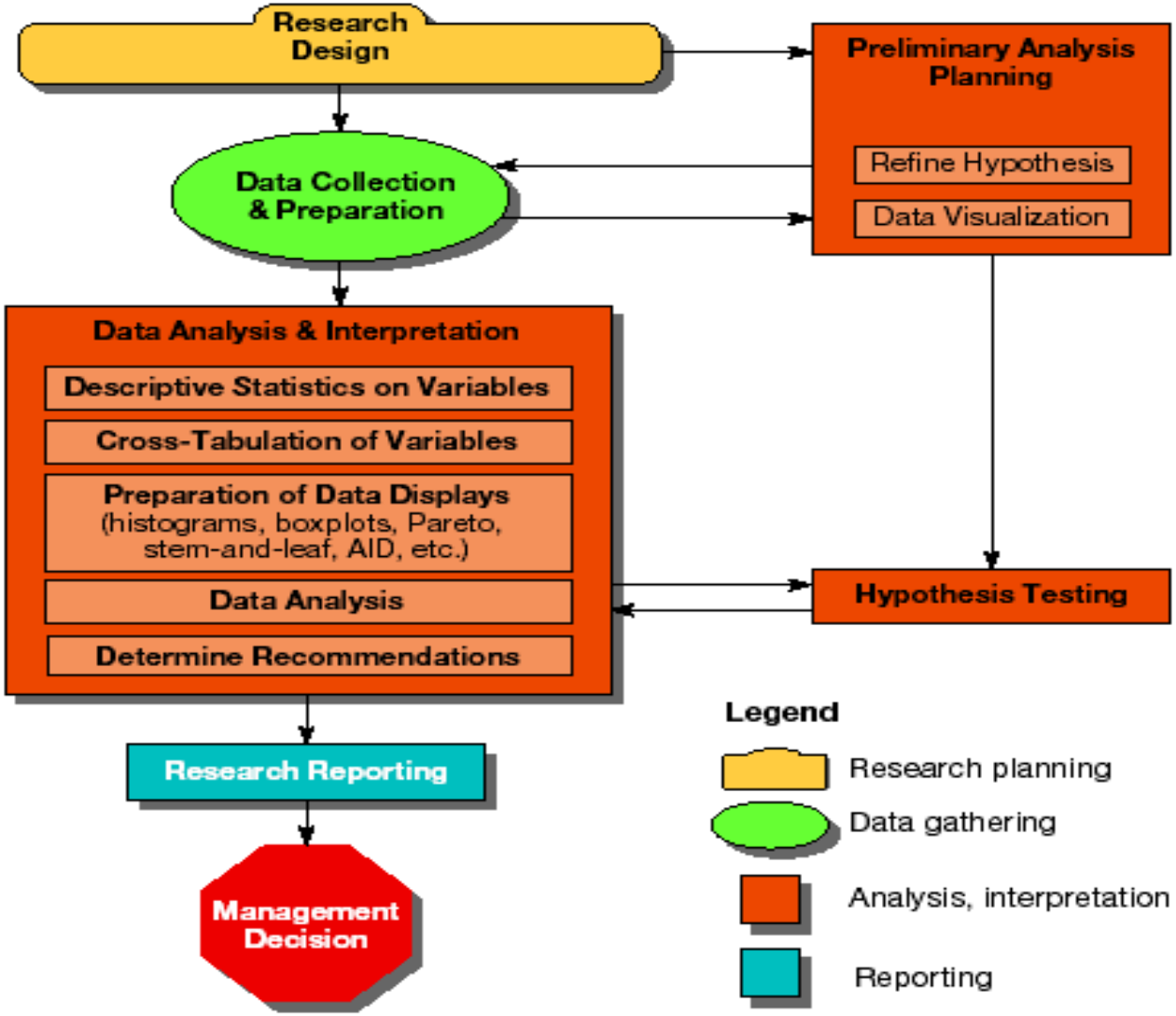
描述性分析

(Descriptive Analysis)

Data Exploration, Examination, and Analysis in the Research Process



Preparation of Data

- 原始数据的准备 Prepare the raw data
  - 编码 Coding
  - 编辑 Editing
- 通过SPSS打开数据文件 Open the data with SPSS
- 数据处理 Process the data
  - 缺省值 Missing value
- 数据分析 Analyze the data

Before the data entry

- 编码手册 You need a code book/scoring guide
- 每个样本需有一个编码号 You give ID number for each case (NOT real identification numbers of your subjects) if you use paper survey.
- 如果是在线调研，需有特定手段来确定样本 If you use online survey, you need something to identify your cases.
- 运用Excel, SPSS或其它软件进行数据录入 You also can use Excel, SPSS or other software to do data entry.

Example of a code book

编码手册需界定 A code book is about how you code your variables. What are in code book?

变量名称 Variable names

对应的应对选项的内涵 Values for each response option

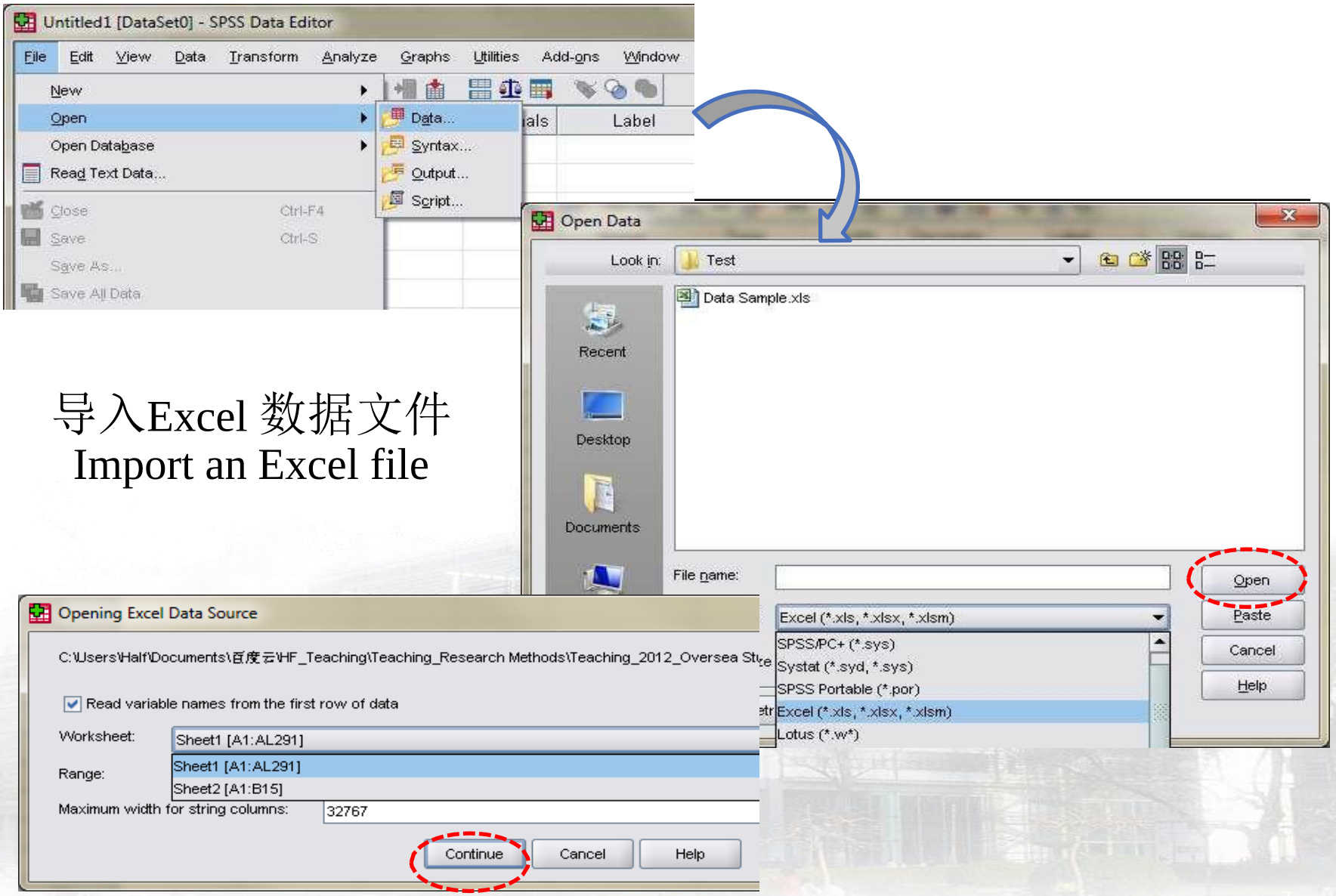
如何对变量重新编码 How to recode variables

Construct	Original Items	Rules
A	A1-A5	1. strongly disagree 5. Strongly agree
B	B1-B5	
C	C1-C6	
D	D1-D6	
E	E1-E6	
Ownership	CHA	1. State-owned; 2. Private owned; 3. foreign-controlled; 4. joint venture; 5. none
Industry	IND	0 Manufacturing; 1 service
Firm Size	Size	1. <100; 2. 100-299; 3. 300-499; 4. 500-999; 5. 1000-1999; 6. >=2000
IT Department Size	ITQ	1. =<2; 2. 3-5; 3. 6-10; 4. 11-15; 5. >16
Position	POS	1. Senior VP of other functions; 2. CFO; 3. CIO/CTO; 4 COO; 5 CEO; 6 Others

不合法变量名示例及分析

变量名	错误原因
NAME_@123	**变量名超过8个字符（11.0及以前版本）**
NAME_1.	**以句点结尾**
_12NAME	**第一个字符不是字母**
NA ME?1	**变量名中有？和空格**
ALL	**用了不能用作变量的关键词**

- **输入不合法的变量名时，会自动显示出错信息**



导入Excel 数据文件
Import an Excel file

member1.sav [DataSet1] - SPSS Statistics Data Editor

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

1 : peer 4.4

	peer	inseek	rb	pfram	future	social	role	task
1	4.40	3.40	2.00	3.33	1.40	3.00	4.25	4.00
2	4.40	3.40	2.00	3.33	1.40	3.00	4.25	4.00
3	4.40	3.40	2.00	3.33	1.40	3.50	4.25	4.00
4	4.40	3.40	2.00	3.33	1.60	3.00	4.25	4.60
5	4.40	3.40	2.00	3.33	1.60	3.00	4.25	4.60
6	4.40	3.40	2.00	3.33	1.60	4.00	4.25	4.60
7	4.40	3.20	2.33	3.00	1.80	5.00	4.25	4.20
8	2.80	4.60	2.33	4.00	2.20	4.00	4.25	4.00
9	4.80	2.80	2.33	2.67	3.00	1.00	3.00	3.00
10	4.80	2.80	2.33	2.67	3.00	1.00	3.00	3.00
11	4.80	2.80	2.33	2.67	3.00	1.00	3.00	3.00
12	4.80	2.80	2.33	2.67	3.00	1.00	3.00	3.00
13	4.80	2.80	2.33	2.67	3.00	1.00	3.00	3.00
14	4.20	4.40	2.67	4.33	3.00	3.25	3.50	3.40
15	2.20	3.80	2.67	3.33	3.80	4.00	4.00	4.00
16	3.80	2.40	2.67	3.33	2.60	2.00	2.00	3.80
17	4.80	2.80	2.67	3.00	3.00	3.50	4.00	4.00
18	4.80	2.80	2.67	3.00	3.00	3.50	4.00	4.00
19	4.80	2.80	2.67	3.00	3.00	3.50	4.00	4.00
20	3.80	4.20	2.67	3.67	3.00	1.75	3.75	3.60
21	4.40	4.20	2.67	4.00	3.20	3.00	3.00	3.60

