00
3.00 3.00 43.00
4.00 3.00 72.00
Weight Cases
row
column
freq
Do not weight cases
Weight cases by
Frequency Variable:
Current Status: Do not weight cases
OK
Paste
Reset
Cancel
Help
Check weight cases by.
Move Freq to Frequency Variable box
Data View
Variable View
SPSS Processor is ready

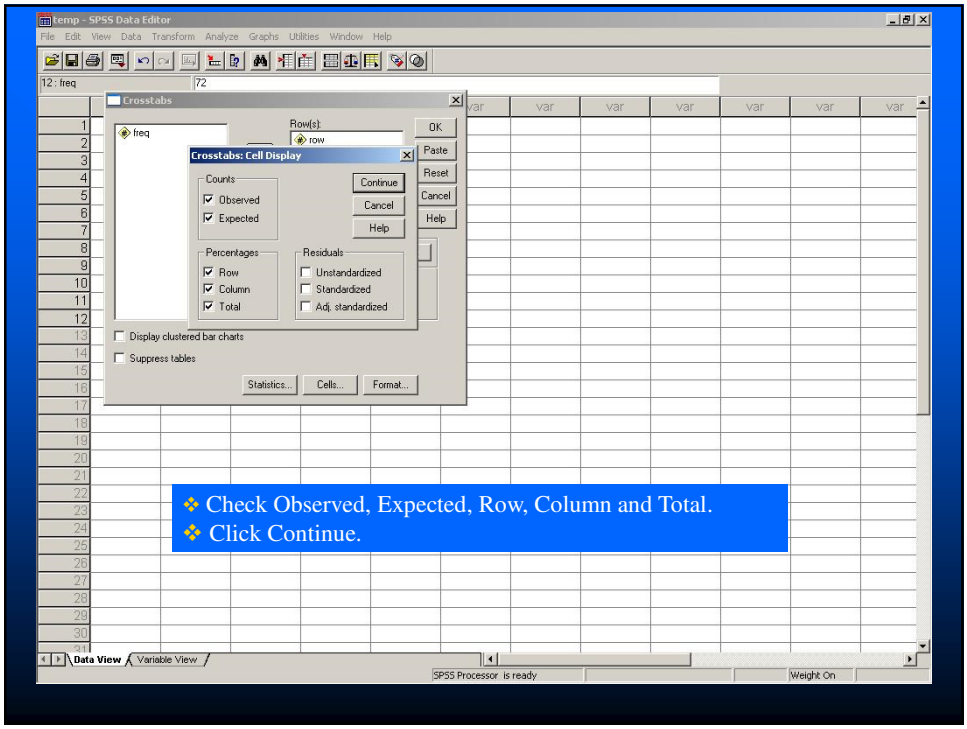

temp - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
12 : freq
72
row column freq var var var var var var var var var
1.00 1.00 65.00
3.00 3.00 43.00
4.00 3.00 72.00
Weight Cases
row
column
Do not weight cases
Weight cases by
Frequency Variable:
freq
Current Status: Do not weight cases
OK
Paste
Reset
Cancel
Help
Click OK.
Data View
Variable View
SPSS Processor is ready


temp - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
Reports
Descriptive Statistics
Custom Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Frequencies...
Descriptives...
Explore...
Crosstabs...
Ratio...
Click Analyze.
Select Descriptive Statistics.
Select Crosstabs.
Data View
Variable View
Crosstabs
SPSS Processor is ready
Weight On

