



# **Spss : Statistical Package For Social Science section2**

**maryam deldar froutagheh**

# روند Explore

یکی از امکانات مفید برای توصیف اطلاعات و آنالیز اکتشافی داده‌ها، روند Explore است که دارای امکانات فراوانی است.

Explore

در روند Explore می‌توانید برای خلاصه کردن و توصیف مشاهدات از نمودارهای مناسب بهره‌گیری نمایید.

جدول‌های خلاصه شده‌ای از اطلاعات مربوط به نمونه‌ها و زیر گروه‌هایی از نمونه‌ها را بدست آورید.

در میان اطلاعات، نمونه‌های غیرعادی را که مقادیر آنها از سایر نمونه‌ها بیشتر یا کمتر هستند، کاوش کنید.

شاخص‌های آماری مناسب را در گروه‌ها محاسبه کرده و فاصله اطمینان برای میانگین به دست آورید .

می‌توانید نمودار ساقه و برگ و هیستوگرام را برای نمایش توزیع فراوانی در بین گروه‌ها رسم کنید.

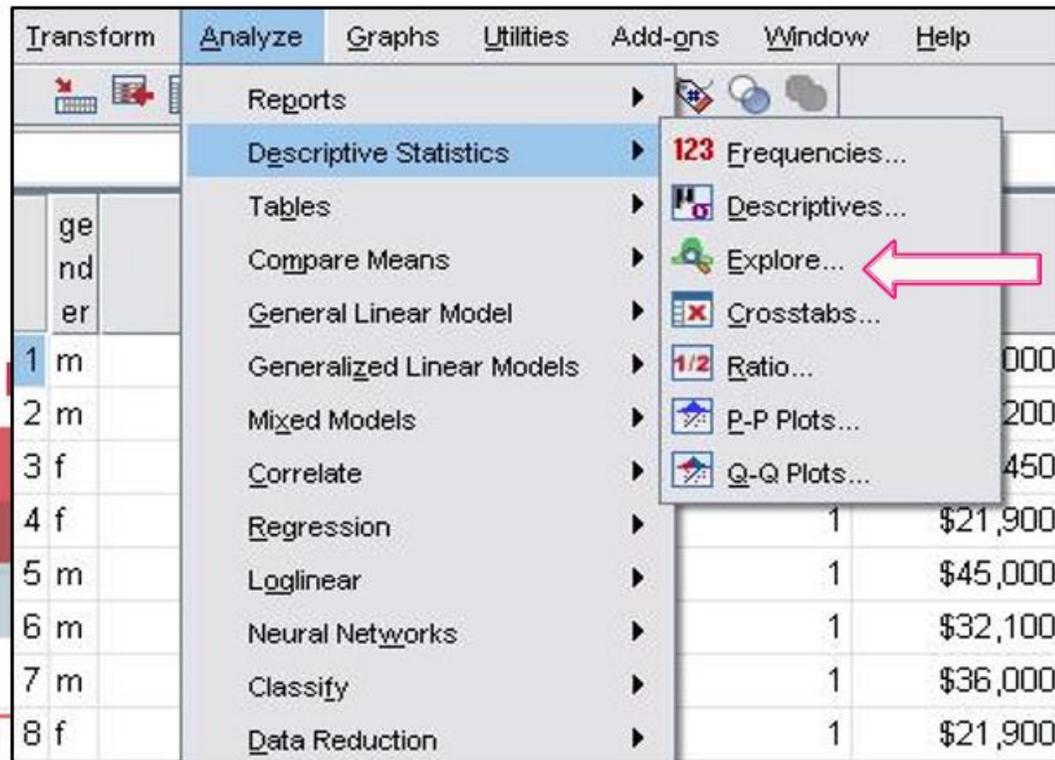
نمودار مفید جعبه‌ای را برای مقایسه شاخص‌های میانه، چارک اول و سوم بین گروه‌ها بکار گیرید.

می‌توانید از نمودار Q-Q برای بررسی نرمال بودن مقادیر متغیر وابسته در هر گروه استفاده کنید.

> SPSS

# روند Explore

اگر می خواهید در بین داده ها، اطلاعات بیشتری از نمونه ها را جستجو کنید و موشکافانه تر به جزئیات داده ها نگاه کنید. دستور Explore را به عنوان یک دستور جامع آنالیز اکتشافی داده ها به کار بگیرید.



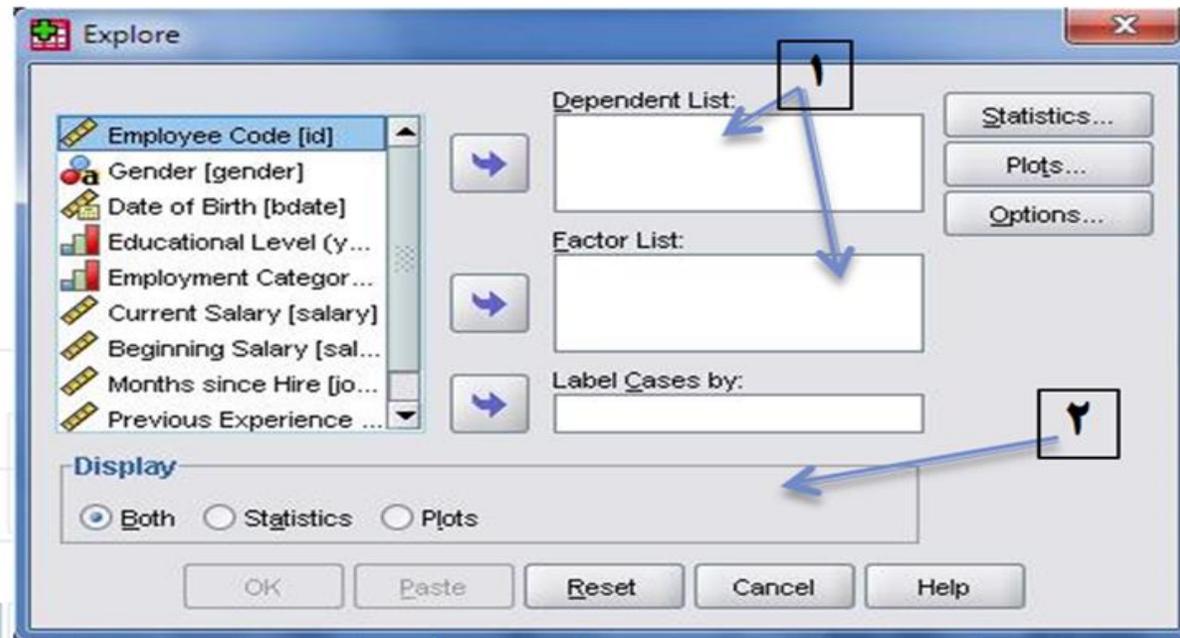
The screenshot shows the SPSS software interface. The 'Analyze' menu is open, and the 'Explore...' option is highlighted with a pink arrow. The background shows a data table with columns for 'gender' and numerical values.

gender	
1	m
2	m
3	f
4	f
5	m
6	m
7	m
8	f

- در روند Explore مقادیر یک صفت کمی را در سطوح یک متغیر کیفی و با استفاده از نمودار یا شاخص های توصیفی مقایسه کرد.
- همچنین می توان از این روند برای مقایسه گروهها نیز استفاده کرد.
- در این روند متغیر کمی (Scale) را به عنوان متغیر وابسته و متغیر کیفی (Nominal , Ordinal) را به عنوان متغیر مستقل در نظر می گیریم.
- می توانید از نمودار Q-Q برای آزمون نرمال بودن متغیر وابسته در هر گروه استفاده کنید.
- از منو اصلی گزینه Analyzed, گزینه descriptive statistics و سپس explore را انتخاب کنید.

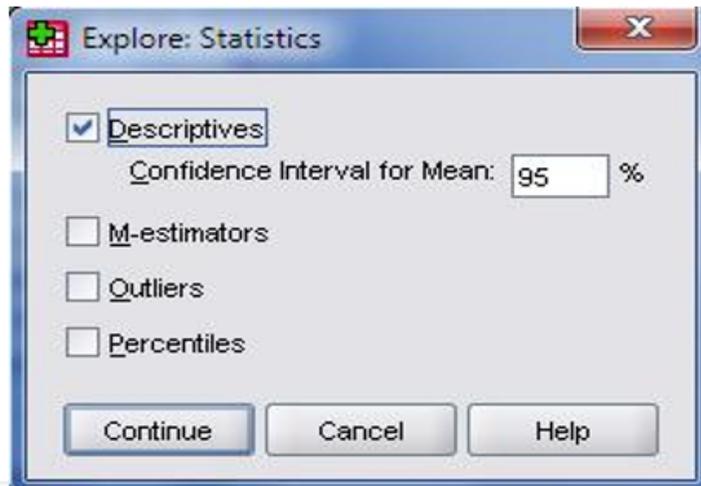
# روند Explore

- ۱- به پنجره dependent list یک متغیر کمی مانند سن / تعداد فرزند و ... و به پنجره factor list یک متغیر کیفی مثل جنس / گروه خون و... منتقل کنید.
- ۲- اگر می‌خواهید فقط شاخص‌ها را محاسبه کنید، در پایین و سمت چپ پنجره، گزینه Statistics و اگر می‌خواهید فقط نمودار داده‌ها را مشاهده کنید گزینه plots و اگر می‌خواهید از هر دو مورد استفاده کنید گزینه Both را علامت‌دار کنید.



