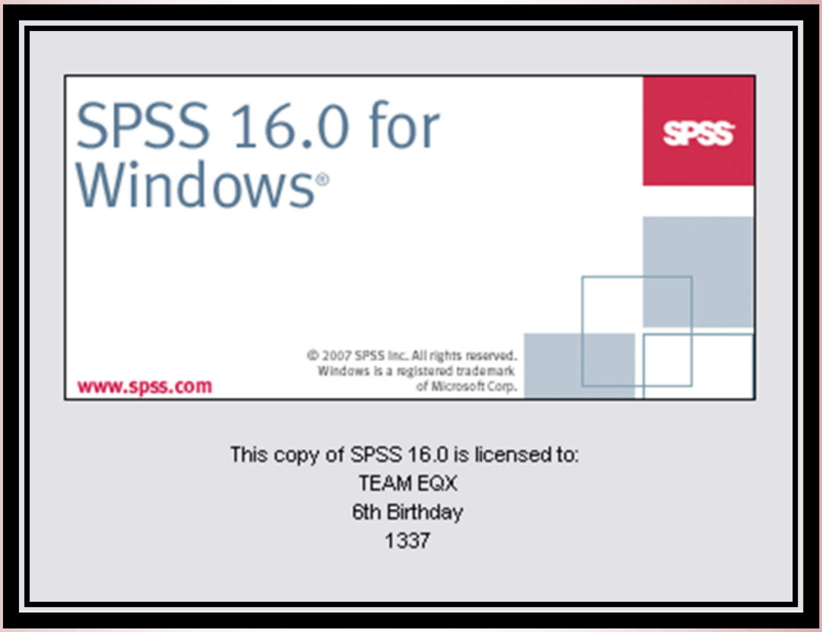
SPSS

16.0



Thun Kosal

2011

មាតិកា

Page

SPSS16.0 .... **1**

Data View 3

Variable View 3

Design Variable 7

Frequencies Analysis **8**

សង់ Pie chart **12**

សង់ Bar chart **17**

Descriptive Analysis **20**

សង់ Histogram **23**

Crosstabs Analysis **25**

Test of hypothesis **29**

One Sample test for mean 30

Independent Sample test 32

Dependent Sample test 34

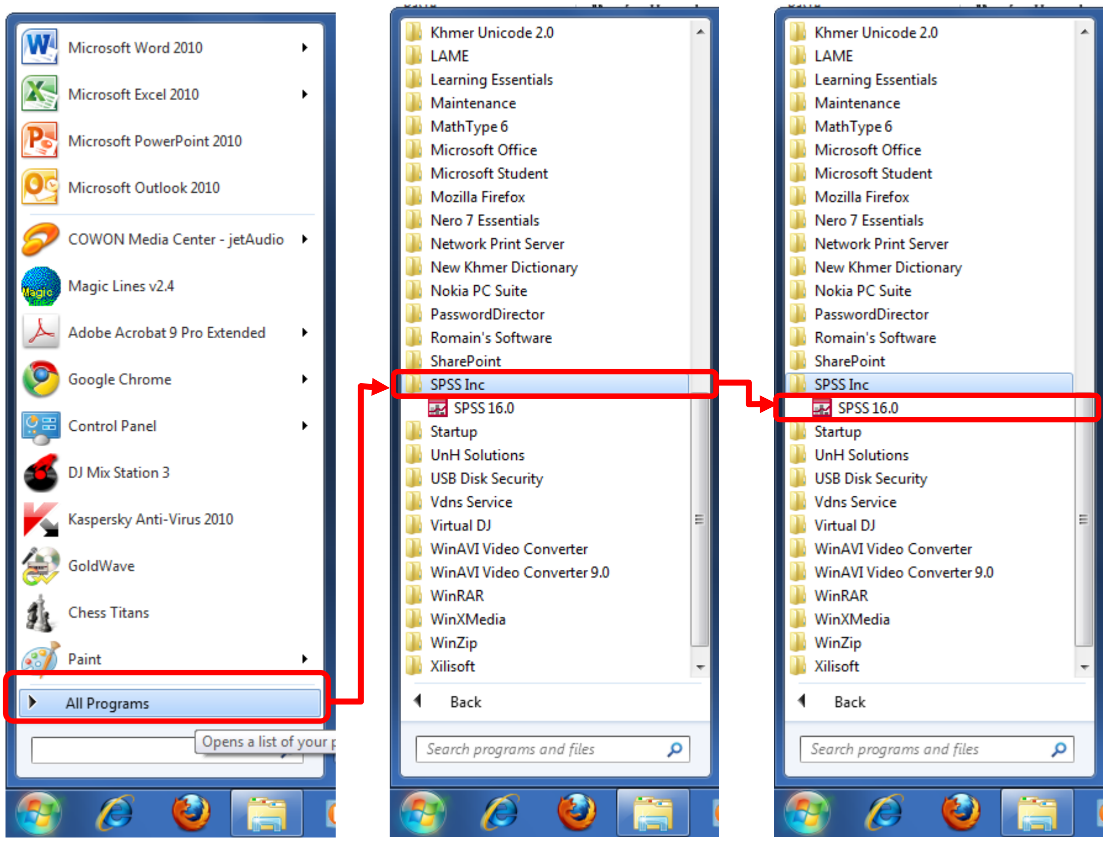
Table: The Normal Distribution **40**

SPSS 16.0

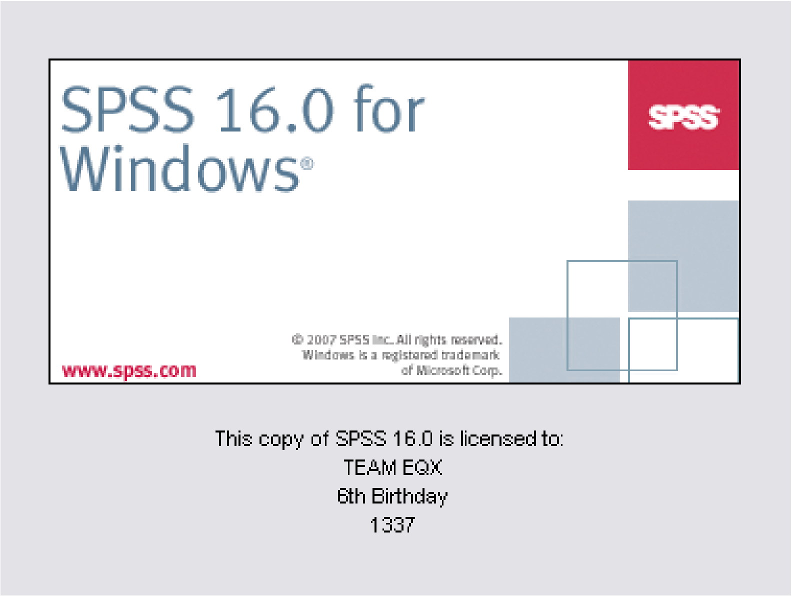
ពាក្យថា SPSS គឺមកពី Statistics Package for the Social Sciences ។

ដើម្បីចូលទៅកាន់កម្មវិធី SPSS សូមចុច:

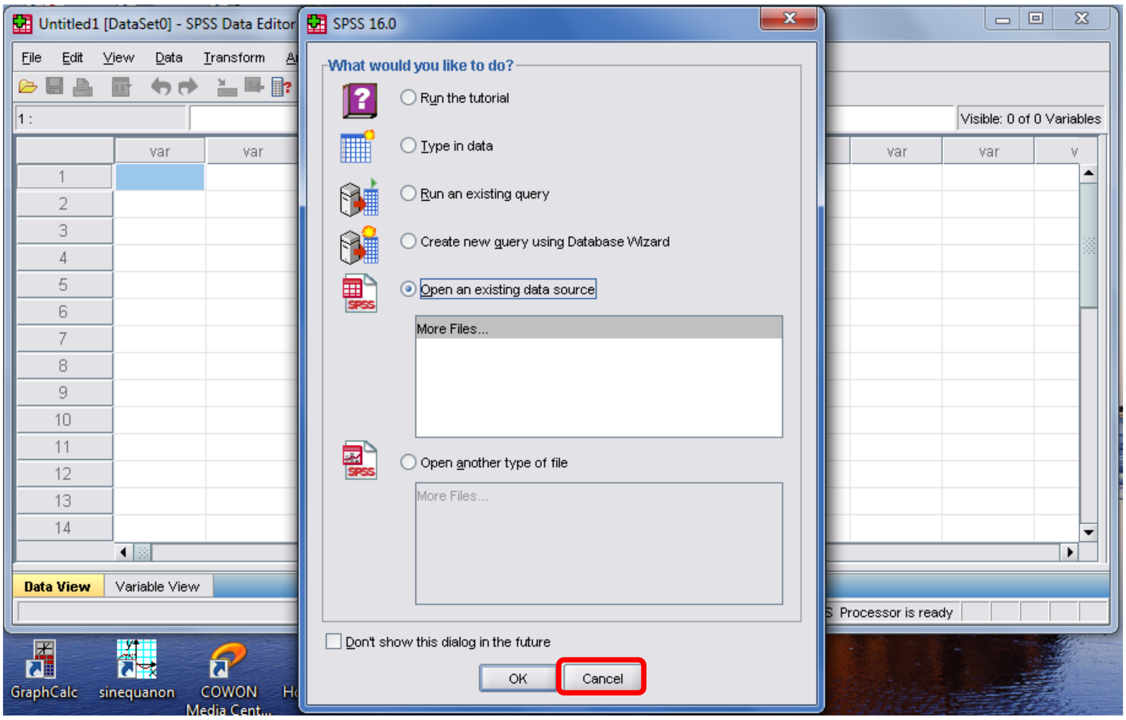
Start ➡ Programs /All programs (Win 7) ➡ SPSS Inc ➡ SPSS 16.0



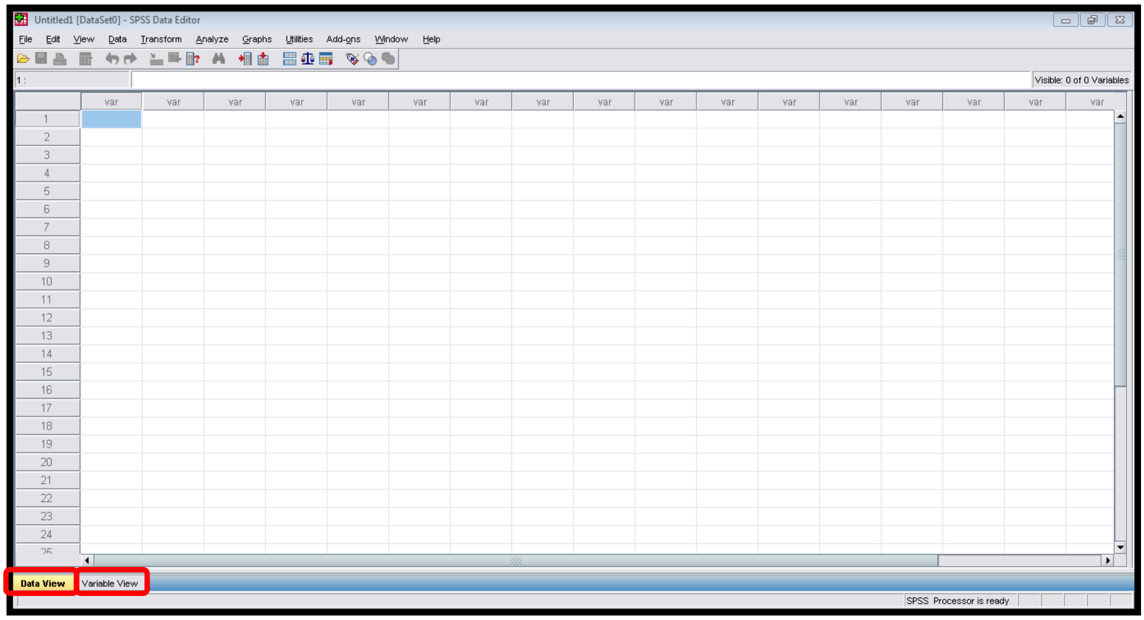
យើងបានផ្ទាំងដូចខាងក្រោម:



បន្ទាប់មកយើងឃើញផ្ទាំង:



ចុចលើពាក្យ Cancel យើងបាន:



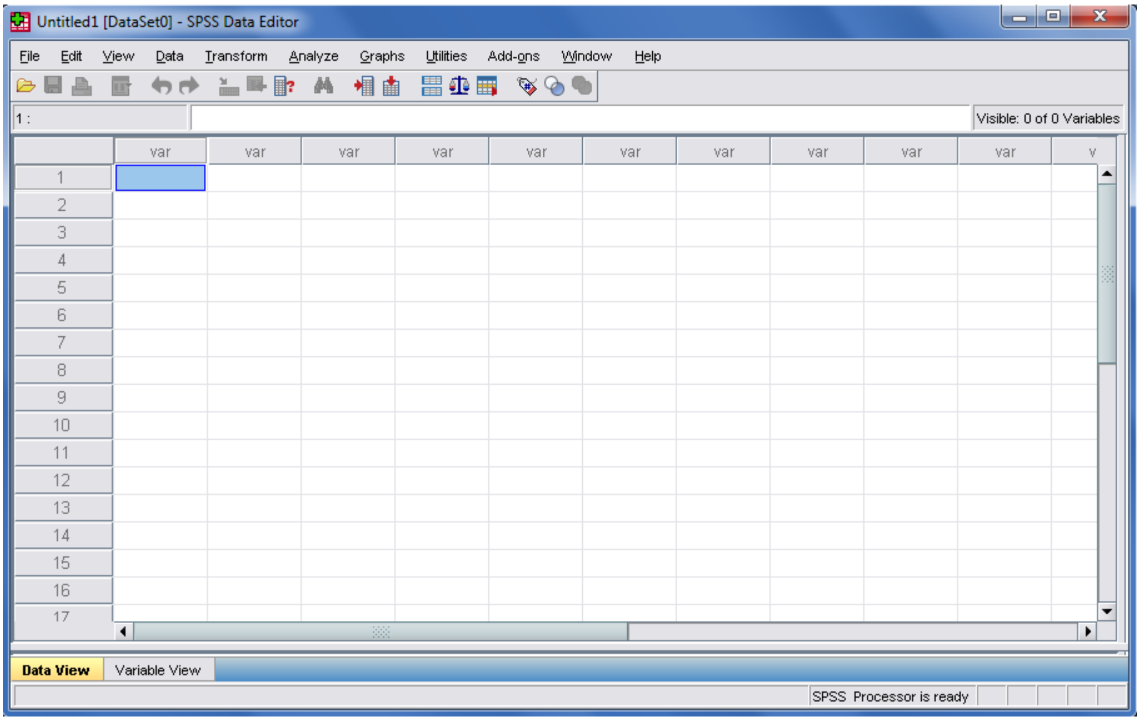
នៅក្នុង Editor ចែកចេញជាពីរផ្នែកគឺ Data View និង Variable View។

I- **Data View**: ប្រើសម្រាប់បញ្ចូលទិន្នន័យទៅតាមលក្ខណៈ design variable នីមួយៗ។

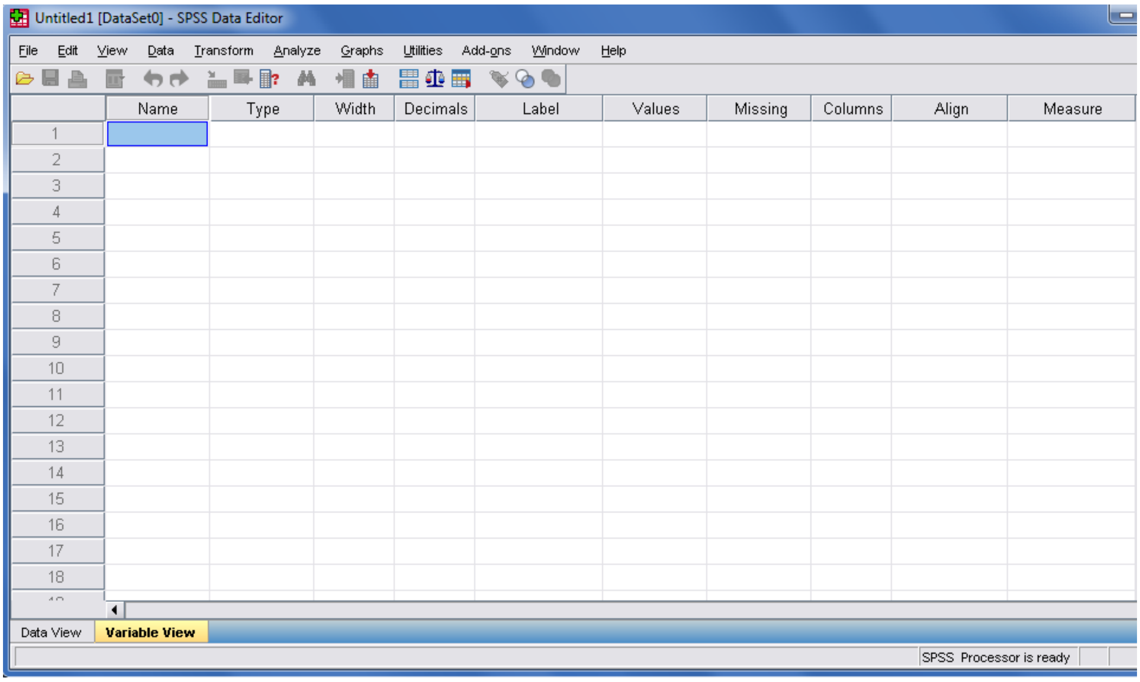
<u>សម្គាល់:</u>

a. យើង design នៅក្នុង Variable View ដើម្បីបាន Data View។

b. Design *ជួរដេកទី 1* ក្នុង Variable View យើងបាន *ជួរឈរទី 1* ក្នុង Data View។

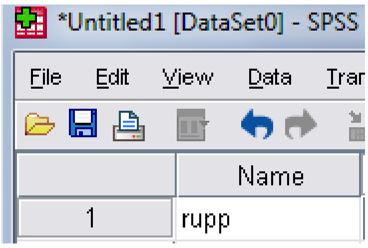


II- Variable View: ប្រើសម្រាប់ design ទិន្នន័យទៅតាមប្រភេទនៃសំណួរអង្កេត។

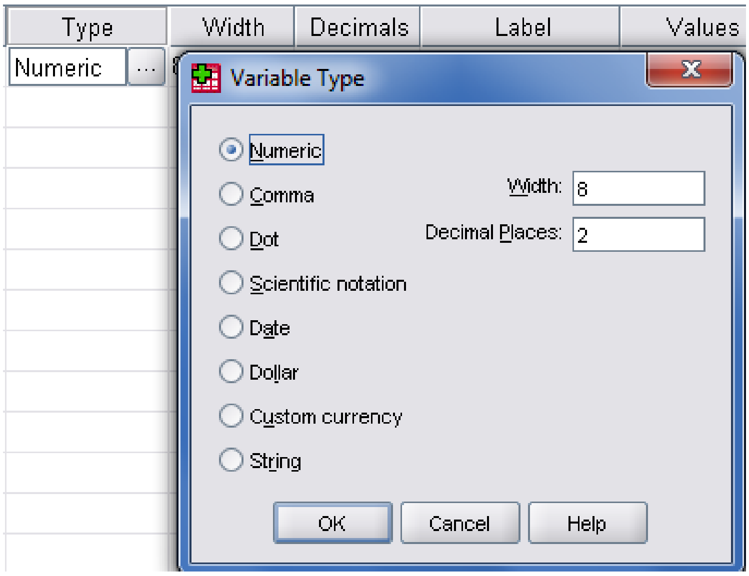


នៅក្នុង Variable View មាន ១០ ចំណុច:

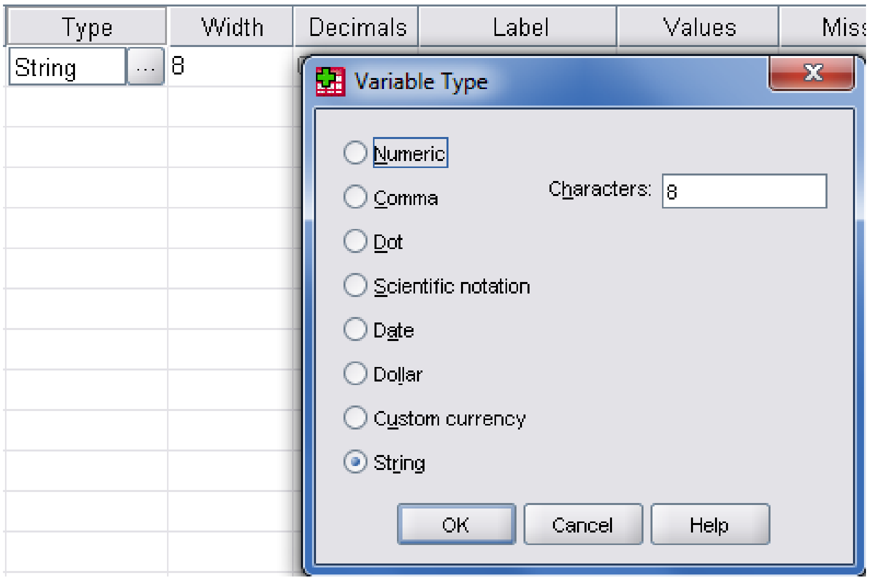
1. Name: ប្រើសម្រាប់ផ្តល់ឈ្មោះឲ្យ Variable Name នីមួយៗ។



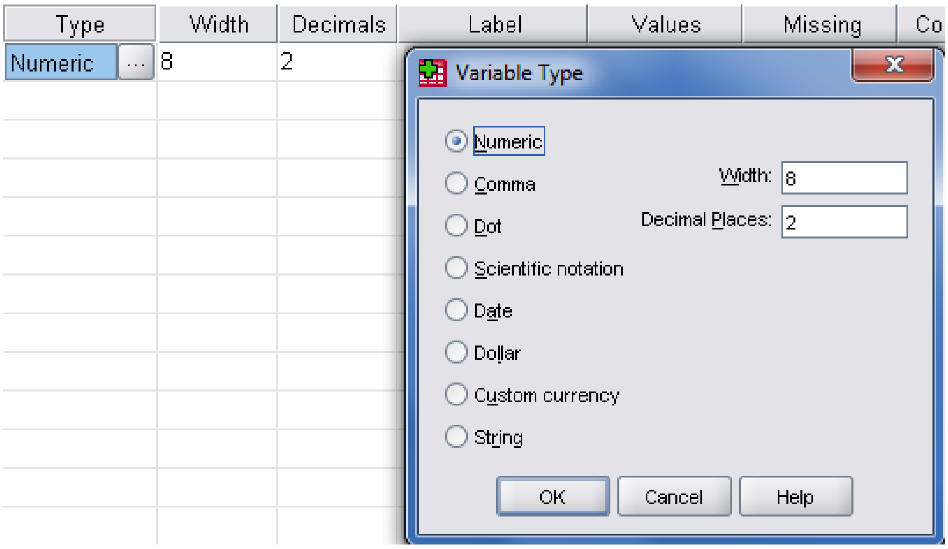
2. Type: ប្រើសម្រាប់សម្គាល់លក្ខណៈបញ្ចូលទិន្នន័យក្នុង Variable នីមួយៗ។



