

بنام یکتای بی همتا

آموزش مقدماتی SPSS

برای دانشجویان کارشناسی جغرافیا

مرتضی اسمعیل نژاد

دانشجوی دکتری جغرافیا

دانشکده جغرافیا و برنامه ریزی محیطی، گرده جغرافیا

واژه آمار از کلمه لاتین **Status** سرچشمه گرفته است که به معنی حالت، وضع یا موقعیت می باشد. از این کلمه به عنوان ریشه واژه **Statistica** (آمار) به وجود آمده است. جمع آوری آمار و اطلاعات را می توان به زمان پیدایش اولین تمدنهای بشری نسبت داد. در زمان داریوش لغت «شاهامار» به معنی سرشماری به کار برده می شد

آمار عمدتاً با وضعیتهایی سر و کار دارد که در آنها وقوع یک پیشامد به طور حتمی قابل پیش بینی نیست. استنتاجهای آماری غالباً "غیر حتمی اند، زیرا مبتنی بر اطلاعات ناکافی هستند: تخمین نرخ بیکاری در یک ناحیه بر مبنای مطالعه روی چند هزار نفر از مردان یا بررسی طرح جدید ترافیک بر اساس نظرسنجی، مثالهایی از این موضوع اند.

آمار چیست؟

آمار علمی است که پیرامون جمع آوری و تنظیم و تحلیل و تفسیر اطلاعات عددی سخن می گوید. موضوع آمار عبارتست از هنر و علم جمع آوری، تعبیر و تجزیه و تحلیل داده ها و استخراج تعمیمهای منطقی در مورد پدیده های تحت بررسی
مباحث اصلی علم آمار کدامند؟

۱- آمار توصیفی که ویژگی های کلی تعدادی از داده ها را (که معمولاً به نمونه تعلق دارند) در قالب یک عدد بیان می کند. آمار توصیفی به هدف اول علم آمار خدمت می کند.

۲- آمار استنباطی که بر اساس ویژگیهای مشاهده شده در نمونه، ویژگی های جامعه را برآورد می کند. آمار استنباطی هدف دوم علم آمار را تأمین می کند.

۳- احتمالات که خطای برآوردهای آمار استنباطی را اندازه می گیرد.

با پیشرفت علوم و گسترش تکنولوژی، اهمیت استفاده از روشهای آماری در علوم مختلف بیش از پیش مورد توجه قرار گرفته است و آموختن آمار کاربردی در هر رشته جزء ملزومات گردیده است. فرآیند آنالیز آماری کمک میکند تا پژوهشگر بتواند از داده های اولیه، اطلاعات مورد نیاز خود را استخراج کند و در صورت لزوم نتایج را تعمیم دهد. اگر حجم داده ها بزرگ باشد، استفاده از

روشهای مختلف آنالیز آماری بسیار خسته کننده و مشکل خواهد بود. امروزه انواع نرم افزارهای مختلف آماری موجود، قادرند انواع آنالیزهای آماری را انجام دهند.

نرم افزار SPSS یکی از قدیمی ترین، برنامه های کاربردی در زمینه تجزیه و تخیلهای آماری است؛ نرم افزاری آماری با قابلیت انجام توصیفی زیبا و گویا از اطلاعات، شامل رسم نمودارها و چارت های گوناگون و محاسبات مربوط به میانگین، انحراف معیار واریانس، میانه و غیره.

قابلیت های نرم افزار SPSS به شرح زیر است:

تهیه خلاصه های آماری مانند گراف ها، جداول، آمارها و ...

انواع توابع ریاضی مانند قدر مطلق، تابع علامت، لگاریتم، توابع مثلثاتی و ...

تهیه انواع جداول سفارشی مانند جداول فراوانی، فراوانی تجمعی، درصد فراوانی و ...

انواع توزیع های آماری شامل توزیع های گسسته و پیوسته

تهیه انواع طرح های آماری

انجام آنالیز واریانس یکطرفه، دوطرفه، چندطرفه و آنالیز کوواریانس

تکنیک های تجزیه و تحلیل سری های زمانی

ایجاد داده های تصادفی و پیوسته

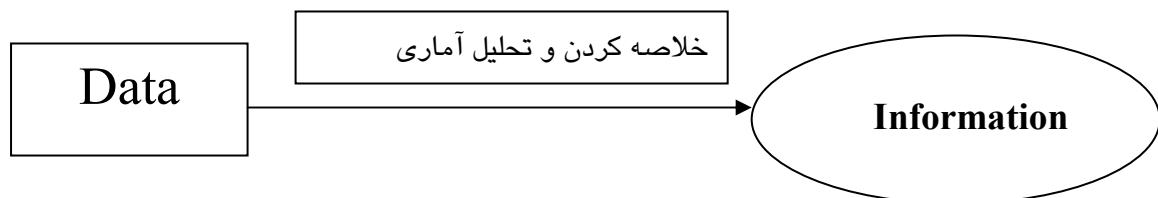
محاسبه انواع آماره های توصیفی

انواع آزمون های مرتبط با مقایسه میانگین بین دو یا چند جامعه مستقل و وابسته

قابلیت مبادله اطلاعات با نرم افزارهای دیگر

برآزش انواع مختل رگرسیون

اگر وزن دانشجویان دانشکده جغرافیا و برنامه ریزی را داشته باشید، فقط داده خام دارید، اما اگر میانگین، انحراف معیار، مد، نما، نمودار..... را داشته باشید، اطلاعات دارید.



شروع کار با SPSS و اعمال مقدماتی

برای ورود به محیط نرم افزار SPSS در صورت وجود Icon نرم افزار SPSS در Desktop روی آن Double-click کنید:





اگر در کادر زیر تیک بگذارید، در دفعات بعدی هنگام ورود به برنامه این سوال پرسیده نخواهد شد و به صورت پیش فرض یک صفحه کار خالی باز خواهد شد (این کار توصیه نمیشود).



با دادن پاسخ مناسب به کادر گفتگوی بالا وارد محیط SPSS خواهید شد. تصویر محیط اولیه SPSS در صفحه بعد آمده است.

➤ SPSS همواره وضعیت جاری خود را در کادری در پایین صفحه به اطلاع شما می‌رساند، توجه کنید که زمانی برنامه به خوبی کار خواهد کرد که عبارت SPSS Processor is ready در این کادر نوشته شده باشد:



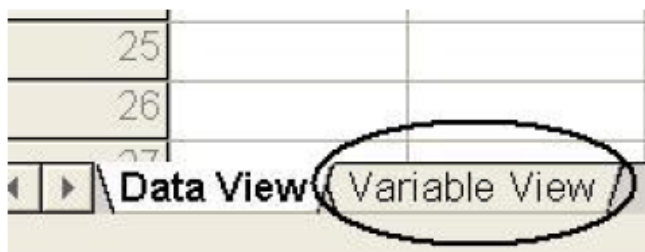
نحوه وارد کردن داده‌ها در SPSS :

معمولاً داده‌هایی که در اختیار داریم عبارتند از مقادیر چندین متغیر برای یک واحد نمونه. اعداد اولین ستون سمت چپ، نماینده واحدهای نمونه‌ای هستند و در هر سطر مقادیر متغیرهای مورد بررسی برای هر واحد نمونه وارد خواهد شد.

به عنوان مثال داده‌های ایستگاههای سینوپتیک را در اختیار داریم در یک ستون نام ایستگاهها و در ستون‌های دیگر متغیرها را وارد می‌نماییم. اما قبل از وارد کردن داده‌ها باید نوع داده را برای نرم افزار تعریف نماییم.

تعریف متغیرها در SPSS:

برای تعریف متغیرها ابتدا روی عبارت Variable view در نوار زیر کلیک کنید:



محیط کار SPSS به صورت زیر تغییر خواهد کرد:

مشخصه های مختلف یک متغیر در SPSS :

Name: در این قسمت نام متغیر را وارد میکنیم. نام یک متغیر در SPSS حداکثر ۸ کاراکتر متشکل از ارقام و حروف است که کاراکتر اول آن نباید رقم باشد. به عنوان مثال داریم:

نامهای مجاز در SPSS	نامهای مجاز در SPSS
2num	Vazn
Mizane masraf	Income
.	Nomre3
.	X
.	.

