پیام نوریها

public channel

✅کانال پیام نوریها در سال 95 با هدف تهیه جزوات و نمونه سوالات افتتاح و از همان ابتدای تاسیس کوشیده است با تکیه بر تلاش بی وقفه ، کارگروهی و فعالیت های بدون چشمداشت کاربران متمایز خود ، قدمی کوچک در راه پیشرفت ارائه خدمات به دانشجویان این مرز و بوم بردارد.

@Payamnoria

telegram.me/Payamnoria

رایگان است و همیشه رایگان میماند

اطلاع از اخبار و دانلود جزوات و نمونه سوالات

برای ورود به کانال تلگرامی پیام نوریها کلیک کنید

" کانال و خانواده تلگرامی پیام نوریها "

با عضویت در کانال و به آرشیو زیر دسترسی پیدا کنید

✓ تمام نمونه سوالات به روز تا آخرین دوره

✓ جزوات درسی

✓ بیش از ۱۰۰ فلش کارت دروس

✓ اخبار به روز پیام نور

✓ فیلم و فایل آموشی اختصاصی

✓ انجام انتخاب واحد و حذف و اضافه

✓ پاسخگویی به سوالات دانشجویان

✓ معرفی گروه و انجمن های پیام نوری

✓ طنز و توییت دانشجویی

به یکی از بزرگترین کانال های پیام نوری بپیوندید

برای ورود به کانال تلگرامی پیام نوریها کلیک کنید

بسم الله الرحمن الرحيم

دانشگاه پیام نور

کارگاه آمار زیستی

دکتر مسعود یارمحمدی

۱۳۹۶

فصل پنجم

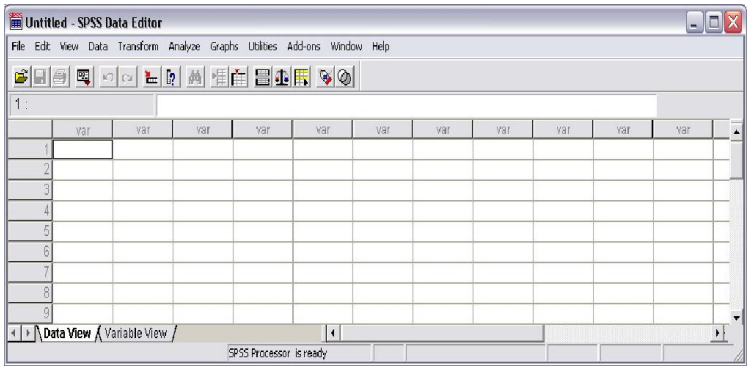
نرم‌افزار آماری SPSS

مقدمه

SPSS برنامه جامعی برای تحلیل داده‌ها می‌باشد. SPSS می تواند داده‌ها را تقریباً از هر نوعی در نظرگرفته و آنها را برای تولید گزارشات جدول دار، نمودارهای مختلف توصیفی و تجزیه وتحلیل آماری به کار گیرد. منوهای ساده و انتخاب هایی که از طریق جعبه های محاوره انجام می گیرند این امکان را فراهم می سازد که بدون نوشتن حتی یک خط دستور، تحلیلهای پیچیده ای برروی داده ها انجام گیرد. در این کتاب از نرم افزار آماری SPSS 13 استفاده می کنیم.

۵-۱ نحوه ورود داده ها

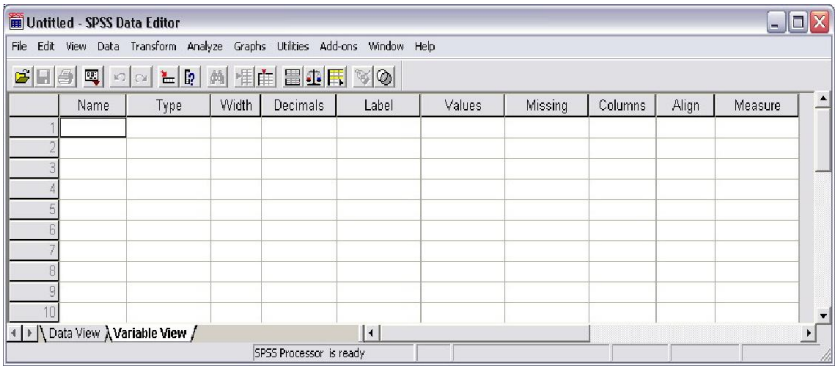
برای وارد کردن داده‌های جدید یا خواندن فایل داده‌های ذخیره شده از پنجره SPSS Data Editor استفاده می-کنیم. در شکل (۵-۱) این پنجره دو گزینه با نامهای Data View (برای ورود داده ها) و Variable View (برای معرفی متغیرها) دیده می شوند. در موقع ورود به نرم افزار SPSS ما در حالت پنجره Data View (شکل (۵-۱)) می باشیم. با انتخاب دکمه variable view پنجره ای به شکل ۵-۲ دیده می شوند.



شکل ۵-۱ پنجره Data Editor برای ورود داده ها

۵-۱-۱ پنجره Variable View

این پنجره شامل ۱۰ ستون است. برای تعریف و شناساندن یک متغیر به کامپیوتر می باید این ۱۰ مشخصه برای هر متغیر براساس نوع آن متغیر تعریف شود.



شکل ۵-۲ پنجره Variable View برای معرفی متغیرها

در زیر به تشریح این ده ویژگی می پردازیم:

نام متغیر (name):

در این گزینه محقق باید نامی را با توجه به ضوابط زیرانتخاب کند:

۱) تعداد کاراکتر حداکثر ٤٥ حرف است. اولین کاراکتر باید حرف و آخرین آن نباید نقطه باشد.از حروف فارسی می توان استفاده کرد ولی توصیه می شود این کار انجام نشود. در قسمت بر چسب می توان از الفباء فارسی استفاده کرد.

۲) نام تکراری نباشد. از اسامی:

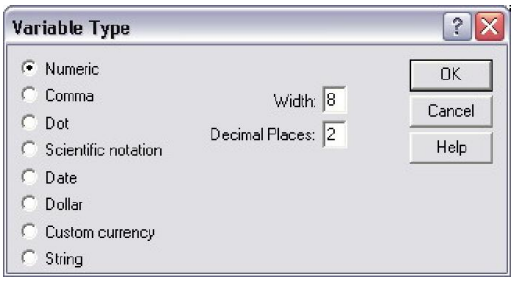
THR,TO, WITH ALL,AND,BY,EQ,GE,GT,LE,LT,NE,NOT,OR,

نباید استفاده نمود.

۳) استفاده از علائم (@ # $ ٪ ^ &) و عملگرهای ریاضی (= + - / *) غیرمجاز است.

نوع متغیر (Type)

پس از تعیین نام متغیر SPSS به طور پیش فرض سایر مشخصات را تعیین می کند. نوع بیش فرض متغیر Numeric است. با انتخاب قسمت خاکستری در این محل پنجره ای به صورت شکل ۵-۳ ظاهر می شود. شما می توانید یکی ازگزینه های مورد نظر را مشخص نموده و دکمه OK را فعال نمایید.



شکل ۵-۳ پنجره تعیین نوع متغیر

مهمترین انواع متغیرها ازبالا به پایین عبارتند از:

۱) Numeric : این گزینه برای معرفی داده های کمی به کار می رود. جمع کل ارقام صحیح و اعشاری در قسمت Width و تعداد ارقام اعشاری در قسمت Decimal Places مشخص می شود. بعنوان مثال اگر داده های ما شامل حداکثر ۵ رقم صحیح و سه رقم اعشاری باشد، در قسمت Width عدد ۸ و در قسمت Decimal Places عدد ۳ را قرار دهید.

۲) Comma : برای وارد کردن داده‌های کمی که قسمت‌های صحیح و اعشاری آنها به کمک کاما (،) از هم جدا شده اند، به‌کار می‌رود.

۳)Dot : برای وارد کردن داده های کمی که قسمت‌های صحیح و اعشاری آنها به کمک نقطه (.) از هم جدا شده‌اند.

٤) Scientific notation : برای ورود داده ها با نماد علمی به کار می رود. به عنوان مثال برای وارد کردن عدد ۲۵٦۰۰۰۰ آن را به صورت 256E+4 می توانیم معرفی کنیم.

۵) Date : این گزینه برای ورود متغیری که تاریخ را نشان می دهد با فرمت های متفاوت به کار می رود. نوع فرمت را محقق می تواند انتخاب کند.

٦) Dollar : این گزینه برای ورود داده های کمی با واحد اندازه گیری دلار به کار می رود.

۷) Custom Currency : با استفاده از این گزینه می توان پنج فرمت مختلف نرخ ارز را بیان نمود.

۸) String : این گزینه برای معرفی متغیرهای رشته ای و یا متغیرهای که حروف در آنها به کار می رود استفاده می شود.

برچسب متغیر (Label)

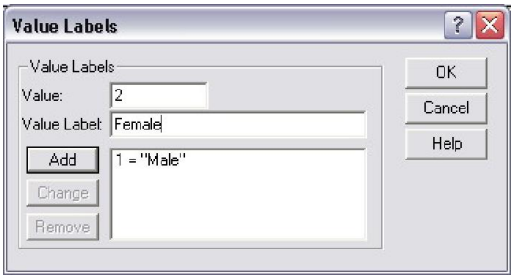
برچسب یک کد شناسایی است و می تواند برای توصیف متغیر مورد استفاده قرار گیرد. این گزینه مخصوصا برای خروجی های اجرایی از قبیل نمودارها و جداول سودمند بوده و می توان از فونتهای فارسی نیز استفاده نمود.

کدگذاری سطوح متغیر های کیفی (values)

برای کد گذاری متغیرهای کیفی با کلیک کردن در قسمت خاکستری پنجره ای به شکل ۵-٤ باز می شود. در فضای خالی مربوط به Values کد مناسب و در فضای خالی مربوط به Value Label برچسب مربوطه را وارد می کنیم.

بعنوان مثال برای معرفی متغیری مانند جنسیت در دو سطح ، زن و مرد نام متغیر را Gender و برچسب را "جنسیت" (sex) معرفی می کنیم. حال در قسمت فضای خالی value عدد یک را وارد کرده و سپس در فضای

خالی گزینه Value Label کلمهٔ "مرد (Male)" را وارد می کنیم و سپس دکمهٔ Add را کلیک کنید. مجدداً به فضای خالی مربوط به Value رفته و عدد ۲ را وارد کنید، سپس در قسمت فضای خالی Value Label عبارت "زن (Female)" را تایپ کنید و دکمه Add را فعال سازید. در قسمت پنجره پایین عبارات "Male=1" و "Female=2" را مشاهده می نمایید.

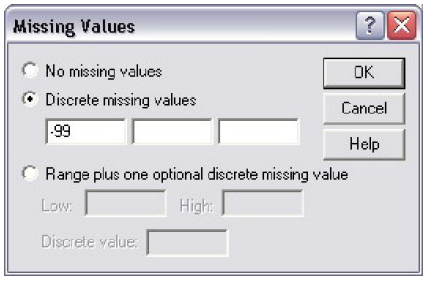


شکل ٥-٤ پنجره مربوط به کدگذاری متغیر های کیفی

در صورت اشتباه وارد نمودن، می توانید از گزینه Remove برای پاک کردن و Change برای تغییر دادن کدها استفاده کنید.

مقادیر گمشده Missing Values

اگر در داده های جمع آوری شده بعضی از متغیرها شامل مشاهدات گمشده باشد، بعنوان مثال پاسخگو به سؤالی پاسخ نداده ویا برخی از داده ها جمع آوری نشده باشد در این صورت برای آن متغیرها به طور جداگانه در محل سلولی که در تقاطع نام متغیر و ستون Missing است قرار بگیرید و سپس روی قسمت خاکستری رنگ این سلول کلیک کنید تا پنجره (شکل ٥-٥) که مربوط به معرفی داده های گمشده (Missing Values) است فعال شود. پیش فرض نرم افزار در حالت عدم وجود داده های گمشده است. بهرحال داده های گمشده را می توان به دو صورت یعنی مجزا با حداکثر سه کد متفاوت ویا با یک کد مجزا و یک دامنه در حدود مشخص تعریف نمود. برای اولین صورت در کادر Missing Values گزینهٔ Discrete Missing Values را فعال کرده و از یک کد خاص نظیر ۹۹- می توانیم استفاده کنیم. در صورت نیاز می توانیم از دو کد دیگر در محلهای خالی باقیمانده استفاده کرد. حال در موقع ورود اطلاعات هرگاه به دادهٔ گمشده ای برخورد نماییم، بجای خالی گذاشتن آن سلول از کد تعریف شده (۹۹-) استفاده می کنیم.



شکل ٥-٥ پنجره معرفی داده های گمشده

در صورت دوم ممکن است محقق بخواهد علاوه بر معرفی یک کد مجزا از یک محدوده خاص نیز برای بیان داده های گمشده استفاده کند، جهت این کار کد گزینه مربوط به Range plus one optional discrete Missing values را فعال کرده در قسمت low حد پایین و در قسمت high حد بالا محدودهٔ داده های گمشده را مشخص کرده و در قسمت Discrete کد مربوط به مشاهدات گمشده مجزا را وارد می کنیم. انتخاب هرکدام از این گزینه ها اختیاری است.

اندازه ستون نمایش دهنده متغیر (Columns)

در این ستون می توانید برای هر کدام از متغیرها، عرض ستون را در پنجرهٔ Data View مشخص نمایید.

چگونگی نمایش هر سلول (Align)

در این ستون می توانید نحوه قرارگرفتن اعداد در قسمت چپ، راست، و وسط سلول را مشخص نمایید.

مقیاس اندازه گیری متغیر (Measure)

در این ستون می توانید مقیاس اندازه گیری متغیرها را مشخص کنید. مقیاسهای Scale, Nominal , Ordinal به ترتیب برای متغیرهایی با مقیاس رتبه ای، اسمی و نسبی به کار برده می شوند.

۵-۱-۲ پنجره Data View و نحوه ورود داده های مربوط به پرسشنامه (۱)

با توجه به مطالب مطرح شده تاکنون داده های پرسشنامه (۱) در فصل دوم را پس از وارد نمودن مشخصات متغیر ها در پنجره Variable View شکل ۵-۶، در پنجره Data View شکل ۵-۷ وارد می کنیم.

porseshnameh1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	x1	String	8	0	NAME	None	-99	8	Center	Nominal
2	x2	Numeric	8	0	SEX	{0, MALE}...	-99	8	Center	Scale
3	x3	Numeric	8	0	EDU	{0, N-EDU}...	-99	8	Center	Scale
4	x4	Numeric	8	0	INCOME_RANK	{1, VERYL}...	-99	8	Center	Scale
5	x5	Numeric	8	0	BG	{1, O}...	-99	8	Center	Scale
6	x6	Numeric	8	2	HIGHT	None	-99.00	8	Center	Scale
7	x7	Numeric	8	2	WIGHT1	None	-99.00	8	Center	Scale
8	x8	Numeric	8	2	WIGHT2	None	-99.00	7	Center	Scale
9	x9	Numeric	8	2	AGE	None	-99.00	8	Center	Scale
10	x10	Numeric	8	2	BLOODP	None	-99.00	9	Center	Scale
11	x11	Numeric	8	0	INCOME	None	-99	8	Right	Scale

شکل ۵-۶ مشخصات متغیرهای پرسشنامه (۱) در پنجره Variable View

porseshnameh1.sav - SPSS Data Editor

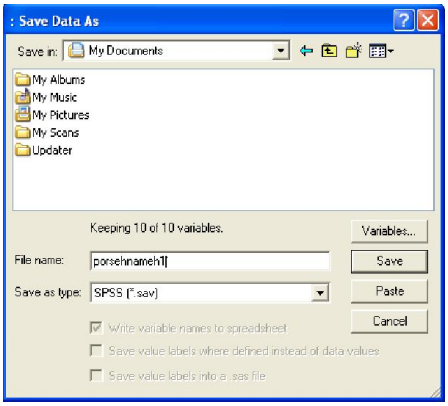
File Edit View Data Transform Analyze Graphs Utilities Window Help

20 : x11 6000

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
1	A	1	0	1	1	170.00	67.00	60.00	50.00	12.00	2000
2	B	0	1	1	1	160.00	56.00	55.00	20.00	11.00	2500
3	C	1	2	1	1	164.00	60.00	57.00	25.00	10.00	3500
4	D	0	1	2	2	177.00	76.00	73.00	30.00	13.00	2700
5	E	1	3	3	2	-99.00	46.00	44.00	17.00	10.00	4000
6	F	0	4	4	1	156.00	55.00	50.00	29.00	12.00	5000
7	G	1	3	3	3	170.00	66.00	66.00	38.00	11.00	4100
8	H	1	4	5	1	180.00	70.00	69.00	77.00	16.00	5500
9	I	1	1	1	4	170.00	65.00	64.00	80.00	16.00	2200
10	J	0	0	4	4	180.00	80.00	76.00	44.00	14.00	1800
11	K	0	1	3	3	188.00	87.00	80.00	55.00	13.00	2300
12	L	1	2	2	2	170.00	79.00	77.00	64.00	14.00	3000
13	M	1	2	1	1	177.00	78.00	79.00	29.00	9.00	3200
14	N	0	3	1	1	171.00	70.00	70.00	64.00	16.00	1900
15	O	1	2	4	2	185.00	86.00	85.00	55.00	17.00	4300
16	P	1	2	3	4	186.00	67.00	60.00	77.00	18.00	2800
17	Q	0	3	2	3	160.00	55.00	50.00	28.00	11.00	4700
18	R	1	3	1	2	171.00	70.00	66.00	29.00	12.00	5100
19	S	0	1	1	1	160.00	56.00	54.00	19.00	12.00	2300
20	T	1	4	2	1	180.00	89.00	80.00	80.00	17.00	6000

شکل ۵-۷ داده های مربوط به پرسشنامه (۱) وارد شده در پنجره Data View

حال پس از تایپ داده ها نخست داده ها را ذخیره می کنیم. جهت اینکار از منو بار صفحهٔ اصلی کلید File و سپس دکمه Save را فعال کرده، در پنجره‌ای نظیر شکل ۵-۸ در محل خالی مربوط به نام فایل (File name) نام دلخواهی را مثلاً porseshnameh1 را تایپ کرده و پس از مشخص کردن اینکه این فایل قرار است در کدام محل حافظه ذخیره شود دکمهٔ Save را فعال می کنیم. فایلی بنام porseshnameh1.sav ایجاد می شود.

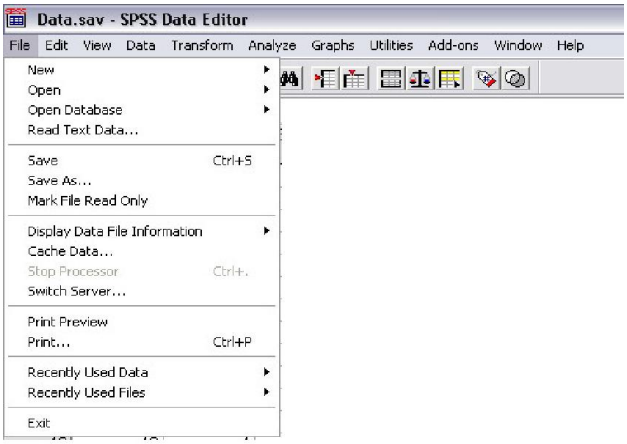


شکل ۵-۸ پنجره مربوط به ذخیره کردن فایل های ایجاد شده

۵-۲ خط فرمان (Menu Bar)

حال به طور خلاصه آیکون های مهم خط فرمان (منو بار) مربوط به پنجره اصلی که انجام کلیه فرامین اجرایی از طریق آن انجام می شود را توضیح می دهیم. این خط فرمان دارای کلیدهای زیر است:

۵-۲-۱ File (فایل) : با توجه به شکل (۵-۹) با کلیک کردن بر روی دکمهٔ File پنجره ای شامل گزینه های متفاوت ظاهر می شود که شبیه بیشتر نرم افزارهای تحت Windows است. مهمترین این گزینه ها به شرح زیر است:



شکل ۵-۹ گزینه های مختلف مربوط به پنجره File

۱) **گزینه‌های New (جدید) و Open (باز کردن)** به ترتیب برای ایجاد و بازکردن فایلهای داده ها (Data)، فرامین برنامه ای (Syntax) ، خروجی های (Output)، خروجی ها به صورت متنی ساده (Draft Output) ، و فایلهای نوع Script استفاده می شود.

