



# مقایسه میانگین تک جامعه و دو جامعه مستقل

جلسه دوم از سری کارگاه های آموزش SPSS برای اساتید

مدرس سرکار خانم دکتر سید سولماز طالبی



# سوال

- محققین ادعا میکنند که برنامه های آگاهی بخش در زمینه سرطان کمک کننده بوده و به همین دلیل متوسط سن تشخیص سرطان کاهش چشم گیری داشته است. متخصصین بیان میکنند اگر میانگین سن تشخیص ۱۰ سال کاهش داشته باشد این ادعا درست است. حال ما بدنبال بررسی این ادعا هستیم. **(آیا میانگین سنی زنان مبتلا به سرطان بین سالهای ۸۰ تا ۸۵، ۴۰ سال است؟)**

- فعالان حوزه زنان ادعا میکنند که زنان شاغل بدلیل ارتباط با همکاران و دریافت سریعتر اطلاعات بیماری را سریعتر تشخیص میدهند. برای آزمون این ادعا باید **فرض برابری میانگین سنی زنان خانه دار و شاغل مبتلا به سرطان را آزمود.**



# پیش فرض ها

- نرمال بودن داده ها ( اگر به صورت چشمی برقرار باشد کافی است)
- در صورت برقرار نبودن این فرض، با تبدیلاتی بر روی داده ها سعی میکنیم تا فرض را برقرار کنیم.
- اگر به هیچ صورتی داده ها نرمال نشدند یا حجم داده ها کم بود، باید از روش های ناپارامتری استفاده کنیم.

# آزمون نرمالیتی داده ها

- آزمون کلموگروف اسمیرنوف:

IBM SPSS Statistics Data Editor

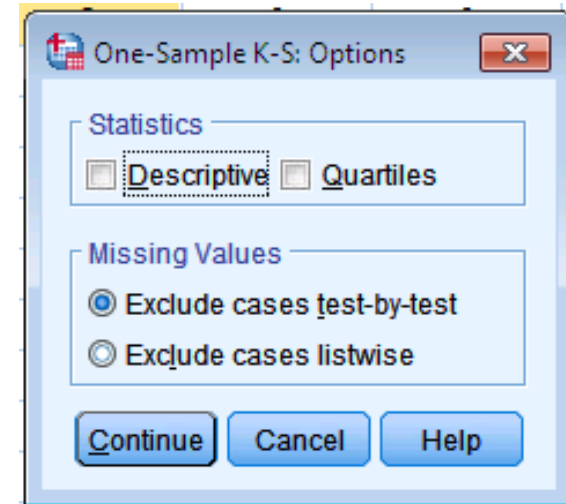
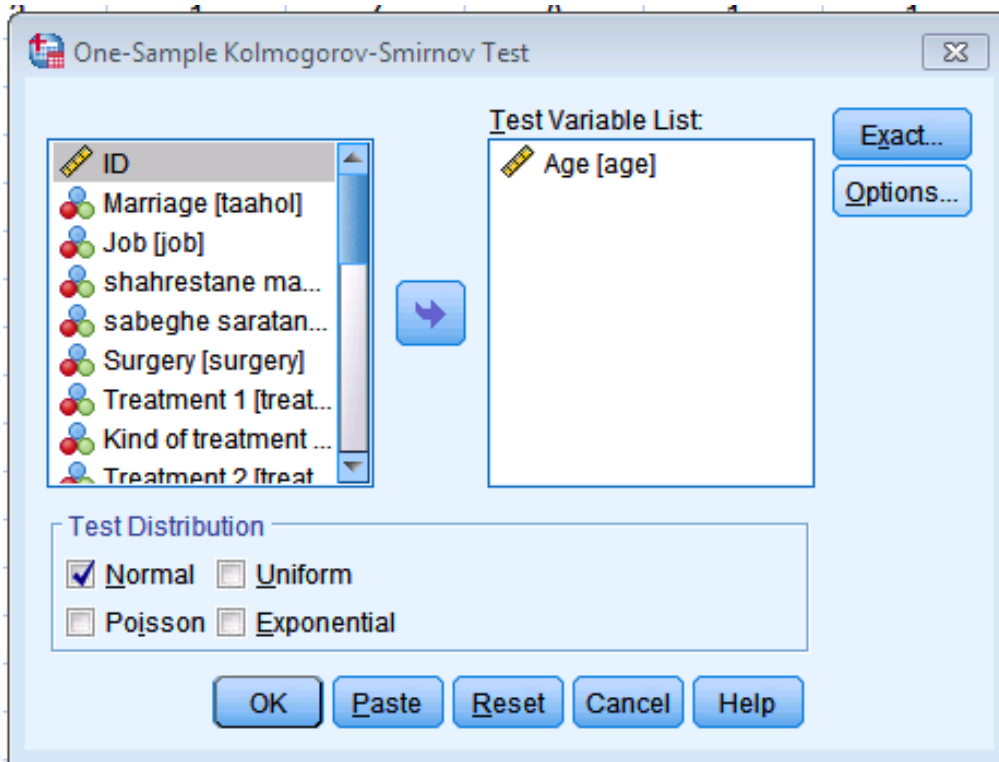
File Edit View Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Reports  
Descriptive Statistics  
Neural Networks  
Classify  
Dimension Reduction  
Scale  
Nonparametric Tests  
Forecasting  
Survival  
Multiple Response  
Missing Value Analysis...  
Multiple Imputation  
Complex Samples  
Simulation...

0	2	1	2	0
0	2	0	0	0
0	1	1	1	0
0	2	1	1	0
			1	1
0	5	0		
0	1	0		
0	2	1		
0	1	1		

Related Samples...  
Independent Samples...  
One Sample...  
Legacy Dialogs  
Chi-square...  
Binomial...  
Runs...  
1-Sample K-S...

