

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



## همبستگی، آزمون ارتباط بین دو متغیر کیفی و آزمون مقایسه میانگین دو جامعه همبسته

جلسه چهارم از سری کارگاه های آموزش SPSS برای اساتید

مدرس سرکار خانم دکتر سیده سولماز طالبی

داده های جمع آوری شده در زمان های مختلف از یک فرد (مطالعات قبل و بعد یا مطالعات کوهورت) مثال: اندازه گیری قند خون در زمان های مختلف

مطالعات مقطعی که گروه ها همسان شده اند (Match Study) مثال: از نظر گروه خونی گروه آزمون و کنترل یکسان اند.

مطالعه مقطعی دانش آموزان یک کلاس، افراد یک خانواده، دوقلوها

مطالعه مقطعی دو اندازه گیری از یک فرد. مثال: نمره کسب شده هر فرد در دو پرسشنامه مختلف

## همبستگی

### • مثال:

میزان همبستگی فشارخون سیستولیک و دیاستولیک مادران در ۳ ماه اول چقدر است؟  
میزان همبستگی فشارخون سیستولیک مادران در ۳ ماه اول و ۳ ماه دوم بارداری چقدر است؟  
میزان همبستگی وزن فرزندان در زایمانهای مختلف یک خانم

$$\begin{cases} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{cases}$$

صفر بودن همبستگی به معنی استقلال است.  
بنابراین اگر فرض صفر رد شود یعنی داده ها  
همبسته اند.

## همبستگی

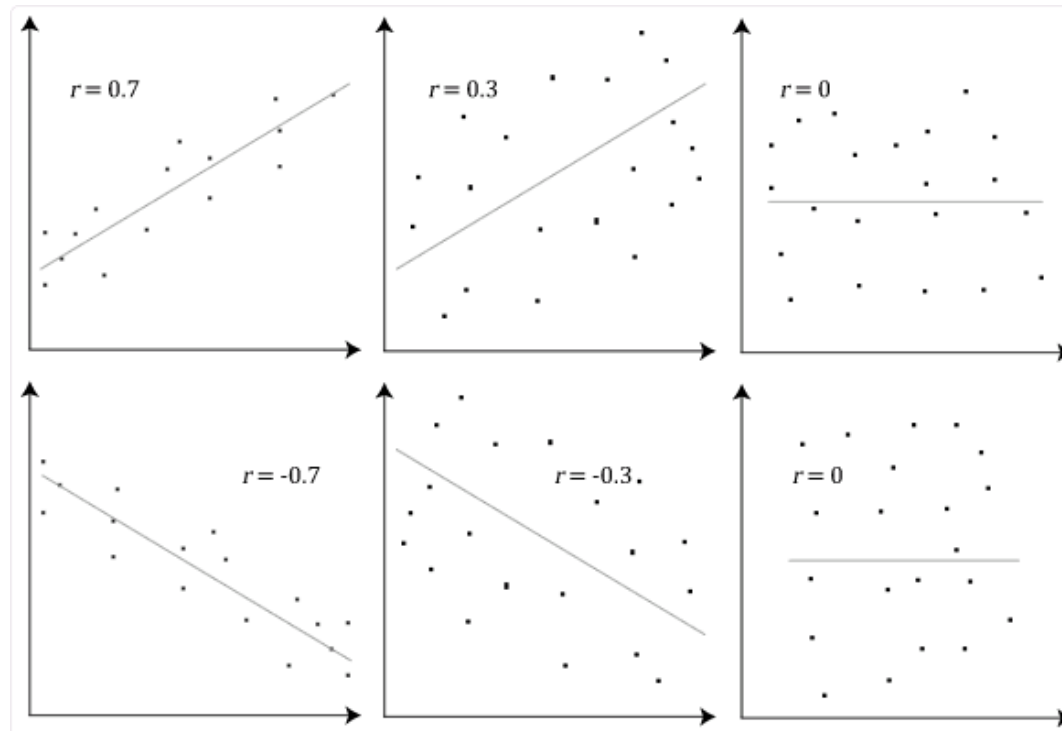
- پیش فرض ها:

بررسی نرمالیتی به منظور انتخاب درست ضریب همبستگی  
عدم وجود داده پرت

داده پرت داده ایی است که از ابر داده ها فاصله داشته باشد. شناسایی و بررسی این داده مهم است زیرا میتواند آنالیز ما را تحت تاثیر قرار دهد. اشتباه ترین عمل در زمان مواجهه با داده پرت حذف آن است.