۲) گزینه Read Text Data برای باز کردن فایلهای داده های ذخیره شده به صورت متنی که با پسوند dat بوده و با نرم افزارهای Word ، Notepad، و WordPad ساخته شده باشد. مشخصه این فایلها آن است که به صورت متنی بوده و با tab از هم جدا شده اند.

۳) گزینه های Save **و** Save As به ترتیب برای ذخیره نمودن فایلهای جدید و فایلهای قدیمی که تغییرات جدید در آنها ایجاد شده است به کار می رود. بعنوان مثال اگر بخواهیم فایلی را با فرمت Excel ذخیره کنیم مراحل زیر را انجام دهید:

- فرمان File-Save as را فعال کنید تا کادر محاوره ای Save Data as باز شود.
- از فهرست پنجره Save As Type گزینهٔ (*.xls) را انتخاب کنید.
- در پنجره File Name نامی دلخواه مثلا" (myfile) را تایپ نمایید.
- در پنجره Save in محل ذخیره این فایل را تعیین نمایید.
- اگر نمی خواهید اسامی متغیرها در صفحه گسترده Excel نوشته شود گزینهٔ Write Variable Names to Spreadsheet را غیر فعال سازید.
- گزینهٔ OK را فعال کنید.

٤) گزینه Display Data Info اطلاعات کاملی در مورد فایل داده، تعداد آنها و متغیرهای موجود می دهد.

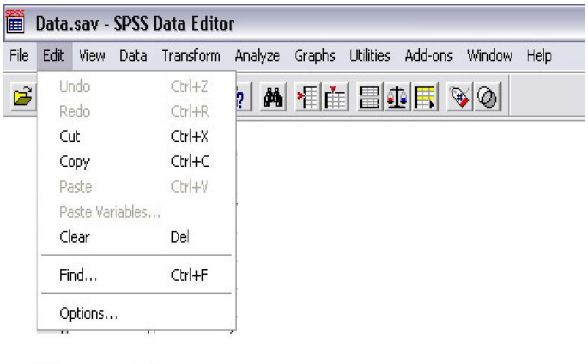
۵) گزینه های Print **و** Print Preview به ترتیب برای چاپ داده ها و نشان دادن متن ارسالی به چاپگر به کار می رود. حال اگر بخواهید تنها بخشی از فایل داده ها را چاپ کنید، لازم است آن بخش را بوسیله ماوس انتخاب کرده سپس فرمان File-Edit را فعال نموده و در کادر Print ، گزینهٔ Selection را انتخاب و دکمهٔ OK را کلیک کنید.

٦) گزینه های Recently Used data **و** Recently Used files به ترتیب برای نشان دادن آخرین داده ها و فایلهای استفاده شده به کار می رود.

۷) گزینهٔ Exit برای خروج از SPSS استفاده می شود.

۵-۲-۲ Edit (ویرایش):

با کلیک کردن بر روی این دکمه پنجره ای به شکل ۵-۱۰ باز می شود، که شامل گزینه های زیر است:



شکل ۵-۱۰ پنجره Edit (ویرایش)

۱) **گزینه های** Cut، Copy، **و** Paste که معمولاً در تمام نرم افزارهای تحت ویندوز به طور یکسان به ترتیب عملیات بریدن، کپی کردن و نصب کردن را انجام می دهد.

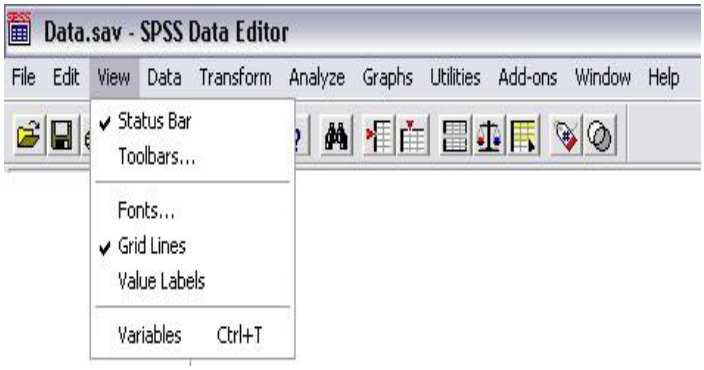
۲) **گزینهٔ** Clear برای از بین بردن متن انتخاب شده است.

۳) **گزینهٔ** Find برای چستجو و یافتن یک عدد دلخواه در متن داده ها استفاده می شود.

٤) **گزینهٔ** Option برای ایجاد انتخابهای متفاوت در بخش ویرایشی است.

۵-۲-۳ View (دیدن):

با کلیک کردن بر روی این دکمه پنجره ای به شکل (۵-۱۱) باز می شود، که شامل گزینه های زیر است:



شکل ۵-۱۱ گزینه های مختلف مربوط به پنجره View

۱)**گزینه های** Status Bar و Toolbars برای تغییر دادن نحوه منو بار است و معمولا" بهتر است در حالت پیش فرض باشد.

۲) **گزینهٔ** Fonts برای تغییر نوع فونت و اندازه آن به طور دلخواه جهت نمایش داده های پنجره Data View به کار می رود.

۳) **گزینهٔ** Grid Lines با فعال کردن این گزینه خط کشی هایی که مرزبندی بین خانه ها را نشان می دهند، پنهان یا آشکار کرد.

٤) **گزینه** Value Labels با فعال شدن آن به جای نمایش کد داده ها، برچسب تعریف شدهٔ آنها نشان داده می شود.

۵) **گزینهٔ** Variables با فعال شدن آن به پنجره Variables View می رویم.

۵-۲-٤ Data (داده ها):

با کلیک کردن بر روی این دکمه پنجره ای باز می شود، که شامل گزینه های مختلفی است که در این کتاب تنها به مهمترین آنها اشاره می کنیم:

۱) Insert Variable: اگر بخواهیم متغیری را بین دو متغیر یک فایل (مثلا" بین متغیر های دوم و سوم) قرار دهیم، یکی از خانه های ستون سوم را فعال کنید. سپس فرمان Data-Insert Variable را انتخاب کنید. این متغیر بین ستونهای مذکور ایجاد می شود. می توان با استفاده از پنجره Variable View مشخصات این متغیر را معرفی کنید.

۲) Insert Cases: اگر بخواهیم بین سطرهای دوم و سوم سطری را اضافه کنیم، یکی از خانه های ردیف سوم را انتخاب کرده و فرمان Data-Insert Cases را فعال می کنیم.

۳) Go to Case: موقعی که با فایل بزرگی سروکار داشته باشیم و بخواهیم به سطر دلخواهی برویم، در این حالت فرمان Data-Go to Case را فعال نموده و در کادر ظاهر شده شماره سطر مورد نظر را وارد نموده و دکمهٔ OK را فعال سازید.

٤) Sort Cases: گاهی مرتب کردن یک فایل داده به صورت صعودی و یا نزولی براساس یک متغیر ضرورت دارد. جهت انجام این کار فرمانِ Data-Sort Cases را فعال کنید تا کادر مربوط به Sort Cases باز شود. در سمت چپ این کادر محاوره ای اسامی تمام متغیر ها فهرست شده اند. متغیری که می خواهیم فایل داده ها بر اساس آن مرتب شوند را انتخاب کرده و با زدن دکمه آن را به بخش Sort by منتقل می کنیم. سپس در بخش Sort Order اگر بخواهیم عمل مرتب کردن به طور صعودی (نزولی) باشد، گزینهٔ Ascending (Descending) را فعال کرده و OK می کنیم. بایستی دقت کنیم که همواره فایل داده ها نسبت به یک متغیر قابل مرتب شدن است.

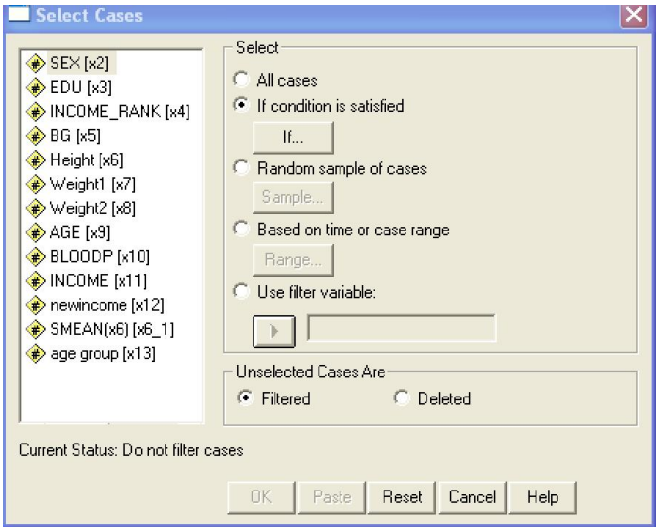
۵) **Merge Files** : برای ادغام کردن دو فایل با متغیرهای یکسان داده‌ای به دو صورت سطری و ستونی.

٦) **Split File** :گاهی نیاز است که پردازش براساس مقادیر خاصی از متغیرهای گروه‌بندی شده مانند جنسیت، گروه‌های خونی، گروه‌های نژادی و غیره، اجرا شود. در چنین مواردی، بهتر است که فایل داده ها را برحسب سطوح آن متغیر تجزیه کنیم و هر بار پردازش را برمبنای یکی از کدها (رده‌ها)ی آن متغیر انجام دهیم.

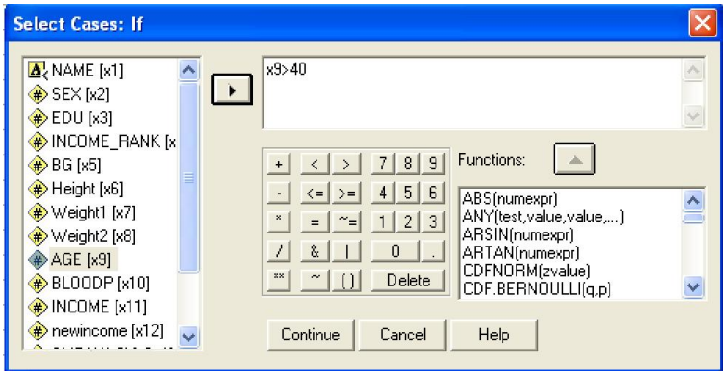
۱۲) *Select Cases*

الف- انتخاب گزینه *All Cases* (که همواره پیش فرض است) باعث انتخاب همه موارد (سطرهای داده) می‌شود.
ب- با انتخاب گزینه *If condition is satisfied*، تنها سطرهایی از داده انتخاب می شود که در شرط خاصی صدق کنند.

بعنوان مثال، اگر بخواهیم یک توصیف آماری تنها بر روی اطلاعات مربوط به افرادی با سن بالای ٤٠ سال انجام گیرد، کافی است که بعد از فعال کردن این گزینه (شکل ۵-۱۲-الف) و فشردن دکمه *if...* ، متغیر *age* را به پنجره مقابل انتقال داده و آن را بزرگتر از ٤٠ قرار داده (شکل ۵-۱۲- ب) ، سپس بر روی کلیدهای *Continue* و *Ok* کلیک کنیم. بعد از اجرای این فرمان و در صفحه *Data Veiw*، بر روی رکوردهایی که در محاسبات لحاظ نمی‌شوند خطی مورب کشیده می‌شود (ضمنا"، یک متغیر جدید در فایل داده‌ها بنام *filter_$* ایجاد می‌شود).



شکل ۵- ۱۲ (الف)



شکل ۵-۱۲ (ب)

ج- با انتخاب گزینه *Random Sample of cases* و فشردن کلید *Sample*، پنجره نمونه تصادفی باز می‌شود (شکل ۵-۱۲-ج)

شکل ۵-۱۲-ج

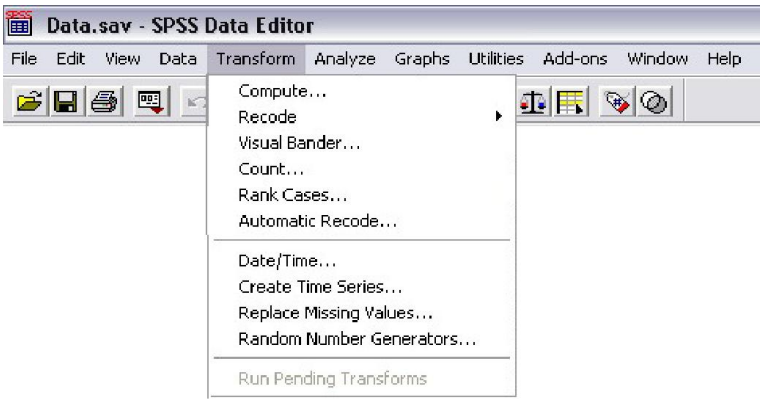
در این پنجره و در قسمت *Approximately* می‌توانید هر درصد دلخواهی از مشاهدات را برای انتخاب تصادفی از میان آنها قرار داده و در قسمت *Exactly* می‌توان دقیقا" تعداد مشخصی از سطرها (که در قسمت خالی اول قرار داده می‌شود) از بین n مشاهده اول (که در قسمت خالی دوم قرار می‌گیرد) به تصادف انتخاب نمود.

د- با انتخاب گزینه *Based on time or Case range* و فشردن دکمه *Range* می‌توان سطرهایی از داده‌ها را که در دامنهٔ مشخصی قرار دارند، انتخاب نمود.

ه – انتخاب گزینه *Use Filter Variable*، برای وقتی است که بخواهید سطرهای فایل داده را بر حسب یک متغیر فیلتری دو حالتی (مانند جنسیت که فقط دو مقدار ۰ و ۱ را می‌گیرد) انتخاب کنید. با فعال کردن این گزینه و انتقال متغیر فیلتری به فضای خالی و سپس فعال نمودن دکمه *OK*، بر روی شماره سطرهایی که شرط مذکور را ندارند، خط کشیده شده و در محاسبات منظور نمی‌شوند (انتخاب گزینه *Deleted* سبب می‌شود که سطرهایی که شرط مورد نظر را ندارند، حذف شوند).

۵-۲-۵ تبدیلات مربوط به داده ها (Transform)

گاهی نیاز است که تبدیلات خاصی بر روی متغیرهای موجود در یک فایل از داده ها بعمل آید. این تبدیلات به طور خلاصه تحت فرمان خاص تبدیل (Transform) در منو اصلی به صورت های مختلف قابل اجرا است. با کلیک کردن بر روی دکمه Transform پنجره ای به شکل ۵-۱۳ باز می شود، که شامل گزینه های مختلفی است که در این کتاب تنها به مهمترین آنها اشاره می کنیم:



شکل ۵-۱۳ گزینه های مختلف مربوط به پنجره Transform

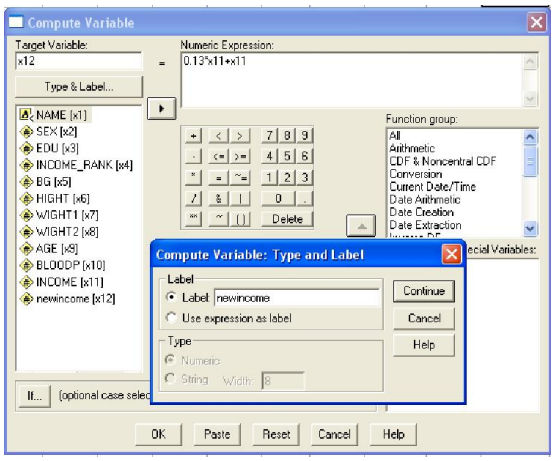
مهمترین گزینه های مربوط به پنجره Transform به شرح زیر است:

۱) دستور محاسبه (Compute):

با استفاده از این دستور می توان با استفاده از متغیرهای موجود متغیر جدیدی را تولید کنیم. به عنوان مثال، فرض کنید بخواهیم میزان درآمد افراد را براساس افزایش ۱۳ درصدی از درآمد گذشته محاسبه و تحت نام متغیر جدیدی بنام Newincome ذخیره نماییم. به عبارت دیگر داریم:

$$\text{Newincome} = (0.13)\text{income} + \text{income}$$

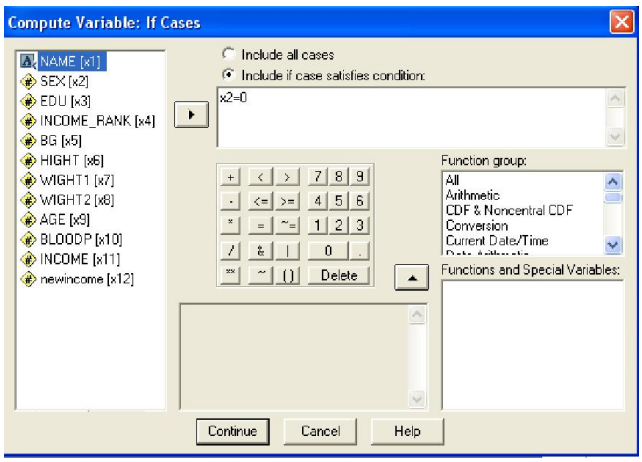
جهت اینکار از منو بار اصلی فرمان Transform را فعال کرده گزینه Compute را فعال کرده ودر پنجره ظاهر شده در قسمت Target Variable نام متغیر جدید x12 با برچسب Newincome را وارد نموده و در پنجره Numeric Expression فرمول محاسباتی درآمد جدید را نظیر شکل (۵-۱٤) وارد کرده و سپس دکمهٔ OK را فعال می نماییم.



شکل ۵-۱۴ پنجره مربوط به ساختن متغیر جدید Income

در نتیجه یک ستون جدید بنام x12 و با برچسب Newincome در پنجره داده ها ایجاد می شود.

اگر بخواهیم که این تغییر درآمد فقط برای زنها اعمال شود, کافی است که در پنجره Compute Variable دکمه If را در بالای دکمه Ok را فعال سازیم. پنجره جدیدی به شکل ۵-۱۵ باز می‌شود. حال گزینه Include if Case satisfies Condition را فعال ساخته عبارت sex=0 (کد جنسیت برای زنان) را در جعبه مستطیل شکل یادداشت کنید. سپس دکمه Continue فعال کنید. حال اگر بخواهید این عمل به عنوان نام یک متغیر جدید ایجاد شود، نام آن متغیر را در قسمت Target Variable وارد نمایید, و سپس Ok را کلیک نمایید. در غیر این صورت تغییرات جدید روی متغیر Newincome اعمال می گردد.



شکل ۵-۱۵ محدود کردن نوع تبدیل با استفاده از If شرطی

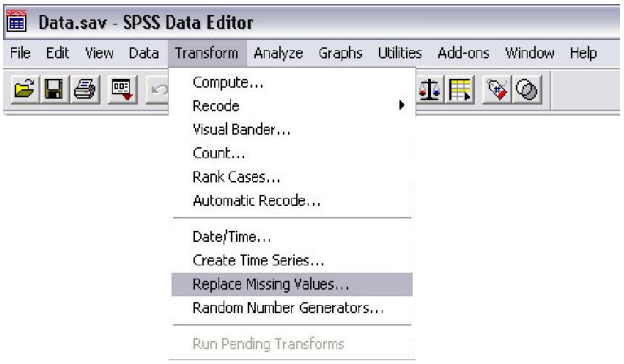
همانطوری که در شکل ۵-۱۵ ملاحظه می فرمایید در قسمت پنجره Function group، تعداد قابل توجهی از توابع حسابی (Arithmetic) ، مقادیر احتمال تجمعی توزیع‌های نرمال, برنولی, بتا, دوجمله ای, و سایر توزیع

های مهم، توابع مثلثاتی, تولید داده‌های تصادفی و غیره، نیز در این پنجره موجود است، که در صورت نیاز قابل استفاده می‌باشد.

۲) دستور Recode برای کد گذاری مجدد یک متغیر کیفی بکار می رود. در بخش ۵-٤-۳ نحوهٔ کاربرد این دستور را در یک مثال کاربردی نشان می دهیم.

