

دانشگاه علوم پزشکی و خدمات بهداشتی،  
درمانی کرمانشاه  
دانشکده پرستاری سنقر



## آموزش مقدماتی نرم افزار SPSS ۱۱/۵

حبیب الله اسماعیلی استادیار آمارزیستی دانشگاه علوم پزشکی مشهد  
سلیمان خیری استادیار آمارزیستی دانشگاه علوم پزشکی شهرکرد

## کارگاه مقدماتی

### محاسبات آماری بوسیله SPSS

#### مقدمه

SPSS نرم افزار تجزیه و تحلیل آماری و مدیریت داده ها می باشد. این نرم افزار در ابتدا برای پژوهشگران علوم انسانی ابداع شد (SPSS مخفف Statistical Package for the Social Sciences است). در حال حاضر با گسترش چشمگیر SPSS، این نرم افزار یکی از کاربردی ترین نرم افزارهای آماری شده است. گرچه قابلیت های تجزیه و تحلیل اطلاعات در SPSS در بسیاری از جهات از نرم افزارهای قدرتمند آماری همچون SAS و S-Plus کمتر می باشد اما سادگی نصب و تسهیلاتی که برای ورود اطلاعات و تجزیه و تحلیل آنها در SPSS وجود دارد باعث استفاده بیشتر این نرم افزار در عرصه کاربرد شده است.

هدف اصلی یک نرم افزار آماری، بکارگیری روش های آماری برای تجزیه و تحلیل اطلاعات می باشد. در کنار این هدف مهم، تولید انواع نمودار به عنوان ابزاری برای انتقال آسان و سریع اطلاعات در این نرم افزارها مورد توجه است. با این وجود برای رسم نمودارهای آماری مناسب، می توان از نرم افزارهای گزارش گیری همچون Excell که قابلیت ویژه ای در تولید انواع نمودارها دارند استفاده نمود.

#### شروع کار با SPSS

چنانچه نرم افزار SPSS روی ویندوز نصب شده باشد با دوبار کلیک (فشار دگمه سمت چپ موس) روی آیکون (نشانه) SPSS، پنجره ای به نام «ویرایشگر داده ها» (SPSS Data Editor)، ظاهر می گردد. اگر آیکون SPSS روی Desktop وجود نداشته باشد با اجرای مسیر زیر SPSS فراخوانی می شود.

Start→Programs→SPSS for Windows→SPSS11.5 for Windows

ویرایشگر داده ها (SPSS Data Editor)، پنجره اصلی در SPSS بوده که کلیه عملیات ورود اطلاعات و اجرای فرامین مربوط به کارهای آماری و تولید نمودار در آن انجام می گیرد. علاوه بر این پنجره، پنجره های "خروجی" (Output) "دستوری" (Syntax) و "ویرایشگر نمودار" (Chart Editor) در SPSS وجود دارند. نتیجه اجرای کلیه فرامین در پنجره خروجی ظاهر می شود.

#### پنجره ویرایشگر داده ها (SPSS Data Editor)

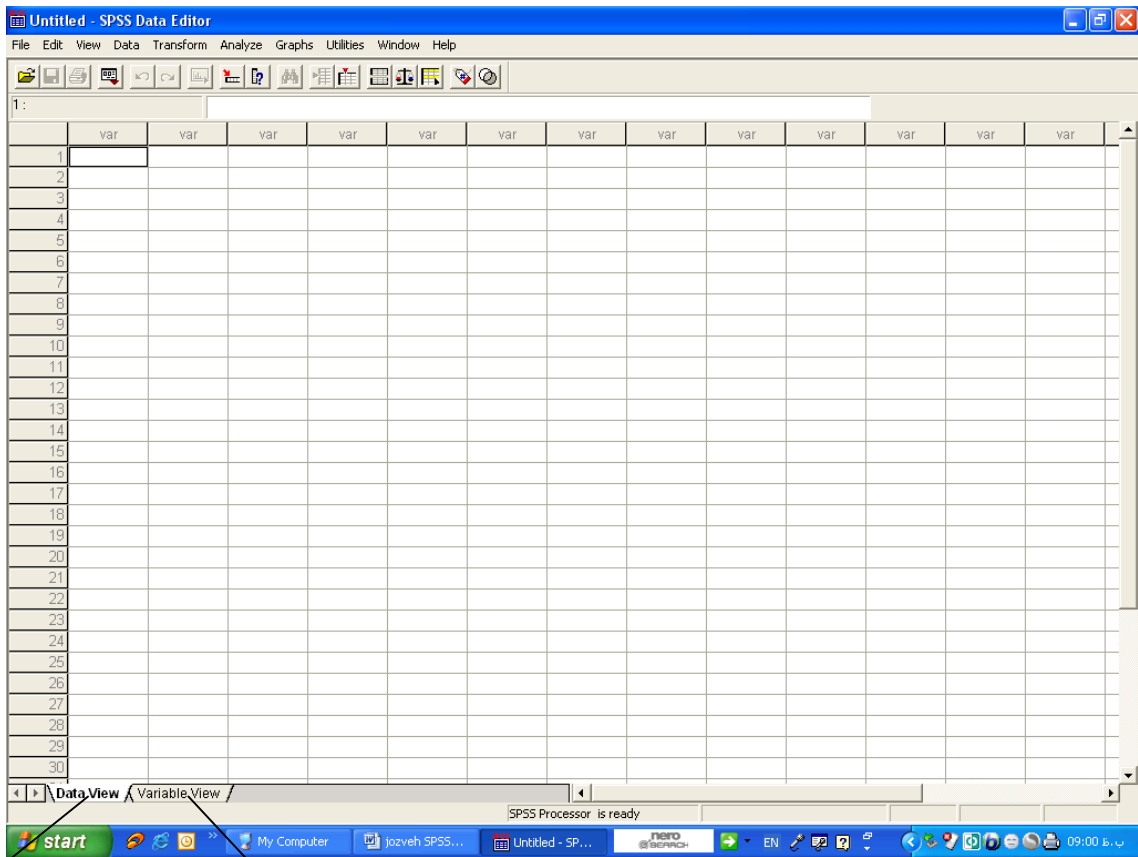
در نوار دستور واقع در قسمت بالای پنجره ویرایشگر داده ها، منوهای File، Edit، View، Data، Transfrom، Analyz، Graphs، Utilities، Windw، Help قرار دارند. منوهای Data،

Transfrom, Analyz, Graphs، منوهای اختصاصی SPSS می باشند. دیگر منوها که در سایر نرم افزارهای تحت ویندوز نیز قابل مشاهده هستند را منوهای عمومی گویند.

- در منوی File، فرمان های مربوط به ایجاد فایل جدید داده ها، فراخوانی یک فایل داده اطلاعات یا خروجی، ذخیره، چاپ و فرمان مربوطه ورود اطلاعات از سایر نرم افزارهای آماری و گزارش گیری وجود دارد.
- در منوی Edit، فرمان های مربوط به ویرایش اطلاعات (از جمله کپی، جایگزینی و نمایش یک داده خاص) وجود دارد.
- در منوی View، فرمان های مربوط به نمایش اطلاعات وجود دارد.
- در منوی Data، می توان از تسهیلاتی که SPSS برای ورود اطلاعات در اختیار کاربر قرار می دهد استفاده نمود.
- در منوی Transform، فرمان های مربوط به تغییر مقدار یا مقادیر یک متغیر، محاسبه متغیر جدید بر اساس تابعی از متغیرهای موجود، کدبندی متغیر کمکی و یا تغییر کد متغیر کیفی وجود دارد.
- در منوی Analyze، فرمان های لازم برای اجرای کلیه کارهای آماری وجود دارد.
- در منوی Graphs، فرمان های مربوط به تولید نمودارهای گوناگون از جمله انواع نمودارهای ستونی، دایره ای، هیستوگرام، جعبه ای و پراکنش وجود دارد.

### تعریف متغیرها و ورود اطلاعات

پنجره ویرایشگر داده ها شامل دو بخش به نام های "نمایش داده ها" (Data view) و "نمایش متغیرها" (Variable view) می باشد که در هر زمان تنها یکی از آنها در معرض نمایش است. بخش نمایش داده ها، صفحه گسترده ای است که جهت ورود اطلاعات بکار می رود. (صفحه قبل) در بخش نمایش داده ها، مشاهدات هر مورد (هر آزمودنی) در یک سطر و مشاهدات هر متغیر در یک ستون قرار می گیرد. نام متغیرها در سر ستون ها و شماره آزمودنی ها در اولین ستون سمت چپ قرار دارند. با وجود این بهتر است در هر تحلیلی، ستون اول نیز به شماره پرسشنامه یا شماره آزمودنی اختصاص یابد. این ستون معمولا با نام ID نام گذاری می گردد. در بخش نمایش متغیرها، (Variable view) کارهای مربوط به نام گذاری و تعیین مشخصات متغیرها انجام می گیرد. (شکل ذیل)

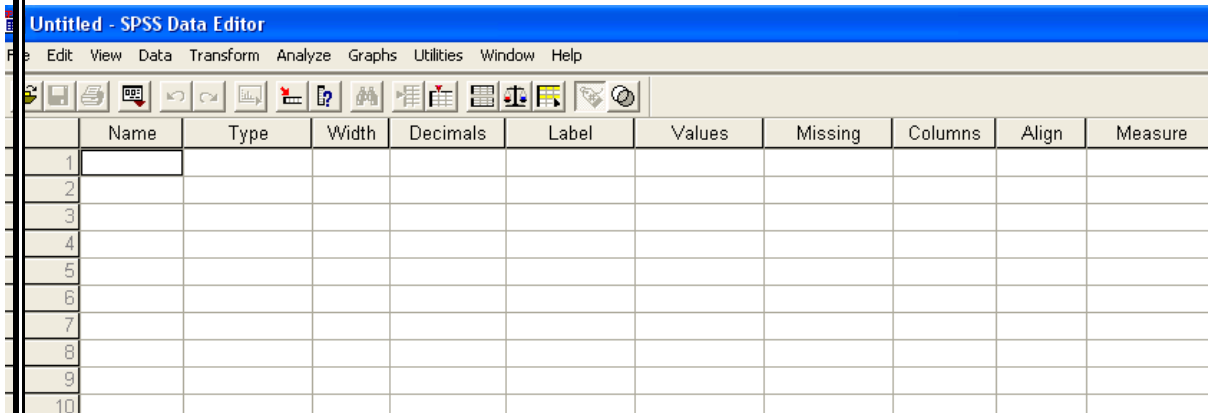


Data View

Variable View

## نمایش متغیرها (Variable View)

با کلیک کردن روی Variable View به صفحه ای می رویم که در آن صفحه، در ستون اول شماره متغیر و در ستون های بعدی، نام و مشخصات مربوط به هر متغیر وارد می گردد.



- در قسمت Name، نام متغیر وارد می شود. توجه کنید که نام متغیر حداکثر از ۸ کاراکتر (حرف) تشکیل شده و نمی توان در نام متغیر از فضای خالی استفاده کرد. بعضی از کارکترهای موجود در صفحه کلید (مثل ؟ و ! ) قابل استفاده در نام متغیر نیستند. نام متغیر در SPSS با حروف کوچک ثبت می شوند.

- در قسمت Type، لازم است نوع متغیر معلوم شود. پیش فرض SPSS این است که متغیر از نوع عددی (Numeric) است. با کلیک در قسمت Type منوی شامل پیشنهادهای دیگر برای نوع متغیر ظاهر می شود. در حالت کلی چنانچه متغیر، نام آزمودنیها (اسامی افراد یا آزمودنی ها) باشد نوع متغیر رشته ای (String) بوده و در سایر موارد می توان نوع عددی را به کار برد. با این حال گزینه های دیگری که در عمل کاربرد کمی دارند پیشنهاد شده است. این گزینه ها عبارتند از گزینه Common، که برای نمایش اعداد چند رقمی که هر سه رقم آنها با یک کاما جدا شود بکار می رود. گزینه Dot، که برای نمایش اعداد چند رقمی که هر سه رقم آنها با یک نقطه جدا شود استفاده می شود. نماد علمی (Scientific notation)، که در این صورت رقم به صورت مضربی از ۱۰ به نمایش در می آید. تاریخی (Date)، که برای نمای اطلاعات زمانی بکار می رود. دلاری (Dollar)، که برای نمایش اطلاعات هزینه ای بر حسب دلار بکار می رود. سفارشی (Custom Carrancy)، که برای ثبت اطلاعات خاص بکار می رود. رشته ای (String)، که برای ثبت اطلاعات کاراکتری همچون اسامی بکار می رود.

- در بخش اعشاری (Decimal)، تعداد رقم های اعشار متغیر عددی مشخص می گردد. پیش فرض SPSS عدد دو می باشد.

- بخش برچسب (Lable)، چنانچه بخواهیم نام کلی متغیر یا توضیحی در مورد آن ثبت شود، با کلیک در قسمت Lable این کار انجام می گیرد. توجه شود که در صورت درج اطلاعاتی در بخش برچسب، این اطلاعات در خروجی به عنوان نام متغیر ظاهر می شود.

- در بخش مقدار (Value)، برچسب کدهای عددی بکار گرفته شده برای متغیرهای کیفی یا اسمی تعریف می گردند. با کلیک در این قسمت اسامی کدها را در قسمت Value Lable، تعریف می کنیم در این صورت در خروجی، برچسب های متغیر ها به جای کدهای عددی گزارش می شوند.

- گمشده (Missing)، اگر لازم باشد مقدار یا مقادیر خاصی از متغیر مورد بررسی به عنوان داده گمشده در نظر گرفته شود آن مقادیر در این قسمت پیشنهاد می شوند. برای انجام این کار با کلیک کردن در قسمت گمشده، منوی ظاهر می شود که می توان با فعال نمودن گزینه Discrete missing سه مقدار خاص و با فعال نمودن گزینه Range of missing value دامنه ای از مقادیر (برای متغیرهای پیوسته) و با فعال نمودن گزینه Range pluse one discrete missing value (برای متغیرهای پیوسته) از دو انتخاب فوق را به عنوان داده گمشده پیشنهاد داد. توجه کنید که مقادیر گمشده در تجزیه و تحلیل اطلاعات مورد استفاده قرار نمی گیرند. علاوه بر این در هنگام ورود اطلاعات، چنانچه مقدار یک متغیر ثبت نشود سیستم SPSS آن را به نقطه نشان خواهد داد که به این نوع گمشده، گمشده سیستمی (System Missing) گویند.

- در بخش ستون (Column)، پهنای هر ستون مشخص می گردد.

- با کلیک در بخش هم ترازی (Alignment) می توان نحوه قرار گرفتن اطلاعات در هر ستون را که می تواند راست، مرکز یا سمت چپ هر ستون باشد پیشنهاد نمود.

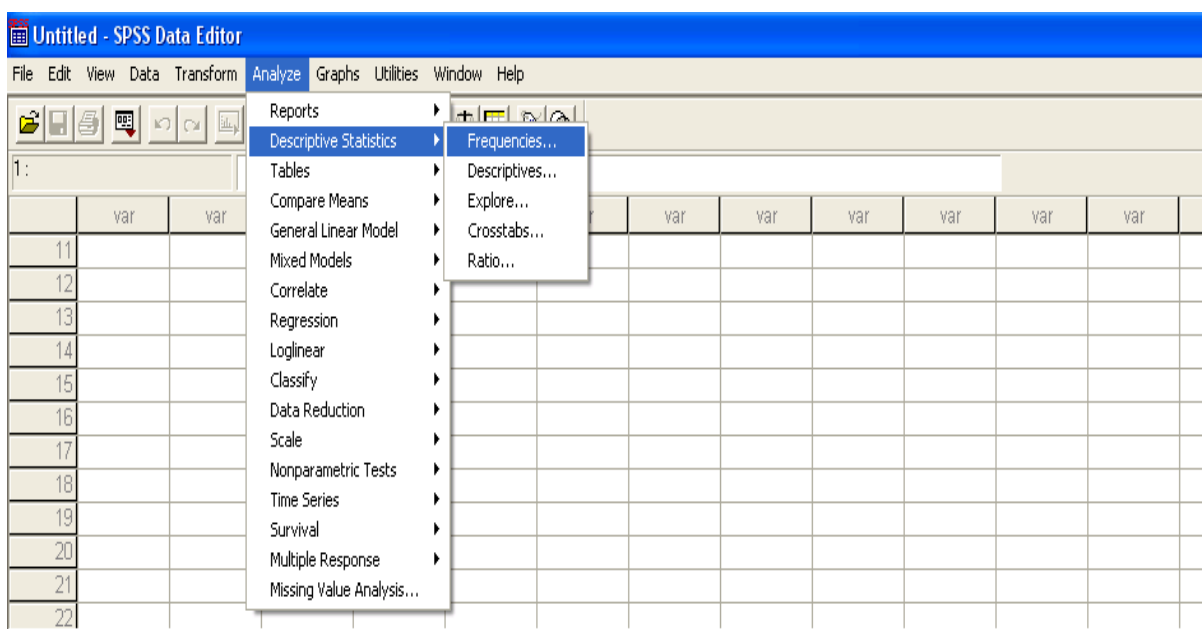
- با کلیک در بخش اندازه گیری (Measure)، می توان مقیاس اندازه گیری متغیر را پیشنهاد داد. پیش فرض SPSS، نسبی (Scale) بودن مقیاس اندازه گیری متغیرها است. علاوه بر این می توان مقیاس های رتبه ای (Ordinal) و اسمی (Nominal) را به ترتیب برای متغیرهای کیفی رتبه ای و اسمی پیشنهاد نمود. توصیه می شود مقیاس ها همان نسبی (Scale) باشد و تغییر ندهید.