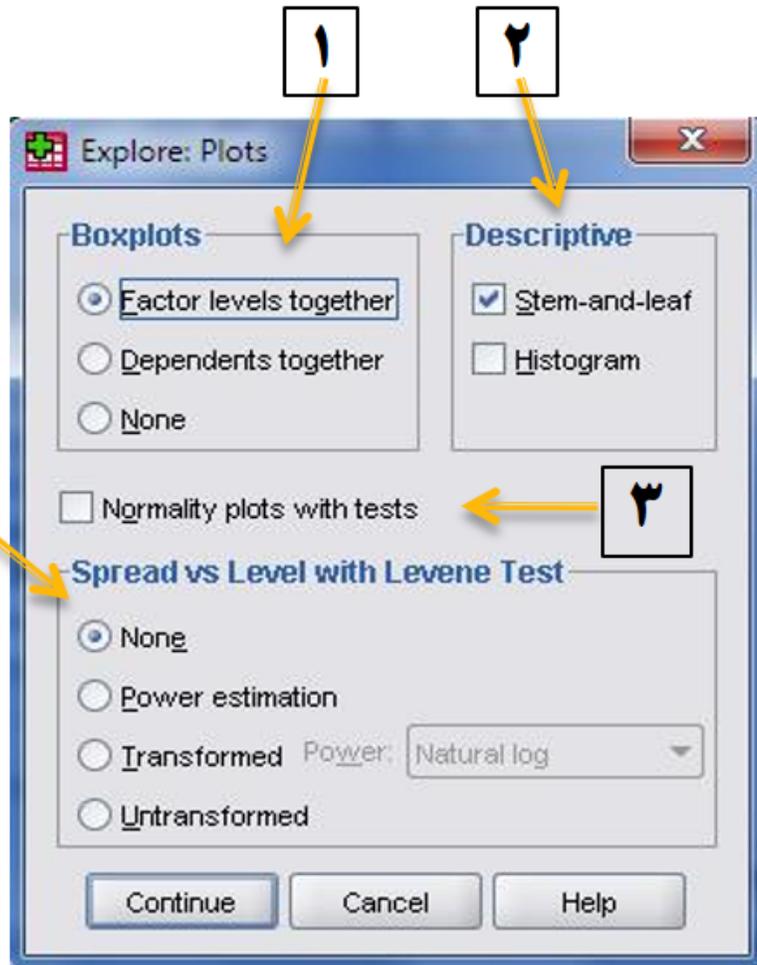
# روند Explore

- روی گزینه statistics کلیک کنید تا کادر محاوره آن (explore: statistics) باز شود. در این کادر محاوره می‌توانید هر یک از گزینه‌های زیر را انتخاب کنید:
- بطور پیش فرض گزینه descriptive علامت دار شده است که منجر به محاسبه شاخص‌های توصیفی و همچنین یک فاصله اطمینان ۹۵٪ برای میانگین خواهد شد.
- گزینه M-estimator، منجر به محاسبه آماره‌ای می‌شود که به هر داده بسته به فاصله آن از میانگین، وزن می‌دهد.
- گزینه Out line پنج مورد از بزرگترین و کوچکترین مقادیر متغیر وابسته را نمایش می‌دهد.
- گزینه percentiles صدکهای ضروری ۵ و ۱۰ و ۲۵ و ۵۰ و ۷۵ و ۹۰ و ۹۵ را محاسبه می‌کند.



توضیح اینکه فاصله‌ی اطمینان مذکور، فاصله‌ای است که با اطمینان ۹۵ درصد، میانگین جامعه را در بر می‌گیرد.

# روند Explore



● اگر گزینه plots را انتخاب کنید پنجره مربوط به explore: plots (مانند شکل) باز شده و در آن می‌توانید نمودارهای زیر را رسم کنید.

- ۱- نمودار جعبه ای با گزینه های :  
الف- سطوح متغیر فاکتور با هم  
ب- سطوح متغیر وابسته با هم  
ج- هیچکدام.

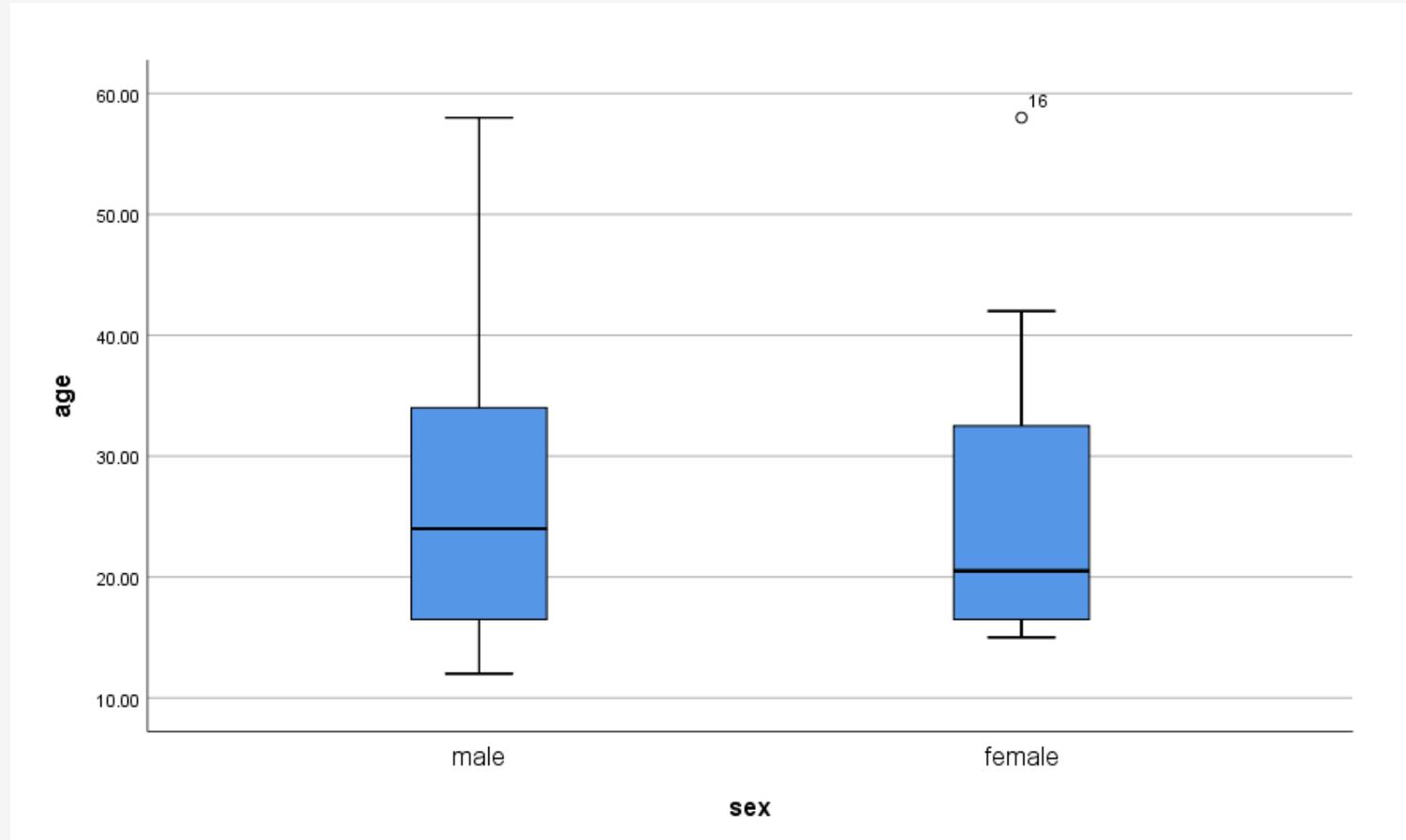
- ۲- نمودار های توصیفی با دو گزینه:  
الف- رسم نمودار ساقه و برگ  
ب- رسم هیستوگرام فراوانی

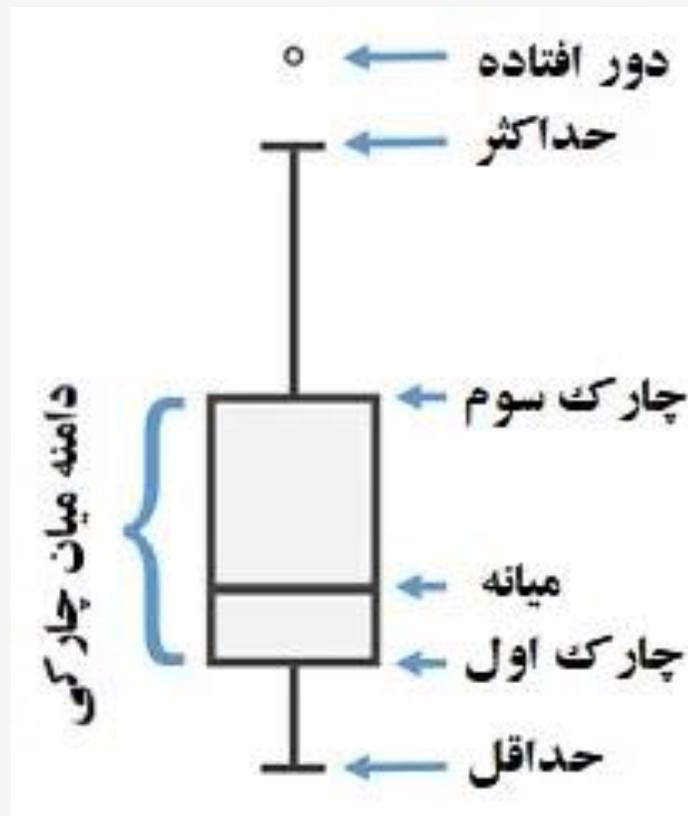
۳- اگر گزینه Normality plots with tests را علامتدار کنید. آزمون نرمال بودن توزیع را می‌توانید با استفاده از نمودار Q-Q انجام دهید.

۴- گزینه Spread vs. level with levene test مربوط به آزمون لون برای یکسان بودن واریانس ها در سطوح هر یک از متغیرهای فاکتور است.