# آزمون نرمالیتی داده ها



# آزمون نرمالیتی

## One-Sample Kolmogorov-Smirnov Test

		Age
N		100
Normal Parameters <sup>a,b</sup>	Mean	50.84
	Std. Deviation	14.551
Most Extreme Differences	Absolute	.083
	Positive	.083
	Negative	-.047
Test Statistic		.083
Asymp. Sig. (2-tailed)		.086 <sup>c</sup>

a. Test distribution is Normal

b. Calculated from data.

c. Lilliefors Significance Correction.

فرض آزمون

$$\begin{cases} H_0 = \text{distribution is Normal} \\ H_1 = \text{distribution is not Normal} \end{cases}$$



## نکته

- آزمون نرمالیتی نیز مانند هر آزمون دیگر به حجم نمونه حساس است. به عبارت دیگر با افزایش حجم نمونه حساست آزمون افزایش میابد و تخطی کم از فرض صفر را نیز شناسایی میکند. در نتیجه آزمون معنادار میشود.



# چک کردن نرمالیتی توسط نمودار

Graphs Utilities Add-ons Window Help

Chart Builder...  
Graphboard Template Chooser...  
+ Weibull Plot...  
+ Compare Subgroups  
+ Regression Variable Plots

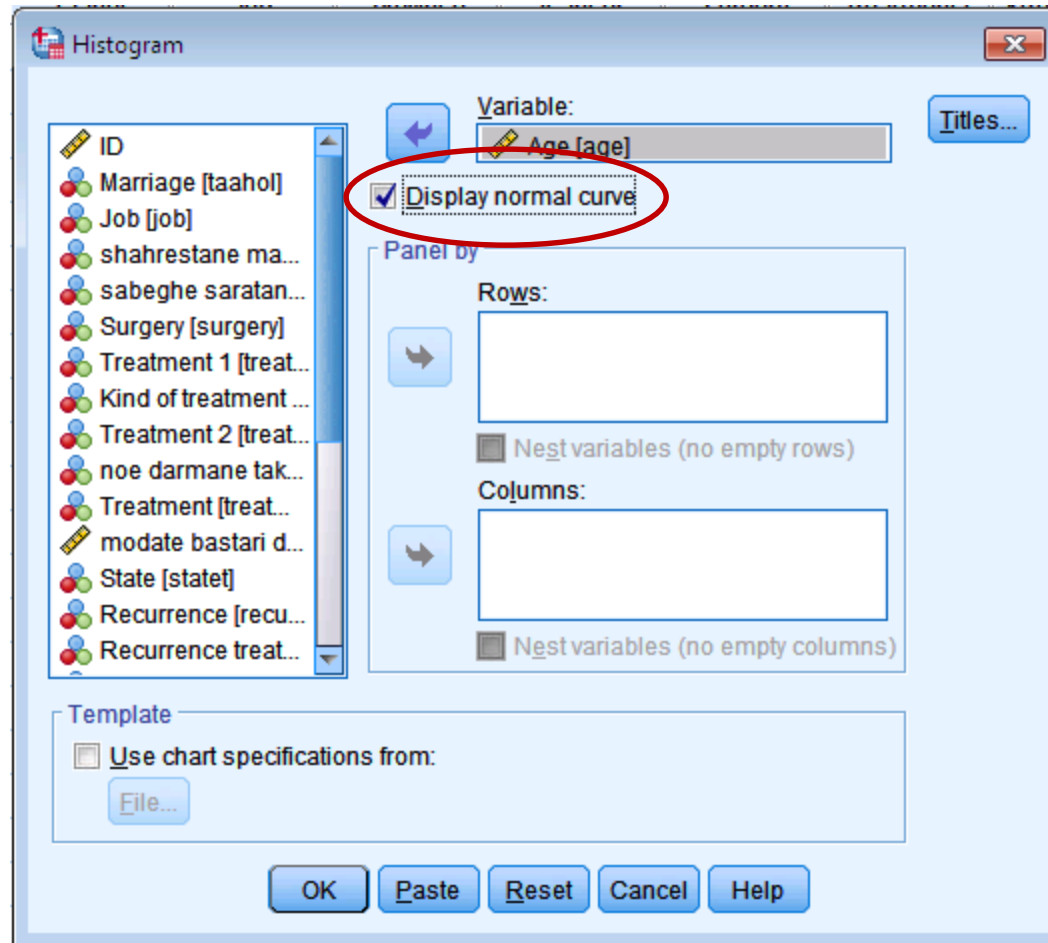
Legacy Dialogs

	atment1	ktreatment1	trea
7	0	1	
7	0	3	
4	0	1	
7	0	4	
3	0	2	
7	0	2	
7	0	1	
2	0	2	
5	0	1	
2	0	1	
2	0	2	
5	0	3	

Bar...  
3-D Bar...  
Line...  
Area...  
Pie...  
High-Low...  
Boxplot...  
Error Bar...  
Population Pyramid...  
Scatter/Dot...  
Histogram...