### توصیف داده ها

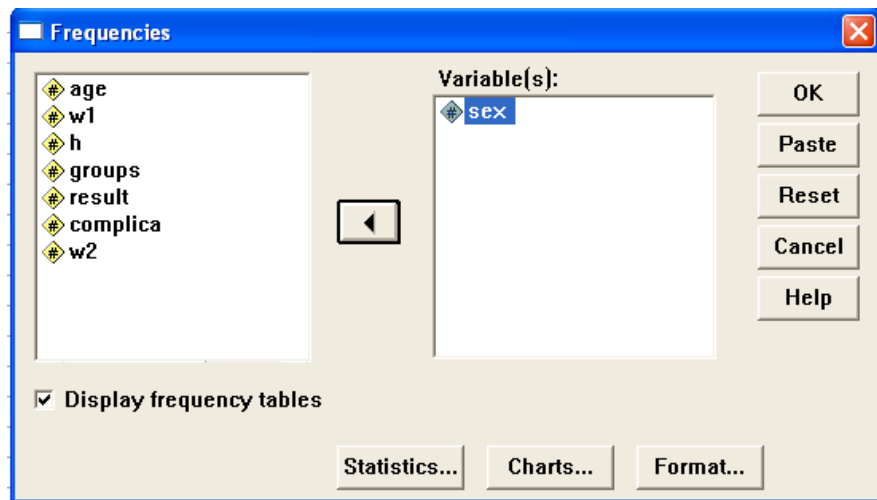
#### جدول فراوانی:

برای توصیف متغیرهای کیفی (مثل جنس، شغل، تحصیلات، شدت بیماری و ...) می توان جدول توزیع فراوانی هر کدام را که شامل فراوانی (Frequency)، درصد فراوانی نسبی (Percent)، درصد فراوانی نسبی معتبر (Valid percent)، و فراوانی تجمعی (Cumulative percent)، می باشد را محاسبه نمود.

برای انجام این کار در SPSS لازم است مسیر زیر را انتخاب نمود.



با این کار منوی Frequency ظاهر گشته پس از انتخاب متغیرهای مورد نظر از لیست متغیرها OK را کلیک می نمائیم.



خروجی به صورت ذیل می باشد.

Output1 - SPSS Viewer

File Edit View Insert Format Analyze Graphs Utilities Window Help

Output

- Descriptives
- Notes
- Frequencies
- Title
- Notes
- Statistics
- SEX

### Frequencies

Statistics

SEX

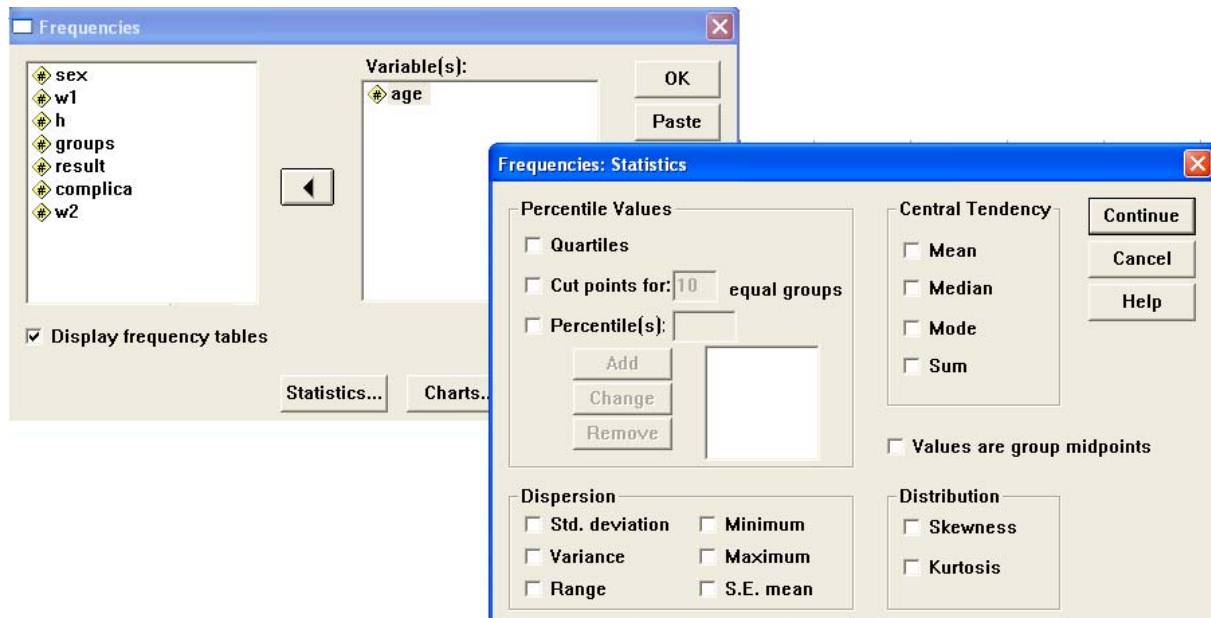
N	Valid	40
	Missing	0

SEX

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid male	20	50.0	50.0	50.0
female	20	50.0	50.0	100.0
Total	40	100.0	100.0	

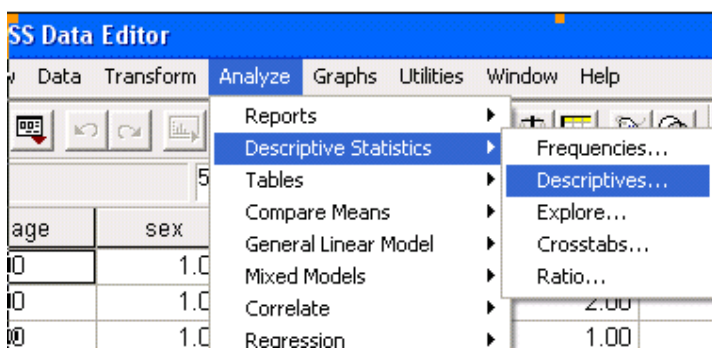
تذکر: گاهی اوقات برخی مقادیر برای یک متغیر نامعلوم است یا اصطلاحاً دارای مقادیر گمشده (Missing Value) می باشیم، در این صورت از (Valid percent) برای تعیین درصد فراوانی استفاده می کنیم که درصد موارد معلوم را محاسبه می کند.

اگر جدول فراوانی ساده برای یک متغیر کمی نیز لازم باشد می توان از مسیر فوق استفاده نمود و با انتخاب گزینه Statistics در این منو شاخص های آماری همچون میانگین، میانه، صدک ها، چندک ها و کشیدگی را نیز بدست آورد. برای محاسبه شاخص ها برای متغیر کمی، از این منو کمتر استفاده می شود.



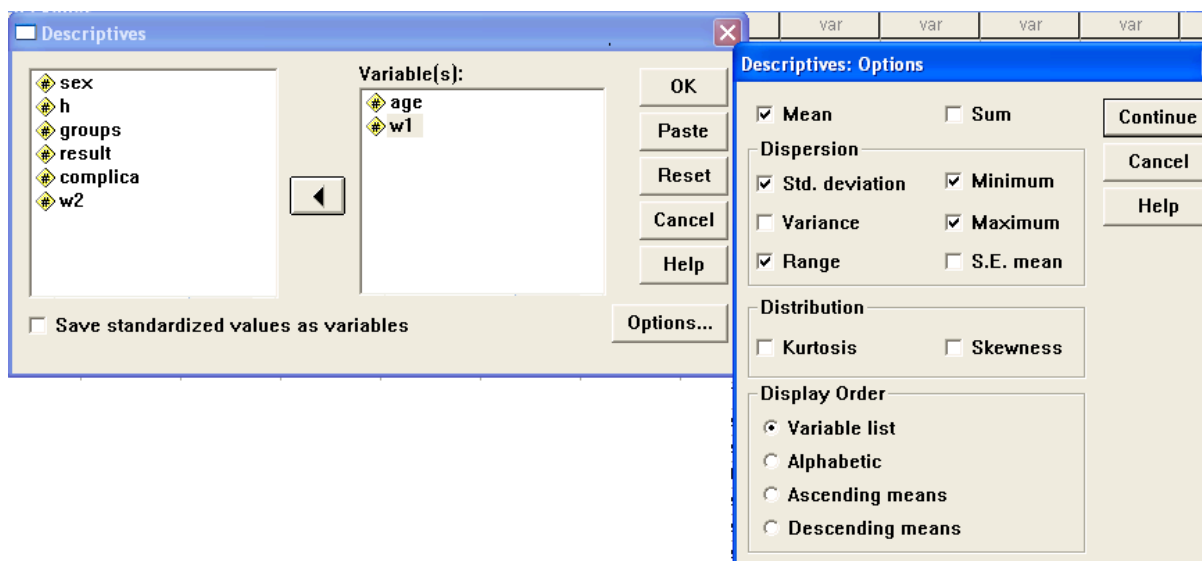
### توصیف متغیرهای کمی

برای متغیرهای کمی (سن، وزن، قد، BMI، فشار خون سیستمولیک و ...) می توان یکی از شاخص های متمایل به مرکز (Central Tendency) یعنی میانگین (Mean)، و شاخص های پراکندگی (Dispersion) شامل: دامنه (Range)، یا واریانس (Variance)، انحراف معیار (Standard Deviation) و خطای معیار (Standard Deviation of Mean) یا (S.E.Mean) را محاسبه کرد. همچنین در این منو می توان شاخص هایی چون حداکثر (Maximum)، حداقل (Minimum)، کشیدگی توزیع (Kurtosis) و چولگی توزیع (Skwness) را بدست آورد. برای محاسبه مقادیر فوق مسیر زیر انتخاب می شود.





با انتخاب مسیر فوق، منوی ظاهر شده و متغیرهای مورد نیاز را از لیست متغیرها انتخاب می‌نمائیم. با انتخاب OK برای متغیر یا متغیرهای انتخاب شده تعداد، حداقل، حداکثر، میانگین و انحراف معیار ظاهر می‌شود برای محاسبه سایر شاخص‌ها از گزینه Option در منوی Descriptive استفاده می‌نمائیم.



پس از کلیک کردن روی continue و سپس OK خروجی به صورت ذیل ظاهر می‌شود. که در آن مقادیر درخواست شده، شامل تعداد، دامنه، حداقل، حداکثر، میانگین و انحراف معیار آمده است.

**Descriptive Statistics**

	N	Range	Minimum	Maximum	Mean	Std. Deviation
AGE	38	32.00	25.00	57.00	40.2632	10.47180
W1	40	32.00	53.00	85.00	68.9750	8.09713
Valid N (listwise)	38					

### رسم نمودار

همانطور که قبلاً گفته شد، نرم افزار SPSS جهت تجزیه و تحلیل داده‌ها بکار برده می‌شود اما می‌توان برخی نمودارها را نیز توسط آن رسم نمود. برای رسم نمودار در منوی Graph انواع نمودارها قرار داده شده است.

## نمودار ستونی (Bar Chart):

یکی از نمودارهایی است که برای متغیرهای کیفی بکار گرفته می شود، برای رسم آن از منوی Graph به قسمت Bar رفته و کلیک می کنیم. پس از define کردن متغیر مورد نظر را از لیست متغیرها انتخاب نموده به قسمت سمت راست می بریم. بهتر است نمودار بر حسب درصد باشد بنابراین عبارت Percent را فعال می کنیم.

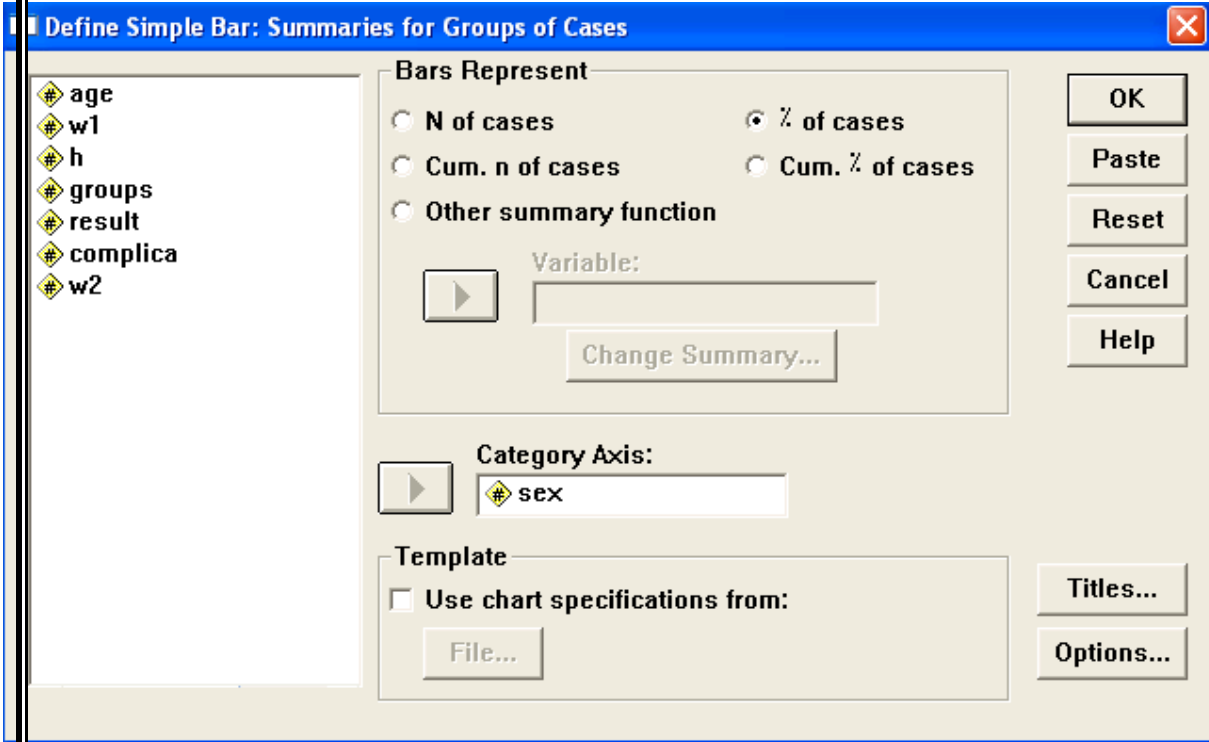
	age	sex	w	groups	result	complica	w2
1	55.00	1.00		1.00	3.00	1.00	58.00
2	30.00	1.00		2.00	2.00	2.00	73.00
3	41.00	1.00		1.00	3.00	2.00	78.00
4	28.00	2.00		1.00	2.00	2.00	64.00
5	44.00	1.00		1.00	3.00	1.00	65.00
6	.	2.00		1.00	2.00	1.00	60.00
7	38.00	2.00		1.00	1.00	2.00	59.00
8	31.00	2.00		2.00	1.00	2.00	62.00
9	29.00	2.00		2.00	1.00	2.00	65.00
10	57.00	1.00		2.00	1.00	2.00	71.00
11	51.00	2.00		2.00	2.00	2.00	66.00
12	29.00	2.00		2.00	2.00	2.00	79.00
13	57.00	1.00		2.00	3.00	2.00	69.00
14	25.00	1.00		2.00	2.00	2.00	70.00
15	49.00	1.00		1.00	1.00	1.00	55.00

**Bar Charts**

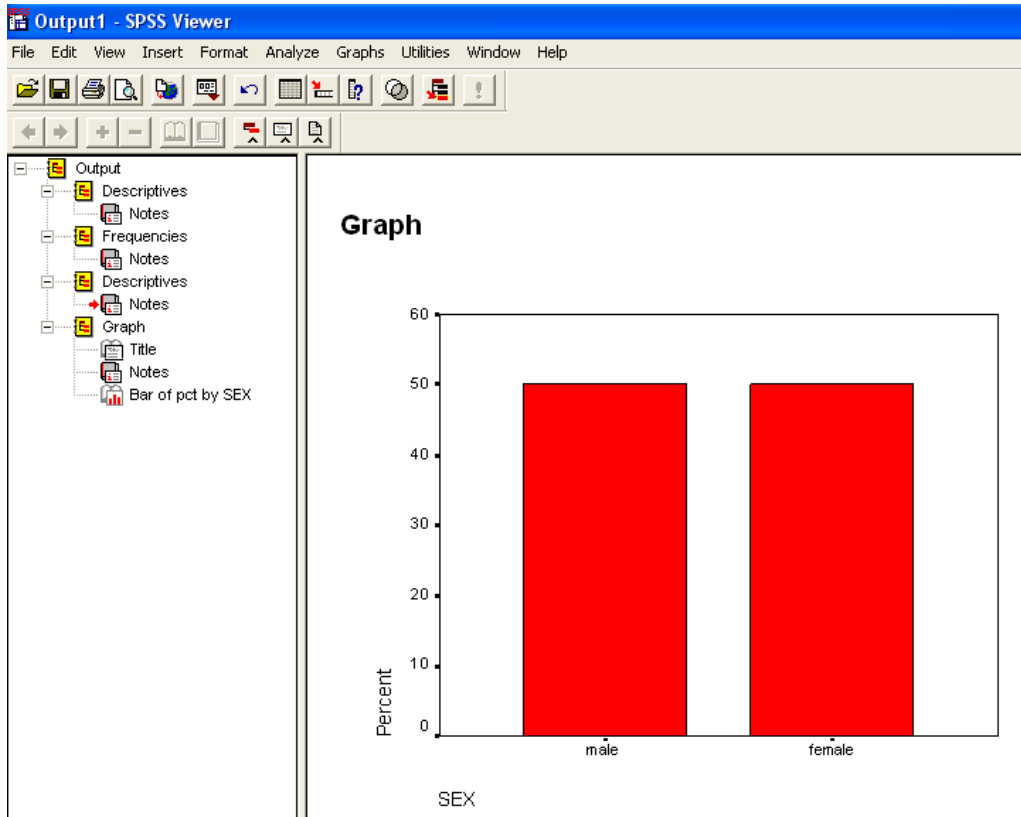
Simple  
 Clustered  
 Stacked

**Data in Chart Are**

- Summaries for groups of cases
- Summaries of separate variables
- Values of individual cases