سعی کنید اسمی که انتخاب میکنید حتی الامکان بامفهوم باشد تا در مراحل بعدی کارتان ساده تر شود.

Type نوع متغیر مورد بررسی را مشخص می کند. Numeric متغیرهای عددی مانند: جمعیت شهری، بارش، دما،... و String متغیرهای رشته ای مانند نام ایستگاه و نام شهر.... بقیه کاربرد چندانی ندارند.

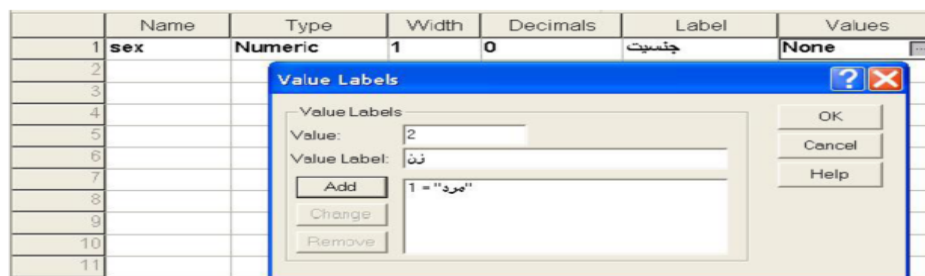
Decimals و Width: اگر نوع متغیر عددی (Numeric) باشد در این دو ستون تعداد کل ارقام عدد (Width) و تعداد رقمهای اعشار (Decimals) را میتوان تعیین کرد.

Label: چون اسم متغیرها در SPSS دارای محدودیت است (تا ۸ کاراکتر و ...) در صورتیکه مایل باشیم میتوانیم در این قسمت یک عنوان (Label) برای متغیر قرار میدهیم. به عنوان مثال اگر متغیر مورد بررسی تعداد افراد هر خانواده باشد میتوانیم در قسمت Label تایپ کنیم:

Number of family members

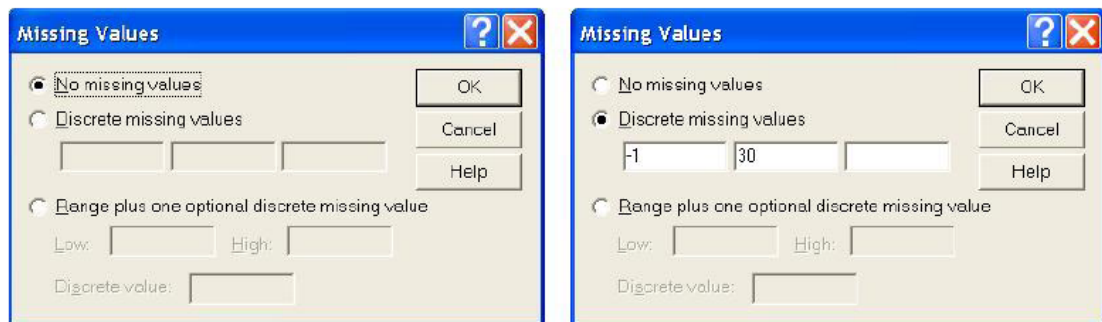
Value: در این قسمت اگر متغیر مورد بررسی دارای مقادیر مشخصی باشد، مقادیر آن و برچسب مقادیر را مشخص میکنیم. به عنوان مثال متغیر جنسیت دارای دو مقدار ۱ با برچسب "مرد" و ۲ با برچسب "زن" است.

➡ برای اضافه کردن یک مقدار و برچسب مربوط به آن مطابق تصویر زیر مقدار و برچسب را در کادرهای مربوطه تایپ کرده و دکمه Add را کلیک کنید:

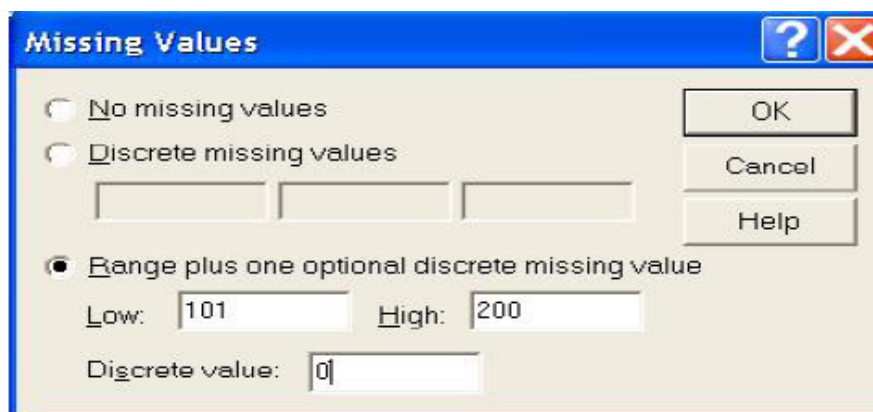


Missing: در این قسمت نحوه معرفی داده های گمشده به SPSS تعیین میشود.

انتخاب گزینه اول به معنی عدم تعریف داده گمشده است. با انتخاب گزینه دوم میتوان چند مقدار مشخص را برای معرفی یک مقدار گمشده تعیین کرد. به عنوان مثال اگر متغیر مورد بررسی تعداد واحد باشد میتوان از مقادیری که خارج از دامنه تغییراتند چند مقدار برای معرفی مقادیر گمشده تعیین کرد. (مثلاً ۳۰ یا -۱)



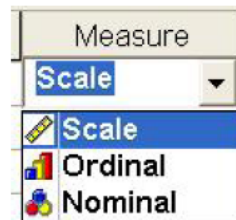
با انتخاب گزینه سوم میتوان یک بازه عددی را برای مشخص کردن مقادیر گم شده تعیین کرد. مثلاً اگر متغیر مورد بررسی درصد طلای موجود در یک انگشتر باشد میتوان بازه ۱۰۱ تا ۲۰۰ و عدد ۰ را برای تعیین مقادیر گم شده تعیین کرد:



Columns: اندازه ستونی که محل وارد کردن داده های متغیر مورد بررسی است را در این قسمت وارد میکنیم.

Align: تعیین میکند که داده ها در سمت راست، چپ یا وسط خانه (cell) قرار گیرند.

Measure: در این قسمت مقیاس اندازه گیری داده ها را تعیین میکنیم:

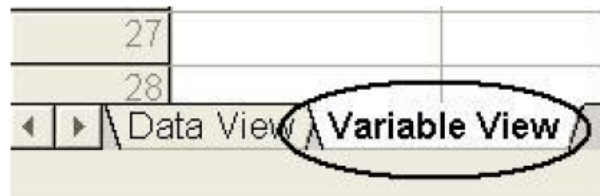


Scale- (فاصله ای): اعداد حقیقی هستند که اعمال جبری روی آنها امکان پذیر است.

Ordinal- (رتبه ای): متغیری عددی است که مجاز به انجام عملیات جبری در مورد آنها نیستیم بلکه تنها ارجحیت یا برتری را نشان میدهد.

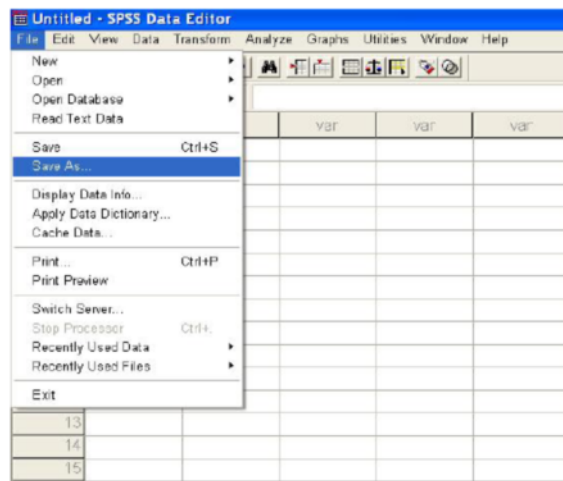
Nominal- (اسمی): فاقد هرگونه ارجحیت میباشد.