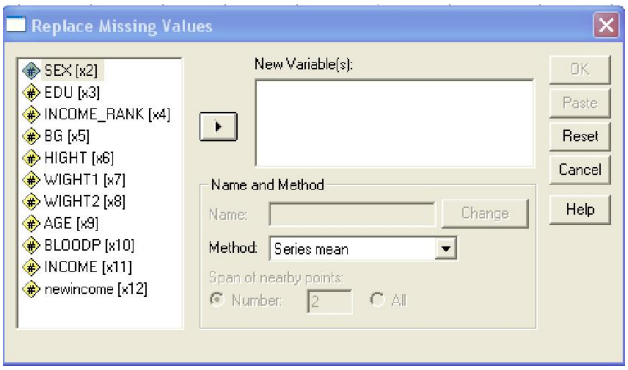
۳) دستور جایگزین کردن مقادیر گمشده (Replace Missing Values…)

اگر در جریان ایجاد فایلها، مقادیری به عنوان مقادیر گمشده تعریف شوند و بعداً قصد جایگزینی آنها را با مقادیر دیگری داشته باشیم، دستور…Replace Missing Values در منوی Transform را فعال می کنیم. بدین صورت که از منوی Transform بر روی …Replace Missing Values نظیر شکل (۵-۱٦) کلیک می نمائیم:



شکل ۵-۱٦ طریقه اجرای دستور جایگزین کردن مقادیر گمشده

که با اجرای فرمان …Replace Missing Values، پنجره شکل ۵-۱۷ باز می شود:



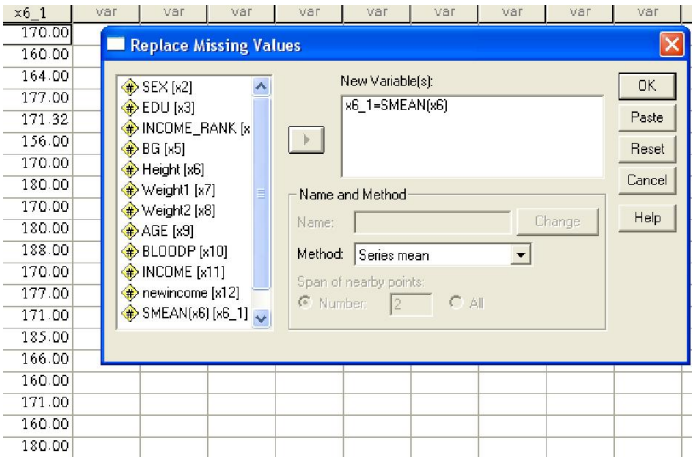
شکل ۵-۱۷ پنجره مربوط به دستور Replace Missing Values

متغیری را که قرار است داده گمشده اش جایگزین گردد به قسمت New Variable منتقل کرده و سپس از جعبه Method روش جایگزینی مقادیر گمشده را انتخاب می کنیم. پیش فرض برنامه استفاده از میانگین مقادیر

این متغیر است. به طور کلی در جعبه Method، روشهای زیر جهت جایگزینی مقادیر گمشده وجود دارند که به ترتیب عبارتند از:

- روش Series Mean (میانگین سری داده ها): در این روش؛ میانگین کلیه مقادیر معتبر متغیر مورد نظر، جایگزین کلیه مقادیر گمشده می شود.
- روش Mean of nearby points (میانگین نقاط نزدیک هم): در این روش؛ میانگین، تنها با شرکت نقاط نزدیک به نقطه گمشده محاسبه می شود. فاصله ای که میانگین در آن محاسبه می شود در پاسخ به جمله Span of Nearby points تعیین می شود: گزینه Number که تعداد معینی را تعیین می کند و All که سبب شرکت تمام داده هاست.
- روش Median of nearby points (میانه نقاط نزدیک هم): این روش کاملاً مشابه روش قبلی است با این تفاوت که بجای میانگین از میانه نقاط انتخابی استفاده می شود.
- روش Linear intorpolation (درونیابی خطی): در این روش، بین اولین نقطه معتبر قبل از نقطه گمشده درونیابی خطی انجام می شود و حاصل، جایگزین مقدار گمشده می شود.
- روش Linear trend and point (یافتن روند خطی نقاط): در این روش، یک روند خطی برحسب کلیه مشاهدات معتبر متغیر و اندیس دهی به مقادیر از کوچک به بزرگ تعیین می شود. اندیس نقطه گمشده در این روند خطی، فرضی محاسبه شده و مقدار نقطه معادل آن، جایگزین مقدار گمشده می شود.

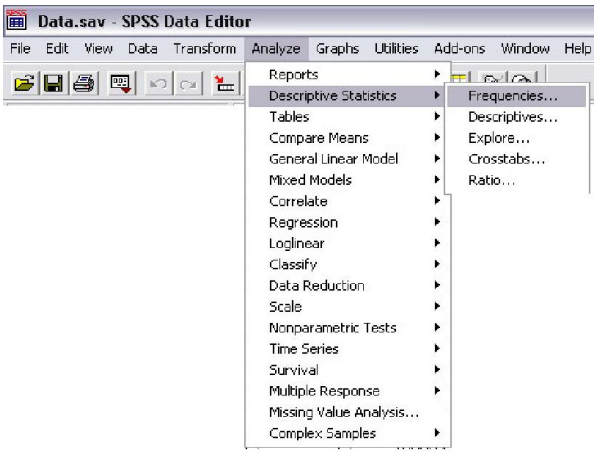
بعنوان مثال، اگر بخواهیم مقدار گمشده متغیر سن (Age) را با میانگین سری داده ها جایگزین کنیم، نظیر شکل ۵-۱۸ ابتدا در پنجره Replace Missing Values... متغیر سن (Age) را وارد بخش New Variable(s) نموده و از بخش Name and Method، گزینه Mean of nearby points را انتخاب و سپس بر روی دکمه OK کلیک می نمائیم. متغیر جدید x6_1 متغیری است که مشاهده گمشده مربوط به فرد پنجم با عدد ۱۷۱/۳۲ که میانگین دیگر افراد است جایگزین شده است.



شکل ۵-۱۸ پنجره مربوط به مقادیر گمشده

۵-۳ آمار توصیفی

برای اجرای آمار توصیفی در نرم افزار SPSS13 کافی است از زیرشاخه Analyze بر روی Descriptive Statistics کلیک نمائید شکل ۵-۱۹ ظاهر می شود.

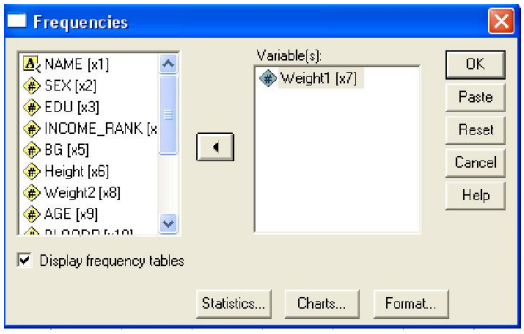


شکل ۵-۱۹ اجرای دستور Analyze

این زیرشاخه خود شامل بخشهای مختلفی است که مهمترین آنها عبارتند از:

۵-۳-۱ فراوانی‌ها (Frequencies…)

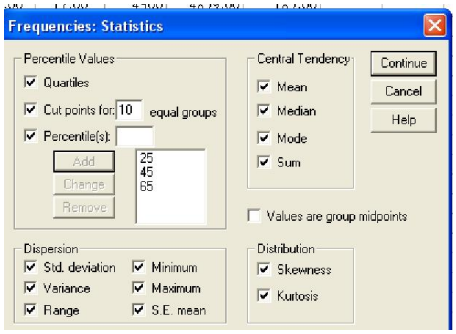
فرمان Frequencies…، آمار توصیفی مربوط به متغیرها را بطور تکی یا گروهی محاسبه، ترسیم یا جدولبندی می نماید. بعنوان مثال، برای محاسبه آمار توصیفی مربوط به متغیر وزن افراد متغیر مربوطه یعنی (Weight1) را از لیست متغیرها انتخاب و به پنجره Varable(s) منتقل می کنیم. شکل ۵-۲۰.



شکل ۵-۲۰ نحوه اجرای آمار توصیفی فراوانی

Statistics…

اگر روی دکمه Statistics کلیک کنید، کادر شکل ۵-۲۱ باز می شود. در این کادر می‌توان هر یک از آماره های مورد نیاز را به شرح زیر فعال کنید.



شکل۵-۲۱ کادر Frequencies

۱) Quartiles مقادیر چارکهای اول تا سوم را می دهد.

۲) Cut Pointd ضریب نقطه برش را مشخص می کند. مثلا اگر ۱۰ انتخاب شود صدکهای مرتبه ۱۰، ۲۰، ۳۰،...و غیره که همگی ضریبی از ۱۰ هستند را می دهد.

۳) Percentiles هر صدک دلخواه را می دهد. به عنوان مثال برای یافتن صدکهای ۲۵، ٤۵، و ٦۵ ام وارد کردن عدد ۲۵ در جعبه مقابل Percentile(s) و سپس با کلیک کردن روی دکمه Add صدکهای دیگر را در جعبه مشخص می کنیم.

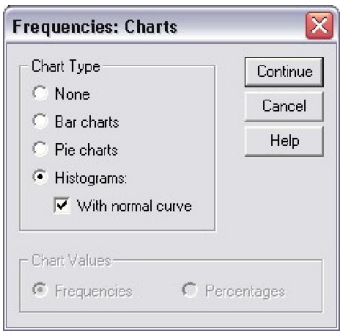
٤) Central Tendency معیارهای متمایل به مرکز داده ها شامل میانگین, میانه، نما و مجموع را می دهد.

۵) Dispersion معیارهای پراکندگی شامل انحراف معیار، واریانس، دامنه، کوچکترین مشاهده، بزرگترین مشاهده و انحراف معیار از میانگین را می دهد.

٦) Distribution معیاری مربوط به توزیع شامل ضریب چولگی و ضریب کشیدگی را می‌دهد.

Charts…

حال اگر بخواهیم علاوه بر مقادیر آماره های مختلف متغیر درآمد، یکی از نمودارهای میله ای (Barchart)، دایره ای (Pie Chart) و یا بافت نگار (Histogram) را رسم نمایم، کافی است در پنجره …Frequencies، وارد قسمت …Charts در شکل ۵-۲۲ شده و نمودار مورد نظر را انتخاب می‌نمائیم. در صورت انتخاب گزینه مربوط به بافت‌نگار با کلیک کردن بر روی دکمه With Normal Curve نمودار توزیع نرمال نیز رسم می‌شود.

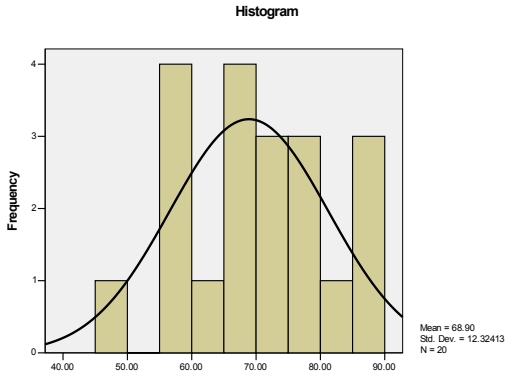


شکل ۵-۲۲ گزینه های Charts

به عنوان مثال برای متغیر وزن Weight نتایج در جدول ۵-۱ و شکل ۵-۲۳ آمده است.

20	Valid	N
0	Missing	
68.9000	Mean	
2.75576	Std. Error of Mean	
68.5000	Median	
55.00(a)	Mode	
12.32413	Std. Deviation	
151.884	Variance	
-.004	Skewness	
.512	Std. Error of Skewness	
-.686	Kurtosis	
.992	Std. Error of Kurtosis	
44.00	Range	
45.00	Minimum	
89.00	Maximum	
1378.00	Sum	
55.0000	10	
55.2000	20	
57.0000	25	
61.5000	30	
66.4000	40	
67.0000	45	
68.5000	50	Percentiles
70.0000	60	
73.9000	65	
77.4000	70	
78.7500	75	
79.8000	80	
87.9000	90	

جدول ۵-۱ نتایج آمار توصیفی متغیر وزن افراد



شکل ۵-۲۳ نمودار بافت نگار مربوط به متغیر وزن

Format…

با کلیک کردن روی این دکمه پنجره ای باز می شود که شامل این گزینه هاست:

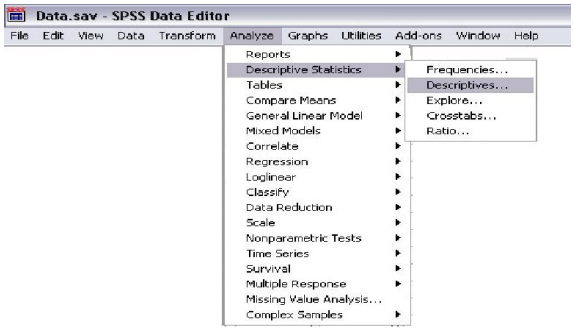
۱) Order by نحوه مرتب کردن سطرهای مختلف جدول فراوانی را تعیین میکند.

۲) اگر چند متغیر وارد کرده اید گزینه Compare Variables سبب می شود که آماره های توصیفی برای مقایسه در یک جدول ثبت شود.

۵-۳-۲ توصیفات (Descriptives…)

برای اجرای این دستور نظیر شکل ۵-۲٤ گزینه

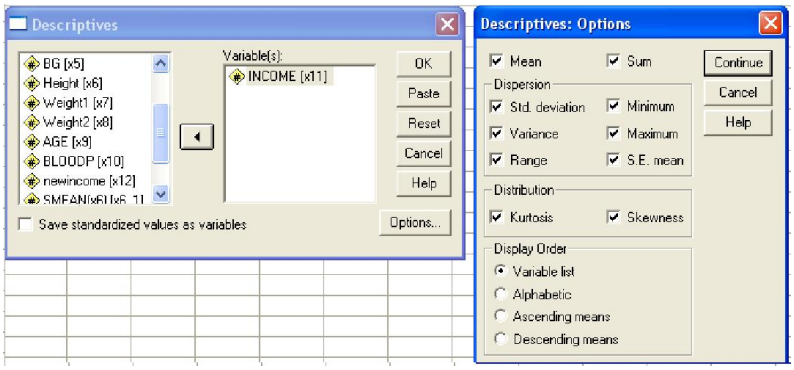
Analyze –Descriptive Statistics–Descriptive… را انتخاب می کنیم.



شکل ۵-۲٤ نحوه اجرای فرمان Descriptives…

برخلاف فرمان Frequencies… که در ابتدا به آن اشاره شد، فرمان Descriptives…، تنها آماره های توصیفی متغیرها را محاسبه و نمایش داده و بیشتر جنبه مقایسه ای میان چندین متغیر از نظر مقدار کمیتهای مرکزی و پراکندگی را دارد. به عنوان مثال، برای محاسبه آمار توصیفی متغیر درآمد (Income) این متغیر را به بخش variable(s) منتقل می نمائیم.

در پنجره Descriptives، گزینه Save standardized values as variables، امکان ساخت و ذخیره سازی متغیرهای استاندارد شده از متغیرهای اصلی را فراهم می کند. لازم به ذکر است که در این پنجره، کلید Options… ، نوع آماره هایی که برای متغیرهای منتخب محاسبه می شود و ترتیب قرار گرفتن آنها را در خروجی، تعیین می نماید. با فعال کردن این کلید پنجره نظیر شکل ۵-۲۵ باز می شود:

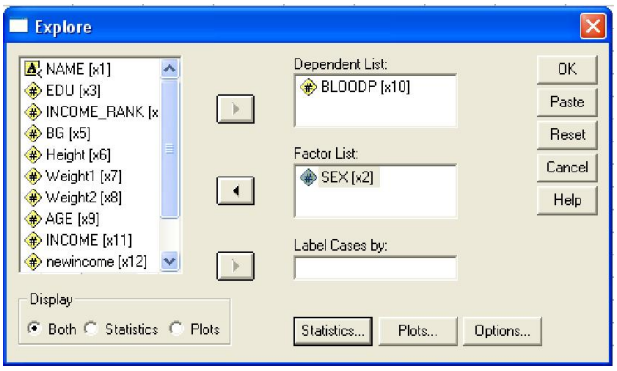


شکل ۵-۲۵

با انتخاب دکمه Continue و سپس دکمه Ok می توان به نتایج توصیف آماری رسید.

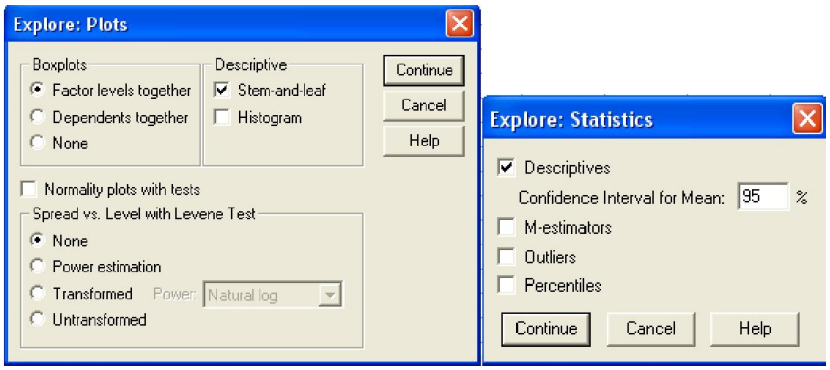
۵-۳-۳ اکتشاف ...Explore

با انجام این فرمان کادر شکل (۵-۲۶) باز می شود.



شکل ۵-۲۶ پنجره مربوط به Explore

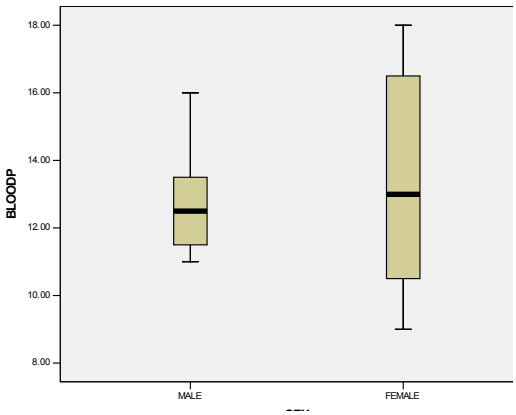
نام یک (یا چند) متغیر را به قسمت Depencet List منتقل کنید.(مثلا متغیر BloodP). سپس نام یک متغیر گروه بندی مثلاً (Sex) به قسمت Factor List منتقل نمایید. با کلیک کردن دکمه های Statistics و سپس Plots (شکل (۵-۲۷) می توانیم گزینه های مورد نظر را فعال کرده و Ok نماییم.



شکل ۵-۲۷ گزینه های Statistics و Plots مربوط به دستور Explore

انتخاب Descriptives آماره های توصیفی متغیر منتقل شده به بخش Depenedent List را فراهم می سازد. با استفاده از گزینه Plots می توان نمودارهای جعبه ای (Boxplots) و ساقه و برگ (Stem-and-leaf) متغیر وابسته برای متغیر گروه بندی را رسم نمود.

به عنوان مثال نمودار جعبه ای مربوط به میزان فشارخون در دو گروه مردان و زنان در شکل(۵-۲۸ الف) و نمودار ساقه و برگ متناظر با آن در شکل(۵-۲۸ ب) آمده است.



شکل (۵–۲۸ الف) نمودار جعبه ای فشار خون برای مردان و زنان

Stem-and-Leaf Plots

BLOODP Stem-and-Leaf Plot for
x2= MALE

Frequency	Stem & Leaf
7.00	1 . 1122334
1.00	1 . 6

Stem width: 10.00
Each leaf: 1 case(s)

BLOODP Stem-and-Leaf Plot for
x2= FEMALE

Frequency	Stem & Leaf
1.00	0 . 9
6.00	1 . 001224
5.00	1 . 66778

Stem width: 10.00
Each leaf: 1 case(s)

شکل(۵– ۲۸ب) نمودار ساقه و برگ فشار خون برای مردان و زنان

۵–۳–٤ فرمان Crosstabs…

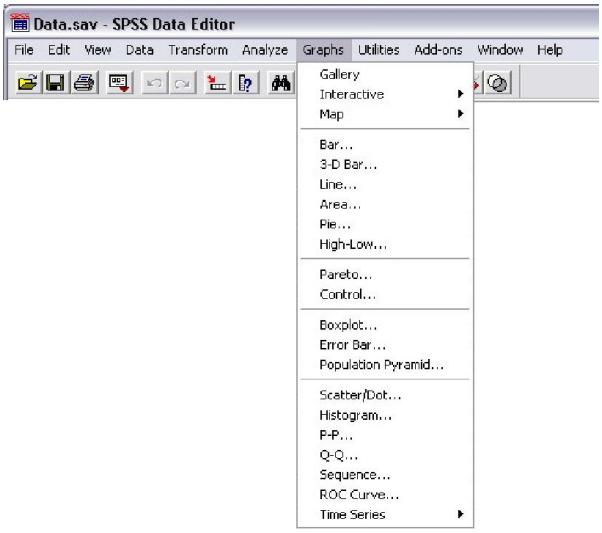
از این فرمان برای ساختن و تحلیل جداول توافقی استفاده می شود. جداول توافقی یا متقاطع، در مورد متغیرهای گروهبندی کننده (کیفی اسمی و یا کیفی رتبه ای) کار برد پیدا می کند. به عنوان مثال اگر بخواهیم ارتباط دو متغیر جنسیت و گروه خونی را برای داده های پرسشنامه (۱) بسنجیم به این فرمان نیازمندیم. با توجه به اهمیت و کاربرد این جداول مخصوصا برای تحلیل پرسشنامه های آماری، بحث مربوط به این موضوع به آینده موکول می گردد.