3. Width: ប្រើសម្រាប់កំណត់ចំនួនតួអក្សរដែលត្រូវបញ្ចូលក្នុង Cell នីមួយៗនៃ Variable ហើយវាមានឥទ្ធិពលចំពោះ Data Type ជា String។



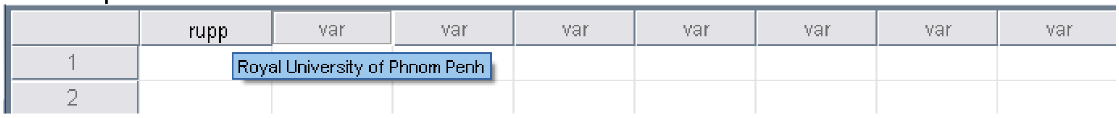
4. Decimals: ប្រើសម្រាប់កំណត់ចំនួនខ្ទង់ក្រោយក្បៀសនៃចំនួនមួយ ហើយវាមានឥទ្ធិពលចំពោះ Data Type ជា Numeric។



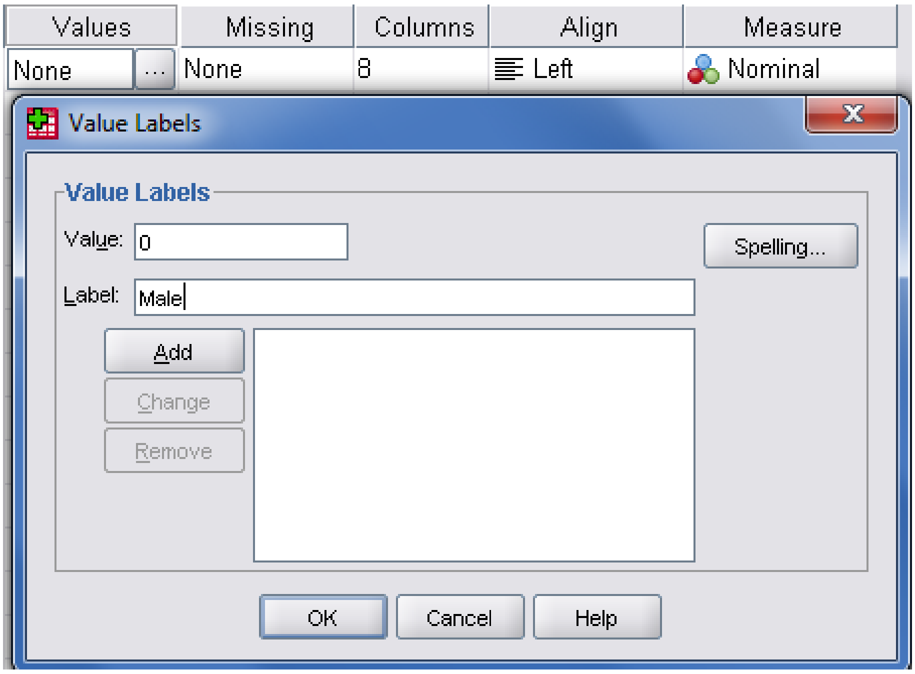
5. Label: ប្រើសម្រាប់ផ្តល់ Text ពន្យល់ឲ្យ Variable Name នីមួយៗ។

Name	Type	Width	Decimals	Label
rupp	Numeric	8	2	Royal University of Phnom Penh

ក្នុង Data View យើងនឹងឃើញ:



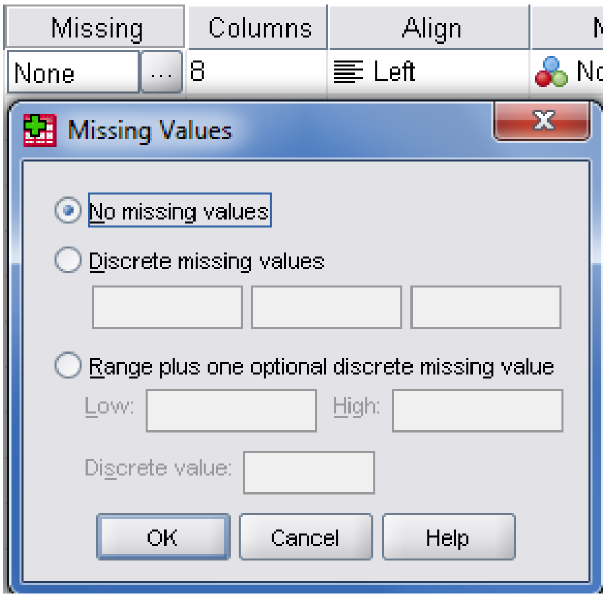
6. Values: ប្រើសម្រាប់កំណត់កូដទិន្នន័យទៅឲ្យ Variable Label។



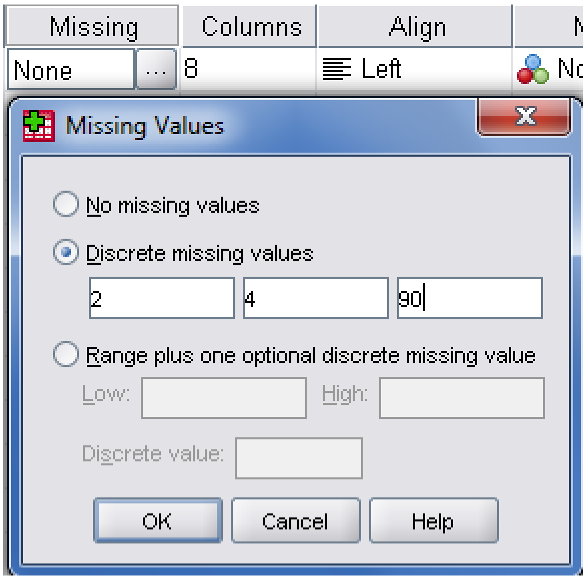
7. Missing: ប្រអប់ Missing Value ប្រើសម្រាប់បើក ឬបិទទិន្នន័យឲ្យចូលរួមវិភាគ។

នៅក្នុងប្រអប់ Missing មាន ៣ចំណុច:

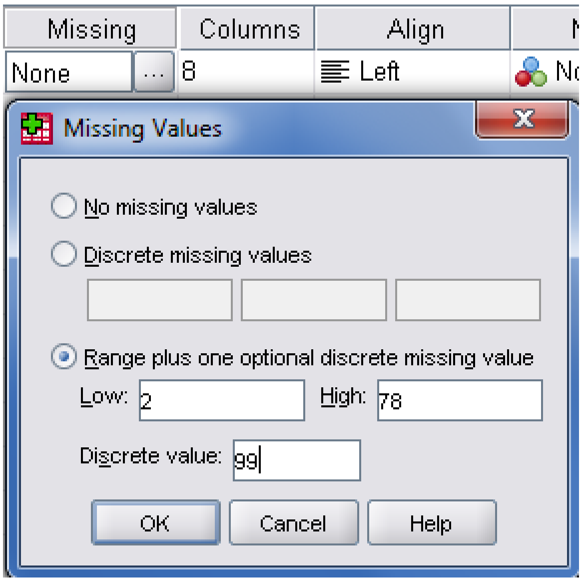
a. No missing values: ប្រើសម្រាប់បើកទិន្នន័យឲ្យចូលរួមវិភាគទាំងអស់។



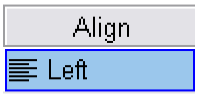
b. Discrete missing values: ប្រើសម្រាប់បិទទិន្នន័យមិនឲ្យចូលរួមវិភាគយ៉ាងច្រើនចំនួន ៣ លេខ។



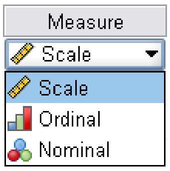
c. Range plus one optional discrete missing value: ប្រើសម្រាប់បិទទិន្នន័យមិនឲ្យចូលរួមវិភាគចំនួនមួយចន្លោះ និងបន្ថែម ១ លេខផ្សេងទៀត។



8. Columns: ប្រើសម្រាប់ពង្រីកជួរឈរនៃ Data View។
9. Align: ប្រើសម្រាប់កំណត់ទីតាំងទិន្នន័យនៃជួរឈររបស់ Data View ឲ្យនៅឆ្វេង ខាងស្តាំ កណ្តាល។



10. Measure: ប្រើសម្រាប់កំណត់រង្វាស់ទិន្នន័យ។ នៅក្នុង Measure មាន 3 ចំណុច:



a. **Scale** ទិន្នន័យជាប្រភេទ Quantitative (លេខ) ដែលអាចធ្វើប្រមាណវិធីពិជគណិតបាន។

b. **Ordinal** ទិន្នន័យជាប្រភេទ Qualitative (សភាពលក្ខណៈអ្វីមួយ) ដែលមិនអាចធ្វើប្រមាណវិធីពិជគណិតបាន ហើយ**មាន**លំដាប់ប្រៀបធៀប។

c. **Nominal** ទិន្នន័យជាប្រភេទ Qualitative (សភាពលក្ខណៈអ្វីមួយ) ដែលមិនអាចធ្វើប្រមាណវិធីពិជគណិតបាន ហើយ**គ្មាន**លំដាប់ប្រៀបធៀប។

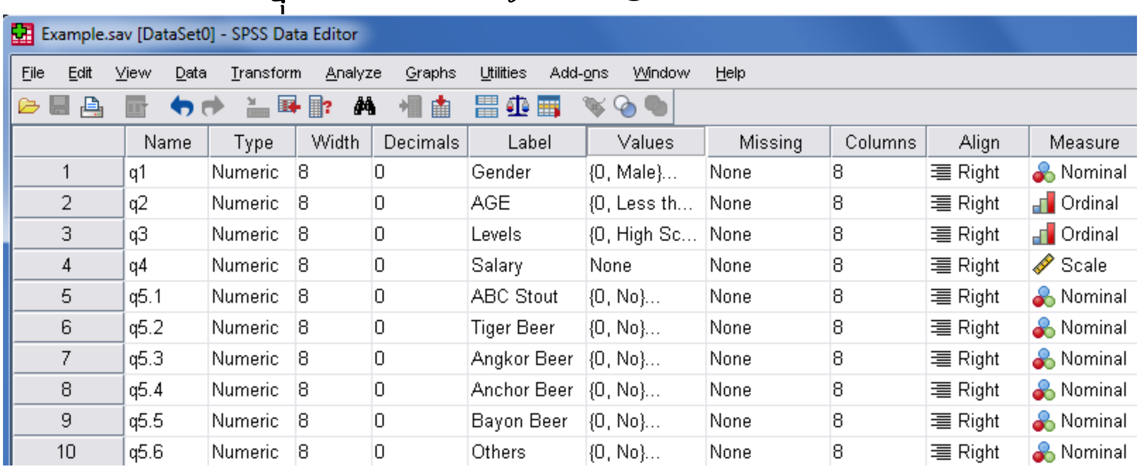
ឧទាហរណ៍: ចូរ design variable របស់ SPSS តាមប្រភេទសំណួរនីមួយៗខាងក្រោម:

1. តើអ្នកមានភេទអ្វី? (Nominal)
   ☐ Male ☐ Female
2. តើអ្នកមានអាយុប៉ុន្មាន? (Ordinal)
   ☐ Less than 30 ☐ 30-40 ☐ 40-50 ☐ More than 50
3. តើអ្នកមានកម្រិតវប្បធម៌អ្វី? (Ordinal)
   ☐ High School ☐ BA ☐ MA ☐ PhD
4. តើអ្នកមាន Salary ប៉ុន្មាន? … (Scale)
5. តើអ្នកចូលចិត្តពិសារស្រាបៀរណាខ្លះ? (Nominal)
   ☐ ABC Stout ☐ Tiger Beer ☐ Angkor Beer
   ☐ Anchor Beer ☐ Bayon Beer ☐ Others

សម្គាល់: នៅក្នុងករណីជួបប្រទះសំណួរអង្កេតដែលមានជម្រើសច្រើន (សំណួរទី ៥) គេត្រូវ:

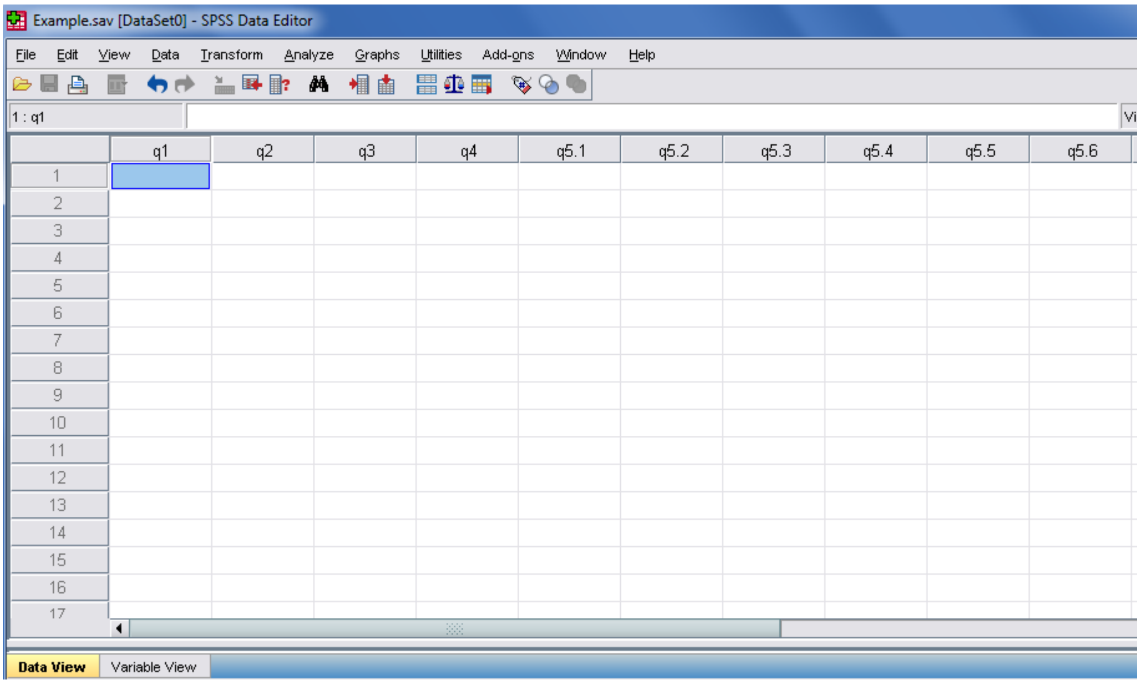
- យកប្រអប់ជម្រើសនីមួយៗមកបង្កើតជា variable។
- Variable នីមួយៗមានលេខកូដ 0=No, 1=Yes។
- Variable នីមួយៗមានរង្វាស់ទិន្នន័យជាប្រភេទ Nominal ។

យើង design ក្នុង Variable View ដូចខាងក្រោម:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	q1	Numeric	8	0	Gender	{0, Male}...	None	8	Right	Nominal
2	q2	Numeric	8	0	AGE	{0, Less th...	None	8	Right	Ordinal
3	q3	Numeric	8	0	Levels	{0, High Sc...	None	8	Right	Ordinal
4	q4	Numeric	8	0	Salary	None	None	8	Right	Scale
5	q5.1	Numeric	8	0	ABC Stout	{0, No}...	None	8	Right	Nominal
6	q5.2	Numeric	8	0	Tiger Beer	{0, No}...	None	8	Right	Nominal
7	q5.3	Numeric	8	0	Angkor Beer	{0, No}...	None	8	Right	Nominal
8	q5.4	Numeric	8	0	Anchor Beer	{0, No}...	None	8	Right	Nominal
9	q5.5	Numeric	8	0	Bayon Beer	{0, No}...	None	8	Right	Nominal
10	q5.6	Numeric	8	0	Others	{0, No}...	None	8	Right	Nominal

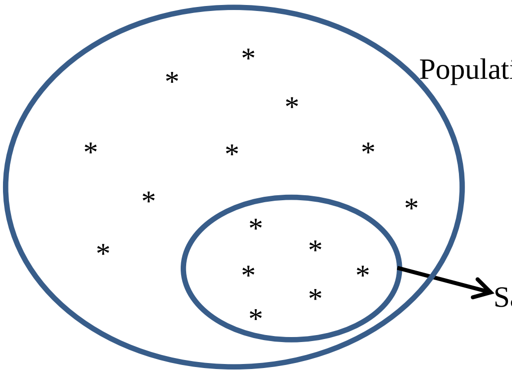
យើងបាន Data View មានទម្រង់ដូចខាងក្រោម:



Frequencies Analysis

Frequencies Analysis គឺជាការសិក្សាវិភាគទៅលើទិន្នន័យប្រភេទ Qualitative ប្រើសម្រាប់វាយតម្លៃជាភាគរយ។

ក្រាបដែលបកស្រាយទៅលើវាមាន: Pie chart, Bar chart។



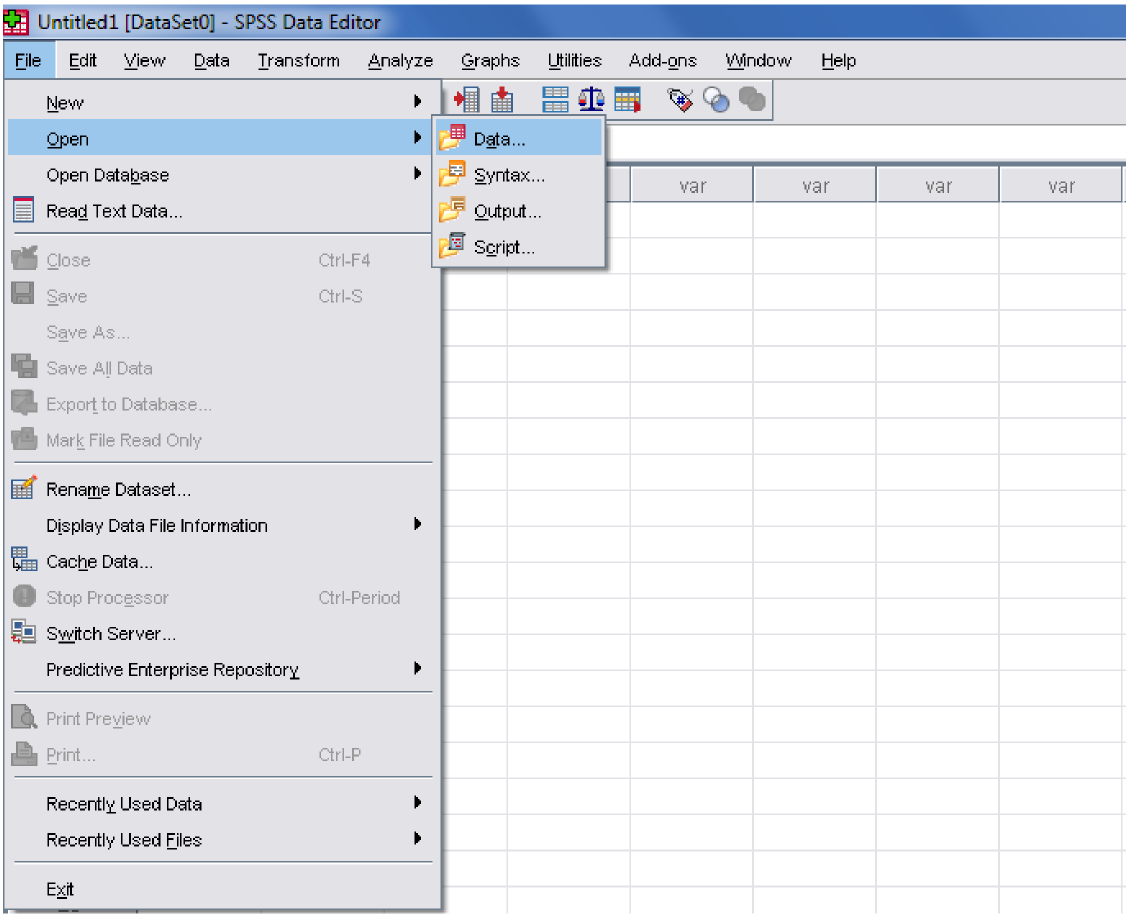
Population គឺជាបណ្តុំនៃព័ត៌មានទាំងឡាយអាចជា មនុស្ស សត្វ វត្ថុ ទីកន្លែង... ។

Sample គឺជាផ្នែកនៃ Population ដែលគេយកមកធ្វើតេស្ត។

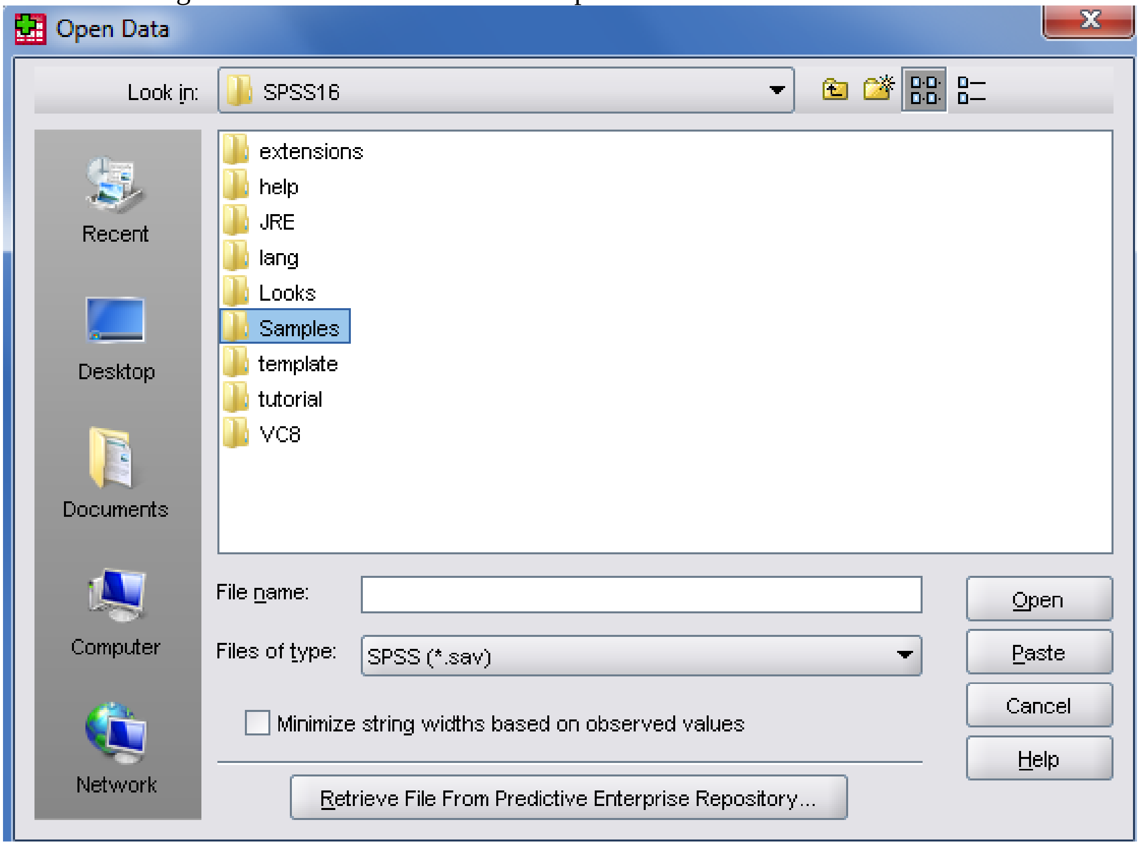
ឧទាហរណ៍: ដោយអនុវត្តទៅលើតារាងទិន្នន័យ Cars។ ចូរវាយតម្លៃជាភាគរយទៅលើប្រភេទរថយន្តដែលគេនិយមប្រើប្រាស់ជាងគេបំផុត? (ប្រើជួរឈរដែលមានឈ្មោះ Origin ➡ Label: Country of origin)

