

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

مؤسسه آموزش عالی آزاد ماهان با افتخار تقدیم می کند

کاربرد آمار و روش تحقیق

مجموعه ارتباطات اجتماعی

از سری کتابهای کمک آموزشی کارشناسی ارشد

حسین صنعتی  
پریسا حاج کریمی

موسسه آموزش عالی آزاد



سرشناسه	صنعتی، حسین
عنوان	کاربرد آمار و روش تحقیق / حسین صنعتی-پریسا حاج کریمی
مشخصات نشر	تهران، مشاوران صعود ماهان، ۱۳۹۵
مشخصات ظاهری	۳۹۲ صفحه رحلی
فروست	سری کتاب‌های کمک آموزشی کارشناسی ارشد
شابک	۹۷۸-۶۰۰-۳۸۹-۲۴۶-۰
وضیعت فهرست‌نویسی	فیبای مختصر
یادداشت	این مدرک در آدرس <a href="http://opac.nali.ir">http://opac.nali.ir</a> قابل دسترسی است
شناسه افزوده	حاج کریمی، پریسا
شماره کتاب‌شناسی ملی	۳۲۰۰۳۴۱



مجموعه ارتباطات اجتماعی

## کاربرد آمار و روش تحقیق

ناشر	مشاوران صعود ماهان
مدیر مسئول	هادی سیاری - مجید سیاری
مدیر برنامه‌ریزی و تولید	محسن صداقت
به قلم	حسین صنعتی - پریسا حاج کریمی
ویراستار ادبی	زهرا سروری
گرافیکست	حامد شاملو
صفحه‌آرا	فاطمه طاهر
طراح جلد	سمیرا خانزاد
نوبت و تاریخ چاپ	اول / ۱۳۹۵
شمارگان	۲۰۰۰ نسخه
قیمت	۳۲۰/۰۰۰ ریال
شابک	۹۷۸-۶۰۰-۳۸۹-۲۴۶-۰

نشانی: تهران، خیابان  
ولی‌عصر - بالاتر از  
تقاطع مطهری - جنب  
بانک ملی پلاک ۲۰۵۰  
شماره تماس:  
۸۸۱۰۰۱۱۳-۴

### «ن والقلم وما یسطرون»

کلمه نزد خدا بود و خدا آن را با قلم بر ما نازل کرد. به پاس تشکر از چنین موهبت الهی، مؤسسه ماهان درصدد برآمده است تا در راستای انتقال دانش و مفاهیم با کمک اساتید مجرب و مجموعه کتب آموزشی خود برای شما داوطلبان ادامه تحصیل در مقطع کارشناسی ارشد، گام مؤثری بردارد. امید است تلاش‌های خدمتگزاران شما در این مؤسسه پایه‌گذار گام‌های بلند فردای شما باشد.

مجموعه کتاب‌های کمک آموزشی ماهان به‌منظور استفاده داوطلبان کنکور کارشناسی ارشد سراسری و آزاد تألیف شده‌است. در این کتاب‌ها سعی کرده‌ایم با بهره‌گیری از تجربه اساتید بزرگ و کتب معتبر داوطلبان را از مطالعه کتاب‌های متعدد در هر درس بی‌نیاز کنیم.

دیگر تألیفات ماهان برای دانشجویان به‌صورت ذیل است: مجموعه کتاب‌های ۸ آزمون: شامل ۵ مرحله کنکور کارشناسی ارشد ۵ سال اخیر به همراه ۳ مرحله آزمون تألیفی ماهان همراه با پاسخ تشریحی است که برای آشنایی با نمونه سؤالات کنکور طراحی شده است. این مجموعه کتاب‌ها با توجه به تحلیل ۳ ساله اخیر کنکور و بودجه‌بندی مباحث در هریک از دروس، اطلاعات مناسبی برای برنامه‌ریزی درسی در اختیار دانشجو قرار می‌دهد.