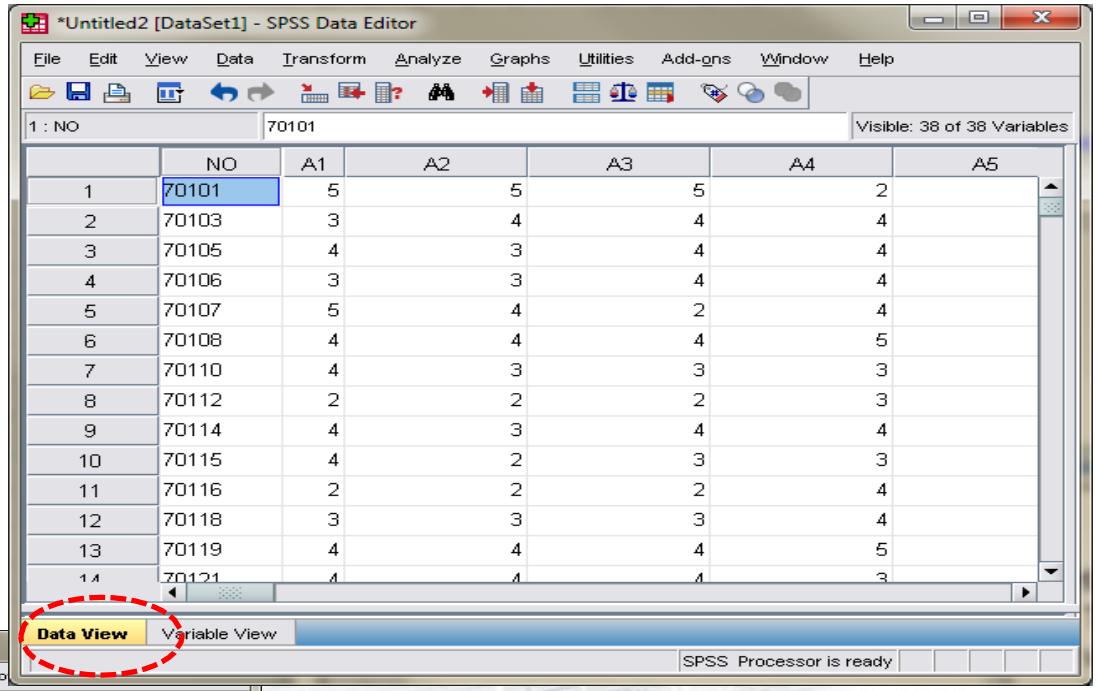
Data View Variable View

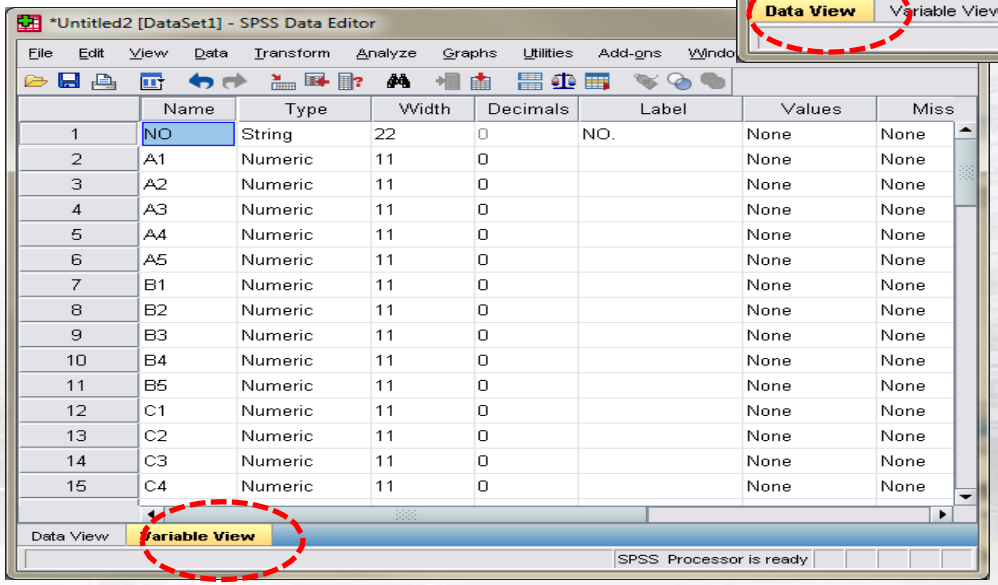
SPSS interface

数据视角 Data view

数据录入 The place to enter data

- 列：变量； 行：样本
  Columns: variables;
  Rows: records



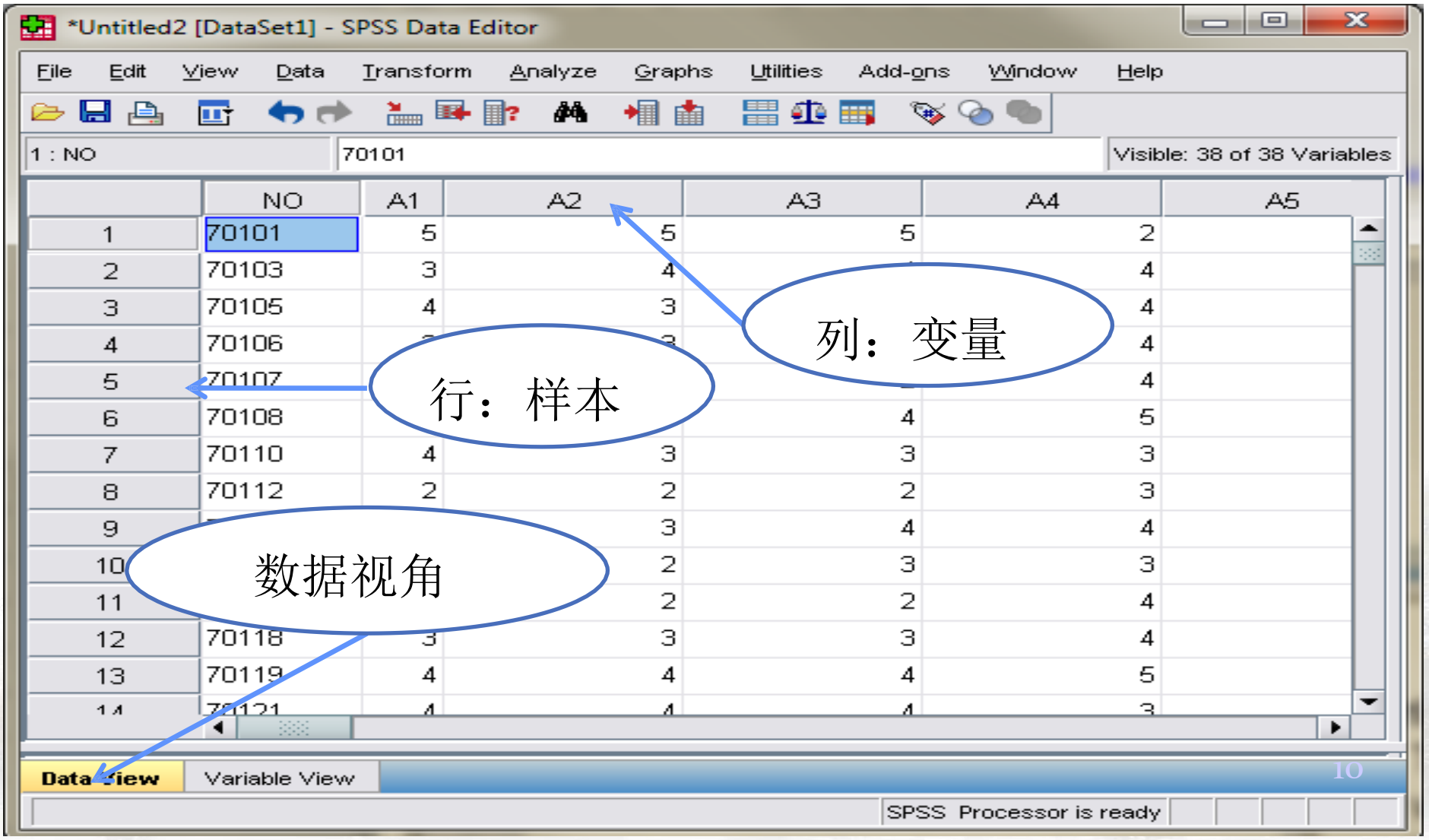


变量视角 Variable view

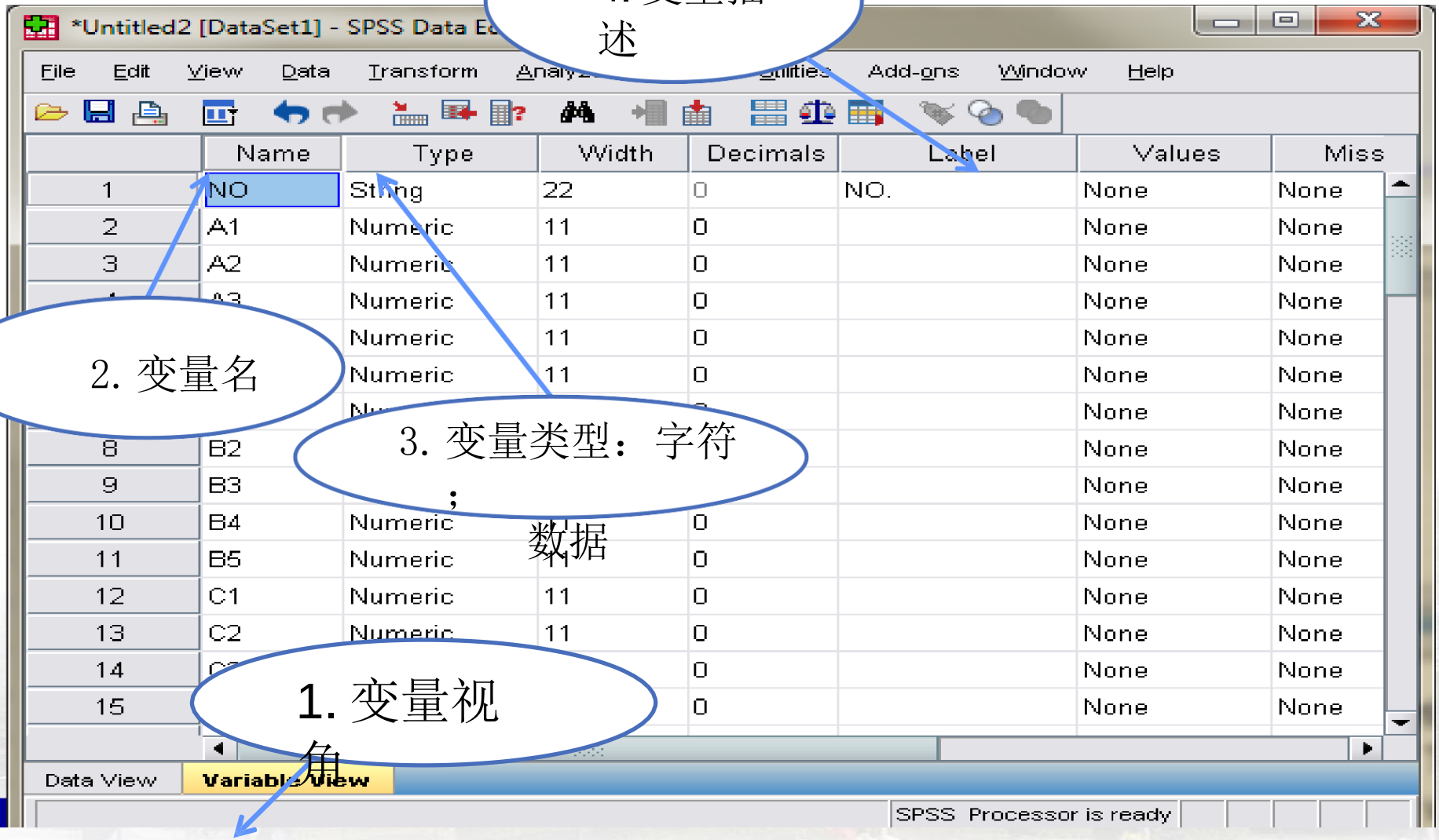
变量的录入 The place to enter variables

所有变量特征的列表
List of all variables
Characteristics of all variables

Enter data in SPSS 16.0

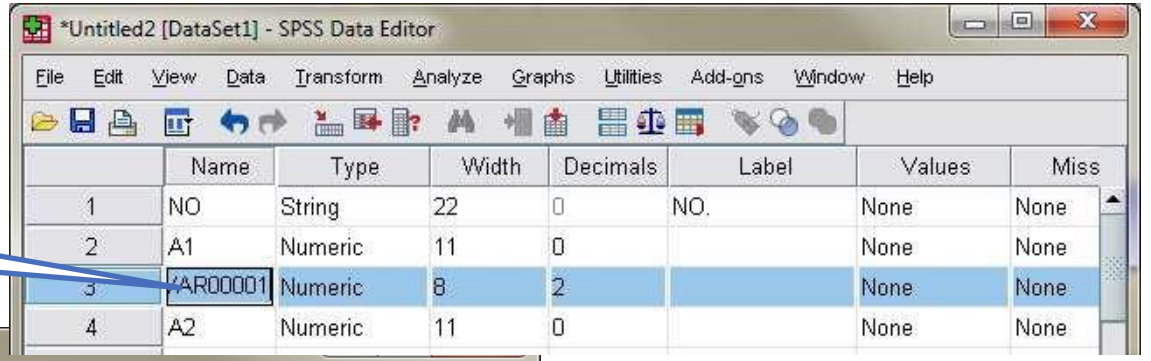


Enter variables

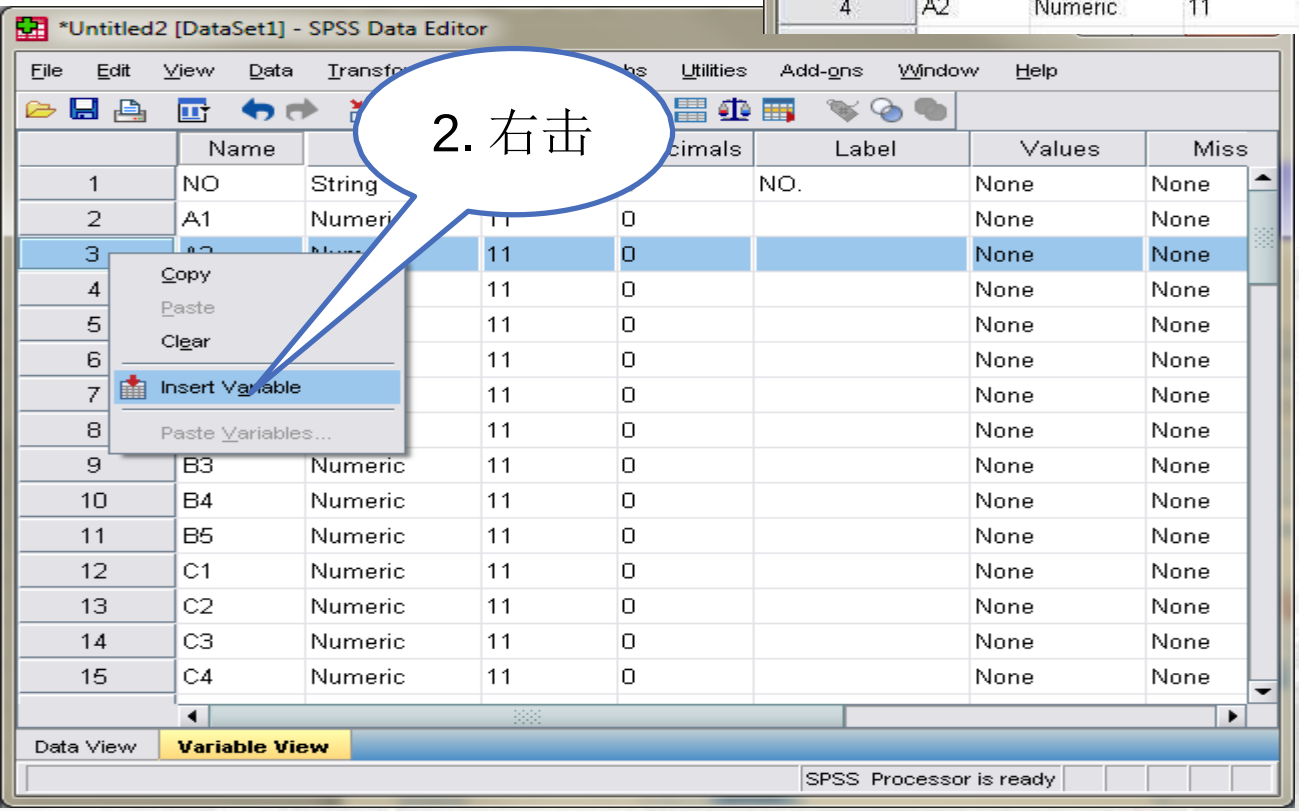


Enter variables

1. 双击改变



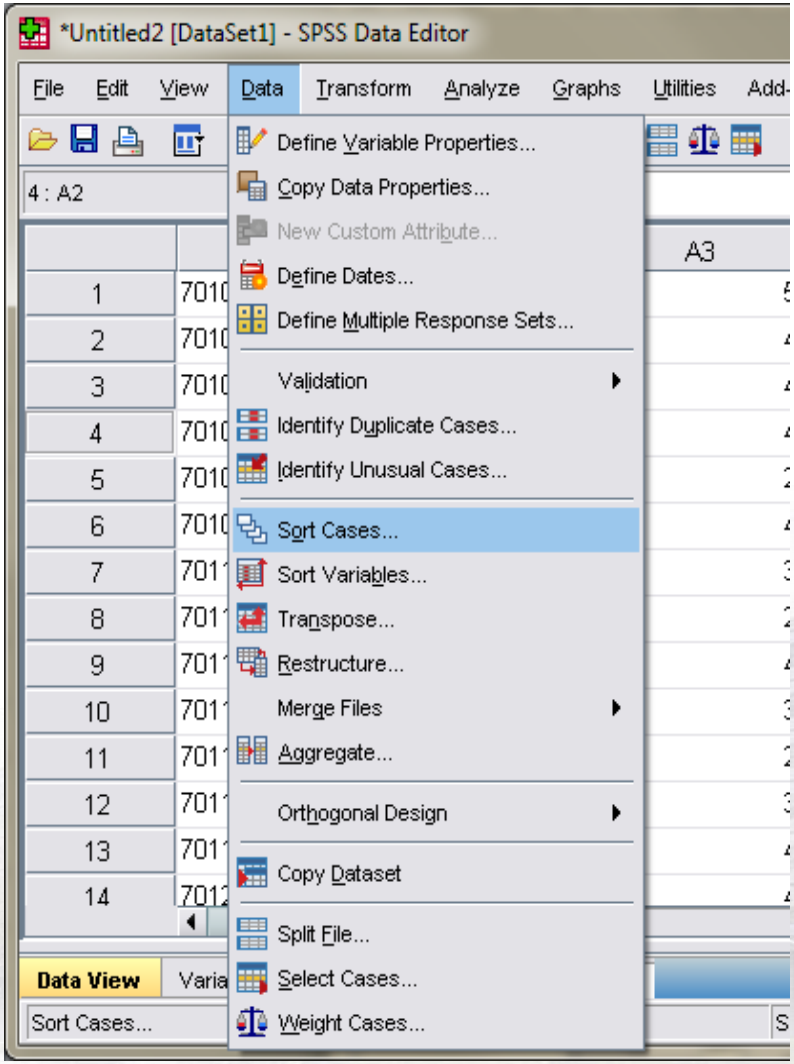
2. 右击

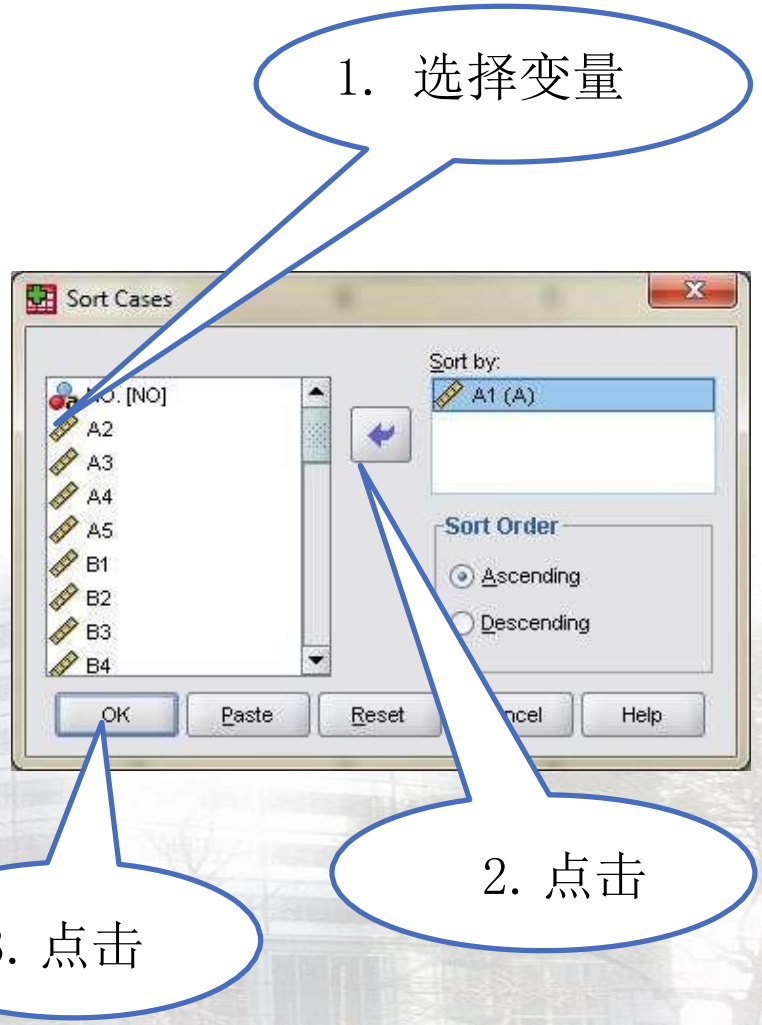


Clean data after import data files

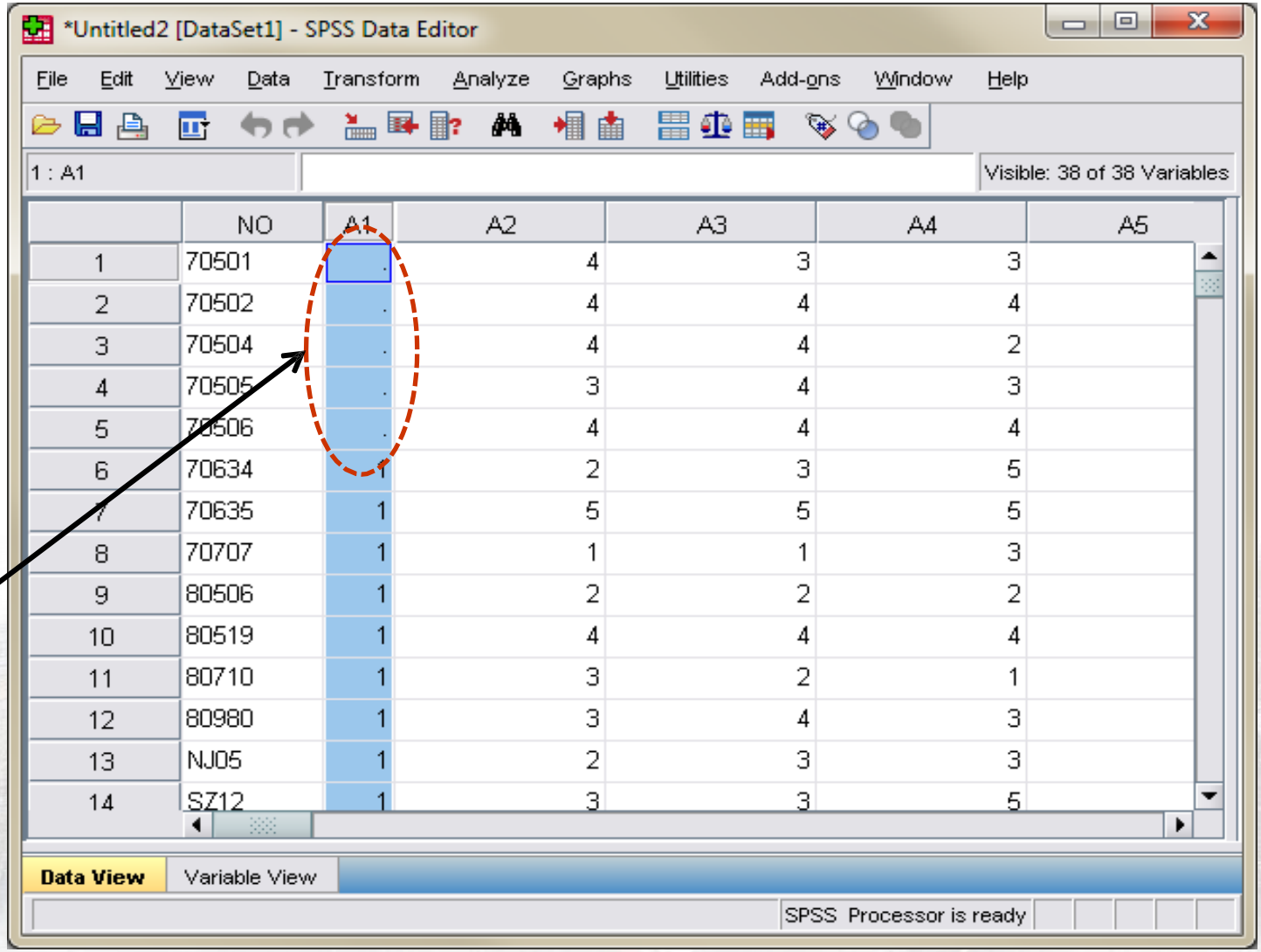
- 输入每个变量的数值和标注 Key in values and labels for each variable
- 数据排序 Sort the data
- 分析每个变量的频次 Run frequency for each variable
- 查看是否有错误数值 Check see if you have variables with wrong values
- 查看缺失值 Check missing values and physical surveys if you use paper surveys,and make sure they are real missing.
- 重新将字符变量变成数值变量 Sometimes, you need to recode string variables into numeric variables