۵–۳–۵ فرمان Ratio…

این فرمان لیست کاملی از آماره ها را در اختیار کاربر می گذارد که برای توصیف نسبت بین دو متغیر کمی استفاده می شوند. در صورت لزوم خوانندگان محترم می‌توانند به قسمت Help ویا مرجع [1] مراجعه نمایند.

۵-٤ رسم نمودارها

یکی از روشهای مهم و اساسی برای توصیف داده ها، استفاده از رسم نمودارهای مختلف است که حتی در آمار استنباطی نیز می توانند به محقق جهت تصور و متعاقبا تحلیل بهتر از جامعه کمک نمایند. در نرم افزار آماری SPSS برای ترسیم یک نمودار آماری، می توان به راحتی با ورود به منوی Graphs و انتخاب نمودار مورد نظر، این کار را انجام داد. شکل ۵-۲۹.



شکل ۵-۲۹ منوی Graphs و فرمانهای مختلف برای رسم انواع نمودار

فرمانهای مختلفی برای رسم انواع نمودار وجود دارد که عبارتند از:

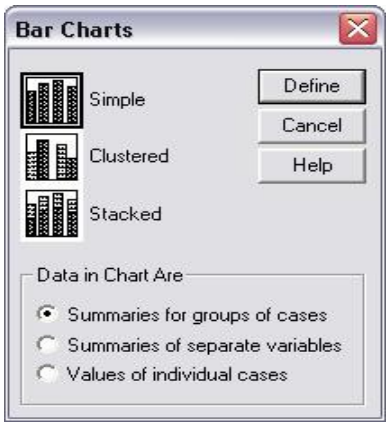
فرمان Gallery

این فرمان صرفا انواع نمودارهای مختلف موجود در SPSS و روش ترسیم آنها را نشان می دهد.با کلیک کردن روی هر کدام از کلید های مشخصه نمودار می توانید از راهنمای کمکی فراهم شده جهت ترسیم داده ها استفاده کرد. دسته ای از این نمودارها برای ترسیم داده های کیفی و دسته دیگر برای رسم داده های کمی طراحی شده است.

الف) نمودار های مربوط به داده های کیفی

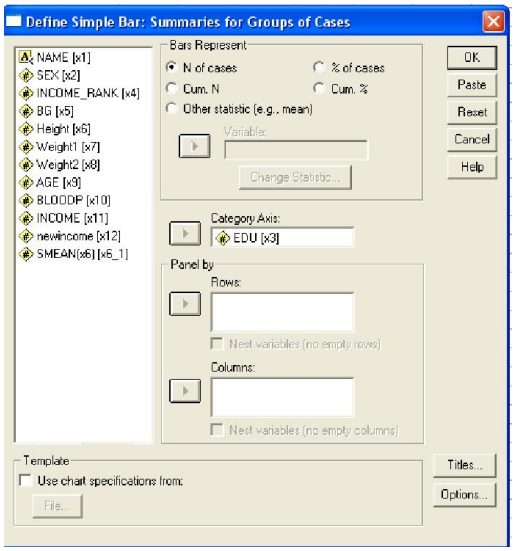
۵-٤-۱ نمودار میله ای ...Graphs-Bar

در این نمودار برای هر رده، مشاهده یا متغیر، بر روی محور افقی، میله ای ترسیم می شود. ارتفاع میله ها می تواند تابعی از خود متغیر رده بندی یا متغیر خلاصه شده دیگری باشد. ترسیم اینگونه نمودارها با فعال نمودن فرمان ...Bar از منوی Graphs قابل انجام می باشد، که همانطوری در شکل ۵-۳۰ دیده می شود امکان انتخاب یکی از سه نوع نمودار ستونی (Simple)، خوشه بندی شده (clustered)، و پشته ای (Stacked) را به کاربر می دهد.

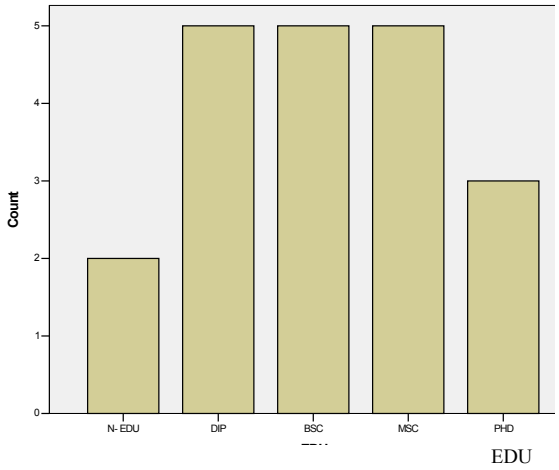


شکل ۵-۳۰ نحوه انتخاب نوع نمودار میله ای

بعنوان مثال اگر بخواهیم نمودار فراوانی یا فراوانی نسبی مربوط به متغیر میزان سواد را رسم کنیم، بعد از انتخاب گزینه Simple ، بر روی دکمه Define کلیک می کنیم. پنجره ای نظیر شکل (۵-۳۱) باز می شود. با وارد کردن متغیر میزان سواد Edu(x3) در قسمت Category Axis و زدن دکمه OK نمودار فراوانی رسم می شود، شکل (۵-۳۲). در قسمت Bars Represent گزینه % of cases برای رسم فراوانی نسبی، گزینه Cum. N برای رسم فراوانی تجمعی، و گزینه Cum % برای رسم فراوانی نسبی تجمعی، به کار می رود. جلوتر کاربرد گزینه Other Statistic(e.g. mean) را با یک مثال عملی نشان می دهیم.



شکل ۵-۳۱ پنجره مربوط به منوی Simple Bar

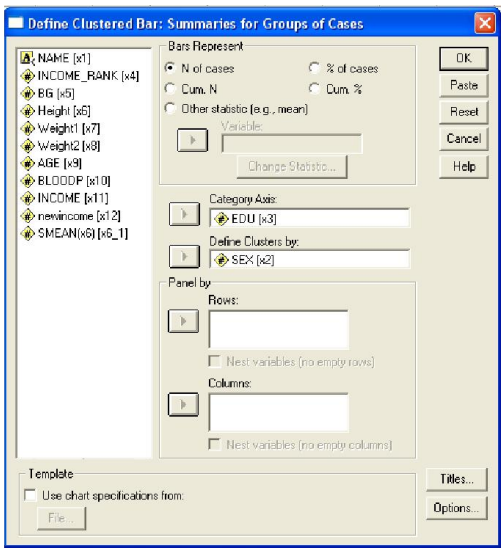


شکل ۵-۳۲ نمودار فراوانی میزان سواد (EDU)

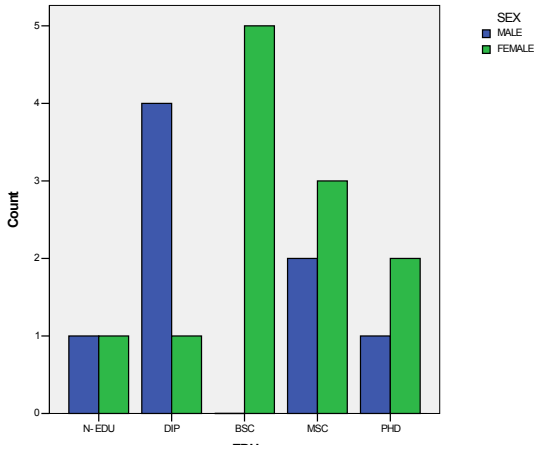
با دو بار کلیک کردن بر روی این نمودار، پنجره Chart Eidotr باز می شود. در این پنجره می توان تغییرات زیادی روی نمودار های تولید شده اعمال کرد. به عنوان مثال می توان تعداد فراوانی ها را در روی نمودار مشخص نمود و یا اینکه رنگ آنها را عوض کرده و یا آنها را سه بعدی نمود.

ب) نمودار میله ای برای دو متغیرکیفی

در صورتی قصد داشته باشیم نمودار یک متغیر کیفی (مثلا میزان سواد) را در مقابل یک متغیر کیفی دیگر (جنسیت) مقایسه کرد مراحل زیر را انجام می دهیم. نخست از منوی Graphs گزینه Bar...، را فعال کرده و روی گزینه Clustered قرار گرفته و دکمه Define را فعال کنید. حال در پنجره Define Clustered Bar متغیر EDu(x3) را در بخش Category Axis و متغیر Sex [x2] را در بخش Define Clusters by نظیر شکل ۵-۳۳ وارد کرده و با فعال کردن OK نمودار را رسم می کنیم که نتیجه در شکل ۵-۳٤ آمده است.

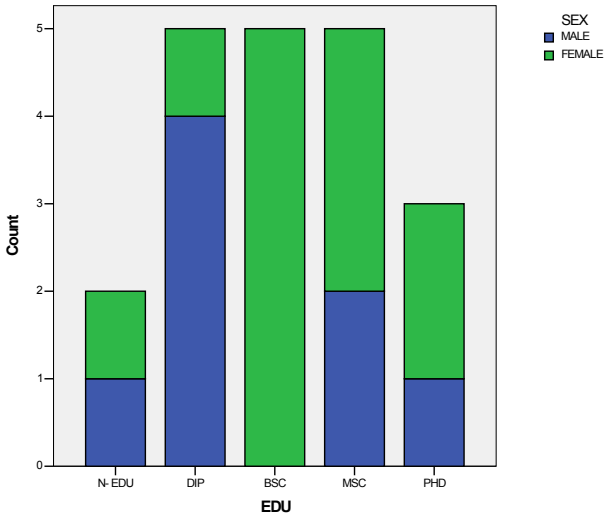


شکل ۵-۳۳ پنجره رسم نمودار دو متغیر کیفی



شکل۵-۳٤ نمودار خوشه بندی شده میزان سواد برای مردان و زنان

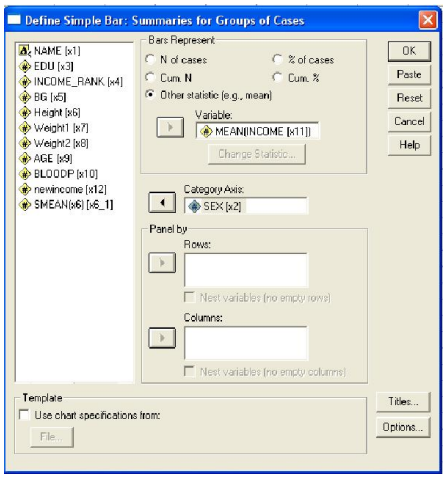
تذکر: اگر در پنجره Bar Charts، از گزینه Stacked به جای Clustered استفاده کرده و بقیه مراحل را نظیر حالت قبل عمل نماییم، نموداری نظیر شکل ۵-۳۵ را خواهیم داشت.



شکل ۵-۳۵ نمودار پشته ای میزان سواد برای مردان و زنان

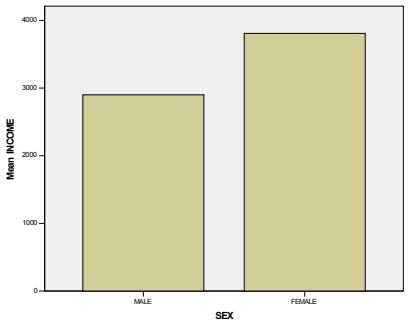
ج) نمودار میله ای آماره یک متغیر پیوسته نسبت به یک متغیر کیفی

با مراجعه به اطلاعات پرسشنامه (۱) فرض کنید بخواهیم میانگین درآمد مردان و زنان را با یکدیگر مقایسه نمائیم، با انتخاب گزینه Simple از پنجره ..Graph- Bar و کلیک بر روی دکمه Define پنجره ای نظیر شکل ۵-۳٦ ظاهر می شود.



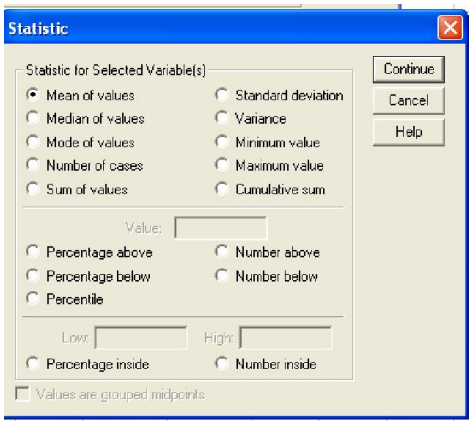
شکل ۵-۳۶ رسم آماره یک متغیر کمی برای یک متغیر کیفی

با وارد نمودن متغیر در آمد (income) در بخش Variable که زیر بخش مربوط به other Statistic(e.g. mean) می باشد و متغیر جنسیت (Sex) را در قسمت Category Axis و زدن کلید OK نمودار نظیر شکل ۵-۳۷ رسم می شود.



شکل ۵-۳۷ مقایسه میانگین درآمد برای گروه مردان و زنان

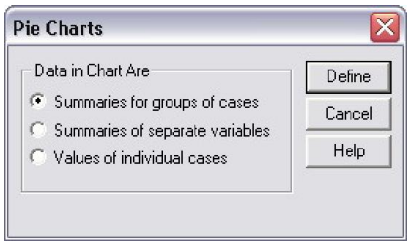
اگر بخواهیم بجای میانگین از معیار آماری دیگری نظیر میانه، مد، و ... استفاده کنید، دکمه Change Statistic... را در موقع ورود متغیر به قسمت Variable فعال کرده پنجره‌ای شبیه شکل ۵-۳۸ حاصل می شود. گزینه مورد نظر را انتخاب و دکمه continue را فعال سازید.



شکل ۵-۳۸ محل انتخاب آماره های مختلف

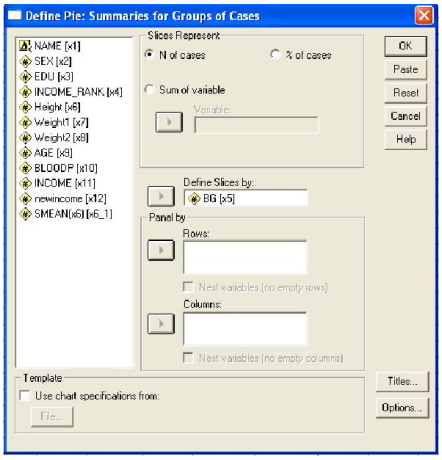
٥-٤-٢ نمودار دایره ای (Pie Chart)

از این فرمان برای رسم نمودار دایره ای استفاده می شود. این نمودار غالبا برای متغیرهای کیفی اسمی و یا رتبه ای مورد استفاده قرار می گیرد. برای رسم این نمودار با انتخاب گزینه Pie... از منوی Graphs، پنجره شکل ٥-٣٩ باز می شود:



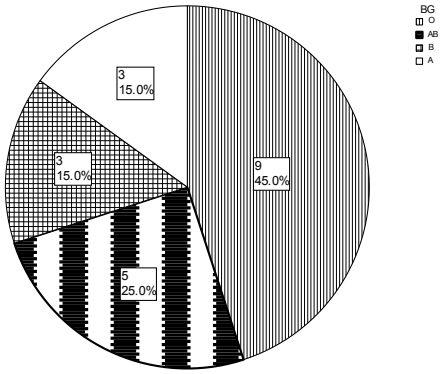
شکل ٥-٣٩

به عنوان مثال، اگر بخواهیم فراوانی و درصد گروه های مختلف خونی را به کمک نمودار دایره ای نشان دهیم ، با انتخاب گزینه Summaries for groups of cases از پنجره Pie Chart و کلیک بر روی دکمه Define پنجره‌ای نظیر شکل ٥-٤٠ باز می‌شود. با وارد نمودن متغیر درآمد (BG) در بخش Define Slices by و زدن کلید OK نمودار دایره‌های مورد نظر در قسمت خروجی برنامه ظاهر می شود.



شکل ٥-٤٠

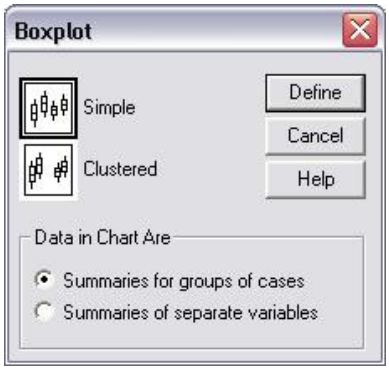
با دو بار کلیلک روی این نمودار و انتقال آن به پنجره Chart Editor تغییرات لازم اعمال شده و در نهایت در شکل ٥-٤١ رسم می شود.



شکل ۵-۴۱ فراوانی و فراوانی نسبی در نمودار دایره ای به نوع گروه خونی

۵-٤-۳ نمودار جعبه ای(Box plot)

این نوع نمودارها، برای نمایش پراکندگی داده ها و مکان‌های مربوط به کوچکترین و بزرگترین مشاهده و چارک-های اول، دوم و سوم و نیز بررسی تقارن در توزیع داده بکار می رود. از این نمودار می‌توان به منظور مقایسه یک متغیر کمی در رده های مختلف با یک یا دو متغیر کیفی استفاده کرد. برای رسم این نمودار کافی است از منوی Graphs بر روی گزینه ...Boxplot کلیک کنید. (شکل ۵-٤۲).



شکل ۵-٤۲ پنجره Boxplot

فرض کنید بخواهیم پراکندگی میزان فشار خون افراد را نسبت به جنسیت و گروه سنی (زیر ٤۰ سال و بالای ٤۰ سال) مقایسه کنیم. برای انجام این کار نخست نیازمند هستیم که یک متغیر جدید کیفی اسمی که نشان دهنده گروه سنی زیر ٤۰ سال و بالای ٤۰ سال است، بسازیم. این کار توسط یک دستور ساده بنام Recode که زیر فرمان Transforms است قابل انجام است.

دستور Recode

۱) از منو بار transform، فرمان Recode را انتخاب کنید. دو گزینه ظاهر می شود:

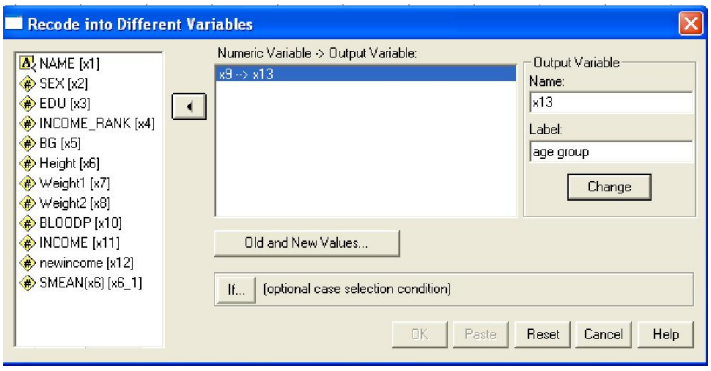
الف: ...Into Same Variable

از بین بردن متغیر قبلی و جایگزین نمودن متغیر recode شده بجای آن

ب: ...Into different Variable

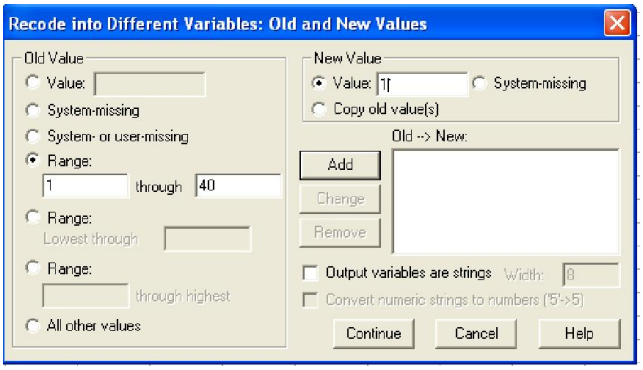
تولید یک متغیر جدید با نام دلخواه که معمولا" گزینه دوم توصیه می شود. با انتخاب این گزینه پنجره ای نظیر شکل ۵-۴۳ باز می شود.