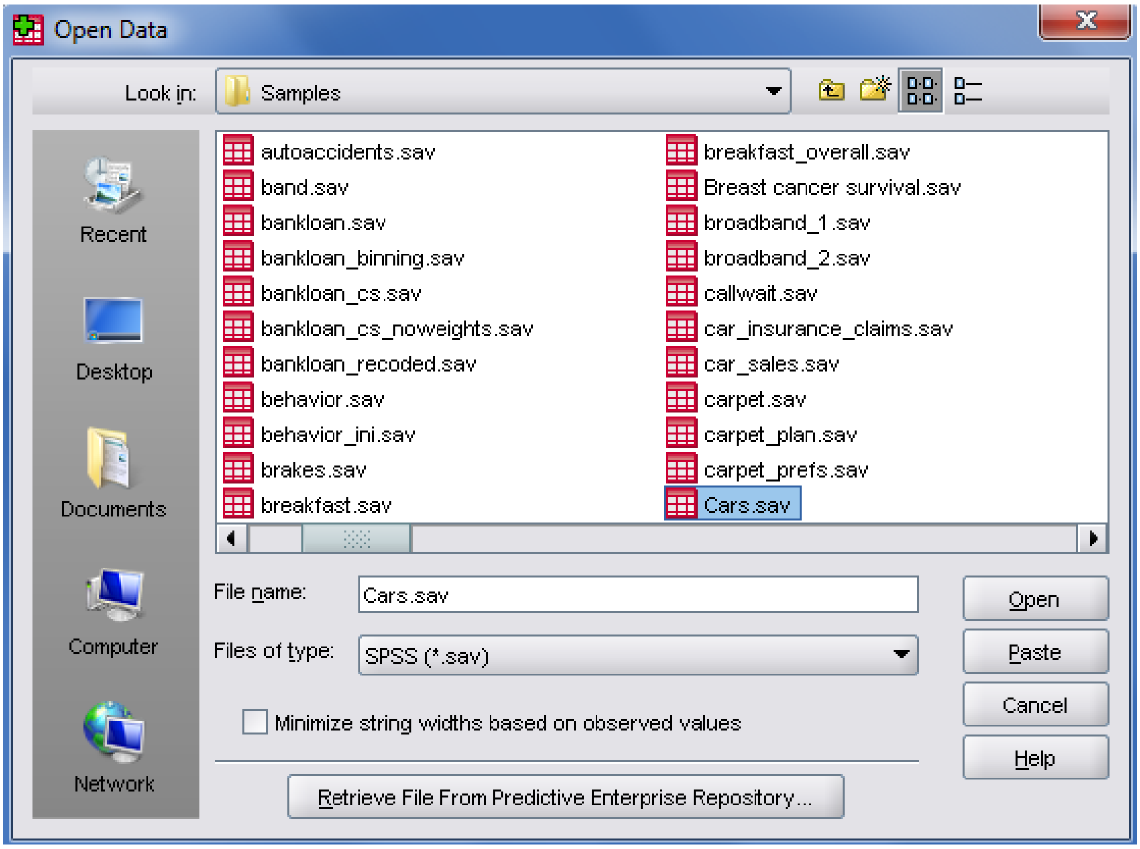
Step 1: “Cars”

File ➡ Open ➡ Data ➡ Cars



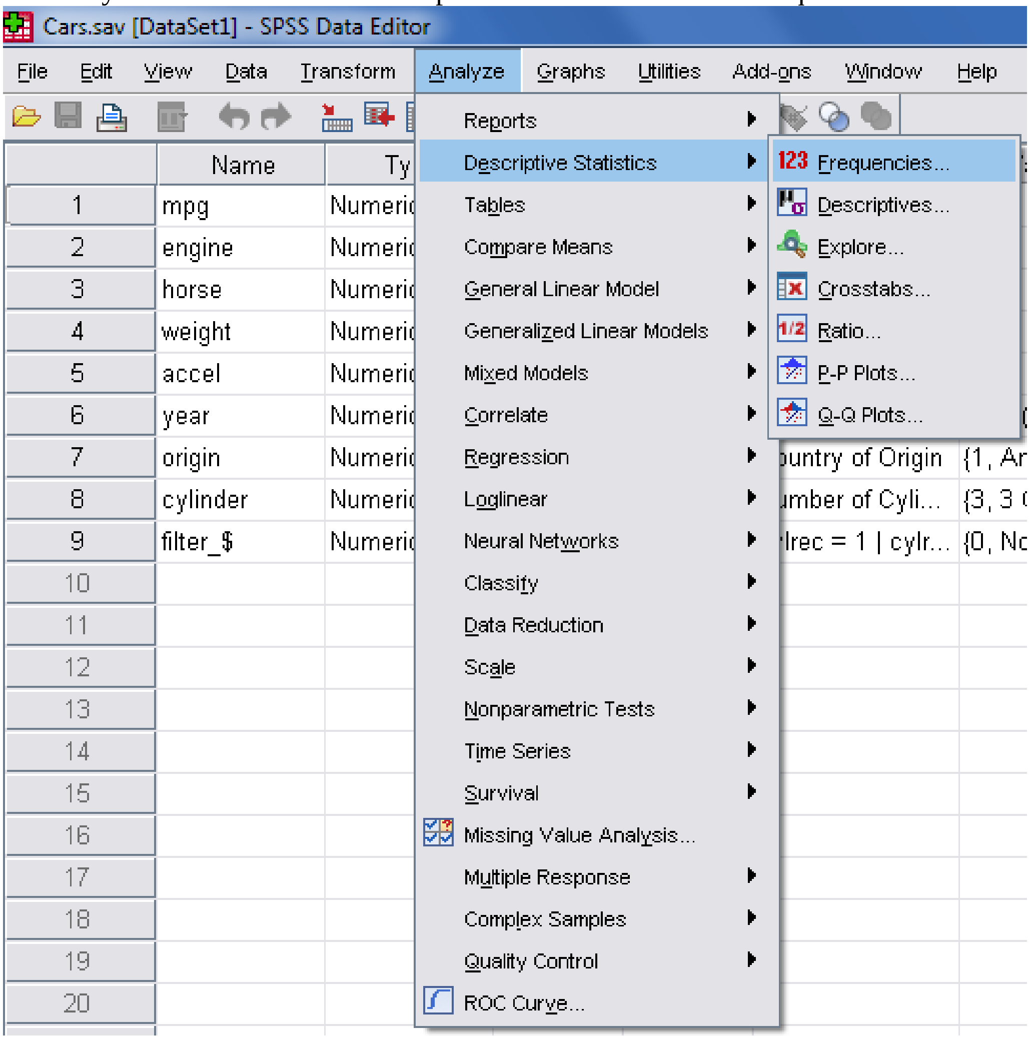
C:\Program Files\SPSSInc\SPSS16\Samples\Cars.sav

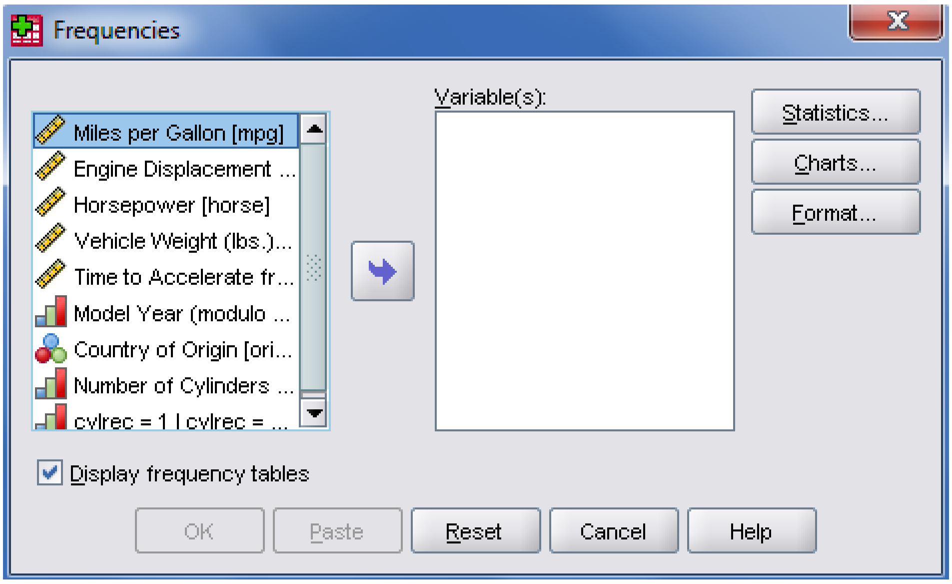


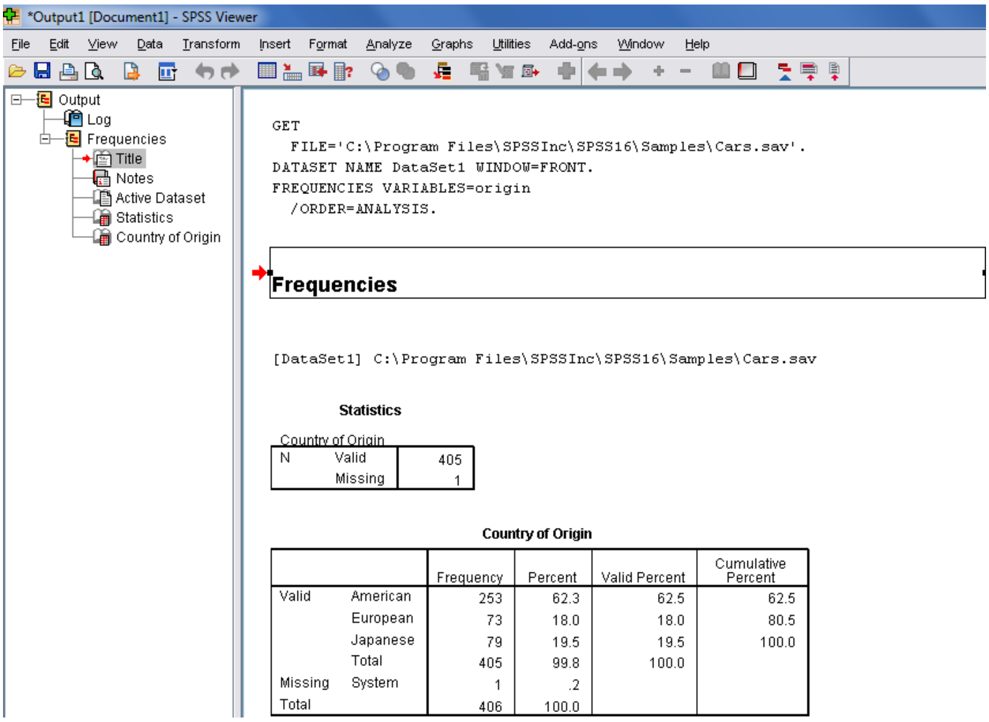


Step 2: Analysis

Analyze ➡Descriptive Statistics ➡ Frequencies








```
GET
  FILE='C:\Program Files\SPSSInc\SPSS16\Samples\Cars.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
FREQUENCIES VARIABLES=origin
  /ORDER=ANALYSIS.
```

Frequencies

```
[DataSet1] C:\Program Files\SPSSInc\SPSS16\Samples\Cars.sav
```

Statistics

Country of Origin

N	Valid	405
	Missing	1

Country of Origin

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	American	253	62.3	62.5	62.5
	European	73	18.0	18.0	80.5
	Japanese	79	19.5	19.5	100.0
	Total	405	99.8	100.0	
Missing	System	1	.2		
Total		406	100.0		

Country of Origin

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	American	253	62.3	62.5	62.5
	European	73	18.0	18.0	80.5
	Japanese	79	19.5	19.5	100.0
	Total	405	99.8	100.0	
Missing	System	1	.2		
Total		406	100.0		

N Valid=ចំនួនទិន្នន័យបានការ (405 គ្រឿងបានទទួលទិន្នន័យច្បាស់លាស់)

Missing=ចំនួនទិន្នន័យមិនបានការ (1 គ្រឿងមិនបានទទួលទិន្នន័យច្បាស់លាស់)

Percent=Frequency/Total
Valid Percent=Frequency/(Total−Missing)
Cumulative Percent=ភាគរយបូកបន្ត (អនុគមន៍របាយ)

ចម្លើយ:

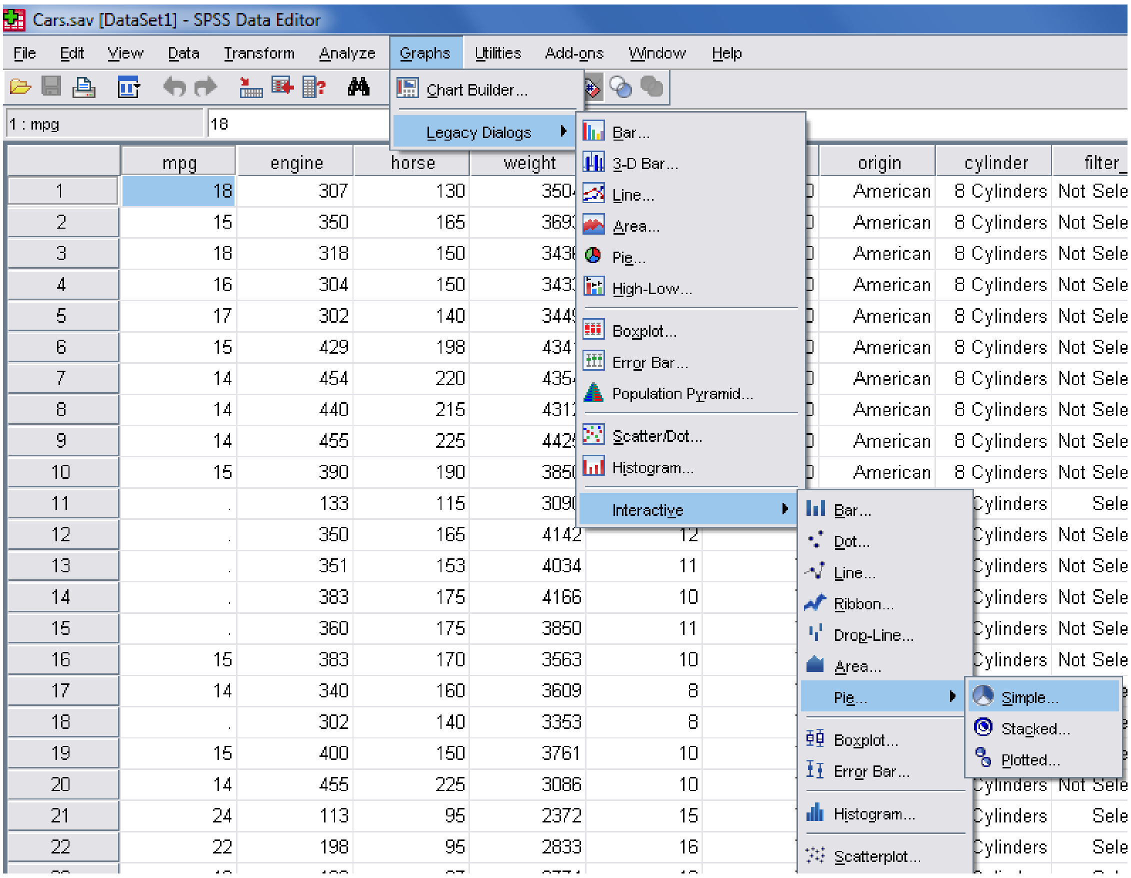
តាមតារាង Output ខាងលើបានបញ្ជាក់ឲ្យឃើញថា នៅក្នុងតំបន់នោះ គេប្រើប្រាស់រថយន្តបីប្រភេទគឺ:

- ប្រភេទរថយន្ត American មាន 62.50%
- ប្រភេទរថយន្ត European មាន 18.00%
- ប្រភេទរថយន្ត Japanese មាន 19.50%

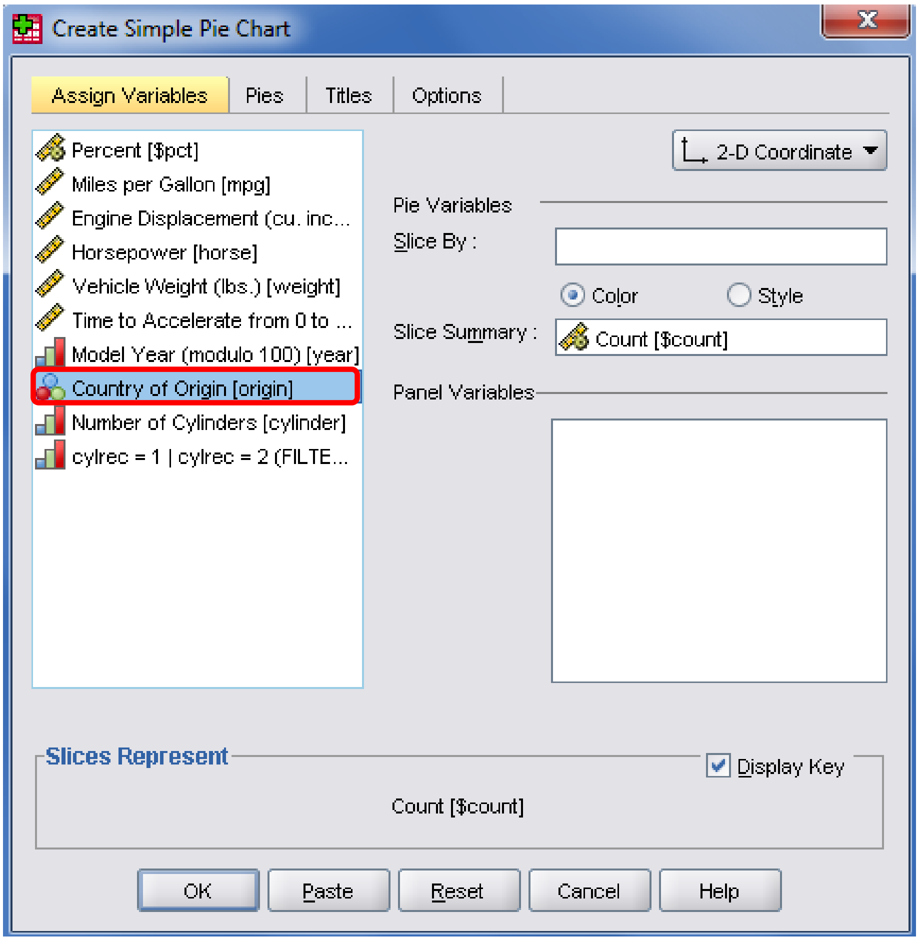
សន្និដ្ឋាន: នៅក្នុងតំបន់នោះ គេនិយមប្រើប្រាស់ប្រភេទរថយន្ត American ច្រើនជាងគេបំផុតដែលមានរហូតដល់ 62.50%។

សង់ Pie chart

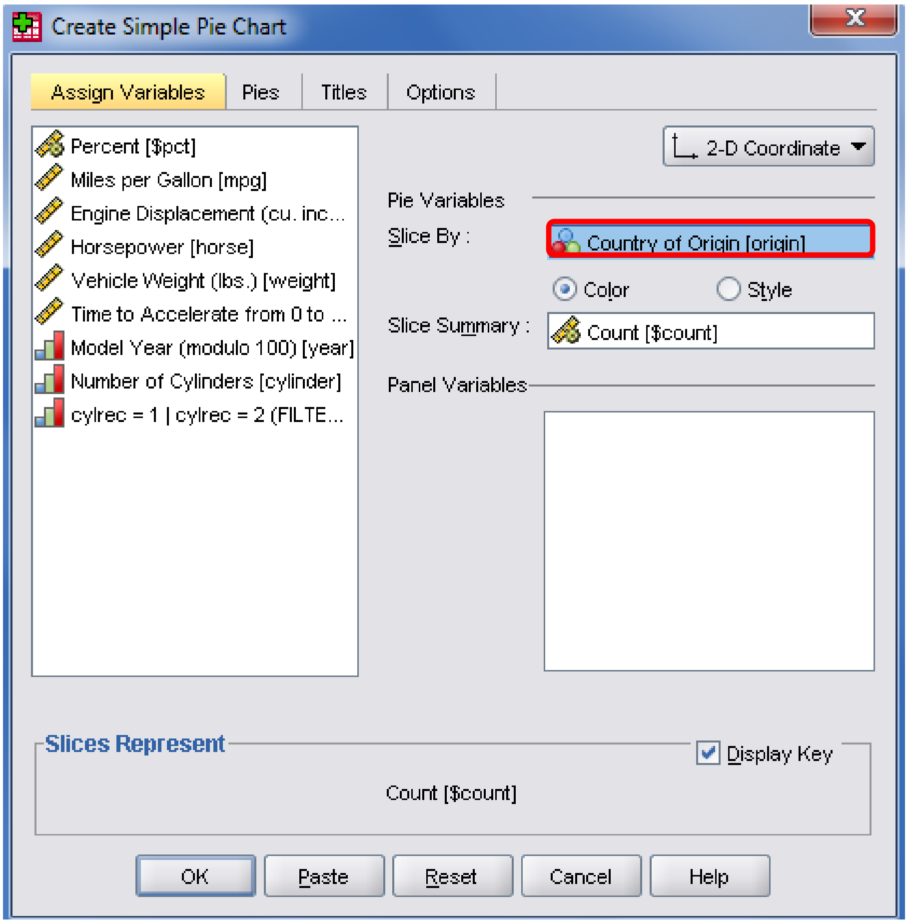
File ➡ Open ➡ "Cars"
Graphs ➡ Legacy Dialogs ➡ Interactive ➡ Pie ➡ Simple