Sorting the data





Missing values

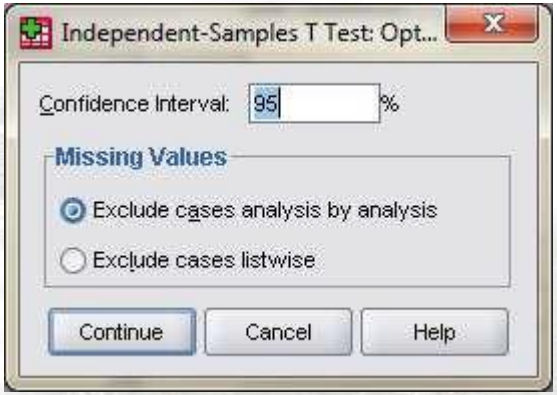


	NO	A1	A2	A3	A4	A5
1	70501	.	4	3	3	
2	70502	.	4	4	4	
3	70504	.	4	4	2	
4	70505	.	3	4	3	
5	70506	.	4	4	4	
6	70634	1	2	3	5	
7	70635	1	5	5	5	
8	70707	1	1	1	3	
9	80506	1	2	2	2	
10	80519	1	4	4	4	
11	80710	1	3	2	1	
12	80980	1	3	4	3	
13	NJ05	1	2	3	3	
14	SZ12	1	3	3	5	

缺失值

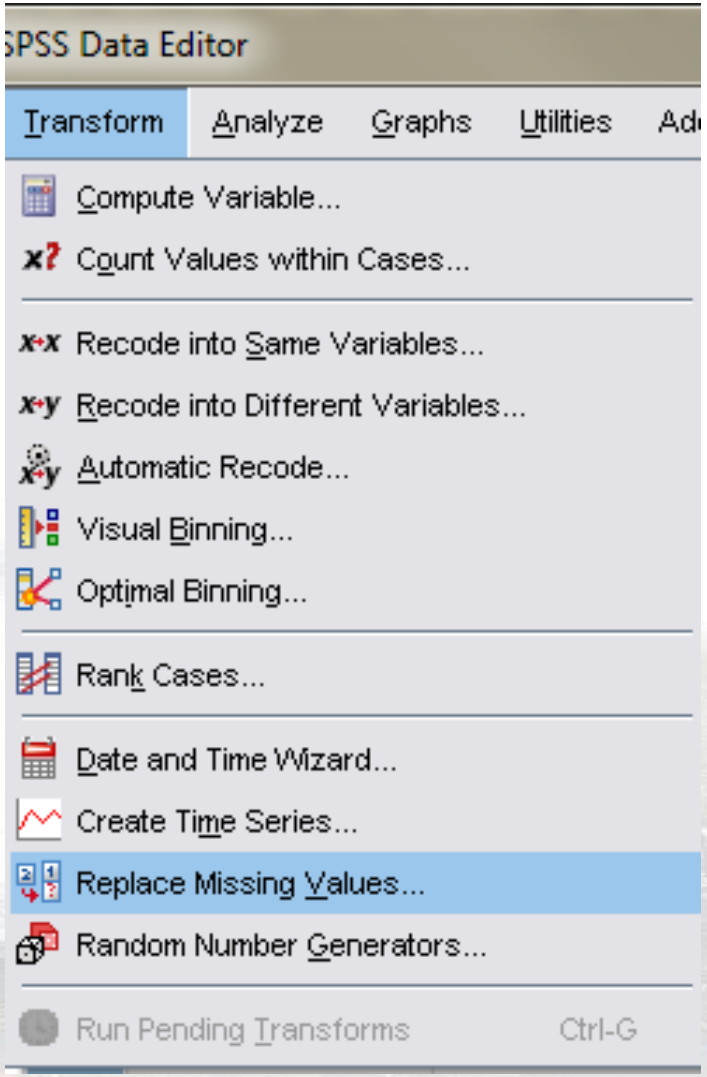
Missing values-methods

- 检查缺失值 Check for missing value
  - 用频次分析分类变量 Use Frequency for categorical variable
  - 用描述统计分析指标变量 Use Descriptive Stat. for measurement variable
- 分类变量 For categorical variables:
  - 如缺失值小于5%， 运用List-wise选项
  - 如缺失值大于于5%，增加一个新类别 If >=5%
- 指标变量变量 For measurement variables:
  - 如缺失值小于5%， 运用List-wise选项
  - 如介于5%和15%间，用替代法来替代缺失值
  - 如大于15%，剔除样本或变量

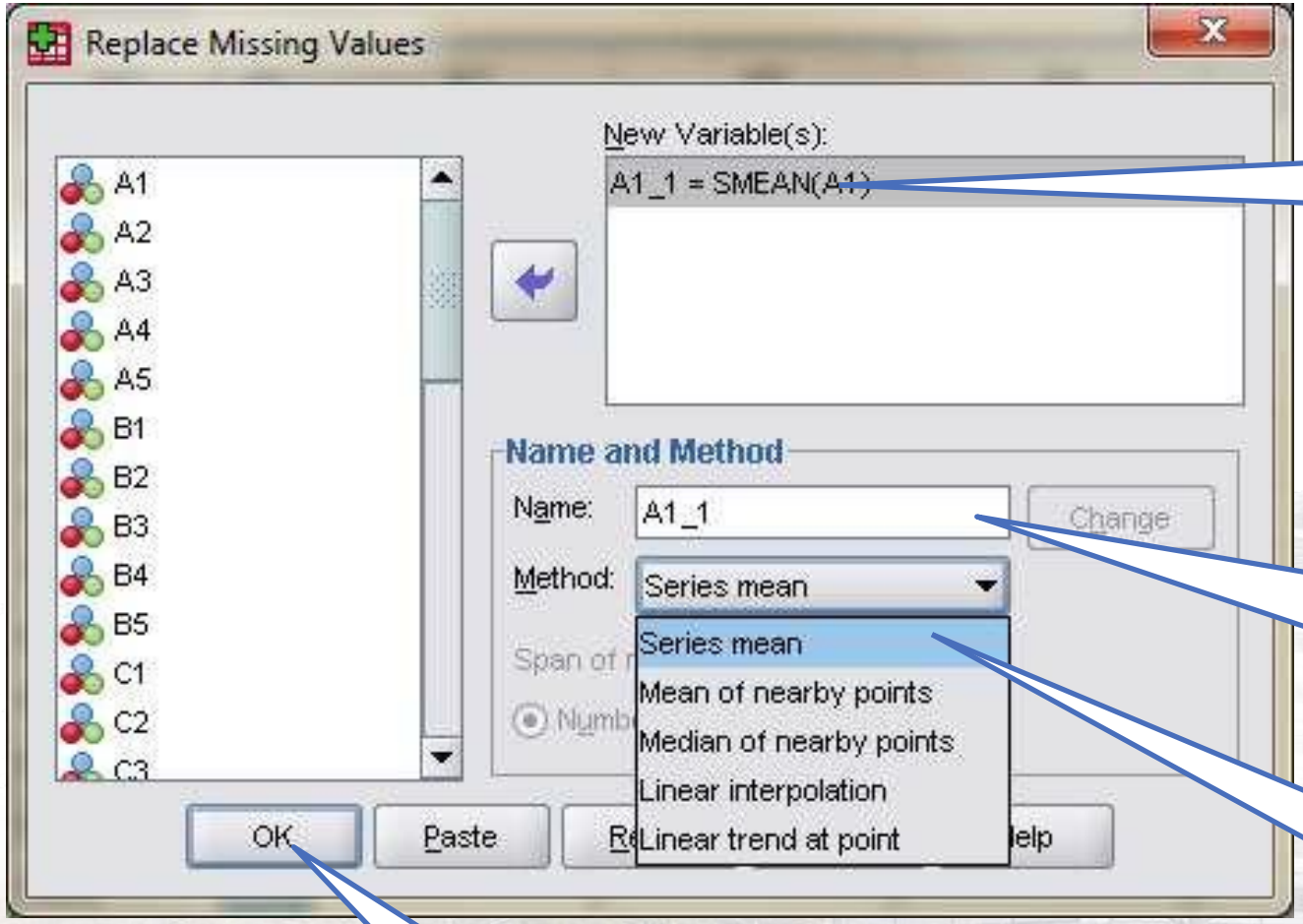


SPSS—Missing value

- Transform：
  - 缺失值 Missing Value
    - 均值替代 series mean
    - 近点有效数值的均值 mean of nearby point
    - 邻近点有效数值的中位数值 median of nearby point
    - 线性插值法 linear interpolation
    - 线性拟合方式 Linear trend at point



Missing value

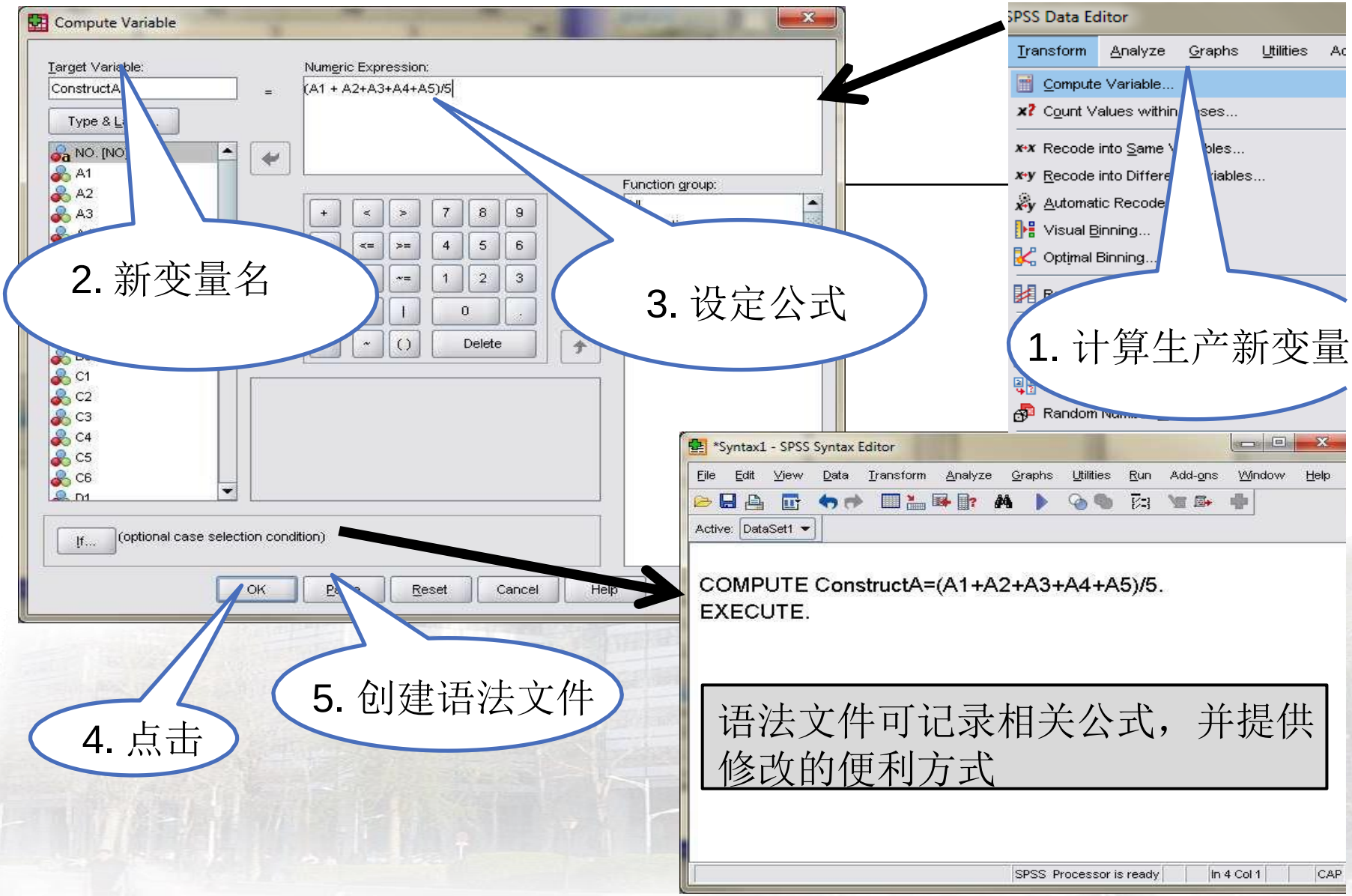


1. 选择有缺失值的变量

2. 如替代原有变量，可将名字改成原有变量名，后点击”change”

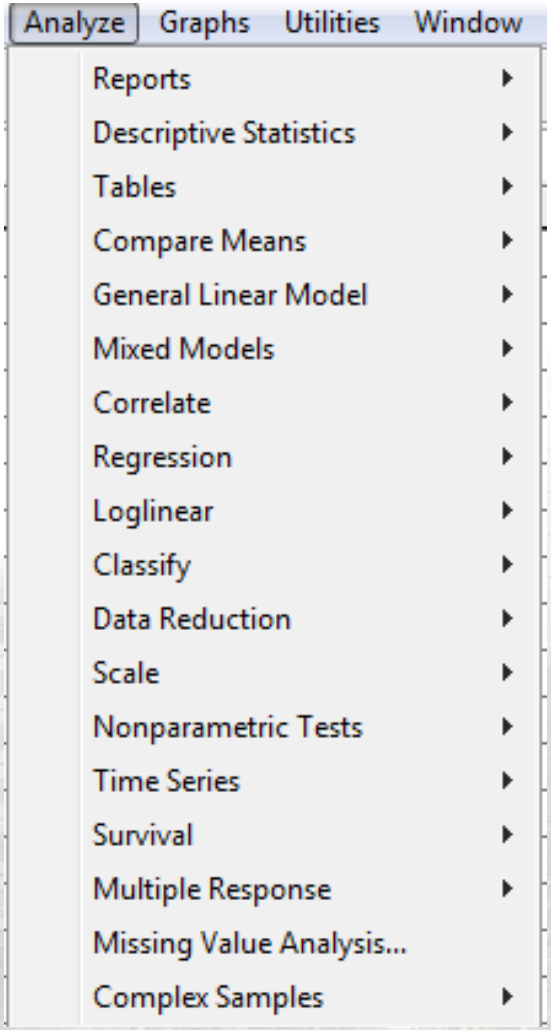
3. 选一方法

4. 点击


Compute Variable
Target Variable:
ConstructA
Numeric Expression:
(A1 + A2+A3+A4+A5)/5
Function group:
Delete
If...
(optional case selection condition)
OK
Reset
Cancel
Help
2. 新变量名
3. 设定公式
SPSS Data Editor
Transform
Analyze
Graphs
Utilities
Compute Variable...
Count Values within
Recode into Same
Recode into Differe
Automatic Recode
Visual Binning...
Optimal Binning...
1. 计算生产新变量
*Syntax1 - SPSS Syntax Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Run
Add-ons
Window
Help
Active: DataSet1
COMPUTE ConstructA=(A1+A2+A3+A4+A5)/5.
EXECUTE.
SPSS Processor is ready
In 4 Col 1
CAP
4. 点击
5. 创建语法文件
语法文件可记录相关公式，并提供修改的便利方式

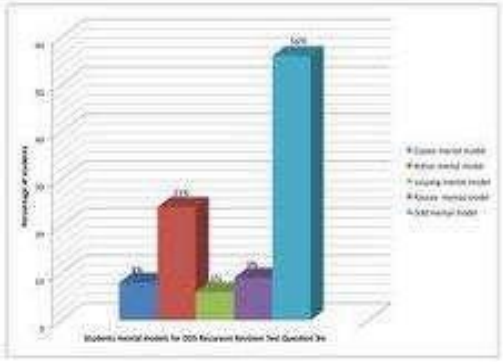
SPSS-- Analyze

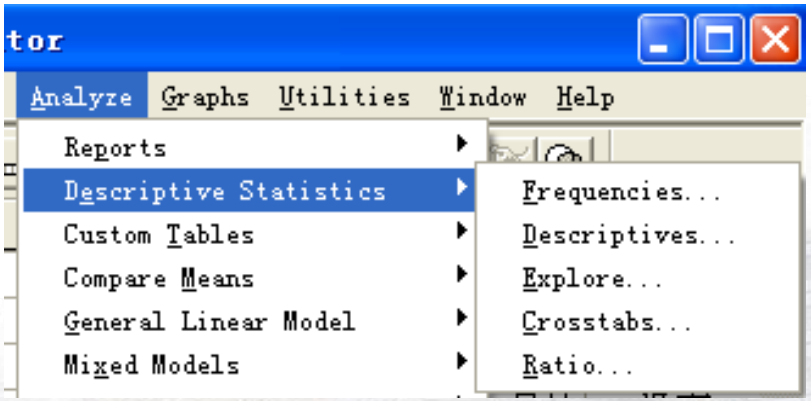
- **Analyze:**
  - Descriptive Statistics
  - Compare means
  - Correlate
  - Data Reduction
  - Scale
  - Regression