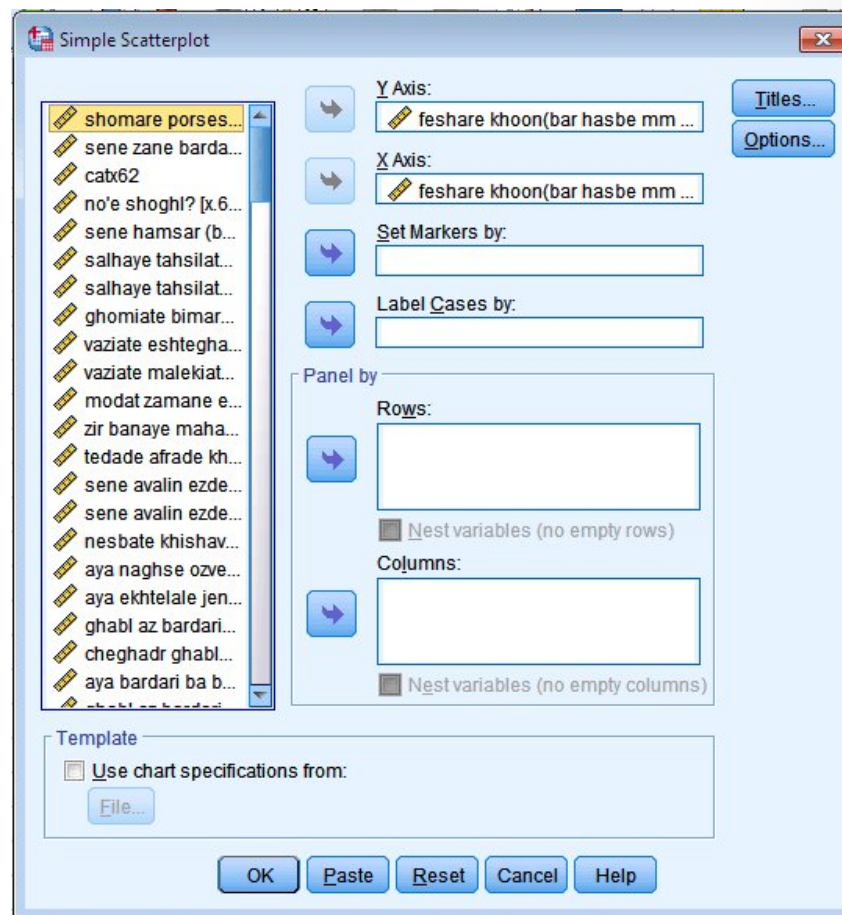
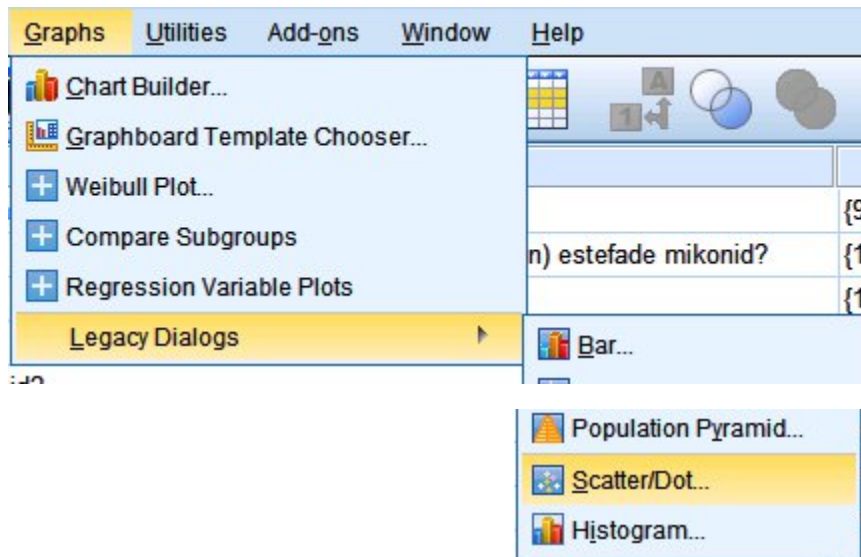
- سوال:

با داده پرت چه کنیم؟ بررسی میکنیم اگر داده صحت داشت وارد آنالیز میشود، اگر اشتباه بود تصحیح میکنیم.

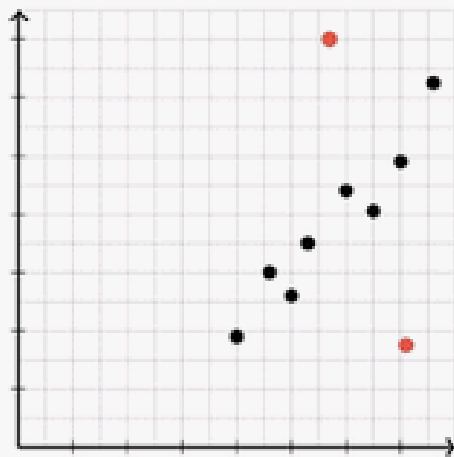
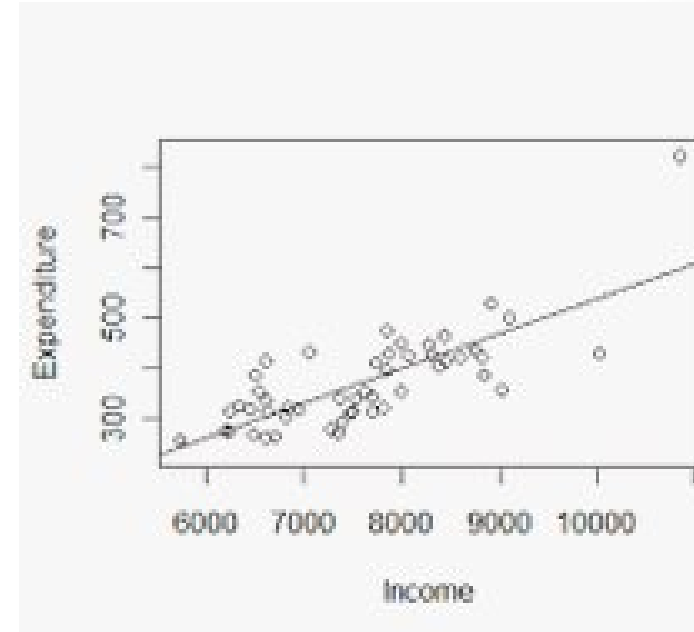
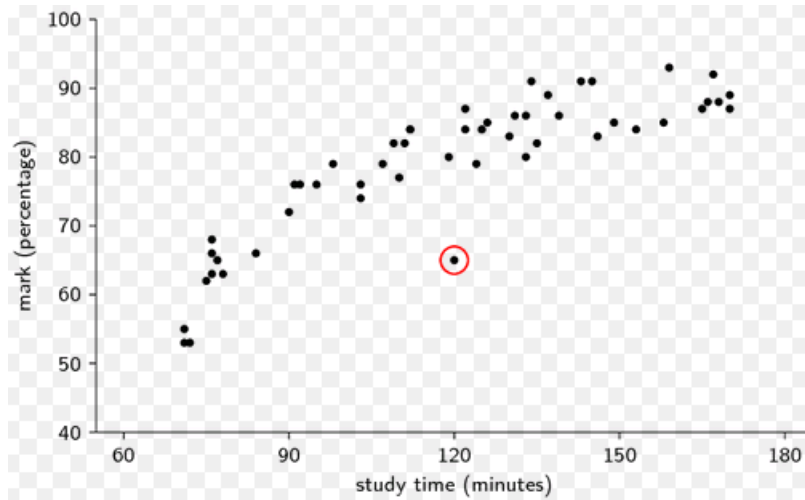