بعد از تعریف تمامی متغیرهای لازم با کلیک روی عبارت Data view به محیط وارد کردن داده ها برگشته و داده ها را وارد کنید.

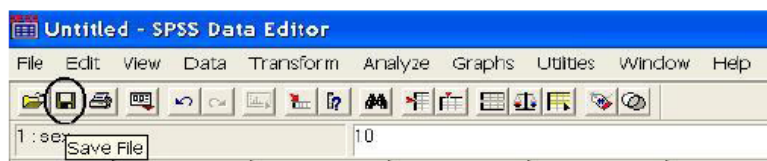


- ذخیره کردن و باز کردن فایلها در SPSS:

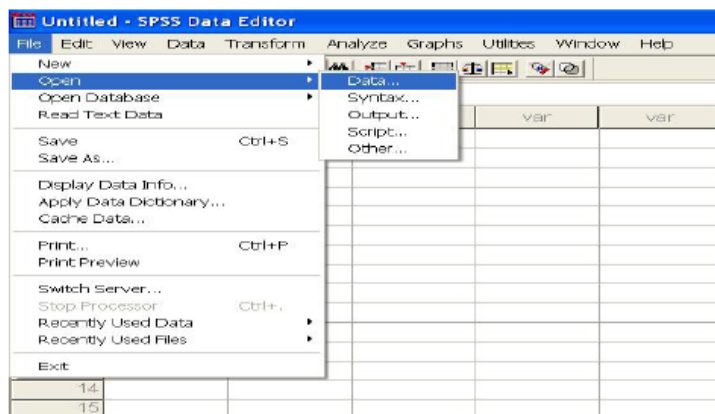
برای ذخیره کردن فایلها در SPSS همانند تمام برنامه های تحت ویندوز از منوی File استفاده می شود:



به جای مسیر فوق میتوانید از کلید میانبر زیر استفاده کنید:



برای باز کردن یک فایل حاوی داده های SPSS از مسیر زیر استفاده کنید:



باز کردن داده ها: ورود داده ها به محیط نرم افزار

به جای مسیر فوق می‌توانید از کلید میانبر زیر استفاده کنید:



تحلیل آماری:

در اصطلاح عامیانه آمار به معنای ثبت و نمایش اطلاعات عددی در مورد یک موضوع مثلاً ثبت و نمایش تعداد بیکاران، تعداد تصادفات رانندگی، میزان محصولات کشاورزی، میزان صدور نفت، جمعیت شهر تهران و غیره می باشد. ولی علم آمار امروزه دارای مفهومی بسیار وسیعتر از این کاربرد عامیانه است. مفاهیم عامیانه آمار زیر مجموعه ای از آمار مصطلح بین آمار دانان است. از نقطه نظر علمی، آمار به مجموعه روشهایی برای جمع آوری تنظیم و خلاصه کردن داده های عددی و غیر عددی و انجام استنباط و نتیجه گیری بوسیله تجزیه و تحلیل آنها، اطلاق می شود.

با بیان دیگری می توان گفت که آمار عبارت است از هنر و علم جمع آوری، تعبیر، تجزیه و تحلیل داده ها و استخراج تعمیمهای منطقی در مورد پدیده های تحت بررسی.

با توجه به تعاریف بالا می توان گفت یک فرآیند تحلیل آماری شامل دو بخش عمده است. اولین قدم نمایش دادن و خلاصه کردن داده ها می باشد تا توجه ما روی ویژگیهای مهم داده ها متمرکز شود و جزئیات غیر ضروری کنار گذاشته شود. اما بخش دوم برای استخراج نکات کلی و استنباط هایی در مورد پدیده تحت مطالعه به کار می رود. بخش اول شامل روشهای آمار توصیفی و بخش دوم در برگیرنده روشهای موسوم به آمار استنباطی است.

آمار توصیفی به آن دسته از روشهای آماری گفته می شود که به پژوهشگر در طبقه بندی، خلاصه کردن، توصیف و تفسیر و برقراری ارتباط از طریق اطلاعات جمع آوری شده کمک می کند. مراحل اساسی توصیف داده ها عبارتست از:

الف) خلاصه کردن و توصیف الگوی کلی

۱) فشرده کردن داده ها در قالب جدول های آماری

۲) نمایش آنها بوسیله نمودار

ب) محاسبه شاخصهای آماری

نقش آمار توصیفی در فرآیند تحلیل آماری بسیار مهم و حیاتی است. آمار توصیفی با خلاصه کردن داده ها، ویژگیهای مهم آنها نمایان می سازد تا ایده های لازم را در ذهن پژوهشگر برای مرحله دوم تحلیل آماری

(آمار استنباطی) ایجاد کند. اینک مراحل مختلف آمار توصیفی را یک به یک و به طور مفصل بررسی می کنیم:

الف) خلاصه کردن و توصیف الگوی کلی:

یک مجموعه داده آماری شامل مجموعه ای از مقادیر یک یا چند متغیر است. متغیرها می توانند عددی یا رسته ای (Categorical) باشند. متغیرهای عددی خود به دو دسته گسسته و پیوسته دسته بندی می شوند.

این دسته بندی، روشهای آماری را که برای داده ها مناسب اند، مشخص می کند. یکی از روشهای خلاصه کردن و توصیف داده ها رسم یک نمودار آماری است. نوع نمودار مورد استفاده به نوع داده ها بستگی دارد و بسته به رسته ای بودن یا عددی بودن، نمودارهای مختلفی به کار برده می شود. جداول فراوانی هم بسته به نوع متغیر، متفاوت خواهند بود، لذا مراحل فوق را برای انواع مختلف متغیر، جداگانه بررسی خواهیم کرد.

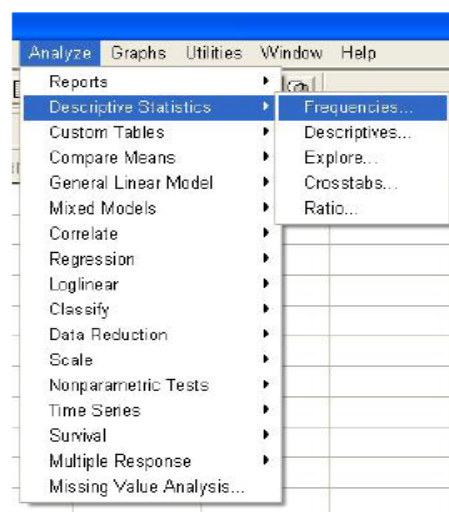
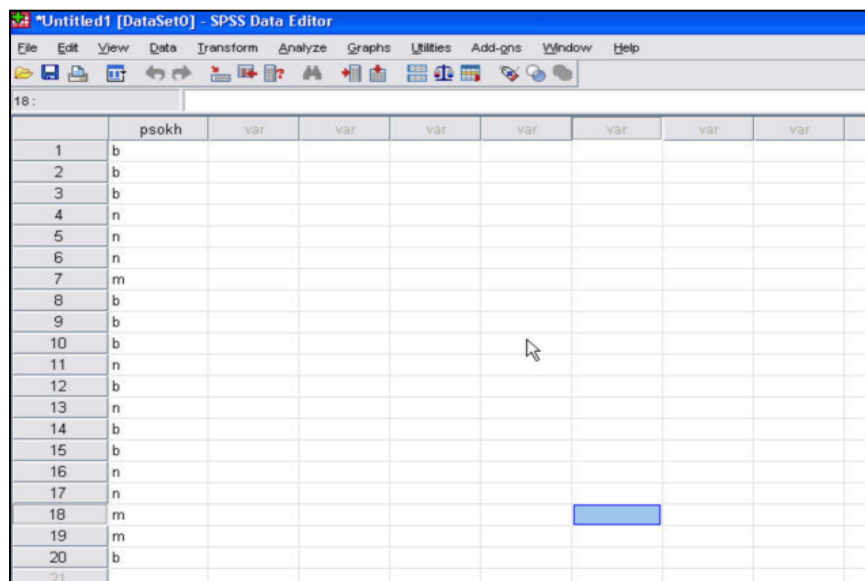
الف - ۱) داده های رسته بندی شده (Categorical Data):

متغیرهای رسته ای به آن دسته از متغیرها اطلاق می شود که از نظر کیفی مقادیر آن به چندین رسته تقسیم می شود. برای مثال جنسیت، رنگ پوست، رشته تحصیلی، رتبه شغلی، شغل و... نمونه هایی از متغیرهای رسته ای هستند. متغیرهای رسته ای به دو دسته کلی کیفی/اسمی و کیفی/رتبه ای تقسیم میشوند.