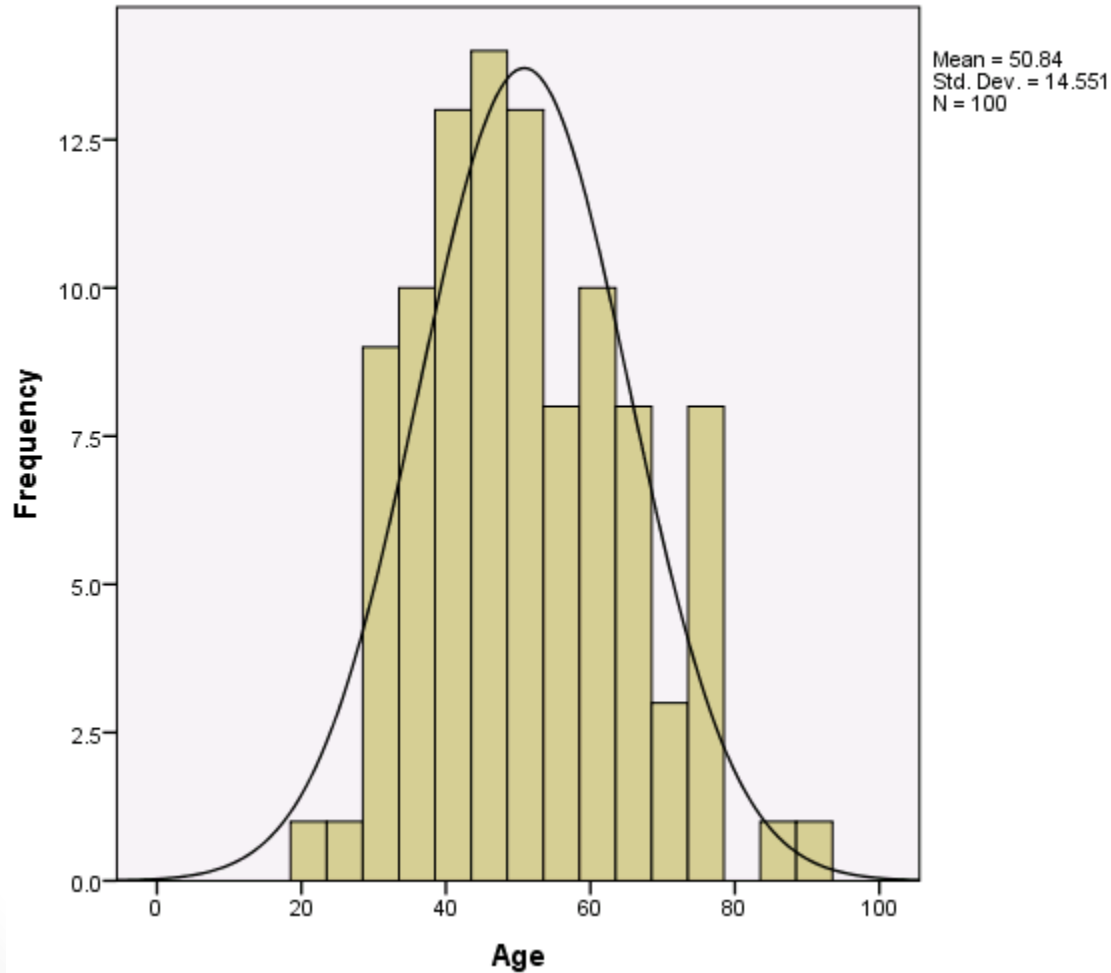


# چک کردن نرمالیتی توسط نمودار





# چک کردن نرمالیتی توسط نمودار





# استفاده از تبدیلات جهت نرمال کردن داده ها

- Transform Analyze Direct M
- Compute Variable...
- Programmability Transforma
- Count Values within Cases...
- Shift Values...

Target Variable: logAGE

Numeric Expression: LN(age)

Type & Label...

ID  
Age [age]  
Marriage [taahol]  
Job [job]  
shahrestane ma...  
sabeghe saratan...  
Surgery [surgery]  
Treatment 1 [treat...]  
Kind of treatment ...  
Treatment 2 [treat...]  
noe darmane tak...  
Treatment [treat...]  
modate bastani d...  
State [statet]  
Recurrence [recu...]  
Recurrence treat...  
Pathology [pathol...]  
Side [side]  
Metastases [met

Function group:  
All  
Arithmetic  
CDF & Noncentral CDF  
Conversion  
Current Date/Time  
Date Arithmetic  
Date Creation

Functions and Special Variables:  
Abs  
Arsin  
Artan  
Cos  
Exp  
Lg10  
Ln  
Lngamma  
Mod  
Rnd(1)

LN(numexpr). Numeric. Returns the base-e logarithm of numexpr, which must be numeric and greater than 0.

If... (optional case selection condition)

OK Paste Reset Cancel Help

دو ستاره یعنی توان



## استفاده از تبدیلات جهت نرمال کردن داده ها

- بعد از هر تبدیل با استفاده از آزمون و نمودار نرمالیتی متغیر را چک میکنیم.
- تبدیلات پرکاربرد:
  - لگاریتم
  - توان دو
  - جذر
  - معکوس