Strength of Association	Coefficient, $r$	
	Positive	Negative
Small	.1 to .3	-0.1 to -0.3
Medium	.3 to .5	-0.3 to -0.5
Large	.5 to 1.0	-0.5 to -1.0

# بررسی عدم وجود داده پرت (Outliers)



# بررسی عدم وجود داده پرت



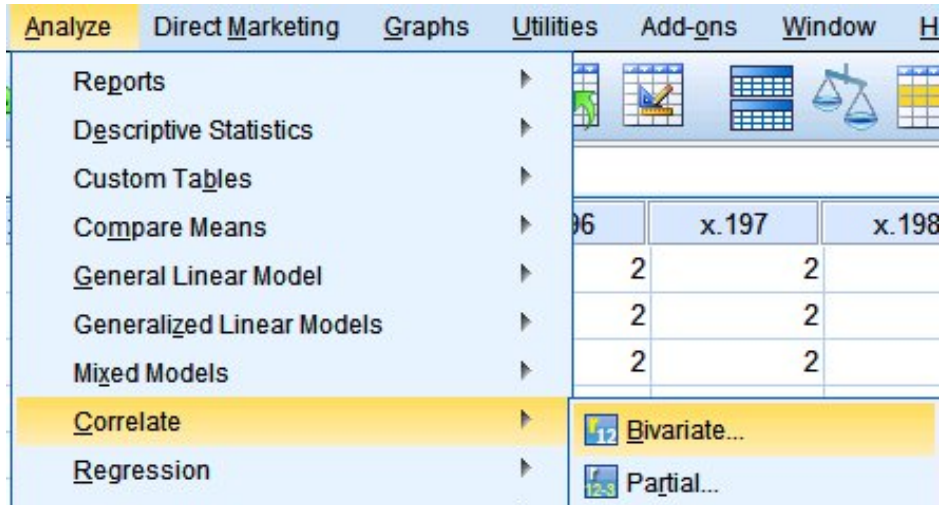




## همبستگی

مثال:

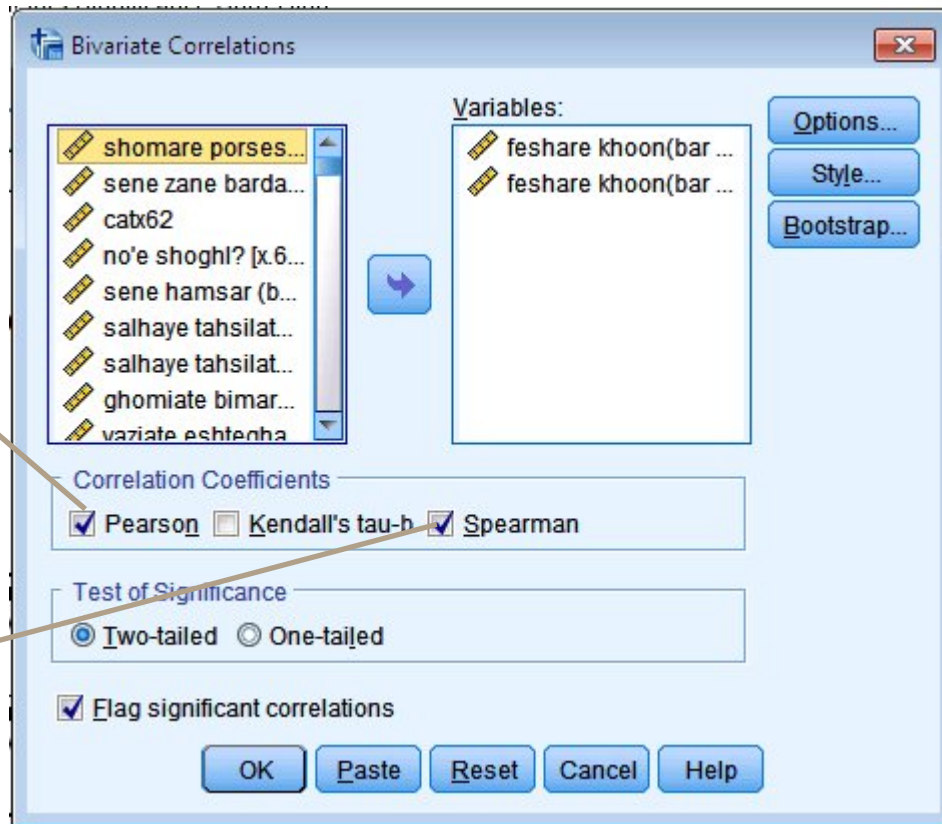
آیا بین فشار خون سیستولیک و دیاستولیک در ۳ ماه اول بارداری همبستگی وجود دارد؟