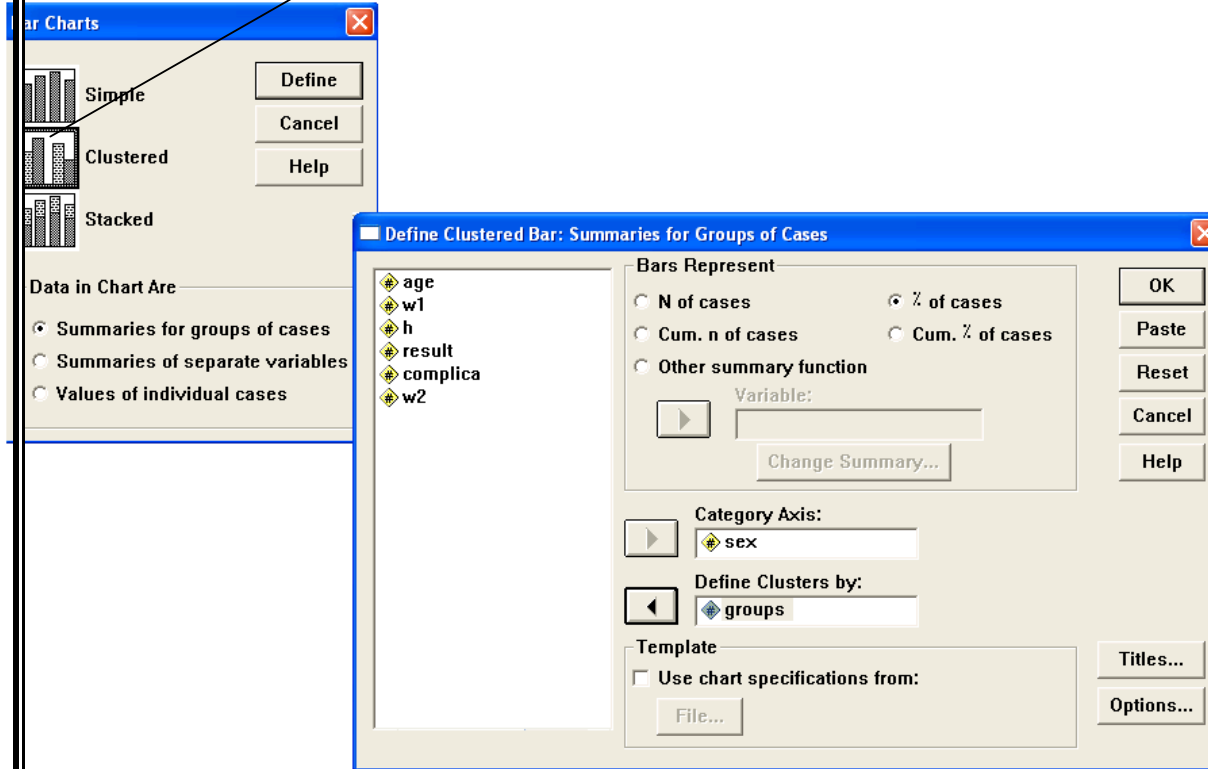


خروجی به صورت ذیل خواهد بود.



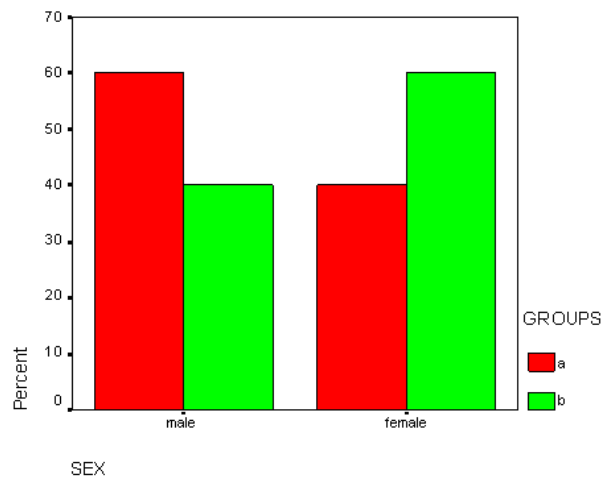
Clustered

در Bar Chart می توان گزینه Cluster را انتخاب و نمودار دو گروه را کنار هم قرار داد تا امکان مقایسه فراهم شود.



خروجی به شکل ذیل می باشد.

Graph

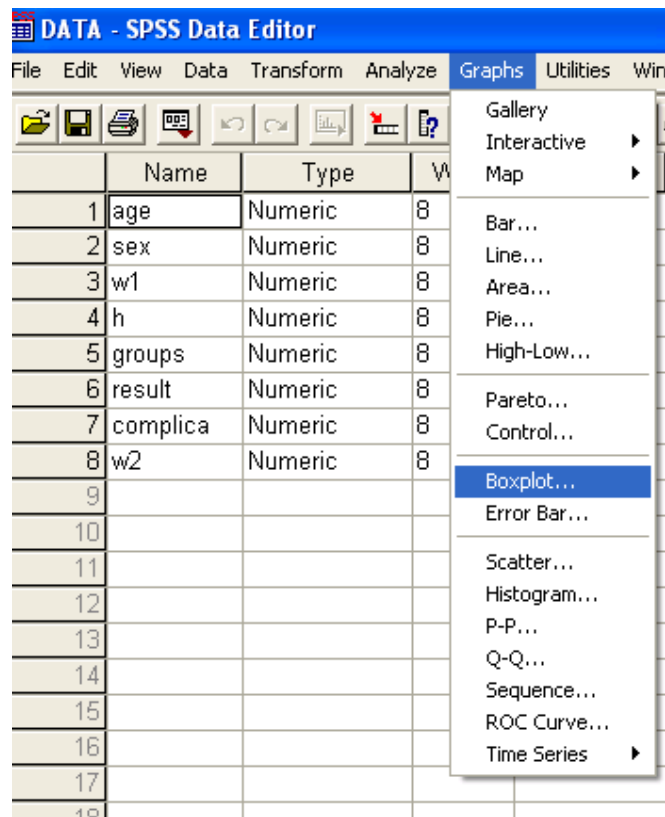


## نمودار دایره ای (Pie Chart):

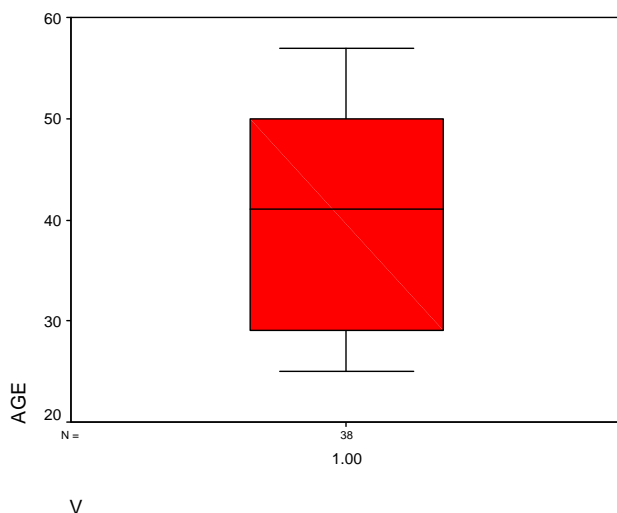
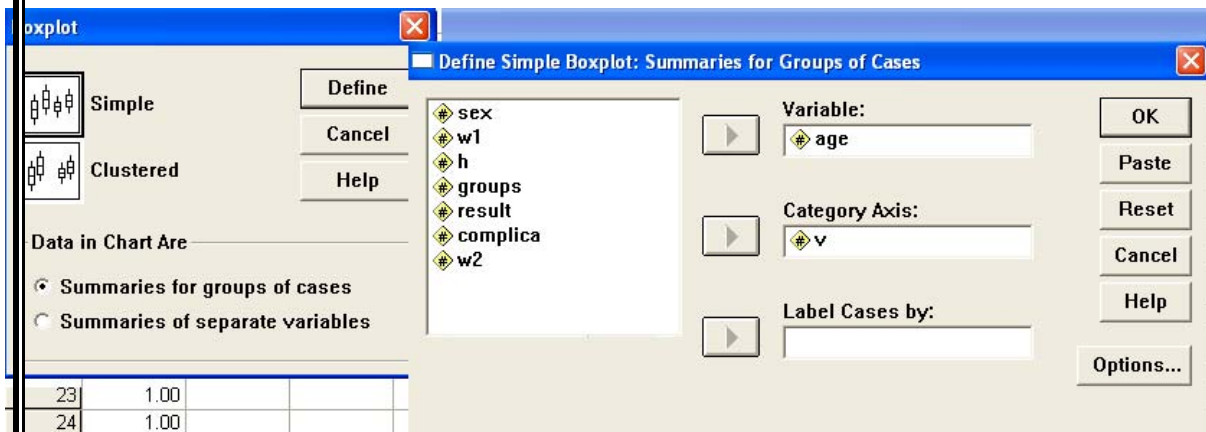
معمولا برای متغیرهای کیفی استفاده می شود. از قسمت Pie به پنجره مربوطه وارد شده و متغیر مورد نظر را انتخاب و Ok می کنیم.

## نمودار جعبه ای (Box plot):

این نمودار اغلب در مواردی رسم می گردد که متغیر دارای توزیع نرمال نیست. در نمودار حداقل، حداکثر، میانه، چارک اول (صدک ۲۵) و چارک سوم (صدک ۷۵) نشان داده می شود.



این نمودار را برای سن افراد رسم کرده ایم. در اینجا V متغیری است کمکی که همه مقادیر آن یک است.



نمودار پراکنش :

این نمودار برای نشان دادن همبستگی بین دو متغیر کمی رسم می شود. برای رسم نمودار پراکنش، گزینه Scatter از منوی Graphs را انتخاب می نماییم. با این کار منوی Scatterplot ظاهر می شود با توجه به اینکه پیش فرض SPSS رسم نمودار پراکنش ساده می باشد گزینه Define را برای رسم نمودار کلیک می کنیم. در نتیجه این کار منوی Simple Scatterplot ظاهر می شود در این منوی یک متغیر را به قسمت Y Axis و متغیر دیگر را به قسمت X Axis وارد نموده، گزینه Ok را برای رسم نمودار کلیک می کنیم.

DATA - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Wi

1 :

	age	sex	w
1	55.00	male	
2	30.00	male	
3	41.00	male	
4	28.00	female	
5	44.00	male	
6	.	female	
7	38.00	female	
8	31.00	female	
9	29.00	female	
10	57.00	male	

Gallery  
Interactive  
Map  
Bar...  
Line...  
Area...  
Pie...  
High-Low...  
Pareto...  
Control...  
Boxplot...  
Error Bar...  
Scatter...  
Histogram...

Scatterplot

Simple Matrix  
Overlay 3-D

Define  
Cancel  
Help

Simple Scatterplot

sex  
h  
groups  
result  
complica  
w2

Y Axis: w1

X Axis: age

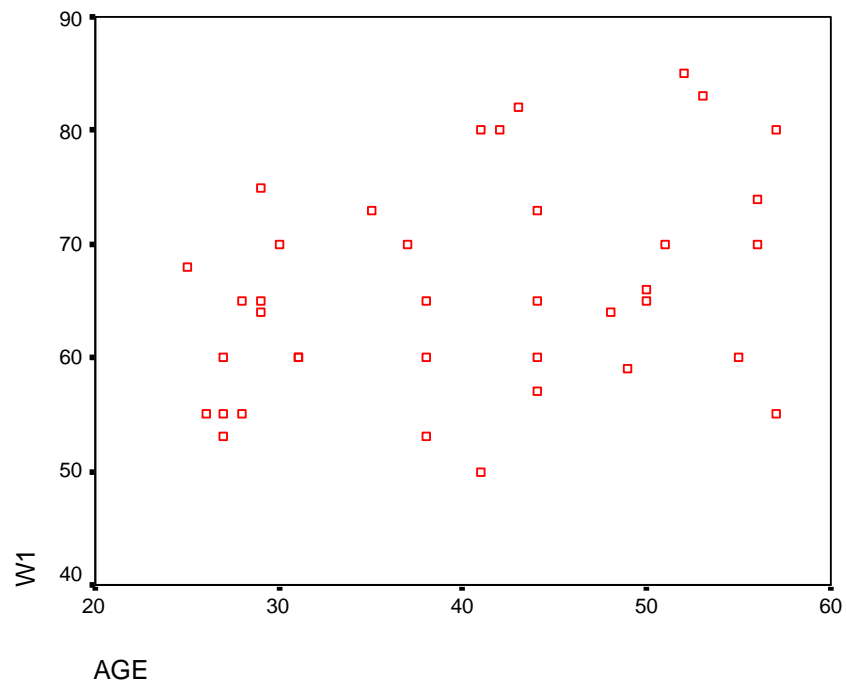
Set Markers by:

Label Cases by:

Template  
 Use chart specifications from:

OK  
Paste  
Reset  
Cancel  
Help

در قسمت خروجی نمودار ذیل به دست می آید.



سایر نمودارها همچون نمودار مستطیلی (هیستوگرام)، نمودار چندضلعی و ... را می توان رسم نمود.

ویرایشگر نمودار:

برای ویرایش نمودن نمودارها، با انجام دو بار کلیک روی نمودار به قسمت ویرایشگر نمودار می رویم. تغییر رنگ، سه بعدی نمودن، تغییرات در فونت، تغییرات در عنوان در این قسمت انجام می شود. سپس ویرایشگر نمودار را می بندیم ☒ تغییرات انجام شده در فایل Output قرار می گیرد.

تذکر: توصیه می شود رسم نمودار با استفاده از نرم افزارهای دیگر مثل Excell صورت پذیرد.

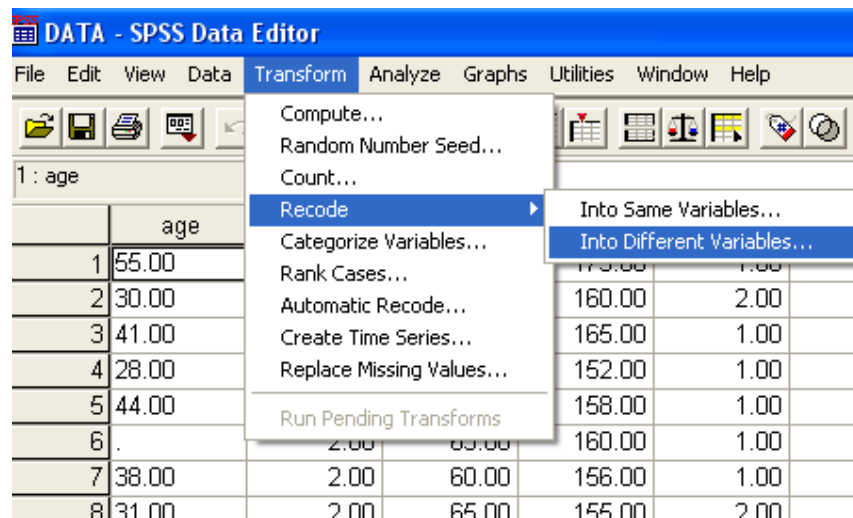
تبدیل داده ها (Data Transformation):

در بسیاری از موارد لازم است که متغیرهای کمی دسته بندی شوند، بعنوان مثال دسته بندی افراد به گروه های سنی ۰-۱۰ سال ۱۱-۲۰ و ... یا ممکن است بخواهیم دو گروه شغلی کارگر و کشاورز را در هم ادغام کنیم. برای این منظور از مطالب ذیل استفاده می نمایم.

دسته بندی کردن

ابتدا به قسمت ویرایشگر داده ها می رویم، از نوار منو، روی کلید Transform رفته، آن را کلیک می کنیم، سپس Recode را انتخاب می کنیم، دو گزینه پیش رو داریم.