-جدول فراوانی برای متغیرهای رسته بندی شده:

جداول فراوانی این نوع متغیرها، با فهرست کردن مقادیر مختلف متغیر، فراوانی مربوط به هر مقدار و درصد فراوانی هر مقدار، بدست خواهد آمد، با یک مثال نحوه ساختن این نوع جداول فراوانی را با SPSS می بینیم:

مثال: پرسشنامه ای از یک روستا در ارتباط با میزان رضایتمندی از وام مسکن روستاییان پرسش شده است. جامعه آماری مورد مطالعه ۲۰ نفر می باشند. ۱۱ نفر پاسخ بلی، ۶ نفر خیر و ۳ نفر پاسخ نداده اند. جدول فراوانی آن را ترسیم نمایید:



برای رسم جدول فراوانی، در کادر کنار عبارت Display frequency tables تیک بگذارید. و در

نهایت دکمه  را کلیک کنید. صفحه جداگانه ای تحت عنوان Output view باز خواهد شد

که خروجی های SPSS همواره در آن ظاهر خواهد شد. خروجی این مثال به صورت زیر خواهد بود:

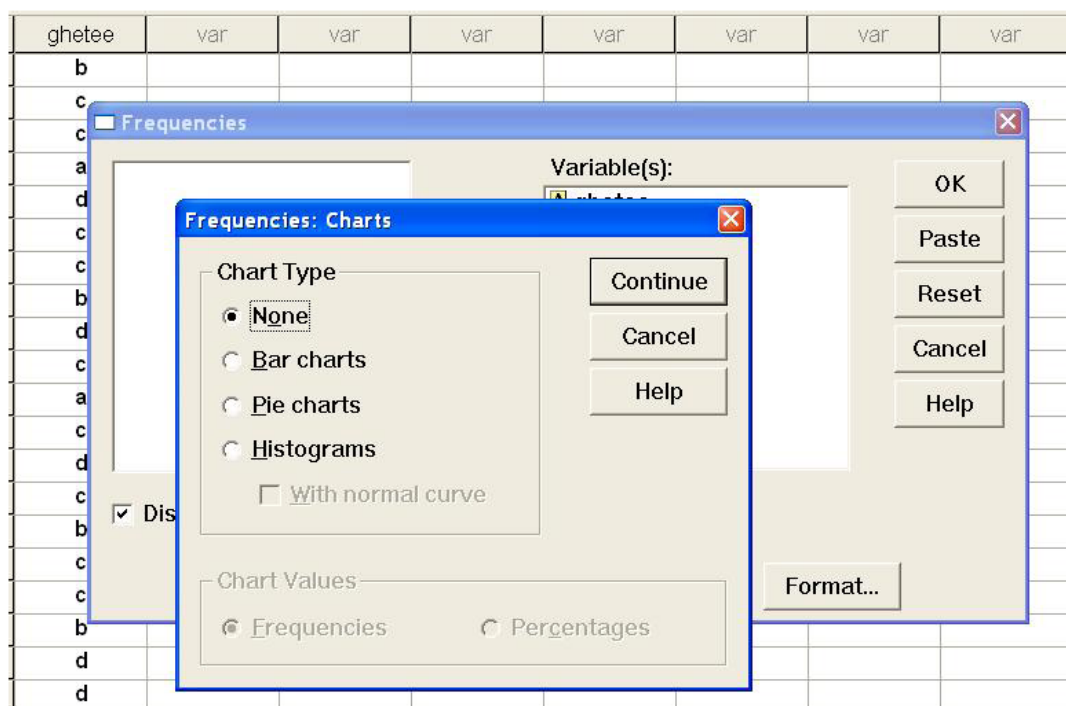
➔ Frequencies

Statistics

N	Valid	20
	Missing	0

psokh					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	b	10	50.0	50.0	50.0
	m	3	15.0	15.0	65.0
	n	7	35.0	35.0	100.0
	Total	20	100.0	100.0	

۱. از مسیری که برای رسم جدول فراوانی طی کردیم رفته روی دکمه **Charts...** کلیک کنید تا کادر گفتگوی زیر باز شود:

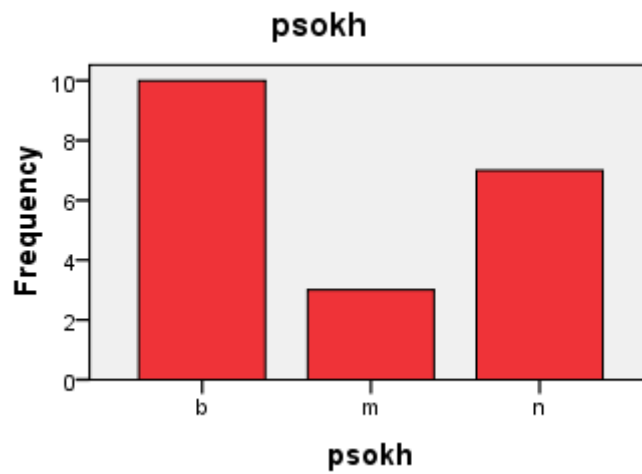
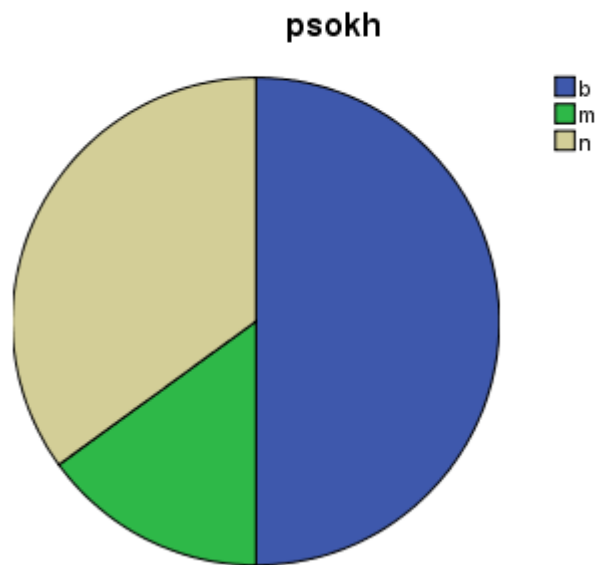


هر کدام از نمودارهای pie chart یا Bar chart را که مایل بودید انتخاب کرده و دکمه

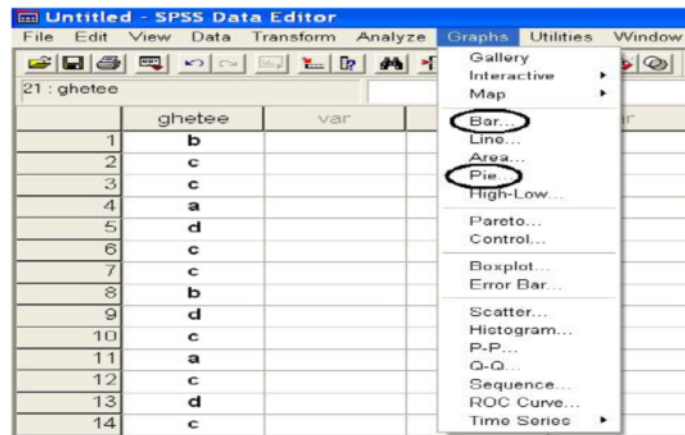
را کلیک کنید. دکمه **Continue** را کلیک کنید.