# Box plot

---

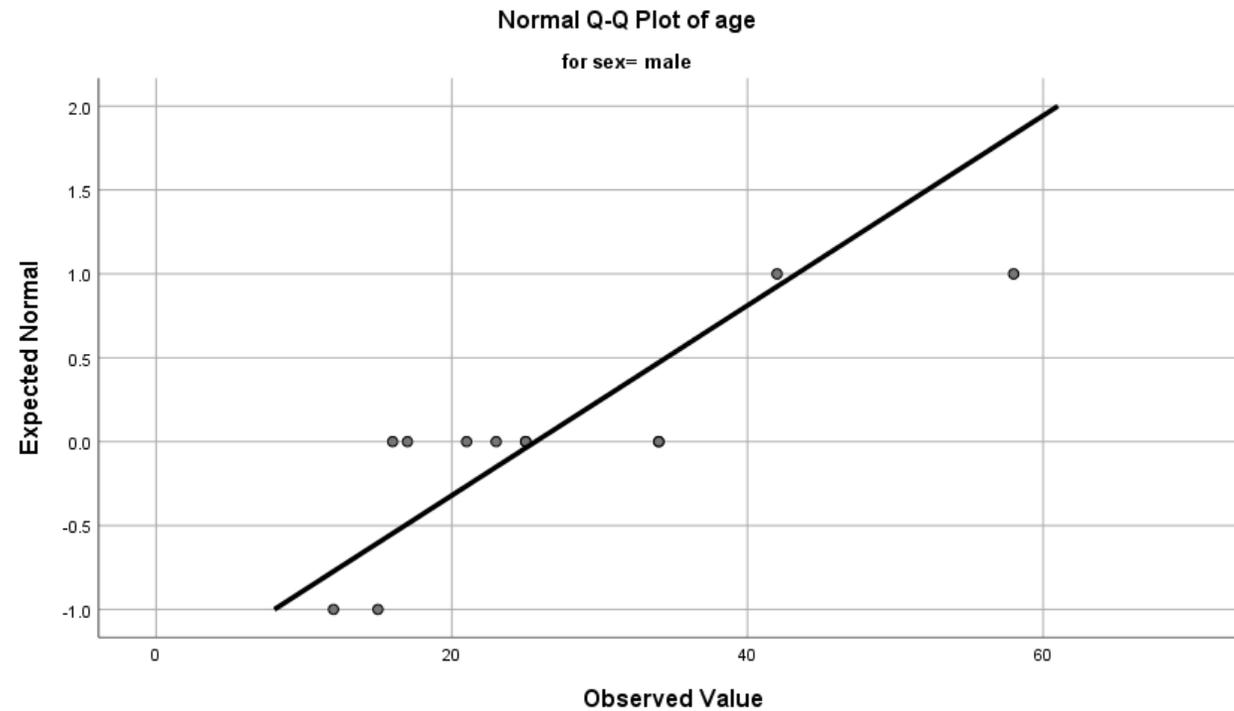




نمودار جعبه‌ای یک روش استاندارد برای نمایش توزیع داده‌ها است که براساس شاخص‌های آماری کوچکترین مقدار، چارک اول، میانه، چارک سوم و بزرگترین مقدار ساخته شده است. همچنین این نمودار می‌تواند در مورد وجود داده‌های دورافتاده یا پرت، اطلاعاتی به شما بدهد و مقدار آن‌ها را تعیین کند. همچنین نشان دادن تقارن در داده‌ها از کارهایی این نمودار است.

# Q-Q Plots

## Normal Q-Q Plots



نمودار چندک چندک  
یا **Q-Q plot** به  
منظور مقایسه دو  
توزیع به کار گرفته  
می شود. از چنین  
نمودارهایی حتی  
می توان مطابقت  
توزیع داده ها را با  
یک توزیع مشخص،  
مورد بررسی قرار داد.

# روند Explore

## تمرین

از مجموعه داده های spss فایل داده ی استخدای (Employee) را معرفی می کنیم. در این فایل اطلاعاتی ۴۷۴ نمونه از بین شاغلین انتخاب شده و از هر نمونه ۱۰ متغیر اندازه گیری شده است. تعدادی از متغیرهای این فایل داده در جدول زیر آورده شده است.

ماهیت	توضیح	نام متغیر
Nominal	جنسیت (مرد - زن)	gender
Nominal	تحصیلات (برحسب سال)	educ
Scale	گروه شغلی (اداری - نگهبانی - مدیریت)	jobcat
Scale	حقوق فعلی	salary
Scale	سابقه کار فعلی	jobtime

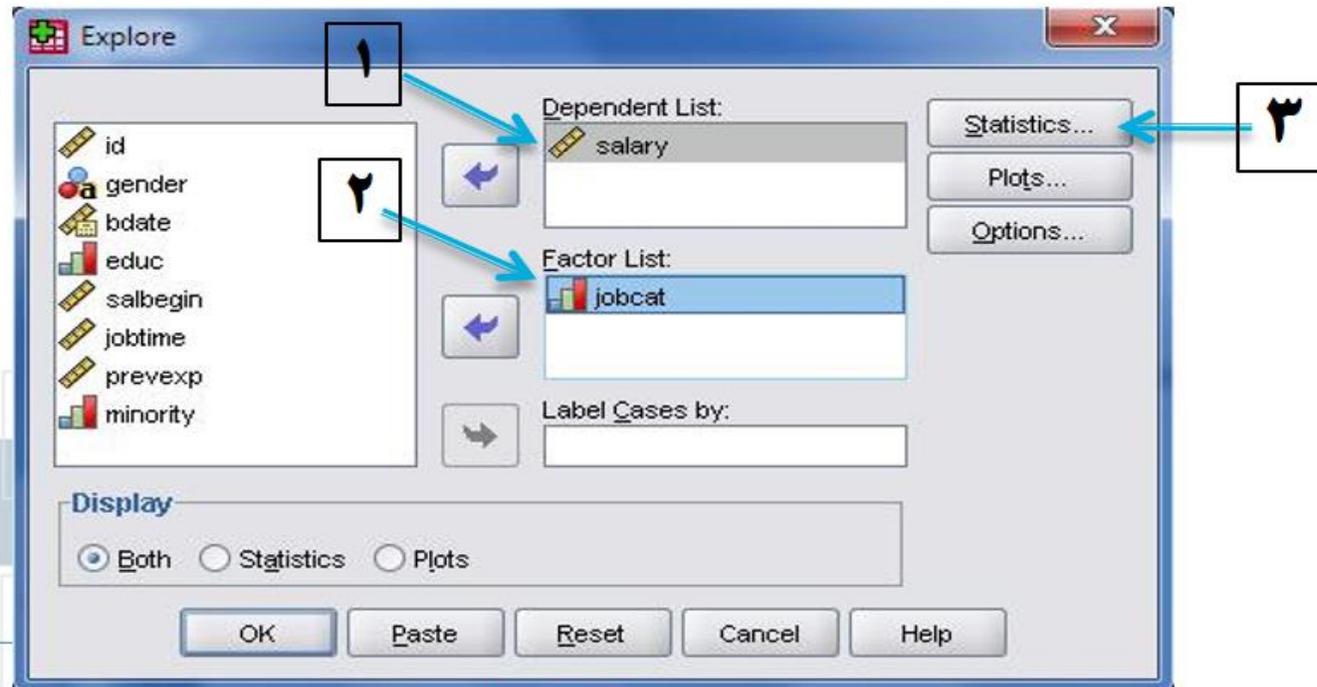
فایل داده های Employee را از مسیر زیر باز کنید و برای مقایسه حقوق شاغلین (Salary) در هر گروه شغلی (Jabcat)، شاخص های توصیفی را بدست آورده نمودار مناسب رسم کنید.

Program file/ SPSSInc/ SPSS17/ Samples/ Employee.sav

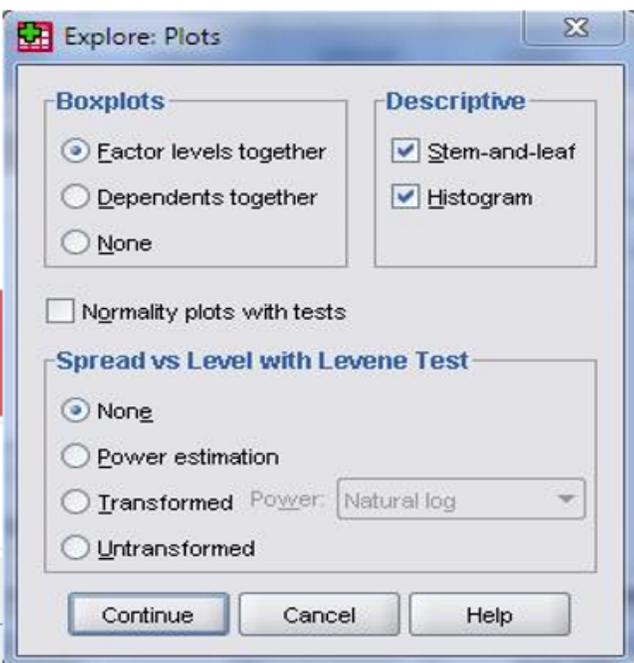
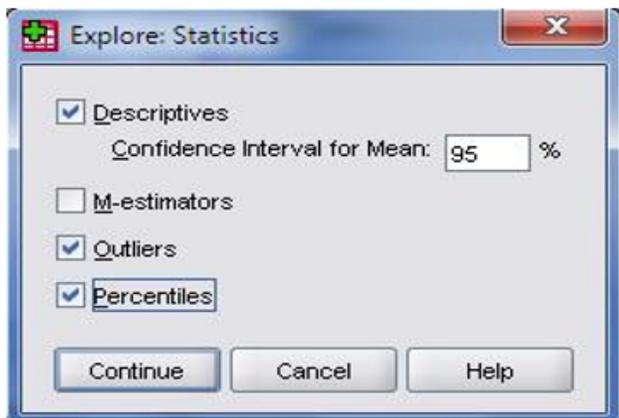
> SPSS

# روند Explore

- ۱- در کادر محاوره Explore متغیر حقوق (Salary) را به کادر متغیرهای وابسته (Dependent list) منتقل کنید.
- ۲- متغیر طبقه شغلی (jobcat) را به کادر متغیرهای مستقل (Independent list) منتقل کنید.
- ۳- از کلید Statistics به کادر محاوره آن منتقل شوید.