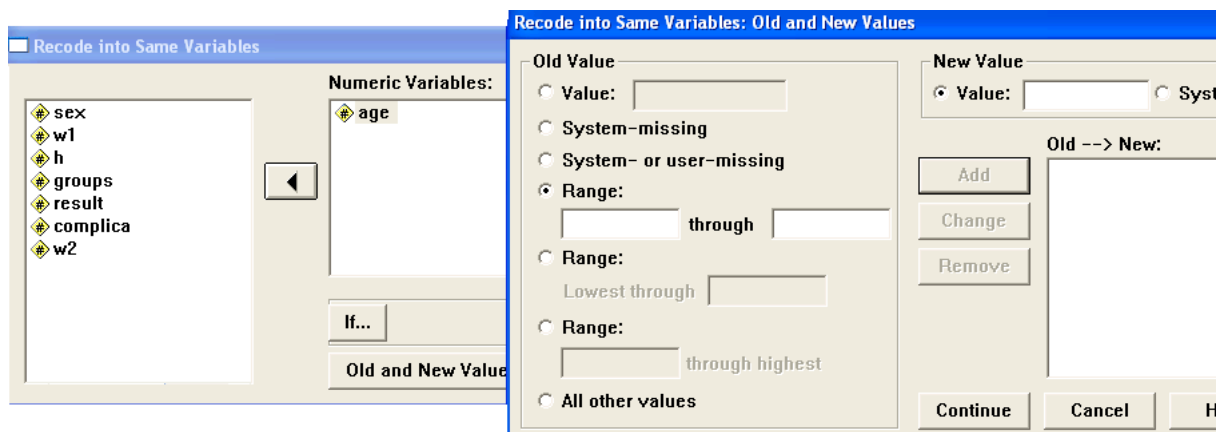




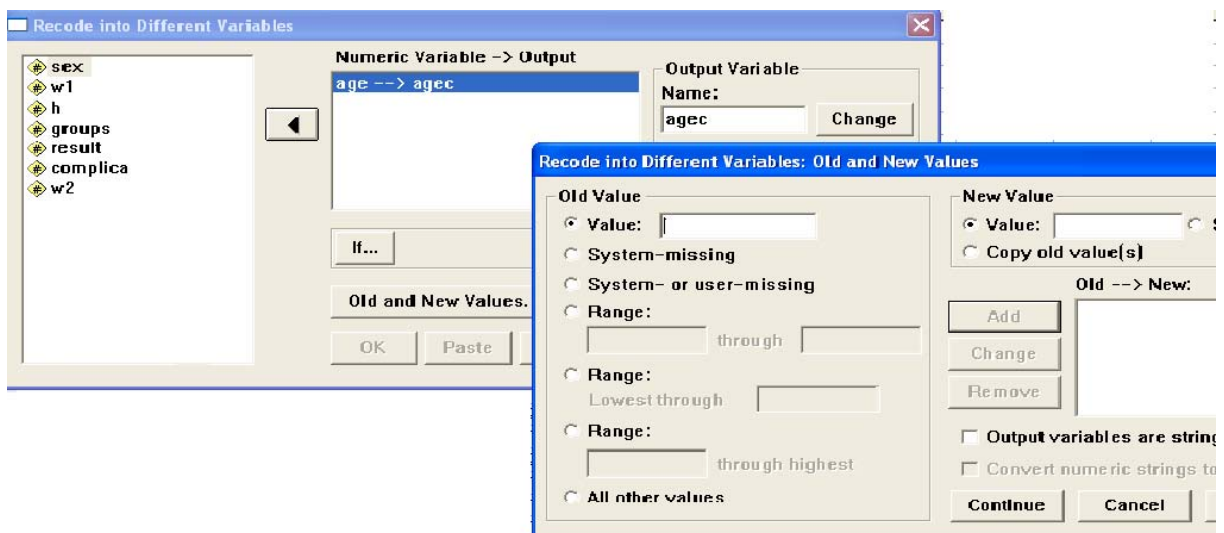
الف) اگر بخواهیم نتیجه دسته بندی روی همان متغیر جایگزین شود (Same Variable). تذکر مهم: با این عمل داده های متغیر قبلی از بین میرود.

ب) اگر بخواهیم نتیجه دسته بندی بصورت یک متغیر دیگر و در ستون جدیدی در ویرایشگر داده ها ثبت شود. (متغیر اول حفظ شده تغییری نمی کند).

الف - روی Same Variable کلیک می کنیم، پنجره ای ظاهر می شود، متغیری که قرار است دسته بندی شود از لیست متغیرها انتخاب و آن را به سمت راست می آوریم، کلید Old and New را انتخاب نموده، پنجره جدیدی ظاهر می شود، قسمت سمت چپ مربوط به Old و سمت راست مربوط به New می باشد. در قسمت Old یکی از گزینه های مورد نظر را انتخاب می کنیم مثلاً Range اولی را فعال می کنیم (با کلیک کردن در دایره توخالی کنار آن) اگر بخواهیم سنین ۰ تا ۱۰ را در یک گروه قرار دهیم عدد صفر را در مستطیل اول و ۱۰ را در مستطیل دوم قرار می دهیم، سپس به قسمت بالای صفحه سمت راست رفته و کد ۱ را برای آن فاصله بر می گزینیم. کلید Add را زده تا تعریف انجام شود، سپس به قسمت Old بر می گردیم و فاصله ۱۱-۲۰ را قرار داده و در قسمت New کد ۲ را برای این فاصله انتخاب و Add می کنیم و به همین ترتیب ادامه داده تا دسته بندی پایان یابد. با انجام Continue به صفحه قبل برگشته و Ok را می زنیم. در این صورت متغیر قبلی به کدهای ۱، ۲ و ۳ ... تبدیل می شود. حال بایستی روی Variable View در پائین سمت چپ ویرایشگر داده ها رفته و این کدها را برای متغیر تعریف کنیم کد ۱ یعنی ۱۰-۰ و کد ۲ یعنی ۱۱-۲۰ و ...



ب- **Different Variable** را بر می‌گزینیم در پنجره ای که ظاهر می‌شود متغیر مورد نظر را انتخاب و به پنجره سمت راست می‌بریم سپس در مستطیلی که ذیل نام متغیر خروجی (Output Variable Name) نام جدیدی می‌گذاریم این نام از قانون اسم متغیرها پیروی می‌کند (یعنی نباید بین حروف جای خالی باشد یا از علائم نباید استفاده کرد، بیشتر از ۸ کاراکتر نباشد و ...) پس از انتخاب نام متغیر جدید، کلید **change** را کلیک می‌کنیم تا در نوار آبی شکل ذیل قرار گیرد و سپس به قسمت **Old and New Variable** رفته و از این قسمت به بعد مانند قسمت الف ادامه می‌دهیم.

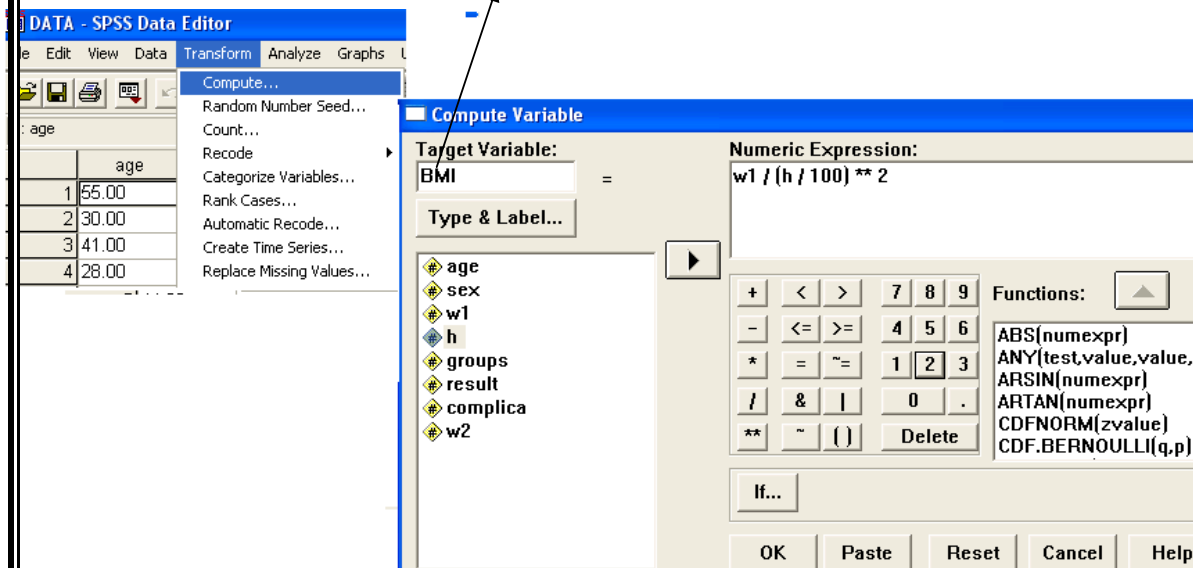


### ایجاد متغیرهای مرکب

اگر بخواهیم از روی دو یا چند متغیر یک اندیکس خاص (مثلا BMI) را تولید کنیم به محاسبات نیاز داریم. یا بخواهیم مجموع نمراتی که هر فرد از چند سؤال بدست می‌آورد را با هم جمع کنیم از قسمت محاسبات استفاده می‌نمائیم. به ویرایشگر داده‌ها رفته و کلید **Transform** را کلیک می‌کنیم در اولین سطر عبارت **Compute** را برمی‌گزینیم، نامی را برای متغیری که می‌خواهیم تولید کنیم انتخاب کنیم (مثلا Score) یا هر اسم دیگر) سپس فرمول مربوطه را با استفاده از لیست متغیرها و کلیده‌های محاسباتی که در همین پنجره وجود دارد می‌نویسیم و سپس کلید **Ok** را می‌فشاریم.

مثال: اگر قد بر حسب سانتیمتر و وزن بر حسب کیلوگرم باشد، BMI که حاصل تقسیم وزن به کیلوگرم بر مربع قد (بر حسب متر) چنین محاسبه می شود.

این نام اختیاری است و شرایط نام متغیر را دارد .



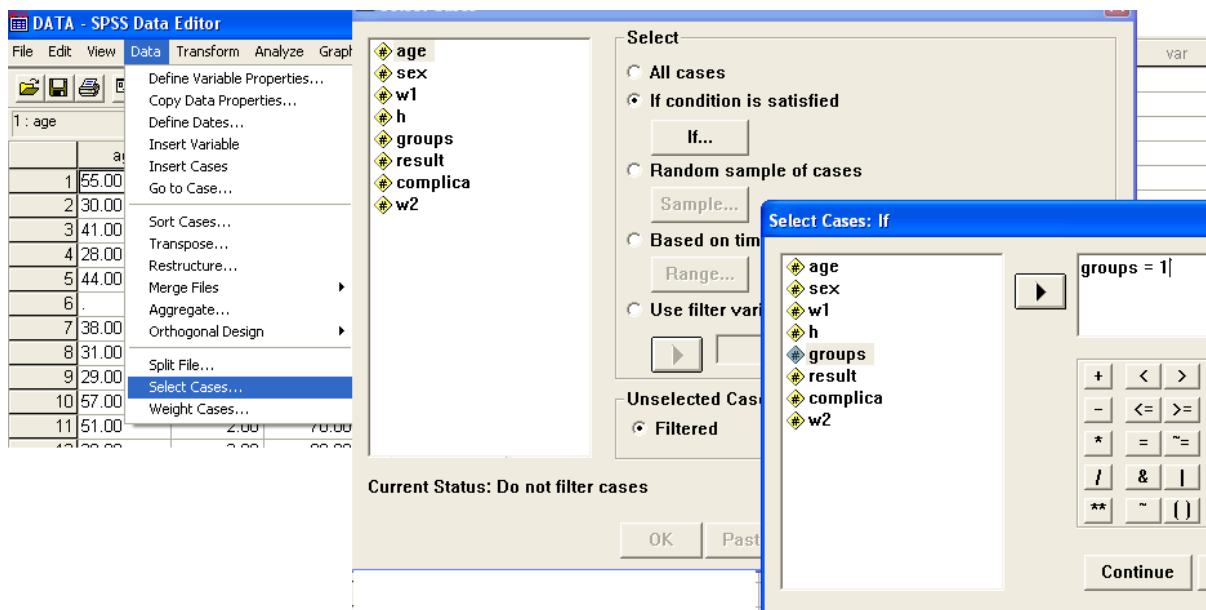
انجام محاسبات بدین ترتیب است که ابتدا از داخلی ترین پرانتز محاسبه آغاز می شود سپس توان و پس از آن ضرب و تقسیم و بعد جمع و تفریق. با انجام یک (Frequency) می توان از صحت محاسبات مطلع شد.

### فیلتر کردن

اگر بخواهیم شاخص های آماری یا توزیع فراوانی را برای گروه هایی خاص انجام دهیم. مثلاً بخواهیم در افرادی که سن بیشتر از ۴۰ سال دارند، میانگین و انحراف معیار BMI را بدست آوریم، یا فقط در افراد مورد بخواهیم برخی شاخص ها را محاسبه کنیم. برای این منظور در ویرایشگر داده ها، از منوی Data وارد می شویم سپس Select case را کلیک می کنیم، پنجره ای ظاهر می شود، در قسمت فوقانی پنجره If را فعال می کنیم و سپس شرط مربوطه را می نویسیم مثلاً  $Age > 40$  یا  $group = 1$  سپس با Ok کردن، بطور اتوماتیک فعالیت های بعد از آن صرفاً برای افراد بالاتر از ۴۰ سال یا برای افرادی که کد متغیر گروه آنها ۱ می باشد، خواهد بود.

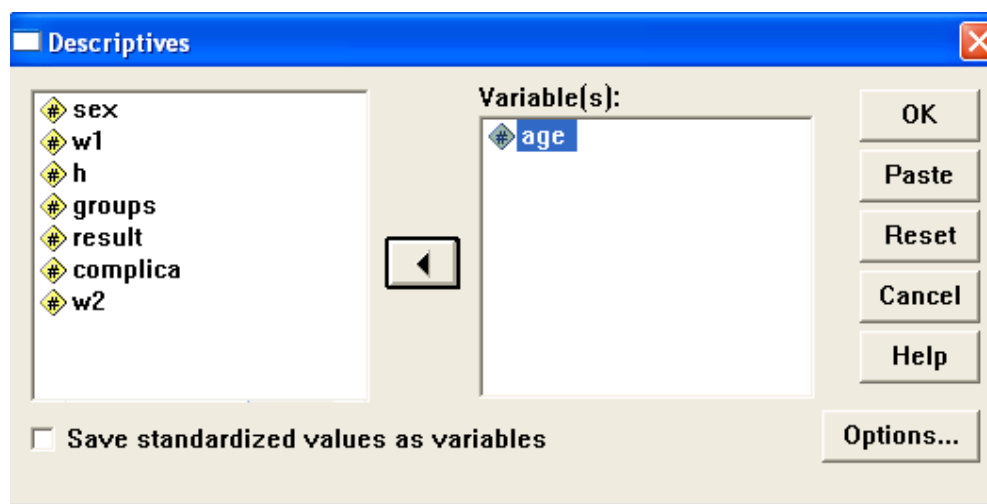
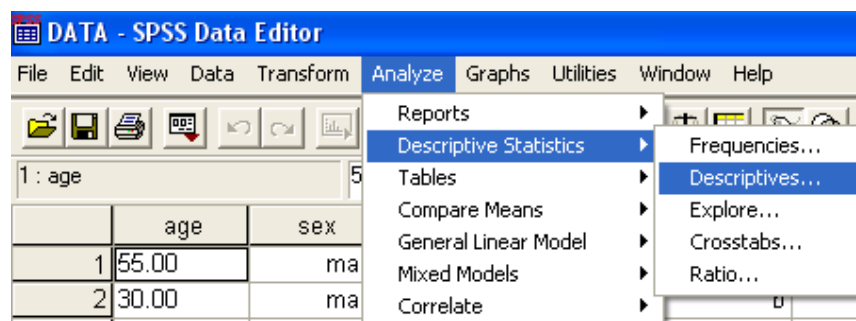
تذکر ۱: اگر بخواهیم محاسباتی را راجع به افرادی که هم در گروه ۱ هستند و هم  $age > 40$  بین آنها عبارت & را قرار می دهیم .

تذکر ۲: پس از محاسبات لازم بر روی این گروه خاص ، با رفتن مجدد به منوی Data و انجام Select case قسمت All Case را فعال کرده و از آن به بعد مجدداً فعالیت ها روی کل افراد صورت خواهد پذیرفت.(این قسمت ممکن است فراموش شود).



مثال: میانگین و انحراف معیار سن را برای افرادی که در گروه ۱ هستند، بدست آورید.

پس از انجام مراحل فوق از منوی اصلی analyze را انتخاب و مسیر ذیل را ادامه می دهیم.



با OK نمودن خروجی فقط برای افرادی محاسبه می شود که در گروه ۱ هستند.

در این بخش برخی از آزمون های ساده آماری را تمرین خواهیم نمود.

## آزمون فرضیه های آماری

به طور کلی هدف از آزمون های آماری تعیین حکمی درباره پارامتر یا پارامترهای مجهول جامعه است که می تواند درست یا نادرست باشد. هرگاه بخواهیم یک ادعا درباره پارامتر مجهول جامعه را از طریق تایید آن به وسیله اطلاعات حاصل از نمونه ثابت کنیم ادعای مورد نظر را به عنوان فرضیه یک (فرض جانشین) (Alternative Hypothesis) و حالت نفی آن ادعا را به عنوان فرضیه صفر (Null Hypothesis) در نظر می گیریم. با توجه به مطالعات گوناگون، فرضیه ها و آزمون های گوناگونی خواهیم داشت. به عنوان مثال در حالت آزمون برابری میانگین دو جامعه مستقل فرضیه جانشین (فرضیه ادعا) این است که میانگین دو جامعه برابر نیستند در مقابل اینکه میانگین های دو جامعه برابر هستند (فرضیه صفر).

در حالت معمول اساس را بر درستی فرضیه صفر قرار می دهیم مگر اینکه شواهد کافی با استفاده از نمونه خلاف آن را قویا نشان دهد که در این صورت گوییم نتیجه معنی دار بدست آمده است. آزمون فرضیه تعیین روشی است که بر اساس اطلاعات بدست آمده از نمونه بتوان نسبت به درستی یا نادرستی فرضیه صفر تصمیم گیری نمود. واضح است که در تصمیم گیری درباره رد فرضیه صفر با دو نوع خطا مواجه هستیم. ممکن است فرضیه صفر واقعا درست باشد اما نتیجه آزمون منجر به رد آن باشد که این خطا را خطای نوع اول می نامیم. همچنین ممکن است آزمون به رد فرضیه صفر نینجامد در حالی که فرضیه صفر واقعا درست نباشد که آن را خطای نوع دوم می نامیم.

فرضیه صفر	فرضیه صفر	واقعیت
درست نیست	درست است	نتیجه آزمون
-	خطای نوع اول	رد فرضیه صفر
خطای نوع دوم	-	رد نکردن فرضیه صفر

احتمال ارتکاب خطای نوع اول را با  $\alpha$  نشان داده و آن را سطح معنی داری آزمون گوئیم. همچنین احتمال ارتکاب خطای نوع دوم را با  $\beta$  نشان می دهیم و مقدار  $(1-\beta)$  را توان آزمون می گوئیم.