مسیر فوق را برای داده های مثال طی کنید و خروجی را ببینید:

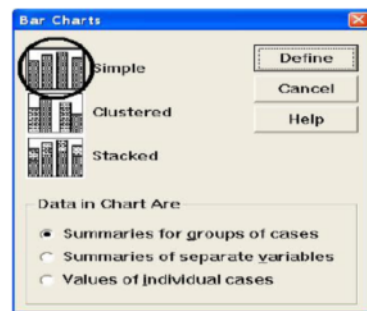
توجه کنید که نمودارهای دایره ای برای متغیرهای کیفی اسمی مناسبترند و نمودارهای میله ای برای متغیرهای کیفی رتبه ای.



۲- راه دوم استفاده از منوی Graph است:



در کادر باز شده روی نوع simple کلیک کرده و دکمه Define را کلیک کنید:



الف - ۲) داده های عددی (Numerical Data)

داده های عددی دو نوع اند: گسسته و پیوسته. مقادیر متغیرهای گسسته اعداد حاصل از شمارش می باشد برای مثال یک خانواده می تواند یک یا دو فرزند داشته باشد اما تعداد فرزندان خانواده نمی تواند عددی ما بین این دو باشد. و در سمت مقابل، متغیرهای پیوسته فاقد واحدهای تفکیک پذیر می باشند. برای مثال وزن یک متغیر پیوسته است.

- متغیر عددی گسسته:

چون مقادیر یک متغیر گسسته، جدا از هم و معمولاً محدود است برای رسم جدول فراوانی یک متغیر عددی گسته همچون حالت متغیر رسته ای عمل می کنیم. اما اگر تعداد مقادیر متفاوتی که یک متغیر گسسته می گیرد زیاد باشد، برای رسم جدول فراوانی، با آن مثل یک متغیر پیوسته رفتار خواهیم کرد.

ب) محاسبه شاخصهای آماری:

مرحله اول آمار توصیفی یعنی تشکیل جداول فراوانی و رسم نمودارهای آماری در SPSS بیان شد. مرحله دوم در آمار توصیفی، خلاصه کردن داده ها در قالب اعدادی است که موسوم به **شاخصهای آماری** هستند.

شاخصهای آماری به دو دسته تقسیم میشوند: شاخصهایی که گرایش به مرکز یا مرکزیت داده ها را اندازه میگیرند (شاخصهای مرکزی) و شاخصهایی که برای اندازه گیری تغییر پذیری داده ها به کار میروند (شاخصهای پراکندگی).

- شاخصهای مرکزی:

شاخصهای مرکزی مهم عبارتند از:

مد (Mode): مد داده ای است که بیشترین فراوانی را دارد. استفاده از این شاخص بیشتر در متغیرهای رسته ای است.

میان (Median): و **چندکها:** میان به داده وسطی داده ها اطلاق میشود و در داده های کم تعداد یک شاخص پر کاربرد و کار آمد است.

میان داده ای است که تقریباً نصف داده ها از آن کمتر و نصف داده ها از آن بیشترند.

تعریف چندکها هم معادل میان است، چندک مرتبه p ، مقداری است که تقریباً $100 \times p$ درصد داده ها از آن کمتری مساوی آن و $100 \times (1-p)$ درصد داده ها از آن بیشترند. ساده ترین نوع چندکها، چارکها (Quartiles) و دهکها هستند.

- چارک اول: مقداری است که یک چهارم داده ها از آن کمتر یا مساوی با آن هستند.

- چارک دوم: معادل میان است.

- چارک سوم: مقداری است که سه چهارم داده ها از آن کمتر یا مساوی با آن هستند.

- دهک اول: مقداری است که یک دهم داده ها از آن کمتر یا مساوی با آن هستند. سایر دهکها هم به همین صورت تعریف می شوند.

میانگین (Mean): پرکاربردترین و کاراترین شاخص برای اندازه گیری مرکزیت داده ها میانگین است. البته در صورتیکه تعداد داده ها کم باشد یا تعدادی داده پرت در میان داده ها مشاهده شود، دقت میانگین کاهش خواهد یافت لذا در صورتیکه یکی از حالات فوق اتفاق بیفتد باید در استفاده از میانگین هوشیار بود.

-شاخصهای پراکندگی:

غیر از شاخصهایی که گرایش داده ها را به یک مقدار مرکزی نشان میدهد، علاقه مند به شاخصهایی هستیم که به نوعی میزان پراکندگی داده ها را بیان کنند. مهمترین شاخصهای آماری پراکندگی عبارتند از:

دامنه تغییرات (Range): تفاضل بزرگترین و کوچکترین داده را دامنه تغییرات می نامند.

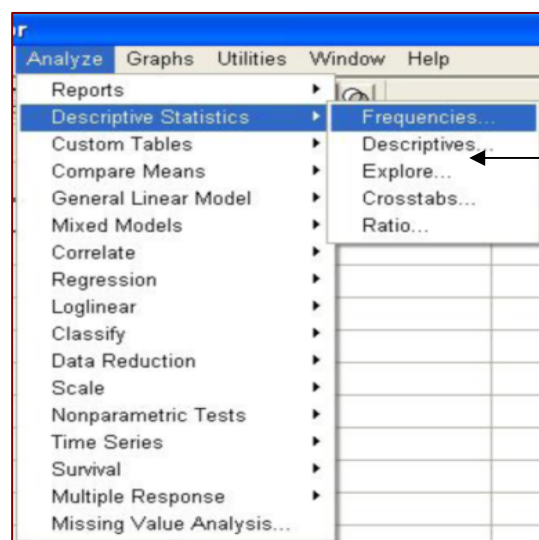
واریانس (Variance): میانگین مربعات تفاضل داده ها از میانگین را واریانس گویند.

انحراف معیار (Standard division): جذر واریانس را انحراف معیار گویند.

انحراف استاندارد میانگین (Standard Error of Mean): جذر حاصل تقسیم واریانس بر

تعداد داده ها را انحراف استاندارد میانگین گویند.

برای محاسبه شاخصهای بالا، ابتدا مسیر زیر را طی کنید:



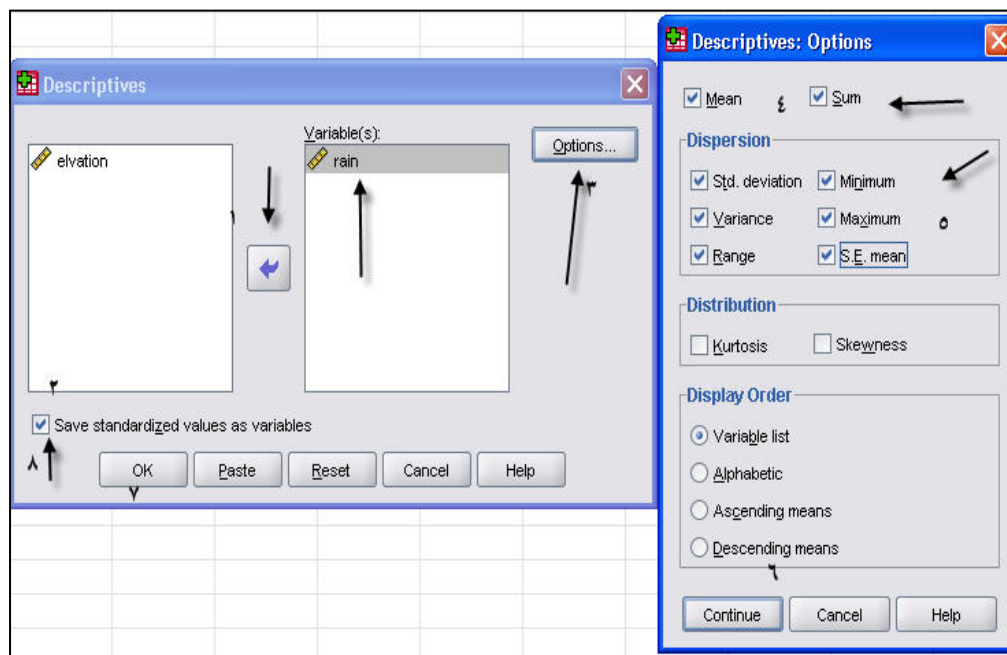
Statistics...