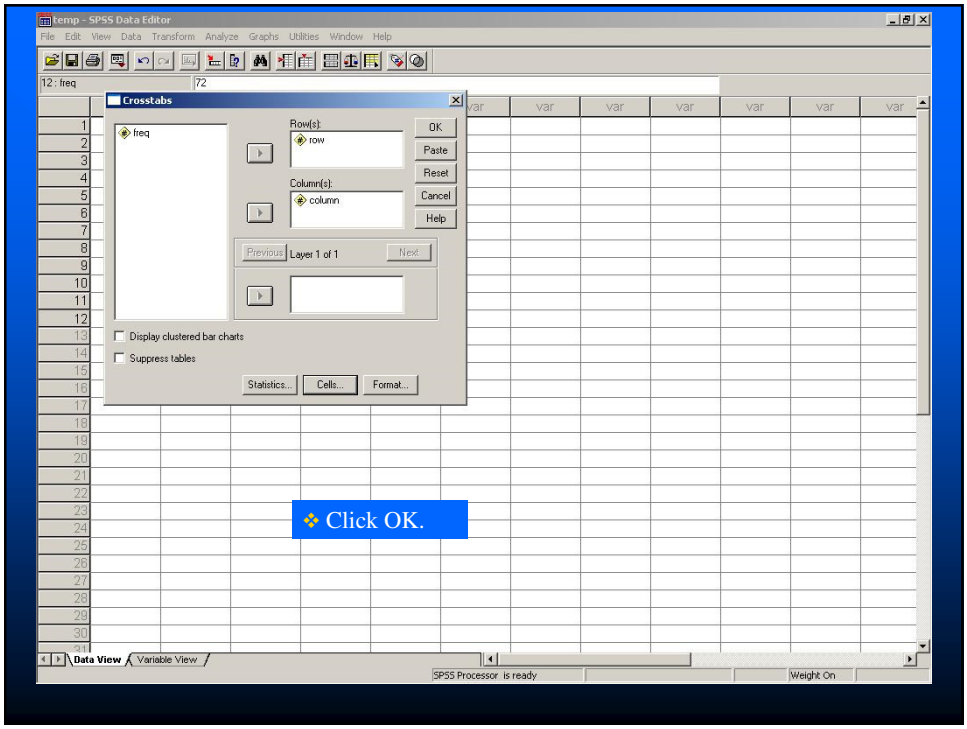

temp - SPSS Data Editor
Crosstabs
freq
Row(s):
row
Column(s):
column
OK
Paste
Reset
Cancel
Help
Previous
Layer 1 of 1
Next
Display clustered bar charts
Suppress tables
Statistics...
Cells...
Format...
Move Row to the Row(s) box.
Move Column to the Column(s) box.
Click Statistics.
Data View
Variable View
SPSS Processor is ready
Weight On


Temp - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
12 : freq
72
Crosstabs
freq
Row(s):
OK
Crosstabs: Statistics
Chi-square
Correlations
Continue
Cancel
Help
Nominal
Contingency coefficient
Phi and Cramér's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
Check Chi-square.
Click Continue.
Data View
Variable View
SPSS Processor is ready
Weight On