واضح است که آزمون فرضیه باید به گونه ای باشد که در آن احتمال ارتکاب خطاهای نوع اول و دوم تا حد ممکن حداقل باشند. یک راه حل برای این منظور افزایش حجم نمونه است. اما چنانچه در انتخاب نمونه محدودیت داشته باشیم لازم است از روش آزمونی استفاده نماییم که با توجه به اهمیت بیشتر خطای نوع اول، احتمال ارتکاب این خطا را ثابت گرفته و همزمان احتمال خطای نوع دوم را حداقل نماید.

در یک تحلیل آماری رد کردن فرضیه صفر به عنوان "نتیجه معنی دار Significance" نامیده می شود. بنابراین یک نتیجه معنی دار نتیجه ای است که علی القاعده بر حسب شانس و تصادف رخ نداده باشد اگر چه سطح معنی دار ( $\alpha$ ) باید مشخص باشد. مقادیر متداول برای سطح معنی دار آزمون ( $\alpha$ ) عبارتند از:  $0/01$ ،  $0/05$  و  $0/1$ . معمولاً وقتی که واژه "نتیجه معنی دار" استفاده می شود، منظور از آن رد کردن فرضیه صفر در سطح خطای در نظر گرفته شده است (مثلاً  $0/05$  یا  $0/01$ ) است. یک نتیجه "غیر معنی دار" یعنی اینکه فرضیه صفر رد نشده است.

در عمل فرض می نماییم که فرضیه صفر صحیح است، پس از آن میزان تطابق احتمالی نمونه بدست آمده با فرضیه صفر را محاسبه می نماییم. این مقدار در نرم افزارهای آماری با عبارت "مقدار احتمالی P-Value" یا عبارت "میزان معنی داری Significance" و در SPSS با Sig. مشخص می شود. در یک قاعده کلی فرضیه صفر را رد می کنیم هر گاه P-Value یا Significance آزمون کمتر از احتمال خطای نوع اول گردد که در این صورت گوئیم آزمون در سطح  $\alpha$  معنی دار است.

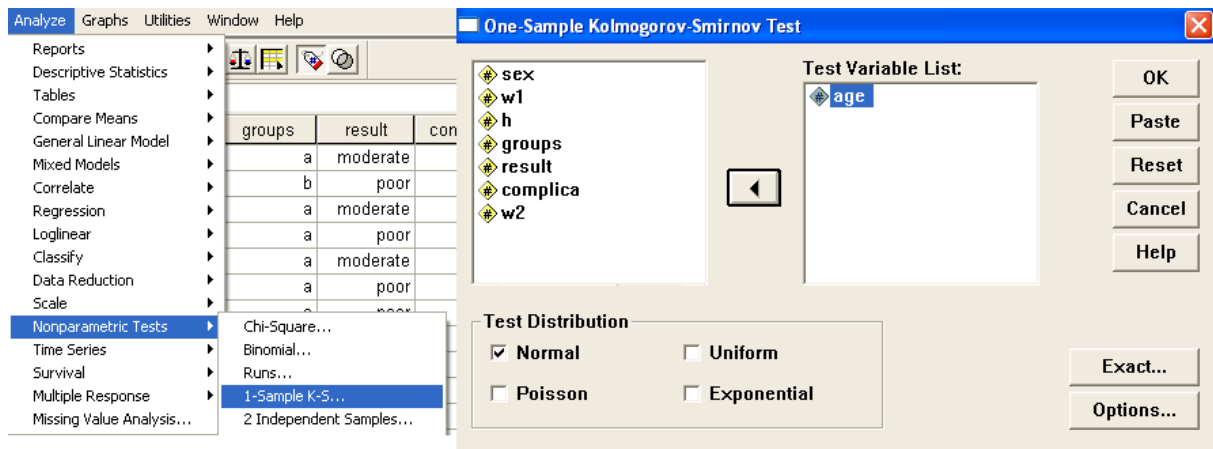
### بررسی فرضیه نرمال بودن مشاهدات

یکی از فرضیات اصلی برای اکثر آزمون های آماری، نرمال بودن توزیع مشاهدات است. اجرای آزمون نرمال بودن استفاده از آزمون کلموگروف-اسمیرنوف (Kolmogorov-smirnov) است. این آزمون یک آزمون ناپارامتری برای بررسی توزیع مشاهدات است. پیش فرض SPSS، انجام آزمون برای بررسی توزیع نرمال است. بررسی توزیع های پواسن (Poisson)، یکنواخت (Uniform)، و نمایی (Exponential) نیز با این آزمون امکان پذیر می باشد. میزان تقریبی معنی داری آزمون (Asymptotic Significance) در انتهای خروجی آزمون کلموگروف-اسمیرنوف می آید که با مقایسه آن با  $\alpha$  می توان آزمون را در سطح معنی داری  $\alpha$  انجام و نسبت به نرمال بودن توزیع مشاهدات تصمیم گیری نمود.

اگر  $\alpha = 0/05$  در نظر بگیریم چنانچه  $P\text{-value} > 0/05$  باشد، می توان فرض کرد، توزیع مشاهدات نرمال است.

مثال: آیا می توان فرض کرد توزیع سن در جامعه ای که از آن نمونه انتخاب کرده ایم نرمال است؟

مسیر ذیل را از منوی اصلی از قسمت analyze انتخاب می کنیم.



با OK کردن خروجی به صورت ذیل ظاهر می شود.

## ► NPar Tests

**One-Sample Kolmogorov-Smirnov Test**

		AGE
N		38
Normal Parameters <sup>a,b</sup>	Mean	40.2632
	Std. Deviation	10.47180
Most Extreme Differences	Absolute	.154
	Positive	.154
	Negative	-.087
Kolmogorov-Smirnov Z		.949
Asymp. Sig. (2-tailed)		.329

a. Test distribution is Normal.

b. Calculated from data.

بزرگتر از ۰,۰۵ پس نرمال است

مقایسه میانگین ها (compare means) :

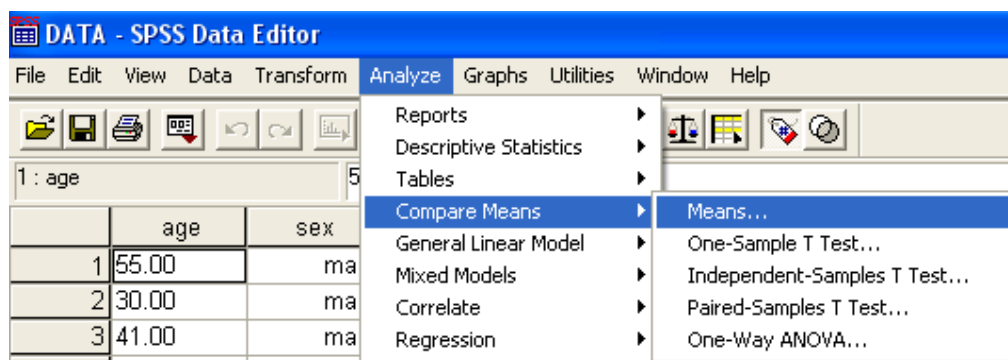
در بسیاری از مطالعات ، هدف مقایسه بین میانگین گروه ها در جامعه می باشد . با انجام نمونه ای که از جامعه انتخاب می شود .

الف ) می توان به صورت توصیفی میانگین های نمونه را در زیر گروه های مختلف نشان داد.

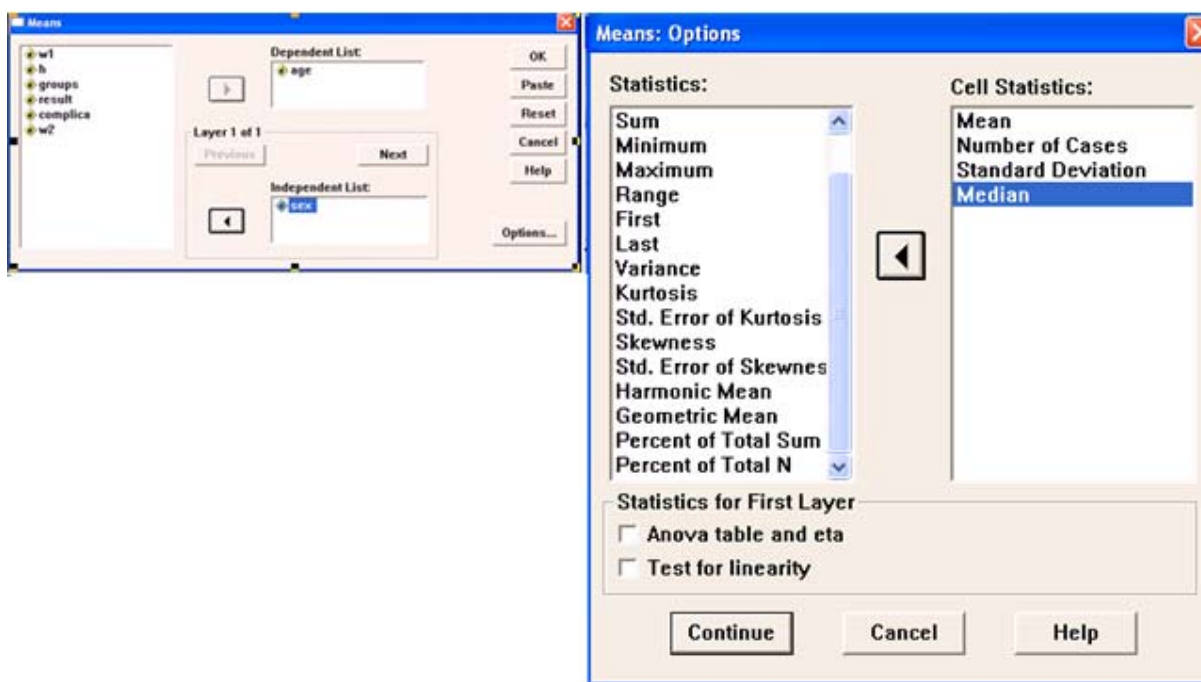
ب ) می توان با آزمون های آماری یکسان بودن آن ها را در جامعه ازمون نمود.

برای قسمت الف می توان از مسیر ذیل اقدام نمود.

مثال: میانگین، میانه و انحراف معیار سن را در افراد مذکر و مونث بدست آورید.



با کلیک کردن روی **Means** پنجره ذیل ظاهر می شود که با انتخاب **Age** و **sex** از لیست متغیرها و **OK** کردن، تعدا، میانگین و انحراف معیار در خروجی ظاهر می شود ولی اگر شاخص های دیگری همچون میانه، حداقل و سایر شاخص ها در دو جنس نیاز باشد با کلیک روی **Option** و انتخاب شاخص ها از لیست آنها می توان مقادیرشان را به دست آورد.



و با انجام **Continue** و سپس **OK** خروجی به صورت ذیل است.



## Report

AGE				
SEX	Mean	N	Std. Deviation	Median
male	41.9474	19	10.93655	41.0000
female	38.5789	19	9.99064	38.0000
Total	40.2632	38	10.47180	41.0000

ب) برای مقایسه میانگین جامعه به چند آزمون مهم اشاره می شود.

### آزمون میانگین یک جامعه با یک عدد ثابت

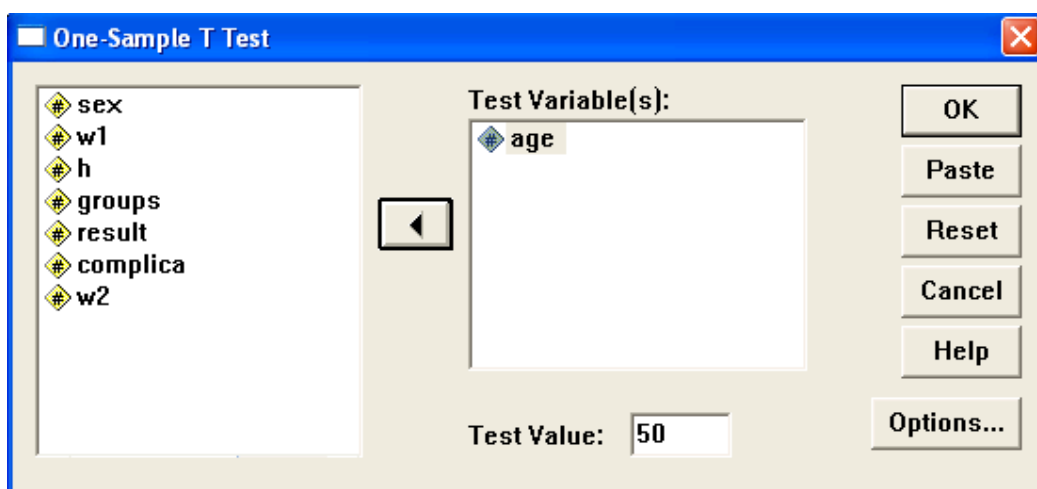
هدف از انجام این آزمون بررسی برابر بودن میانگین یک متغیر کمی جامعه با یک مقدار مشخص می باشد

$$(H_0: \mu = \mu_0)$$

مثال: آیا میانگین سن در افراد جامعه برابر با عدد ۵۰ می باشد. ( $H_0: \mu = 50$ )

پیش فرض انجام آزمون این است که متغیر مورد مطالعه دارای توزیع نرمال است (توسط کلموگروف - اسمیرنوف آزمون می شود که قبلا گفته شد). برای اجرای آزمون مقایسه میانگین با یک عدد ثابت مسیر زیر را در SPSS انتخاب می کنیم.

با این کار منوی One Sample T Test ظاهر می گردد. در این منو لازم است متغیر مورد مطالعه را از لیست متغیرها انتخاب و وارد قسمت Test Variable نمود. آنگاه مقدار  $\mu_0$  را وارد گزینه Test Value می کنیم. (در مثال بالا عدد ۵۰) با مشاهده مقدار Sig. در جدول خروجی آزمون و مقایسه آن با  $\alpha$ ، نسبت به درستی یا رد فرض صفر تصمیم گیری می نماییم. پیش فرض SPSS برای مقدار  $\mu_0$  عدد صفر می باشد یعنی چنانچه مقدار  $\mu_0$  را وارد Test Value ننماییم، SPSS آزمون برابری میانگین متغیر مورد مطالعه را با عدد صفر انجام می دهد.



→ T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
AGE	38	40,2632	10,47180	1,69875

One-Sample Test

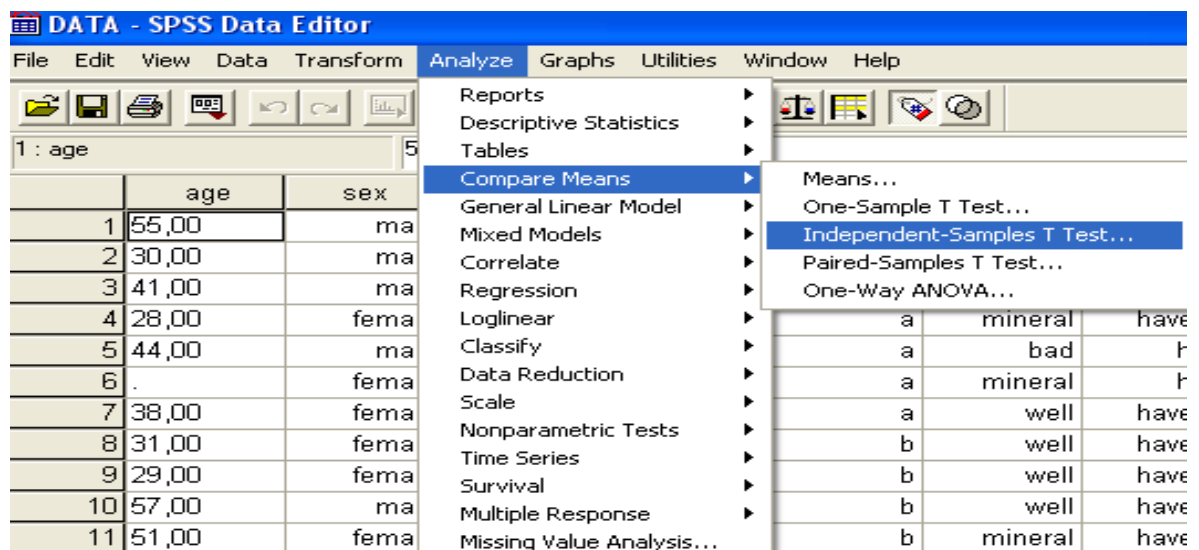
	Test Value = 50					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
AGE	-5,732	37	,000	-9,7368	-13,1788	-6,2948

p-value < 0.05

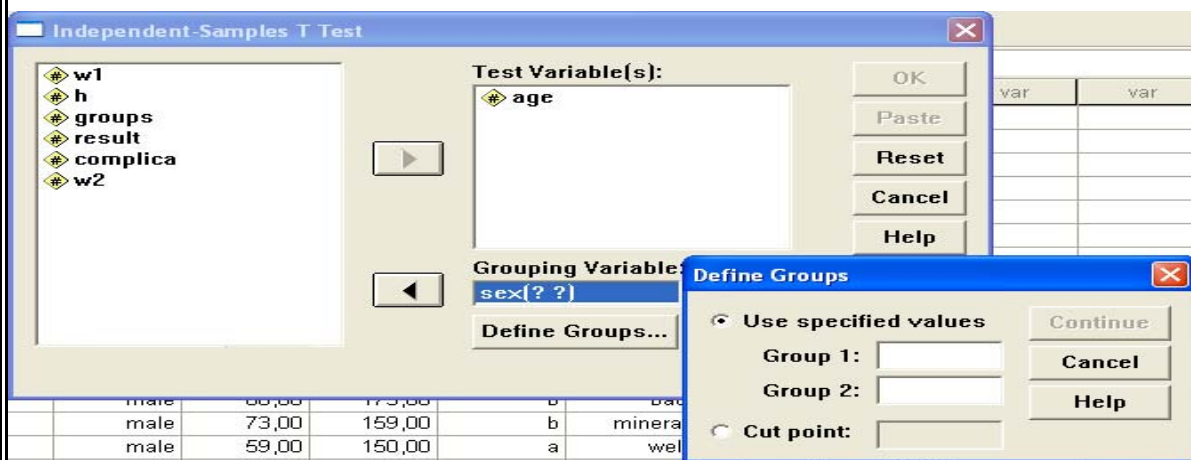
مقدار sig.(2-tailed) که همان p-value می باشد کمتر از ۰/۰۵ است لذا میانگین سن جامعه ای که از آن نمونه گرفته ایم با ۵۰ اختلاف معنی داری دارد.