حال فرض کنید بخواهیم متغیری بنام Age group بسازیم که یک متغیر کیفی با دو حالت ۱ (نشان دهنده سن زیر ۴۰ سال) و ۲ (نشان دهنده سن بالای ۴۰ سال) باشد. جهت اینکار متغیر age (x9) را به قسمت Input variable منتقل می کنیم.اینکار باعث می شود که x9->? در پنجره ایجاد شده و برنامه از ما نام متغیر خروجی را درقسمت Name و برچسب آن را در قسمت Label می خواهد.نام آن را x13 و برچسب آن را age group تایپ کنید و سپس دکمه Change را فعال کنید، ملاحظه می کنید که در قسمت پنجره قبلی بجای علامت ؟ نام x13 می نشیند.



شکل ۵-۴۳ پنجره Recode یا تولید متغیر های جدید کیفی

حال برای تولید متغیر x13، دکمه Old and New Values را فعال نمایید. پنجره ای نظیر شکل ۵-۴۴ ظاهر می شود.



شکل ۵-۴۴ پنجره تولید متغیر جدید کیفی

در قسمت Old Value قسمت range عدد ۱ و در قسمت through عدد ۴۰ را تایپ کنید. حال به قسمت new Value رفته و در قسمت Value عدد ۱ را بعنوان کد سن زیر چهل سال معرفی کرده و دکمه فعال Add را کلیک کنید. خواهید دید که در قسمت Old--> New عبارت 1 thru 40-->1 ظاهر می شود. همین کار را

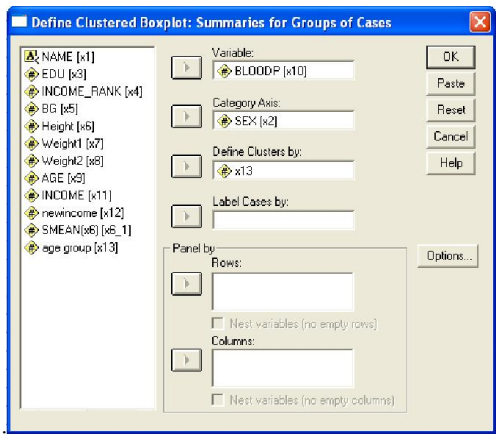
برای معرفی کد ۲ برای سن بین ٤١ تا ٨٠ سال تکرار کنید تا عبارت 41 thru 80→2 در زیر عبارت قبلی بیاید. دکمه Continue و سپس دکمه ok را فعال کنید، متغیر جدید x13 ایجاد می شود.

نمودار جعبه ای

همانطوری که در قسمت قبل مطرح شد فرض کنید بخواهیم توسط نمودار جعبه ای، پراکندگی میزان فشار خون افراد را نسبت به جنسیت و گروه سنی (زیر ٤٠ سال و بالای ٤٠ سال) مقایسه کنیم. جهت اینکار مراحل زیر را انجام دهید:

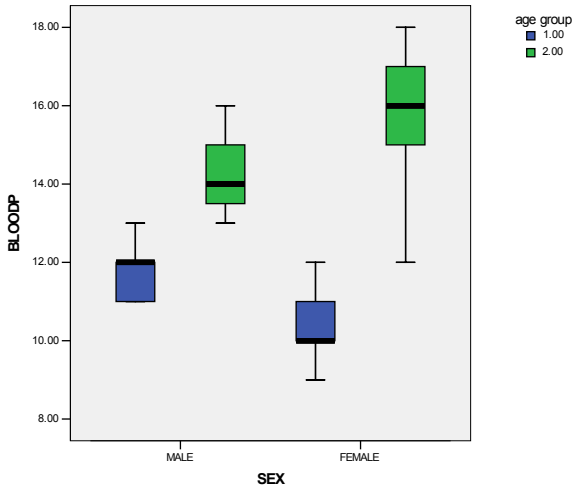
١) از منوی Graph-Box plot در شکل ٥-٤٢ گزینه clustered را انتخاب کرده و دکمهdefine را فعال کنید. پنجره ای نظیر شکل ٥-٤٥ ظاهر می شود.

٢) متغیرهای Bloodp ، Sex و x13 را به ترتیب به قسمت های variable، Category Axis و define Clusters by منتقل کنید.



شکل ٥-٤٥ پنجره Clustered Boxplot

سپس دکمه ok را فعال نمایید شکل ٥-٤٦ حاصل کار است.



شکل ٥-٤٦ نمودار جعبه ای فشار خون نسبت به سن و جنسیت

۵-٤-٤ نمودار پراکنش Graphs- Scatter/Dot…

از این فرمان برای رسم نمودار پراکنش داده ها استفاده می شود، تا نوع رابطه آنها را حدس بزنند. این نمودار در بحث مدلهای رگرسیون ساده خطی که در آینده مورد بررسی قرار می گیرد از اهمیت زیادی برخوردار است. با اجرای این فرمان کادر شکل ۵-٤۷ باز می شود. که شامل پنج گزینه زیر است:

۱) Simple Scatter : برای رسم نمودار پراکنش دو متغیر کمی پیوسته.

۲) Matrix Scatter : برای رسم نمودار پراکنش بیش از دو متغیر به صورت ماتریسی

۳) Simple Dot : برای رسم نمودار نقطه ای ساده.

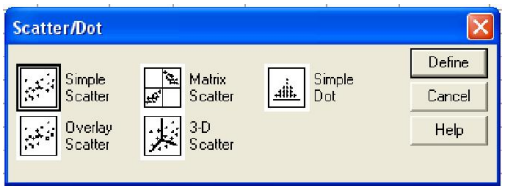
٤) Overlay Scatter : برای رسم نمودار پراکنش روی هم

۵) 3-D Scatter : برای رسم نمودار سه بعدی

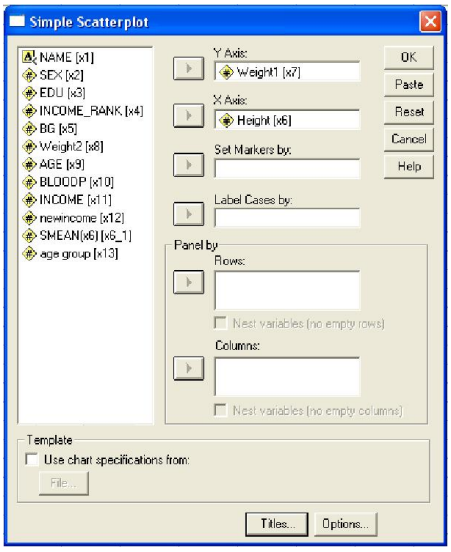
رسم نمودار پراکنش دو بعدی

به عنوان مثال فرض کنید بخواهیم نمودار پراکنش متغیرهای وزن نسبت به قد را رسم کنیم. برای این منظور مراحل زیر را انجام می دهیم.

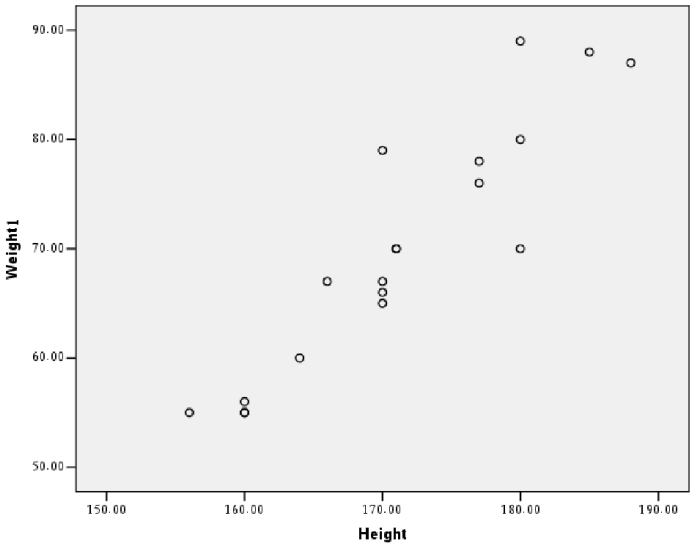
۱- فرمان Graphs-Scatter/Dot… را فعال کنید تا کادر شکل ۵-٤۷ باز شود. روی آیکون Simple سپس دکمه Define کلیک کنید، کادر شکل ۵-٤۸ باز می شود. متغیرهای weight1 و Height را به ترتیب به پنجره های Yxis و Xxis منتقل کنید و روی دکمه Ok کلیک کنید. نمودار شکل ۵-٤۹ حاصل می شود.



شکل ۵-٤۷ کادر Scatter/Dot

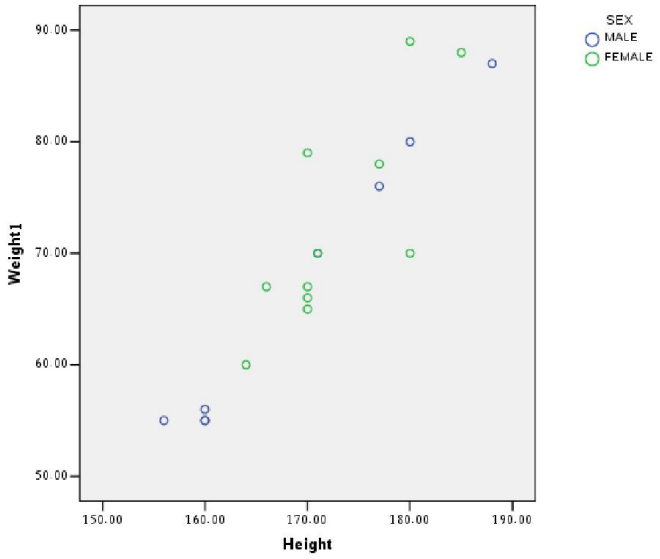


شکل ۵-٤۸ کادر مربوط به Simple Scatterplot



شکل ۵-۴۹ نمودار پراکنش وزن نسبت به قد

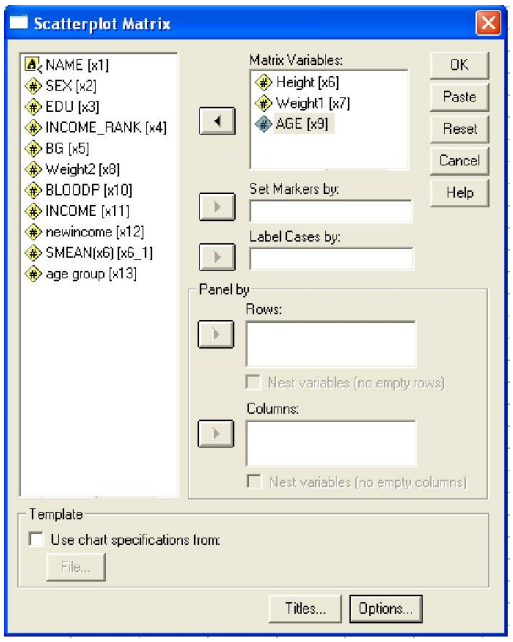
در صورتی‌که بخواهیم نمودار پراکنش برحسب رده‌های مختلف متغیر جنسیت با رنگ خاص نشان داده شود، کافی است در پنجره Simple Scatter plot شکل ۵-۴۸، متغیر Sex را وارد بخش Set Markers by نمائیم. نمودار شکل ۵-۵۰ حاصل می شود.



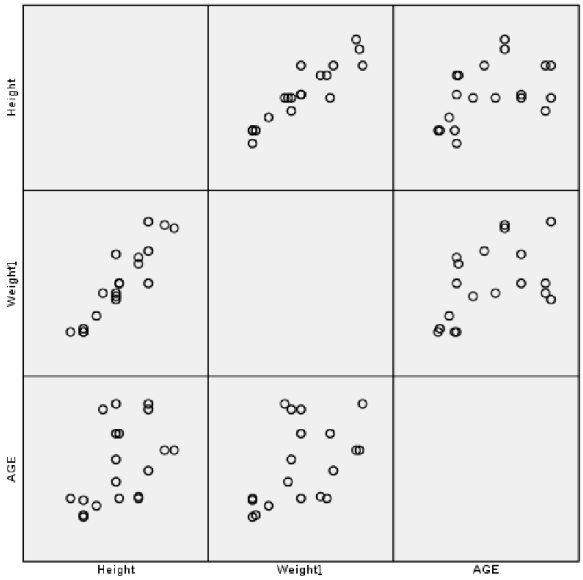
شکل ۵-۵۰ نمودار وزن نسبت به قد بر حسب متغیر جنسیت

رسم نمودار ماتریسی

برای رسم نمودار ماتریسی سه متغیر سن، وزن، و قد از منوی Graph-Scatter/Dot... کادر شکل ۵-۴۷ گزینه Matrix Scatter را انتخاب و سپس دکمه Define را فعال می کنیم. پنجره ای نظیر شکل ۵-۵۱ باز می شود. متغیرهای مربوط به قد (height)، وزن (Weight1)، سن (age) را به کادر مربوط به Matrix Variable منتقل کرده و دکمه OK را فعال می کنیم. نمودار ماتریسی در شکل ۵-۵۲ آمده است.



شکل ۵-۵۱ پنجره مربوط به Scatter plot Matrix



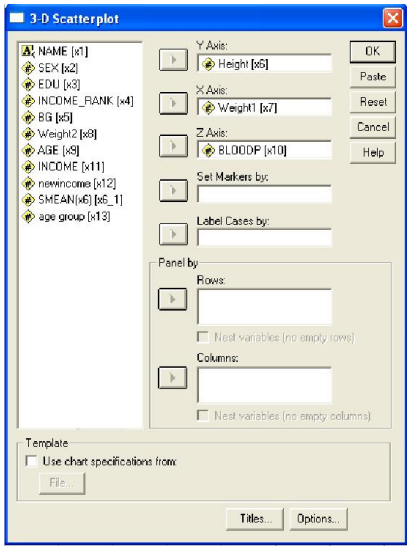
شکل ۵-۵۲ نمودار ماتریسی برای سه متغیر وزن، قد و سن

تذکر: در صورتی‌که بخواهیم نمودار پراکنش برحسب رده‌های مختلف یک متغیر کیفی و با رنگ خاص نشان داده شود، کافی است در پنجره Simple Scatter plot شکل ۱-۵۱، متغیر مربوطه را وارد بخش Set Markers by نمائیم.

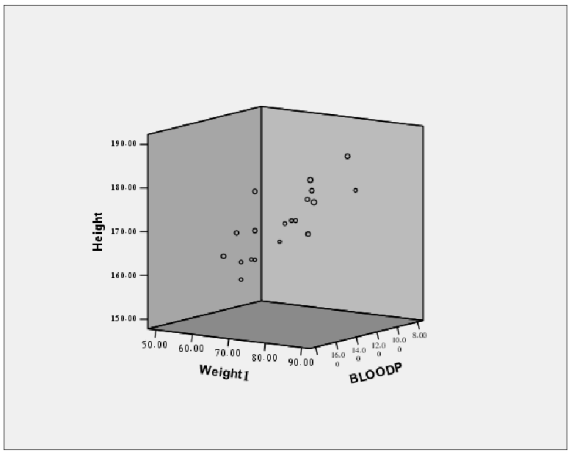
نمودار سه بعدی

برای رسم نمودار سه بعدی مربوط به سه متغیر سن، وزن، و قد از منوی Graph-Scatter/Dot... کادر در شکل ۵-۴۷ گزینه 3-D Scatter را انتخاب کرده و سپس دکمه Define را فعال می کنیم. پنجره ای نظیر

شکل ۵-۵۳ باز می شود. متغیرهای مربوط به قد (height)، وزن (Weight1)، فشار خون (BloodP) را به ترتیب به کادر های YAxis، XAxis، و ZAxis منتقل کرده و دکمه OK را فعال می کنیم. نمودار ماتریسی در شکل ۵-۵٤ آمده است.



شکل ۵-۵۳ کادر 3-D Scatterplot



شکل ۵-۵٤ نمودار سه بعدی مربوط به متغیرهای قد، وزن و سن

فصل سیزدهم
تجزیه و تحلیل آماری داده ها در نرم افزار آماری SPSS

مقدمه

نرم افزار آماری SPSS برنامهٔ جامعی برای تجزیه و تحلیل آماری داده ها می باشد. SPSS می تواند هرنوع داده ای از هر نوع فایل بخصوص گرفته و آنها را برای تولید گزارشات شامل توصیفات آماری مناسب و تحلیل های پیچیده مورد نیاز به کار گیرد. منوهای ساده وانتخابهایی که از طریق پنجره های محاوره ای انجام می گیرند این امکان را فراهم می آورند که بدون تایپ کردن حتی یک خط دستور تحلیلهای پیچیده ای بر روی داده ها انجام گیرد. در این فصل کاربرد نرم افزار SPSS برای انجام موارد مشروحه زیر با ذکر مثالهای کاربردی معرفی می کنیم:

- انجام آزمون فرضهای آماری مربوط به میانگین (ها) در یک جامعه نرمال شامل آزمون برابری میانگین یک جامعه با یک عدد دلخواه، آزمون برابری میانگین های دو جامعه مستقل، آزمون برابری میانگین های دو جامعه وابسته (زوج شده)، آزمون برابری میانگین های چند جامعه مستقل نرمال (تحلیل واریانس یکطرفه).
- انجام تحلیل واریانس دو طرفه شامل طرح بلوکهای تصادفی و تحلیل واریانس دوطرفه (عامله) با اثر متقابل.
- ضریب همبستگی پیرسن (برای دو متغیر کمی) و ضریب همبستگی اسپیرمن (برای دو متغیر کیفی رتبه ای) و رگرسیون ساده خطی.
- معیارهای همایندی برای دو متغیر کیفی اسمی، دو متغیر کیفی رتبه ای، و یک متغیر کیفی اسمی و یک متغیر کمی و آزمون استقلال در جداول توافقی

۱۳-۱ آزمون میانگین یک جامعه نرمال

مثال ۱۳-۱. محققی مدعی است که میانگین آلبومین در گردش خون بر حسب گرم مردان سالم ۲۰ تا ۲۹ ساله، ۱۳۲ گرم است، ادعای او را در سطح معنی داری $\alpha = 0/05$ مورد بررسی قرار دهید.

۱۱٦، ۱٤٦،۱۳٦، ۱۱۹، ۱۰٦، ۱۱۸، ۱۱۸، ۱۵۳، ۱۲۳، ۱۲۲، ۱۳۹، ۱۱٦، ۱۲۷، ۱۰٦، ۱٤۵، ۱۲۹، ۱۲۰، ۱۲۲، ۱۳۰، ۱۱٤»، ۱٤٦، ۱۳۳، ۱۲٤، ۱٤۱، ۱۳۳، ۱۳۱، ۱۳۳، ۱٤٤، ۱٤٦، ۱۳۳، ۱٤۱

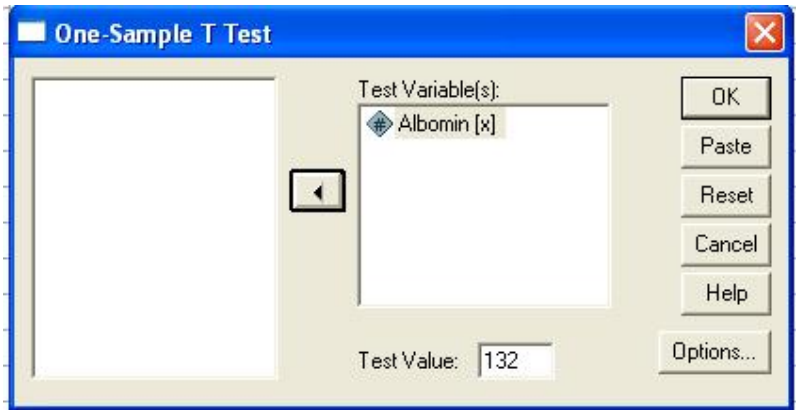
$H_0 : \mu = 132$

$H_1 : \mu \neq 132$

برای انجام این آزمون نخست داده ها (متغیر x) را وارد می کنیم. سپس فرمان

Analyze – Compare Means – One Sample T test

را وارد کنید تا پنجره مربوط به آن باز شود شکل (۱۳-۱). متغیر x را بـه بخـش test variable(s) وارد کرده و در محل test Value عدد ۱۳۲ را تایپ نماییـد. فاصـله اطمینـان ۹۵ در صـد بـرای میـانگین بـه صورت پیش فرض محاسبه می شود. در صورتی که بخواهید میزان سطح اطمینان را تغییر دهید کافی است دکمه Option را فعال کرده و تغییرات لازم را اعمال نمایید.



شکل (۱۳-۱)

با فشردن دکمه ok خروجی زیر را خواهیم داشت.