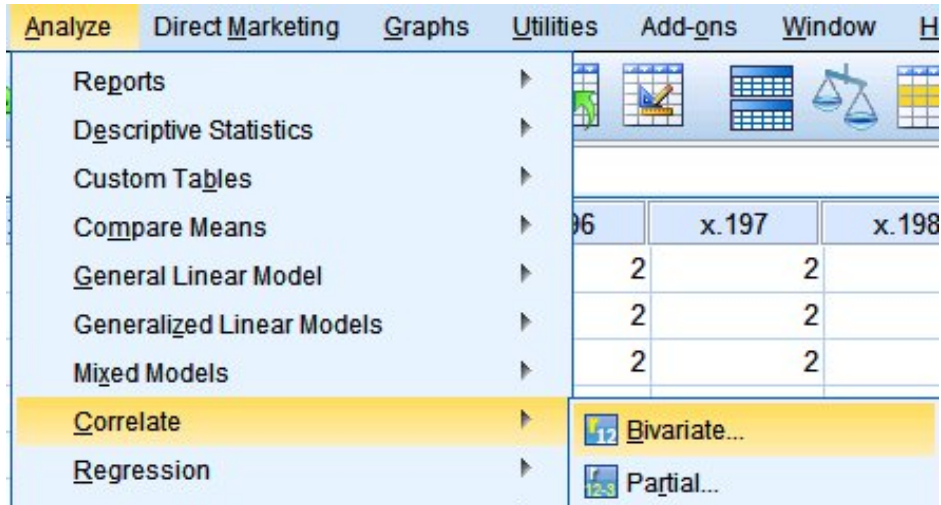


# همبستگی

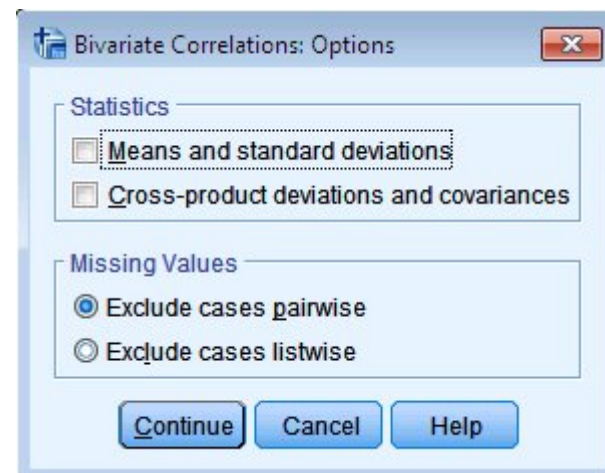
زمانی که داده ها کمی نرمال باشند

زمانی که داده ها کمی غیر نرمال،  
رتبه ایی یا فاصله ایی باشند





# همبستگی



## Correlations

خروجی

Correlations

		feshare khoon(bar hasbe mm jive, systol?)	feshare khoon(bar hasbe mm jive, diastol?)
feshare khoon(bar hasbe mm jive, systol?)	Pearson Correlation	1	.565**
	Sig. (2-tailed)		.000
	N	156	156
feshare khoon(bar hasbe mm jive, diastol?)	Pearson Correlation	.565**	1
	Sig. (2-tailed)	.000	
	N	156	156

میزان همبستگی بین دو متغیر

\*\* . Correlation is significant at the 0.01 level (2-tailed).

$$\begin{cases} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{cases}$$

## Nonparametric Correlations

Correlations

		feshare khoon(bar hasbe mm jive, systol?)	feshare khoon(bar hasbe mm jive, diastol?)
Spearman's rho	feshare khoon(bar hasbe mm jive, systol?)	Correlation Coefficient	1.000
		Sig. (2-tailed)	.000
		N	156
feshare khoon(bar hasbe mm jive, diastol?)	feshare khoon(bar hasbe mm jive, diastol?)	Correlation Coefficient	.586**
		Sig. (2-tailed)	.000
		N	156

\*\* . Correlation is significant at the 0.01 level (2-tailed).

## نکته:

- این آزمون ها میزان همبستگی **خطی** را میسنجند. یعنی اگر فرض صفر رد نشد، تنها میتوان گفت همبستگی خطی بین دو متغیر وجود ندارد. احتمال دارد همبستگی از درجات دیگر وجود داشته باشد.
- میزان همبستگی نشان دهنده میزان تغییر یک متغیر به ازای تغییرات متغیر دیگر است.
- میزان همبستگی بین یک متغیر کمی و کیفی مفهوم ندارد. برای آزمون همبستگی در این حالت از رگرسیون یا آزمون های مقایسه میانگین استفاده میشود.

میزان همبستگی بین وزن مادر در ۳ ماه اول و وزن مادر در ۳ ماه دوم را بدست آورید.  
(متغیرهای x.104 و x.132)

آیا بین افزایش وزن مادر در بارداری (به ۳ گروه کم، خوب، زیاد تقسیم کنید) و وزن کودک (کم وزن، نرمال، چاق) همبستگی وجود دارد؟ (متغیر وزن مادر در ۳ ماه اول = x.104، متغیر وزن مادر در ۳ ماه سوم = x.206 و متغیر وزن کودک = x.230.2)



# آزمون سنجیدن ارتباط بین دو متغیر کیفی



## $\chi^2$ test

- به این آزمون، آزمون استقلال و با مناسبت مدل میگویند.
- زمانی که هر دو متغیر کیفی باشند برای سنجش ارتباط بین آنها از این آزمون استفاده میشود.

$$\begin{cases} H_0: \text{استقلال (عدم وجود ارتباط)} \\ H_1: \text{عدم استقلال (وجود ارتباط)} \end{cases}$$

- پیش فرض ها:
- مقدار مورد انتظار برای هیچ طبقه ای نباید کمتر از ۵ باشد.
- داده ها مستقل باشند.