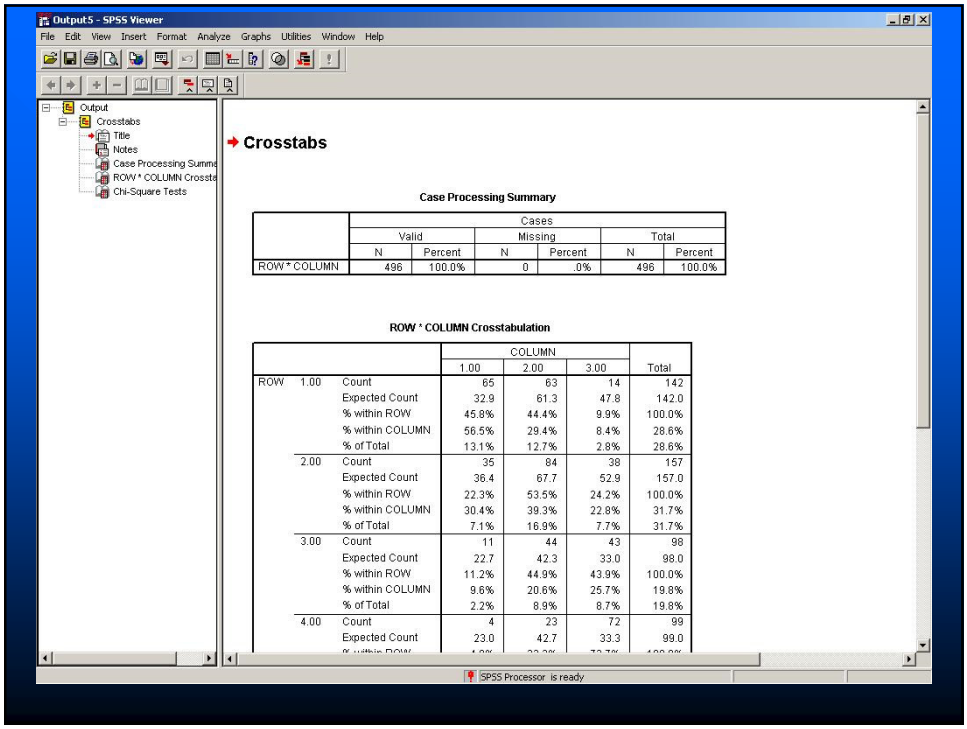

Temp - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
12 : freq
72
Crosstabs
freq
Row(s):
row
Column(s):
column
OK
Paste
Reset
Cancel
Help
Previous
Layer 1 of 1
Next
Display clustered bar charts
Suppress tables
Statistics...
Cells...
Format...
Click Cells.
Data View
Variable View
SPSS Processor is ready
Weight On




- Check Observed, Expected, Row, Column and Total.
- Click Continue.




- Click OK.




**VAR00001 * VAR00002 Crosstabulation**

| | | | VAR00002 | | | |
|---|---|---|---|---|---|---|
| | | | 1.00 | 2.00 | 3.00 | Total |
| VAR00001 | 1.00 | Count | 65 | 63 | 14 | 142 |
| | | Expected Count | 32.9 | 61.3 | 47.8 | 142.0 |
| | | % within VAR00001 | 45.8% | 44.4% | 9.9% | 100.0% |
| | | % within VAR00002 | 56.5% | 29.4% | 8.4% | 28.6% |
| | | % of Total | 13.1% | 12.7% | 2.8% | 28.6% |
| | 2.00 | Count | 35 | 84 | 38 | 157 |
| | | Expected Count | 36.4 | 67.7 | 52.9 | 157.0 |
| | | % within VAR00001 | 22.3% | 53.5% | 24.2% | 100.0% |
| | | % within VAR00002 | 30.4% | 39.3% | 22.8% | 31.7% |
| | | % of Total | 7.1% | 16.9% | 7.7% | 31.7% |
| | 3.00 | Count | 11 | 44 | 43 | 98 |
| | | Expected Count | 22.7 | 42.3 | 33.0 | 98.0 |
| | | % within VAR00001 | 11.2% | 44.9% | 43.9% | 100.0% |
| | | % within VAR00002 | 9.6% | 20.6% | 25.7% | 19.8% |
| | | % of Total | 2.2% | 8.9% | 8.7% | 19.8% |
| | 4.00 | Count | 4 | 23 | 72 | 99 |
| | | Expected Count | 23.0 | 42.7 | 33.3 | 99.0 |
| | | % within VAR00001 | 4.0% | 23.2% | 72.7% | 100.0% |
| | | % within VAR00002 | 3.5% | 10.7% | 43.1% | 20.0% |
| | | % of Total | .8% | 4.6% | 14.5% | 20.0% |
| Total | | Count | 115 | 214 | 167 | 496 |
| | | Expected Count | 115.0 | 214.0 | 167.0 | 496.0 |
| | | % within VAR00001 | 23.2% | 43.1% | 33.7% | 100.0% |
| | | % within VAR00002 | 100.0% | 100.0% | 100.0% | 100.0% |
| | | % of Total | 23.2% | 43.1% | 33.7% | 100.0% |

**Chi-Square Tests**

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 142.102[a] | 6 | .000 |
| Likelihood Ratio | 144.634 | 6 | .000 |
| Linear-by-Linear Association | 125.225 | 1 | .000 |
| N of Valid Cases | 496 | | |

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 22.72.