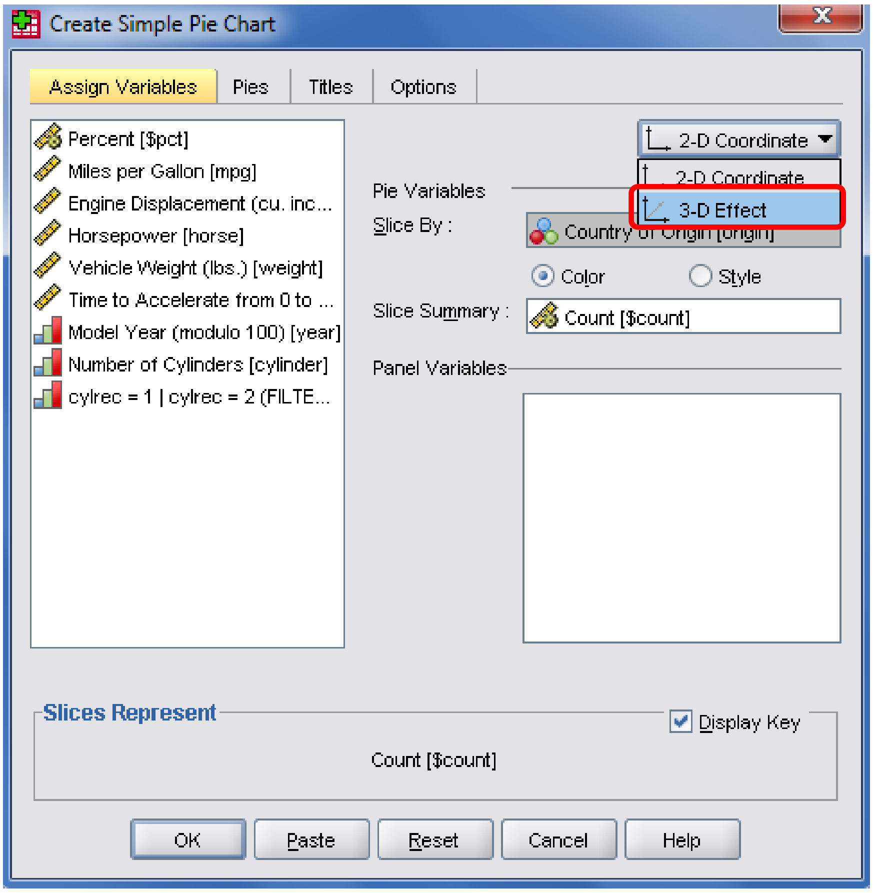
បន្ទាប់មក វានឹងចេញផ្ទាំងខាងក្រោម:



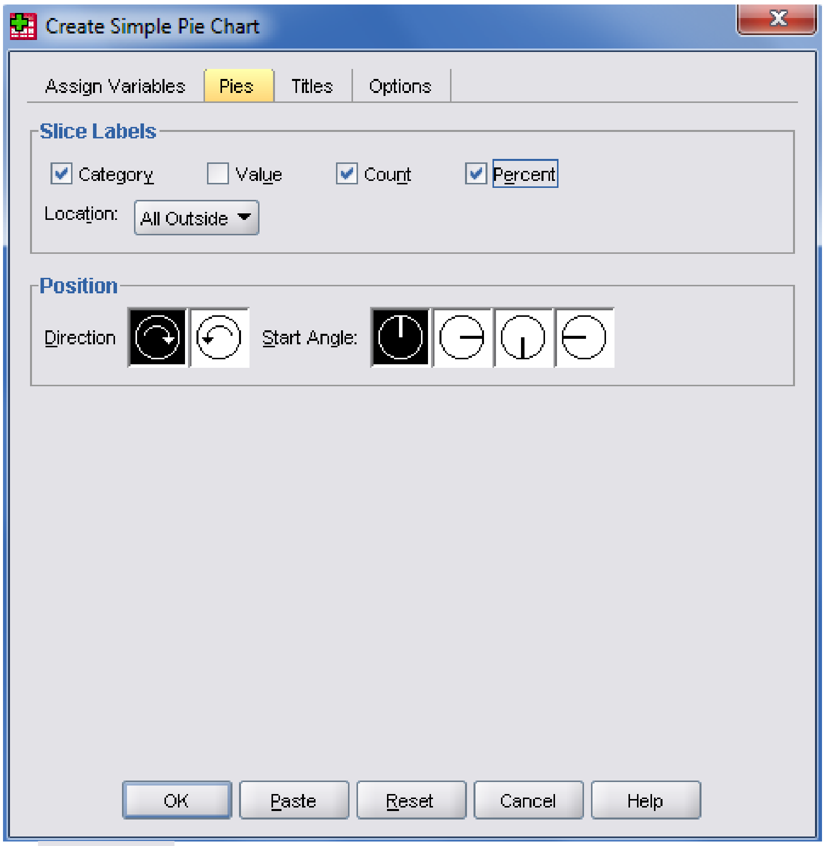
ចុចលើ Country of Origin [origin] ទាញយកទៅដាក់ក្នុង Slice By:



ជ្រើសយក 3D-Effect



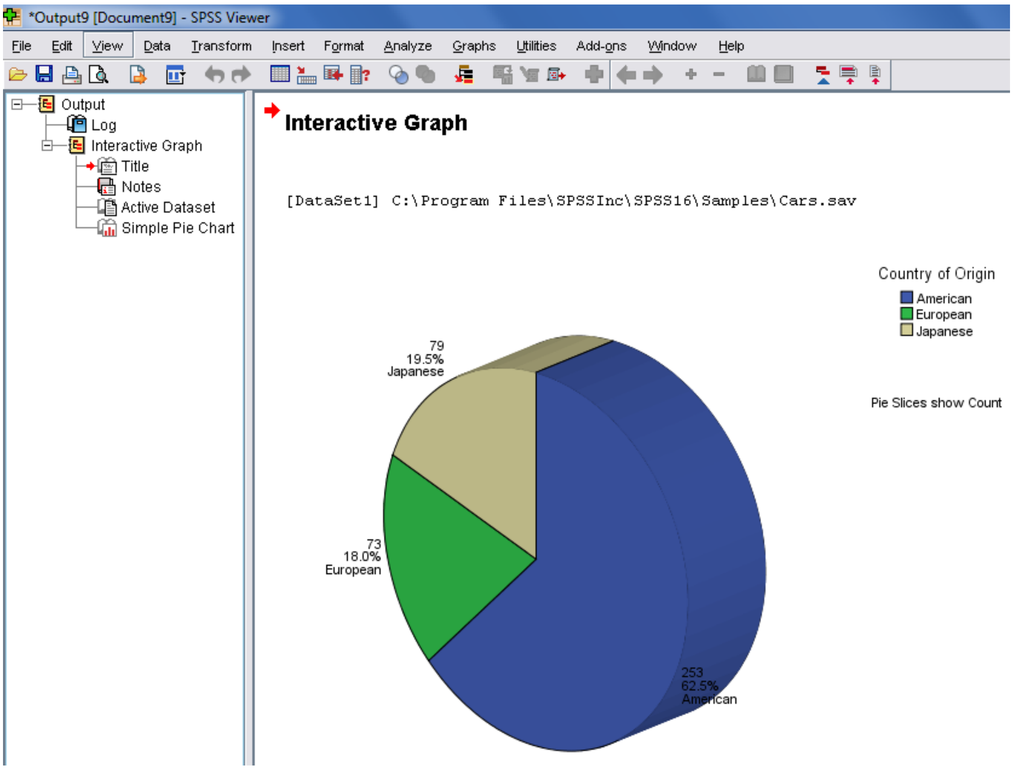
ចុចលើ Pies Tab:



សម្គាល់:

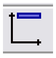
☐ Category	ដាក់ឈ្មោះ
☐ Value	តម្លៃ
☐ Count	រាប់ចំនួន
☐ Percent	ភាគរយ

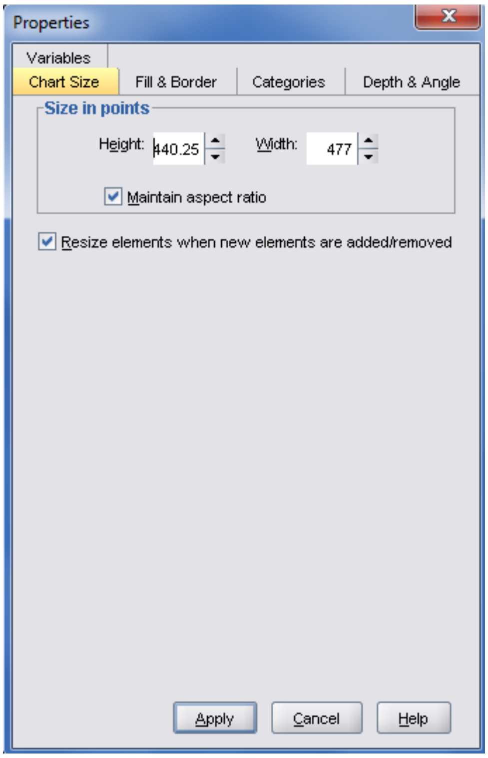
កំណត់រួចរាល់ ចុច OK យើងនឹងឃើញរូបដូចខាងក្រោម:

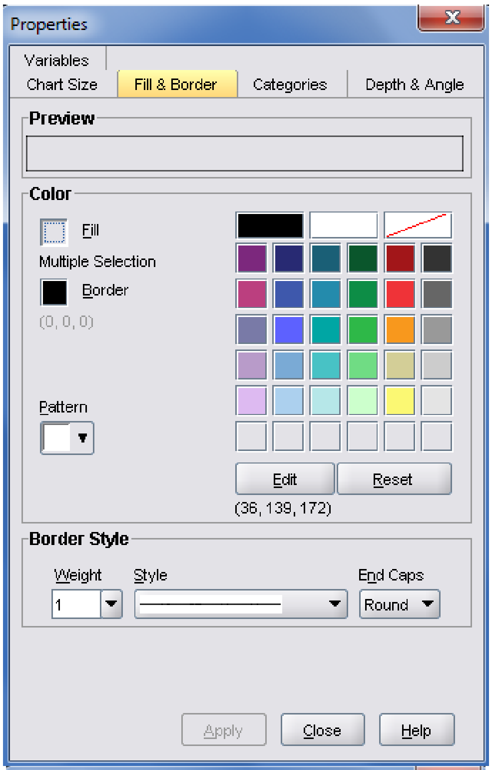


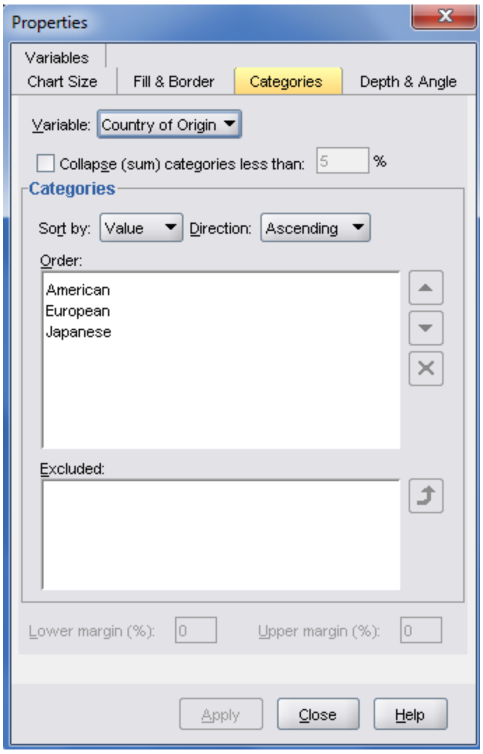
សម្រាប់ការ design ផ្សេងៗ : Double Click

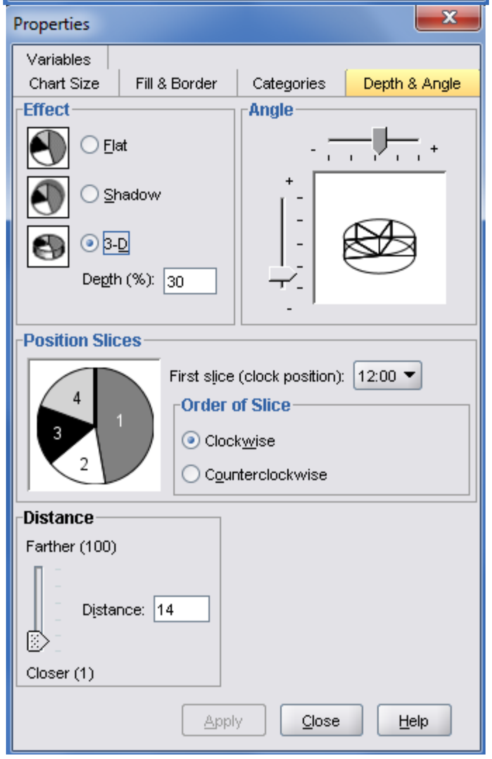
ក្នុង Chart Editor យើងអាច:

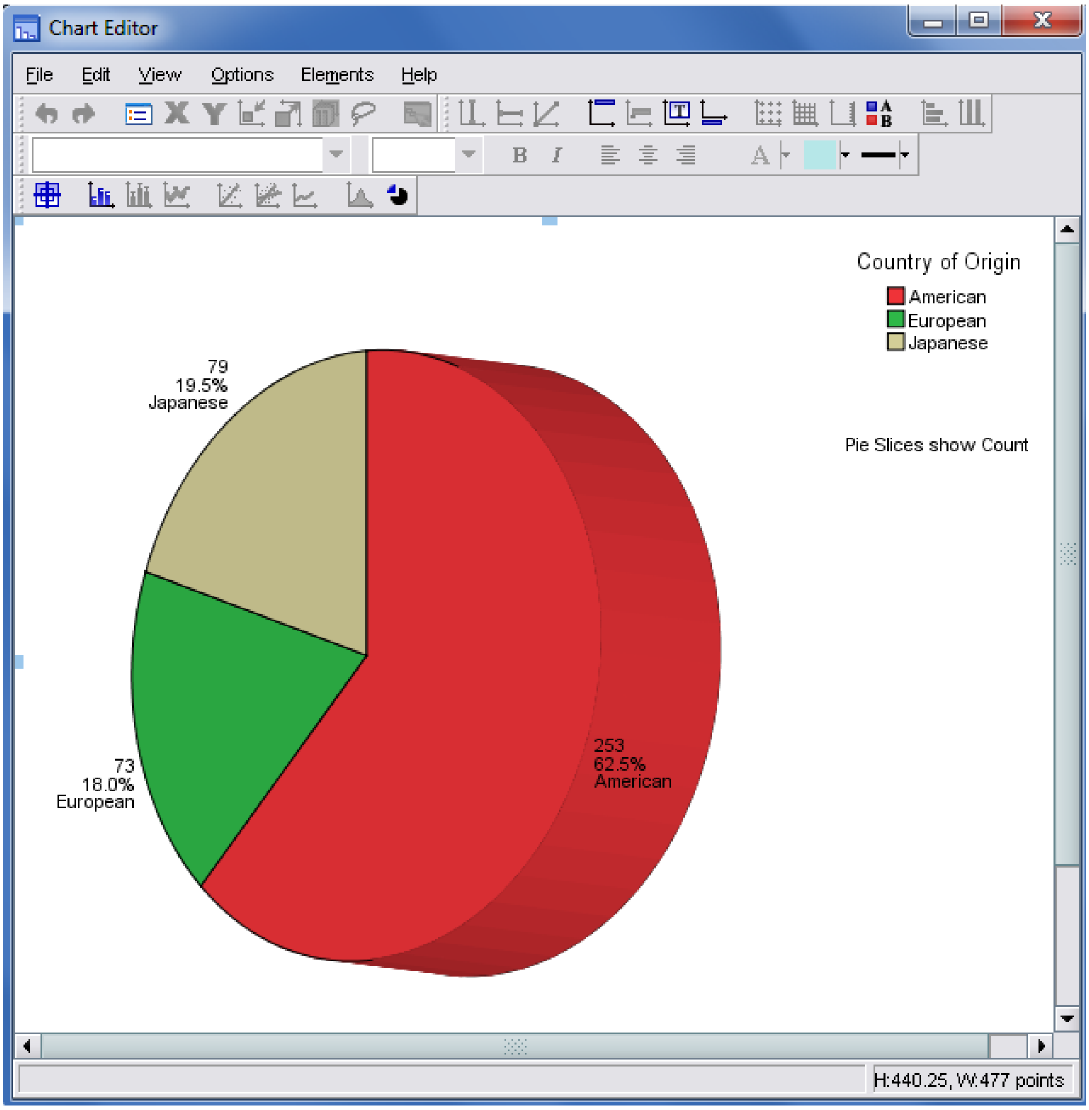
- ដាក់ចំណងជើងលើ chart 
- ប្តូរពណ៌ផ្នែកនីមួយៗ ដោយគ្រាន់តែ select ផ្នែកដែលចង់ប្តូរ ហើយ Right click យកពាក្យ Properties Window (Ctrl+T)
- បង្វិល pie chart …









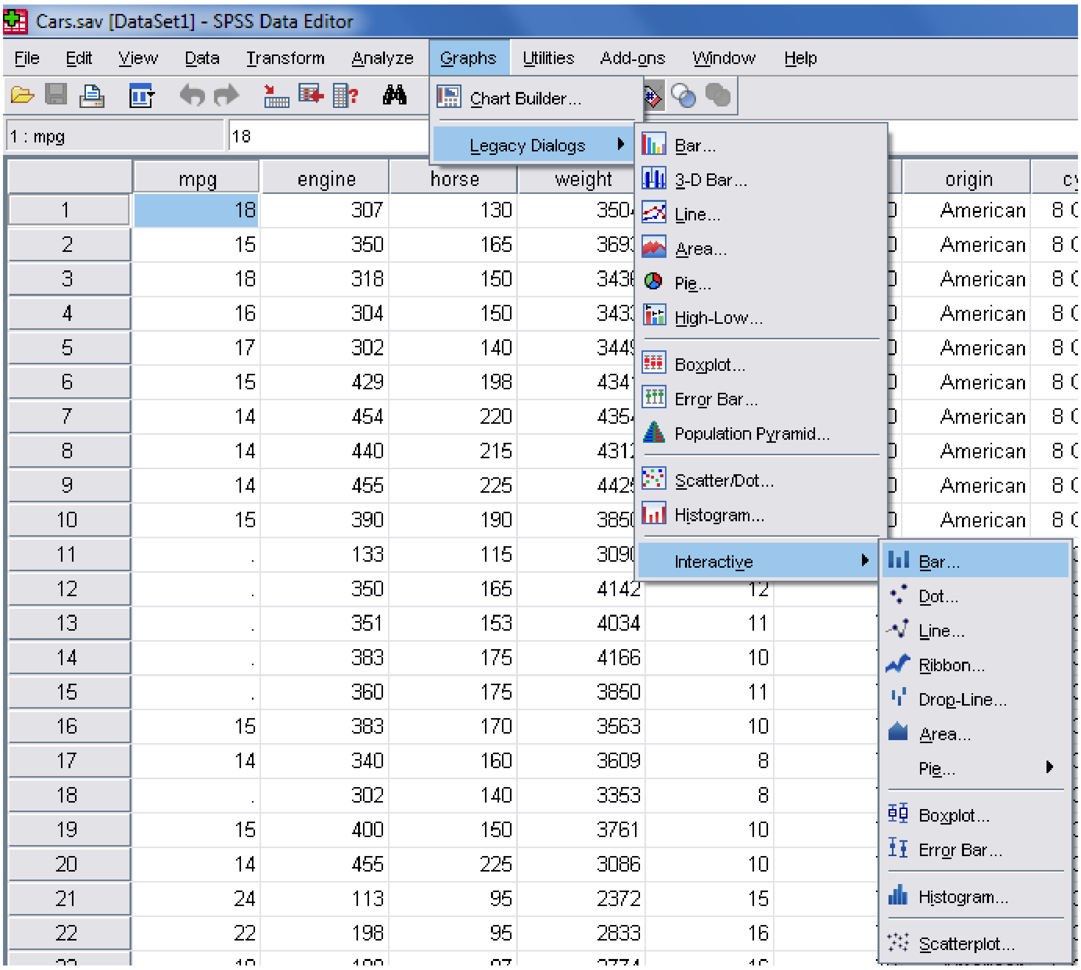


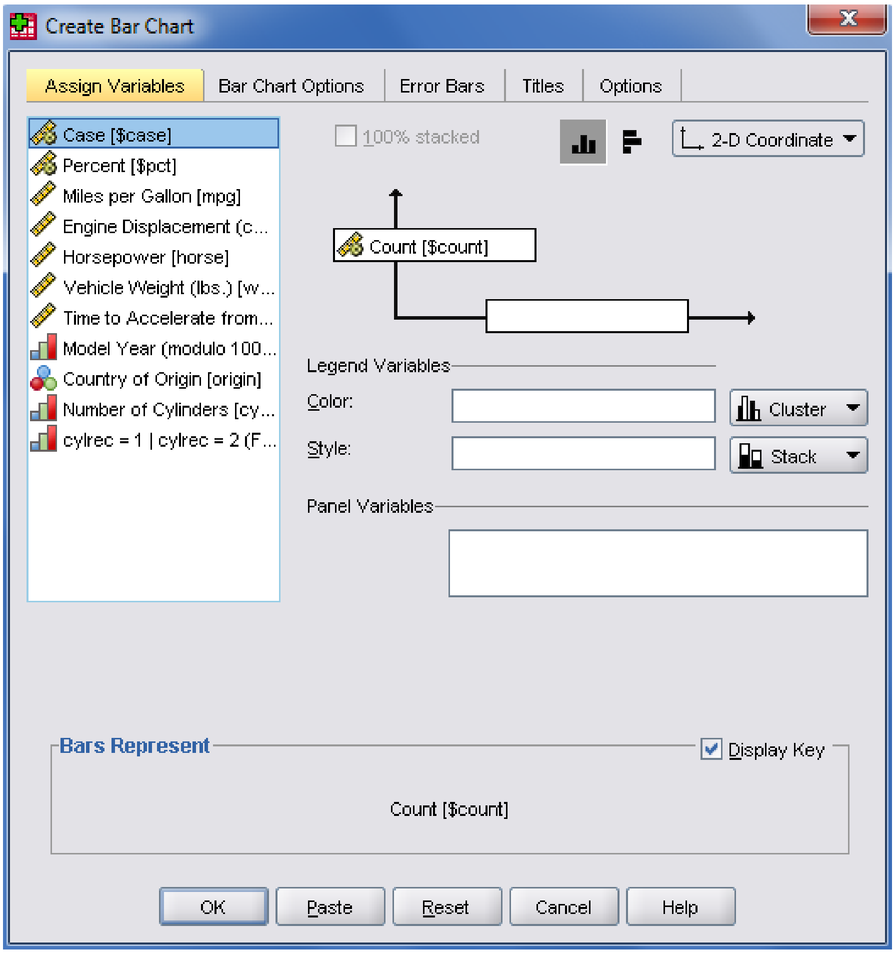
ចំណកស្រាយ: តាមក្រាភិចខាងលើបង្ហាញឲ្យឃើញថា:

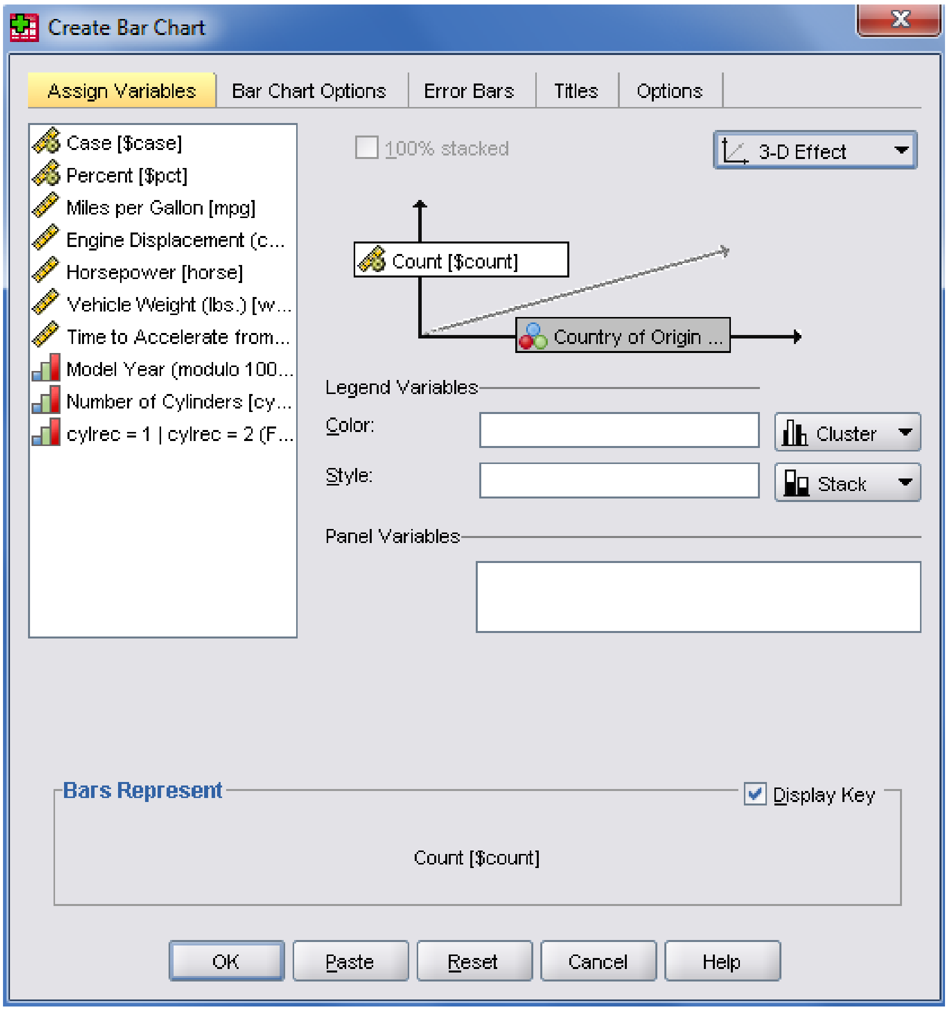
- រថយន្ត American មានចំនួន 253 គ្រឿងត្រូវជា 62.5%
- រថយន្ត European មានចំនួន 73 គ្រឿងត្រូវជា 18%
- រថយន្ត Janpanese មានចំនួន 79 គ្រឿងត្រូវជា 19.5%

៥ Bar chart

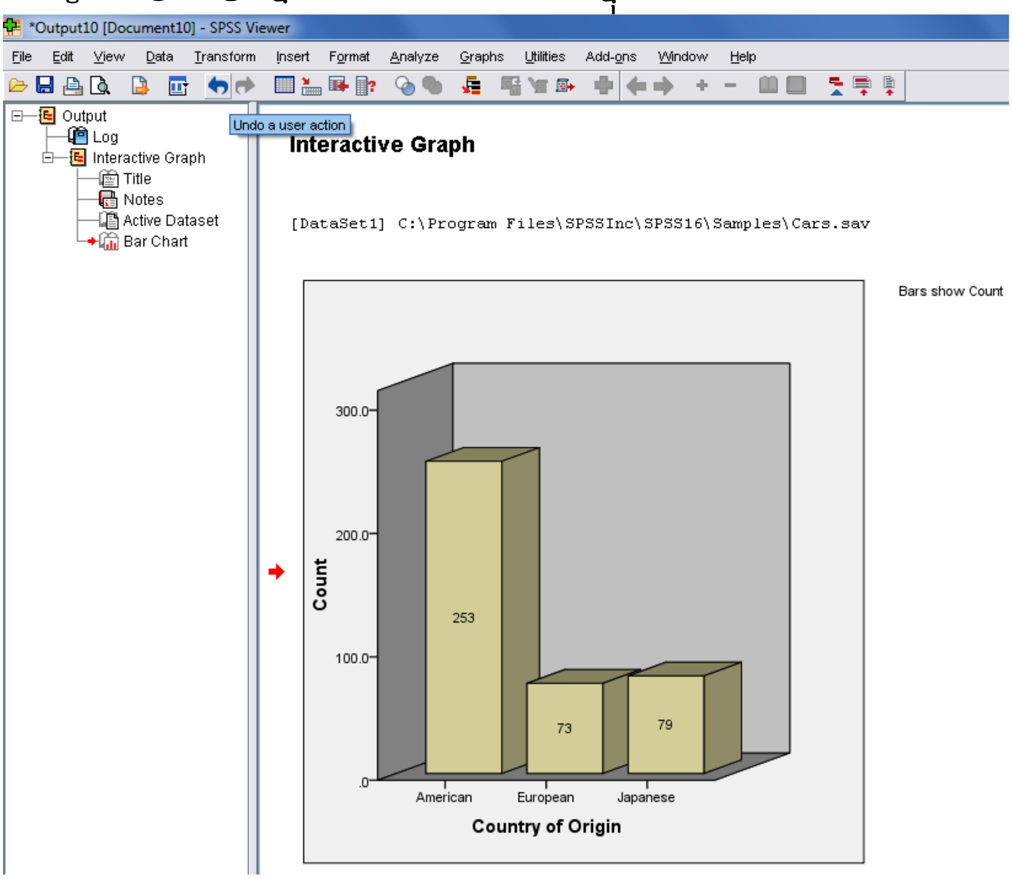
Graphs ➡ Legacy Dialogs ➡ Interactive ➡ Bar


Cars.sav [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
Chart Builder...
Legacy Dialogs
Bar...
3-D Bar...
Line...
Area...
Pie...
High-Low...
Boxplot...
Error Bar...
Population Pyramid...
Scatter/Dot...
Histogram...
Interactive
Bar...
Dot...
Line...
Ribbon...
Drop-Line...
Area...
Pie...
Boxplot...
Error Bar...
Histogram...
Scatterplot...
1 : mpg
mpg engine horse weight origin
American


Create Bar Chart
Assign Variables
Bar Chart Options
Error Bars
Titles
Options
Case [$case]
Percent [$pct]
Miles per Gallon [mpg]
Engine Displacement (c...
Horsepower [horse]
Vehicle Weight (lbs.) [w...
Time to Accelerate from...
Model Year (modulo 100...
Country of Origin [origin]
Number of Cylinders [cy...
cylrec = 1 | cylrec = 2 (F...
100% stacked
2-D Coordinate
Count [$count]
Legend Variables
Color:
Cluster
Style:
Stack
Panel Variables
Bars Represent
Display Key
Count [$count]
OK
Paste
Reset
Cancel
Help



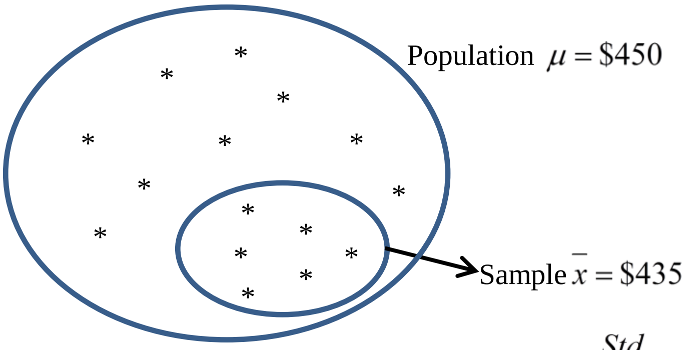
វាមានលក្ខណៈស្រដៀងគ្នាទៅនឹងការ Design នៅក្នុង Pie chart ដែរ។



Descriptive Analysis

Descriptive Analysis គឺជាការសិក្សាវិភាគទៅលើទិន្នន័យ Quantitative (លេខ) ប្រើសម្រាប់វាយតម្លៃមធ្យម។ Output របស់វាមាន: Min, Max, Mean, SEmean (Standard Error), …។

ក្រាបដែលបកស្រាយលើវាមាន: Histogram & Boxplot។



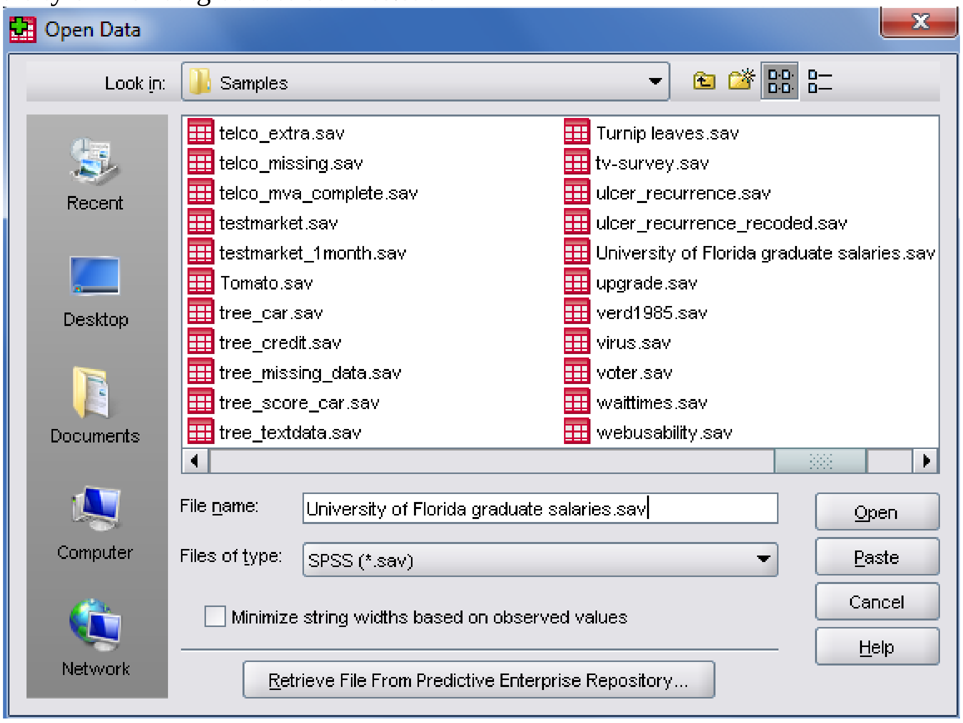
$$\text{SEmean} = \frac{Std}{\sqrt{n}} = \$15$$

SEmean គឺជា Error ជាមធ្យមដែលបានកើតឡើងក្នុងការវាយតម្លៃមធ្យម Sample ទៅឲ្យមធ្យម Population។

<u>លំហាត់</u>

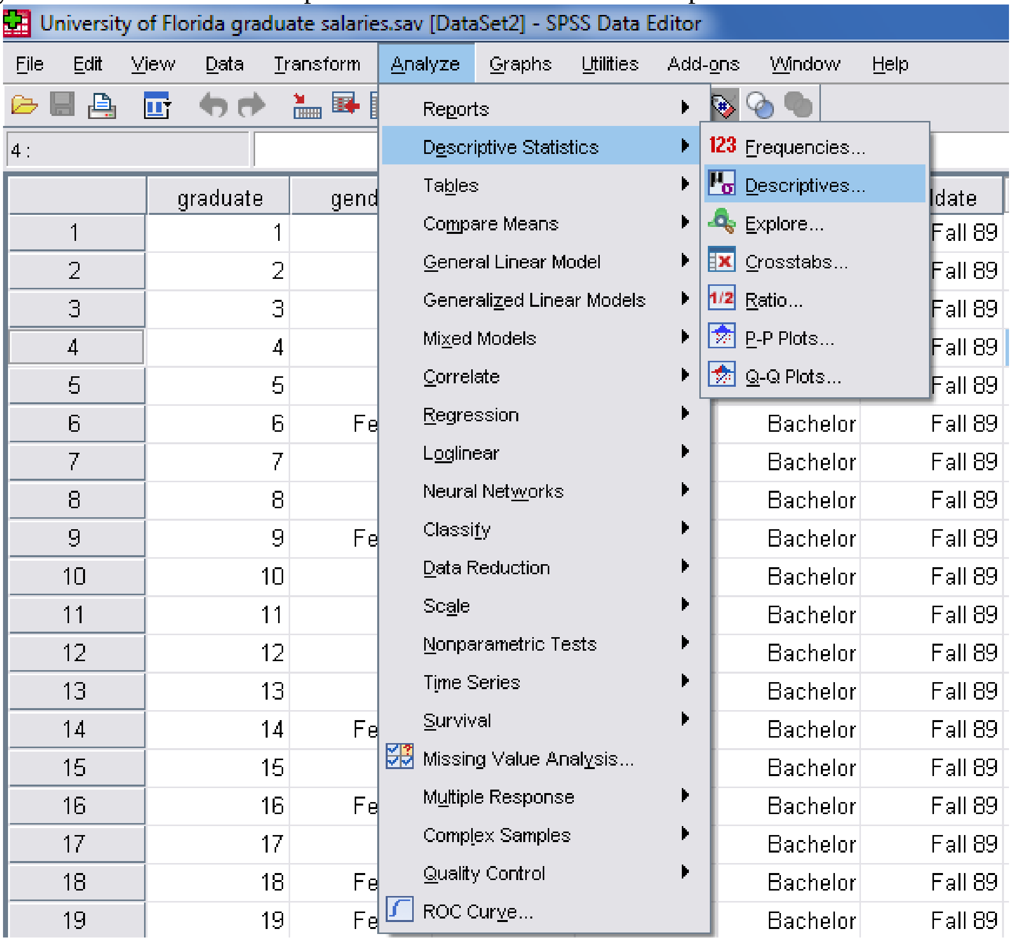
ដោយអនុវត្តទៅលើតារាងទិន្នន័យ “University”។ ចូរវាយតម្លៃមធ្យមទៅលើ Salary ប្រចាំឆ្នាំរបស់និស្សិតដែលបានបញ្ចប់ការសិក្សា។

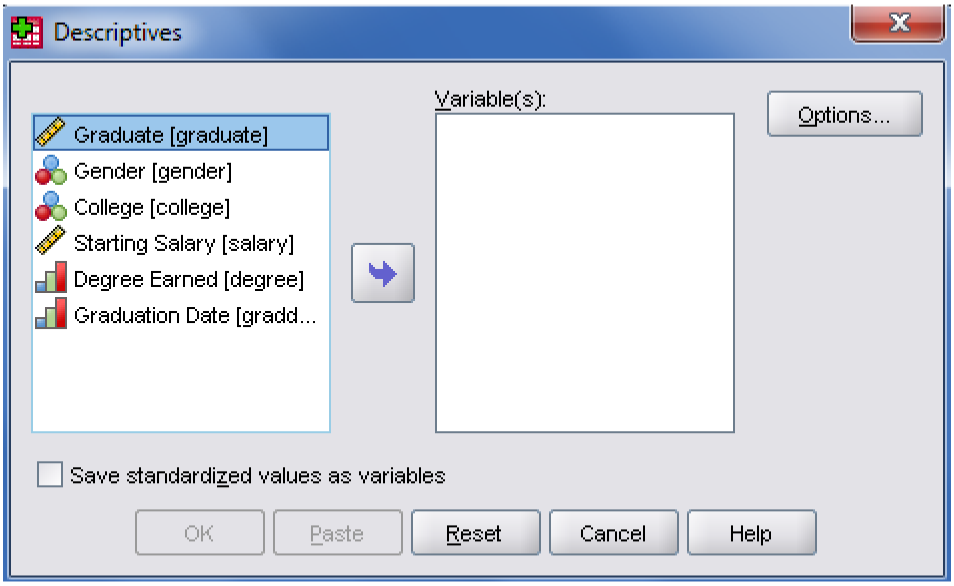
<u>Step1</u>: File ➡ Open ➡ Data ➡ C: Program files\SPSSInc\SPSS16\Sample ➡University of Florida graduate salaries.sav

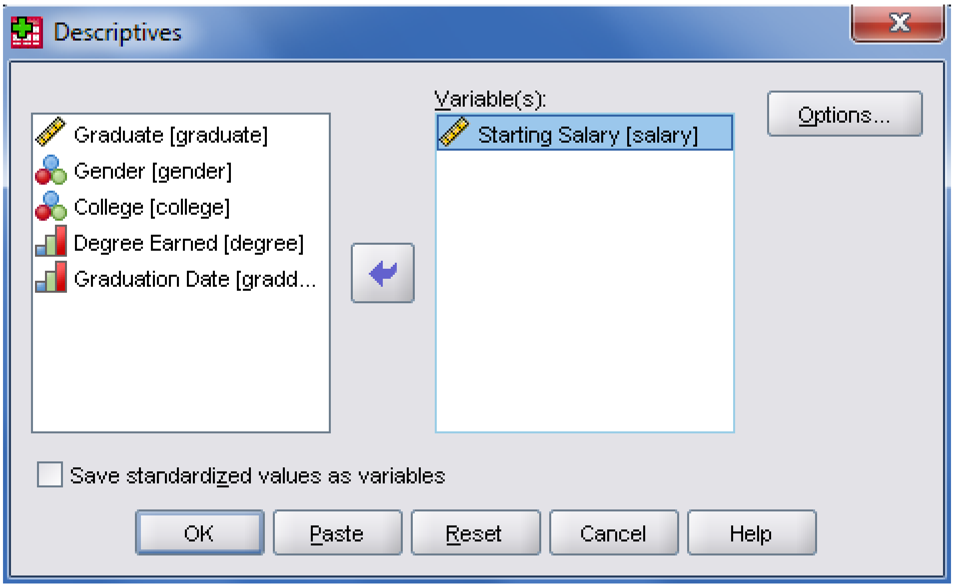


Step2: Analysis

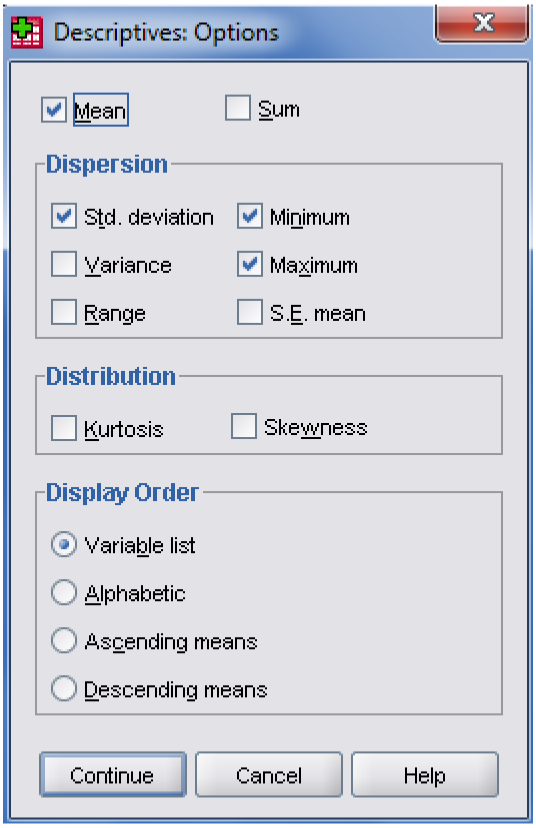
Analyze ➡ Descriptive Statistics ➡ Descriptives



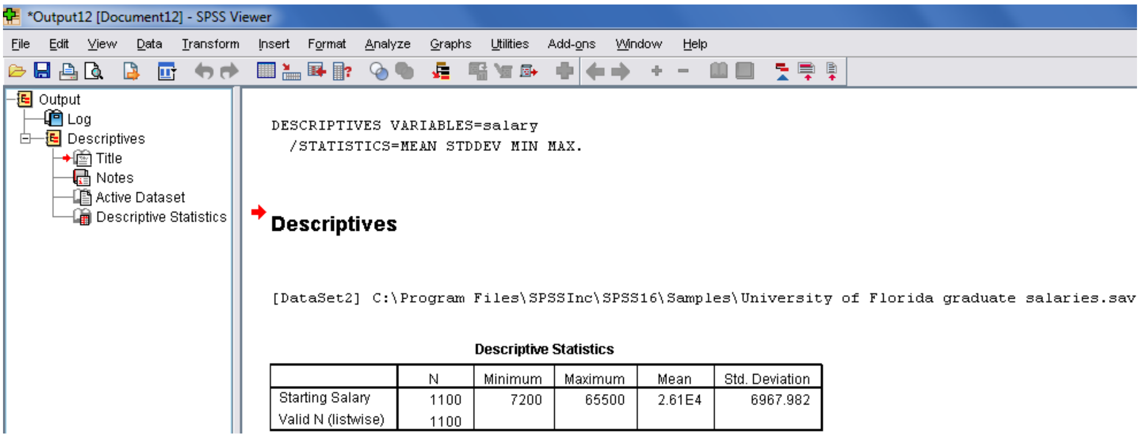




Click Options



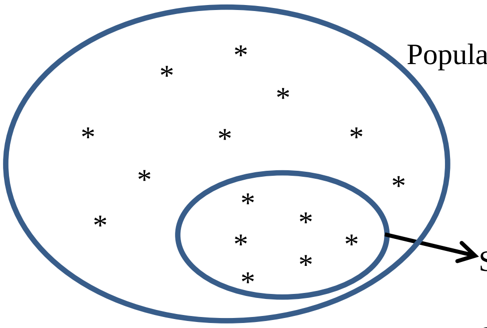
ចុច Continue រួចចុច OK



Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Starting Salary	1100	7200	65500	2.61E4	6967.982
Valid N (listwise)	1100				

ចម្លើយ:



Population $\mu = ?$

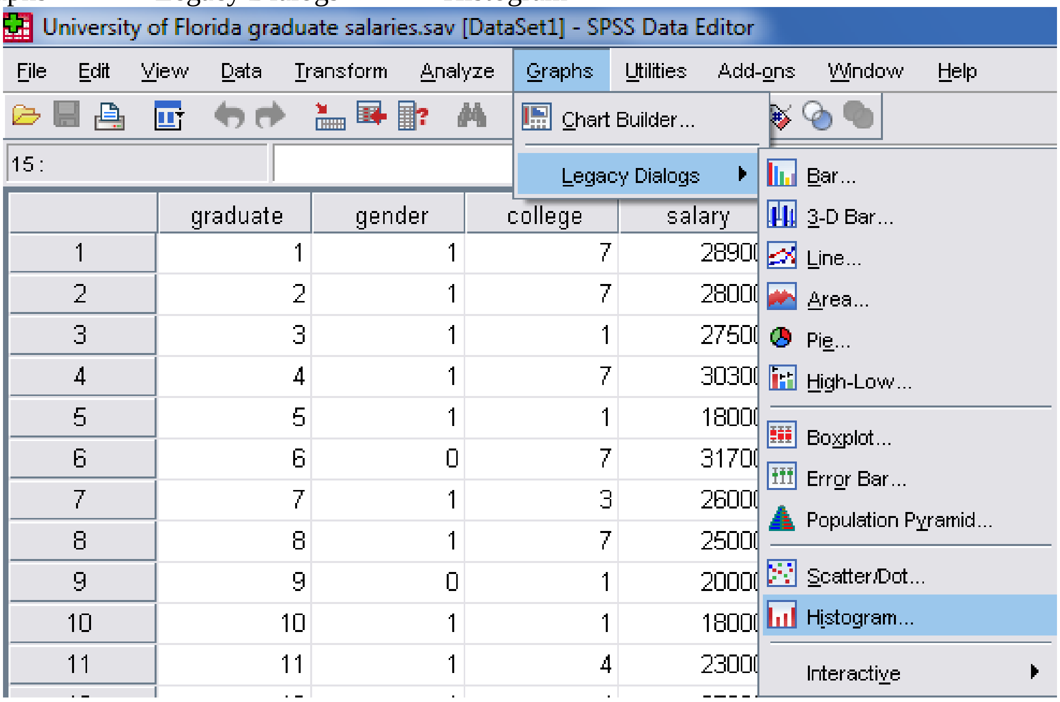
Sample $n = 1100,\ \bar{x} = \$2.61 \times 10^4$