# $\chi^2$ test

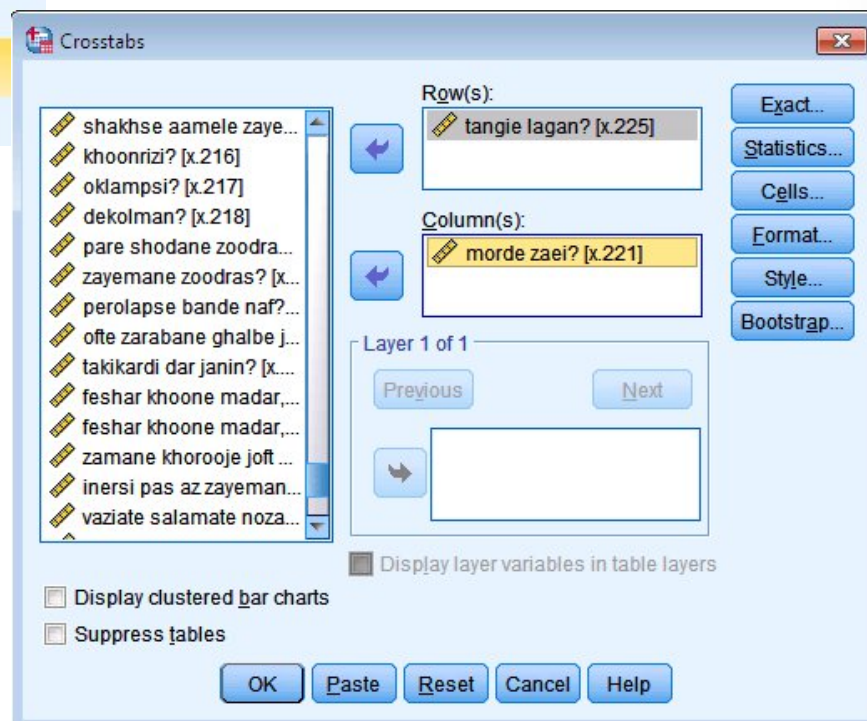
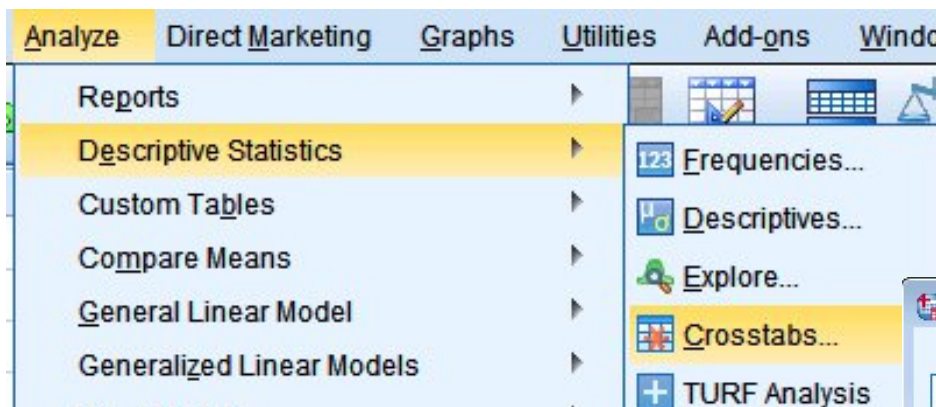
• مثال:

آیا بین سطح تحصیلات و دیابتی بودن یا نبودن ارتباط وجود دارد؟  
آیا بین استعمال دخانیات و سقط ارتباط وجود دارد؟  
آیا بین مصرف دخانیات و جنس ارتباط وجود دارد؟

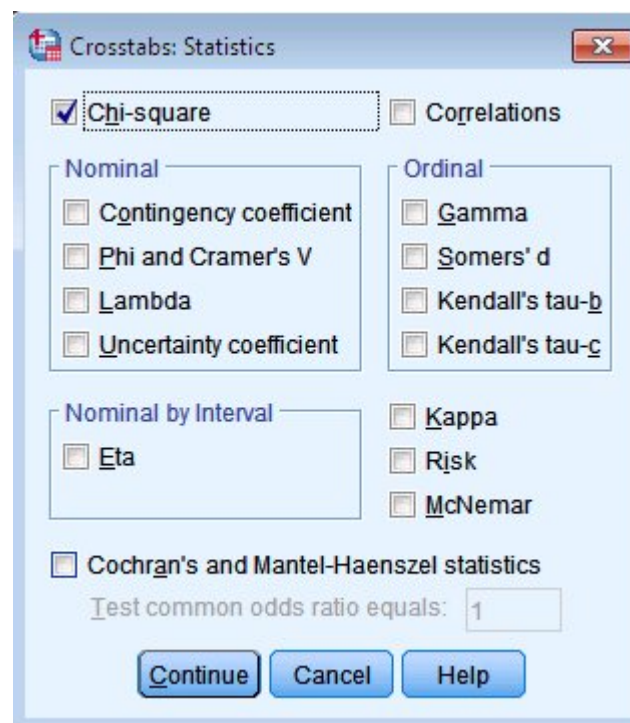
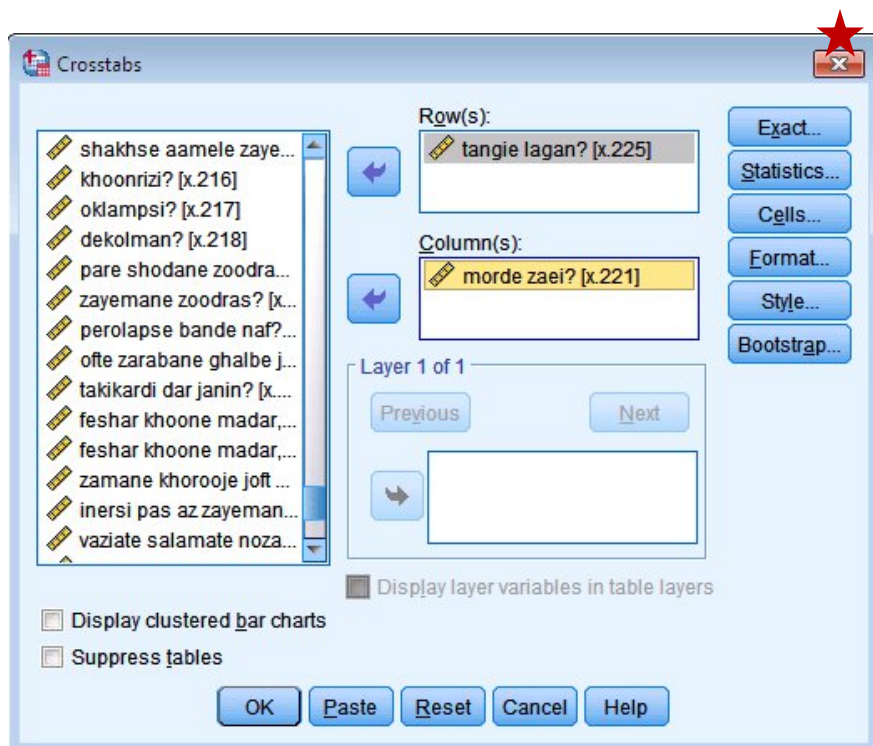