One-Sample Statistics

Std. Error Mean	Std. Deviation	Mean	N	
2.33826	12.80719	129.9000	30	Albomin

One-Sample Test

Test Value = 132						
95% Confidence Interval of the Difference		Mean Difference	Sig. (2-tailed)	df	T	
Upper	Lower					
2.6823	-6.8823	-2.10000	.377	29	-.898	Albomin

جدول ۱۳-۱

تفسیر نتایج: در قسمت اول جدول ۱۳-۱ آمار توصیفی ساده ارایه شده است. در قسمت دوم ایـن جـدول مقدار آماره آزمون t=-0.898 با p-مقـدار (Sig. (2-tailed) برابـر بـا ۰/۳۷۷ کـه چـون از ۰/۰۵ بزرگتر است لذا فرضیه H_0 رد نمی شود.

۱۳-۲ آزمون میانگین های دو نمونه مستقل

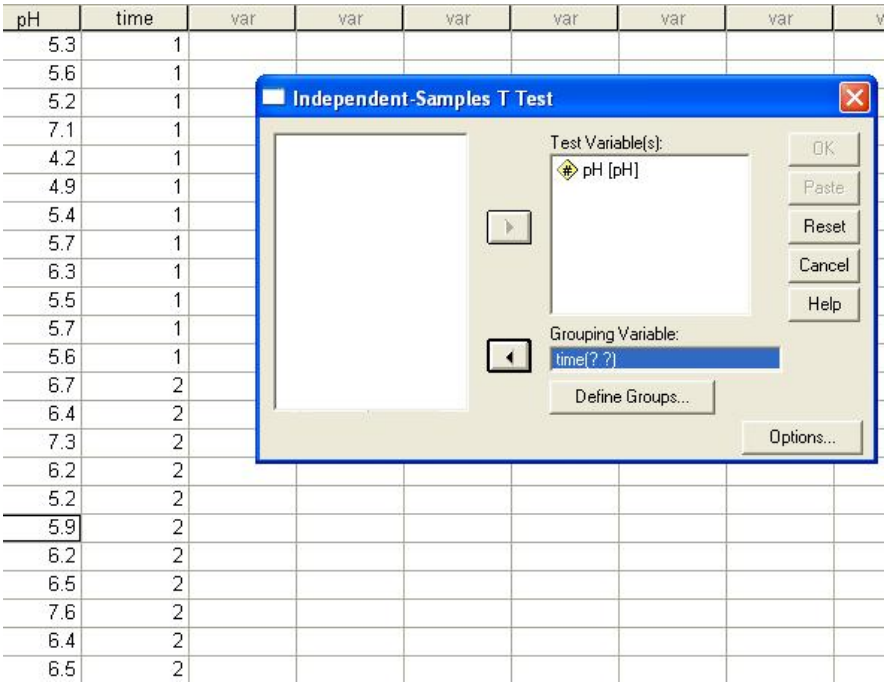
مثال۱۳-۲ . مقدار pH سلولهای کاکتوس در سپیده دم و شامگاه با استفاده از pH سنج اندازه گیری شده است. و نتایج زیر بدست آمده است. در سطح معنی داری $\alpha = 0/05$ آیا بین میانگینها تفاوت معنی داری موجود است.

سپیده دم (Dawn) : ۵/۳، ۵/٦، ۵/۲، ۷/۱، ٤/۲، ٤/۹، ۵/٤، ۵/۷، ٦/۳، ۵/۵، ۵/۷، ۵/٦

شامگاه (Dusk) : ٦/۷، ٦/٤، ۷/۳، ٦/۲، ۵/۲، ۵/۹، ٦/۲، ٦/۵، ۷/٦، ٦/٤، ٦/۵

برای انجام این آزمون در قسمت معرفی متغیرها دو متغیر Ph (کمی) و time (کیفی با دو برچسب ۱ (Dawn) و ۲ (Dusk) تعریف می کنیم. در قسمت ورود داده ها در ستون مربوط به pH نخست pH مربوط به سپیده دم و سپس pH مربوط به شامگاه را وارد نمایید. در ستون Time در مقابل اطلاعات مربوط به مقدار pH در سپیده عدد ۱ و در مقابل مقادیر مربوط به شامگاه عدد ۲ را تایپ کنید.

سپس فرمان Analyze- Compare Means- Independent Samples T- Test را وارد کنید تا کادر شکل(۱۳-۲) ظاهر می شود. نام متغیر pH را به بخش test Variable(s) و نام متغیر time را به بخش Grouping Variables منتقل نمایید.



شکل (۱۳-۲)

حال روی دکمه Define Groups کلیک کنید تا کادری با همین نام ظاهر شود (شکل ۱۳-۳) در چهارگوش اول عدد ۱ و در چهارگوش دوم عدد ۲ را وارد کرده دکمه continue را فعال کنید.

شکل (۱۳-۳)

در صورتی که بخواهید میزان سطح اطمینان را برای یافتن فاصله اطمینان تفاضل دو میانگین از مقدار ۰/۹۵ تغییر دهید کافی است دکمه Option را فعال نمایید. با زدن دکمه Ok نتایج (جدول ۱۳-۲) را خواهید دید.

Group Statistics

	time	N	Mean	Std. Deviation	Std. Error Mean
pH	Dawn	12	6.542	.7051	.2036
	Tusk	11	6.445	.6409	.1932

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
pH	Equal variances assumed	.017	.896	-3.206	21	.004	-.9038	.2819	-1.4900	-.3176
	Equal variances not assumed			-3.220	21.000	.004	-.9038	.2807	-1.4875	-.3201

جدول (۱۳-۲)

تفسیر نتایج:

در قسمت اول جدول ۱۳-۲ آمار توصیفی ساده ارائه شده است.

در قسمت دوم جدول ۱۳-۲، با توجه به اینکه در خط اول مقدار آماره F برای آزمون برابری واریانس‌های دو جامعه (Leven's Test Equality of Variance) برابر با ۰/۰۱۷ بوده و P-مقدار (Sig) متناظر با آن ۰/۸۹۸ است، لذا فرض برابری واریانس‌های دو جامعه پذیرفته می شود. در صورتی‌که فرض برابری واریانس‌ها رد می‌شد، می‌بایست نتایج مربوط به خط دوم را درنظر می‌گرفتیم.

حال با توجه به مقدار آماره آزمون $t = -3.206$ و p -مقدار (Sig =۰/۰۰٤) متناظر با آن در خط اول چون داریم $(sig = 0/004) < (\alpha = 0/05)$ پس فرض برابری میانگین‌ها رد می‌شود. بعبارت دیگر بین مقادیر pH تفاوت معنی داری وجود دارد.

۱۳-۳ آزمون میانگین های زوج شده

مثال ۱۳-۳. در بیمارستان کودکان مفید تهران برای بررسی تاثیر استامینفون در پایین آوردن میزان درجهٔ حرارت بدن ۱۰ کودک را در نظر گرفته و درجه حرارت تب آنها را قبل و یک ساعت پس از مصرف اندازه گیری نمودند. نتایج در چدول زیر آمده است.

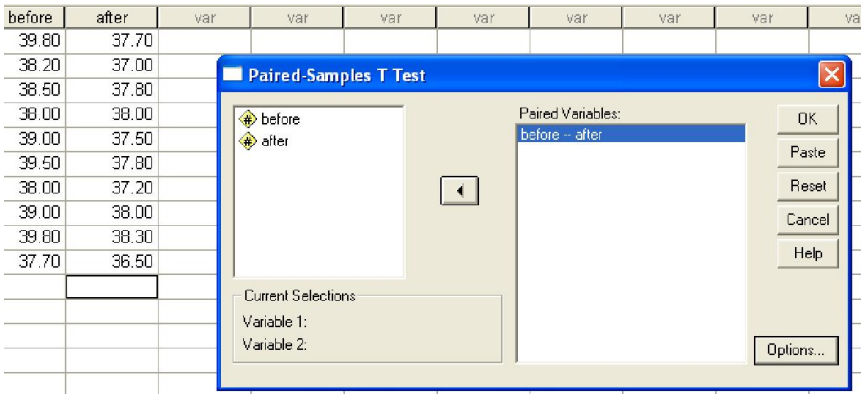
قبل	۳۹/۸	۳۸/۲	۳۸/۵	۳۸	۳۹	۳۹/۵	۳۸	۳۹	۳۹/۸	۳۷/۷
بعد	۳۷/۷	۳۷	۳۷/۸	۳۸	۳۷/۵	۳۷/۸	۳۷/۲	۳۸	۳۸/۳	۳۶/۵

در سطح معنی داری $\alpha = 0/01$ آیا دارو موثر بوده است یا خیر؟

برای انجام این آزمون در قسمت معرفی متغیرها دو متغیر کمی با دو رقم اعشارbefore و after را معرفی نموده و ارقام مربوطه را در قسمت داده ها وارد کنید. سپس فرمان

Anayze- Compare Means- Paird- Sample T Test

را وارد نموده و متغیر های before و after را به پنجره Paired Variables منتقل کنید. کادر شکل (۱۳-٤) را خواهید داشت.



شکل (۱۳-٤)

دکمه Options نظیر حالتهای قبل است. با فشردن دکمه Ok، این آزمون اجرا می شود. نتایج در جدول (۱۳-۳) ارائه شده است.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	38.7500	10	.78067	.24687
	after	37.5800	10	.54119	.17114

Paired Samples Correlations

		N	Correlation	Sig.
Pair1	before & after	10	.657	.039

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	before - after	1.17000	.58888	.18622	.74874	1.59126	6.283	9	.000

جدول (۱۳-۳)

تفسیر نتایج: در قسمت اول جدول (۱۳-۳) آمار توصیفی، در قسمت دوم این جدول میزان ضریب همبستگی پیرسن بین دو متغیر محاسبه شده است. در قسمت سوم این جدول نتیجه آزمون مقایسه دو میانگین زوج شده (قبل و بعد) ارائه شده است. با توجه به اینکه P –مقدار(Sig) در ستون آخر نزدیک صفر می باشد، پس فرضیه صفر (بی تأثیر بودن مصرف آستامینفون :H_0)، در سطح معنی داری $\alpha = 0/01$ رد می شود. بعبارت دیگر، مصرف قرص آستامینفون، در کاهش میزان تب کودکان بیمار تأثیر داشته است.

۱۳-٤ آزمون برابری میانگین های چند جامعه مستقل نرمال (تحلیل واریانس یکطرفه)

اگر تعداد جوامعی که می خواهیم برابری میانگین های آنها را مقایسه کنیم بیش از دو تا باشد از روش تحلیل واریانس یکطرفه استفاده می کنیم.

مثال ۱۳-٤. اثرسه نوع آنتی بیوتیک بر روی رشد یک باکتری با اضافه کردن این آنتی بیوتیکها به ظروف محیط کشت که توسط این باکتری ها تلقیح شده اند مورد بررسی قرار می گیرد. قطر کلونی ها[1] (به میلی متر) را سه روز بعد اندازه گرفتیم. یک ظرف نیز بدون استفاده از انتی بیوتیک به عنوان شاهد نیز در نظر گرفته شده است. نتایج در جدول زیر آمده است.

تیمار	قطر کلونی
شاهد	٦/٥، ٧/٨، ٥/٩، ٦/٧، ٦/٩، ٧، ٦/٣، ٦/٥، ٦/٦،٩/٧٨، ٧/٢، ٦/٧، ٨، ٧/٩
آنتی بیوتیک A	٤/٥، ٥/٦، ٥/٤، ٤/٩، ٤/٨، ٤/٦، ٥/١ ، ٥/٣، ٥/٣، ٥/٨، ٥/١، ٦، ٤/٤ ، ٥/٣
آنتی بیوتیک B	٢/٧، ٣/٢، ٢/٣،٤/٤، ٤/١، ٢/٨، ٣/٣، ٣/٥، ٢/٧، ٣/٢٣
آنتی بیوتیک C	**٤/٣، ٥/٧، ٥/٣، ٥/٦، ٤/٥، ٤/٩، ٥/١، ٥/٣، ٤/٧، ٦/٣، ٤/٨، ٤/٩، ٥/٢، ٥/٤، ٤/٨، ٥**

در سطح معنی داری $\alpha = 0/01$ یکسان بودن اثرچهار نوع تیمار مختلف را آزمون کرده در صورتیکه فرض برابری میانگین ها رد شود، آزمون چند دامنه ای دانکن را جهت مقایسهٔ دقیق تر به کار ببرید.

نحوهٔ ورود داده ها در *SPSS*

جهت ورود داده ها در قسمت تعریف متغیر ها نخست یک متغیرکیفی بنام Treatment (در ٤ سطح ۰، ۱، ۲، ۳ به ترتیب برای معرفی سطوح شاهد و آنتی بیوتیک های C,B,A) و سپس یک متغیر کمی بنام Diameter (برای معرفی مقدار قطر کلونی) تعریف کنید.

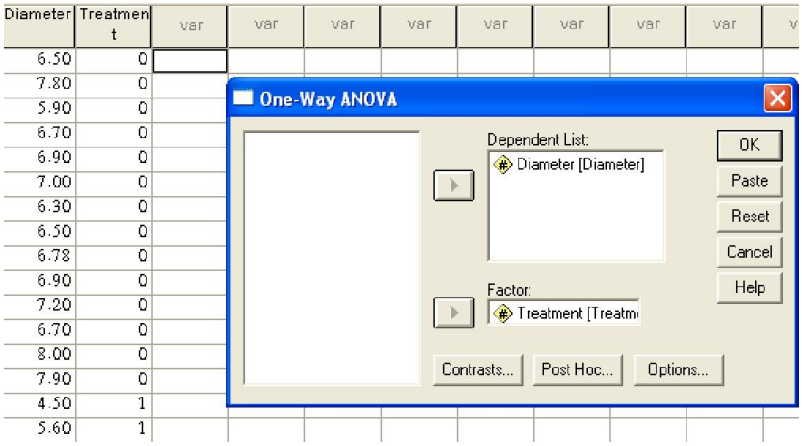
[1] colonies

در قسمت ورود داده ها در ستون مربوط به Diameter به ترتیب اطلاعات مربوط به قطر را به ترتیب برای تیمارهای شاهد، آنتی بیوتیک های A, B و C وارد می کنیم. در ستون مربوط به Treatment ، ١٤ بار عدد صفر (تعداد مشاهدات شاهد) ، ١٤ بار عدد یک (تعداد مشاهدات آنتی بیوتیک *A*) ، ١٠ بار عدد دو (تعداد مشاهدات آنتی بیوتیک *B*) ، و ١٦ بار عدد سه (تعداد مشاهدات آنتی بیوتیک *C*) را وارد نمایید.

سپس فرمان

Analyze-Compare Means- One way ANOVA

کادر محاوره ای شکل (١٣-٥) باز می شود.متغیر *Diameter* را به بخش *Dependent List* و متغیر *Treatment* را به بخش *Factor* منتقل نمایید.



شکل (١٣-٥)

با کلیک کردن روی دکمهٔ *Ok* نتایج جدول (١٣-٤) را خواهیم داشت.

Oneway

ANOVA

Diameter

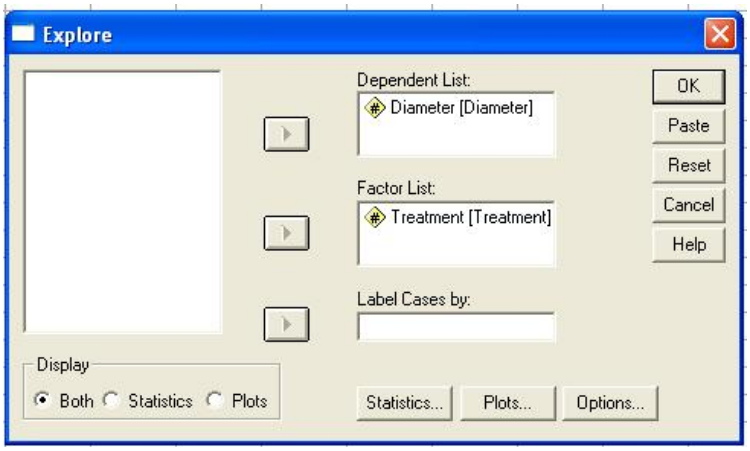
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	84.949	3	28.316	103.371	.000
Within Groups	13.696	50	.274		
Total	98.645	53			

جدول (١٣-٤)

که با توجه به اینکه *P* –مقدار بدست آمده برای آزمون F نزدیک صفر می باشد لذا فرضیه برابری میانگین ها را در سطح معنی داری پنج درصد .رد می کنیم. حال سؤالی که مطرح است آن است که کدامیک از تیمار ها رفتارش با بقیهٔ تیمارهها متفاوت بوده و وجودش باعث عدم رد فرضیهٔ H_0 می شود. جهت پاسخ به این سؤال نخست فرمان زیر را اجرا می کنیم:

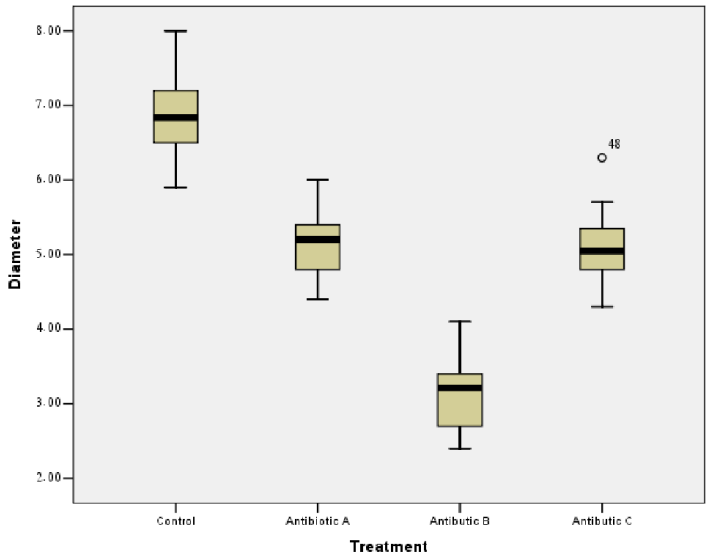
Analyze - Descriptive Statistics – Explore

در پنجره باز شده مربوط به *Explore* (شکل ۱۳-٦) در قسمت *Dependent List* متغیر *Diameter* و در پنجره *Factor List* متغیر *Treatment* را وارد کنید.



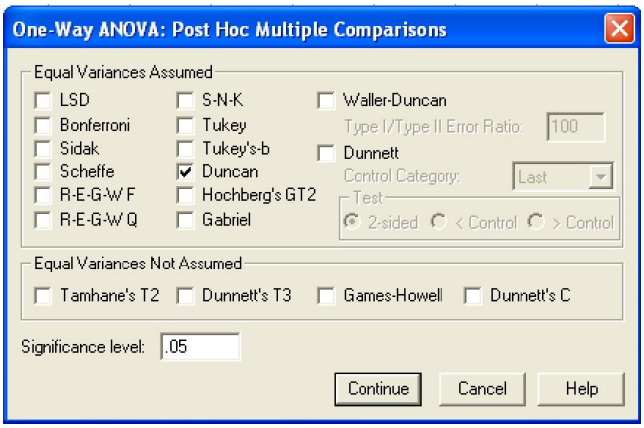
شکل ۱۳-٦

بدون هیچگونه تغییری در این پنجره دکمه *OK* را فعال نمایید. علاوه بر آمار توصیفی، نمودارشاخه و برگ ونمودار جعبه ای (شکل ۱۳-۷) رسم می گردد.



شکل (۱۳-۷)

بنا براین نمودار کاملا مشخص است که مقادیر قطری کلونی مربوط به تیمار کنترل بالاترین مقدار و مقادیر مربوط به آنتی بیوتیک نوع *B* کمترین مقدار را داراست. آنتی بیوتیکهای نوع *A* و *C* به طور تقریبی نظیر هم عمل می کنند. حال این موضوع را می توانیم با استفاده از آزمون چند دامنه ای دانکن (و یا هر آزمون دلخواه دیگر) با فعال نمودن دکمهٔ *Post Hoc...* در پنجره *One Way ANOVA* مورد بررسی قرار داد. با فعال کردن این دکمه پنجره شکل (۱۳-۸) را خواهیم داشت.



شکل (۱۳–۸)

با فعال کردن هر کدام از آزمونهای مورد نظر می توان فرض یکسان بودن تیمارها را بررسی نمود. به عنوان مثال دکمه مر بوط به *Duncan* را فعال نموده و سپس با زدن دکمهٔ *Continue* و *Ok* این فرمان را اجرا کنید. نتیجه در جدول (۱۳–۵) نشان دهندهٔ یکسان بودن تقریبی برای انتی بیوتیک های نوع A و نـوع C و بالا بودن رشد باکتری برای تیمار شاهد (بدون آنتی بیوتیک) و کم بودن رشـد بـاکتری در ظـرف C است.