روی دکمه کلیک کنید. کادر زیر باز خواهد شد. شاخصهای مرکزی را میتوانید از کادر

Central Tendency و شاخصهای پراکندگی را از کادر Dispersion انتخاب کنید.

پس از انتخاب شاخصهای مورد نظر دکمه Continue و سپس OK را کلیک کنید.

روی دکمه **Statistics...** کلیک کنید. کادر زیر باز خواهد شد. شاخصهای مرکزی را میتوانید از کادر Central Tendency و شاخصهای پراکندگی را از کادر Dispersion انتخاب کنید. پس از انتخاب شاخصهای مورد نظر دکمه Continue و سپس OK را کلیک کنید.

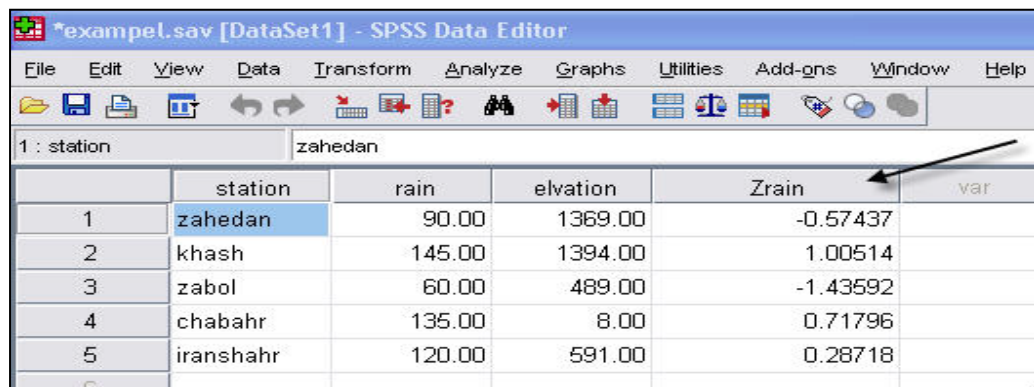


نکته: گزینه شماره ۸ جهت محاسبه نمره استاندارد داده ها انتخاب می گردد.

پس از طی مراحل بالا جدول که شامل شاخص های پراکندگی و مرکزی است، خروجی نرم افزار آشکار می گردد.

Descriptive Statistics									
	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
rain	5	85.00	60.00	145.00	550.00	1.1000E2	15.57241	34.82097	1.212E3
Valid N (listwise)	5								

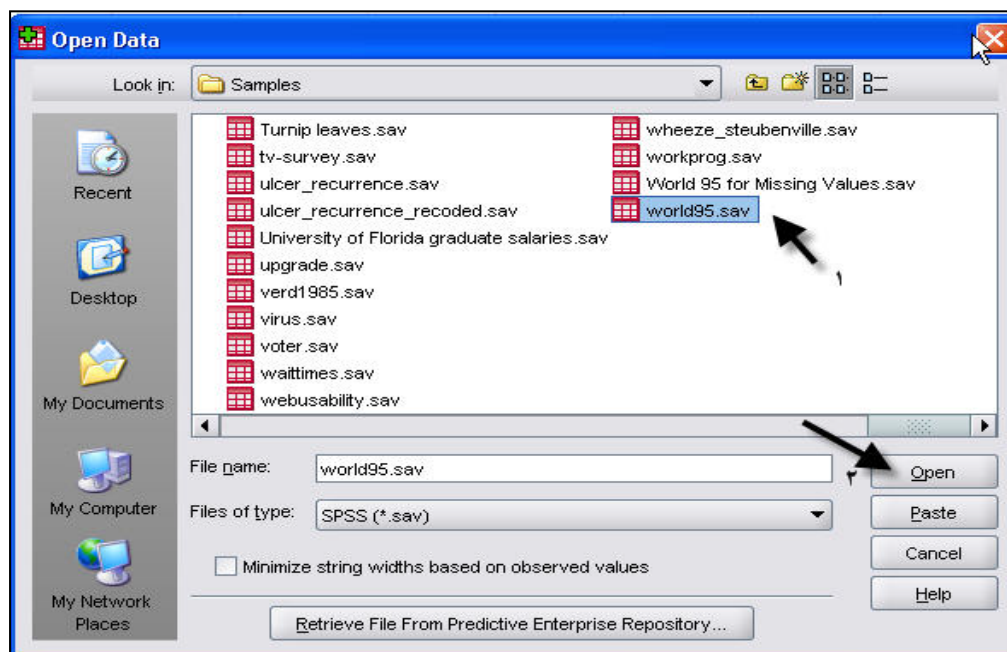
نمره استاندارد محاسبه شده برای داده ها در یک ستون در data view ایجاد می گردد. شکل زیر



	station	rain	elvation	Zrain	var
1	zahedan	90.00	1369.00	-0.57437	
2	khash	145.00	1394.00	1.00514	
3	zabol	60.00	489.00	-1.43592	
4	chabahr	135.00	8.00	0.71796	
5	iranshahr	120.00	591.00	0.28718	

دادهای وارد شده مربوط به متغیرهای مختلف کشورهای جهان می باشد که از مثال های نرم افزار بوده و از مسیر زیر وارد می شوند.

C:\Program Files\SPSSInc\SPSS16\Samples\world95



country	country	populn	density	urban	religion	lifeexpf	lifeexpm	literacy	pop_incr	babymort	gdp_cap	region	calories	aids
1	Afghanistan	20500	25.0	18	Muslim	44	45	29	2.8	168.0	205	3	.	.
2	Argentina	33900	12.0	86	Catholic	75	68	95	1.3	25.6	3408	6	3113	390
3	Armenia	3700	126.0	68	Orthodox	75	68	98	1.4	27.0	5000	5	.	.
4	Australia	17800	2.3	85	Protstnt	80	74	100	1.4	7.3	16848	1	3216	472
5	Austria	8000	94.0	98	Catholic	79	73	99	0.2	6.7	18396	1	3495	115
6	Azerbaijan	7400	86.0	54	Muslim	75	67	98	1.4	35.0	3000	5	.	.
7	Bahrain	600	828.0	83	Muslim	74	71	77	2.4	25.0	7875	5	.	.
8	Bangladesh	125000	800.0	16	Muslim	53	53	35	2.4	106.0	202	3	2021	.
9	Barbados	256	605.0	45	Protstnt	78	73	99	0.2	20.3	6950	6	.	41
10	Belarus	10300	60.0	65	Orthodox	76	66	99	0.3	19.0	6500	2	.	.
11	Belgium	10100	329.0	96	Catholic	79	73	99	0.2	7.2	17912	1	.	160
12	Bolivia	7900	6.9	51	Catholic	64	59	78	2.7	75.0	730	6	1916	6
13	Bosnia	4600	87.0	36	Muslim	78	72	86	0.7	12.7	3098	2	.	.
14	Botswana	1359	2.4	25	Tribal	66	60	72	2.7	39.3	2677	4	2375	141
15	Brazil	156600	18.0	75	Catholic	67	57	81	1.3	66.0	2354	6	2751	4931
16	Bulgaria	8900	79.0	68	Orthodox	75	69	93	0.2	12.0	3631	2	.	2
17	Burkina Faso	10000	36.0	15	Animist	50	47	18	2.8	118.0	357	4	2288	415
18	Burundi	6000	216.0	5	Catholic	50	46	50	2.3	105.0	208	4	1932	722
19	Cambodia	10000	55.0	12	Buddhist	52	50	35	2.9	112.0	260	3	2166	.
20	Cameroon	13100	27.0	40	Animist	59	55	54	2.9	77.0	593	4	2217	303

برای محاسبه همبستگی از مسیر زیر وارد می شویم و سپس پنجره زیر باز می شود.