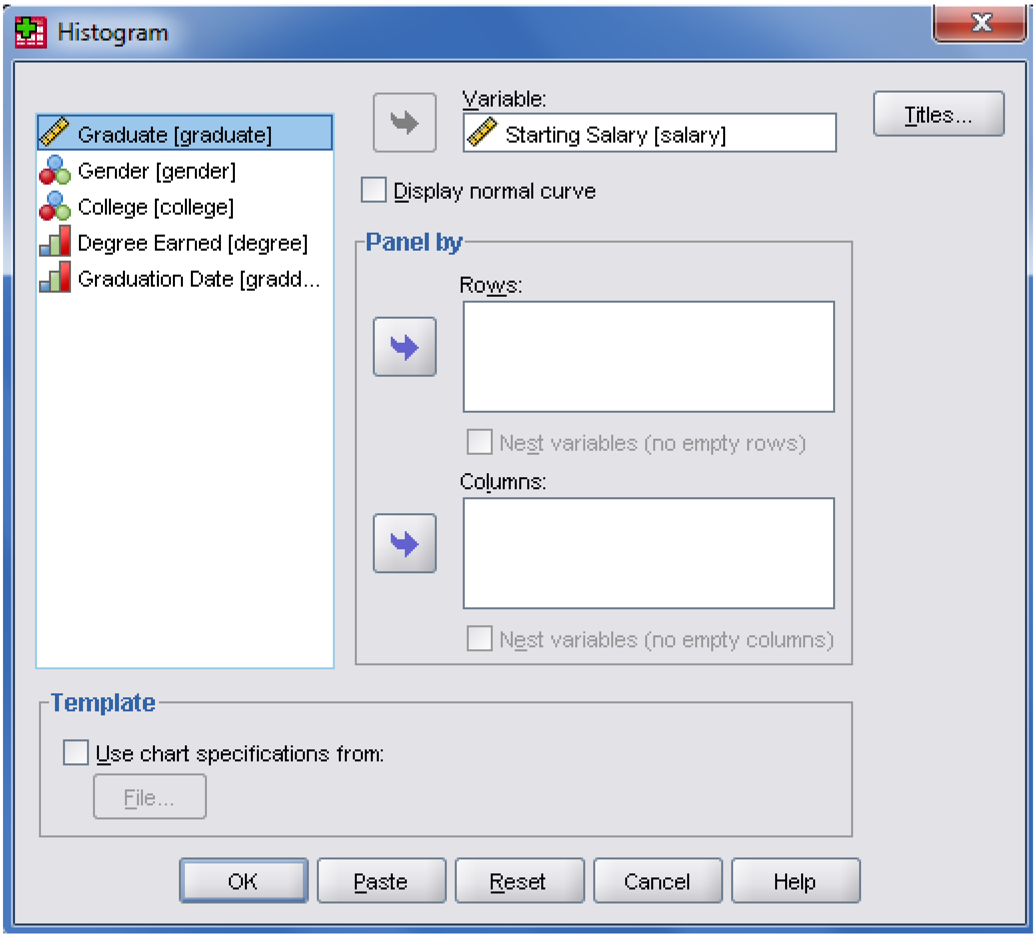
SEmean$= = \$210.09$

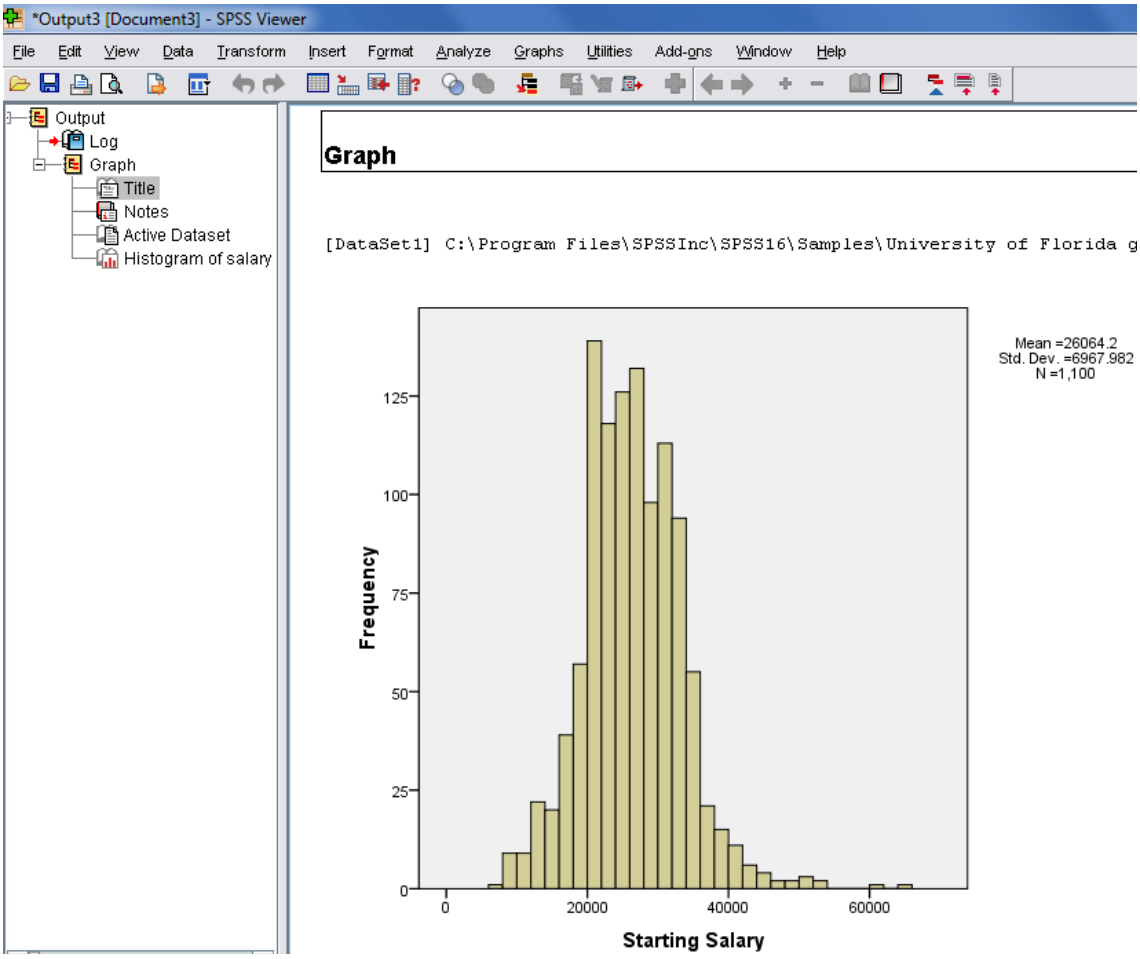
តាមតារាង Output ខាងលើបានបញ្ជាក់ឲ្យឃើញថា និស្សិតដែលបានបញ្ចប់ការសិក្សាហើយមានប្រាក់បៀវត្សប្រចាំឆ្នាំទាបជាងគេស្មើ $7200 និងខ្ពស់ជាងគេស្មើ $65500។ ម្យ៉ាងទៀត គេដឹងថា និស្សិតដែលបានបញ្ចប់ការសិក្សាទាំងនេះមានប្រាក់បៀវត្សជាមធ្យមប្រចាំឆ្នាំស្មើ $26000 ដែលការវាយតម្លៃនេះមាន Error ជាមធ្យមស្មើ $210.09។

សង់ Histogram

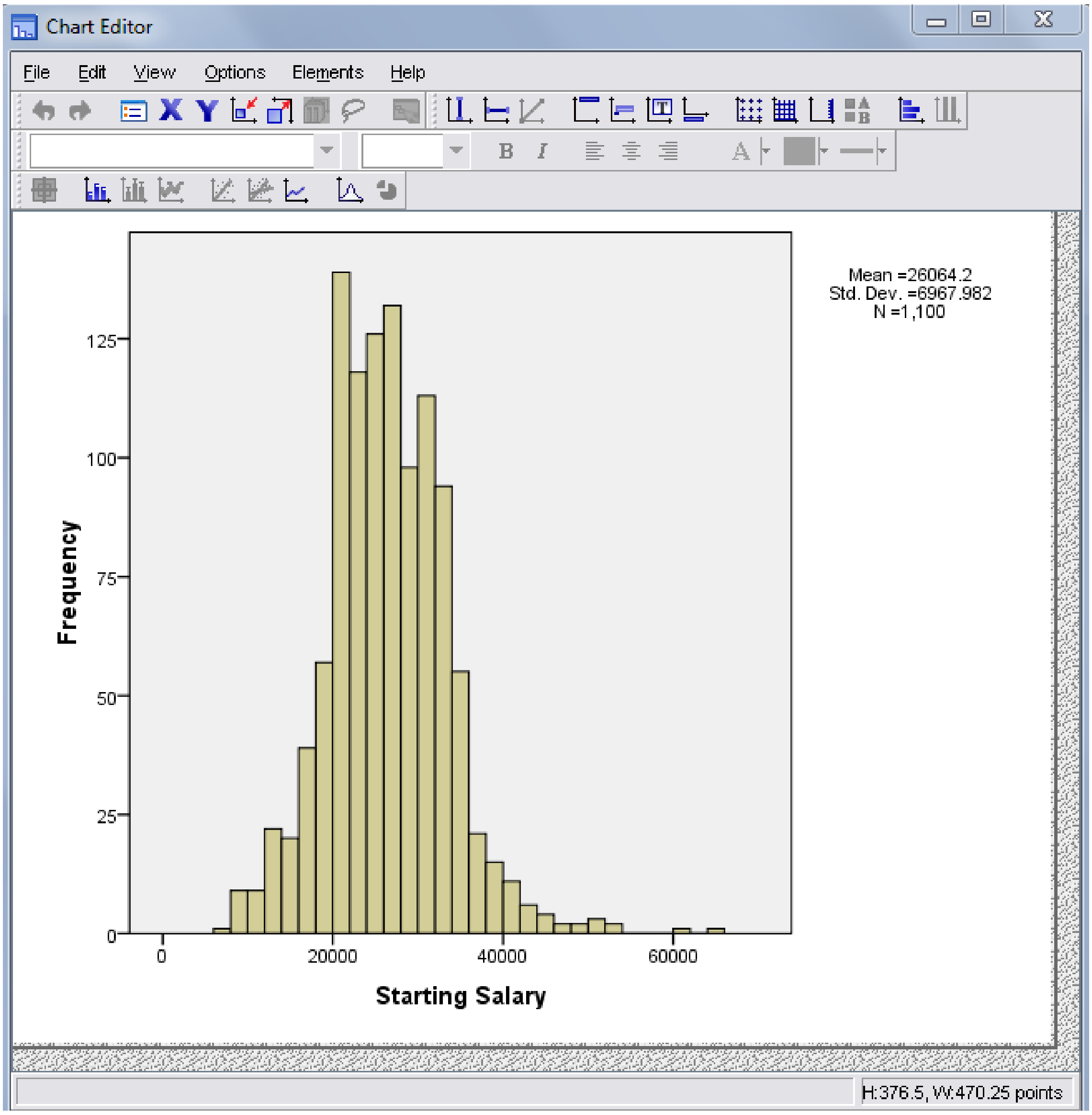
Graphs ➡ Legacy Dialogs ➡ Histogram




Histogram
Graduate [graduate]
Gender [gender]
College [college]
Degree Earned [degree]
Graduation Date [gradd...
Variable:
Starting Salary [salary]
Titles...
Display normal curve
Panel by
Rows:
Nest variables (no empty rows)
Columns:
Nest variables (no empty columns)
Template
Use chart specifications from:
File...
OK
Paste
Reset
Cancel
Help


*Output3 [Document3] - SPSS Viewer
File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help
Output
Log
Graph
Title
Notes
Active Dataset
Histogram of salary
Graph
[DataSet1] C:\Program Files\SPSSInc\SPSS16\Samples\University of Florida g
Mean =26064.2
Std. Dev. =6967.982
N =1,100
Frequency
0
25
50
75
100
125
0
20000
40000
60000
Starting Salary

បន្ទាប់ពី Double click លើ Histogram យើងឃើញរូបដូចខាងក្រោម:



Crosstabs Analysis

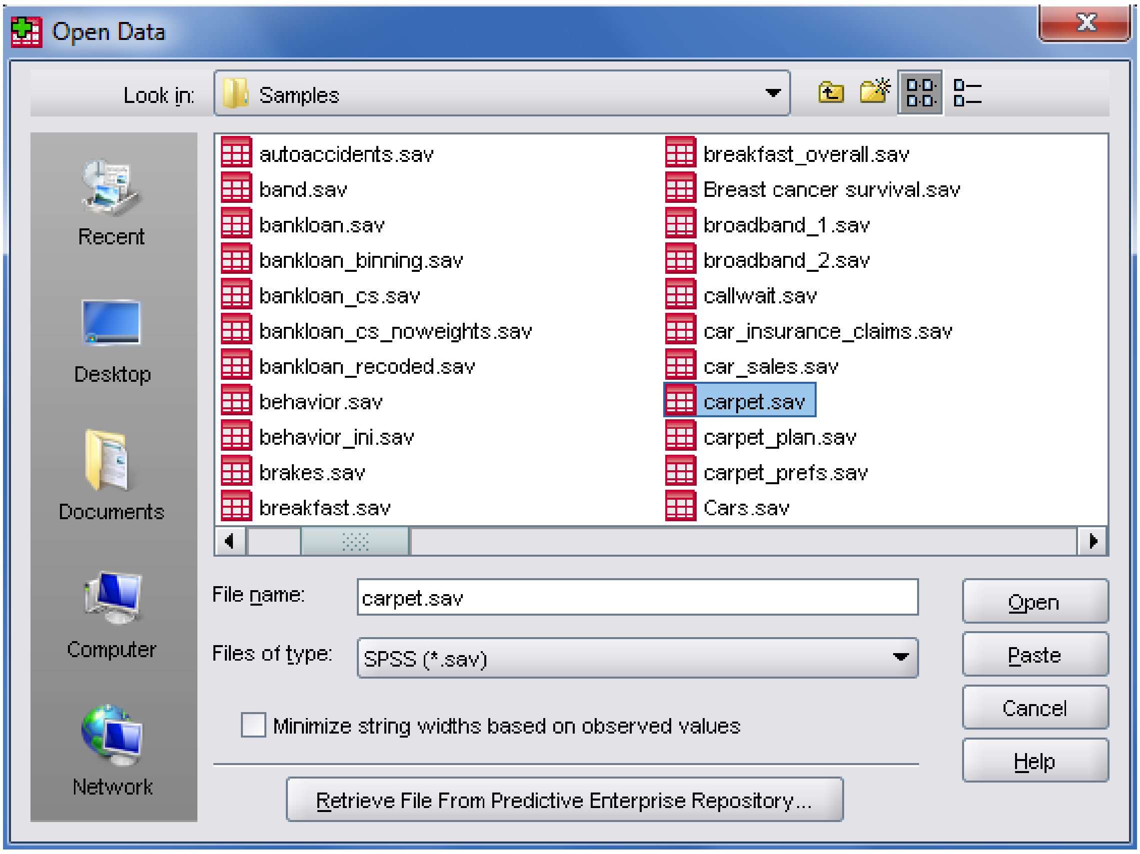
Crosstabs Analysis ប្រើសម្រាប់សិក្សាវិភាគទៅលើទិន្នន័យ Qualitative ដែលមានពីរ Variables (ប្រើសម្រាប់រាប់ចំនួន និងរកភាគរយ)។

ក្រាបដែលបកស្រាយរបស់វាមាន: Clustered (ក្រាបសសរភ្លោះ), Bar chart។

ឧទាហរណ៍: ដោយអនុវត្តទៅលើតារាងទិន្នន័យ "carpet.sav"។ ចូរបង្កើតតារាង Crosstabs វាយតម្លៃជាភាគរយធៀបនឹង Row ដែលមានដូចខាងក្រោម:

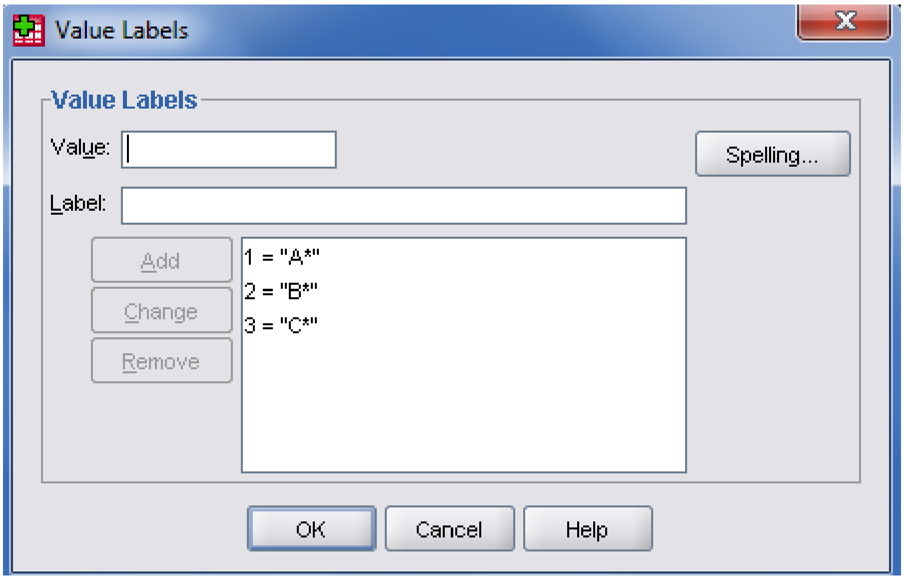
	A*	B*	C*	Row Total
Glory	?	?	?	100%
Bissell	?	?	?	100%
Column Total	?	?	?	100%

Step1: "carpet" (File➡Open➡Data➡..➡carpet.sav)

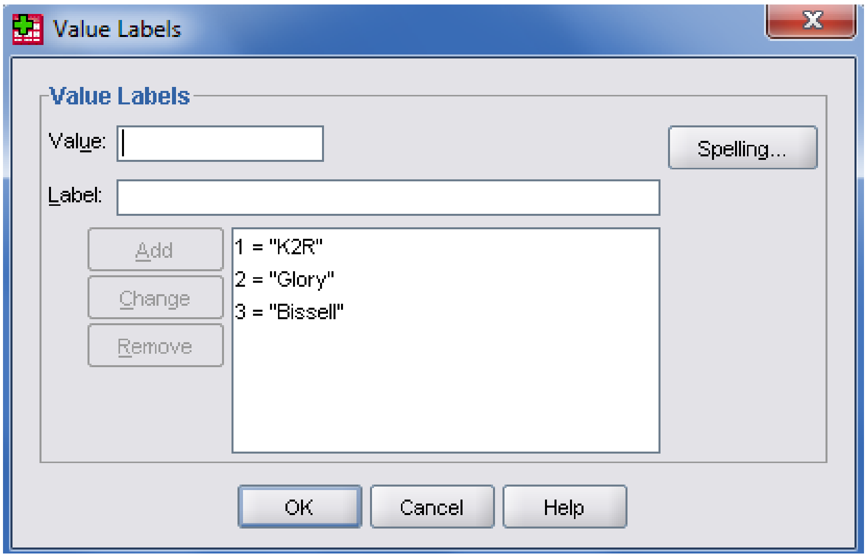


Step2:

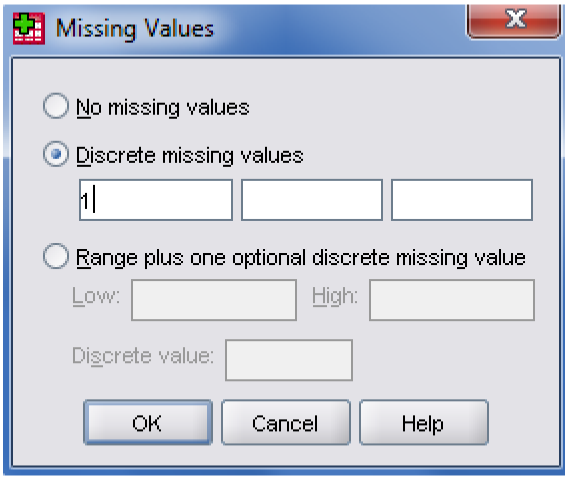
- Variable: package $\begin{cases} 1 = "A*" \\ 2 = "B*" \\ 3 = "C*" \end{cases}$



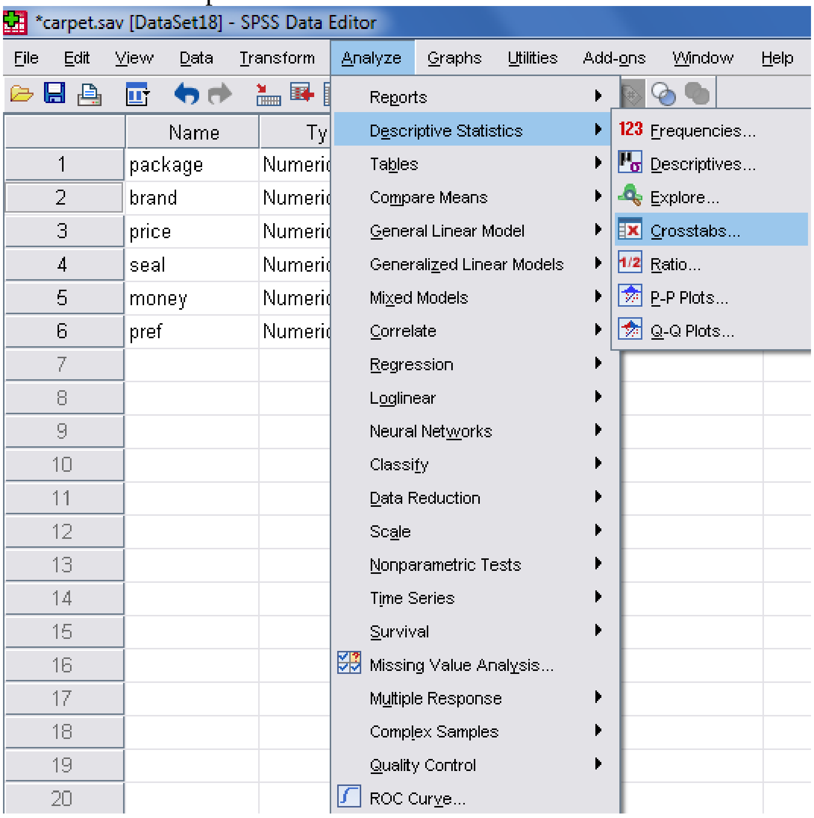
- Variable: brand $\begin{cases} 1 = "K2R" \\ 2 = "Glory" \\ 3 = "Bissell" \end{cases}$