Descriptors of Research Design-- Purpose of the Study

Descriptive





- Who
- What
- Where
- When
- How much

Not WHY & HOW

TABLE 1
Summary Statistics

	Mean	s.d.	Minimum	Maximum
1. Early termination	0.13	0.34	0.00	1.00
2. Extend	0.14	0.35	0.00	1.00
3. Time	8.12	2.90	0.00	12.00
4. Time squared	74.38	39.95	0.00	144.00
5. Fixed fee	0.56	0.50	0.00	1.00
6. Time and materials	0.31	0.46	0.00	1.00
7. Hybrid	0.10	0.30	0.00	1.00
8. Compustar hardware	0.24	0.43	0.00	1.00
9. Mainframe	0.26	0.44	0.00	1.00
10. Proprietary	0.15	0.35	0.00	1.00
11. Other hardware	0.09	0.29	0.00	1.00
12. Programming	0.46	0.50	0.00	1.00
13. Interdependence	0.12	0.33	0.00	1.00
14. Innovation	2.53	1.20	1.00	6.00
15. Prior projects	7.90	6.10	0.00	15.23
16. Duration	2.11	1.06	0.18	4.95
17. Minimum distance	4.14	1.76	0.69	6.56
18. Any office 50	0.45	0.50	0.00	1.00
19. Measurement	0.44	0.50	0.00	1.00
20. Breadth	3.96	1.85	1.00	7.00
21. Lines	4.19	2.74	0.00	9.00
22. Reusability	0.39	0.49	0.00	1.00

Weber, L., Mayer, K.J., and Macher, J.T. (2011), "An Analysis of Extendibility and Early Termination Provisions: The Importance of Framing Duration Safeguards", *Academy of Management Journal, Vol.* 54, 182-202

Sample: Angst et al. 2010

Table 2 Descriptive Statistics for Full Sample and by U.S. Regions

Variable	Descriptives: Mean ± std. dev. or count (% of total) All	NE	MW	South	West	Description
HOSPITALS	3,989	641	1,056	1,571	720	Number of hospitals
HITC	34.43 ± 6.56	34.82 ± 5.09	32.99 ± 7.36	35.35 ± 6.31	34.17 ± 6.64	Concentration of HIT apps
PROFIT	665 (16.7%)	61 (9.5%)	120 (11.4%)	371 (23.6%)	113 (15.7%)	*For-profit* variable
TEACH	334 (8.4%)	121 (19.2%)	80 (7.6%)	88 (5.6%)	43 (6.0%)	Academic or not
HITINIT	16.74 ± 14.47	23.81 ± 14.47	16.24 ± 12.57	9.15 ± 3.46	27.72 ± 20.63	State HIT initiatives
IS BUDGET	2.84 ± 3.04	2.76 ± 2.60	3.07 ± 3.24	2.71 ± 3.01	2.85 ± 3.14	Percentage of budget for IS (scale from 1 to 10)[a]
SIZE	183.31 ± 163.62	229.35 ± 172.85	168.09 ± 167.28	180.84 ± 164.06	170.23 ± 140.04	Number of staffed beds
AGE	29.37 ± 34.09	38.35 ± 43.01	33.14 ± 36.27	24.18 ± 28.19	27.12 ± 31.19	When hosp. opened relative to 2005 (2001 is four years)
AGE × SIZE	6,635 ± 12,667	9,329 ± 14,616	6,820 ± 13,172	5,734 ± 11,987	5,933 ± 11,075	Number of staffed beds × age of hospital
WIRED[b]	0.24 ± 1.03	0.36 ± 1.40	0.23 ± 0.98	0.18 ± 0.80	0.15 ± 0.78	Number of times on Most Wired list in past seven years

[a]The mean of 2.84 falls into 1.5%–1.99% of total budget range.

[b]Although the award is entitled "The 100 Most Wired Hospitals," the total count on average exceeds 100 hospitals per year. This occurs for two reasons: (1) more than 100 hospitals made the list because of statistical ties, and/or (2) a health system, or a portion of a health system that includes multiple hospitals, was given the award.

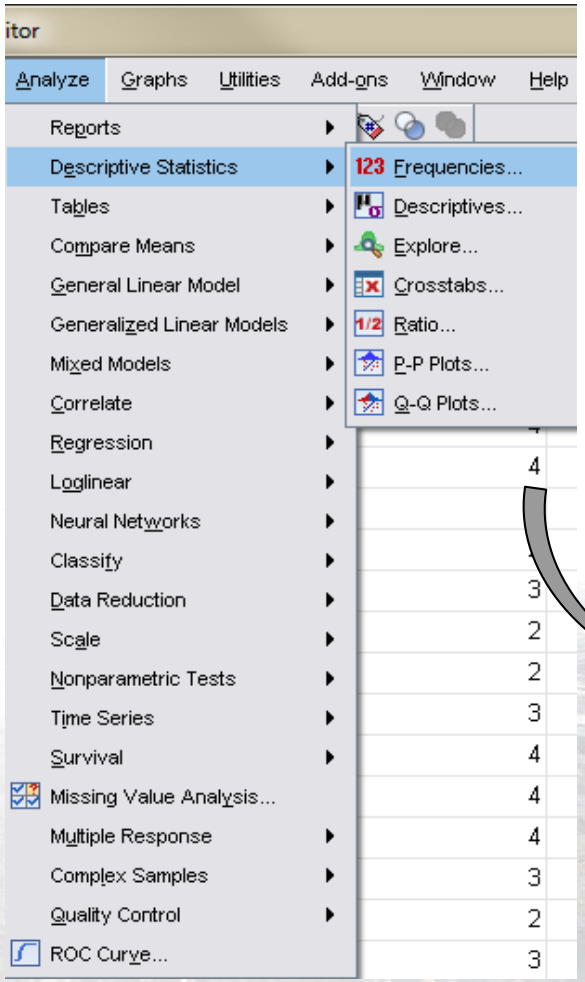
Angst, C.M., Agarwal, R., Sambamurthy, V., Kelley, K., 2010. Social Contagion and Information Technology Diffusion: The Adoption of Electronic Medical Records in US Hospitals. Management Science 56, 1219-1241

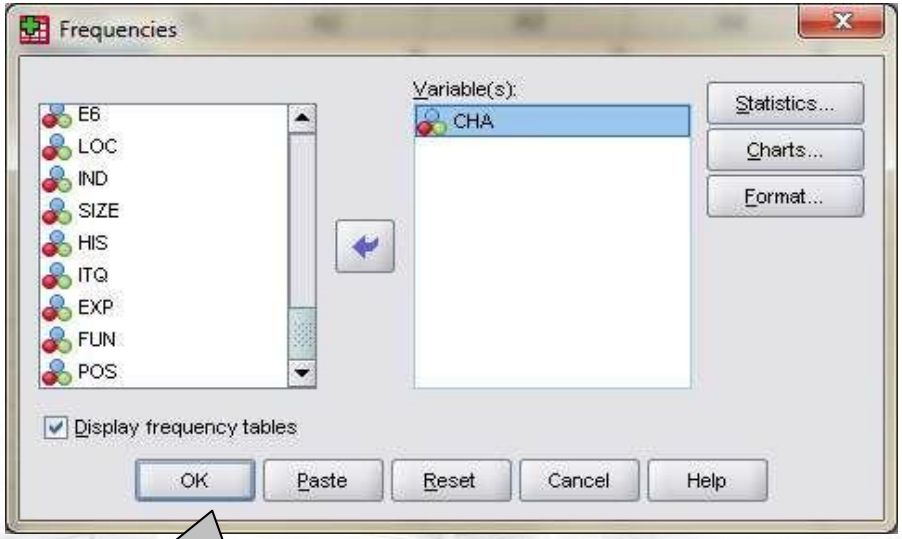
Table 3. Sample Characteristics (N = 211)

	Frequency	Percent
Revenues (2001)		
Less than $100 million (m)	15	6.2
$100 m - $250 m	75	31.1
$250 m - $500 m	54	22.4
$500 m - $1billion (b)	44	18.3
$1b - $2b	36	14.9
More than $2b	17	7.1
Industry Group		
Electronics and Computing Machinery	65	27.0
Wholesale and Retail	46	19.1
Financial Services	43	17.8
Business and Professional Services	25	10.4
Metals and Plastics	17	7.1
Pharmaceuticals and Healthcare	12	5.0
Other	33	13.6
Respondents		
IT Executive Survey		
Chief Information Officer	116	46.2
IT Director	50	20.7
SVP/VP, Information Technology	49	20.3
IT Manager	26	10.8
Strategic Planner Survey		
SVP/VP Corporate Development	113	46.9
Business Development Officer	60	24.9
VP Strategic Planning	37	15.3
Chief Financial Officer	31	12.9

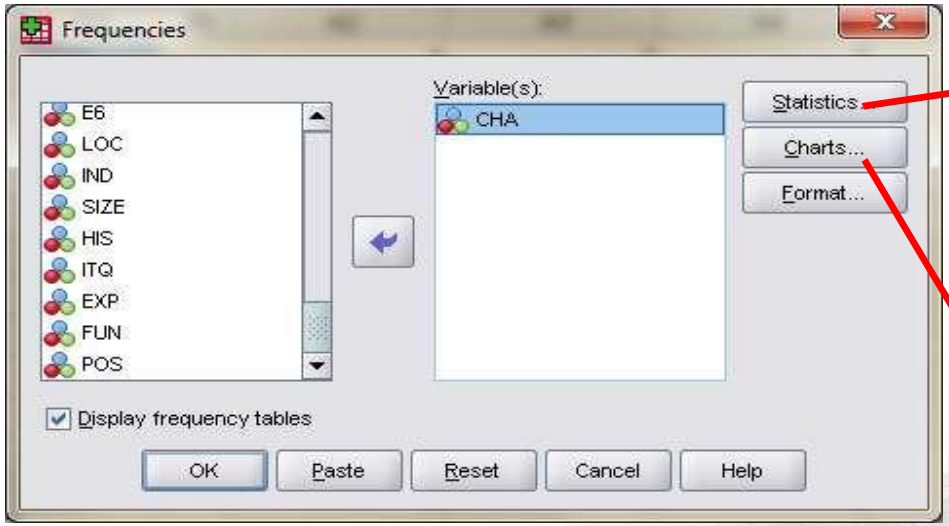
Tallon, P.P., and Pinsonneault, A. (2011), "Competing Perspectives on the Link between Strategic Information Technology Alignment and Organizational Agility: Insights from a Mediation Model", *MIS Quarterly, Vol.* 35, 463-484.

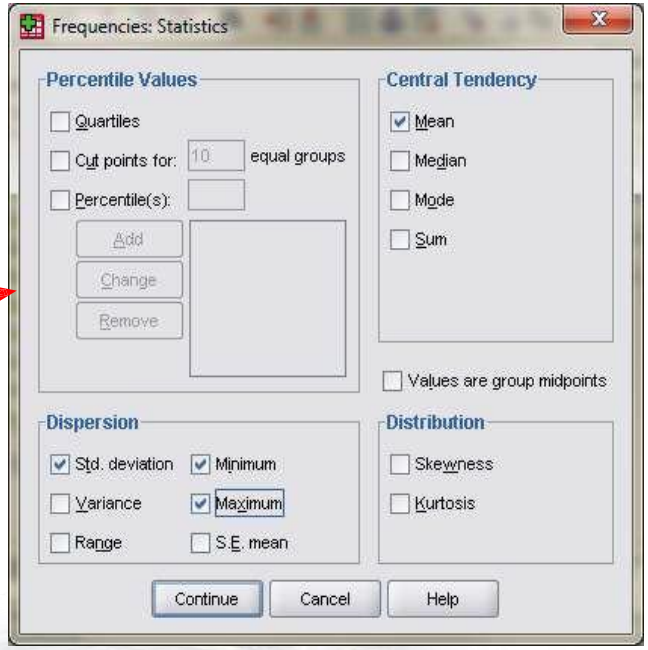
Descriptive Statistics——Frequencies

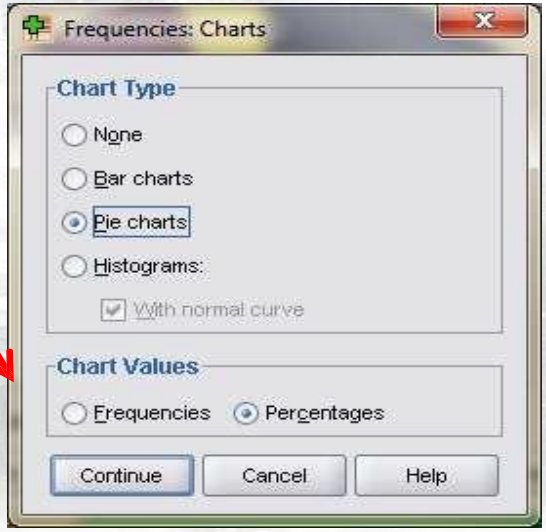




Descriptive——Frequencies







Descriptive——Frequencies

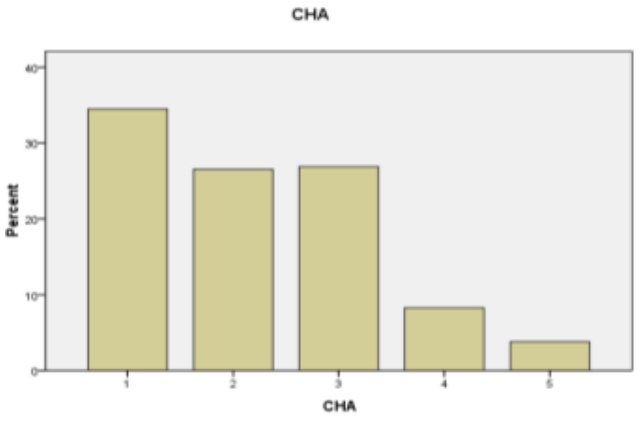
Mean		均数
Std. Error of Mean		标准误
Median		中位数
Mode		众数
Std. Deviation		标准差
Variance		方差
Skewness		峰度系数
Std. Error of Skewness		峰度系数的标准误
Kurtosis		偏度系数
Std. Error of Kurtosis		偏度系数的标准误
Range		全距
Minimum		最小值
Maximum		最大值
Sum		合计
Percentiles	**25**	**25%位数**
	50	**50%位数**
	75	**75%位数**

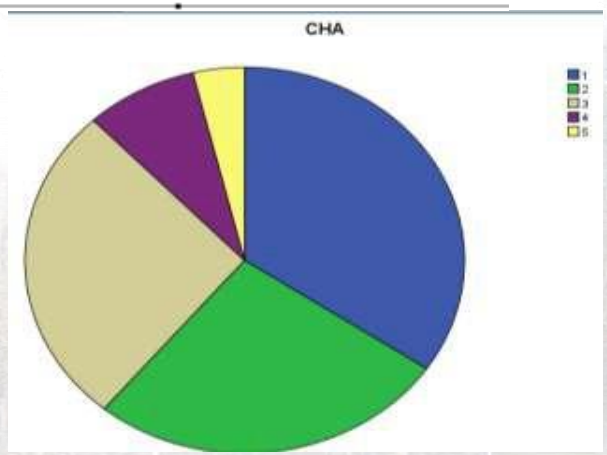
Statistics

身高

N	V alid	999
	M issing	1
M ean		160.639
S td. E rror of M ean		.2659
M edian		160.100
M ode		161.0
S td. Dev iation		8.4038
V ariance		70.6235
S kew ness		.226
S td. E rror of S kew ness		.077
Kurtosis		.405
S td. E rror of Kurtosis		.155
Range		69.1
M inimum		136.4
M aximum		205.5
S um		160478.0
P ercentiles	25	154.700
	50	160.100
	75	166.500

Frequency table





*Output4 [Document4] - SPSS Viewer

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

[DataSet1]