Homogeneous Subsets

Diameter

Duncan[a,b]

Treatment	N	Subset for alpha = .05		
		1	2	3
Antibutic B	10	3.1330		
Antibutic C	16		5.1125	
Antibiotic A	14		5.1500	
Control	14			6.9343
Sig.		1.000	.855	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 13.099.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

جدول (۱۳–۵)

۱۳–۵ طرح بلوکهای کاملا تصادفی

مثال ۱۳–۵ به عنوان بخشی از مطالعه ای در بارهٔ کیفیت مواد غذایی جو صحرایی، ٦ نوع مختلـف از دانه های جو که پوستهٔ آنها را جدا کرده اند از نظر میزان موادشان تحت تجزیه و تحلیـل قـرار دادنـد.

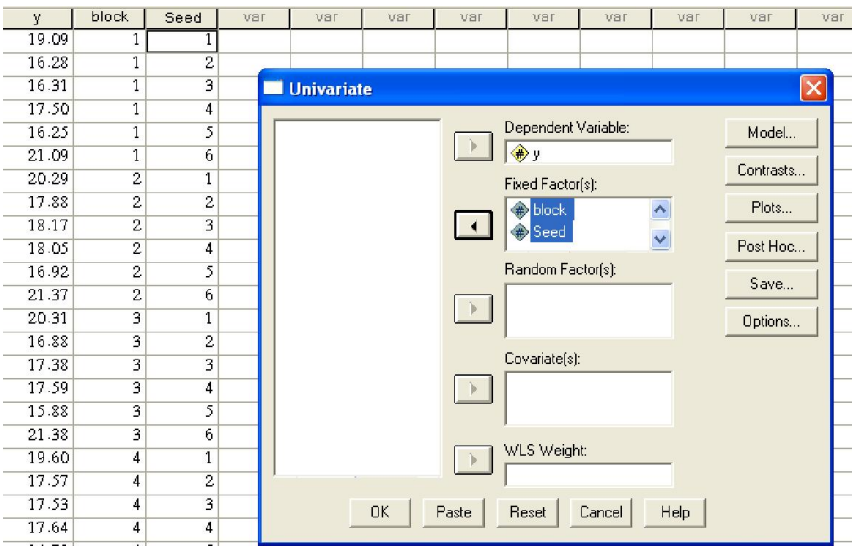
گیاهان را بر اساس یک طرح بلوکی تصادفی شده پرورش داده اند. اندازه های زیر میزان پروتئین ٦ نوع دانه را نشان می دهد. تجزیه و تحلیل داده ها را از طریق نرم افزار آماری SPSS ارائه می کنیم.

بلوک انواع دانه	١	٢	٣	٤	٥	٦
١	١٩/٠٩	١٦/٢٨	١٦/٣١	١٧/٥٠	١٦/٢٥	٢١/٠٩
٢	٢٠/٢٩	١٧/٨٨	١٨/١٧	١٨/٠٥	١٦/٩٢	٢١/٣٧
٣	٢١/٣١	١٦/٨٨	١٧/٣٨	١٧/٥٩	١٥/٨٨	٢١/٣٨
٤	١٩/٦٠	١٧/٥٧	١٧/٥٣	١٧/٦٤	١٤/٧٨	٢٠/٥٢
٥	١٨/٦٢	١٦/٧٢	١٦/٣٤	١٧/٣٨	١٥/٩٧	٢١/٠٩
٦	٢٠/١٠	١٧/٣٢	١٧/٨٨	١٨/٠٤	١٦/٦٦	٢١/٥٨

نحوه ورود داده ها: تمام داده های مربوط به مقادیر متغیر وابسته (میزان پروتئین) را سطر به سطر در ستون اول وارد کرده و نام این متغیر را y می نامیم. در ستون دوم به تعداد انواع دانه ها که ٦ می باشد، شش بار عدد یک، شش بار عدد ٢، شش بار عدد ٣، شش بار عدد ٤، شش بار عدد ٥ و شش بار عدد ٦ را تایپ نمایید، و نام این متغیر را *Block* بنامید. در ستون سوم که آن را *Seed* نام گذاری می کنیم، اعداد یک تا شش را ٦ بار تایپ نمایید. پس از ذخیره نمودن این اطلاعات با نام دلخواه، فرمان زیر را اجرا نمایید. تا کادر شکل (١٣-٩) باز شود:

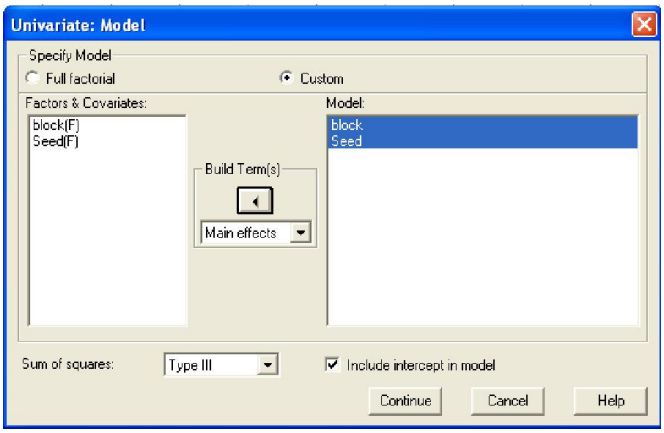
Analyze- General Linear model-Univariate…

متغیر y را به بخش *Dependent Variable* و متغیرهای *Block* و *Seed* را به بخش *Fixed Factor* منتقل نمایید.



شکل (١٣-٩)

حال روی دکمهٔ Model کلیک کنید تا کادر شکل (١٣-١٠) ظاهر شود.



شکل (۱۳-۱۰)

گزینهٔ Custom را انتخاب کنید. از لیست پایین افتادنی بخش Build Term(s) گزینهٔ Main effects را انتخاب کنید. هر یک از فاکتورهای block و seed را انتخاب و با فشار دکمهٔ ⟶ آنها را به بخش Model منتقل کنید. روی دکمهٔ Continue کلیک کنید تا به کادر Univariate برگردید. روی دکمهٔ OK کلیک کنید. خروجی این دستورات در جدول (۱۳-٦) ارائه شده است.

Tests of Between-Subjects Effects

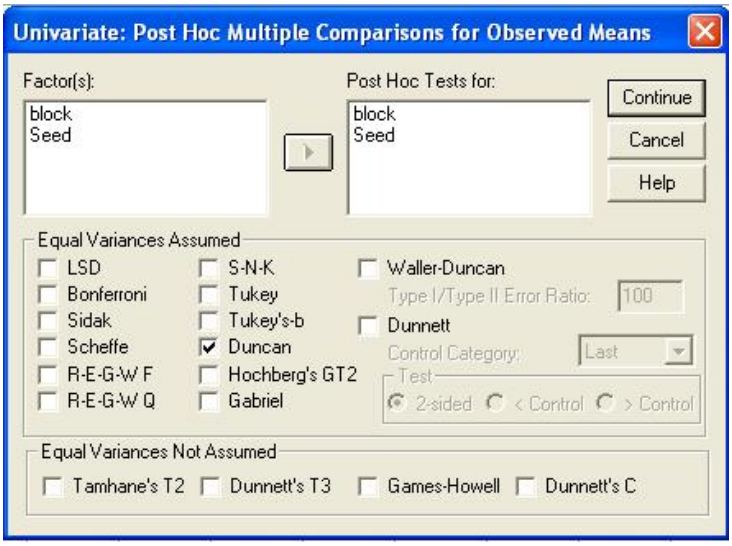
Dependent Variable: y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	112.899[a]	10	11.290	56.467	.000
Intercept	11879.547	1	11879.547	59415.43	.000
block	6.112	5	1.222	6.113	.001
Seed	106.788	5	21.358	106.820	.000
Error	4.999	25	.200		
Total	11997.445	36			
Corrected Total	117.898	35			

a. R Squared = .958 (Adjusted R Squared = .941)

جدول (۱۳-٦)

با توجه به اینکه p-مقدار برای فرض یکسان بودن بلوکها (سطر سوم جدول ۱۳-٦) برابر ۰/۰۰۱ و کمتر از ۰/۰۱ می باشد، لذا فرض مذکور رد می شود. از طرف دیگر p -مقدار برای فرض یکسان بودن انواع دانه ها (سطر چهارم جدول ۱۳-٦) نزدیک صفر بوده و بالطبع این فرض نیز رد می شود. جهت بررسی اینکه کدام یک از بلوکها و کدام یک از تیمارها با یکدیگر تفاوت دارند، آزمون مقایسه ای دانکن را می توان با فعال نمودن دکمهٔ Post Hoc در پنجره شکل (۱۳-۹) انجام داد. در نتیجه پنجره ای شکل (۱۳-۱۱) ظاهر می شود.



شکل (۱۳-۱۱)

با فعال نمودن گزینهٔ مربوط Duncan و سپس دکمه های Continue و OK ، جداول (۱۳-۷-الف) که مقایسه میانگین های بلوکها و (۱۳-۷-ب) که مقایسهٔ میانگین های مربوط به انواع دانه ها را نشان می دهد، ظاهر می شوند.

y

Duncan[a,b]

block	N	Subset 1	Subset 2
5	6	17.6867	
1	6	17.7533	
4	6	17.9400	
3	6	18.2367	18.2367
6	6		18.5967
2	6		18.7800
Sig.		.061	.056

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = .200.

a. Uses Harmonic Mean Sample Size = 6.000.

b. Alpha = .05.

جدول ۱۳-۷-الف

y

Duncan[a,b]

Seed	N	Subset 1	2	3	4	5
5	6	16.0767				
2	6		17.1083			
3	6		17.2683	17.2683		
4	6			17.7000		
1	6				19.6683	
6	6					21.1717
Sig.		1.000	.541	.107	1.000	1.000

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = .200.

a. Uses Harmonic Mean Sample Size = 6.000.

b. Alpha = .05.

جدول ۱۳-۷-ب

تفسیر نتایج:

با توجه به جدول (۱۳-۷-الف) داریم که بلوکهای ۵، ۱، ٤، شبیه به یکدیگر عمل کرده و تشکیل یک گروه مشابه را می دهند. به همین ترتیب بلوکهای ۳، ٦ و ۲ شبیه به یکدیگر عمل کرده و آنها نیز یک گروه مشابه را رقم می زنند.

با توجه به جدول (۱۳-۷- ب) عملکرد دانه های ۲و۳، شبیه به یکدیگر بوده و به همین ترتیب دانه های ۳ و ٤ نیز نظیر هم عمل می کنند.

۱۳-٦ طرح تحلیل واریانس دوطرفه (عامله) با اثر متقابل

مثال ۱۳-٦ به منظور بررسی نوعی حساسیت، چهار حیوان آزمایشگاهی خرگوش، موش، خوکچه هندی، و گربه را در سه زمان صبح، عصر و شب تحت مطالعه قرار می دهند. به عنوان قسمتی از تحقیق، تعیین ائوزینوفیل خون (تعداد در هر میلی متر مکعب) روی چهار حیوان صورت گرفته و نتایج در جدول زیر ارائه شده است.

زمان / حیوان	صبح	عصر	شب
خرگوش	۲۸۵ ۳۱۵ ۲۱۷	۲۹٦ ۲۰۲ ۲۸۳	۳۰۸ ۲۹۵ ۲٤۳
موش	۱۸۳ ۱۹٦ ۲۰٤	۱٦۲ ۱۷۹ ۱٥٤	۲۰۳ ۱٦۳ ۱۲٥
خوکچه	۳۱۰ ۳۲۳ ۲۸٦	۳۱۷ ۳٤٥ ۳۱٥	۲۹٦ ۳۱۷ ۳۰٦
گربه	۲۱٥ ۲۲۱ ۲۱۷	۲۲۳ ۲۰۳ ۲٦۳	۱۹٦ ۲۱۹ ۲٤٥

الف- جدول تحلیل واریانس دو طرفه راتشکیل دهید.

ب- آیا میانگین ائوزینوفیل در این حیوانات متفاوت است؟ (سطح معنی داری ۰/۰۱)

ج- آیا میانگین ائوزینوفیل در ساعات مختلف روز متفاوت است؟ (سطح معنی داری ۰/۰۱)

د- آیا اثر متقابل زمان – حیوان در سطح معنی داری ۰/۰۱ وجود دارد؟

نحوهٔ ورود داده ها:

ستونهای اول، دوم، سوم به ترتیب نشاهندهٔ متغیر های میزان ائوزینوفیل خون (کمی)، حیوان، و زمان (هر دو کیفی)می باشد. لذا:

در ردیف اول ستونهای اول، دوم، و سوم به ترتیب اعداد ۲۱۷ ، ۱ ، ۱

در ردیف دوم ستونهای اول، دوم، و سوم به ترتیب اعداد ۳۱۵ ، ۱ ، ۱

در ردیف سوم ستونهای اول، دوم، و سوم به ترتیب اعداد ۲۸۵ ، ۱ ، ۱

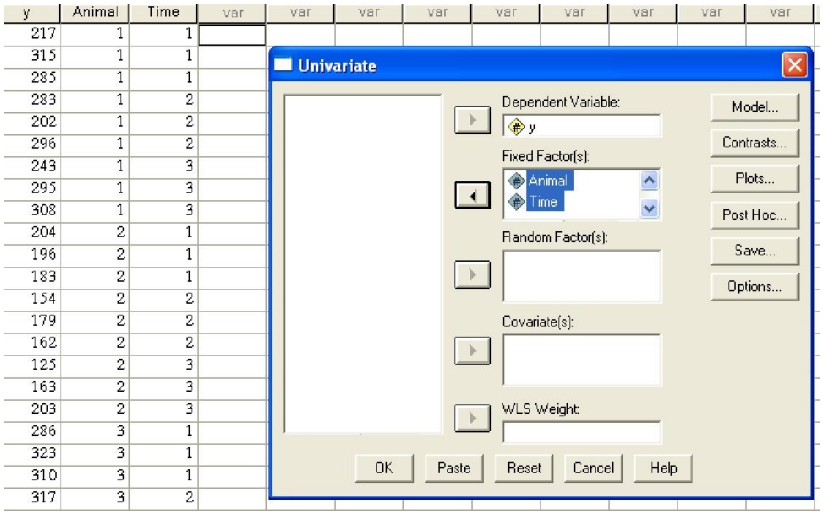
در ردیف جهارم ستونهای اول دوم، و سوم به ترتیب اعداد ۲۸۳ ، ۱ ، ۲

.....

و الخ نظیر شکل (۱۳-۱۲) داده ها را وارد می نماییم.

پس از ذخیره نمودن این اطلاعات با نام دلخواه، فرمان زیر را اجرا نمایید. تا کادر شکل (۱۳-۱۳) باز شود:

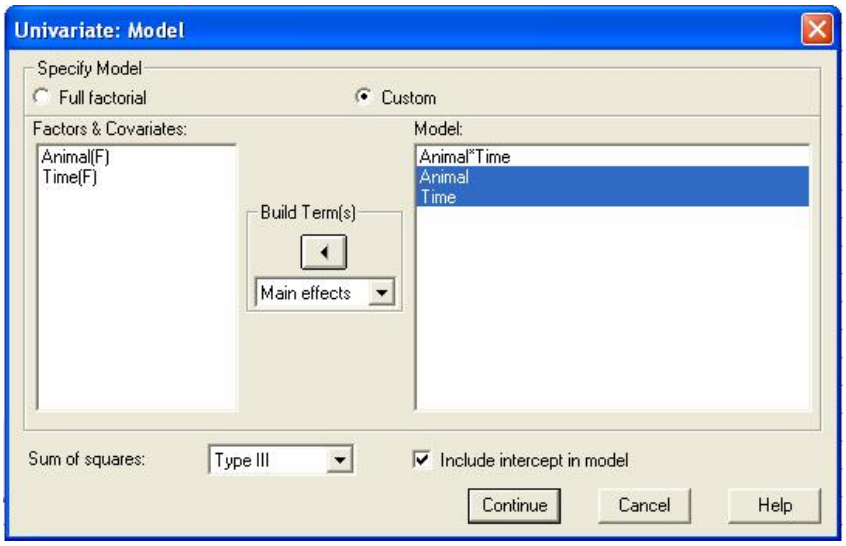
Analyze- General Linear model-Univariate…



شکل(۱۳-۱۲)

متغیر y را به قسمت Dependent Variable و متغیرهای Animal, Time را به بخش Fixed Factors منتقل کرده دکمهٔ model را فعال نمایید.پنجره ای نظیر شکل (۱۳-۱۳) باز می شود.

گزینهٔ Custom را انتخاب کنید. از لیست پایین افتادنی بخش Build Term(s) یک بار گزینهٔ Main effects (برای ارسال اثرات اصلی) و بار دیگر گزینه Interaction (برای ارسال اثر متقابل) را انتخاب کنید. هر دو فاکتورهای Animal و Time را انتخاب و با فشار دکمهٔ ← آنها را به بخش Model منتقل کنید. سه اثر در بخش Model ظاهر می شوند.



شکل (۱۳-۱۳)

روی دکمهٔ Continue کلیک کنید تا به کادر Univariate برگردید. روی دکمهٔ OK کلیک کنید. خروجی این دستورات در جدول (۱۳-۸) ارائه شده است.

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	100711.222[a]	11	9155.566	10.581	.000
Intercept	2165802.78	1	2165802.78	2503.095	.000
Animal * Time	3369.111	6	561.519	.649	.691
Animal	97211.222	3	32403.741	37.450	.000
Time	130.889	2	65.444	.076	.927
Error	20766.000	24	865.250		
Total	2287280.00	36			
Corrected Total	121477.222	35			

a. R Squared = .829 (Adjusted R Squared = .751)

جدول (۱۳-۸)

با توجه به مقادیر p–مقدار بزرگتر از ۰/۰۵ برای اثر متقابل (Animal*Time) فرض عدم وجود اثر متقابل در مدل رد نمی شود.

با توجه به مقدار نزدیک به صفر برای p–مقدار برای اثر حیوان (Animal) فرض یکسان بودن میانگین ائوزینوفیل برای حیوانات مختلف رد می شود.

با توجه به مقدار نزدیک به یک (۰/۹۲۷) برای p–مقدار برای اثر زمان (Time) فرض یکسان بودن میانگین ائوزینوفیل در ساعات مختلف روز متفاوت رد نمی شود.

۱۳-۷ ضریب همبستگی و رگرسیون

مثال ۱۳-۷ به منظور بررسی تاثیر یک داروی معین روی کنترل ضربان قلب دارو به ۱۰ نفر بیمار قلبی داده شد. جدول زیر نتایج میزان دوز مصرفی (برحسب میلی گرم) و واکنش در ضربان قلب (در دقیقه) را نشان می دهد.

x دوز mg	۰/۱	۰/۲	۰/۳۰	۰/۷	۰/۹	۱/۲	۱/٤٥	۱/۷۰	۲/۲٥	۲/۷٥
Y ضربان(دقیقه)	۳	٤	٥	٦	۷	۸	۹	۱۰	۱۱	۱۲

الف: ضریب همبستگی پیرسون و اسپیرمن را بدست آورده و در سطح معنی داری ۰/۰۱ آزمون کنید.
ب) معادلهٔ رگرسیون ساده خطی را بدست آوردید.

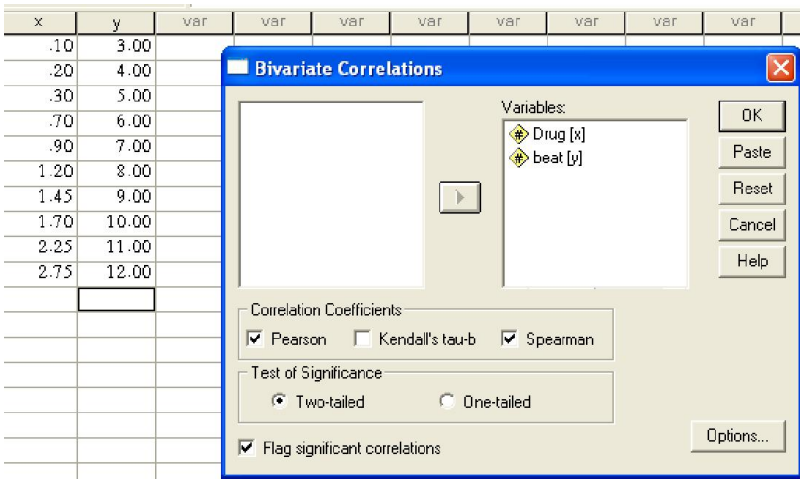
نحوهٔ ورود داده ها:

دو متغیر x و y در قسمت variable view معرفی کرده و سپس داده ها وارد می کنیم.