# روند Explore



۱- در این کادر محاوره گزینه Descriptive از قبل علامتدار شده تا بعضی از آماره های توصیفی مهم به همراه یک فاصله اطمینان ۹۵٪ برای میانگین حقوق کارمندان در هر گروه شغلی محاسبه شود.

۲- گزینه Outliers را انتخاب کنید تا بیشترین و کمترین حقوق بگیران در هر طبقه شغلی مشخص شوند.

۳- گزینه Percentiles را انتخاب کنید تا چندکهای حقوق کارمندان در هر گروه معلوم شود.

کلید continue را کلیک کنید تا به کادر محاوره اصلی برگردید...

برای رسم نمودار دلخواه، گزینه Plots... را کلیک کنید تا به کادر محاوره Explore: Plots وارد شوید

دو گزینه Stem-and-Leaf و Histogram را به منظور رسم نمودار ساقه و برگ و هیستوگرام انتخاب کنید.

کلیدهای continue و ok را کلیک کنید و نتایج را در خروجی مشاهده نمایید.

**مشاهده نتایج**

# Count values

Statistics Data Editor

Transform Analyze Graphs Utilities Extensions Window Help

Count Values within Cases...

Count Occurrences of Values within Cases

Target Variable: ny

Target Label:

Numeric Variables: sex, cens

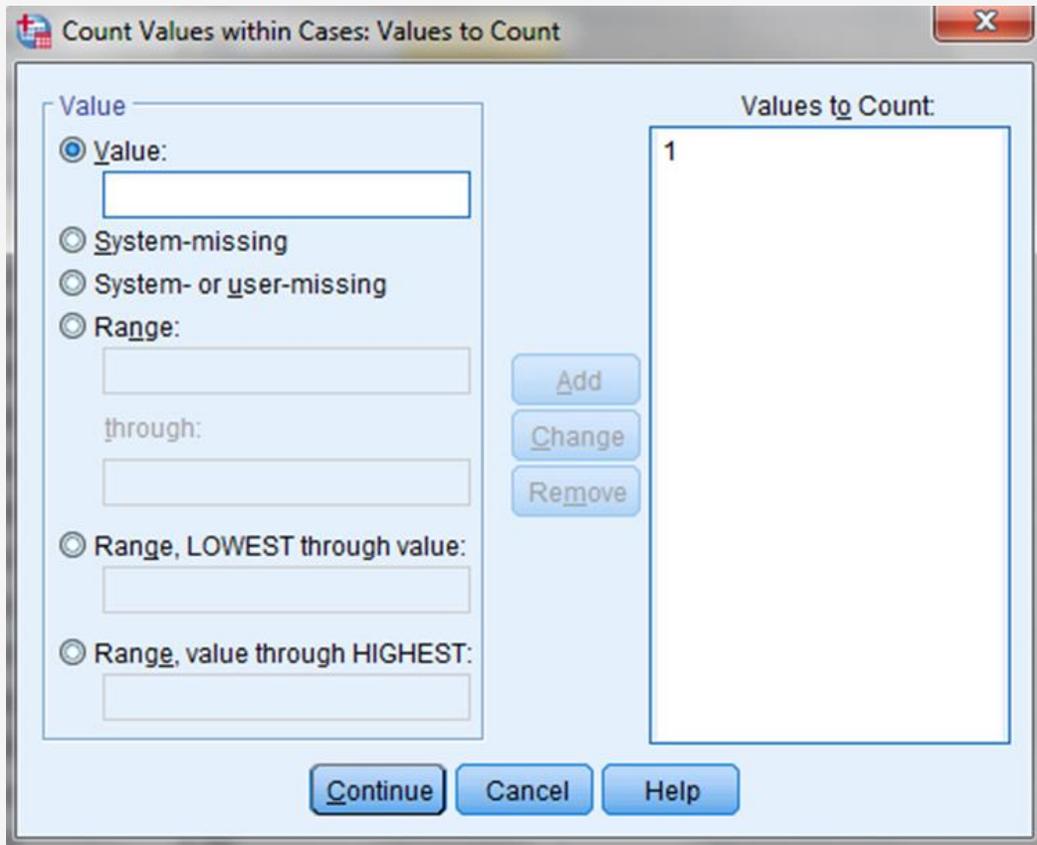
Define Values...

If... (optional case selection condition)

OK Paste Reset Cancel Help

گزینه ی شمارش مقادیر عددی در موردها، شباهت های بسیاری با گزینه های دیگر در آمار توصیفی از قبیل **Frequency** یا **Recode** در سربرگ **Transform** دارد.

برای مثال، به کمک این گزینه می توانیم تعداد نمره های خاصی را که مد نظر داریم تحت عنوان یک متغیر جدید نشان دهیم.



**Statistics**

ny

N	Valid	1048
	Missing	0

**ny**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	961	91.7	91.7	91.7
	1.00	72	6.9	6.9	98.6
	2.00	15	1.4	1.4	100.0
Total		1048	100.0	100.0	

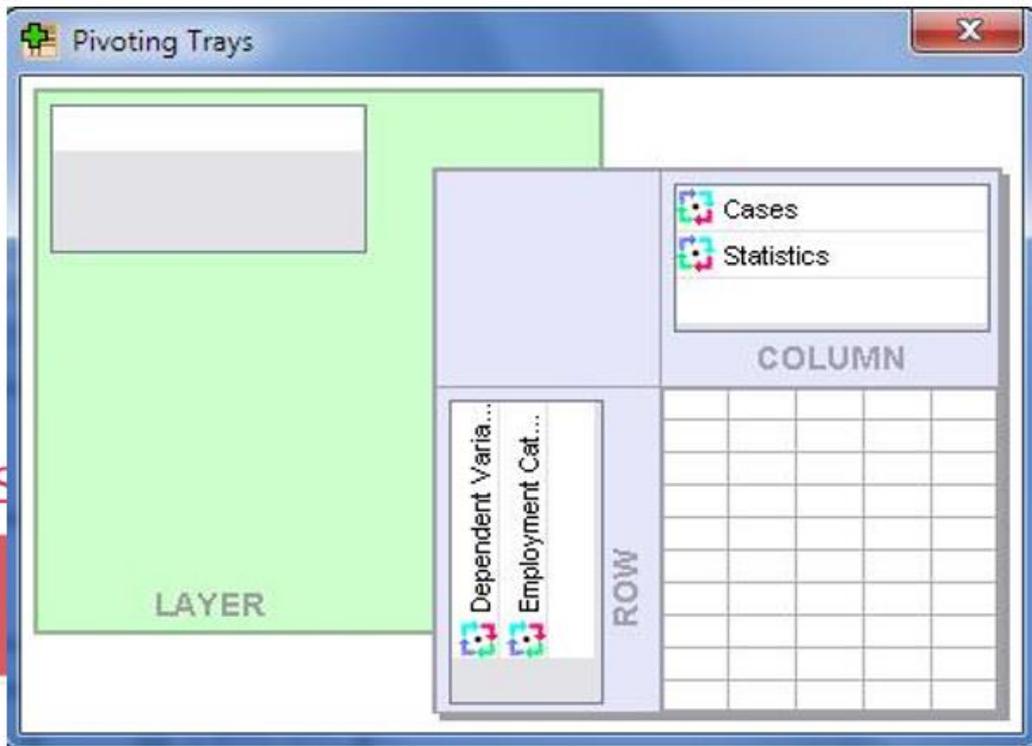
# جدول گردان (pivot table)

آیکون جدول گردان



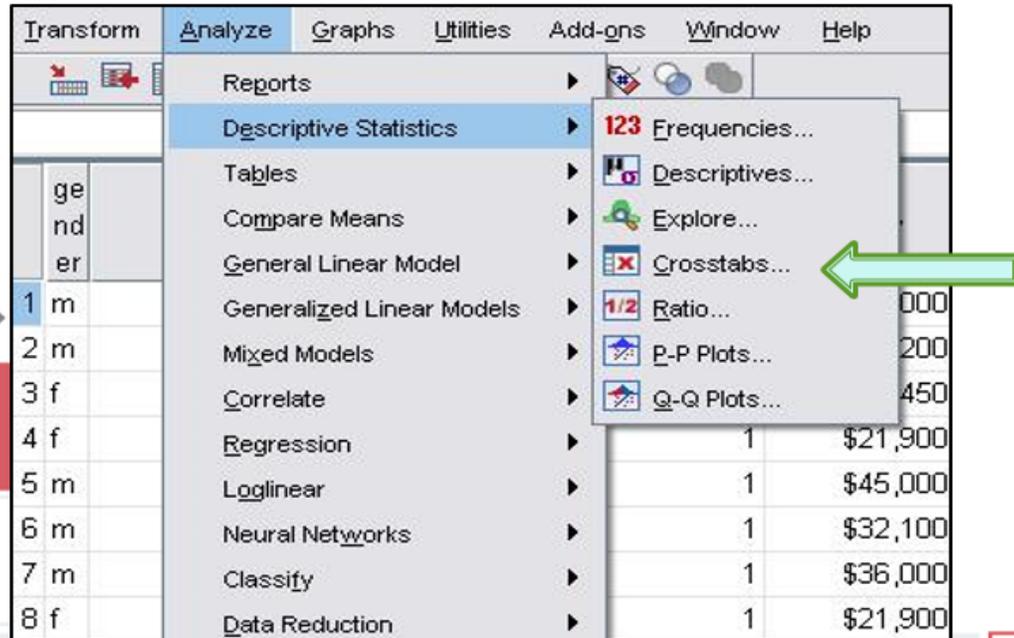
● اگر یک جدول دو طرفه در هر یک از روند های Explore یا cross tab برای مقایسه سطوح متغیرها در خروجی spss ایجاد کرده‌اید، می‌توانید به سادگی جای سطرها و ستون‌ها را عوض کنید و معانی مختلفی از جدول استنباط نمایید. برای این کار روی جدول دو بار کلیک نمایید تا ویراشگر جدول (شکل بالا) را مشاهده کنید.