Statistics

CHA

N	Valid	290
	Missing	0
Mean		2.20
Std. Deviation		1.118
Minimum		1
Maximum		5

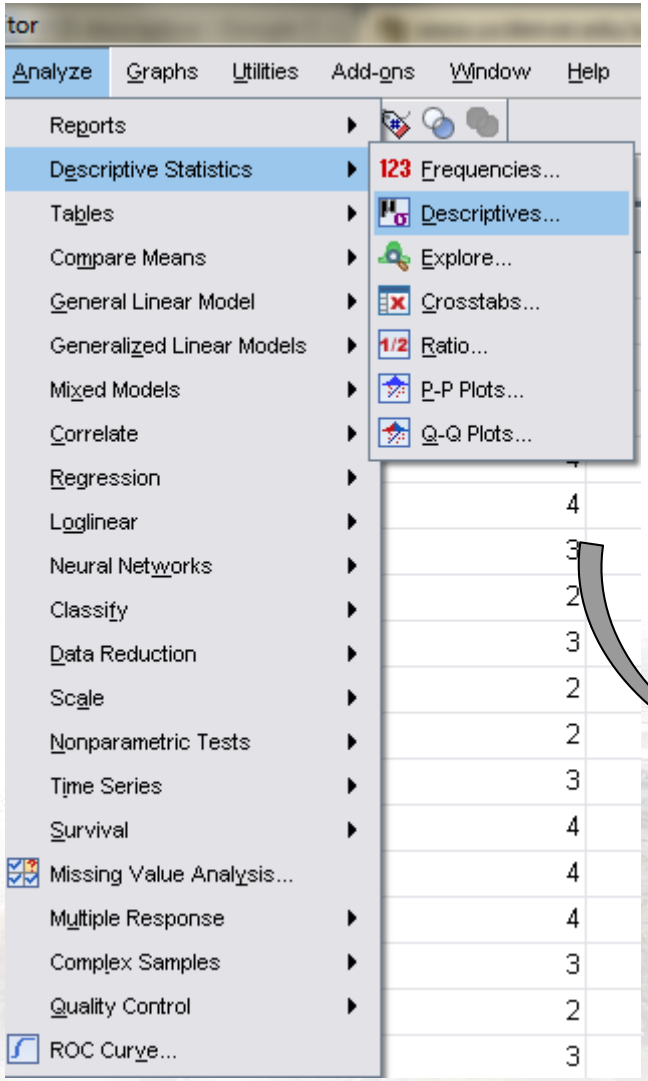
CHA

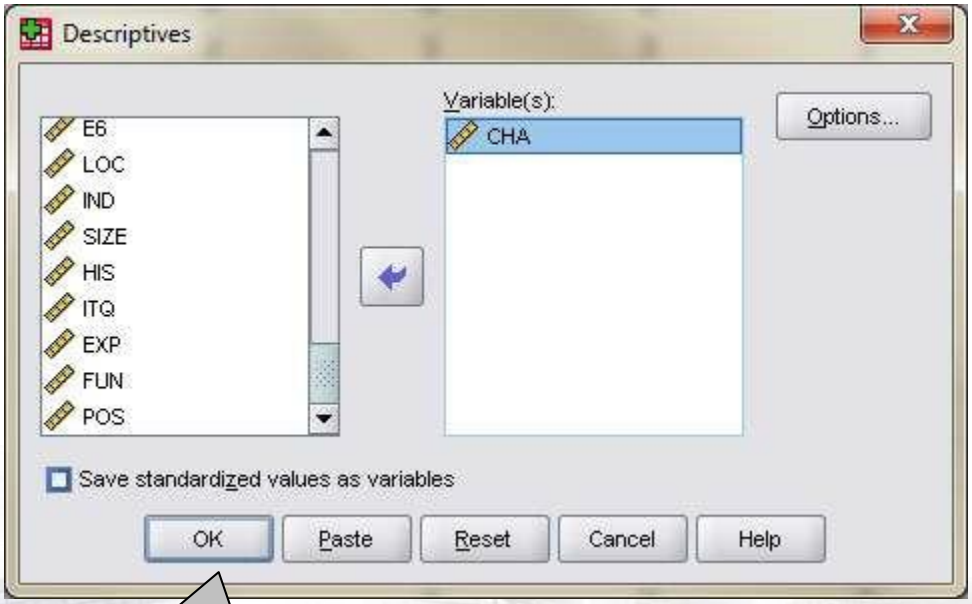
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	100	34.5	34.5	34.5
	2	77	26.6	26.6	61.0
	3	78	26.9	26.9	87.9
	4	24	8.3	8.3	96.2
	5	11	3.8	3.8	100.0
	Total	290	100.0	100.0	

CHA

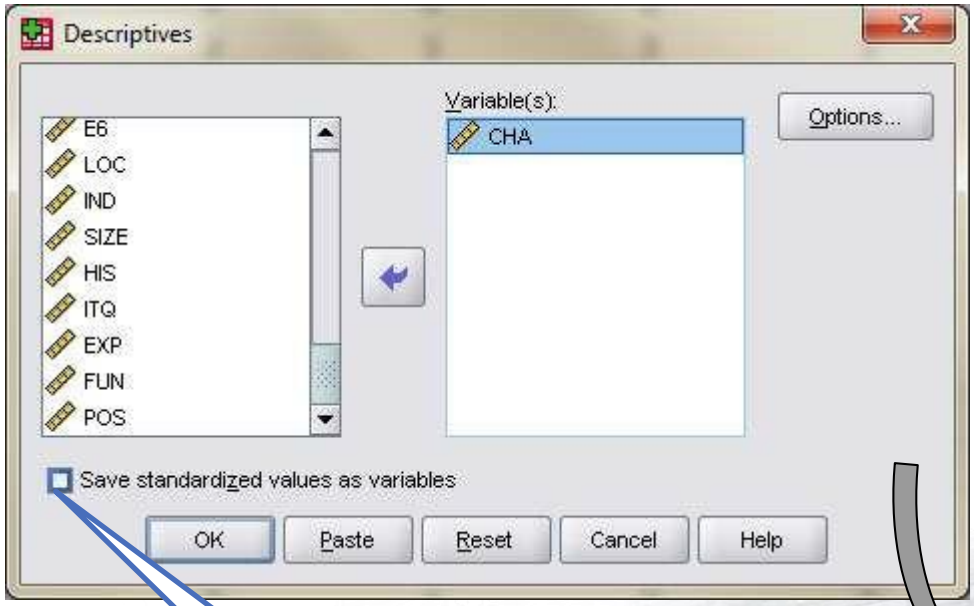
SPSS Processor is ready H: 502, W: 627 pt.

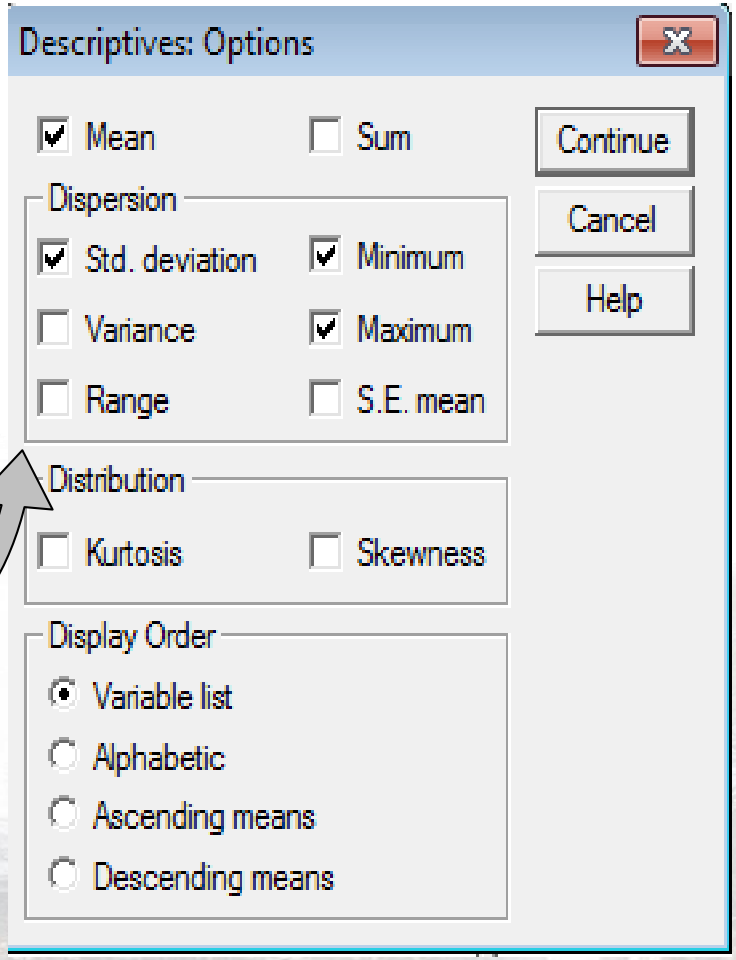
Descriptive——descriptives





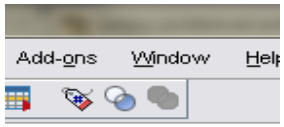
Descriptive——descriptives





选择标准数据

Descriptive——descriptives



同时进行多变量的基础性描述统计分析，只需将它们同时从左边选入右边框即可

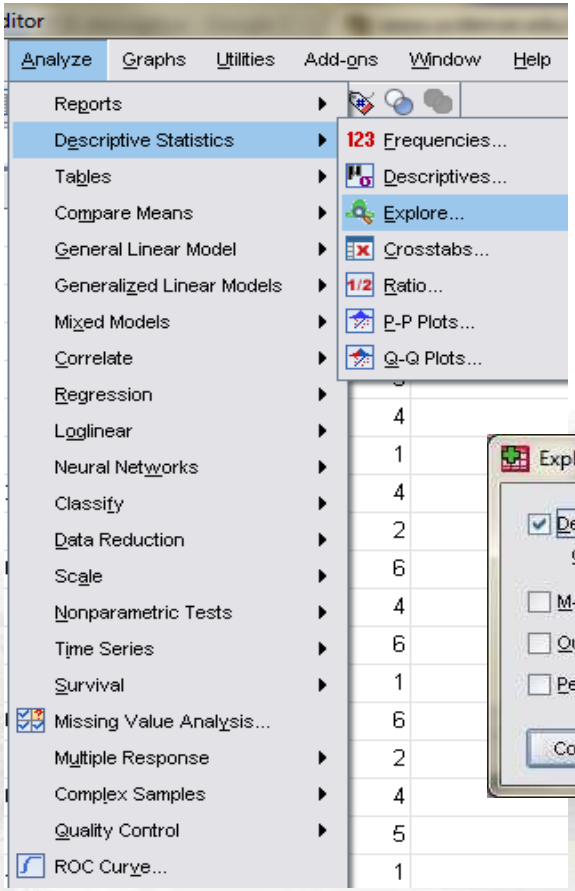
对原有变量标准化之后形成的新变量

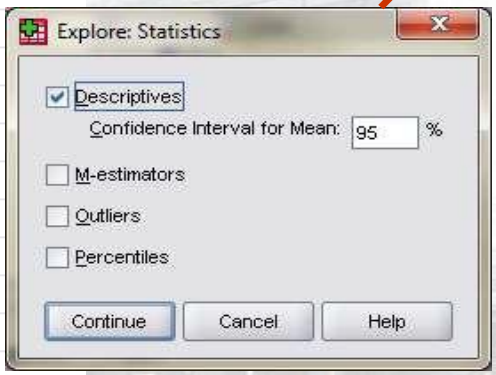
ZCHA
-1.07677
-0.18203
-1.07677
-1.07677
-1.07677
-1.07677
-0.18203
-0.18203
-1.07677
-1.07677
-1.07677
0.71271
1.60745
-0.18203
-0.18203
-1.07677
-1.07677
0.71271
-1.07677
-0.18203
-1.07677
-1.07677
-1.07677
-1.07677
-0.18203
-1.07677

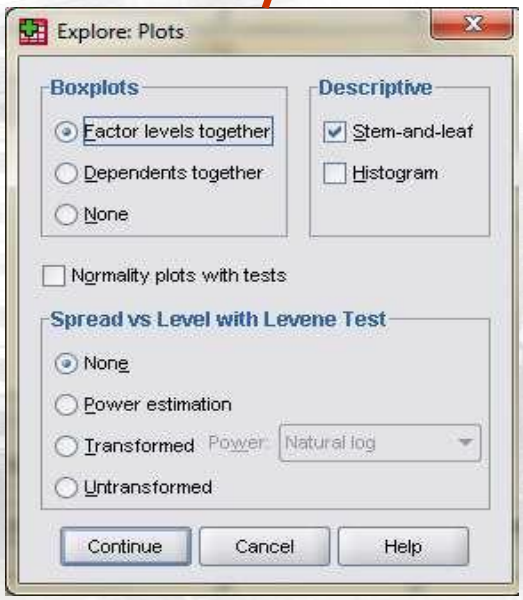
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
CHA	290	1	5	2.20	1.118
Valid N (listwise)	290				

Descriptive——Explore







Descriptive——Explore

增加 “INDS”变量后，结果将分成两部分—— 一组对应 制造业，一组对应服务业

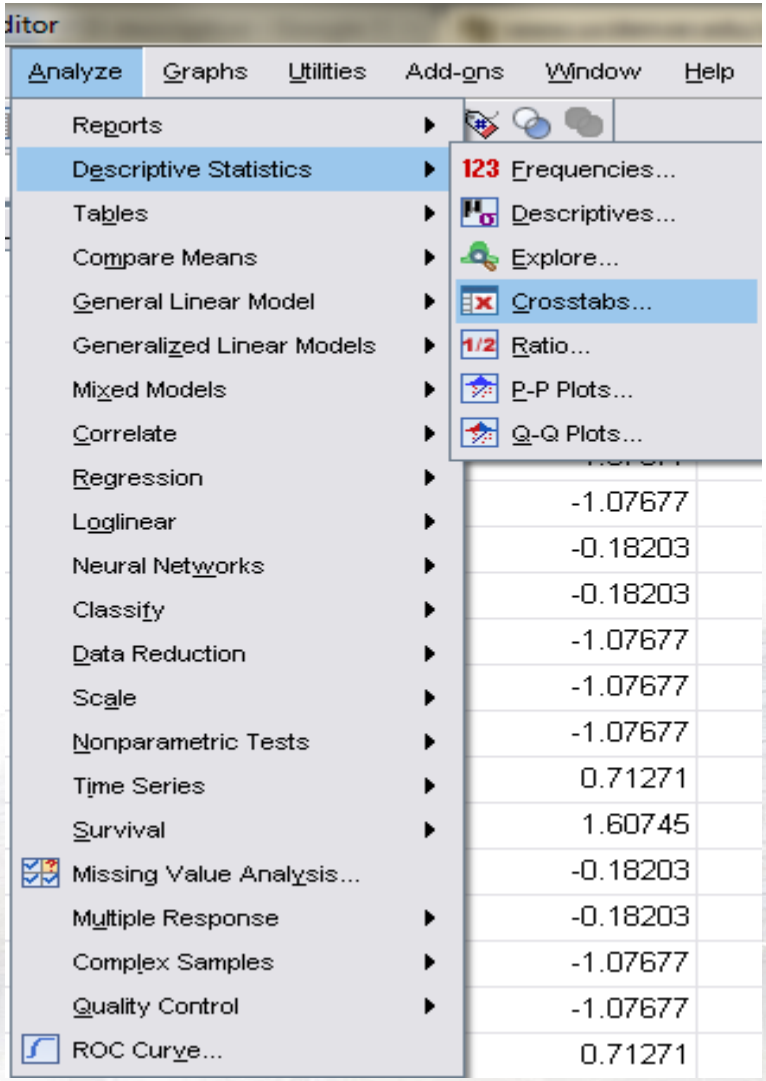
Descriptives

	INDS			Statistic	Std. Error
CHA	0	Mean		2.20	.104
		95% Confidence Interval for Mean	Lower Bound	2.00	
			Upper Bound	2.41	
		5% Trimmed Mean		2.12	
		Median		2.00	
		Variance		1.333	
		Std. Deviation		1.155	
		Minimum		1	
		Maximum		5	
		Range		4	
		Interquartile Range		2	
		Skewness		.595	.217
		Kurtosis		-.521	.431
	1	Mean		2.20	.085
		95% Confidence Interval for Mean	Lower Bound	2.04	
			Upper Bound	2.37	
		5% Trimmed Mean		2.13	
		Median		2.00	
		Variance		1.194	
		Std. Deviation		1.093	
		Minimum		1	
		Maximum		5	
		Range		4	
		Interquartile Range		2	
		Skewness		.628	.188
		Kurtosis		-.267	.375

Case Processing Summary

		Cases					
		Valid		Missing		Total	
	INDS	N	Percent	N	Percent	N	Percent
CHA	0	124	100.0%	0	.0%	124	100.0%
	1	166	100.0%	0	.0%	166	100.0%

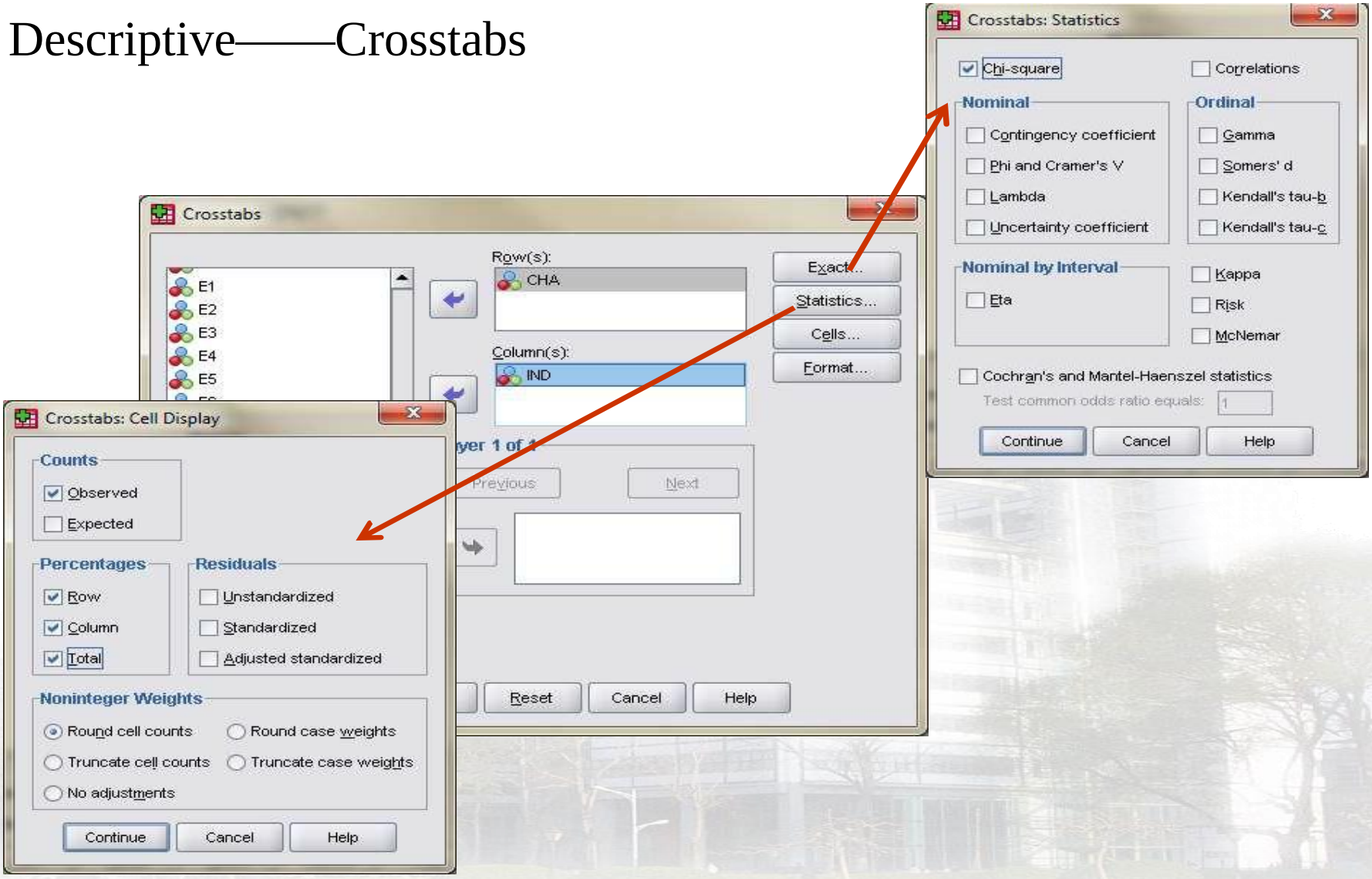
Descriptive——Crosstabs



在“crosstabs”分析中，可输入一个或更多变量作为频数表中的“行”，其它变量作为“列”

The “Crosstabs” dialog box lets you enter one variable (or more) as “rows” in a frequency table, and another variable as the “columns” in the same table.