آزمون برابری میانگین دو گروه مستقل

هدف از انجام این آزمون مقایسه میانگین های دو گروه مستقل در جامعه می باشد. (مثلا ، مقایسه میانگین سن در دو گروه مذکر و مونث) فرضیه انجام این آزمون نرمال بودن توزیع متغیر مورد مطالعه در دو گروه می باشد. در اینجا دو متغیر داریم، یک متغیر همان است که می خواهیم از آن میانگین بگیریم مثل سن که حتما باید کمی باشد و متغیر دیگر کیفی دو حالت است (مثل جنس). برای اجرای آزمون مسیر زیر را در SPSS انتخاب می نماییم.



با انتخاب مسیر فوق منوی Independent-Samples T-Test ظاهر می شود. حال متغیر مورد آزمون را از لیست متغیرها به قسمت Test Variable و متغیر گروه بندی را به قسمت Group Variable فراخوانی نموده و مقادیر متغیر گروه بندی را تعریف می نماییم. مثلا برای جنس کد ۱ و ۲ داشتیم.



با توجه به این که مقدار آماره آزمون و درجه آزادی آن بستگی به برابری یا نابرابری واریانس های متغیر مورد مطالعه در دو گروه دارد، لازم است آزمون برابری واریانس ها نیز به موازات این آزمون صورت گیرد. نتایج آزمون برابری واریانس های تحت عنوان آزمون لون (Leven's Test) در ابتدای جدول خروجی آزمون برابری میانگین ها ارائه می گردد. با مشاهده میزان معنی داری آن می توان نسبت به برابری واریانس ها تصمیم گیری نمود. در نهایت چنانچه فرض برابری واریانس ها پذیرفته شود از نتایج موجود در سطر اول و چنانچه فرض برابری واریانس ها رد گردد از نتایج موجود در سطر دوم آزمون میانگین استفاده و بر اساس مقدار Sig. آزمون میانگین، نسبت به برابری میانگین ها تصمیم گیری می شود. خروجی SPSS به صورت ذیل است.

Group Statistics

SEX	N	Mean	Std. Deviation	Std. Error Mean
AGE male	19	41,9474	10,93655	2,50902
AGE female	19	38,5789	9,99064	2,29201

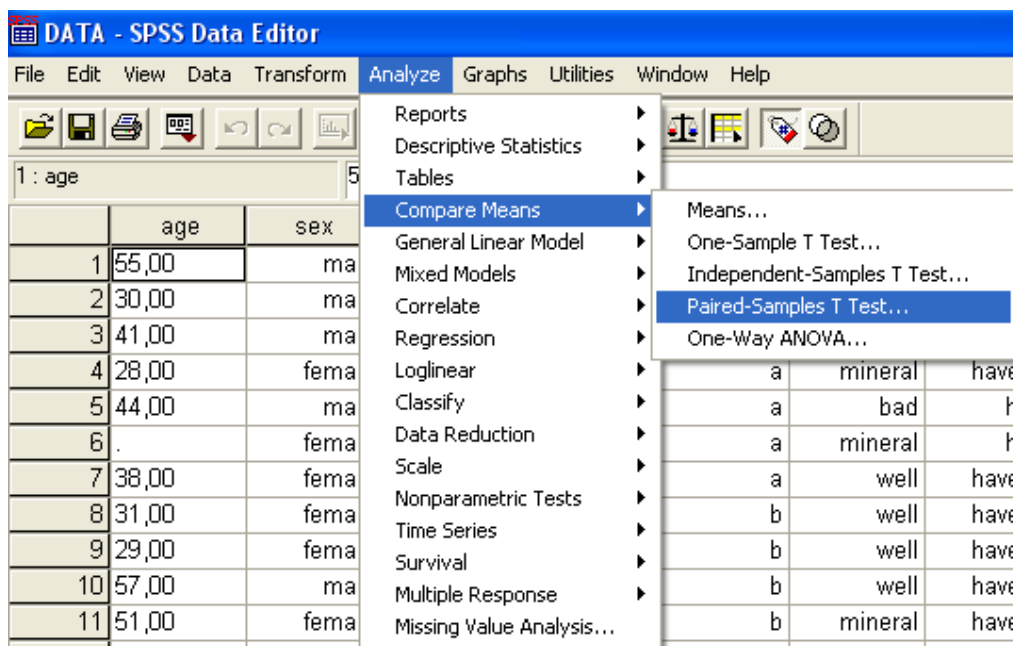
p-value > 0.05 معنی داری از خط اول

Independent Samples Test

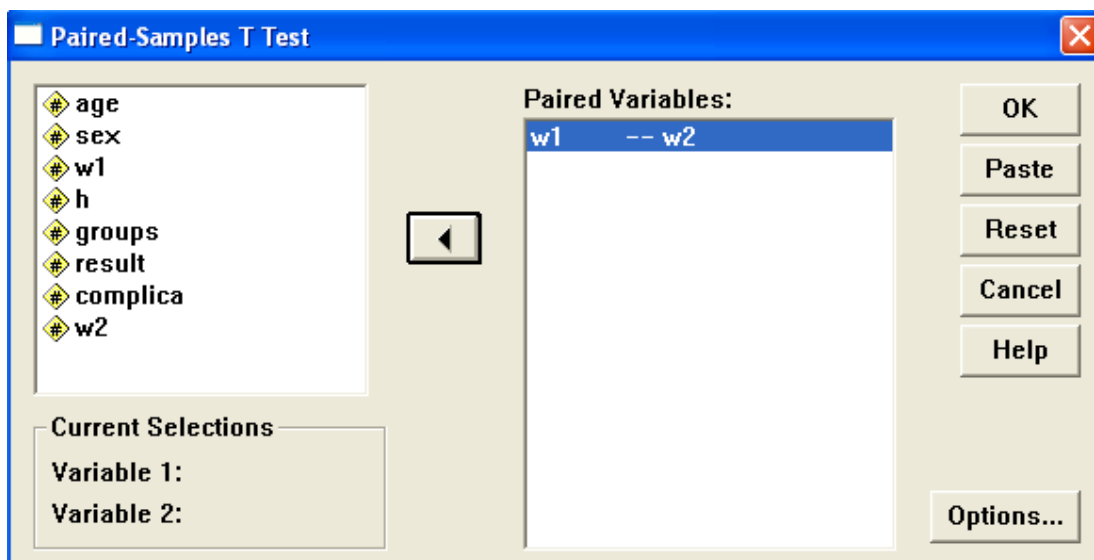
	Levene's Test for Equality of Variances	t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
AGE	Equal variances assumed	,093	,762	,991	36	,328	3,3684	3,3684
	Equal variances not assumed			,991	35,709	,328	3,3684	3,3684

## آزمون برابری میانگین های دو گروه وابسته (مشاهدات زوجی)

هدف از انجام این آزمون که به آزمون قبل و بعد نیز معروف می باشد برابری میانگین های دو گروه وابسته می باشد. پیش فرض انجام آزمون، نرمال بودن متغیر مورد مطالعه در دو گروه می باشد. در این آزمون برای هر آزمودنی دو مشاهده وجود دارد. برای انجام آزمون لازم است گروه اول و گروه دوم (مشاهدات قبل و بعد) را به صورت موازی وارد نمود به گونه ای که مشاهدات هر آزمودنی در مقابل یکدیگر باشند. (مشاهدات قبل به عنوان یک متغیر مثل وزن اولیه و مشاهدات بعد متغیری دیگر مثل وزن ثانویه) برای انجام آزمون لازم است مسیر زیر را انتخاب نمود.



با انتخاب مسیر فوق منوی Paired-Samples T-Test ظاهر می گردد در این منو لازم است دو متغیر قبل و بعد را همزمان انتخاب نموده و آنها را به قسمت Paired Variable فراخوانی نمود.



در خروجی این آزمون، ابتدا مقادیر میانگین و انحراف معیار دو متغیر جهت توصیف داده ها ظاهر می شود و سپس نتیجه آزمون همبستگی بین مقادیر قبل و بعد (این قسمت در تفسیر مقایسه میانگین قبل و بعد تأثیری ندارد)، سپس نتیجه آزمون برابری میانگین دو گروه که بر اساس برابری میانگین اختلافات قبل و بعد با مقدار صفر می باشد، ارائه می گردد. فرض برابری میانگین قبل و بعد در سطح خطای  $\alpha$  رد می گردد اگر مقدار Sig. آزمون کمتر از  $\alpha$  باشد.

**Paired Samples Statistics**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	W1	65.6750	40	9.17462	1.45064
	W2	61.8750	40	9.51433	1.50435

**Paired Samples Correlations**

		N	Correlation	Sig.
Pair1	W1 & W2	40	.837	.000

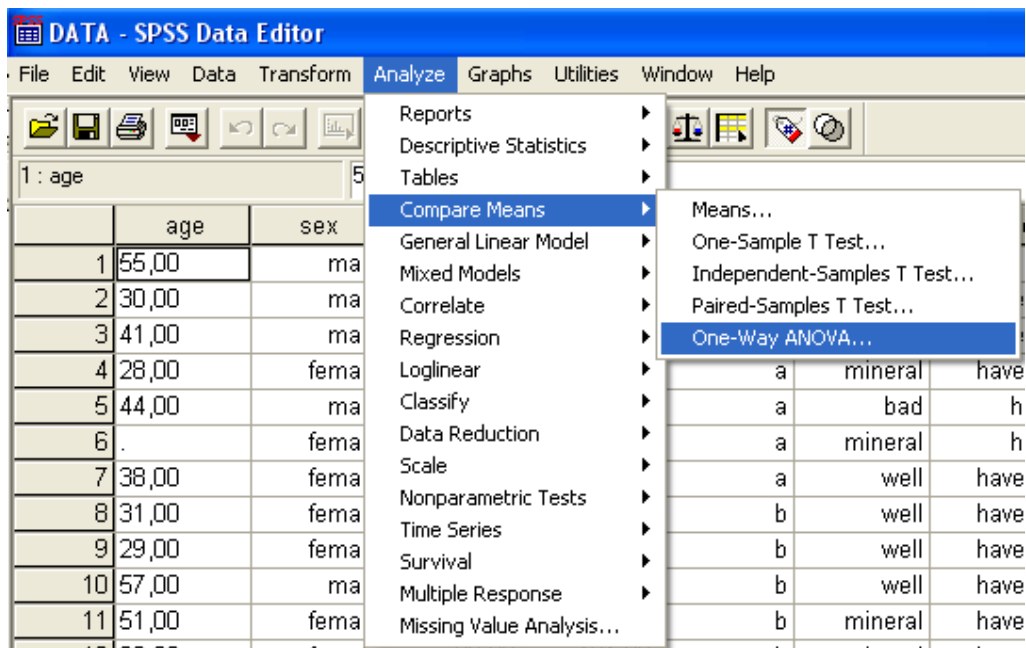
**Paired Samples Test**

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair1	W1 - W2	3.8000	5.34070	.84444	2.0920	5.5080	4.500	39	.000

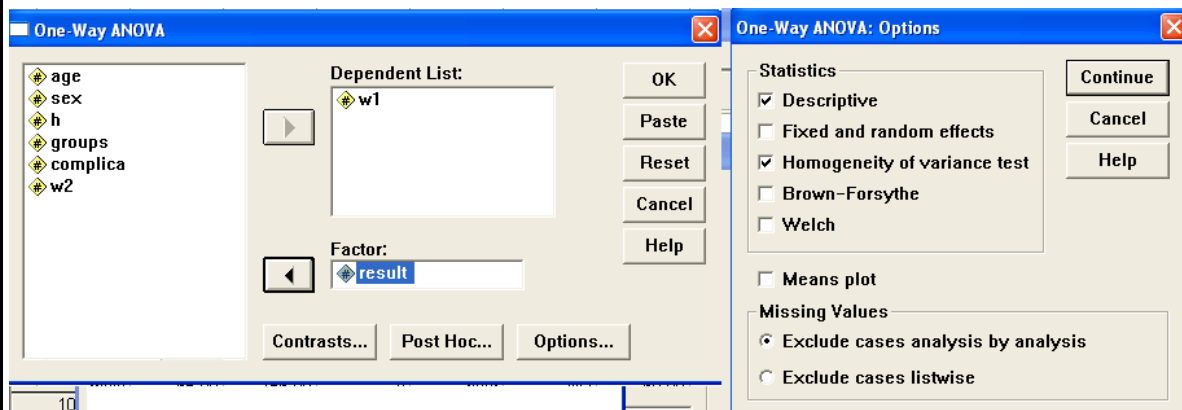
p-value < 0.05 بین میانگین قبل و بعد اختلاف معنی دار وجود دارد

### آزمون برابری میانگین های چند گروه مستقل (آنالیز واریانس یک راهه) (One-Way ANOVA)

پیش فرض انجام آزمون برابری میانگین چند گروه این است که توزیع مشاهدات در هر گروه نرمال و واریانس صفت در گروه ها برابر باشند. برای انجام آنالیز واریانس یک راهه، دو متغیر داریم. یکی متغیر کمی که همان متغیر مورد مطالعه است و متغیر دیگر یک متغیر کیفی که همان متغیر گروه بندی می باشد. (مثلا، متغیر کمی مثل وزن اولیه و یک متغیر کیفی که بیش از دو گروه دارد مثل نتیجه درمان که سه حالت خوب، متوسط و بد دارد). برای آزمون، مسیر زیر را انتخاب می کنیم.



در منوی آنالیز واریانس یک راهه متغیر کمی مورد مطالعه را به قسمت Dependent Variable و متغیر گروه بندی را به قسمت Factor فراخوانی می نماییم.



بطور کلی در ANOVA تغییرات متغیر کمی به دو بخش تغییرات درون گروهی و تغییرات بین گروهی تقسیم شده و بر اساس مقایسه این تغییرات، آماره آزمون محاسبه می گردد. با مشاهده مقدار Sig. خروجی جدول آنالیز واریانس می توان نسبت به برابری میانگین ها تصمیم گیری کرد. چنانچه فرض برابری میانگین ها رد شود نمایانگر این است که حداقل یکی از میانگین ها با بقیه متفاوت است.

پس از کلیک روی Continue و سپس OK در خروجی سه جدول ظاهر می شود.

۱- توصیف داده ها

۲- آزمون برابری واریانس ها (از شرایط آنالیز واریانس می باشد و بایستی برقرار باشد).