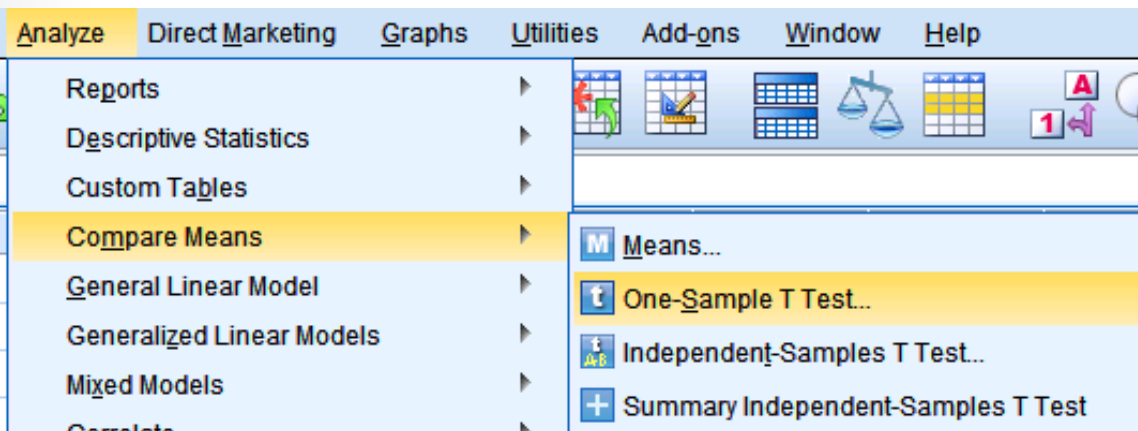


# سوال

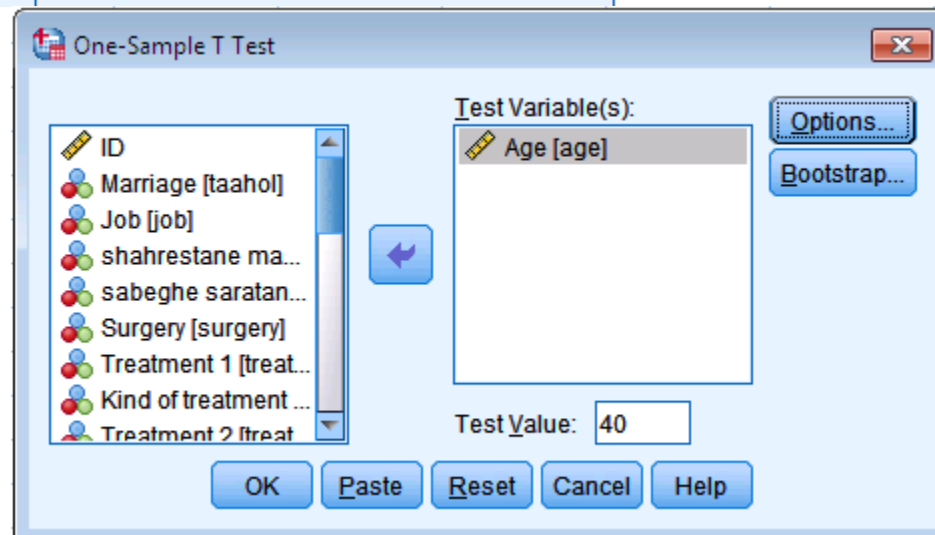
- نرمالیتی متغیر مدت زمان بستری در بیمارستان (bastari) را آزمون کنید.
- در صورت نرمال نبودن تبدیل لگاریتمی و توان دو را اعمال کنید.

# مقایسه میانگین یک جامعه با مقدار پیش فرض

- آیا میانگین سنی زنان مبتلا به سرطان بین سالهای ۸۰ تا ۸۵، ۴۰ سال است؟



$$\begin{cases} H_0: \mu_{age} = 40 \\ H_1: \mu_{age} \neq 40 \end{cases}$$





## مقایسه میانگین یک جامعه با مقدار پیش فرض

### One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Age	100	50.84	14.551	1.455

### One-Sample Test

	Test Value = 40					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Age	7.450	99	.000	10.840	7.95	13.73



# مقایسه میانگین یک جامعه با مقدار پیش فرض

• سوال:

آزمون نشان داد که میانگین متغیر سن ۴۰ سال نبوده، بیشتر از ۴۰ است یا کمتر؟

## One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Age	100	50.84	14.551	1.455

## One-Sample Test

	Test Value = 40					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Age	7.450	99	.000	10.840	7.95	13.73





## مقایسه میانگین یک جامعه با مقدار پیش فرض

• نکته:

اگر متغیر مورد نظر نرمال نبود و تبدیل برای نرمالیتی اعمال شده بود، باید همان تبدیل روی مقدار مورد نظر برای آزمون نیز اعمال شود.

• مثال:

اگر هدف مقایسه میانگین متغیر سن با عدد ۴۰ باشد ولی تبدیل لگاریتمی بر روی این متغیر اعمال شده باشد باید میانگین متغیر تبدیل یافته را با لگاریتم ۴۰ مقایسه کنیم.



## مقایسه میانگین یک جامعه با مقدار پیش فرض

• سوال:

مدیریت بیمارستان به منظور بهبود وضعیت بیمارستان قصد خرید تجهیزات و تخت بیمارستانی را دارد. برای برآورد تعداد تخت مورد نیاز دانستن میانگین مدت زمان بستری مهم است. مدیریت این بیمارستان میخواهد بداند آیا میانگین مدت زمان بستری ۱۰ روز است؟



# مقایسه میانگینهای دو جامعه مستقل

- از این آزمون زمانی استفاده میشود که هدف بررسی برابری میانگین یک متغیر در دو گروه مستقل است. بنابراین متغیر وابسته باید کمی نرمال و متغیر مستقل طبقه ایی دو حالتی باشد.