Descriptive——Crosstabs



Descriptive——Crossta

基于双变量分组后每组的样本数 The total number of firms that have each individual combination of the various levels of the two variables.

基于某一变量分组的样本数和占 比The far-right column shows you the total for each diagnosis, as a percent of all firms.

基于另一变量分组的样本数和占 比 The bottom-most row slices the data the other way, showing you the total for each type of industry, as a percent of all firms.

CHA * INDS Crosstabulation

			INDS		
			0	1	Total
CHA	1	Count	46	54	100
		% within CHA	46.0%	54.0%	100.0%
		% within INDS	37.1%	32.5%	34.5%
		% of Total	15.9%	18.6%	34.5%
	2	Count	28	49	77
		% within CHA	36.4%	63.6%	100.0%
		% within INDS	22.6%	29.5%	26.6%
		% of Total	9.7%	16.9%	26.6%
	3	Count	34	44	78
		% within CHA	43.6%	56.4%	100.0%
		% within INDS	27.4%	26.5%	26.9%
		% of Total	11.7%	15.2%	26.9%
	4	Count	11	13	24
		% within CHA	45.8%	54.2%	100.0%
		% within INDS	8.9%	7.8%	8.3%
		% of Total	3.8%	4.5%	8.3%
	5	Count	5	6	11
		% within CHA	45.5%	54.5%	100.0%
		% within INDS	4.0%	3.6%	3.8%
		% of Total	1.7%	2.1%	3.8%
Total		Count	124	166	290
		% within CHA	42.8%	57.2%	100.0%
		% within INDS	100.0%	100.0%	100.0%
		% of Total	42.8%	57.2%	100.0%

This can be used to test the non-response bias

Manager = 131; and Director = 29. Non-response bias was examined using a Chi-square test; the non-significant χ^2 test results indicate the representativeness of the respondents for the sampling frame.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.863[a]	4	.761
Likelihood Ratio	1.880	4	.758
Linear-by-Linear Association	.001	1	.981
N of Valid Cases	290		

a. 1 cells (10.0%) have expected count less than 5. The minimum expected count is 4.70.

Table 2
Comparisons of sample and respondents

Variables	Sam		ts
SIC			
34	760 (26%)	70	83
35	680 (23%)	62	65
36	599 (20%)	55	58
37	490 (16%)	45	38
[illegible]	450 (15%)	41	29
Chi-square test ($\chi^2 = 7.6$, d.f. = 4, $P > 0.10$)			
Number of employees			
100–249	1280 (43%)	117	135
250–499	650 (22%)	60	63
500–999	419 (15%)	38	35
>1000	630 (20%)	58	40
Chi-square test ($\chi^2 = 8.6$, d.f. = 3, $P > 0.02$)			
Job title			
CEO/President	680 (23%)	62	70
Vice-president	459 (15%)	42	43
Manager	1610 (54%)	148	131
Director	230 (8%)	21	29
Chi-square test ($\chi^2 = 5.3$, d.f. = 3, $P > 0.10$)			
Total	2979[a] (100)	273	273

Zhang, Q. Y., Vonderembse, M. A., & Lim, J. S. (2003). Manufacturing flexibility: defining and analyzing relationships among competence, capability, and customer satisfaction. *Journal of Operations Management*, 21(2), 173-191.

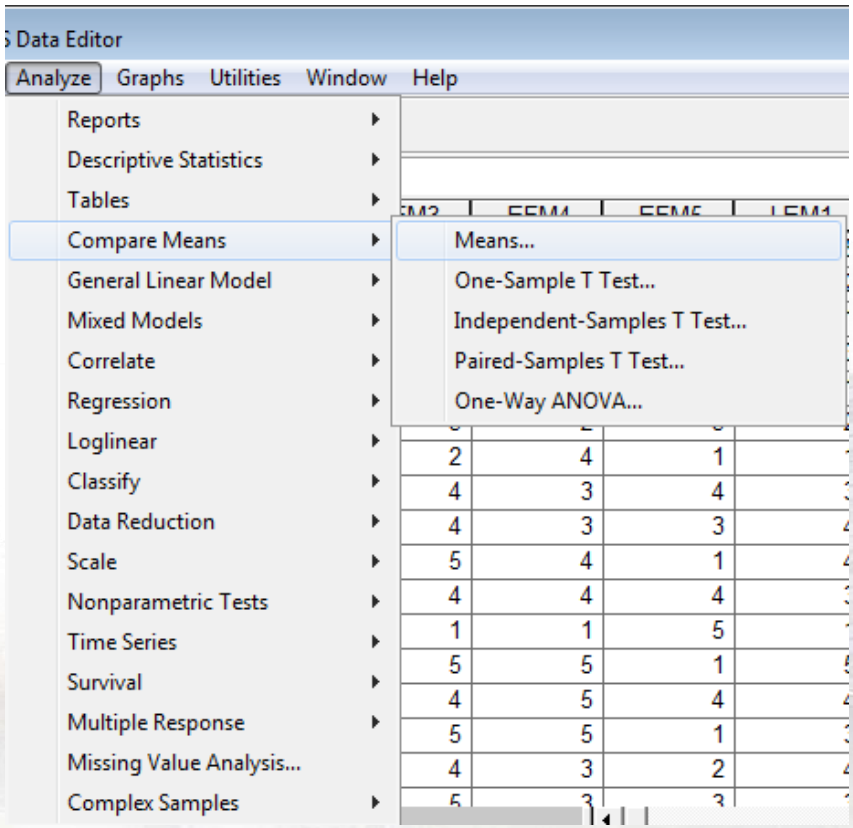
SPSS的均值检测(Compare Means)

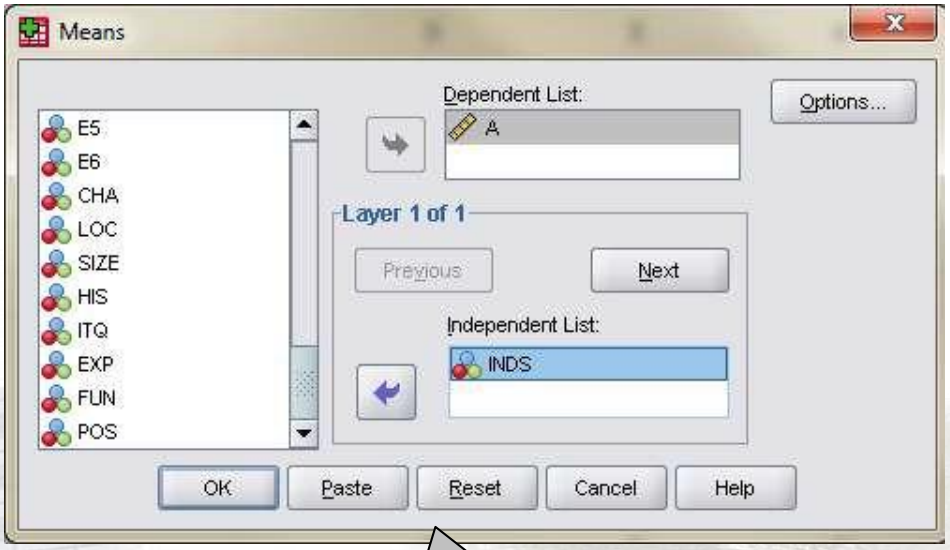
- Means：平均数
- One-samples T test：单一样本检测
- Independent-samples T test：独立样本检测
- Paired-samples T test：成对样本检测
- One-way ANOVA：单因素方差分析

平均数(Means)

分析基于产业类别分类的特定构念的均值差异

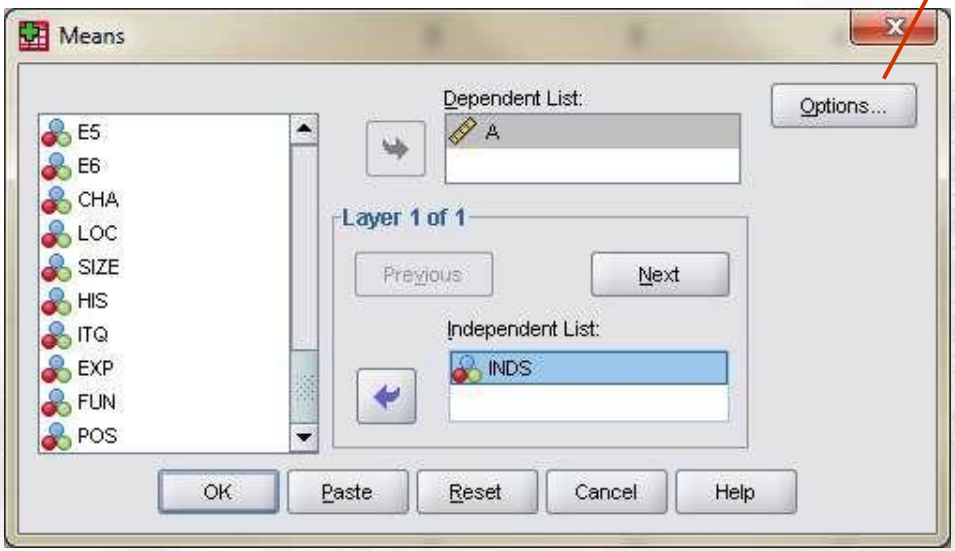
Analyze the difference of mean of a specific construct based on Industry type

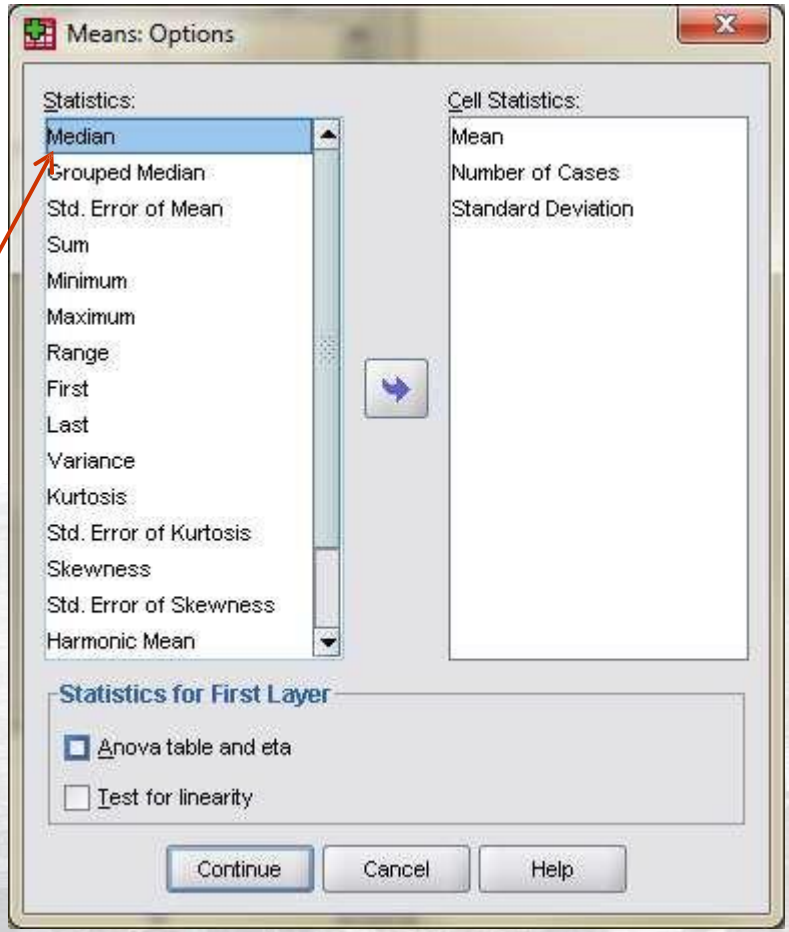




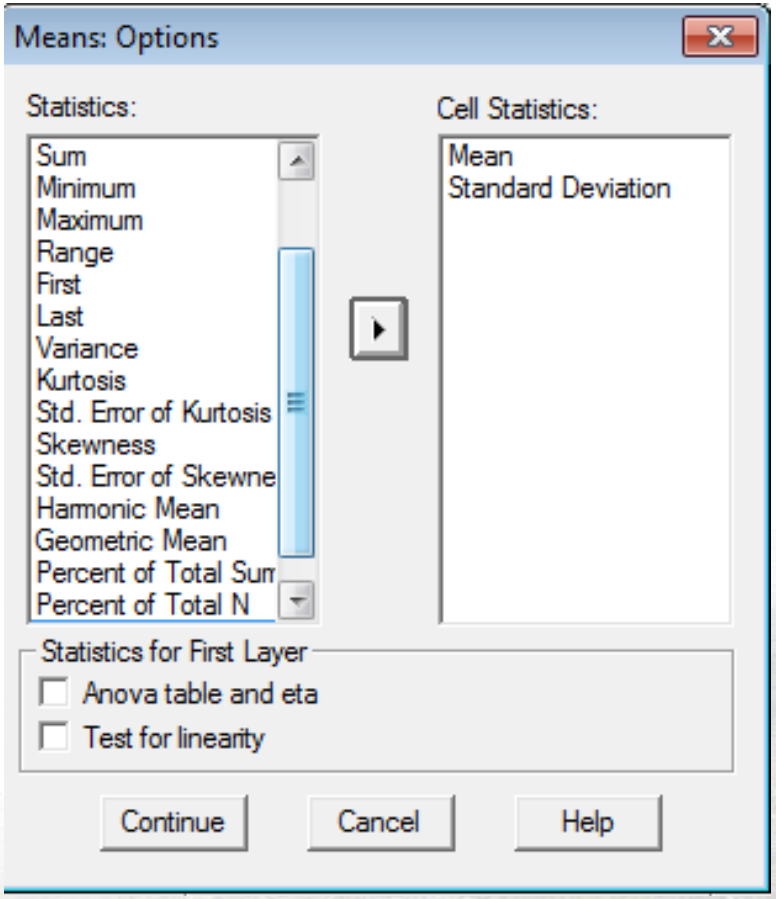
平均数(Means)

1. 将感兴趣的变量选入因变量框 Move the variables of interest [A] to the Variables list box.
2. 将分组变量选入自变量框 To subdivide the dataset into groups, move the grouping variable [INDS] to the Grouping Variable(s) list box.





平均数(Means)



Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
A * INDS	290	100.0%	0	.0%	290	100.0%

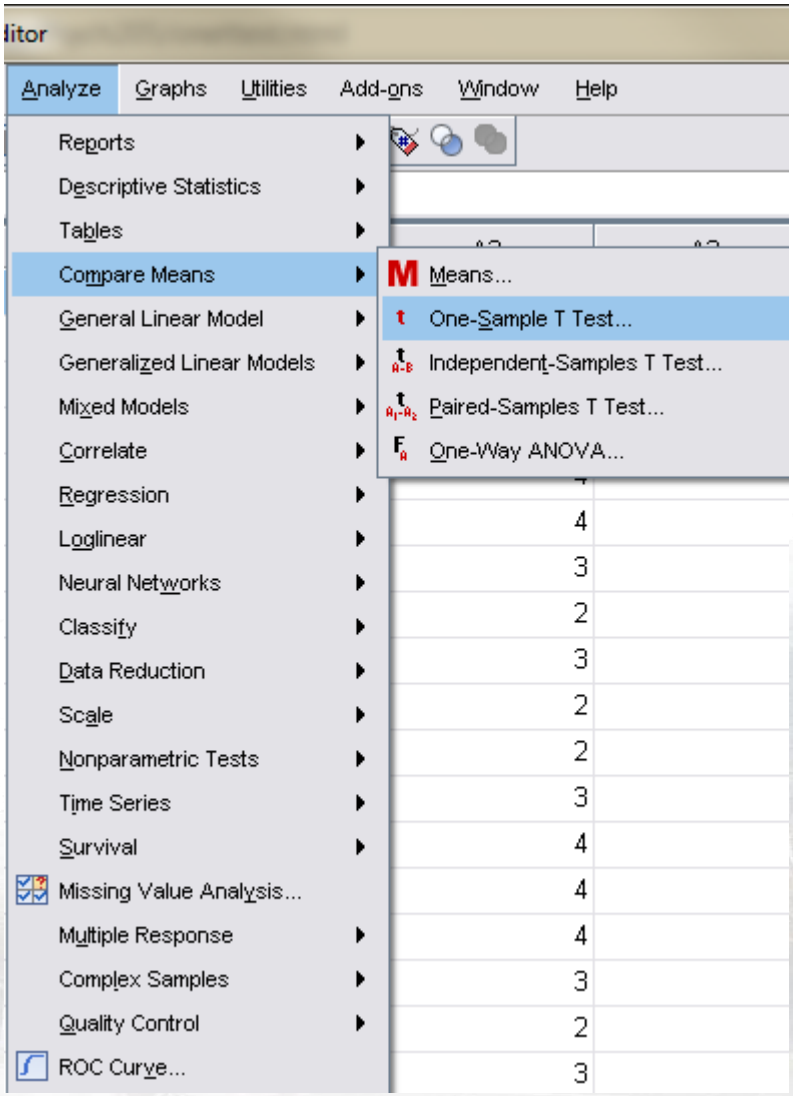
Report

A

INDS	Mean	N	Std. Deviation
0	3.5774	124	.63870
1	3.5940	166	.70201
Total	3.5869	290	.67458

The mean of A is close between INDS1 and INDS0

单一样本平均数检测(One-Sample T Test)



比较样本均值与某一特定值的差异。特定值通常是总体的均值

compares the mean score of a sample to a known value.