# $\chi^2$ test

- آیا بین تنگی لگن و مرده زایی ارتباط وجود دارد؟



# $\chi^2$ test

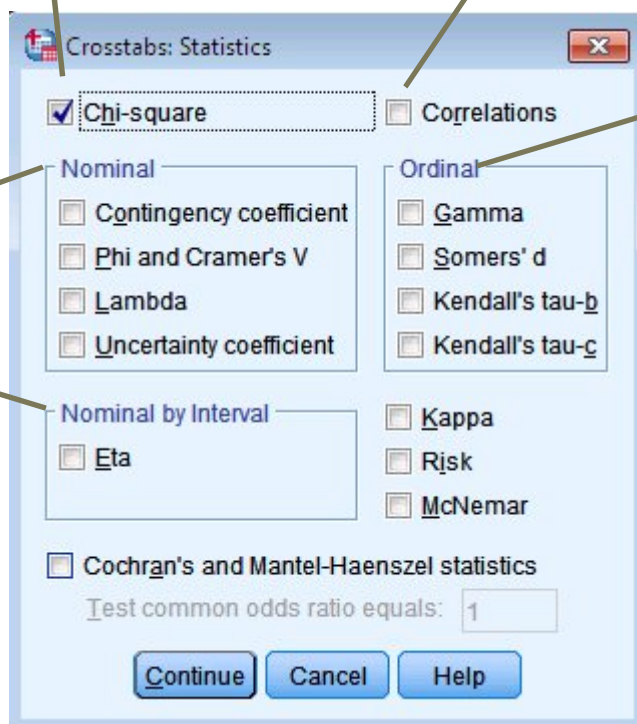


# $\chi^2$ test

آزمون استقلال یا  
ارتباط را میسجد

زمانی که دو متغیر اسمی  
باشند گزینه ای از این کادر  
انتخاب میکنیم

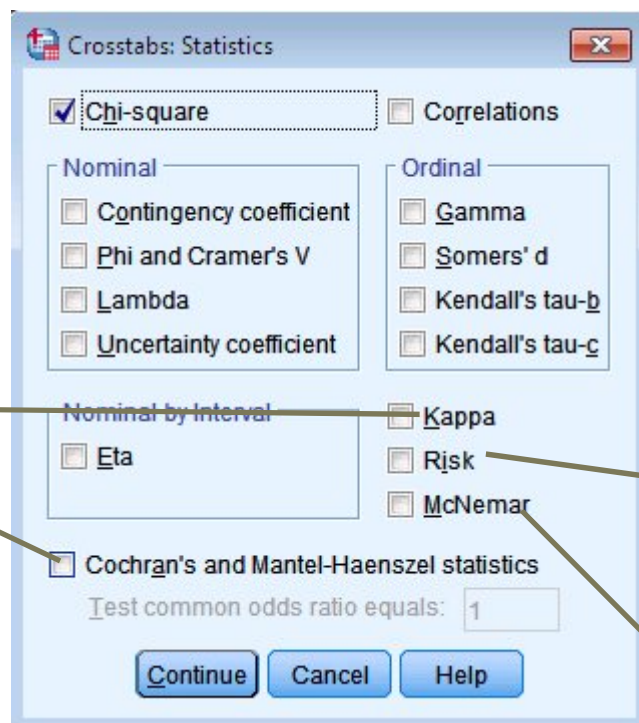
زمانی که یکی از متغیرها  
اسمی و دیگری رتبه ایی  
باشد گزینه ای از این کادر  
انتخاب میکنیم



ضریب همبستگی اسپیرمن را  
زمانی که متغیرهای انتخاب شده  
رتبه ایی باشند را محاسبه میکند.