• مثال:

- بررسی یکسانی میانگین قد در دختران و پسران نوجوان
  - بررسی یکسانی میانگین سطح هموگلوبین در مردان و زنان
  - بررسی یکسانی میانگین مدت زمان تحرک در زنان خانه دار و شاغل
- فرض آزمون:

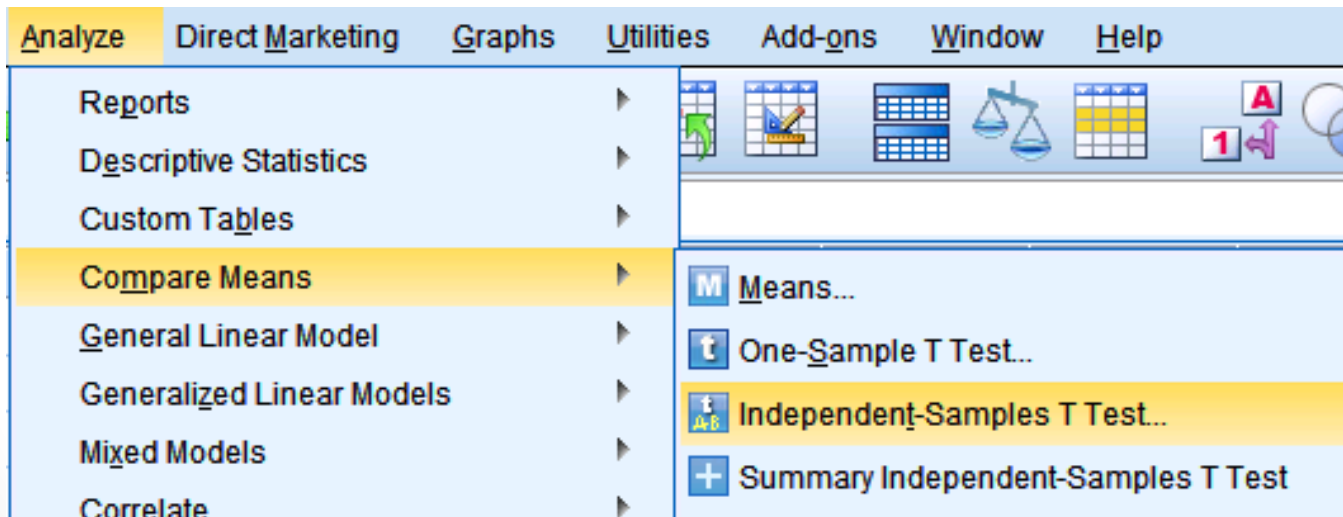
نرمال بودن داده ها و یکسانی واریانس متغیر در گروه ها



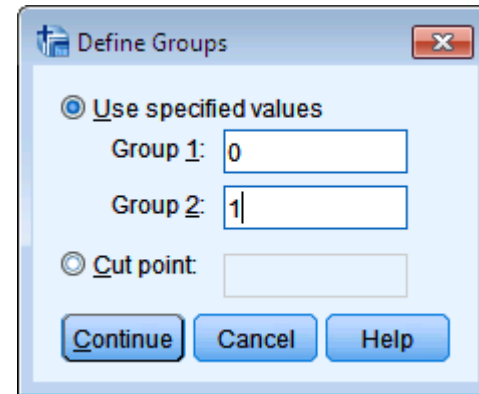
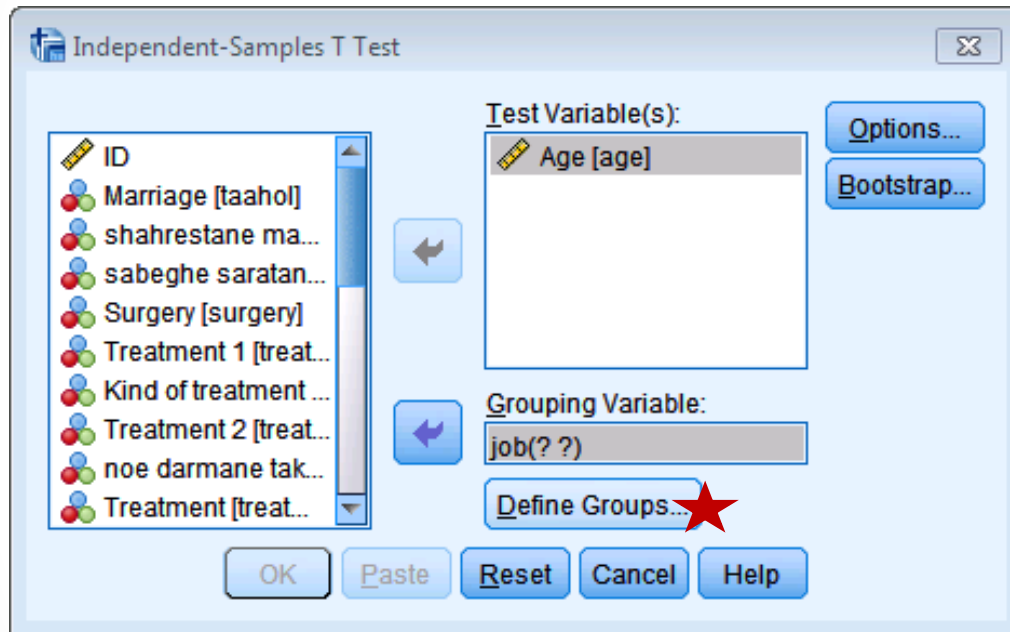
# مقایسه میانگینهای دو جامعه مستقل

آیا میانگین سنی زنان خانه دار و شاغل مبتلا به سرطان با هم برابر است؟

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$



# مقایسه میانگینهای دو جامعه مستقل





# مقایسه میانگینهای دو جامعه مستقل

Group Statistics

Job	N	Mean	Std. Deviation	Std. Error Mean
Age khanedar	84	52.73	14.659	1.599
shaghel	16	40.94	9.161	2.290

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Age	Equal variances assumed	5.374	.023	3.096	98	.003	11.789	3.807	4.233	19.344
	Equal variances not assumed			4.220	31.829	.000	11.789	2.793	6.097	17.480



# مقایسه میانگینهای دو جامعه مستقل

Group Statistics

Job	N	Mean	Std. Deviation	Std. Error Mean
Age khanedar	84	52.73	14.659	1.599
shaghel	16	40.94	9.161	2.290

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Age	Equal variances assumed	5.374	.023	3.096	98	.003	11.789	3.807	4.233	19.344
	Equal variances not assumed			4.220	31.829	.000	11.789	2.793	6.097	17.480



## مقایسه میانگینهای دو جامعه مستقل

- با توجه به آزمون فرض برابر میانگین های سن تشخیص در زنان خانه دارو شاغل رد شد. اینک با توجه به میانگین سنی در هر گروه میتوان گفت که میانگین سن تشخیص در زنان شاغل به صورت معنا داری کمتر از زنان خانه دار است.