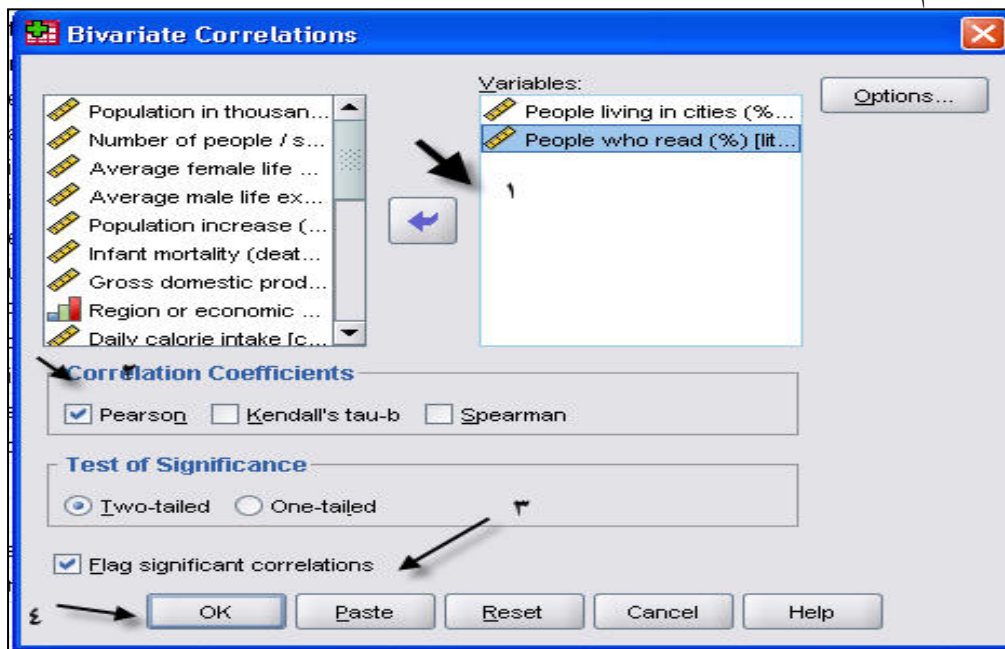
Analysis\Bivariate\Correlation

در پنجره زیر داده ها را به سمت راست منتقل و سپس نوع ضریب همبستگی را انتخاب و

گزینه flag را نیز فعال می نمایم.

در این مثال می خواهیم ضریب همبستگی بین میزان سواد و شهرنشینی را در کشورهای جهان

محاسبه نمایم.



پس از محاسبه جدول زیر در خروجی باز می شود. میزان همبستگی ۶۵ درصد محاسبه گردیده است

که رابطه بالایی بین میزان سواد و شهرنشینی در کشورهای جهان وجود دارد.

Correlations

[DataSet1] C:\Program Files\SPSSInc\SPSS16\Samples\world95.sav

Correlations		People living in cities (%)	People who read (%)
People living in cities (%)	Pearson Correlation	1	.650**
	Sig. (2-tailed)		.000
	N	108	107
People who read (%)	Pearson Correlation	.650**	1
	Sig. (2-tailed)	.000	
	N	107	107

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

جهت فراوانی گرفتن از داده های پیوسته باید آنها را Recode نمود تا بتوان فراوانی آنها را مشاهده نمود.

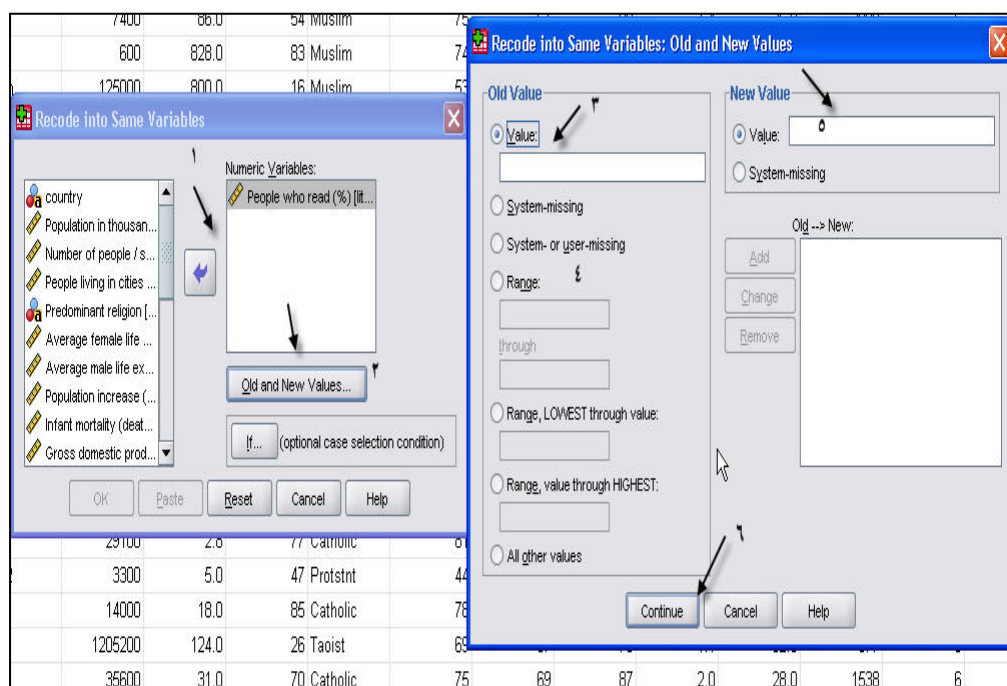
مثال: فراوانی درصد سواد کشورها.

در نرم افزار از مسیر زیر Recode را انتخاب می نمایم.

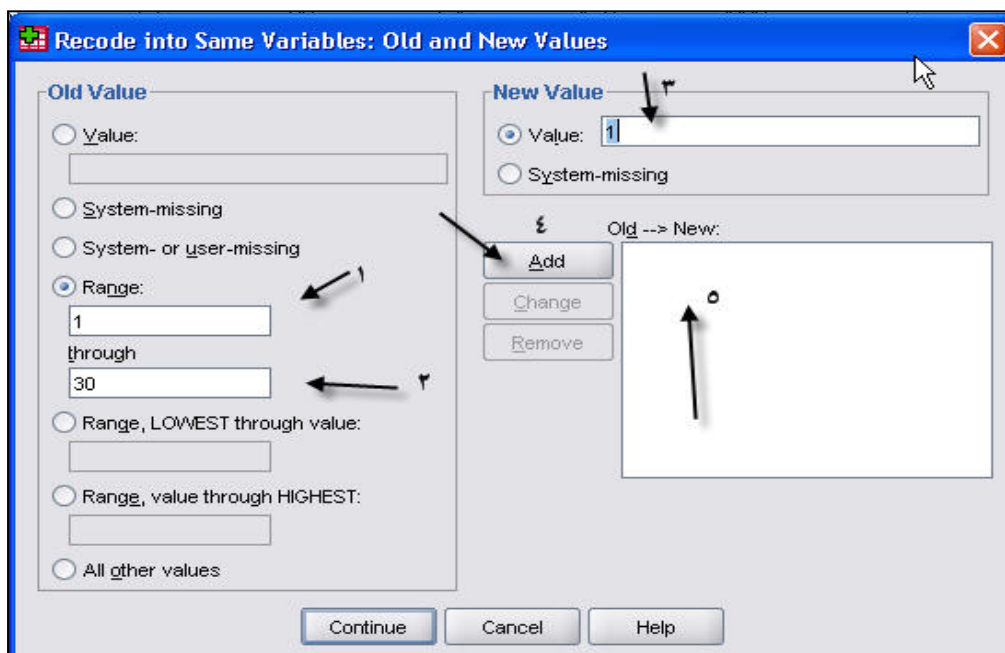
Transform\Recode into same variable

پنجره زیر باز شده در قسمت ۱ داده را به سمت راست انتقال داده و سپس قسمت ۲ گزینه old and new value را

انتخاب کنید. پنجره سمت راست باز می شود.