الف– برای بدست آوردن **ضرایب همبستگی پیرسن و اسپیرمن** فرمان زیر را وارد می کنیم.

Analyze-correlate-Bivariate…

تا کادر شکل (۱۳-۱٤) باز شود. متغیرهای x و y را به بخش Variableمنتقل می کنیم. سپس گزینه های Pearson و Spearman را فعال نموده و دکمهٔ OK را فعال سازید.



شکل (۱۳-۱٤)

جدول (۱۳-۹) میزان همبستگی پیرسن را برابر با ۰/۹۸۳ (با p- مقدار نزدیک به صفر sig=0) که نشان می دهد در سطح ۰/۰۱ معنی دار است.

Correlations

		Drug	beat
Drug	Pearson Correlation	1	.983**
	Sig. (2-tailed)		.000
	N	10	10
beat	Pearson Correlation	.983**	1
	Sig. (2-tailed)	.000	
	N	10	10

**. Correlation is significant at the 0.01 level

جدول (۱۳-۹)

جدول(۱۳-۱۰) ضریب همبستگی اسپیرمن برای این دو متغیر را می دهد، که مقداری برابر یک دارد.

Correlations

			Drug	beat
Spearman's rho	Drug	Correlation Coefficient	1.000	1.000**
		Sig. (2-tailed)	.	.
		N	10	10
	beat	Correlation Coefficient	1.000**	1.000
		Sig. (2-tailed)	.	.
		N	10	10

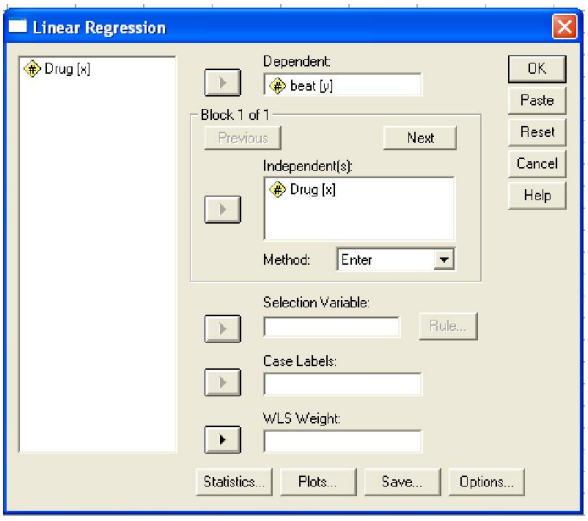
**. Correlation is significant at the 0.01 level (2-tailed).

جدول (۱۳-۱۰)

ب) برای انجام **برازش خط رگرسیون** فرمان زیر را اجرا کنید:

Analyze – Regression-Linear

تا کادر شکل (۱۳-۱۵) باز شود. متغیر y (تعداد ضربان) را به دریچهٔ Dependent و متغیر x (میزان دوز دارو) را به دریچهٔ Independent وارد نمایید.



شکل (۱۳-۱۵)

دکمه Ok را فعال کنید. نتیجهٔ محاسبات در جداول (۱۳-۱۱) تا (۱۳-۱۳) ارائه شده است.

تفسیر نتایج:

۱) در جدول (۱۳-۱۱) مقدار ضریب تعیین که برابر با توان دوم ضریب همبستگی است، برابر با $R^2 = 0.983$ شده است. با توجه به اینکه مقدار نزدیک به یک R^2 یکی از ملاک های مهم نیکویی برازش در مدل رگرسیونی است، لذا نتیجه می گیریم که مدل رگرسیون برازش داده شده مناسب است.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.983[a]	.965	.961	.59773

a. Predictors: (Constant), Drug

جدول (۱۳-۱۱)

۲) در جدول ANOVA (جدول ۱۳-۱۲) با توجه به p-مقدار بدست آمده نزدیک به صفر برای مقدار آزمون F، فرض

H_0 : مدل رگرسیونی مناسب نیست

را رد می کنیم. به عبارت دیگر مدل برازش داده شده معتبر می باشد.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	79.642	1	79.642	222.907	.000[a]
	Residual	2.858	8	.357		
	Total	82.500	9			

a. Predictors: (Constant), Drug

b. Dependent Variable: beat

جدول (۱۳-۱۲)

۳) در جدول (۱۳-۱۳) مقادیر برآورد پارامترهای مدل رگرسیون و آزمون مربوط به آن آمده است.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.647	.320		11.402	.000
	Drug	3.336	.223	.983	14.930	.000

a. Dependent Variable: beat

جدول (۱۳-۱۳)

مقدار برآورد پارامتر عدد ثابت مدل رگرسیون در ستون دوم با عنوان B روبروی ردیف constant و برابر با ۳/٦٤٧ می باشد. با توجه به مقدار آماره $t = 11/402$ و p-مقدار نزدیک به صفر فرضیه صفر $H_0 : a = 0$ رد می شود.

مقدار برآورد پارامتر عرض از مبدا مدل رگرسیون در ستون دوم با عنوان B روبروی ردیف Drug و برابر با ۳/۳۳۶ می باشد. با توجه به مقدار آماره $t = 14/930$ و p–مقدار نزدیک به صفر فرضیه صفر $H_0 : b = 0$ رد می شود.

به عبارت دیگر برآورد مدل رگرسیون خطی ساده عبارت است از:

$$\hat{y} = 3/647 + 3/336\,x$$

۱۳–۸ آزمون استقلال و معیارهای همایندی

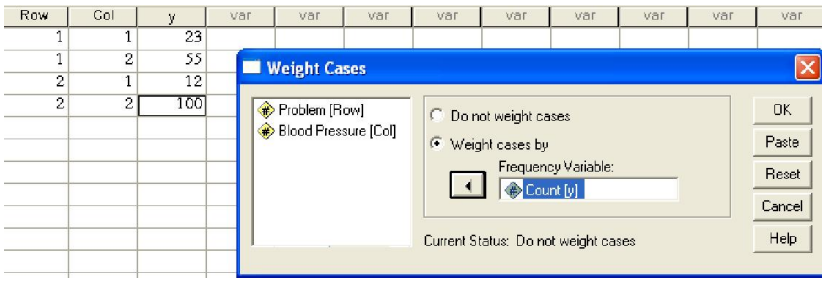
همانطوری که در فصل دوازدهم اشاره شد، برای بررسی میزان ارتباط متقابل (همایندی) دو متغیر کیفی نیازمند محاسبه معیارهای مختلفی هستیم که بسته به نوع دو متغیر کیفی تفاوت داشته و در زیر با ذکر مثال کاربردی مورد مطالعه قرار می گیرند.

الف. دو متغیر کیفی اسمی

مثال ۱۳–۸ مطالعه ای بر روی ۱۹۰ نفر زن باردار نتایج زیر را در مورد رابطهٔ بین فشارخون مادران و یک نوع عارضهٔ حاملگی ارائه داده است. معیارهای همایندی مناسب و آزمون استقلال دو متغیر را بررسی کنید.

عارضه حاملگی ↓ فشار خون →	دارد	ندارد
دارد	۲۳	۵۵
ندارد	۱۲	۱۰۰

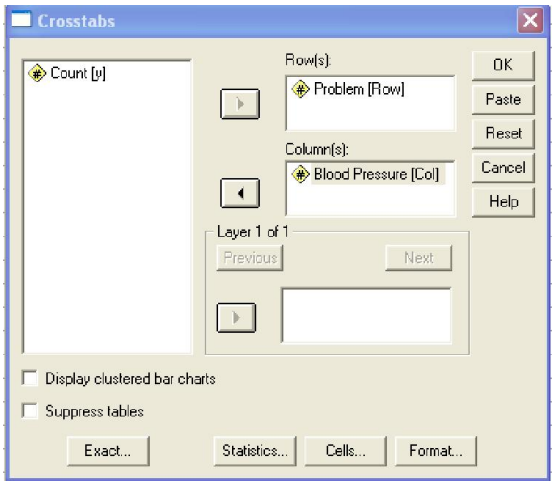
نحوه ورود داده ها: سه متغیر Row، Col، و y در بخش variable view تعریف کرده و برای متغیر Row برچسب problem با سطوح وجود عارضه با کد ۱ و عدم وجود عارضه با کد ۰ و متغیر Col با برچسب Blood Pressure با سطوح وجود فشارخون با کد ۱ و عدم وجود آن با کد ۰ تعریف می کنیم. سپس داده را به صورت شکل (۱۳–۱۶) وارد نمایید. حال از پنجره Data در منو بار اصلی گزینهٔ آخر (Weight Case) را فعال کرده و پس انتقال y به دریچهٔ Fresquency Variable از طریق گزینهٔ Weight cases by دکمهٔ Ok را فعال می نماییم. انجام این کار برای تشکیل جدول توافقی ضروری است.



شکل (۱۳-۱۶)

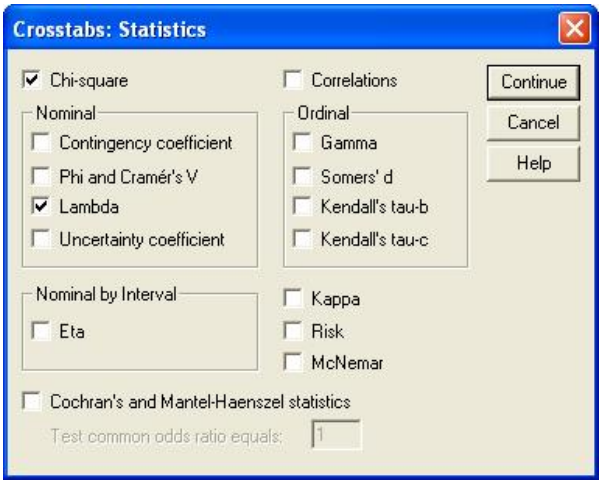
حال فرمان زیر را اجرا نمایید: Analyze- Descriptive statistics- Cross tabs…

کادر شکل (۱۳-۱۷) ظاهر می شود. متغیرRow را به بخش Row(s) و متغیر Col را به بخش Column منتقل کنید.



شکل(۱۳-۱۷)

حال اگر دکمه Statistcs را در کادر شکل (۱۳-۱۷) فعال نمایید، پنجره ای نظیر شکل (۱۳-۱۸) خواهید دید.



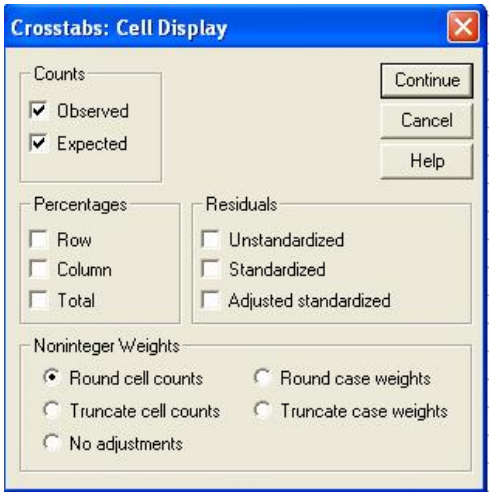
شکل (۱۳-۱۸)

گزینهٔ Chi-square برای انجام آزمون استقلال χ^2 را فعال کنید.

برای بررسی میزان ارتباط دو متغیر کیفی اسمی می توان از معیارهای بخش Nominal استفاده کرد که در این جا ما از گزینهٔ Lambda استفاده کرده ایم.

تذکر مهم: در صورتیکه دو متغیر کیفی رتبه ای باشند از معیارهای بخش Ordinal منجمله معیار Gamma استفاده می شود. اگر مقایسهٔ یک متغیر کیفی اسمی و یک متفیر کمی مورد نظر باشد از آمارهٔ

Eta در بخش Nominal by Interval استفاده کنید. سپس گزینهٔ Continue را فعال کنید. حال گزینهٔ Cells در پنجره شکل (۱۳-۱۷) را فعال کنید تا پنجره شکل (۱۳-۱۹) ظاهر شود.



شکل (۱۳-۱۹)

بخش Counts شامل دو گزینه مقادیر مشاهده شده (Observed) و مقادیر مورد انتظار (Expected)می باشد که در صورت فعال شدن این مقادیر را چاپ می کند. بخش Percentages شامل سه گزینه است که در صورت فعال شدن به ترتیب درصد فراوانی یک سلول جدول توافقی نسبت به سطر (Row)، نسبت به ستون (Column) و نسبت به کل (Total) را چاپ می کند. حال با انتخاب دکمهٔ Continue به پنجره اصلی برگشته و Ok کنید.

تفسیر نتایج : در جدول (۱۳-۱٤) مقادیر مشاهده شده و مورد انتظار جدول توافقی ارائه شده است.

Problem * Blood Pressure Crosstabulation

			Blood Pressure		Total
			1	2	
Problem	1	Count	23	55	78
		Expected Count	14.4	63.6	78.0
	2	Count	12	100	112
		Expected Count	20.6	91.4	112.0
Total		Count	35	155	190
		Expected Count	35.0	155.0	190.0

جدول (۱۳-۱٤)

درجدول (۱۳-۱۵) مقادیر چند آماره مهم منجمله آمارهٔ χ^2 برای آزمون استقلال دو متغیر سطر و ستون ارائه شده است. از آنجا که مقدار Asymp. Sig. (2-sided) متناظر با این آماره برابر به ۰/۰۰۱ بوده و کوچکتر از ۰/۰۱ است لذا فرض استقلال دو متغیر سطر و ستون با اطمینان ۹۹ درصد رد می شود.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	10.783[b]	1	.001		
Continuity Correction[a]	9.570	1	.002		
Likelihood Ratio	10.654	1	.001		
Fisher's Exact Test				.002	.001
Linear-by-Linear Association	10.726	1	.001		
N of Valid Cases	190				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 14.37.

جدول (۱۳-۱۵)

جدول (۱۳-۱۶) مقادیر آماره λ (Lambda) در حالت متقارن، و شرطی (عارضه حاملگی وابسته و فشار خون مستقل) و برعکس ارائه شده است.

Directional Measures

			Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Nominal by Nominal	Lambda	Symmetric	.097	.048	1.876	.061
		Problem Dependent	.141	.070	1.876	.061
		Blood Pressure Dependent	.000	.000	.[c]	.[c]
	Goodman and Kruskal tau	Problem Dependent	.057	.033		.001[d]
		Blood Pressure Dependent	.057	.034		.001[d]

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

c. Cannot be computed because the asymptotic standard error equals zero.

d. Based on chi-square approximation

جدول (۱۳-۱۶)

با توجه به مقدار $\lambda = 0/097$ داریم که میزان پیش بینی هر کدام از این دو متغیر نسبت به دیگری زیاد بالا نمی باشد. پیش بینی متغیر وجود عارضه حاملگی بوسیله دانستن فشارخون بالا به اندازهٔ حدود ۱٤ درصد نسبت به حدس کورکورانه، خطای پیش بینی را کاهش می دهد. و سر انجام اینکه متغیر فشار خون در زنان نمی تواند پوسیله دانستن وضعیت عارضه حاملگی پیش بینی شود.

ب. دو متغیر کیفی رتبه ای

مثال ۱۳-۹ اطلاعاتی از یک نمونهٔ ۲٦٦ نفری از بیماران راجع به فشار روحی محسوس و فراوانی بیماری در میان آنها گرد آوری شد. مسئله این است که آیا مطابقتی میان این دو متغیر رتبه ای وجود دارد؟ آزمون استقلال را به کار ببرید.مشاهدات به شرح زیر است.

فراوانی بیماری↓ درجهٔ فشار روحی←	بسیار زیاد	زیاد	متوسط	کم	خیلی کم
بسیار فراوان	٣٢	١٩	١٣	٨	٢
فراوان	٢٦	١٤	٩	٧	٦
گاه گاهی	٤	٩	١٤	١٩	٢١
نادر	٥	٧	١١	١٦	٢٤

با توجه به اینکه دو متغیر سطر و ستون رتبه ای می باشند معیار گاما (Gamma) و آزمون استقلال را انجام می دهیم. مراحل ورود داده ها نظیر مثال ١٣-٨ می باشد.

تفسیر نتایج:

در سلولهای جدول(١٣-١٧) به ترتیب مقادیر مشاهده شده، مقادیر مورد انتظار و نسبت فراوانی (به کل) ارائه شده است.

Frequency of sickness * Rank of stress Crosstabulation

			Rank of stress					Total
			1	2	3	4	5	
Frequency of sickness	1	Count	32	19	13	8	2	74
		Expected Count	18.6	13.6	13.1	13.9	14.7	74.0
		% of Total	12.0%	7.1%	4.9%	3.0%	8%	27.8%
	2	Count	26	14	9	7	6	62
		Expected Count	15.6	11.4	11.0	11.7	12.4	62.0
		% of Total	9.8%	5.3%	3.4%	2.6%	2.3%	23.3%
	3	Count	4	9	14	19	21	67
		Expected Count	16.9	12.3	11.8	12.6	13.3	67.0
		% of Total	1.5%	3.4%	5.3%	7.1%	7.9%	25.2%
	4	Count	5	7	11	16	24	63
		Expected Count	15.9	11.6	11.1	11.8	12.6	63.0
		% of Total	1.9%	2.6%	4.1%	6.0%	9.0%	23.7%
Total		Count	67	49	47	50	53	266
		Expected Count	67.0	49.0	47.0	50.0	53.0	266.0
		% of Total	25.2%	18.4%	17.7%	18.8%	19.9%	100.0%

جدول(١٣-١٧)

در قسمت اول جدول (١٣-١٨) مقادیر چند آماره منجمله آمارهٔ معروف χ^2 برای آزمون استقلال دو متغیر سطر و ستون ارائه شده است. از آنجا که مقدار Asymp. Sig. (2-sided) متناظر با این آماره نزدیک به صفر وکوچکتر از ٠/٠١ است لذا فرض استقلال دو متغیر با اطمینان ٩٩ درصد رد می شود. در قسمت بعدی این جدول مقدار آماره Gamma برابر ٠/٥١٤ شده است که نشان دهندهٔ میزان همایندی بالای این دو متغیر کیفی رتبه ای است. بعبارت دیگر برای پیش بینی یکی از این دو متغیر با دانستن متغیر دیگر میزان خطای پیش بینی ما ٠/٥١٤ کاهش می یابد.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	78.119[a]	12	.000
Likelihood Ratio	87.085	12	.000
Linear-by-Linear Association	66.153	1	.000
N of Valid Cases	266		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.95.

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Gamma	.514	.048	10.326	.000
N of Valid Cases		266			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

جدول (۱۳-۱۸)

ج. یک متغیر کیفی اسمی و یک متغیر با مقیاس نسبی

مثال ۱۳-۱۰ (مثال ۱۲-۸ در فصل دوازدهم را مجددا ارائه می نماییم) میزان فشارخون سیستولیک (میلی متر جیوه) برای دو گروه از بیماران فشارخونی مرد و زن در جدول زیر آمده است. می خواهیم ببینیم آیا بین جنس و میزان فشارخون ارتباط وجود دارد

مردان	۱۳۰	۱۳٤	۱۳۲	۱٤۰	۱۵۰	۱٤٤	۱۵۰	۱۲۲	۱٦۰	۱۵۰
زنان	۱۳۰	۱۳۸	۱٤٤	۱٤٦	۱٤۸	۱۵۲	۱۵۸	۱۳۰	۱٦۲	۱٦۸

در این فرضیه، جنس یک متغیر کیفی با مقیاس اسمی و میزان فشارخون متغیری کمی با مقیاس فاصله ای (نسبی) می باشد. نسبت همبستگی اتا برای بررسی این ارتباط بکار می رود.

جهت انجام آزمون نخست یک متغیر جنسیت (Sex) در دو سطح (صفر و یک به ترتیب برای مردان و زنان) تعریف می کنیم. سپس یک متغیر دیگر بنام فشار خون (BP) تعریف نموده و کلیه موارد فشار خون ثبت شده را برای مردان و زنان در این ستون تایپ می کنیم. حال فرمان زیر را اجرا کنید.

Analyze- Descriptive statistics- Cross tabs…

با توجه به نتیجه ارائه شده در جدول (۱۳-۱۹) داریم که میزان همبستگی بین دو متغیر جنس و فشارخون برابر مقدار $\eta = 0/265$ شده، که نشان دهندهٔ وجود همبستگی نسبتا بالایی نیست.

Directional Measures

			Value
Nominal by Interval	Eta	sex Dependent	.876
		Blood Pressure Dependent	.265

جدول (۱۳-۱۹)