● اینک روی آیکون جدول گردان کلیک کنید تا کادر محاوره pivoting Trays باز شود. در مکان سطر و ستون این پنجره یک یا چند مربع کوچک به رنگ سبز و قرمز مشاهده می‌کنید که اگر آنها را جا به جا کنید، جدول خروجی تغییر خواهند کرد. با مشاهده آنها می‌توانید برای سهولت استفاده از اطلاعات جدول، بهترین حالت را به دلخواه انتخاب نمایید.



# روند Crosse tab

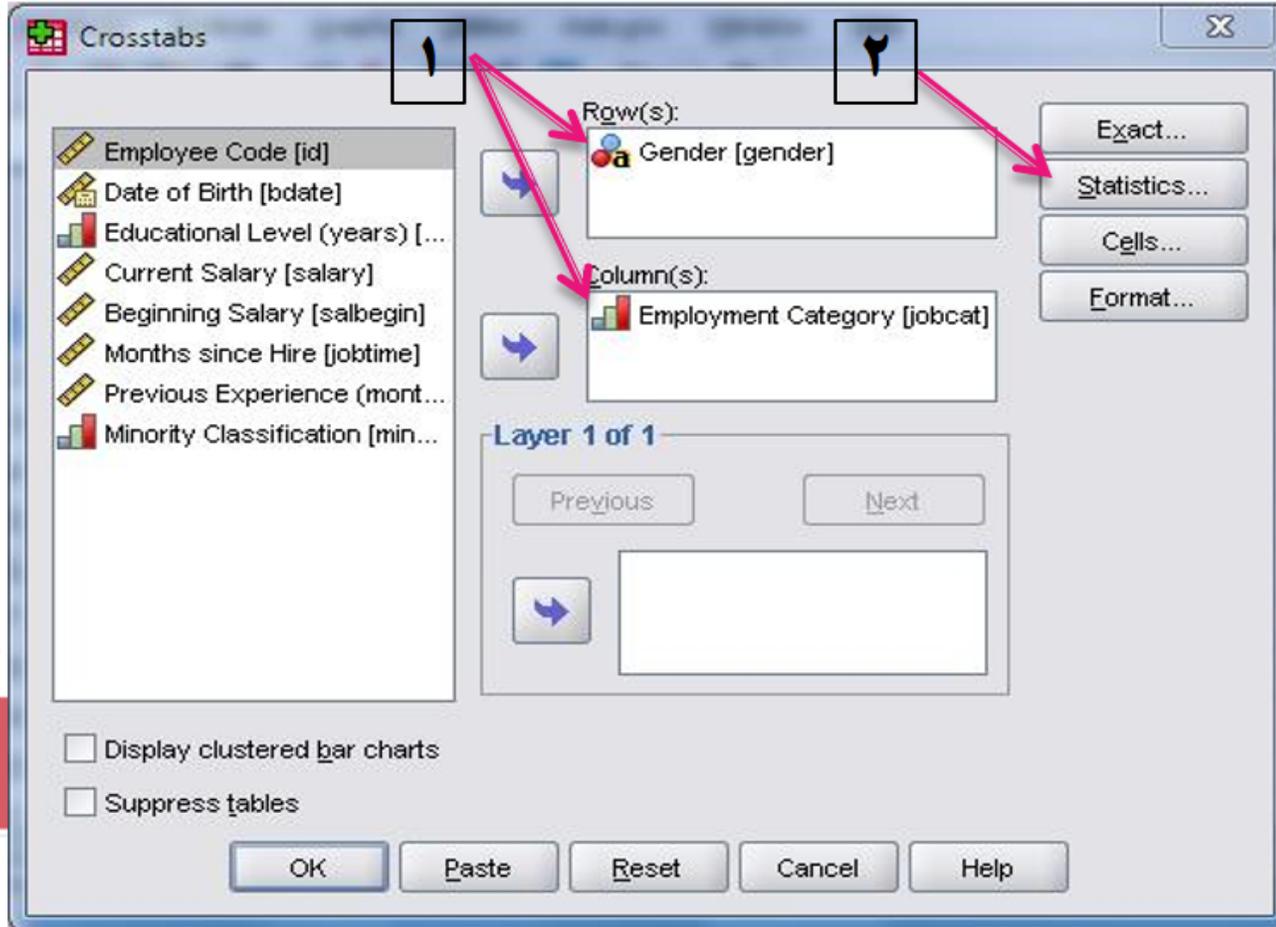
- جدولی را که در آن داده های آماری در  $k$  سطر و  $L$  ستون قرار دارند، یک جدول توافقی است.
- در این جدول می توان تعداد  $k \times l$  ترکیب مختلف از مقادیر را مشاهده کرد.
- در چنین جدولی می توان تأثیر یک متغیر کیفی (اسمی یا ترتیبی) را بر روی یک متغیر کیفی دیگر بررسی کرد.



● در spss، برای ساختن یک جدول فراوانی دو طرفه که به جدول توافقی معروف است، راه های زیادی وجود دارد. یکی از این راهها استفاده از دستور Cross Tab است.

● می توانید از منوی اصلی گزینه Analyze و سپس Descriptive Statistics را انتخاب کرده و سپس از دستور Crosse Tab به کادر محاوره آن وارد شوید.

# روند Crosstab



- ۱- در کادر محاوره Cross tab یک متغیر کیفی به کادر مربوط به سطر Row(s) و یک متغیر کیفی دیگر را به کادر مربوط به ستون Column(s) منتقل کنید. توجه داشته باشید که متغیرهای انتخابی شما باید از نوع رتبه‌ای باشند و یا اگر متغیر کمی پیوسته هستند آنها را به دسته‌های کوچکتر طبقه بندی کنید.
- ۲- اگر می‌خواهید ضرایب همبستگی و مقادیر کای اسکور را محاسبه کنید، گزینه‌های مربوط به آنها را در پنجره Statistics انتخاب کنید.

# روند Crosse tab

## تمرین

برای انجام این تمرین، داده های ۱۹۹۱ U.S. General Social Survey را از پوشه Sample به آدرس زیر باز کنید.

Program file/ SPSSInc/ SPSS16/ Samples

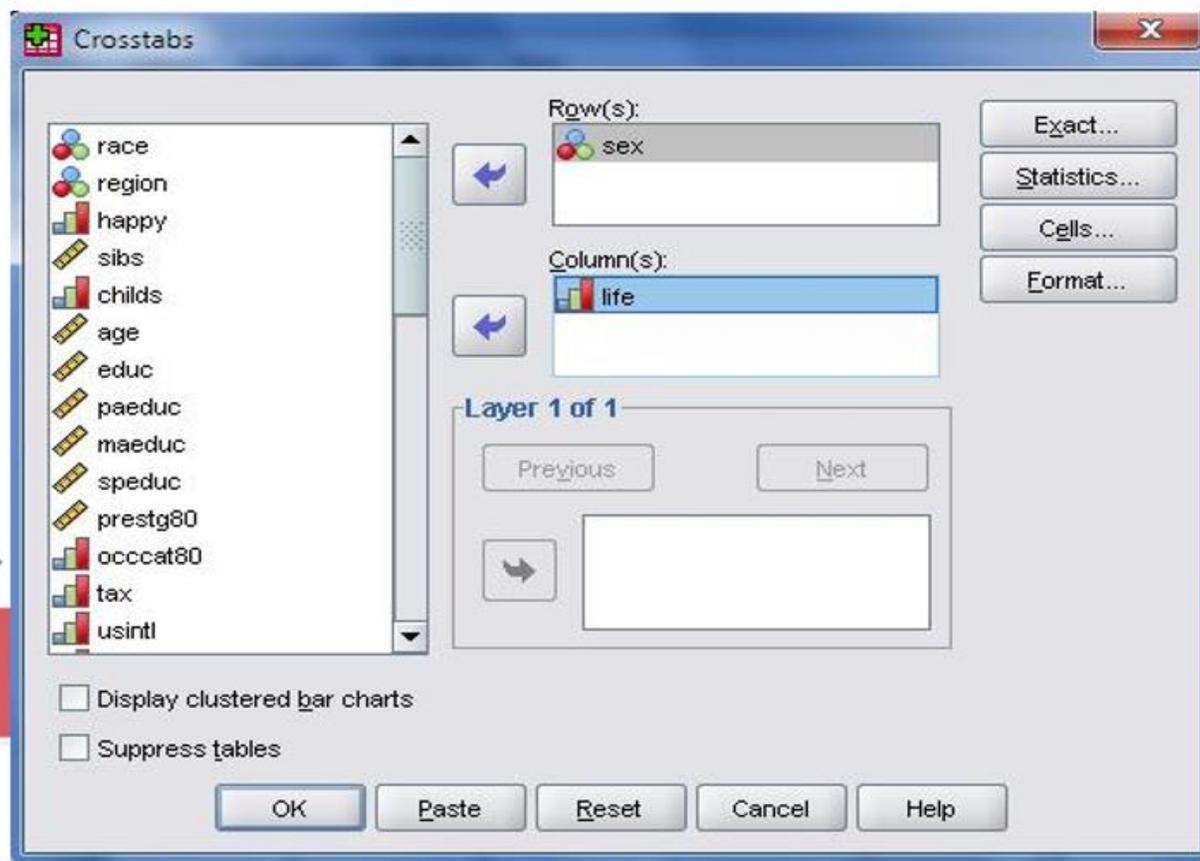
در این فایل، متغیرهای متفاوتی وجود دارند که تعدادی از آنها را در زیر معرفی می کنیم. از این به بعد این فایل داده را با نام مختصر ۱۹۹۱ بکار می بریم.