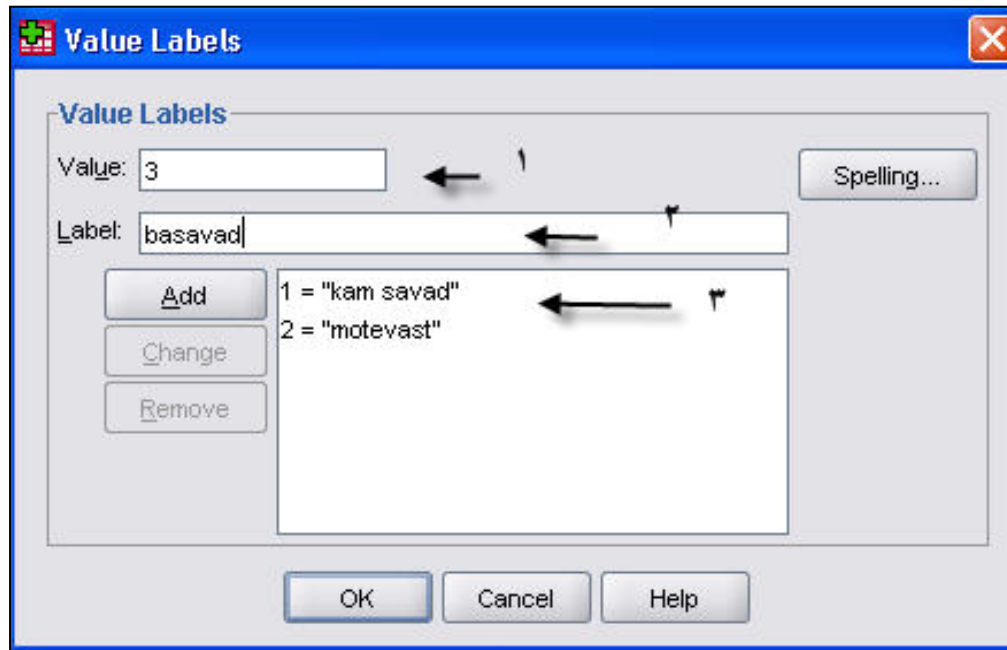


۱. دامنه داده ها را باید تعریف نماییم. برای طبقه بندی داده ها باید حداقل و حداکثر داده را بدانیم و به تعداد طبقات آگاهی داشته باشیم. در این مثال حداقل ۱ و حداکثر ۱۰۰ می خواهیم آنها را به سه طبقه تقسیم نماییم. ۱ تا ۳۰، ۳۰ تا ۶۰ و ۶۰ تا ۱۰۰.
۲. در کادر ۱ عدد ۱ و در کادر ۲ عدد ۳۰ را وارد می کنیم.
۳. به طبقه یک ارزش ۱ می دهیم و در value وارد می کنیم.
۴. گزینه ۴ را کلیک می کنیم تا طبقه با ارزش آن به کادر ۵ وارد شود.
۵. همین کار را برای دو طبقه دیگر انجام می دهیم.



پس از انجام مراحل بالا گزینه continue را کلیک می کنیم. در ستون داده ارزش ها وارد می گردد.

در این مرحله باید برای ارزشها برچسب تعریف نماییم. در پنجره variable data در ردیف داده مورد نظر در ستون value برای هر یک از ارزشها برچسب تعریف می نماییم. در پنجره زیر ابتدا در قسمت value ارزش و در قسمت label نام طبقه را وارد می کنیم و گزینه add را کلیک می کنیم. برای این مثال ارزش ۱ برچسب کم سواد، ۲ کشورهای با سواد متوسط و ۳ باسواد.



پس از اضافه کردن ارزشها و برچسب ها از داده فوق از دستور frequency فراوانی گرفته و جدول زیر باز می شود. طبق آن ۵ درصد کشورها که شامل ۶ کشور هستند کم سواد و ۷۵ درصد باسواد می باشند.

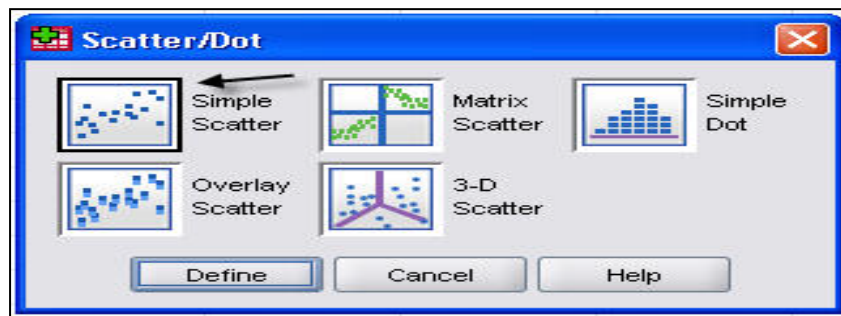
People who read (%)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	kam savad	6	5.5	5.6	5.6
	motevast	19	17.4	17.8	23.4
	basavad	82	75.2	76.6	100.0
	Total	107	98.2	100.0	
Missing	System	2	1.8		
Total		109	100.0		

رسم نمودار Scatter plot:

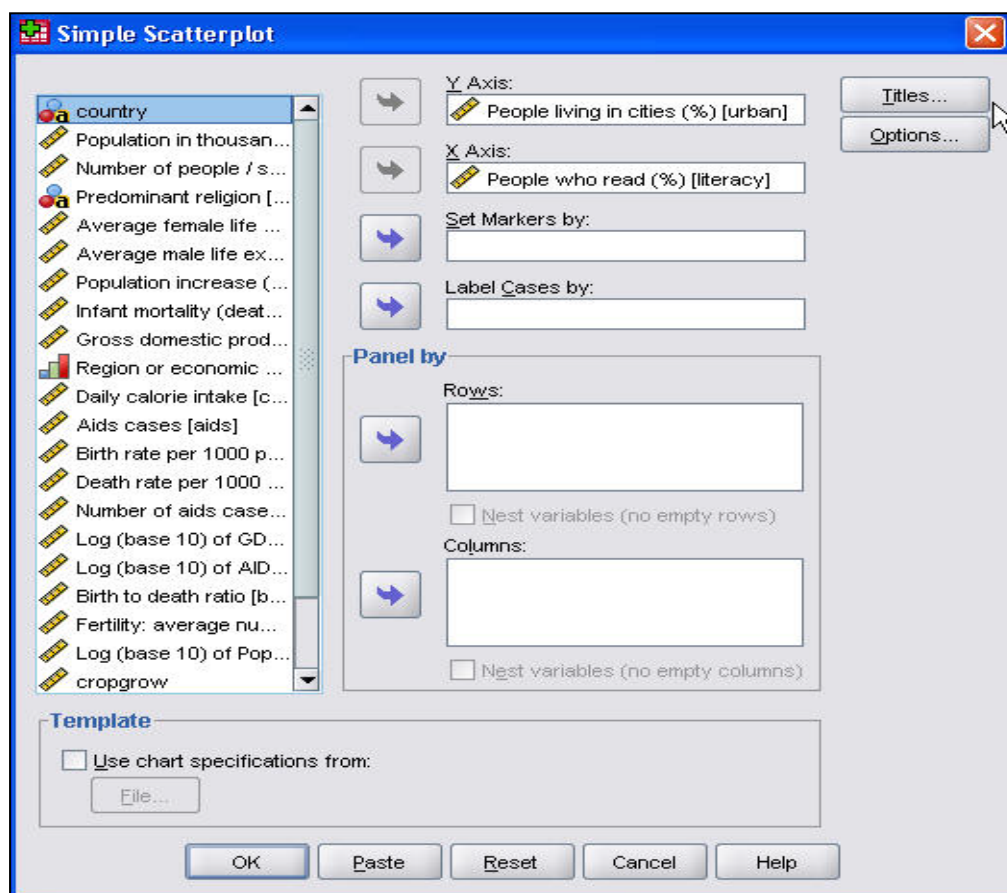
از این نمودار برای رابطه بین دو متغیر استفاده می شود. برای ترسیم آن از مسیر زیر وارد می شود.

Graphs\legacy Dialogs\Scatter plot

پنجره زیر باز می شود.



متغیرها را طبق شکل زیر وارد نمایید. میزان سواد را در X و میزان شهرنشینی را در Y سپس گزینه Ok را کلیک کنید.



نمودار در خروجی نرم افزار ظاهر شده . روی آن دابل کلیک کرده و در پنجره باز شده گزینه Fit line را کلیک کنید. مشاهده می کنید که رابطه مستقیمی بین افزایش شهرنشینی و افزایش میزان سواد وجود دارد.