زمانی که دو متغیر رتبه  
ایی باشند گزینه ای از  
این کادر انتخاب میکنیم

# $\chi^2$ test



ضریب توافق دو اندازه گیری را میسجد و مقداری بین ۰ و ۱ است

خطر را در یک جدول ۲\*۲ محاسبه میکند. در واقع احتمال رخ دادن پیامد را در حضور یک عامل میسجد

آماره منتل هنزل را برای سنجیدن استقلال یک متغیر اسمی دو حالتی و یک پاسخ اسمی دو حالتی محاسبه میکند

آزمون ناپارامتری برای داده های اسمی زوجی میباشد. فرض برابر فراوانی سطرو ستون را میسجد

tangie lagan? \* morde zaei? Crosstabulation

Count

		morde zaei?		Total
		bale	kheir	
tangie lagan?	bale	0	30	30
	kheir	2	128	130
Total		2	158	160

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.467 <sup>a</sup>	1	.494		
Continuity Correction <sup>b</sup>	.000	1	1.000		
Likelihood Ratio	.836	1	.360		
Fisher's Exact Test				1.000	.659
Linear-by-Linear Association	.464	1	.496		
N of Valid Cases	160				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is .38.

b. Computed only for a 2x2 table

در جداول ۲\*۲ که پیش فرض برقرار نباشد اهمیت دارد. در غیر این صورت sig کای دو گزارش میشود.

آزمون روند: برای آزمون روند داشتن بین یک متغیر اسمی دو حالتی و یک متغیر چند سطحی رتبه ای استفاده میشود.

## سوال

- آیا بین نوع زایمان و تنگی لگن ارتباط وجود دارد؟ (متغیرهای ۲۱۲ و ۲۲۵)
- آیا مبتلا بودن مادر به دیابت بارداری در ۳ ماه آخر بر وزن نوزاد اثر گذار است؟  
(متغیرهای ۲۳۰/۲ و ۱۹۹)
- آیا روندی بین تغییرات وزن نوزاد با توجه به مبتلا بودن مادر به دیابت بارداری وجود دارد؟
- آیا مبتلا بودن مادر به فشار خون بارداری بر مرده زایی موثر است؟ (متغیرهای ۱۹۶ و ۲۲۱)



# آزمون مقایسات زوجی



One Sample t-test

داده ها مستقل X یک سطحی Y کمی نرمال

Independent Samples t-test

داده ها مستقل X دو سطحی Y کمی نرمال

One way ANOVA

داده ها مستقل X چند سطحی Y کمی نرمال

Two way ANOVA

داده ها مستقل X1 & X2 چند سطحی Y کمی نرمال

ANCOVA

داده ها مستقل X1 & X2 چند سطحی و X3 کمی Y کمی نرمال

آزمون های ناپارامتری

داده ها مستقل در سطوح مختلف X Y کمی غیر نرمال یا کیفی

Paired t- test

داده ها همبسته X دو سطحی Y کمی نرمال

Repeated Measures Analysis

داده ها همبسته X چند سطحی Y کمی نرمال

GEE

داده ها همبسته در سطوح مختلف X Y کمی غیر نرمال یا کیفی



## Paired t-test

- زمانی استفاده میشود که دو اندازه گیری به گونه ایی همبسته باشند.
- پیش فرض ها:  
متغیرهای وابسته کمی و اختلاف آنها نرمال باشد.
- متغیر مستقل دو سطحی باشد که میتواند نشان دهنده زمان اندازه گیری و یا معرف match ها باشد.

- مثال  
فشار خون قبل از دارو و بعد از دارو  
قدرت بینایی چشم چپ و راست  
مقایسه قند خون دو گروه که با توجه به میزان فعالیت ورزشی همسان شده اند.

# Paired t-test

- برای انجام این آنالیز نحوه وارد کردن داده ها مهم است.

اندازه گیری ها در زمان های مختلف یا شرایط مختلف باید در ستون های متفاوت وارد شوند.

	y1	y2
1	12	54
2	12	34
3	34	65
4	65	75
5	87	45
6	90	67
7	45	45
8	76	34
9	89	70
10	76	65
11		

در آنالیزهای داده های مستقل یک متغیر مستقل معرف گروه میباشد و داده های گروه های مختلف دنبال هم وارد میشوند.

	y1	x
1	12	1
2	12	1
3	34	1
4	65	1
5	87	1
6	90	2
7	45	2
8	76	2
9	89	2
10	76	2
11		



## Paired t-test

مثال:

آیا اختلاف معناداری بین فشار خون سیستولیک در ۳ ماه اول بارداری و فشار خون سیستولیک در ۳ ماه دوم بارداری وجود دارد؟