نام متغیر	توضیح	نوع
Sex	جنسیت (مرد - زن)	Nominal
Race	نژاد (سیاه - سفید - سایر نژادها)	Nominal
Region	منطقه محل سکونت (شمال شرقی - شمال غربی - غرب)	Nominal
Life	زندگی را چگونه می بینید (هیجان انگیز - معمولی - خسته کننده)	Ordinal
Sibs	تعداد خواهر و برادر	Scale
Childs	تعداد فرزند	Scale
Age	سن	Scale
Educe	تحصیلات (برحسب سال)	Scale

> SPSS

این داده ها مربوط به یک مطالعه اجتماعی است که بر روی ۱۵۱۷ نفر از جمعیت ۱۸ سال به بالا در سال ۱۹۹۱ در ایالات متحده امریکا انجام شده است.

# روند Crosstab



فایل داده های 1991 را باز کنید و روند Crosstab را از فرمان زیر فراخوانی کنید:

Analyze/Descriptive Statistics/Cross Tab

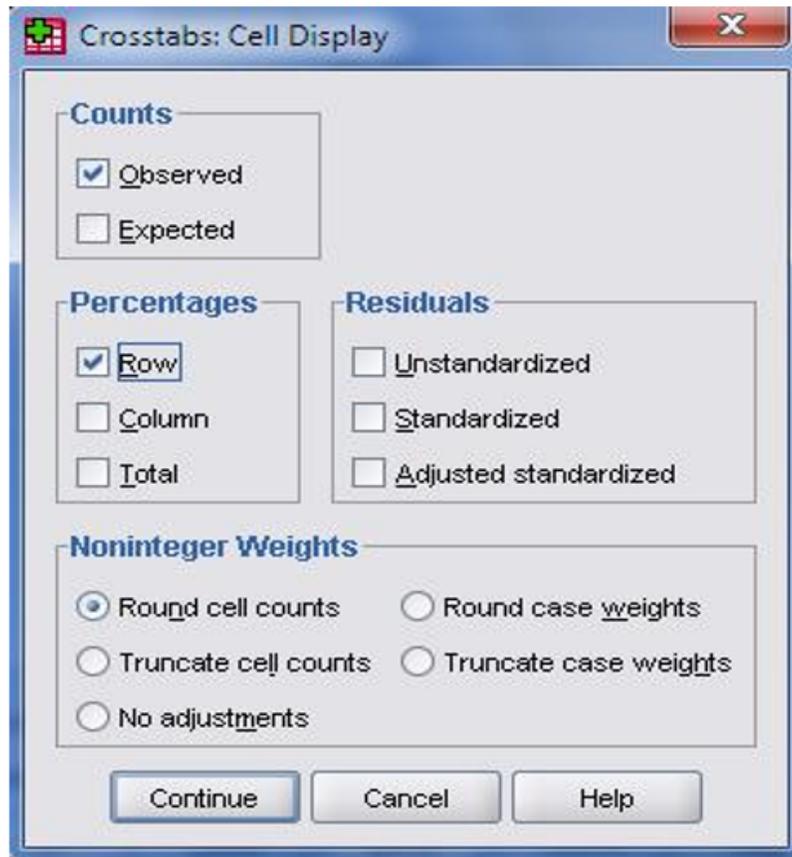
سپس مراحل زیر را دنبال کنید :

۱- در کادر محاوره Crosstab متغیر جنسیت را به کادر سطرها Row(s) و متغیر Life را به کادر مربوط به ستون‌ها column(s)، وارد کنید.

۲- برای رسم نمودار ستونی خوشه‌ای این دو متغیر، گزینه Display Clustered Bar Charts را علامت دار کنید.

گزینه Cells را انتخاب کنید تا به کادر محاوره آن منتقل شوید...

# روند Crosstab



- ۱- در کادر محاوره باز شده در قسمت Counts گزینه Observed علامتدار شده است، آن را تغییر ندهید.
- ۲- در بخش Percentages گزینه Row را علامت دار کنید.
- ۳- در پایان کلید Continue و Ok را به ترتیب کلیک کنید و نتایج را در خروجی SPSS مشاهده کنید.

**مشاهده نتایج**

> SPSS

# تغییر در داده ها (Recode)

## تمرین

در مجموعه‌ی داده‌های ۱۹۹۱، با بررسی متغیر سن، در خواهید یافت که کمترین سن ۱۸ و بیشترین سن ۸۹ سال ثبت شده است. می‌خواهیم این متغیر را به ۷ سطح به صورت زیر طبقه‌بندی کنیم و جدول توزیع فراوانی را کامل نماییم.

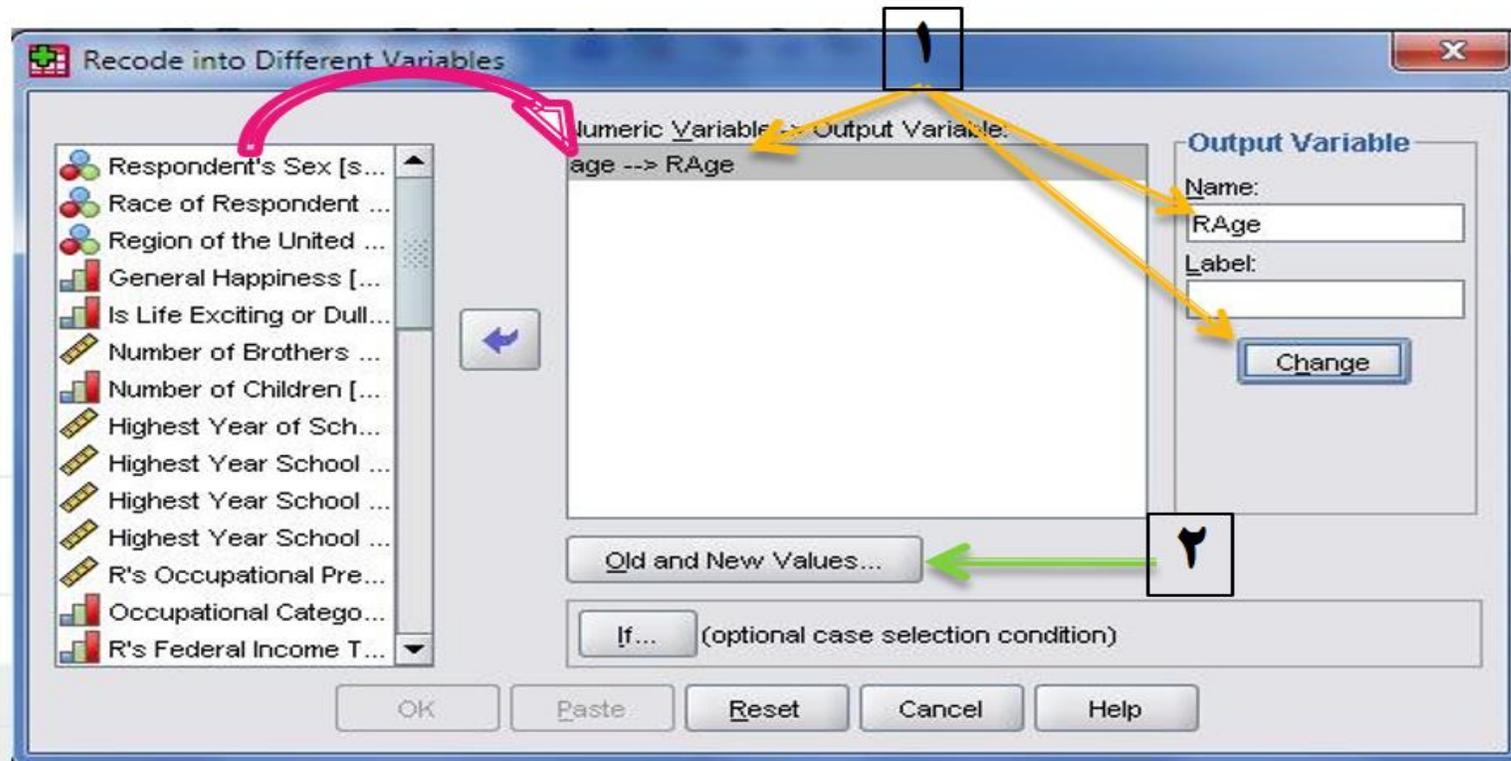
محدوده طبقه	فراوانی طبقه
۱۸ - ۳۰	؟
۳۰ - ۴۲	؟
۴۲ - ۵۴	؟
۵۴ - ۶۶	؟
۶۶ - ۷۸	؟
۷۸ - ۸۰	؟
۸۰ - ۹۲	؟

از نوار منو گزینه Transform و سپس گزینه Recode into Different Variable را انتخاب کنید تا پنجره Recode باز شود...