Group Statistics

	Job	N	Mean	Std. Deviation	Std. Error Mean
Age	khanedar	84	52.73	14.659	1.599
	shaghel	16	40.94	9.161	2.290





# مقایسه میانگینهای دو جامعه مستقل

مثال:

آیا میانگین تعداد غدد بررسی شده در جراحی ماستکتومی و ماستکتومی و اگزیلاری با سایر جراحی ها متفاوت است؟

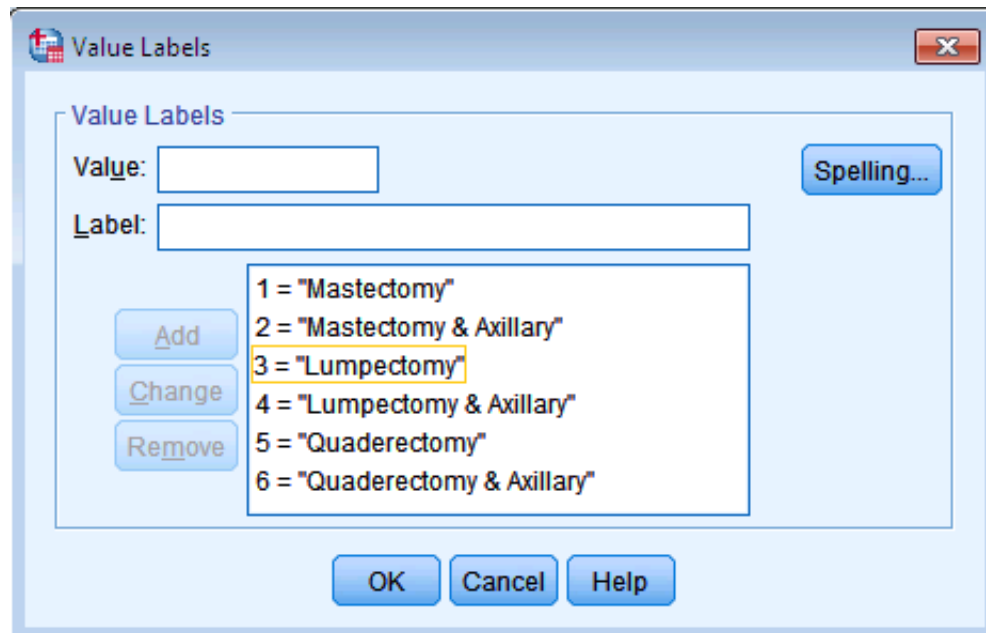
برای انجام این آزمون باید:

۱. نرمالیتی متغیر تعداد غدد بررسی شده، مورد ارزیابی قرار گیرد. با توجه به نمودار به نظر میرسد تبدیل جذر مناسب باشد.

۲. متغیر نوع جراحی به متغیری تبدیل شود که فقط دو نوع جراحی (ماستکتومی یا ماستکتومی و اگزیلاری و سایر جراحی ها) را داشته باشد.

# مقایسه میانگینهای دو جامعه مستقل

- هدف ما ساختن متغیری است که به جای ۶ نوع عمل جراحی دو عمل جراحی داشته باشد.