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

# Paired t-test

• بررسی پیش فرض های آزمون

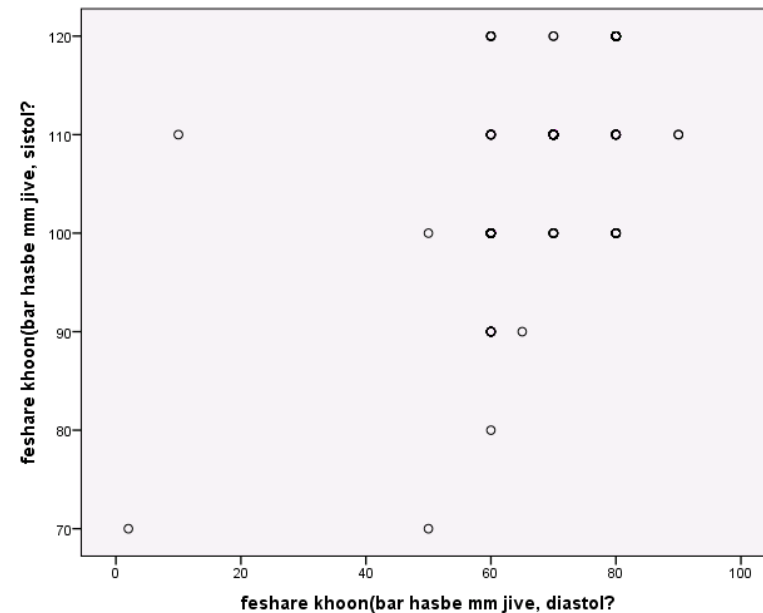
بررسی نرمالیتی اختلاف دو اندازه گیری

بررسی وجود داده پرت

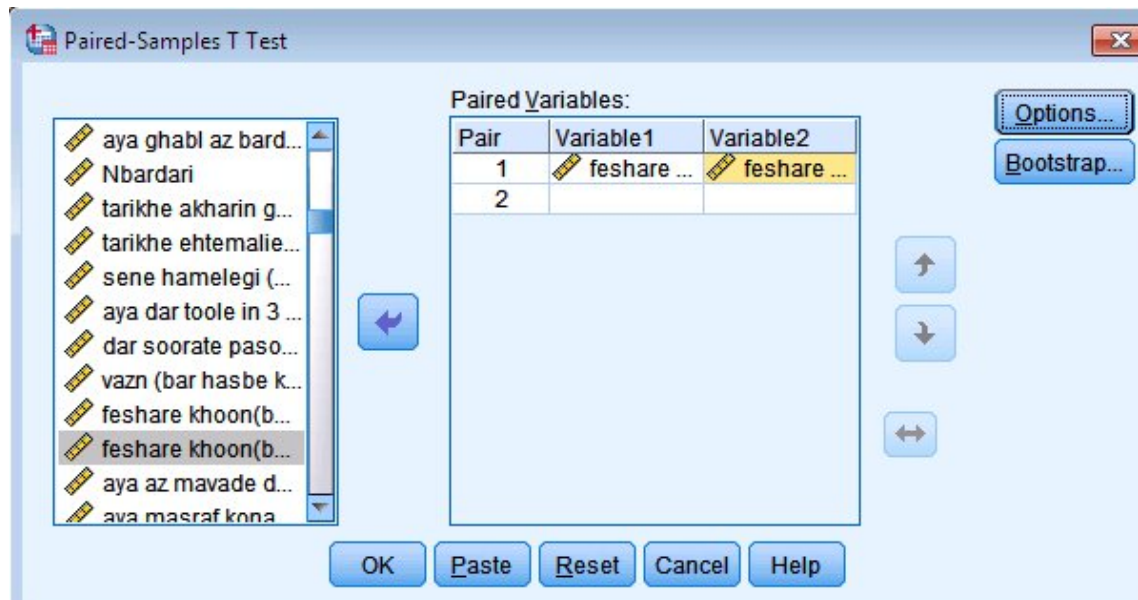
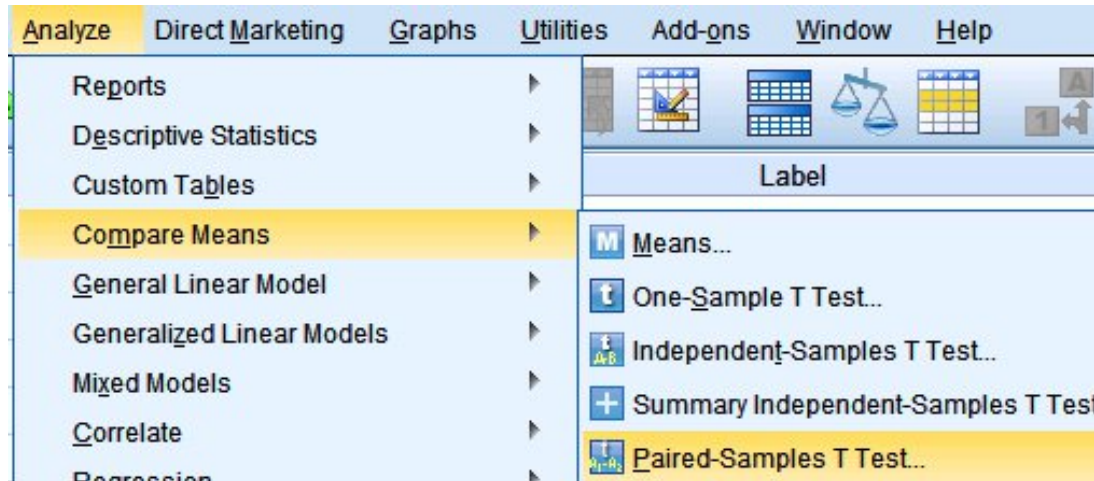
## One-Sample Kolmogorov-Smirnov Test

		diff
N		155
Normal Parameters <sup>a,b</sup>	Mean	38.6645
	Std. Deviation	9.84706
Most Extreme Differences	Absolute	.323
	Positive	.323
	Negative	-.302
Test Statistic		.323
Asymp. Sig. (2-tailed)		.000 <sup>c</sup>

- Test distribution is Normal.
- Calculated from data.
- Lilliefors Significance Correction.



# Paired t-test



**Paired Samples Statistics**

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 feshare khoon(bar hasbe mm jive, sistol?	107.40	154	9.756	.786
feshar khoon(bar hasbe mm jive) sistol?	116.27	154	64.234	5.176

ضریب همبستگی منفی  
و ضعیف است

**Paired Samples Correlations**

	N	Correlation	Sig.
Pair 1 feshare khoon(bar hasbe mm jive, sistol? & feshar khoon(bar hasbe mm jive) sistol?	154	-.029	.725

فرض عدم وجود  
همبستگی رد نمیشود.

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

**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 feshare khoon(bar hasbe mm jive, sistol? - feshar khoon(bar hasbe mm jive) sistol?	-8.864	65.246	5.258	-19.251	1.523	-1.686	153	.094

# سوال

- آیا اختلاف معناداری بین فشار خون دیاستولیک در ۳ ماه اول بارداری و فشار خون دیاستولیک در ۳ ماه دوم بارداری وجود دارد؟ (متغیرهای ۱۰۵ و ۱۳۴)