> SPSS

# تغییر در داده ها (Recode)

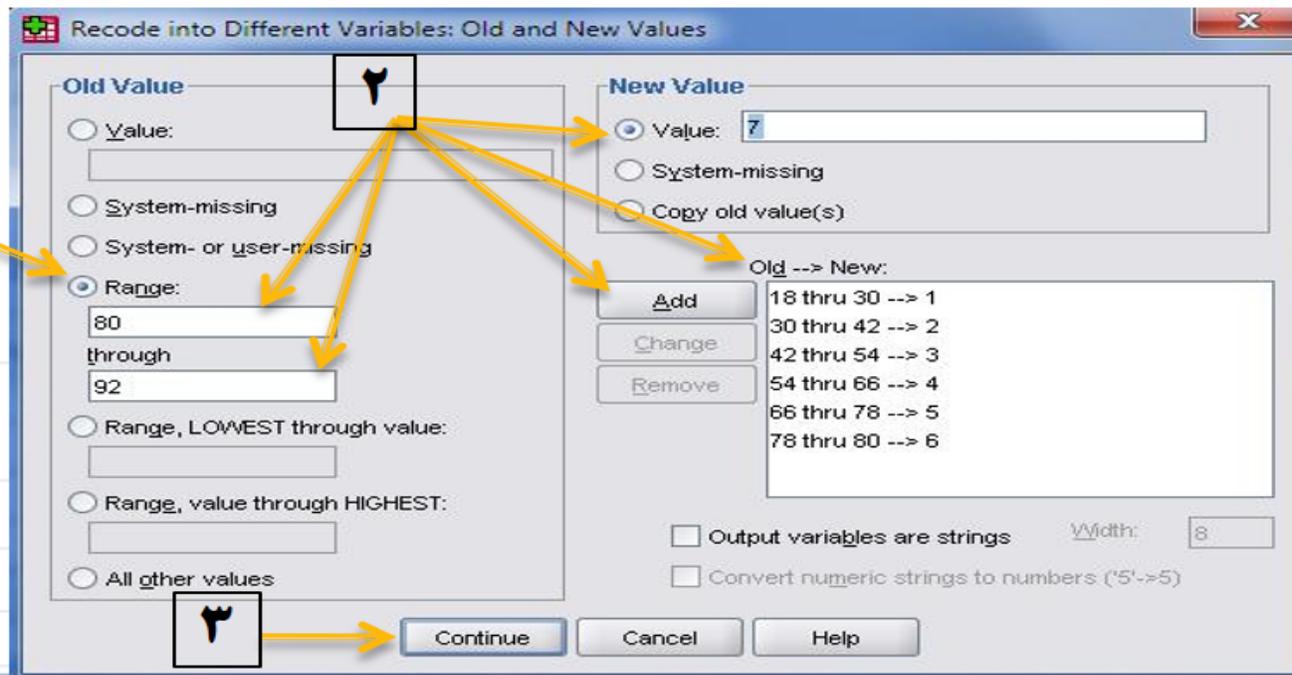
- ۱- متغیر سن را به سمت راست منتقل کنید و در کادر Name، نام RAge را برای متغیر جدیدی که می‌خواهید مقادیر تغییر یافته در آن ثبت شود، تعیین کنید و گزینه Change را کلیک کنید.
- ۲- برای اعمال تغییرات گزینه Old and New values را انتخاب کنید تا به کادر محاوره دیگری وارد شوید...



> SPSS

# تغییر در داده ها (Recode)

- ۱- در کادر محاوره باز شده، گزینه Range را انتخاب کنید و در این قسمت اولین فاصله را وارد کنید.
- ۲- با وارد کردن اولین فاصله (یعنی ۱۸ تا ۳۰)، در قسمت Value نیز مقدار ۱ را وارد کنید و دکمه Add را نیز به عنوان تایید، کلیک نمایید تا تغییرات به کادر Old → New: اضافه شود.
- ۳- در پایان کلید Continue و Ok را به ترتیب کلیک کنید



به شکل مقابل توجه کنید، فاصله آخرین طبقه (۸۰-۹۲) وارد شده و مقدار ۷ برای این فاصله در نظر گرفته شده است، فقط کافی است کلید Add فشار داده شود تا طبقه هفتم به فهرست تغییرات اضافه شود.

# تغییر در داده ها (Recode)

نتیجه این تغییرات اضافه شدن متغیری به نام RAge به انتهای فهرست متغیرها است.

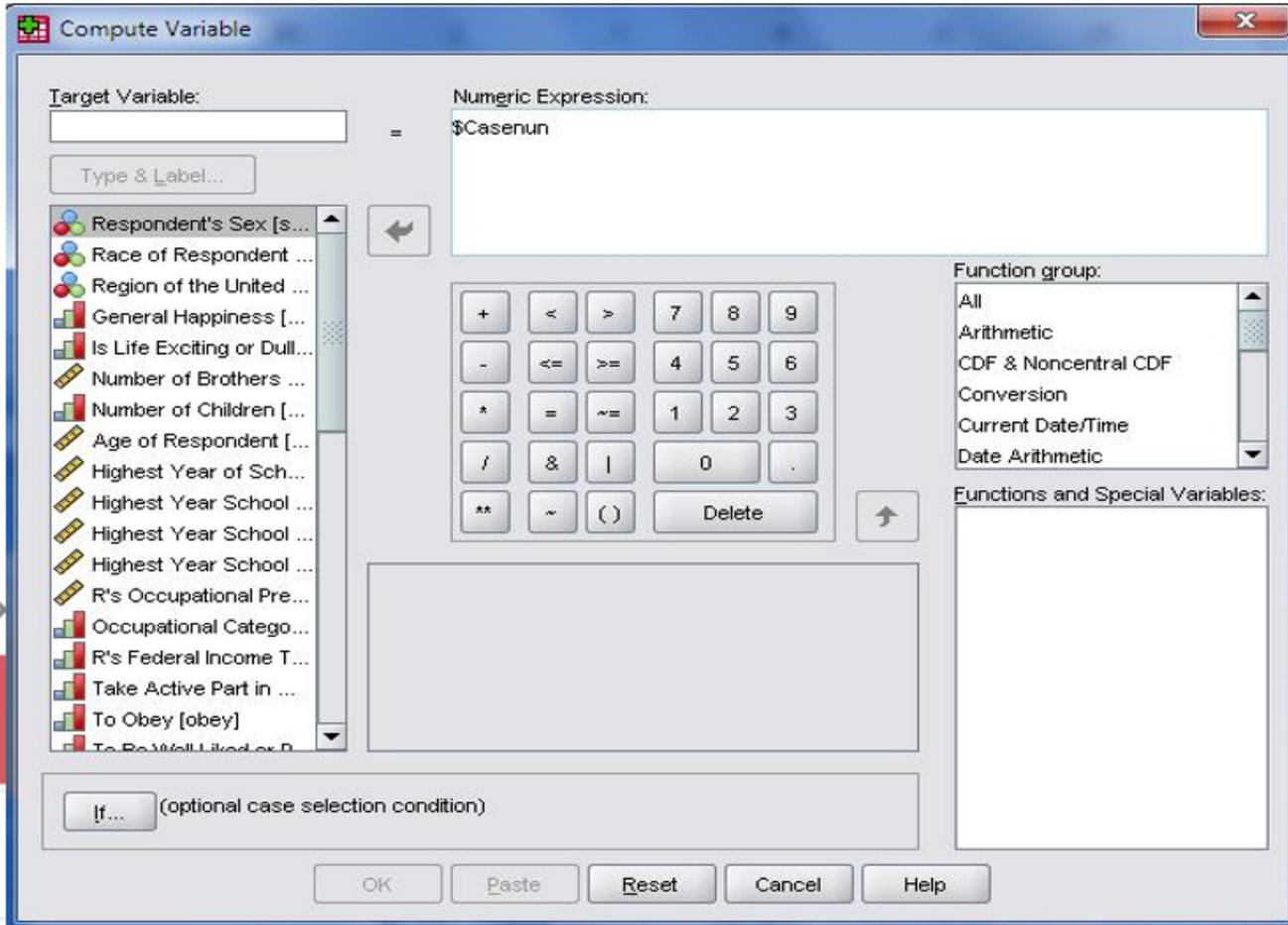
اگر از روند Frequencies برای این متغیر جدول توزیع فراوانی به دست آورید، به صورت زیر است.

توجه: حدود طبقات را بعدا خودمان به این جدول اضافه کرده ایم.

Rage						
جدول توزیع فراوانی برای متغیر دسته بندی شده سن در قابل داده های ۱۹۹۱						
	شماره طبقه	حدود طبقات	Frequency فراوانی مطلق	Percent فراوانی درصدی	Valid Percent فراوانی درصدی معتبر	Cumulative Percent فراوانی تجمعی درصدی
Valid	1	۱۸ - ۳۰	340	22.4	22.5	22.5
	2	۳۰ - ۴۲	451	29.7	29.8	52.2
	3	۴۲ - ۵۴	263	17.3	17.4	69.6
	4	۵۴ - ۶۶	196	12.9	12.9	82.6
	5	۶۶ - ۷۸	194	12.8	12.8	95.4
	6	۷۸ - ۸۰	18	1.2	1.2	96.6
	7	۸۰ - ۹۲	52	3.4	3.4	100.0
	Total		1514	99.8	100.0	
Missing	System		3	.2		
	Total		1517	100.0		