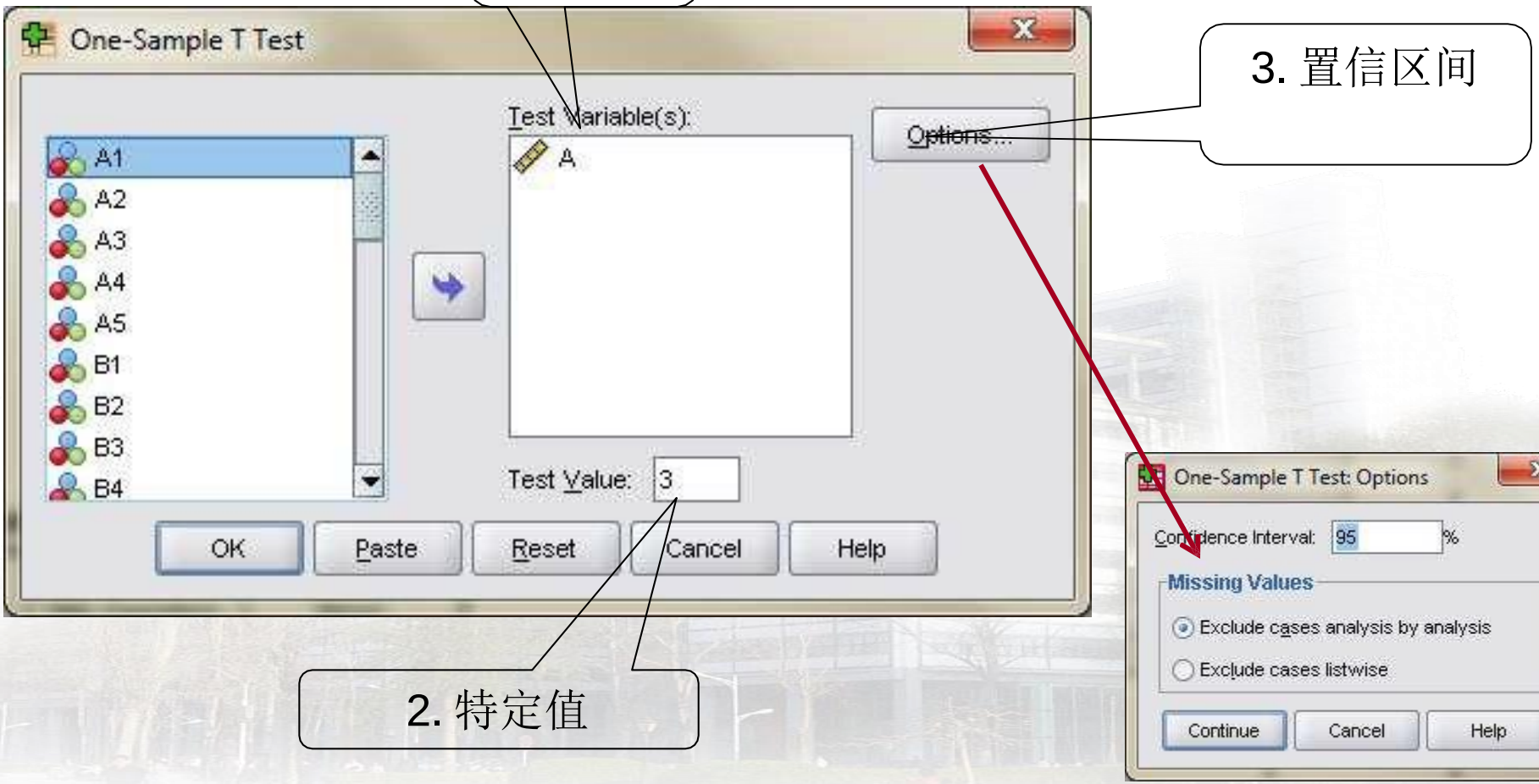
Usually, the known value is a population mean.

Hypotheses:

•**Null:** 样本均值和总体均值不存在显著差异There is no significant difference between the sample mean and the population mean.

•**Alternate:**样本均值和总体均值存在 显著 差 异 There is a significant difference between the sample mean and the population mean.

单一样本平均数检测(One-Sample T Test)



单一样本平均数检测(One-Sample T Test)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
A	290	3.5869	.67458	.03961

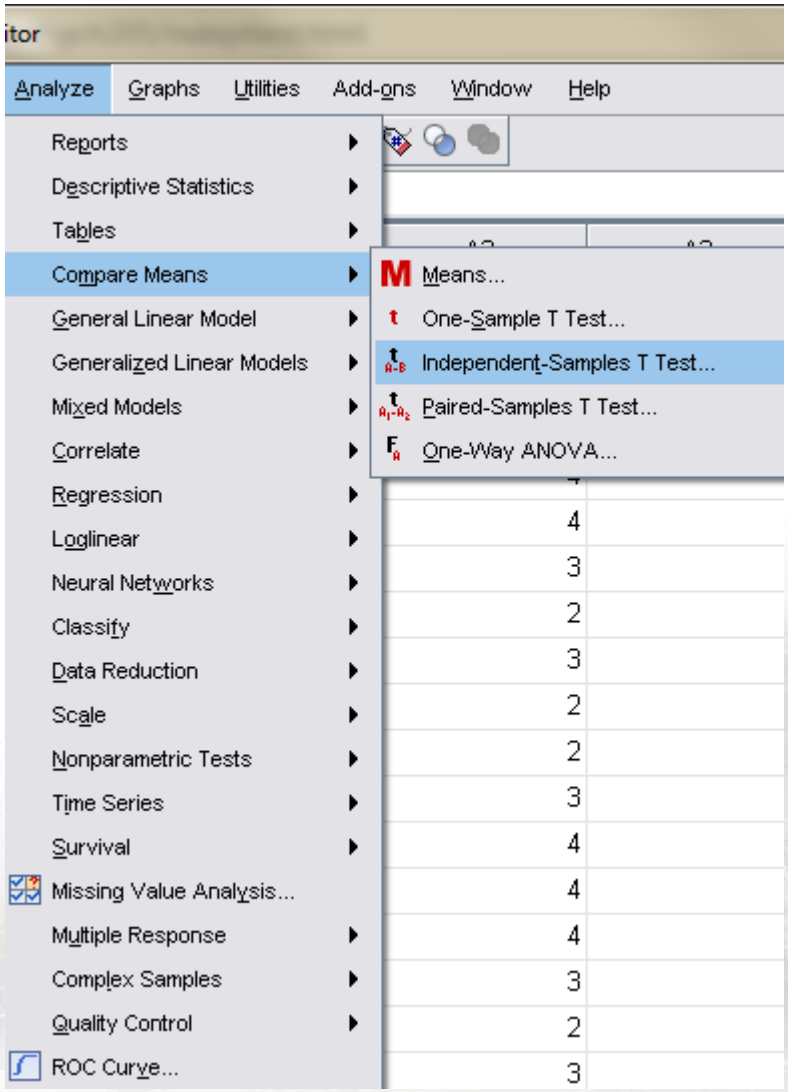
样本均值为3.588, 略高于总体均值3

One-Sample Test

	Test Value = 3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
A	14.816	289	.000	.58690	.5089	.6649

两者之间有显著差异，即样本均值3.5869 显著高于总体均值3.

独立样本检测（Independent-samples T test）



比较某一变量在两组数据中的均值

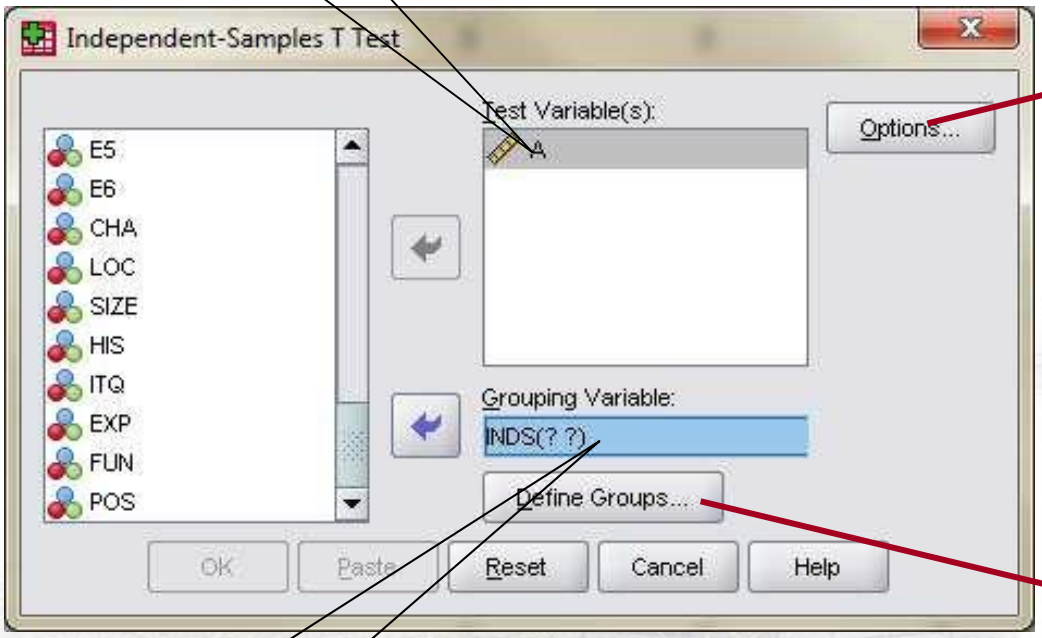
compares the mean scores of two groups on a given variable

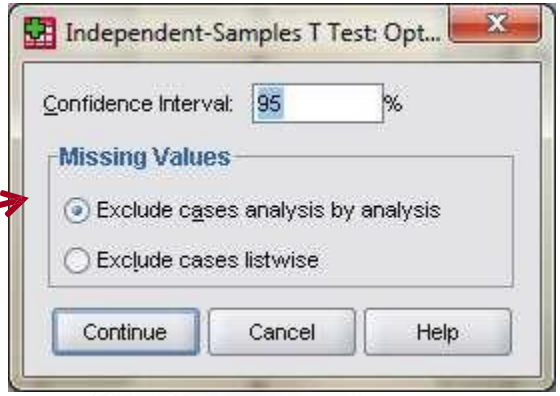
Hypotheses:

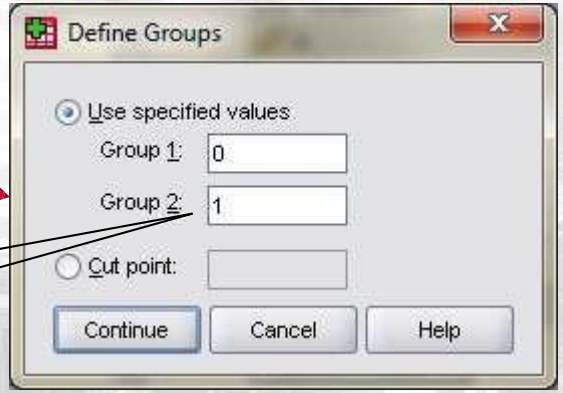
- **Null:** 均值在两组中不存在显著差 异 The means of the two groups are not significantly different.
- **Alternate:** 均值在两组中存在显著 差 异 The means of the two groups are significantly different.

独立样本检测（Independent-samples T test）

1. 分析变量







2. 分组变量

0: Manufacturing
1: Service

独立样本检测（Independent-samples T test）

“Manufacturing”组
高于 “Service”组

Group Statistics

	INDS	N	Mean	Std. Deviation	Std. Error Mean
A	0	124	3.5774	.63870	.05736
	1	166	3.5940	.70201	.05449

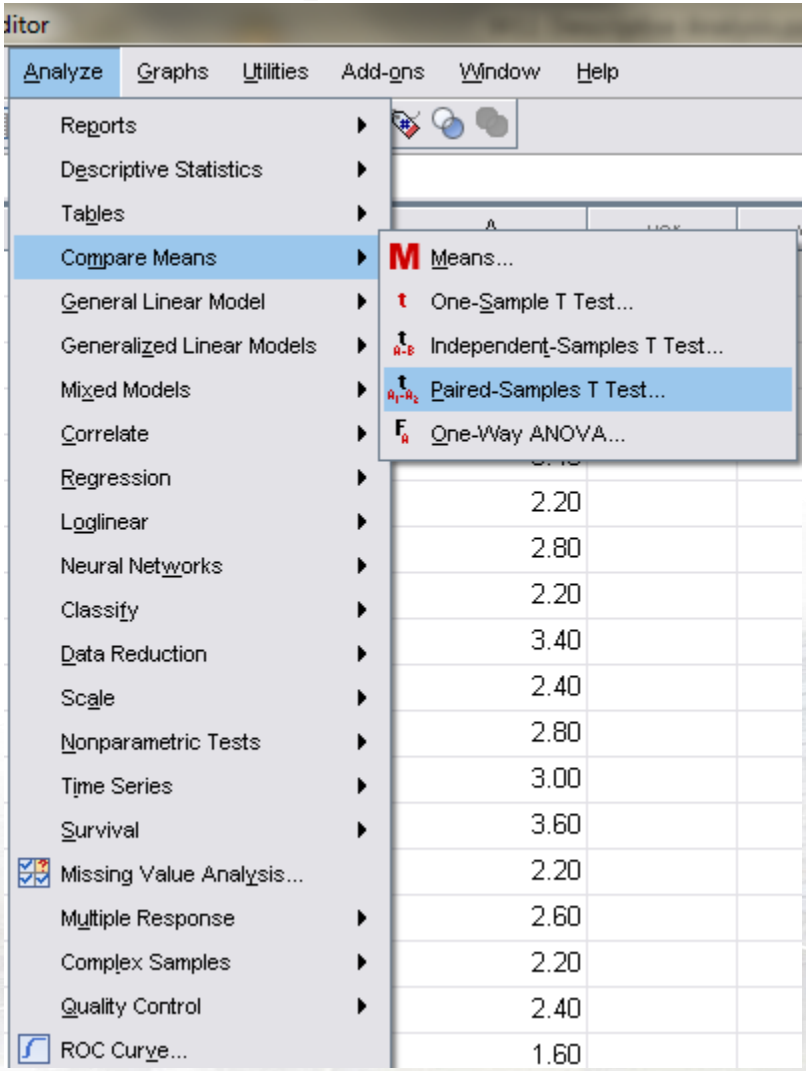
Indeper

		Levene's Test for Equality of Variances		
		F	Sig.	t
A	Equal variances assumed	1.619	.204	-.206
	Equal variances not assumed			-.209

the Levene's Test for Equality of Variances: Significant (<.05) = significantly different.
Not significant= are equal
we have met our second assumption.

Independent Samples T Test:
如方差显著差异，则看第一行数据
如方差无显著差异，则看底行数据

Paired Samples T Test



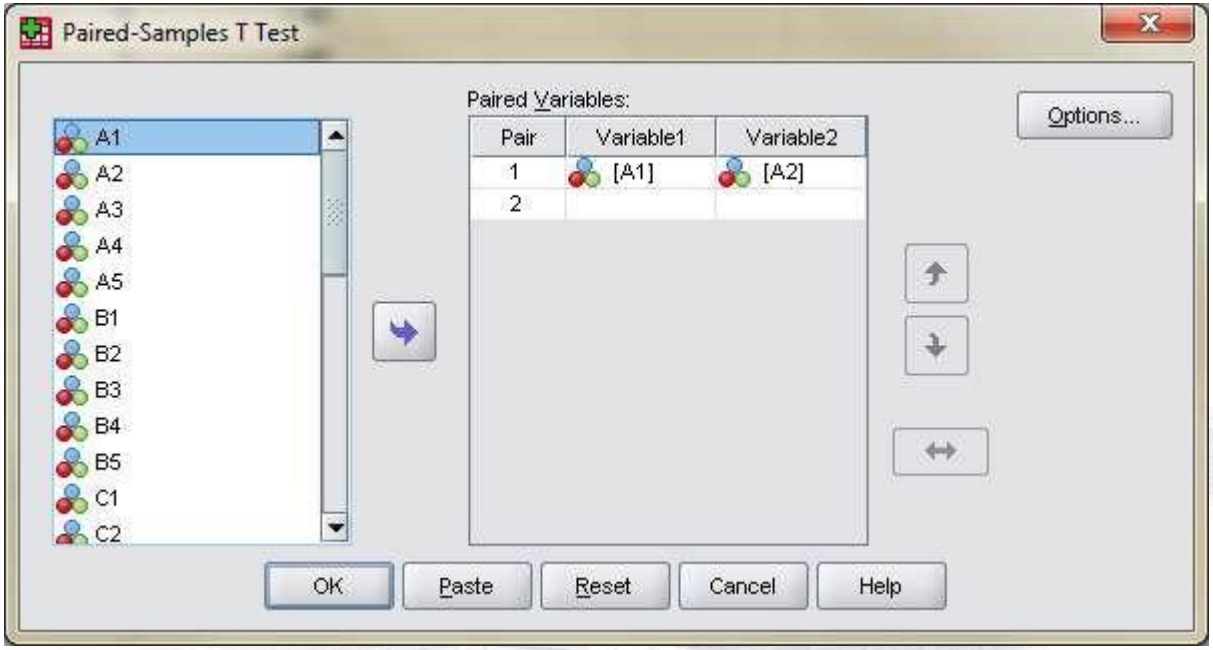
分析两个变量的均值。比较每个样本的 两个变量间的差异，分析其差异相比0 值是否有显著差异

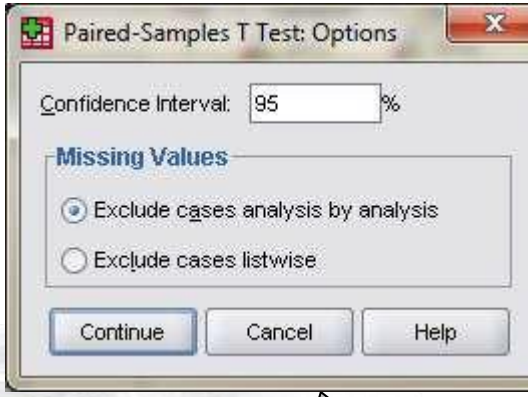
Hypothesis:

•**Null:** 两个变量的均值没有显著差异 There is no significant difference between the means of the two variables.