۳- جدول ANOVA

## Oneway

### Descriptives

W1

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
poor	8	66.5000	8.12404	2.87228	59.7081	73.2919	55.00	80.00
moderate	12	66.4167	8.94893	2.58333	60.7308	72.1025	53.00	80.00
good	20	64.9000	10.04149	2.24535	60.2004	69.5996	50.00	85.00
Total	40	65.6750	9.17462	1.45064	62.7408	68.6092	50.00	85.00

### Test of Homogeneity of Variances

W1

Levene Statistic	df1	df2	Sig.
.179	2	37	.837

چون  $p\text{-value} > 0.05$  می باشد بنابراین واریانس در گروه ها یکسان است

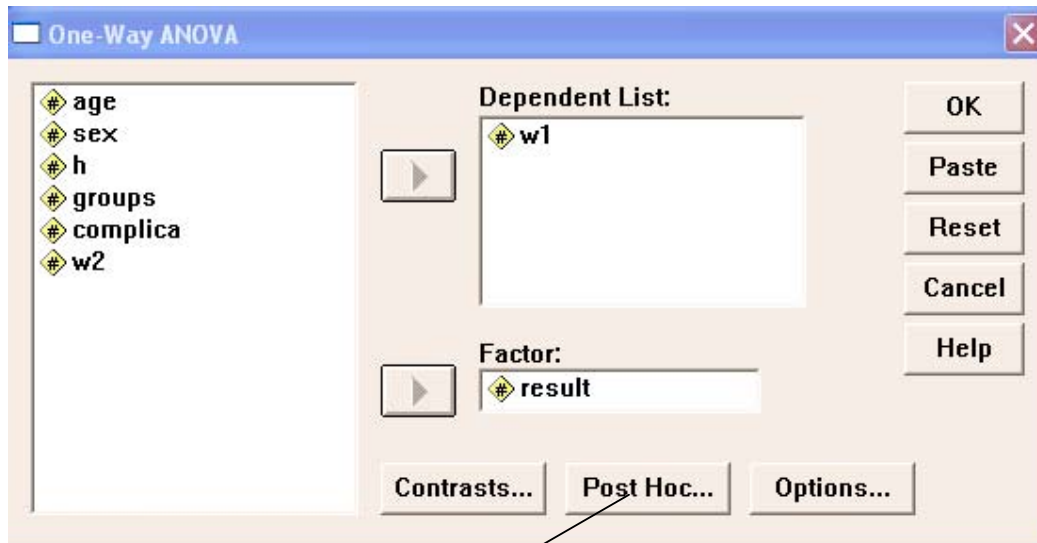
### ANOVA

W1

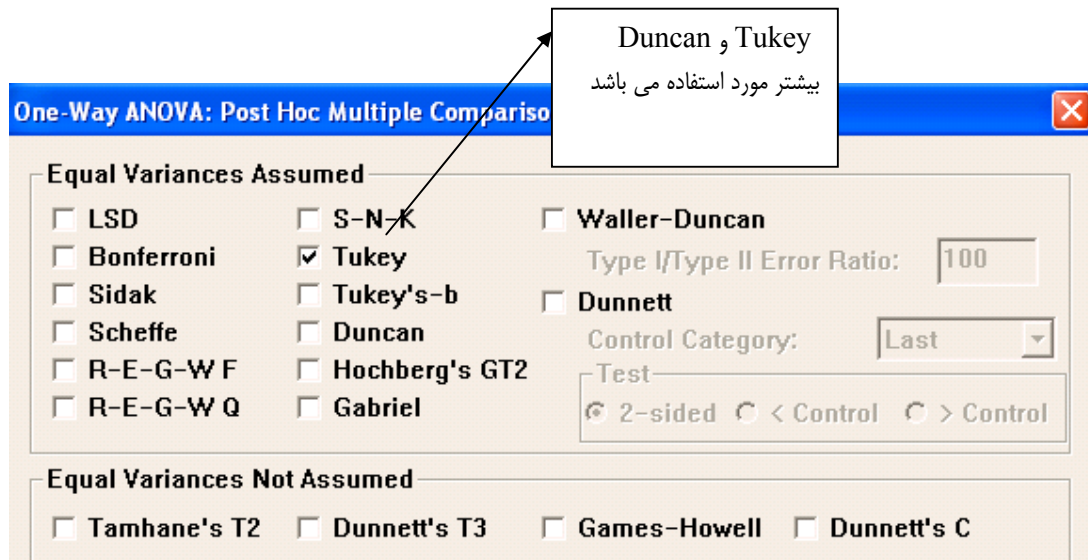
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	24.058	2	12.029	.137	.873
Within Groups	3258.717	37	88.073		
Total	3282.775	39			

چون  $p\text{-value} > 0.05$  می باشد بنابراین میانگین در گروه ها یکسان است

اگر نتیجه آنالیز واریانس نمایانگر تفاوت بین گروه ها باشد، ( $p\text{-value} < 0.05$ ) برای بررسی اینکه میانگین ها کدام یک از گروه ها با بقیه متفاوت است، می توان از مقایسات چندگانه Post Hoc استفاده نمود. از مهمترین روش های مقایسه می توان گزینه Post Hoc را از منوی One-Way ANOVA کلیک و سپس یکی از روش های فوق را برای مقایسه میانگین ها انتخاب نمود.



کلیک روی آن برای مقایسه دو به دو



پس از انتخاب تست توکی جدول ذیل نیز به جدول های قبلی اضافه می شود .

## Post Hoc Tests

**Multiple Comparisons**

Dependent Variable: W1  
Tukey HSD

(I) RESULT	(J) RESULT	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
poor	moderate	.0833	4.28353	1.000	-10.3748	10.5415
	good	1.6000	3.92592	.913	-7.9851	11.1851
moderate	poor	-.0833	4.28353	1.000	-10.5415	10.3748
	good	1.5167	3.42682	.898	-6.8499	9.8832
good	poor	-1.6000	3.92592	.913	-11.1851	7.9851
	moderate	-1.5167	3.42682	.898	-9.8832	6.8499

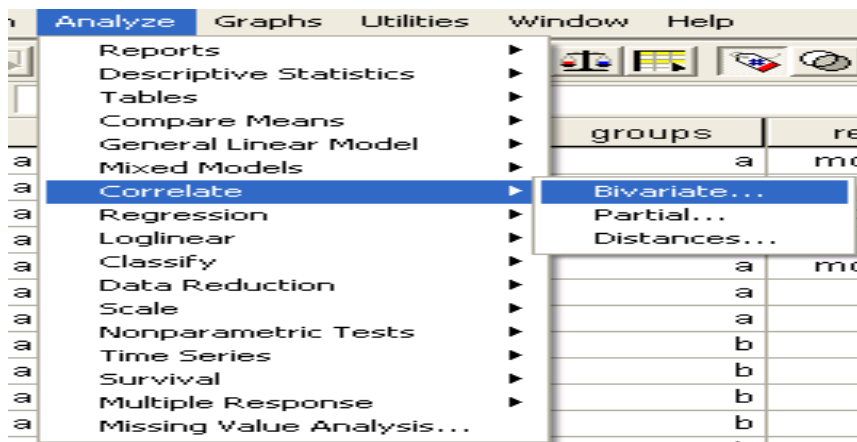
p-value را برای مقایسه دو به دو نشان می دهد

## آزمون همبستگی دو متغیر کمی

اولین قدم در بررسی همبستگی دو صفت کمی، رسم نمودار پراکنش (Scatter Plot) آن دو متغیر است. که در قسمت نمودار ها به آن اشاره شد. اما برای بررسی شدت و جهت رابطه بین متغیر کمی از ضریب همبستگی استفاده می شود. ضریب همبستگی دارای مقداری بین -۱ و ۱ است. چنانچه ضریب همبستگی برابر با صفر باشد نشان دهنده عدم ارتباط خطی دو متغیر با یکدیگر بوده و اگر قدرمطلق آن برابر با یک باشد نشان دهنده وجود یک رابطه قطعی بین دو متغیر است. برای بررسی همبستگی بین دو متغیر کمی از آزمون همبستگی استفاده می شود. چنانچه آزمون همبستگی معنی دار باشد نشان دهنده ارتباط خطی دو متغیر با یکدیگر است. در این صورت اگر ضریب همبستگی مثبت باشد نشان دهنده یک رابطه تقریبی مستقیم بین دو متغیر و اگر ضریب همبستگی منفی

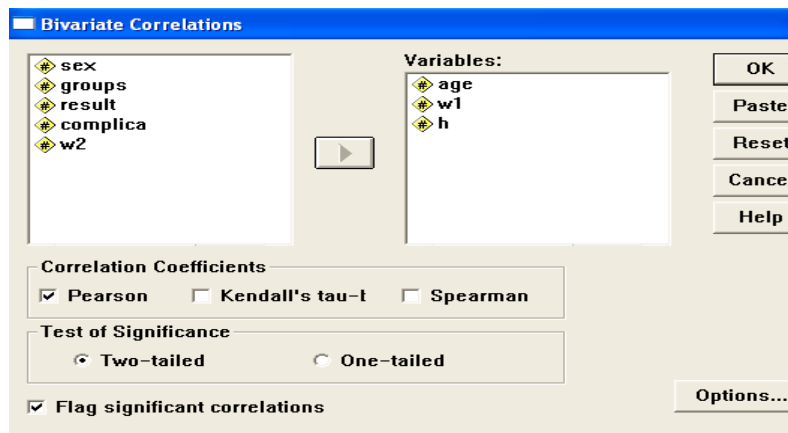


باشد نشان دهنده یک رابطه تقریبی معکوس بین دو متغیر است. هرچه قدرمطلق ضریب همبستگی بیشتر باشد نشان دهنده رابطه شدیدتر بین دو متغیر می باشد. چنانچه توزیع توام دو متغیر نرمال باشد از ضریب همبستگی پیرسن (Pearson Coefficient of Correlation) و در غیر این صورت از ضریب همبستگی اسپیرمن (Spearman) استفاده می شود. ضریب همبستگی پیرسن تنها وجود یا عدم وجود رابطه خطی بین دو متغیر را بررسی می کند در حالیکه ضریب همبستگی اسپیرمن وجود هرگونه رابطه بین دو متغیر کمی را نشان می دهد. آزمون ضریب همبستگی پیرسن پارامتری و آزمون ضریب همبستگی اسپیرمن ناپارامتری است. برای محاسبه ضریب همبستگی و آزمون مربوطه مسیر زیر را در SPSS انتخاب می کنیم.



در این صورت منوی Bivariate Correlation ظاهر می شود. در این منو آن دسته از متغیرهای کمی را که برای آنها قصد انجام آزمون همبستگی داریم وارد قسمت Variables می کنیم. پیش فرض SPSS انجام آزمون ضریب همبستگی پیرسن می باشد. در صورت نیاز می توان ضریب همبستگی اسپیرمن را به جای آن انتخاب نمود. با کلیک Ok ماتریس ضرایب همبستگی و میزان معنی داری آنها در خروجی ارائه می گردد. مشابه دیگر آزمون ها، فرض صفر (عدم وجود رابطه بین دو متغیر) در سطح  $\alpha$  رد می شود هرگاه میزان معنی داری آزمون کمتر از  $\alpha$  باشد.

مثال: آیا ارتباطی بین سن، وزن اولیه و قد وجود دارد؟ شدت ارتباط چقدر است؟



در خروجی نتیجه معنی داری و مقدار همبستگی ظاهر می شود.

مقدار همبستگی

Correlations

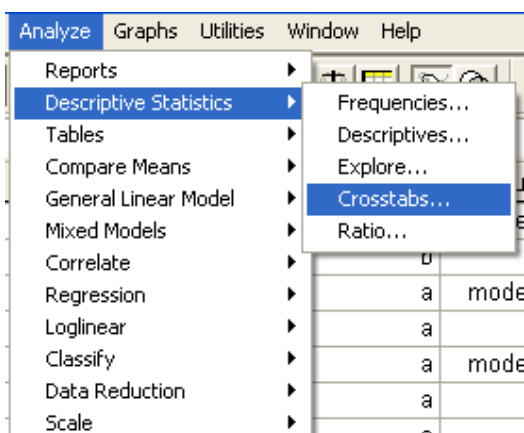
		AGE	W1	H
AGE	Pearson Correlation	1	.344*	.242
	Sig. (2-tailed)	.	.034	.144
	N	38	38	38
W1	Pearson Correlation	.344*	1	.098
	Sig. (2-tailed)	.034	.	.549
	N	38	40	40
H	Pearson Correlation	.242	.098	1
	Sig. (2-tailed)	.144	.549	.
	N	38	40	40

P-value

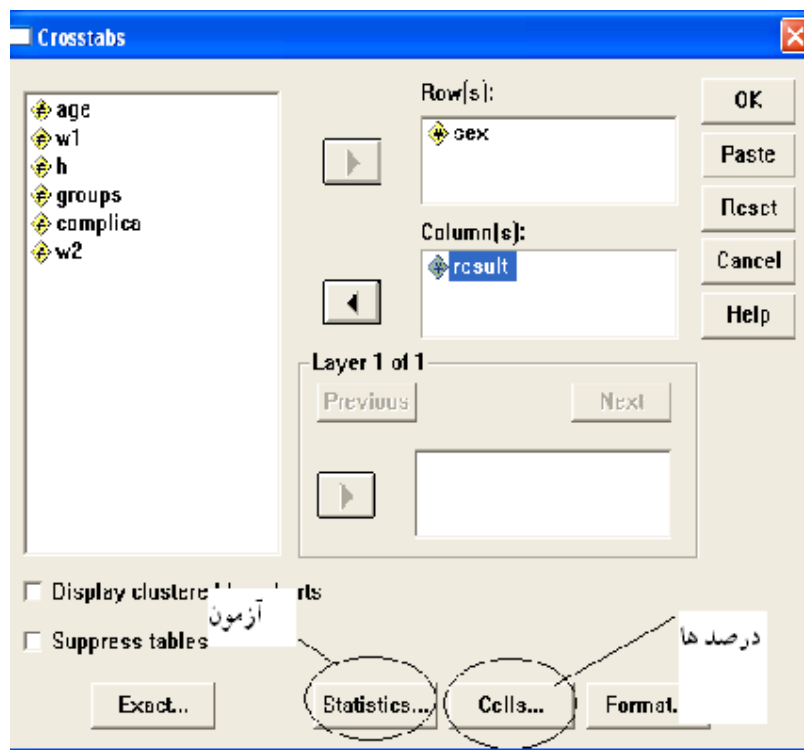
\*. Correlation is significant at the 0.05 level (2-tailed).

### آزمون استقلال جدول توافقی

جدول فراوانی توام دو متغیر کیفی که نشان دهنده وابستگی دو متغیر می باشد، جدول توافقی نامیده می شود. هدف از انجام آزمون استقلال جدول توافقی، بررسی استقلال یا وابستگی متغیرهای کیفی، سطری و ستونی می باشد. این آزمون به آزمون "کای دو" یا "خی دو" نیز مشهور است. آماره آزمون بر اساس مقایسه فراوانی های مورد انتظار و فراوانی های مشاهده شده می باشد. پیش فرض انجام آزمون این است که فراوانی های مورد انتظار در هر خانه جدول بیشتر از ۵ باشند. چنانچه فراوانی های مورد انتظار بیش از ۲۰٪ خانه های جدول کمتر از ۵ باشند مقدار معنی داری آزمون درست نخواهد بود. یکی دیگر از پیش فرض ها آن است که فراوانی مورد انتظار کمتر از یک نباشد. در این صورت یک راه این است که سطرها یا ستون هایی را که فراوانی مورد انتظار کمتر از ۵ دارند در صورت امکان با یکدیگر ادغام نمود. علاوه بر این، در صورتی که جدول توافقی یک جدول دو در دو باشد، می توان از آزمون دقیق فیشر که میزان معنی داری آن در جدول خروجی خواهد آمد استفاده نمود. چنانچه هیچ کدام از راه های قبل انجام پذیر نباشد لازم است از مدل های آماری دیگر برای بررسی رابطه استفاده نمود. برای انجام آزمون استقلال لازم است ابتدا مسیر زیر را در SPSS انتخاب کرد.



با این کار منوی Crosstabs ظاهر می گردد. در این منو لازم است یکی از متغیرهای کیفی را به قسمت Row و دیگری را به قسمت Column فراخوانی نمود. با انتخاب گزینه Cells و سپس انتخاب یکی از گزینه های Percentages می توان محاسبه درصد سطری، ستونی و یا درصد از کل فراوانی را برای خانه های جدول توافقی پیشنهاد کرد. همچنین برای انجام آزمون استقلال لازم است در منوی Crosstabs، گزینه Statistics و سپس گزینه Chi-square را انتخاب نمود.



سطر اول خروجی آزمون کای اسکور بر اساس آماره پیرسن می باشد. فرض استقلال جدول توافقی در سطح خطای ۰.۰۵٪ رد می شود هرگاه میزان تقریبی معنی داری آزمون کمتر از ۰/۰۵ باشد. در این صورت می توان گفت که دو متغیر کیفی با یکدیگر در ارتباط هستند. به این معنی که فراوانی حالات خاصی از سطوح متغیرهای سطر و ستون کمتر از حد معمول و فراوانی حالات خاصی از سطوح متغیرهای سطر و ستون بیشتر از حد معمول دیده شده است.

**RESULT \* SEX Crosstabulation**

			SEX		Total
			male	female	
RESULT	poor	Count	12	16	28
		% within SEX	60.0%	80.0%	70.0%
	moderate	Count	8	4	12
		% within SEX	40.0%	20.0%	30.0%
Total		Count	20	20	40
		% within SEX	100.0%	100.0%	100.0%

کای دو

p-value

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.905 <sup>b</sup>	1	.168		
Continuity Correction <sup>a</sup>	1.071	1	.301		
Likelihood Ratio	1.933	1	.164		
Fisher's Exact Test				.301	.150
Linear-by-Linear Association	1.857	1	.173		
N of Valid Cases	40				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.00.

شرایط:

- ۱- نباید بیش از ۲۰٪ خانه ها مورد انتظار کمتر از ۵ داشته باشد.
- ۲- حداقل مقدار مورد انتظار کمتر از یک نباشد.

در مثال فوق هر دو شرط برقرار است . در صورتی که یکی از دو شرط فوق برقرار نبود با ادغام سطر یا ستون های کنار هم ، امکان برقراری شرایط وجود دارد . در صورت عدم برقراری شوط فوق ادغام کردن را آنقدر ادامه میدهیم تا به یک جدول 2x2 برسیم (دو سطر و دو ستون) ، چنانچه هنوز هم شرط برقرار نبود از آزمون دقیق فیشر که در سطر چهارم ( جدول فوق) استفاده می نماییم.

**آزمون های ناپارامتری (Nonparametric Test)**

چنانچه در آزمون های قبل فرض نرمال بودن برقرار نباشد لازم است به جای آزمون های پارامتری فوق از آزمون های ناپارامتری معادل استفاده کرد. در حالت کلی آزمون های پارامتری به خاطر استفاده از توزیع جامعه دارای قدرت و توان بیشتری در نشان دادن یک نتیجه معنی دار هستند اما چنانچه فرضیات آزمون های پارامتری برقرار نباشند چاره ای جز استفاده از آزمون های ناپارامتری نیست. بیشتر آزمون های ناپارامتری بر اساس مقایسه میانگین رتبه ها می باشند.

## آزمون من- ویتنی (Mann-Whitney Test)

چنانچه در آزمون مقایسه میانگین در گروه مستقل، حجم نمونه کم و توزیع مشاهدات نرمال نباشد به جای استفاده از آزمون Independent sample T-Test از آزمون من- ویتنی استفاده می کنیم. چنانچه نمونه های دارای توزیع نرمال باشند توان آزمون من- ویتنی با توان آزمون t برابر خواهد شد و لذا نتایج دو آزمون در این حالت یکسان خواهند گشت. برای انجام این آزمون مسیر زیر را در SPSS انتخاب می کنیم.