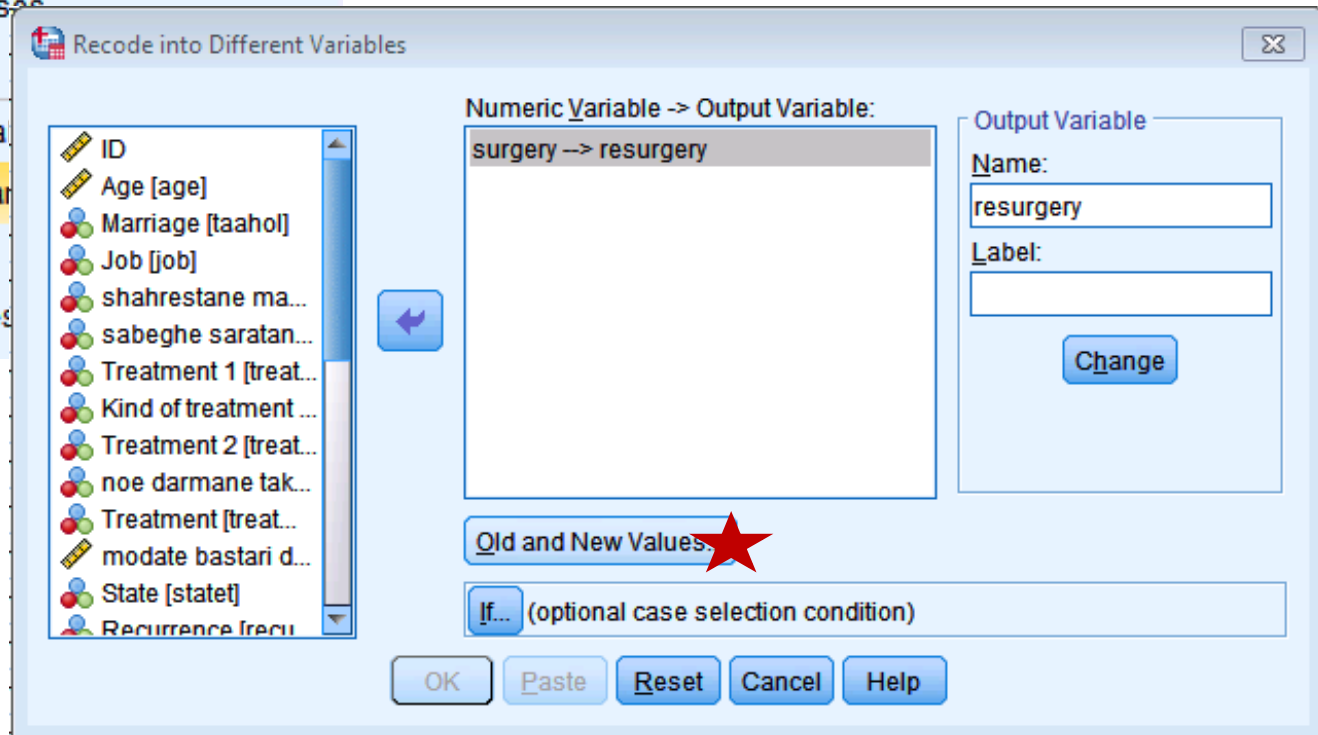
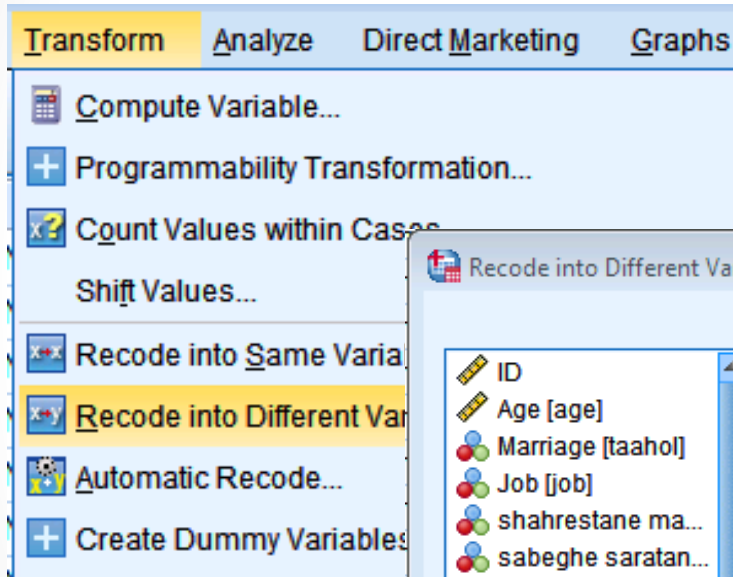


The image shows a 'Value Labels' dialog box from a statistical software package. It contains a list of six categories with their corresponding labels. The third category, '3 = "Lumpectomy"', is highlighted with a yellow background. The dialog box also includes input fields for 'Value' and 'Label', a 'Spelling...' button, and 'Add', 'Change', and 'Remove' buttons. At the bottom, there are 'OK', 'Cancel', and 'Help' buttons.

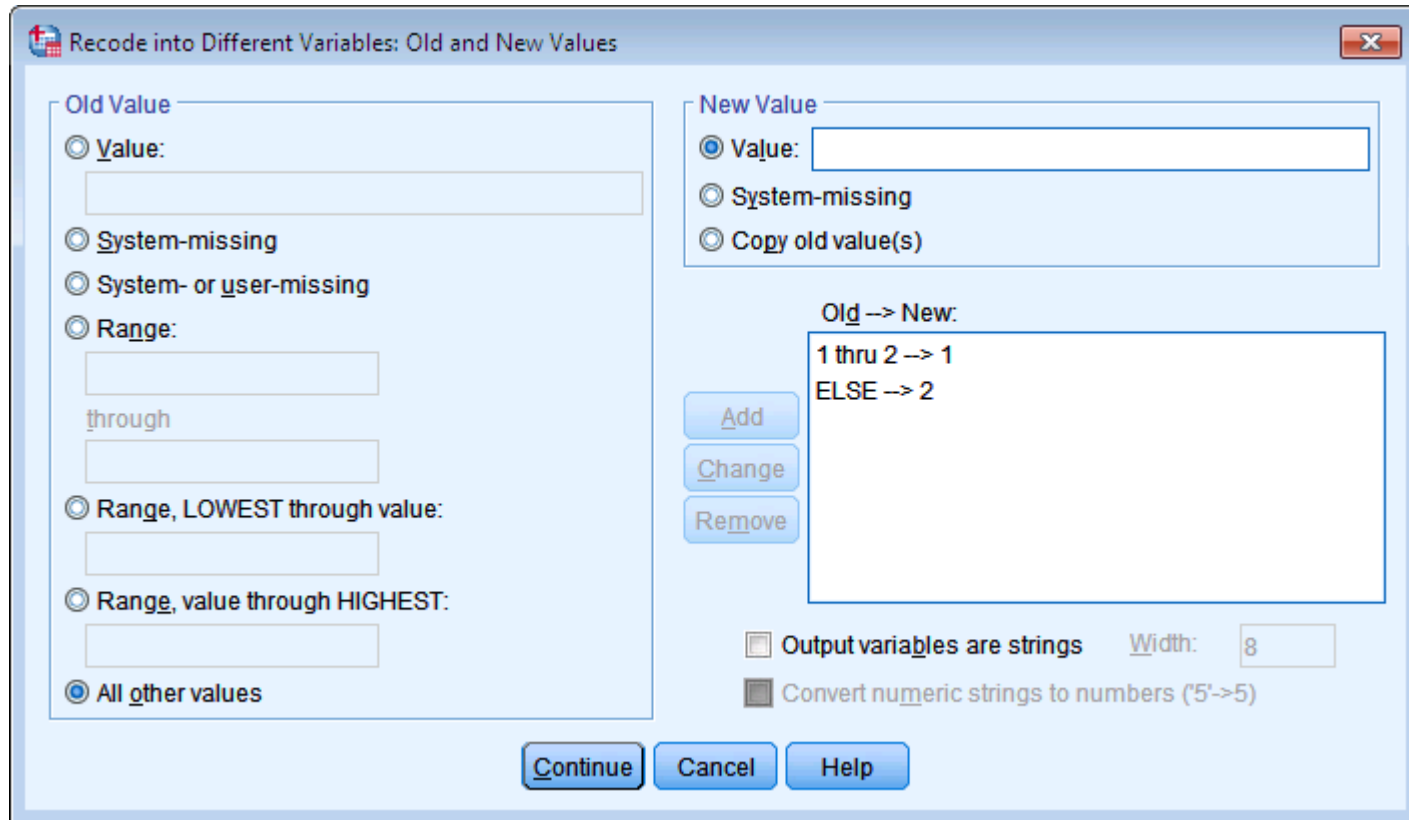
Value	Label
1	"Mastectomy"
2	"Mastectomy & Axillary"
3	"Lumpectomy"
4	"Lumpectomy & Axillary"
5	"Quaderectomy"
6	"Quaderectomy & Axillary"