•**Alternate:**两个变量的均值有显著差 异 There is a significant difference between the means of the two variables.

Paired Samples T Test





Paired Samples T Test

A1的均值更高

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	A1	3.59	290	.978	.057
	A2	3.57	290	.906	.053

两个变量的相关性显著the correlation between the two variables is significant

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	A1 & A2	290	.563	.000

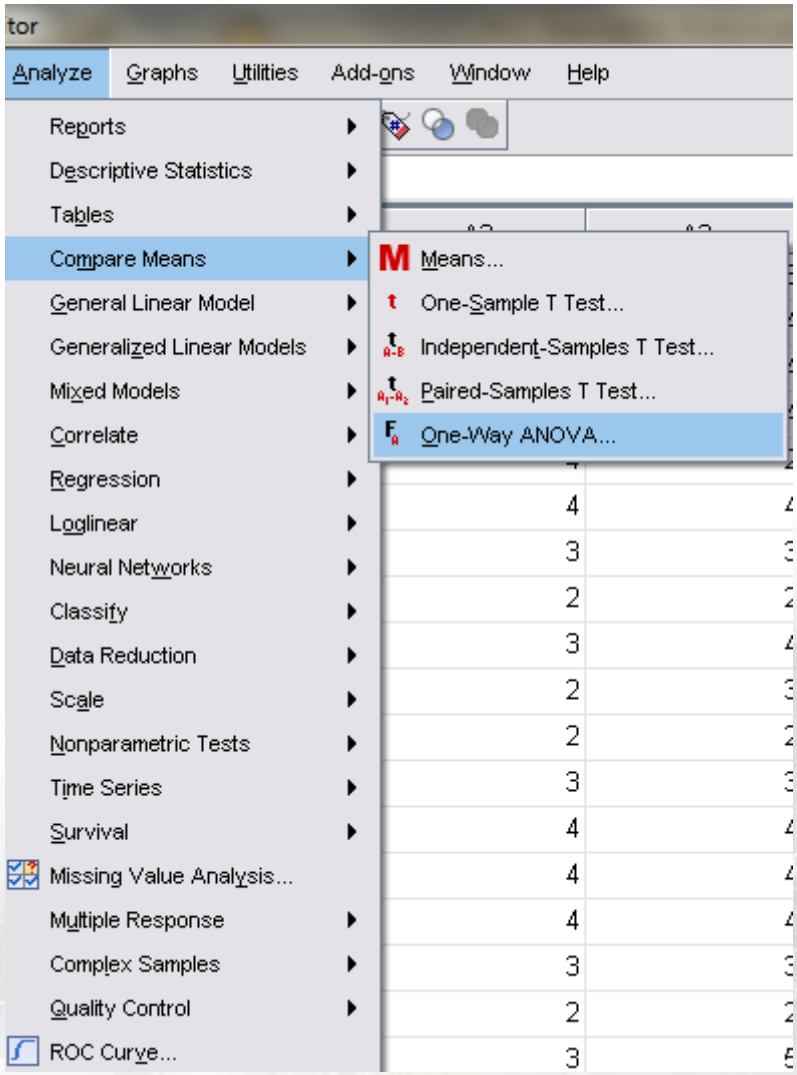
两个变量的差异的描述统计 the descriptive statistics for the difference between the two variables

如果显著值小于0.05，则表示差异显著 If the significance value is less than .05, there is a significant difference.

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	A1 - A2	.017	.882	.052	-.085	.119	.333	289	.740

One-Way ANOVA 单因素方差分析



分析多组自变量分组的均值是否有显著差异

is used to determine whether there are any significant differences between the means of three or more independent (unrelated) groups.

Hypotheses:

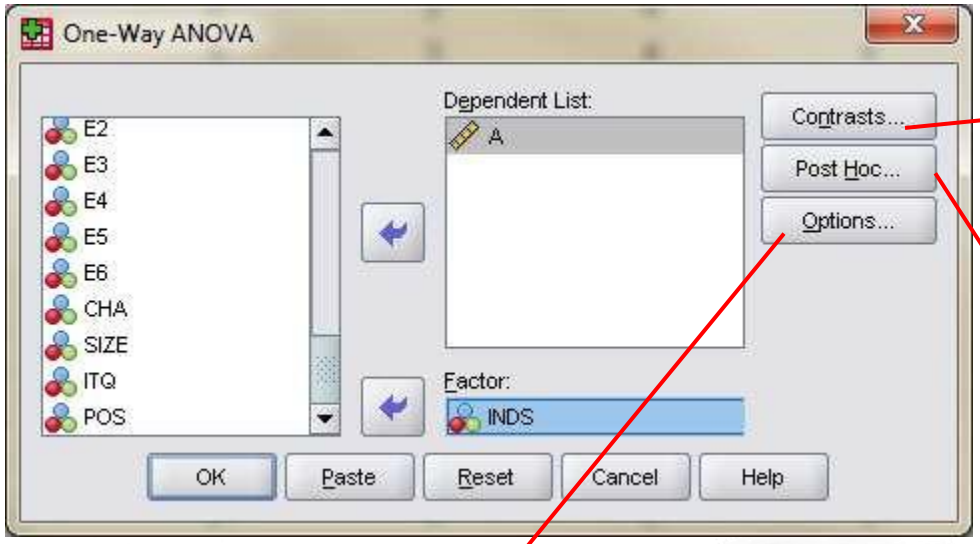
•**Null:** 组内均值没有显著差异

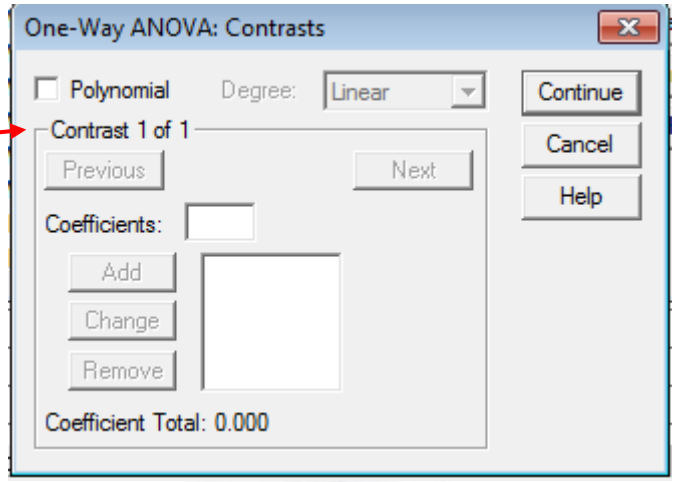
There are no significant differences between the groups' mean scores.

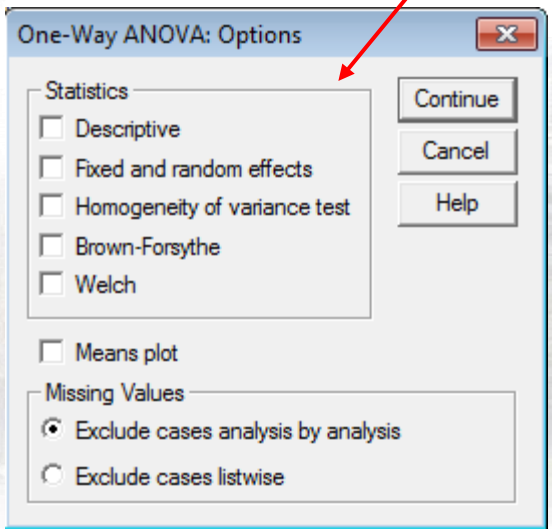
•**Alternate:**组内均值有显著差异

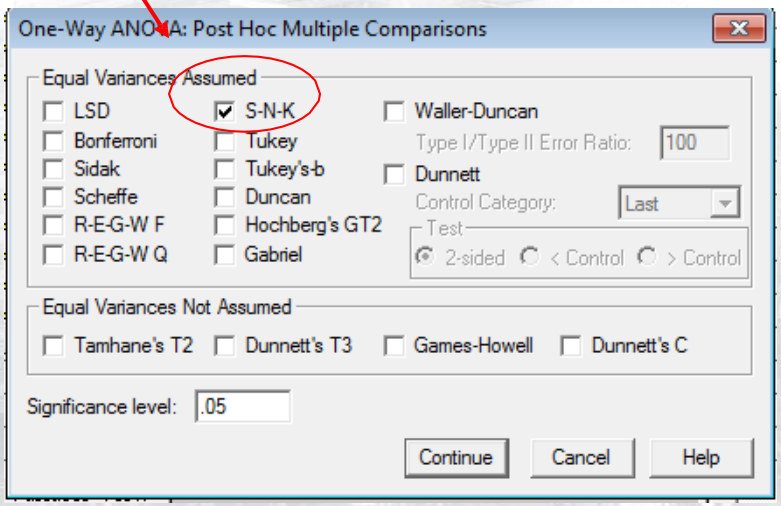
There is a significant difference between the groups' mean scores.

One-Way ANOVA 单因素方差分析









One-Way ANOVA 单因素方差分析

Descriptives

A

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	100	3.4860	.65026	.06503	3.3570	3.6150	2.00	5.00
2	77	3.6805	.72784	.08294	3.5153	3.8457	1.60	5.00
3	78	3.6821	.65222	.07385	3.5350	3.8291	2.20	5.00
4	24	3.4667	.63154	.12891	3.2000	3.7333	2.00	4.60
5	11	3.4364	.66824	.20148	2.9874	3.8853	2.20	4.20
Total	290	3.5869	.67458	.03961	3.5089	3.6649	1.60	5.00

One-Way ANOVA 单因素方差分析

Test of Homogeneity of Variances

A

Levene Statistic	df1	df2	Sig.
.265	4	285	.900

如果 the Levene‘s Test 显著，两个变量的方差差异显著。如不显著，两者方差则相同
is significant, the two variances are significantly different. It is not significant, the two variances are approximately equal.

ANOVA

A

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.995	4	.749	1.661	.159
Within Groups	128.515	285	.451		
Total	131.510	289			

两组间不存在显著差异。因此，表明A在不同所有制类型中不存在显著差异 There is not a significant difference between the two groups. Therefore, we can say that there is a significant difference between the four ownership types on A.

Multiple Comparisons

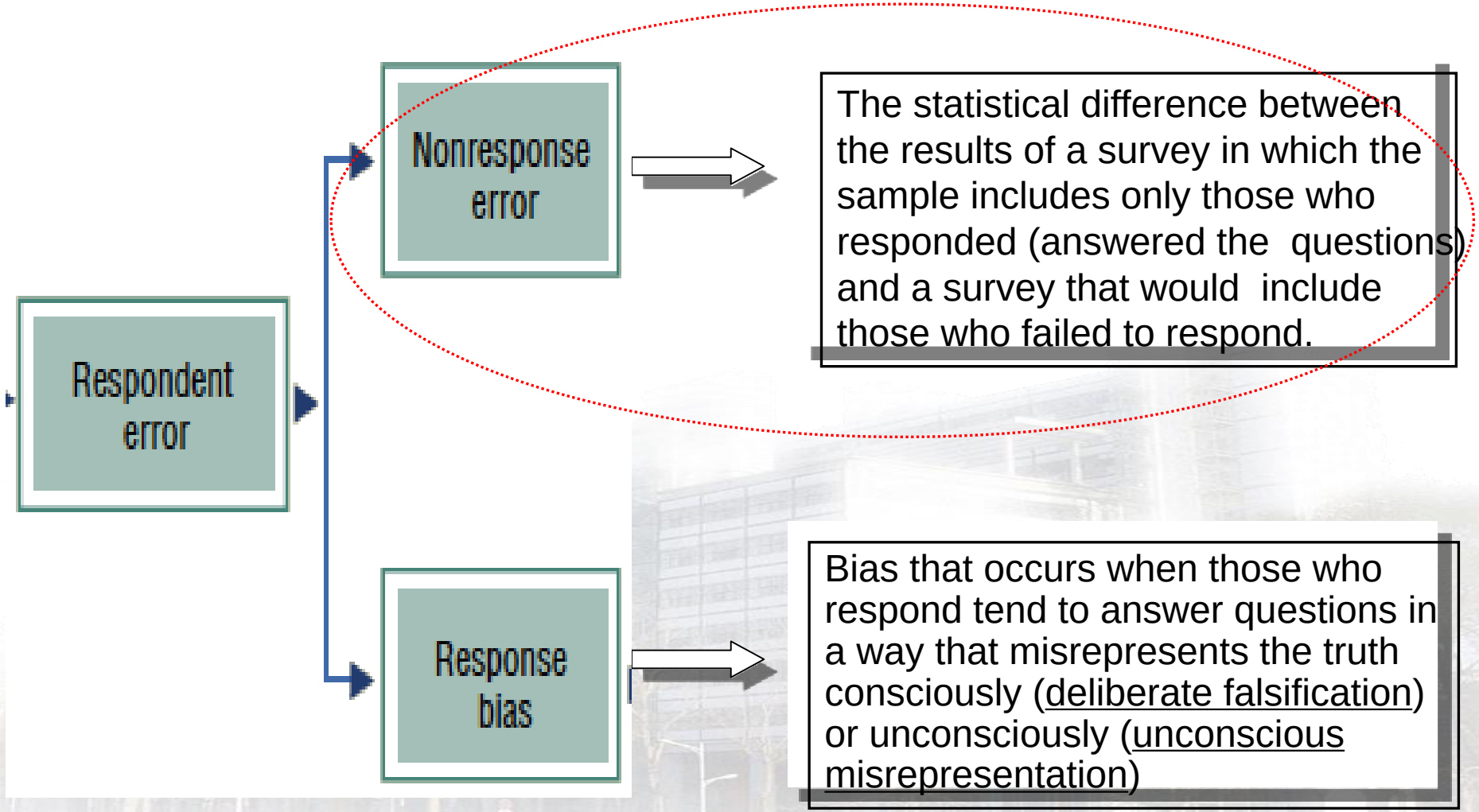
Dependent Variable:A

	(I) CHA	(J) CHA	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	-.19452	.10181	.314	-.4740	.0850
		3	-.19605	.10144	.302	-.4745	.0824
		4	.01933	.15264	1.000	-.3997	.4384
		5	.04964	.21331	.999	-.5360	.6353
	2	1	.19452	.10181	.314	-.0850	.4740
		3	-.00153	.10788	1.000	-.2977	.2946
		4	.21385	.15699	.652	-.2171	.6448
		5	.24416	.21645	.792	-.3501	.8384
	3	1	.19605	.10144	.302	-.0824	.4745
		2	.00153	.10788	1.000	-.2946	.2977
		4	.21538	.15675	.645	-.2149	.6457
		5	.24569	.21627	.787	-.3481	.8394
	4	1	-.01933	.15264	1.000	-.4384	.3997
		2	-.21385	.15699	.652	-.6448	.2171
		3	-.21538	.15675	.645	-.6457	.2149
		5	.03030	.24450	1.000	-.6409	.7015
	5	1	-.04964	.21331	.999	-.6353	.5360
		2	-.24416	.21645	.792	-.8384	.3501
		3	-.24569	.21627	.787	-.8394	.3481
		4	-.03030	.24450	1.000	-.7015	.6409
Games-Howell	1	2	-.19452	.10540	.351	-.4855	.0964
		3	-.19605	.09840	.274	-.4674	.0753
		4	.01933	.14438	1.000	-.3954	.4340
		5	.04964	.21172	.999	-.6236	.7229
	2	1	.19452	.10540	.351	-.0964	.4855
		3	-.00153	.11106	1.000	-.3082	.3051
		4	.21385	.15329	.634	-.2222	.6499
		5	.24416	.21789	.793	-.4373	.9256
	3	1	.19605	.09840	.274	-.0753	.4674
		2	.00153	.11106	1.000	-.3051	.3082
		4	.21538	.14857	.600	-.2093	.6401
		5	.24569	.21459	.781	-.4312	.9226
	4	1	-.01933	.14438	1.000	-.4340	.3954
		2	-.21385	.15329	.634	-.6499	.2222
		3	-.21538	.14857	.600	-.6401	.2093
		5	.03030	.23919	1.000	-.6909	.7515
	5	1	-.04964	.21172	.999	-.7229	.6236
		2	-.24416	.21789	.793	-.9256	.4373
		3	-.24569	.21459	.781	-.9226	.4312
		4	-.03030	.23919	1.000	-.7515	.6909

可分析每组样本 的均值的显著差异明细
of groups are significantly differen.

对有显著差异的 值会通过*显示
SPSS notes a significant difference with an asterisk (*).

The role of "compare mean"



ANOVA_ Non-response bias_ Example

3.4. Non-response bias

This study employs Analysis of variance (ANOVA) to check for differences in annual sales and number of employees between early and late respondents to measure non-response bias and ensure that the sample firms were representative of the population (Armstrong and Overton, 1977). Responses return within four weeks of the first mailing were classified as early ($n = 182$), while those received after four weeks were classified as late ($n = 71$) (cf. Mishra et al., 1998). The ANOVA indicates no significance difference between these two groups in annual turnover ($p = 0.91$) and number of employees ($p = 0.89$).

Wu, L.-Y., 2010. Applicability of the resource-based and dynamic-capability views under environmental volatility. Journal of Business Research 63, 27-31


Q&A


THANK YOU