### Analyze→Nonparametric Test→2Independent Samples

با این کار منوی Two Independent Samples T-Test ظاهر می گردد. در این منو، متغیر مورد آزمون را به قسمت Test Variable List و متغیر گروه بندی را به قسمت Grouping Variable فراخوانی نموده و گزینه Ok را برای انجام آزمون کلیک می نماییم. فرض برابری میانه های دو گروه در سطح خطای ۰/۰۵٪ رد می شود هر گاه مقدار تقریبی معنی داری آزمون (Asymp. Sig.) کمتر از ۰/۰۵ باشد.

## آزمون های علامت و رتبه علامت دار ویلکاکسون (Sign and Wilcoxon Signed Ranks Tests)

این آزمون معادل آزمون t زوجی در حالت ناپارامتری می باشند. در آزمون علامت اختلاف مشاهدات قبل و بعد را به دست آورده و تعداد علامت مثبت و منفی را با یکدیگر مقایسه می کنند. در آزمون رتبه علامتدار ویلکاکسون، قدرمطلق اختلافات مشاهدات قبل و بعد را مرتب نموده و رتبه گذاری می کنند سپس میانگین رتبه های مثبت و منفی با یکدیگر مقایسه می شوند. برای اجرای این آزمون ها مسیر زیر را در SPSS انتخاب می کنیم.

### Analyze→Nonparametric Test→2 Related Samples

با این کار منوی Two Related Samples Tests ظاهر می شود. در این منو لازم است دو متغیر قبل و بعد را همزمان انتخاب و به قسمت Test Pair(s) List فراخواند. پیش فرض SPSS انجام آزمون رتبه علامتدار ویلکاکسون است در صورت نیاز گزینه Sign را در قسمت Test Type برای انجام آزمون علامت انتخاب می نماییم.

## آزمون کروسکال والیس (Kruskal-Wallis Test)

این آزمون معادل آزمون آنالیز واریانس یک راهه در حالت ناپارامتری می باشد. مسیر زیر را برای انجام آزمون انتخاب می کنیم.

### Analyze→Nonparametric Test→K Independent Samples

با این کار منوی Tests for several Independent Samples ظاهر خواهد گشت. حال متغیر کمی مورد مطالعه را به قسمت Test Variable List و متغیر گروه بندی را به قسمت Grouping Variable فراخوانی می نماییم. فرضیه برابری میانه های گروه ها در سطح خطای  $\alpha$  رد می شود هر گاه Sig. آزمون، کمتر از

$\alpha$  باشد. با توجه به اینکه توان آزمون علامتدار ویلکاکسون بیشتر از توان آزمون علامت است، توصیه می شود برای انجام آزمون برابری میانگین قبل و بعد از آزمون رتبه علامتدار ویلکاکسون استفاده شود.

### رگرسیون خطی ساده (Simple Linear Regression)

در مبحث رگرسیون هدف این است که یک رابطه ریاضی بین یک یا چند متغیر مستقل و یک متغیر وابسته به منظور پیش بینی مقدار متغیر وابسته با استفاده از مقدار یا مقادیر متغیرهای مستقل بدست آوریم. بنابراین در رگرسیون، علاوه بر این که همبستگی بین چند متغیر مستقل و یک متغیر وابسته بررسی می گردد، نوع و شکل رابطه ریاضی نیز تعیین می گردد. در رگرسیون ساده خطی رابطه بین متغیر مستقل  $X$  و متغیر وابسته  $Y$  به شکل یک خط راست است که معادله آن بصورت  $y = \alpha x + \beta$  نوشته می شود.  $\alpha$  و  $\beta$  را پارامترهای رگرسیون نامیده،  $\alpha$  مقدار ثابت و  $\beta$  شیب خط می باشد. برآورد  $\alpha$  نشان می دهد که پیش بینی مقدار  $Y$  به ازای  $x=0$  چه میزانی است و برآورد  $\beta$  نشان می دهد که در ازای افزایش یک واحد به مقدار  $X$ ، مقدار  $Y$  چه اندازه تغییر می کند. در رگرسیون خطی فرض می شود که  $Y$  دارای توزیع نرمال می باشد. برای انجام رگرسیون خطی مسیر زیر را در SPSS انتخاب می کنیم.

#### Analyze → Regression → Linear

با این کار منوی Linear Regression ظاهر می شود. در این منو متغیر  $X$  را به قسمت Independent Variable و متغیر  $Y$  را به قسمت Dependent Variable وارد می کنیم. با انتخاب گزینه Ok رگرسیون اجرا می گردد. مقدار معنی داری جدول ANOVA در خروجی رگرسیون نشان می دهد که آیا رگرسیون ارزش دارد یا خیر؟ با در نظر گرفتن خطای ۰/۵، چنانچه میزان معنی داری این جدول کمتر از ۰/۰۵ باشد حاکی از اهمیت رگرسیون است. برآورد پارامترهای رگرسیون در جدول Coefficients ارائه می گردد. برآورد پارامترهای مدل در ستون B و میزان معنی داری آنها در ستون Sig. قرار دارد. برآورد مقدار ثابت مدل (برآورد  $\alpha$ ) با Constant و برآورد شیب مدل (برآورد  $\beta$ ) با X مشخص می گردد.

برای رسم خط رگرسیون ابتدا نمودار پراکنش دو متغیر  $X$  و  $Y$  را رسم می کنیم. با دو بار کلیک کردن روی نمودار پراکنش، نمودار وارد منوی جدیدی به نام SPSS Chart Editor می گردد. حال در منوی جدید گزینه Chart و سپس گزینه Options را انتخاب می کنیم. با این کار منوی Scatterplot Options ظاهر می گردد. در این منو در قسمت Fit Line گزینه Total را کلیک کرده، گزینه Ok را برای اجرا کلیک می کنیم.

توجه به این نکته ضروری است که بررسی فرضیات مدل رگرسیون که با بررسی باقیمانده های مدل صورت می گیرد از اهمیت خاصی برخوردار است. درست نبودن فرضیات مدل رگرسیون منجر به استنباط ناصحیح درباره رابطه بین متغیرهای مستقل و متغیر وابسته می گردد. علاقمندان به این مبحث می توانند به کتب آماری مناسب در زمینه رگرسیون مراجعه نمایند.

موفق باشید

حبیب الله اسماعیلی

بسمه تعالی

پرسشنامه

نام و نام خانوادگی: \_\_\_\_\_ شماره پرونده: \_\_\_\_\_

سن: \_\_\_\_\_

جنس: مذکر  مونث

وزن: \_\_\_\_\_

قد: \_\_\_\_\_

روش درمانی: روش A  روش B

عارضه: بله  خیر

نتیجه درمان: ضعیف  متوسط  خوب

وزن ثانویه: \_\_\_\_\_


#### بسمه تعالی

#### کاربرد نرم افزار SPSS در تحلیل داده ها

محقق می‌مایل است پیامد دو روش پرتودرمانی و عوارض ناشی از بکارگیری این دو روش را مورد مطالعه قرار دهد. نمونه‌ای از بیماران مورد نظر را بطور تصادفی به دو گروه تقسیم می‌کند، گروه اول با روش ۱ و گروه دوم با روش ۲ مورد مداوا قرار می‌گیرند. نتایج حاصل از اندازه‌گیری متغیرها در جدولی به شکل ذیل ثبت گردیده است.

شماره	سن	جنس	وزن اولیه	قد	گروه درمانی	نتیجه درمان	عارضه	وزن ثانویه
۱	۵۵	male	۶۰	۱۷۵	A	poor	yes	۵۸
۲	۳۰	male	۷۵	۱۶۰	B	moderate	no	۷۳



۳	۴۱	male	۸۰	۱۶۵	A	poor	no	۷۸
۴	۲۸	female	۷۰	۱۵۲	A	moderate	no	۶۴
۵	۴۴	male	۷۳	۱۵۸	A	poor	yes	۶۵
۶	-	female	۶۳	۱۶۰	A	moderate	yes	۶۰
۷	۳۸	female	۶۰	۱۵۶	A	good	no	۵۹
۸	۳۱	female	۶۵	۱۵۵	B	good	no	۶۲
۹	۲۹	female	۷۰	۱۶۵	B	good	no	۶۵
۱۰	۵۷	male	۷۵	۱۷۱	B	good	no	۷۱
۱۱	۵۱	female	۷۰	۱۵۰	B	moderate	no	۶۶
۱۲	۲۹	female	۸۰	۱۶۱	B	moderate	no	۷۹
۱۳	۵۷	male	۶۲	۱۷۵	B	poor	no	۶۹
۱۴	۲۵	male	۷۳	۱۵۹	B	moderate	no	۷۰
۱۵	۴۹	male	۵۹	۱۵۰	A	good	yes	۵۵
۱۶	-	male	۶۰	۱۶۵	A	good	no	۵۷
۱۷	۳۸	male	۶۵	۱۵۶	A	moderate	no	۶۰
۱۸	۲۷	male	۵۸	۱۶۲	B	good	no	۵۲
۱۹	۴۴	female	۶۵	۱۶۸	A	moderate	yes	۵۷
۲۰	۵۰	female	۶۵	۱۶۵	B	good	no	۵۹
۲۱	۵۶	male	۷۰	۱۷۰	A	moderate	yes	۷۵
۲۲	۳۱	male	۶۵	۱۷۵	A	poor	no	۶۰
۲۳	۴۲	female	۸۰	۱۶۰	A	moderate	no	۷۵
۲۴	۲۷	male	۶۵	۱۸۰	B	moderate	no	۶۵
۲۵	۴۳	female	۸۲	۱۵۵	B	good	no	۸۰
۲۶	۴۸	female	۶۴	۱۵۷	A	poor	no	۵۹
۲۷	۳۸	female	۵۳	۱۵۹	B	poor	yes	۵۰
۲۸	۳۵	male	۷۳	۱۷۰	B	good	no	۷۰
۲۹	۲۸	female	۶۰	۱۶۰	A	moderate	no	۵۱
۳۰	۵۶	female	۷۴	۱۶۵	A	poor	no	۶۸
۳۱	۵۰	male	۶۶	۱۶۷	A	poor	no	۶۰
۳۲	۲۹	female	۶۹	۱۶۲	B	good	yes	۶۵
۳۳	۵۲	female	۸۵	۱۶۱	B	good	yes	۸۰
۳۴	۲۷	female	۶۰	۱۵۹	B	moderate	no	۵۰
۳۵	۴۴	male	۶۹	۱۶۹	A	poor	no	۶۸
۳۶	۴۱	male	۵۰	۱۶۸	A	moderate	no	۷۳
۳۷	۳۷	male	۷۰	۱۷۰	A	poor	no	۶۲
۳۸	۲۶	female	۶۰	۱۶۰	B	good	no	۵۵
۳۹	۴۴	female	۵۷	۱۵۷	B	poor	no	۵۴
۴۰	۵۳	male	۸۳	۱۸۵	B	moderate	no	۷۸

Sex: (1=male , 2=female) group : (1=A , 2=B) result: (1=poor ,2=moderate 3=good)  
 complication (0=no, 1=yes)

جلسه اول: تعریف متغیرها و ورود داده ها

۱- نرم افزار SPSS را اجرا نموده و متغیرها را تعریف کنید.

۲- برای متغیرهای کیفی کدها را تعریف کنید.

۳- داده ها را وارد نرم افزار SPSS نمایید.

- ۴- فایل ایجاد شده را در محل مناسبی به نام Tamrin1 ذخیره نمائید.
- ۵- از SPSS خارج شده و مجدداً SPSS را اجرا کنید.
- ۶- فایل داده ها را بخوانید.

### جلسه دوم: توزیع فراوانی و شاخص های آماری

- ۱- توزیع فراوانی را برای متغیرهای جنس، گروه درمانی و عارضه بدست آورید.
- ۲- میانگین، وایانس، انحراف معیار، حداقل، حداکثر، میانه، چارک اول، چارک سوم، صدک ۹۷ و صدک سوم را برای متغیر وزن ثانویه بدست آورید.
- ۳- میانگین، انحراف معیار، حداقل و حداکثر را برای متغیرهای کمی (سن، وزن اولیه و وزن ثانویه) بدست آورید.
- ۴- برای متغیرهای جنس و نتیجه درمان نمودارستونی Bar chart رسم کنید. (درصدی)
- ۵- نتیجه درمان را در دو گروه در یک نمودارستونی نشان دهید.
- ۶- هیستوگرام داده های سن را رسم کنید.
- ۷- نمودار جعبه ای (Box-plot) را برای وزن ثانویه در دو گروه درمانی رسم کنید. (در یک نمودار)

### جلسه سوم: تبدیل داده ها

- ۱- افراد را در گروه های سنی زیر دسته بندی کنید و آن را age\_c بنامید. (طبقه بندی کردن داده های کمی)

(۲۵-۳۴)، (۳۵ - ۴۴)، (۴۵ - ۵۴)، (۵۵ - ۶۴)

۲- کدها را تعریف کنید.

۳- جدول فراوانی را بدست آورید

۴- گروه های سنی (۴۵ - ۵۴) را با گروه سنی (۵۵ - ۶۴) ادغام کنید و نام آن را age\_c2 بنامید.

۵- مجددا کدها را برای گروه بندی جدید تعریف کنید.

۶- توزیع فراوانی age\_c2 را بدست آورید.

۷- شاخص BMI را با توجه به فرمول  $BMI = \frac{weight(kg)}{(height(m))^2}$  یا  $BMI = \frac{وزن\ به\ کیلوگرم}{(قد\ به\ متر)^2}$  BMI

برای وزن اولیه محاسبه کنید و آن را BMI1 بنامید.

۸- میانگین، انحراف معیار، حداقل و حداکثر را برای BMI1 بدست آورید.

۹- با توجه به BMI1 افراد را بصورت لاغر، طبیعی و چاق طبقه بندی کنید و آن را BMI1\_C

بنامید. (فاصله ها بصورت ذیل باشد)

- کوچکتر یا مساوی ۱۸/۵ را لاغر

- بین ۱۸/۵ تا ۲۵ را طبیعی

- بیشتر از ۲۵ را چاق

۱۰- میانگین و انحراف معیار سن را برای کسانی که نتیجه درمان آنها خوب بوده است، بدست آورید.

۱۱- میانگین میانه و انحراف معیار سن را برای کسانی که چاق و نتیجه درمان آنها خوب بوده است بدست

آورید.

### جلسه چهارم: آزمون نرمالیتی و مقایسه میانگین ها در جامعه (t-test)

۱- میانگین، انحراف معیار، حداقل، حداکثر و میانه وزن اولیه را در دو گروه درمانی بدست آورید.

۲- آیا وزن اولیه، وزن ثانویه، سن و BMI دارای توزیع نرمال می باشند.

- ۳- آیا می توان گفت میانگین سن جامعه ای که از آن نمونه گرفته ایم با ۳۶ اختلاف دارد؟. (آزمون کنید)
- ۴- میانگین سن و همچنین BMI را در دو گروه درمانی مقایسه کنید. (آزمون کنید)
- ۵- میانگین وزن اولیه را در دو گروه مقایسه کنید. (آزمون کنید)
- ۶- میانگین وزن ثانویه را در دو جنس مقایسه کنید. (آزمون کنید)
- ۷- آیا بین وزن اولیه و ثانویه تفاوت وجود دارد؟
- ۸- آیا در گروه درمانی A ، تفاوتی بین وزن اولیه و ثانویه وجود دارد؟
- ۹- آیا در گروه درمانی B، تفاوتی بین وزن اولیه و ثانویه وجود دارد؟
- ۱۰- کاهش وزن را بدست آورده و بعنوان یک متغیر محاسبه و به نام dw ثبت کنید.
- ۱۱- آیا کاهش وزن در دو گروه درمانی تفاوت دارد؟
- ۱۲- داده ها را Save کنید.
- ۱۳- از SPSS خارج شده و مجدداً آن را اجرا کنید.
- ۱۴- داده ها را بخوانید و مراحل ۱ تا ۱۱ را سریعاً اجرا کنید.

### جلسه پنجم: آنالیز واریانس یک طرفه (one-way ANOVA)

- ۱- آیا میتوان گفت میانگین سن افراد جامعه در سطوح مختلف BMI\_c یکسان است. (ابتدا نرمال بودن متغیر کمی را کنترل کنید). (همگنی واریانس ها فراموش نشود)

۳- آیا کاهش وزن در رده های مختلف سنی یکسان است؟

۴- فرض کنید کاهش وزن در گروه های سنی تفاوت معنی داری دارد، نشان دهید این تفاوت مربوط به کدامیک از گروههاست. (توکی)

### همبستگی

۱- آیا بین سن و کاهش وزن همبستگی وجود دارد؟

۲- آیا بین BMI و کاهش وزن همبستگی وجود دارد؟

۳- همبستگی بین BMI و نتیجه درمان چقدر است؟

۴- همبستگی بین سطوح BMI (BMI1\_C) و نتیجه درمان چقدر است؟

جلسه ششم: آزمون کای دو

۱- آیا جنس در دو گروه درمان یکسان توزیع شده است؟  $\equiv$  ( آیا دو گروه از نظر جنس همگن هستند؟)

۲- آیا تفاوتی بین وجود عارضه در دو گروه وجود دارد؟

۳- آیا تفاوتی بین نتیجه درمان در دو گروه وجود دارد؟

۴- آیا تفاوتی بین سطوح BMI ( BMI1\_C ) در دو گروه وجود دارد؟

۵- آیا بین گروه سنی (age\_c2) و عوارض رابطه ای وجود دارد؟

۶- آیا بین سطوح اولیه BMI1\_c و گروه سنی (age\_c2) ارتباطی وجود دارد؟

۷- در افراد مذکر چه ارتباطی بین سطوح BMI1\_c و نتیجه درمان وجود دارد؟ (نتیجه را با

سوال ۴ جلسه پنجم مقایسه کنید).