> SPSS

# ساختن ردیف



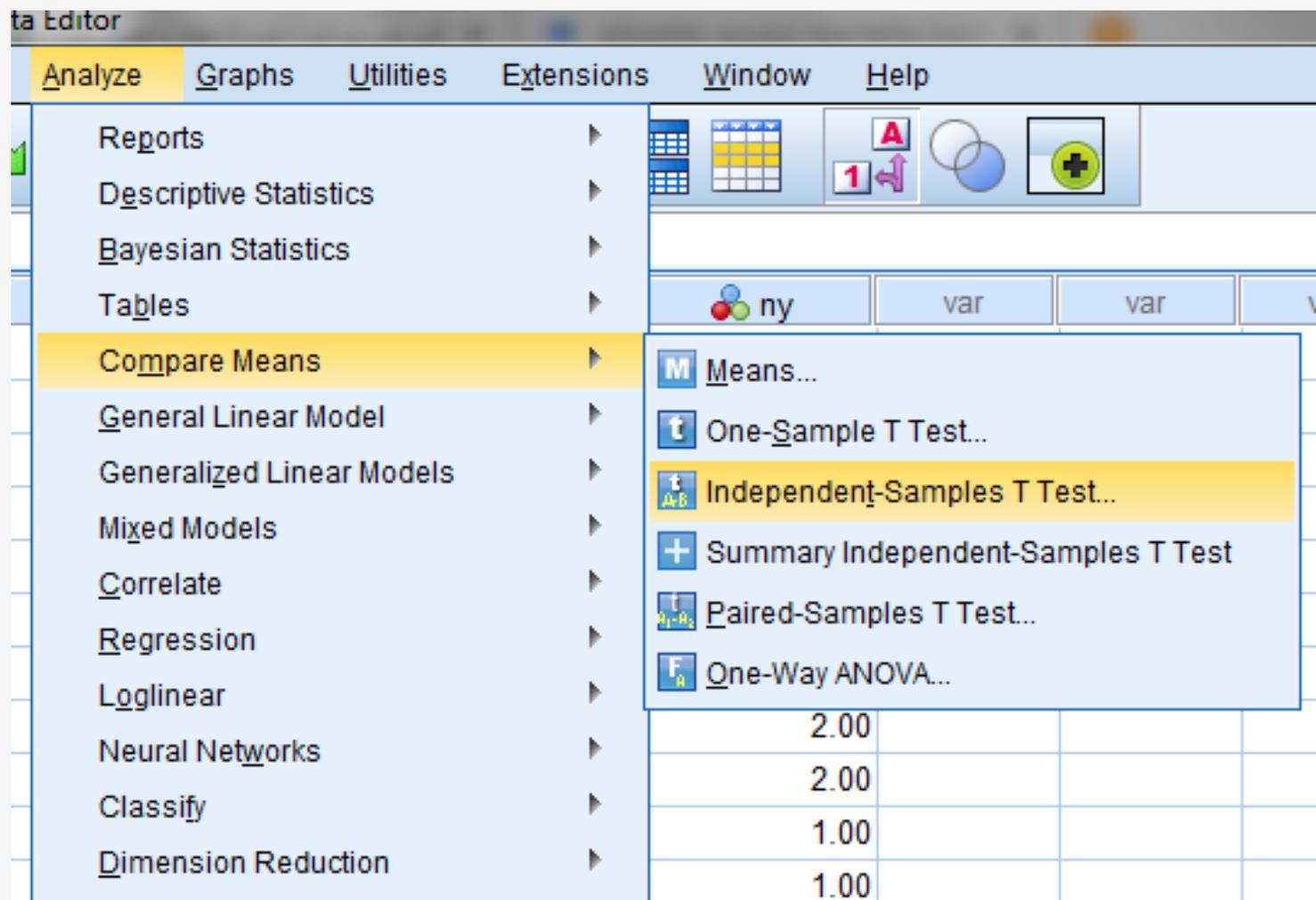
معمولا در یک مجموعه داده، داشتن یک ستون که ردیف هر مورد را نمایش دهد، الزامی به نظر می‌رسد. برای ساختن چنین ستونی باید از دستور Compute استفاده کرد.

از نوار منو گزینه Transform را انتخاب کنید و سپس گزینه compute را کلیک کنید تا پنجره آن باز شده در مستطیل target variable یک اسم دلخواه وارد کنید. سپس در مستطیل Numeric Expression عبارت \$Casenum را تایپ کرده و سپس OK را بزنید.

اینک یک ستون از شماره ۱ تا تعداد نمونه‌هایی که دارید در فهرست متغیرها اضافه شده است.

توجه کنید که نمی‌توان برای فایل داده خالی، شماره ردیف ایجاد کرد.

# آزمون تی مستقل



زمانی که بخواهید میانگین نمره را بین دو گروه مختلف برای یک متغیر پیوسته مقایسه کنید، از آزمون  $t$  مستقل استفاده می شود. به کمک این آزمون، می توانیم میانگین دو جامعه مستقل را با هم مقایسه کرده و نسبت به تفاوت آماری آنها، تصمیم بگیریم.

---

57.00	2.00								
50.00	2.00								
63.00	1.00								
69.00	2.00								

**Independent-Samples T Test**

Test Variable(s):  
age

Grouping Variable:  
sex(??)

Options...  
Bootstrap...

Define Groups

Use specified values  
Group 1: 1  
Group 2: 2

Cut point:

Continue Cancel Help

OK Paste Reset Cancel Help

## T-Test

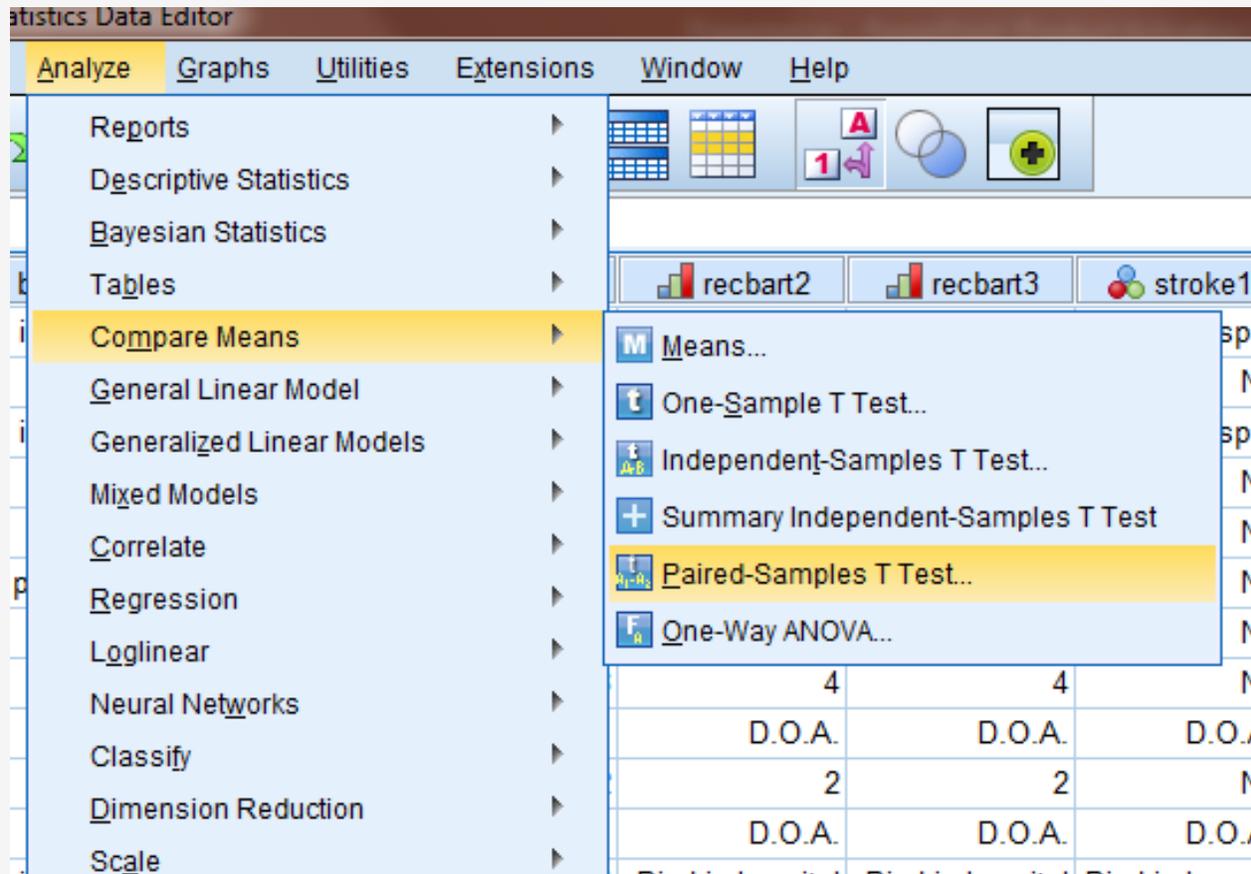
[DataSet1] C:\Users\maryam\Desktop\sick.sav

### Group Statistics

	sex	N	Mean	Std. Deviation	Std. Error Mean
age	female	15	39.33	10.939	2.824
	male	11	29.45	6.714	2.024

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
age	Equal variances assumed	6.136	.021	2.644	24	.014	9.879
	Equal variances not assumed			2.843	23.424	.009	9.879



اگر شما یک گروه داشته باشید و یک متغیر را در دو بازه زمانی در این گروه اندازه گرفته باشید می‌توانید با تحلیل  $t$  زوجی آن را مقایسه کنید. بیشتر کاربرد این تحلیل زمانی است که یک گروه تحت یک مداخله درمانی قرار گرفته باشد و ما قصد داشته باشیم نمره این گروه را قبل و بعد از مداخله درمانی با یکدیگر مقایسه کنیم و ببینیم تفاوتی در این نمره ایجاد شده است یا خیر.

Paired Variables:

Pair	Variable1	Variable2
1	Stroke b...	Stroke b...
2		

Options...  
Bootstrap...

OK Paste Reset Cancel Help

۹- بر اساس یک دوره ورزش مرتب به ۸ دانش آموز طی یک ماه، می خواهیم بدانیم آیا این نوع ورزش تفاوتی معنی دار در وزن دانش آموزان یک مدرسه ایجاد کرده است یا نه؟ داده های زیر وزن دانش آموزان را در قبل از انجام دوره و بعد از آن نشان می دهد.

قبل از دوره	۴۸	۵۳	۳۵/۵	۴۶	۵۹	۶۱	۴۲	۴۸
بعد از دوره	۴۶/۵	۵۲	۳۶	۴۲/۵	۵۶/۵	۶۰	۴۰/۵	۴۷

### Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Stroke between release and 1 month	.03	752	.176	.006
	Stroke between 1 and 3 months	.05	752	.216	.008

### Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Stroke between release and 1 month & Stroke between 1 and 3 months	752	.169	.000

### Paired Samples Test

		Mean	Std. Deviation	Std. Error Mean	Paired Differences		t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Stroke between release and 1 month - Stroke between 1 and 3 months	-.017	.255	.009	-.036	.001	-1.860	751	.063