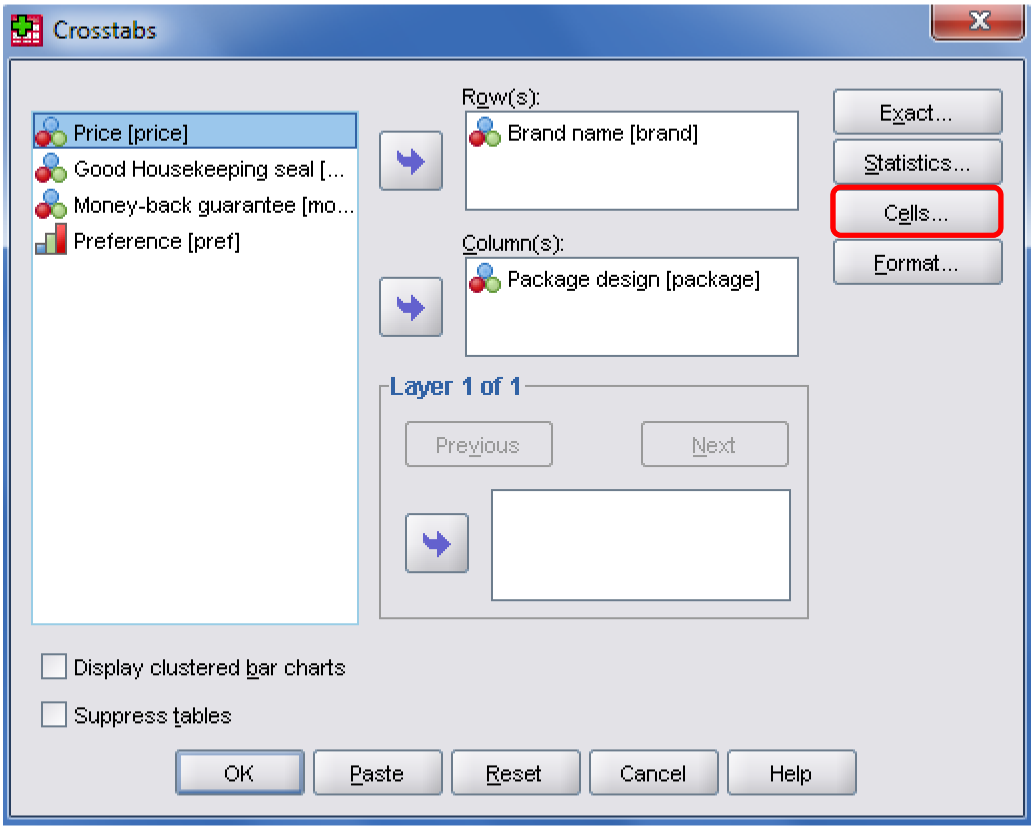


- ប្រើប្រអប់ Missing

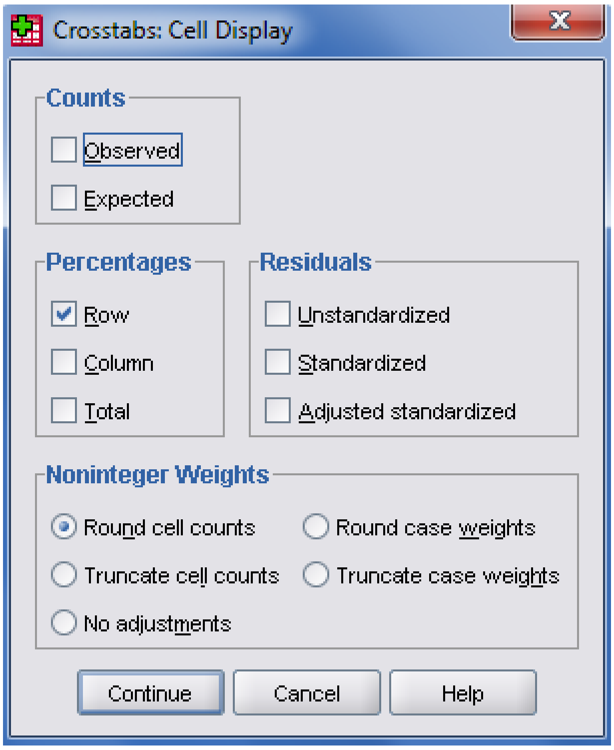


Step3: Analyze ➡ Descriptive Statistics ➡ Crosstabs

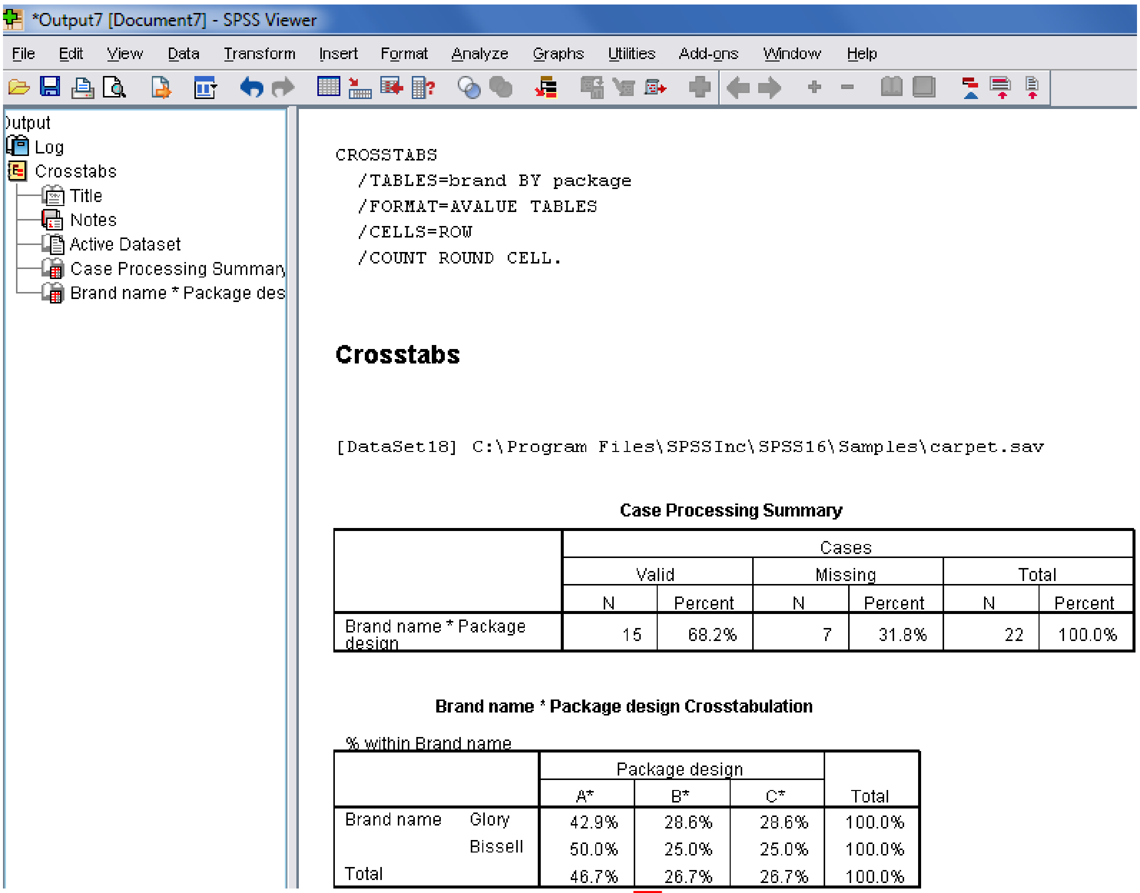




ចុចលើ Cells:



រួចចុច Continue ហើយ ចុច OK យើងបាន Crosstabs ដូចខាងក្រោម:



		Package design			Total
		A*	B*	C*	
Brand name	Glory	42.9%	28.6%	28.6%	100.0%
	Bissell	50.0%	25.0%	25.0%	100.0%
Total		46.7%	26.7%	26.7%	100.0%

Test of Hypothesis

ការតេស្តសម្មតិកម្ម

1. និយមន័យ:

Test of hypothesis គឺជាការសិក្សាផ្ទៀងផ្ទាត់សម្មតិកម្មទៅលើអំណះអំណាងដែលបានលើកឡើងថា«**ពិត**» ឬ«**មិនពិត**»។

ដើម្បីផ្ទៀងផ្ទាត់សម្មតិកម្ម គេបានបង្កើតសម្មតិកម្មថ្មីពីរ:

- H_0 : ឲ្យឈ្មោះថា សម្មតិកម្មសូន្យ (Null hypothesis)
- H_1 ឬ H_A : ឲ្យឈ្មោះថា សម្មតិកម្មប្រឆាំង (Alternative hypothesis)

ចំណាំ:

- H_0 :អាចទទួលយកសញ្ញា: $=, \geq, \leq$
- H_1 :អាចទទួលយកសញ្ញា: $\neq, <, >$
- អំណះអំណាងដែលលើកឡើងអាចនៅក្នុង H_0 ឬ H_1 ។

ឧទាហរណ៍1: សាស្ត្រាចារ្យម្នាក់បានអះអាងថា និស្សិត IT នៅក្នុងសាកលវិទ្យាល័យភូមិន្ទភ្នំពេញមានអាយុជាមធ្យមស្មើ 22 ឆ្នាំ។

ចូរលើកទម្រង់តេស្តដើម្បីផ្ទៀងផ្ទាត់សម្មតិកម្មខាងលើ។

<u>ចម្លើយ</u>

លើកទម្រង់តេស្ត:

$$VS\begin{cases} H_0 : \mu = 22 \\ \boxed{H_1 : \mu \neq 22} \end{cases}$$

← 2 tailed=តេស្តសងខាង

VS=Versus

ឧទាហរណ៍2: អ្នកអង្កេតការណ៍ម្នាក់បានអះអាងថា ប្រាក់ចំណូលជាមធ្យមរបស់ប្រជាជនដែលរស់នៅក្នុងតំបន់នេះទាបជាង \$270 ក្នុងមួយខែ។ ចូរលើទម្រង់តេស្តនៃសម្មតិកម្ម។

<u>ចម្លើយ</u>

លើកទម្រង់តេស្ត:

$$VS\begin{cases} H_0 : \mu \geq \$270 \\ \boxed{H_1 : \mu < \$270} \end{cases} \quad \text{ឬ} \quad VS\begin{cases} H_0 : \mu = \$270 \\ \boxed{H_1 : \mu < \$270} \end{cases}$$

1 tailed less=តេស្តខាងតូច

ឧទាហរណ៍3: នាយកផលិត Disc មួយប្រភេទបានអះអាងថា ការផលិត Disc មានខូចយ៉ាងច្រើន 2%។ ចូរលើកទម្រង់តេស្តនៃសម្មតិកម្ម។

<u>ចម្លើយ</u>

លើកទម្រង់តេស្ត:

$$VS\begin{cases} H_0 : P \leq 2\% \\ \boxed{H_1 : P > 2\%} \end{cases} \quad \text{ឬ} \quad VS\begin{cases} H_0 : P = 2\% (i.e. <) \\ \boxed{H_1 : P > 2\%} \end{cases}$$

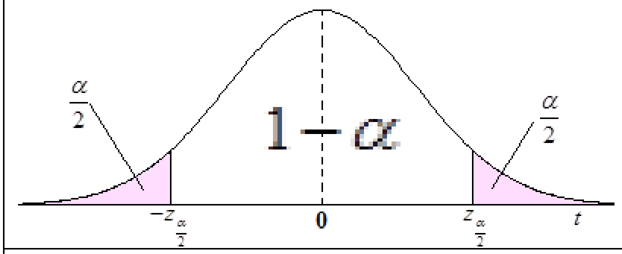
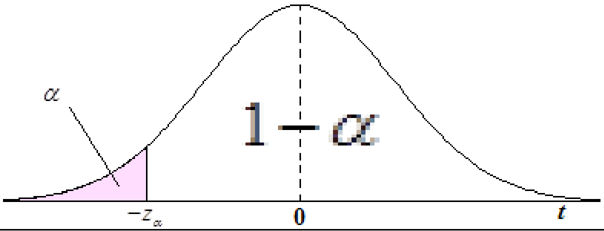
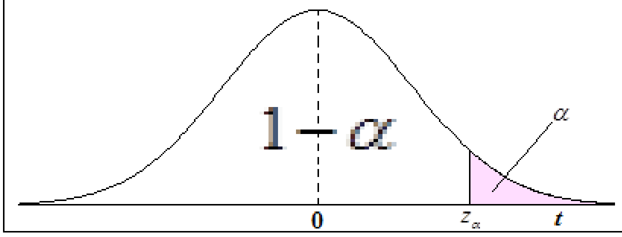
1 tailed greater=តេស្តខាងធំ

2. **<u>One Sample test for mean:</u>**

គឺជាការសិក្សាផ្ទៀងផ្ទាត់សម្មតិកម្ម ជាតម្លៃមធ្យម μ ដែលមាន t-Statistic តាងដោយ:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$ ដែល $\begin{cases} \bar{x} : \text{ជាតម្លៃមធ្យម} \\ \mu_0 : \text{ជា } Test\ Value \\ s : \text{ជា } Std \\ n : \text{ជា } Sample\ size \end{cases}$

តារាងសង្ខេប:

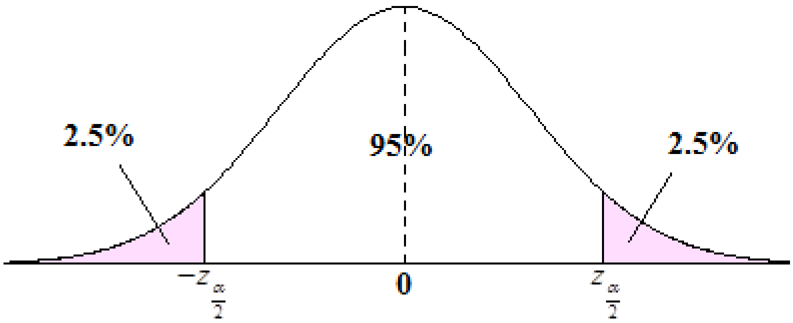
Test of hypothesis	Decision
2 tailed, $VS \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$ 	• Accept H_0 : if $-z_{\frac{\alpha}{2}} \leq t \leq z_{\frac{\alpha}{2}}$ $(\lvert t \rvert \leq \lvert z_{\frac{\alpha}{2}} \rvert)$ • Reject H_0 : if $t < -z_{\frac{\alpha}{2}}$ or $t > z_{\frac{\alpha}{2}}$ $(\lvert t \rvert > \lvert z_{\frac{\alpha}{2}} \rvert)$
1 tailed less, $VS \begin{cases} H_0 : \mu \geq \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ 	• Accept H_0 : if $t \geq -z_{\alpha}$ • Reject H_0 : if $t < -z_{\alpha}$
1 tailed greater, $VS \begin{cases} H_0 : \mu \leq \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$ 	• Accept H_0 : if $t \leq z_{\alpha}$ • Reject H_0 : if $t > z_{\alpha}$

ចំណាំ:

- បើ $\alpha = 5\% \Rightarrow \begin{cases} z_{\frac{\alpha}{2}} = 1.96 \\ z_{\alpha} = 1.64 \end{cases}$
- បើ $\alpha = 10\% \Rightarrow \begin{cases} z_{\frac{\alpha}{2}} = 1.64 \\ z_{\alpha} = 1.28 \end{cases}$

ឧទាហរណ៍: គណនា $z_{\frac{\alpha}{2}}$ បើ $\alpha = 5\%$

$P\left(0 < t < z_{\frac{\alpha}{2}}\right) = 0.4750 \Rightarrow z_{\frac{\alpha}{2}} = 1.96$ (ប្រើតារាងនៅ Normal នៅខាងចុងសៀវភៅ)



S-plus:

t.test(data,alternative="…………",conf.level=1-α)

two.sided

less

greater

លំហាត់

អភិបាលស្រុកម្នាក់បានអះអាងថា នៅក្នុងស្រុករបស់គាត់ គ្រួសារនីមួយៗ ទទួលបានផលស្រូវជាមធ្យមខ្ពស់ជាង10t។ដើម្បីផ្ទៀងផ្ទាត់អំណះអំណាងនេះ គេបាន ធ្វើការអង្កេតទៅលើប្រជាជនចំនួន 100 គ្រួសារ ហើយគណនាបាន $\bar{x} = 14t, s = 4t$ ។

a. ចូរលើកទម្រង់តេស្តនៃសម្មតិកម្មនេះ។

b. តើអំណះអំណាងនេះត្រឹមត្រូវដែរឬទេ បើ $\alpha = 5\%$?

ចម្លើយ

a. លើកទម្រង់តេស្ត:

$VS\begin{cases} H_0 : \mu \le 10t \\ H_1 : \mu > 10t \end{cases}$ ឬ $VS\begin{cases} H_0 : \mu = 10t\,(i.e.<) \\ H_1 : \mu > 10t \end{cases}$

b. Decision of test:

យើងមាន $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{14 - 10}{4/\sqrt{100}} = 10$

$\alpha = 5\% \Rightarrow z_\alpha = 1.64$

គេបាន $t = 10 > z_\alpha = 1.64$

ដូចនេះ Reject H_0 មានន័យថា គ្រួសារនីមួយៗនៅក្នុងតំបន់នោះទទួលបាន ទិន្នផលស្រូវជាមធ្យមខ្ពស់ជាង 10t នាំឲ្យអំណះអំណាងរបស់អភិបាលស្រុកពិតជា ត្រឹមត្រូវដែលអាចជឿជាក់បាន 95%។

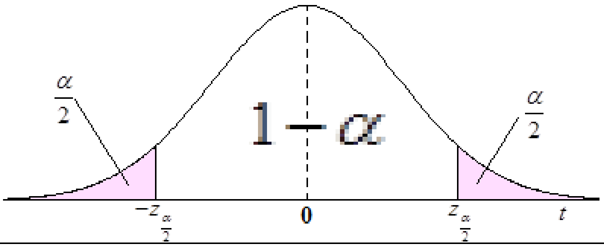
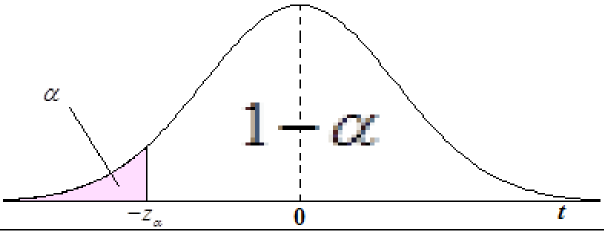
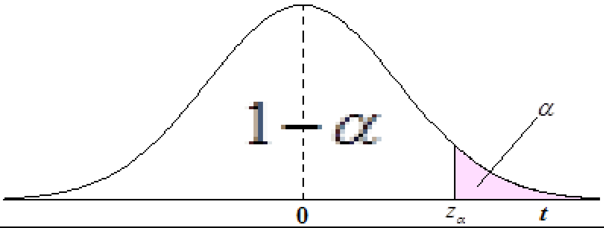
3. Independent Sample test

គឺជាការសិក្សាប្រៀបធៀបមធ្យមពីរ មិនអាស្រ័យគ្នាជ្រើសចេញពី population ផ្សេងគ្នា។ ទិន្នន័យអង្កេតជាប្រភេទ Quantitative។

t-Statistic គាងដោយ:

$$t=\frac{\overline{x}_1-\overline{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}\quad \text{ដែល}\begin{cases}\overline{x_1},\ \overline{x_1}\ \text{ជាមធ្យមគម្រូទី } 1,\text{ ទី}2\\ s_1\ ,\ s_2\ \text{ជា}\ Std\ \text{ទី}1,\ \text{ទី}2\\ n_1,n_1\ \text{ជា}\ Sample\ size\ \text{ទី}1,\text{ ទី}2\end{cases}$$

<u>តារាងសង្ខេប:</u>

Test of hypothesis	Decision
2 tailed, $VS\begin{cases}H_0:\mu_1=\mu_2\\ H_1:\mu_1\neq\mu_2\end{cases}$ 	• Accept H_0 : if $-z_{\frac{\alpha}{2}}\le t\le z_{\frac{\alpha}{2}}$ $(\lvert t\rvert\le\lvert z_{\frac{\alpha}{2}}\rvert)$ • Reject H_0 : if $t<-z_{\frac{\alpha}{2}}$ or $t>z_{\frac{\alpha}{2}}$ $(\lvert t\rvert>\lvert z_{\frac{\alpha}{2}}\rvert)$
1 tailed less, $VS\begin{cases}H_0:\mu_1\ge\mu_2\\ H_1:\mu_1<\mu_2\end{cases}$ 	• Accept H_0 : if $t\ge -z_{\alpha}$ • Reject H_0 : if $t<-z_{\alpha}$
1 tailed greater, $VS\begin{cases}H_0:\mu_1\le\mu_2\\ H_1:\mu_1>\mu_2\end{cases}$ 	• Accept H_0 : if $t\le z_{\alpha}$ • Reject H_0 : if $t>z_{\alpha}$

S-plus:
t.test(data1,data2,alternative= "…………",conf.level=1-α)
two.sided
less
greater

<u>លំហាត់</u>

អ្នកអង្កេតការណ៍ម្នាក់ទៅលើប្រាក់បៀវត្សម៉ោងរបស់កម្មករសហគ្រាសមួយ បានអះអាងថា ប្រាក់បៀវត្សម៉ោងរបស់កម្មការិនីមិនលើសពីប្រាក់បៀវត្សម៉ោងរបស់

កម្មករ។ ដើម្បីផ្ទៀងផ្ទាត់អំណះអំណាងនេះ គេបានជ្រើសរើស Sample ដោយចៃដន្យពីរមិនអាស្រ័យគ្នា:

- Sample ប្រាក់បៀវត្សម៉ោងរបស់កម្មការិនី គេជ្រើសយក 172 នាក់ ហើយគណនាបាន $\overline{x_1} = \$3.68, s_1 = \$0.966$ ។
- Sample ប្រាក់បៀវត្សម៉ោងរបស់កម្មករ គេជ្រើសយក 186 នាក់ ហើយគណនាបាន $\overline{x_2} = \$4.059, s_2 = \$0.839$ ។

a. ចូរលើកទម្រង់តេស្ត។

b. តើអំណះអំណាងខាងលើត្រឹមត្រូវដែរឬទេ បើ $\alpha = 5\%$?