# مقایسه میانگینهای دو جامعه مستقل

- برای این منظور از مسیر زیر استفاده میکنیم.



# مقایسه میانگینهای دو جامعه مستقل



Recode into Different Variables: Old and New Values

**Old Value**

Value:

System-missing

System- or user-missing

Range:

Range, LOWEST through value:

Range, value through HIGHEST:

All other values

**New Value**

Value:

System-missing

Copy old value(s)

Old -> New:

1 thru 2 -> 1  
ELSE -> 2

Add  
Change  
Remove

Output variables are strings Width: 8

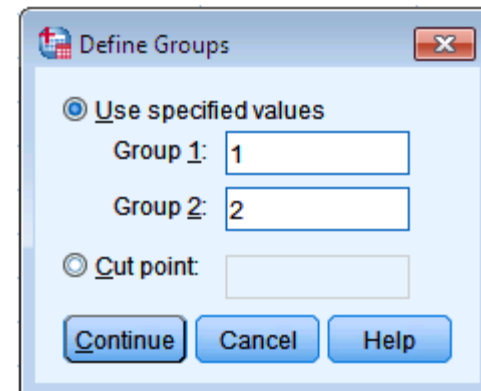
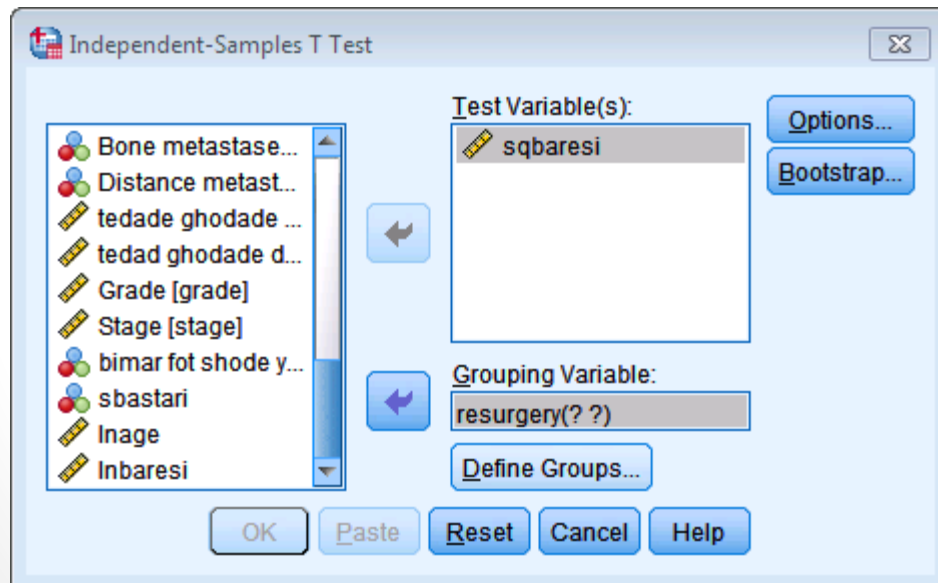
Convert numeric strings to numbers ('5'->5)

Continue Cancel Help

• متغیر جدید در انتهای داده ها ساخته شد.

# مقایسه میانگینهای دو جامعه مستقل

- آیا میانگین تعداد غدد بررسی شده در جراحی ماستکتومی و ماستکتومی و اگزیلاری با سایر جراحی ها متفاوت است؟





# مقایسه میانگینهای دو جامعه مستقل

## Group Statistics

	resurgery	N	Mean	Std. Deviation	Std. Error Mean
sqbares1	1.00	81	3.3569	.77651	.08628
	2.00	19	3.3678	.72640	.16665

## Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
sqbares1	Equal variances assumed	.026	.872	-.056	98	.956	-.01094	.19565	-.39920	.37733
	Equal variances not assumed			-.058	28.482	.954	-.01094	.18766	-.39504	.37317



# مقایسه میانگینهای دو جامعه مستقل

- سوال:
- آیا میانگین تعداد غدد درگیر در زنان با متاستاز و بدون متاستاز برابر است؟