ចម្លើយ

a. លើកទម្រង់តេស្ត:

តាង μ_1 ជាប្រាក់បៀវត្សម៉ោងជាមធ្យមរបស់កម្មការិនី

μ_2 ជាប្រាក់បៀវត្សម៉ោងជាមធ្យមរបស់កម្មករ

$$VS\begin{cases} H_0 : \mu_1 \le \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

b. Decision of test:

យើងមាន $t = \dfrac{\overline{x_1} - \overline{x_2}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \dfrac{3.68 - 4.059}{\sqrt{\dfrac{(0.966)^2}{172} + \dfrac{(0.839)^2}{186}}} = -3.98$

តែ $\alpha = 5\% \Rightarrow z_\alpha = 1.64$

គេបាន $t = -3.98 < z_\alpha = 1.64$

ដូចនេះ Accept H_0 មានន័យថា ប្រាក់បៀវត្សម៉ោងជាមធ្យមរបស់កម្មការិនីមិនលើសពីប្រាក់បៀវត្សម៉ោងជាមធ្យមរបស់កម្មករ នាំឲ្យអំណះអំណាងរបស់អ្នកអង្កេតការណ៍ពិតជាត្រឹមត្រូវ ហើយអាចជឿជាក់បាន 95%។

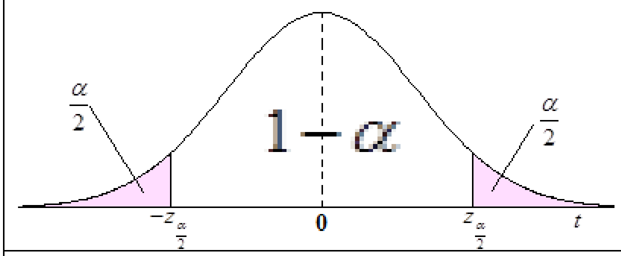
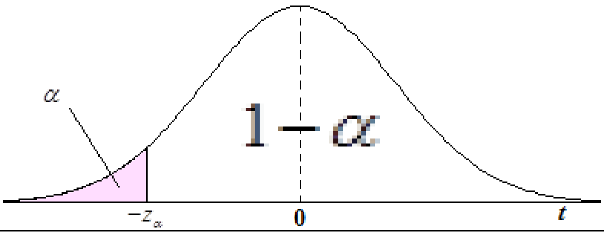
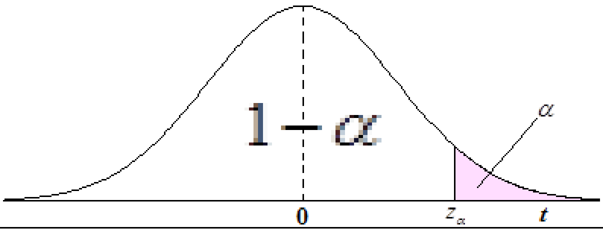
4. Dependent Sample test:

គឺជាការសិក្សាប្រៀបធៀបមធ្យមពីរអាស្រ័យគ្នាជ្រើសចេញពី population តែមួយមុន និងក្រោយ។ទំហំ Sample ស្មើគ្នាហើយទិន្នន័យអង្កេតជាប្រភេទ Quantitative។

t-Statistic តាងដោយ:

$$t = \frac{\overline{D}}{s/\sqrt{n}} \quad \text{ដែល} \begin{cases} s = \sqrt{\dfrac{1}{n-1}\sum\left(d_i - \overline{D}\right)^2} \\ n : Sample\,size \\ \overline{D} = \dfrac{1}{n}\sum d_i \qquad , d_i = x_i - y_i\left(x_i = before, y_i = after\right) \end{cases}$$

តារាងសង្ខេប:

Test of hypothesis	Decision
2 tailed, $VS\begin{cases}H_0:\mu_d=0,\mu_d=\mu_1-\mu_2\\H_1:\mu_d\neq 0\end{cases}$ 	• Accept H_0 : if $-z_{\frac{\alpha}{2}}\le t\le z_{\frac{\alpha}{2}}$ $(\lvert t\rvert\le\lvert z_{\frac{\alpha}{2}}\rvert)$ • Reject H_0 : if $t<-z_{\frac{\alpha}{2}}$ or $t>z_{\frac{\alpha}{2}}$ $(\lvert t\rvert>\lvert z_{\frac{\alpha}{2}}\rvert)$
1 tailed less, $VS\begin{cases}H_0:\mu_d\ge 0\\H_1:\mu_d<0\end{cases}$ 	• Accept H_0 : if $t\ge -z_\alpha$ • Reject H_0 : if $t<-z_\alpha$
1 tailed greater, $VS\begin{cases}H_0:\mu_d\le 0\\H_1:\mu_d>0\end{cases}$ 	• Accept H_0 : if $t\le z_\alpha$ • Reject H_0 : if $t>z_\alpha$

S-plus:
t.test(data1,data2,alternative="…………",conf.level=1-α)
two.sided
less
greater

លំហាត់

គេមានបំណងចង់ដឹងពីប្រសិទ្ធភាពនៃវគ្គបណ្តុះបណ្តាលលើមុខវិជ្ជាស្ថិតិ។ គេបានធ្វើការអង្កេតទៅលើនិស្សិតមួយក្រុមចំនួន 12 នាក់មុន និងក្រោយវគ្គបណ្តុះបណ្តាល ហើយគេទទួលបានទិន្នន័យដូចខាងក្រោម:

- ទិន្នន័យមុនការអង្កេត: 60,50,45,35,80,85,75,65,55,90,95,55។
- ទិន្នន័យក្រោយការអង្កេត: 65,90,50,55,80,90,85,80,85,90,95,75។

តើគេអាចអះអាងថា វគ្គបណ្តុះបណ្តាលនេះមានប្រសិទ្ធភាពបានដែរឬទេ បើ $\alpha=5\%$?

ចម្លើយ

លើកទម្រង់តេស្ត:

តាង x_i ជាទិន្នន័យមុនការអង្កេតដែល $x_i=60,50,45,35,80,85,75,65,55,90,95,55$

y_i ជាទិន្នន័យក្រោយការអង្កេតដែល $y_i = 65, 90, 50, 55, 80, 90, 85, 80, 85, 90, 95, 75$

គេបាន: $d_i = x_i - y_i = -5, -40, -5, -15, 0, -5, -10, -15, -30, 0, 0, -20$

តាង μ_1 ជាមធ្យមពិន្ទុរបស់និស្សិតមុនវគ្គបណ្តុះបណ្តាល

μ_2 ជាមធ្យមពិន្ទុរបស់និស្សិតក្រោយវគ្គបណ្តុះបណ្តាល

$$VS \begin{cases} H_0 : \mu_d \geq 0 \\ H_1 : \mu_d < 0 \end{cases}$$

Decision:

យើងមាន $t = \dfrac{\overline{D}}{s/\sqrt{n}}$

ដោយ $\overline{D} = \dfrac{1}{n} \sum d_i$

$$\overline{D} = \frac{1}{12}\left[(-5)+(-40)+(-5)+(-15)+0+(-5)+(-10)+(-15)+(-30)+0+0+(-20)\right]$$

$$= -\frac{145}{12} = -12.08$$

$$s = \sqrt{\frac{1}{n-1}\sum\left(d_i - \overline{D}\right)^2}$$

$$= \sqrt{\frac{1}{12-1}\left[3\left(-5+\frac{145}{12}\right)^2 + \left(-40+\frac{145}{12}\right)^2 + 2\left(-15+\frac{145}{12}\right)^2 + 3\left(0+\frac{145}{12}\right)^2 + \left(-10+\frac{145}{12}\right) + \left(-30+\frac{145}{12}\right)^2 + \left(-20+\frac{145}{12}\right)^2\right]}$$

$$= \sqrt{\frac{1}{11}\left[3\left(\frac{85}{12}\right)^2 + \left(\frac{-335}{12}\right)^2 + 2\left(\frac{-35}{12}\right)^2 + 3\left(\frac{145}{12}\right)^2 + \left(\frac{-25}{12}\right)^2 + \left(\frac{-215}{12}\right)^2 + \left(\frac{-95}{12}\right)^2\right]}$$

$$= \sqrt{\frac{1}{11}\cdot\frac{1}{12^2}\left[3(85)^2 + 335^2 + 2(35)^2 + 3(145)^2 + 25^2 + 215^2 + 95^2\right]}$$

$$= \sqrt{\frac{1}{1584}\left(21675 + 112225 + 2450 + 63075 + 625 + 46225 + 9025\right)}$$

$$= \sqrt{\frac{1}{1584}\left(255300\right)} = 12.69$$

$$\Rightarrow t = \frac{-12.08}{\frac{12.69}{\sqrt{12}}} = -3.29$$

តែ $\alpha = 5\% \Rightarrow z_\alpha = 1.64$

គេបាន $t = -3.29 < -z_\alpha = -1.64$

ដូចនេះ Reject H_0 មានន័យថា វគ្គបណ្តុះបណ្តាលនេះមានប្រសិទ្ធភាព។

នៅក្នុង SPSS យើងត្រូវបញ្ចូលទិន្នន័យមុន និងក្រោយការអង្កេតដូចខាងក្រោម

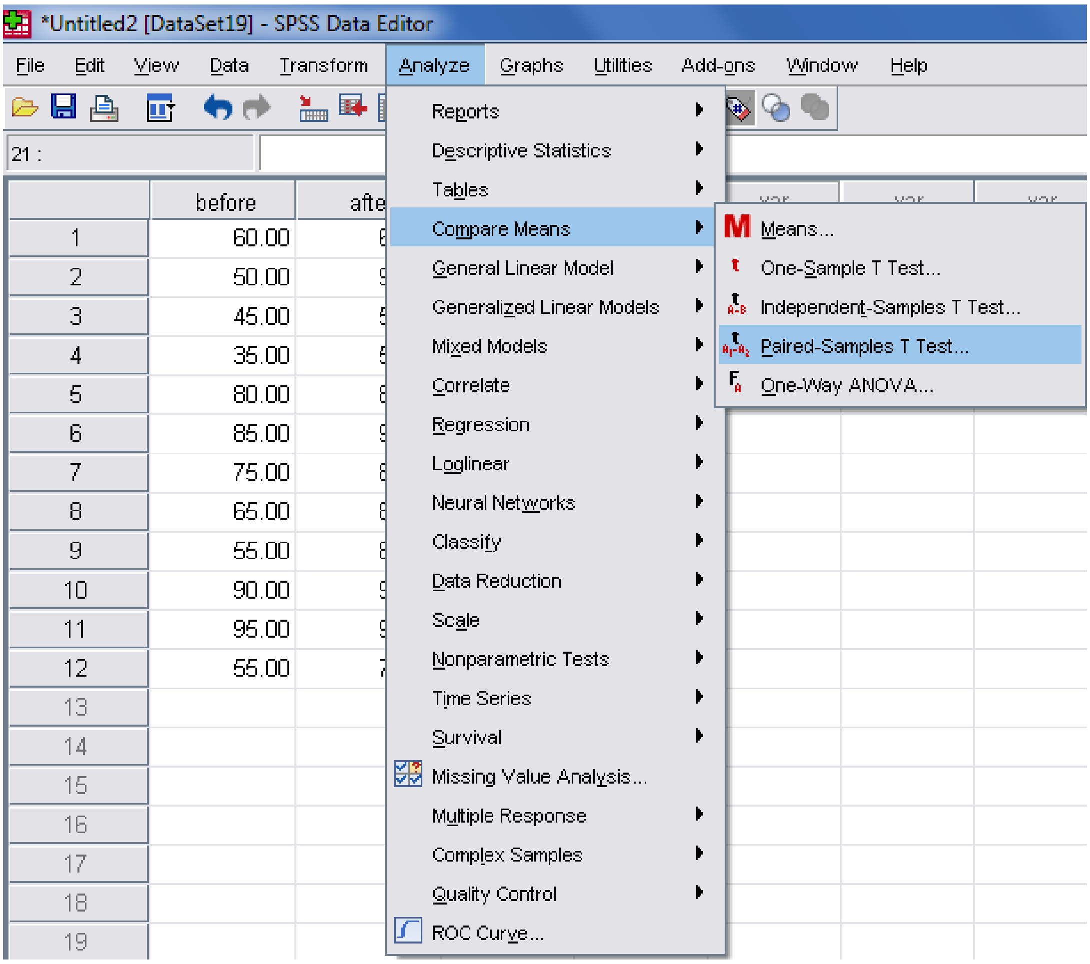
- ក្នុង Variable View:

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
before	Numeric	8	2	Before	None	None	8	Right	Scale
after	Numeric	8	2	After	None	None	8	Right	Scale

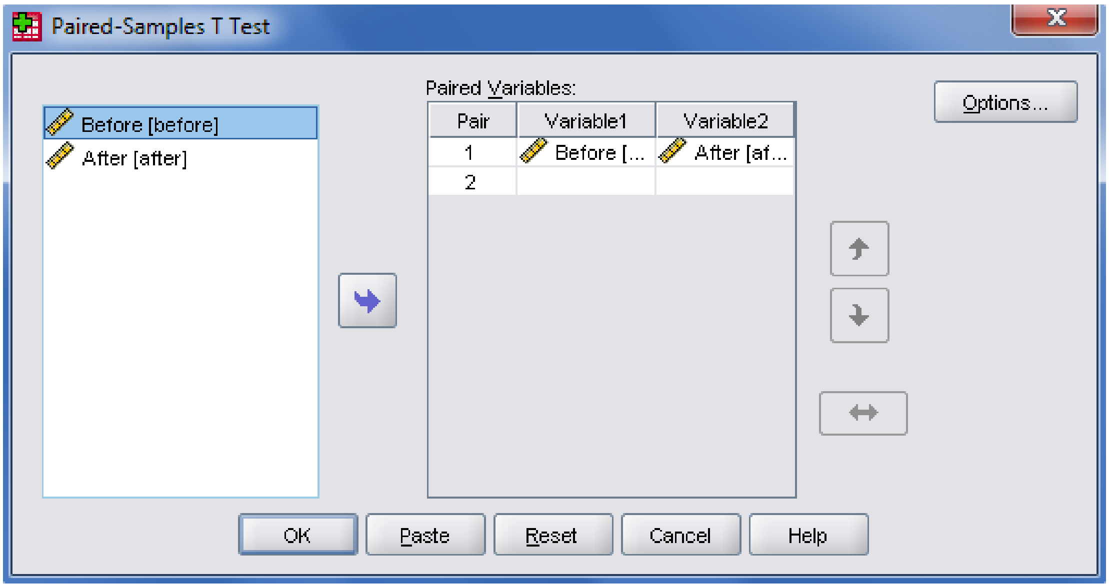
- ក្នុង Design View:

before	after
60.00	65.00
50.00	90.00
45.00	50.00
35.00	50.00
80.00	80.00
85.00	90.00
75.00	85.00
65.00	80.00
55.00	85.00
90.00	90.00
95.00	95.00
55.00	75.00

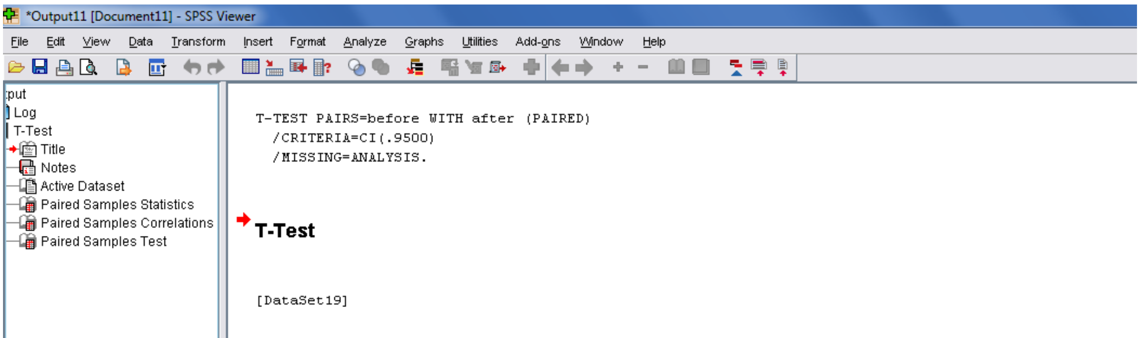
បន្ទាប់មកចុច Analyze →Compare Means→Paired-Samples T Test



បន្ទាប់មកវានឹងចេញផ្ទាំង



ចុច OK យើងបាន Output ដូចខាងក្រោម:



Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before	65.8333	12	19.04938	5.49908
	After	77.9167	12	15.29384	4.41495

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Before & After	12	.748	.005

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Before - After	-1.208E1	12.69544	3.66486	-20.14963	-4.01703	-3.297	11	.007

Select លើ Paired Sample Test

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Before - After	-1.208E1	12.69544	3.66486	-20.14963	-4.01703	-3.297	11	.007

Double Click លើ Paired Sample Test

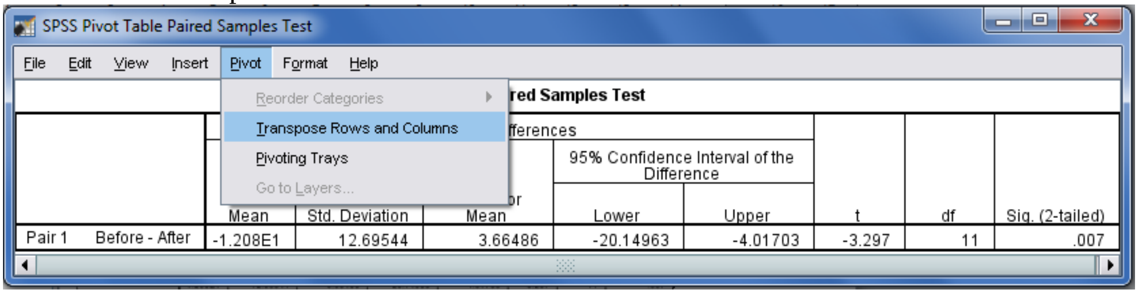
SPSS Pivot Table Paired Samples Test

File Edit View Insert Pivot Format Help

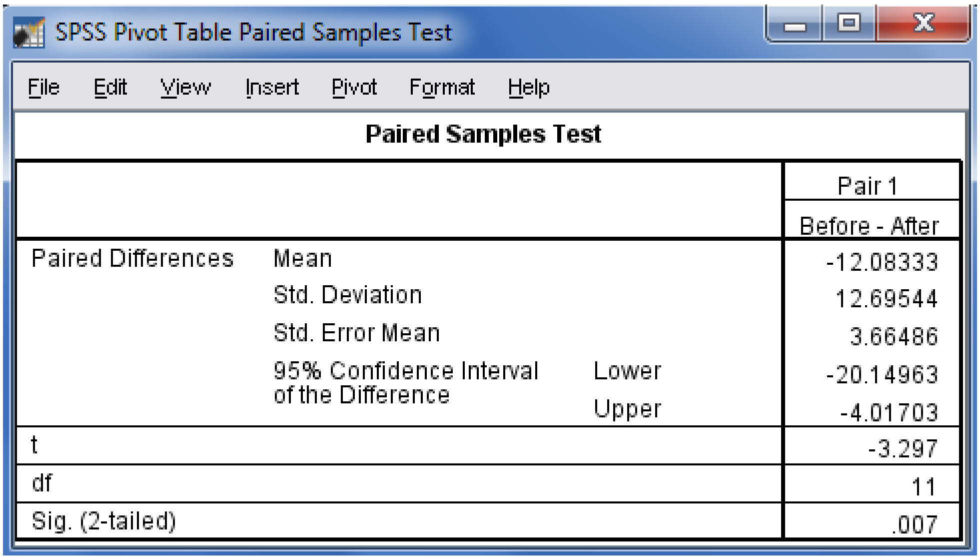
Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Before - After	-1.208E1	12.69544	3.66486	-20.14963	-4.01703	-3.297	11	.007

Click Pivot→Transpose Rows and Columns



យើងបាន:



SPSS Pivot Table Paired Samples Test

Paired Samples Test

			Pair 1 Before - After
Paired Differences	Mean		-12.08333
	Std. Deviation		12.69544
	Std. Error Mean		3.66486
	95% Confidence Interval of the Difference	Lower	-20.14963
		Upper	-4.01703
t			-3.297
df			11
Sig. (2-tailed)			.007

បកស្រាយតារាង Output

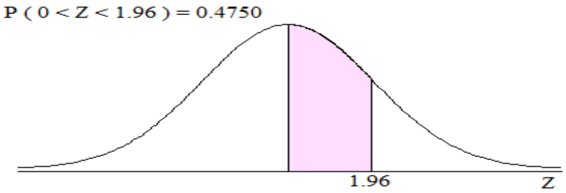
- Mean=-12.08 មានន័យថា វគ្គបណ្តុះបណ្តាលមានប្រសិទ្ធភាពដែលមានគំលាតមធ្យមស្មើ 12.08។
- Std. Deviation= $s=\sqrt{\frac{1}{n-1}\sum\left(d_i-\overline{D}\right)}=12.69$
- Std. Error Mean=3.66486។ ដូចនេះ ការវាយតម្លៃទៅលើមធ្យមគំលាតពិន្ទុរបស់និស្សិតមាន Error ជាមធ្យមស្មើ 3.66486។
- 95% Confidence Interval of the Difference:

  $Lower \le \mu_1-\mu_2 \le Upper$ ឬ $-20.14963 \le \mu_1-\mu_2 \le -4.01703$

ដូចនេះ វគ្គបណ្តុះបណ្តាលមានប្រសិទ្ធភាពដែលមានគំលាតពិន្ទុជាមធ្យមប្រែប្រួលពី 4 ទៅ 20។

- t=-3.297
- df=11, (df=$n-1$)
- Sig. (2-tailed)=.007

Standard of Normal Distribution $$P\left(0<Z<z\right)=\int_0^z \frac{1}{2\pi} e^{-\frac{1}{2}z^2}\, dz$$



The Normal Distribution

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3079	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4430	0.4441
1.6	0.4452	0.4463	0.4474	0.4485	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4700	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4762	0.4767
2.0	0.4773	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4865	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4980	0.4980	0.4981
2.9	0.4981	0.4982	0.4983	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.5000	0.5000	0.5000
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

សម្គាល់:

ដើម្បីបានតារាង Normal បង្រួមកណ្តាលយើងប្រើ S-plus programming ដែលអនុញ្ញាតឲ្យយើងអាចគណនា $P(0<Z<z)$ ដែល Z ជាអថេរ Normal បង្រួមកណ្តាល:

```
function()
{
z1<-seq(0,0.09,0.01)
z2<-seq(0,3.9,0.1)
z<-seq(0,3.99,0.01)
prob<-pnorm(z)-0.5
Table<-t(array(prob,dim=c(length(z1),length(z2))))
Table<-cbind(c(z2),Table)
Table<-rbind(c(NA,z1),Table)
return(Table)
}
```

ភ្នំពេញ, ថ្ងៃទី ២៨ ខែមករា ឆ្នាំ ២០១១

ធុន កុសល