مجموعه کتاب‌های کوچک: شامل کلیه نکات کاربردی در گرایش‌های مختلف کنکور کارشناسی ارشد است که برای دانشجویان به‌منظور جمع‌بندی مباحث در ۲ ماهه آخر قبل از کنکور مفید است.

بدین‌وسیله از مجموعه اساتید، مؤلفان و همکاران محترم خانواده بزرگ ماهان که در تولید و بروزرسانی تألیفات ماهان نقش مؤثری داشته‌اند، صمیمانه تقدیر و تشکر می‌نماییم. دانشجویان عزیز و اساتید محترم می‌توانند هرگونه انتقاد و پیشنهاد در خصوص تألیفات ماهان را از طریق سایت ماهان به آدرس [mahan.ac.ir](http://mahan.ac.ir) با ما در میان بگذارند.

مؤسسه آموزش عالی آزاد ماهان

## مقدمه مؤلف

دوست گرامی!

با عرض سلام و احترام

سعی ما در این کتاب بر این بوده است که مهم‌ترین مطالب را از منابع آزمون‌های درس کاربرد آمار و روش تحقیق در علوم ارتباطات اجتماعی فراهم آوریم و با زبانی روان تشریح کرده و تقدیم شما نماییم.

هریک از بخش‌ها و یا فصل‌های این کتاب دارای تست‌های تمرینی است که هدف از طرح آنها کمک به خواننده در افزایش تسلط بر مطالب مربوطه از زوایای مختلف است. در آخر هر کدام از بخش‌ها نیز سوالات کنکوری سراسری و آزاد مربوط به همان بخش درج شده و به صورت مشروح پاسخ داده شده‌اند.

در تدوین کتاب حاضر، آخرین تغییرات در منابع طرح سوالات آزمون‌های لحاظ شده و مطالب آزمون‌های مفید موجود در منابع جدید نیز درج شده‌اند. چنانچه شما عزیزان مایل به حصول اطلاع از آخرین تغییرات در منابع این درس و یا رفع سوالات و ابهامات خود در خصوص این کتاب باشید، می‌توانید از طریق پست الکترونیکی با نگارنده این سطور تماس حاصل فرمایید.

صمیمانه آرزوی موفقیت شما را داشته و انتقادات و پیشنهاداتتان برای ارتقای کیفیت این مجموعه را به دیده منت پذیرایییم.

حسین صنعتی

Hosein.sanati62@gmail.com

# فهرست مطالب

## بخش اول: آمار

### فصل اول: مفاهیم اساسی آمار

تعریف علم آمار	۹
آمار توصیفی و آمار استنباطی	۹
تعریف برخی از مفاهیم آماری	۱۰
متغیر و انواع آن	۱۰
تحلیل مسیر	۱۱
مقیاس اندازه‌گیری	۱۲
کارکردهای آمار توصیفی	۱۳
سوالات چهارگزینه‌ای فصل اول	۱۴

### فصل دوم: آماده‌سازی، سازمان‌دهی و نمایش داده‌ها (نمودارها)

آماده‌سازی داده‌ها	۱۵
تشکیل نظام کدگذاری	۱۵
دستکاری داده‌ها پیش از تحلیل	۱۷
سازماندهی داده‌ها	۱۷
نمایش داده‌ها	۲۰
سوالات چهارگزینه‌ای فصل دوم	۲۳

### فصل سوم: نشان دادن مرکزیت داده‌ها (پارامترهای مکان مرکزی)

نمایا مد	۲۵
میانگین	۲۷
میانگین	۲۹
معادلات پیرسون	۳۱
رابطه بین نما، میانگین، میانگین	۳۳
شاخص‌های وضعی	۳۴
چارک‌ها	۳۴
محاسبه صدک‌ها و دهک‌ها	۳۵
رتبه درصدی	۳۸
سوالات چهارگزینه‌ای فصل سوم	۴۰

### فصل چهارم: پارامترهای پراکندگی

دامنه تغییرات	۴۵
انحراف چارکی	۴۶
انحراف متوسط (انحراف میانگین = $M.d$ یا $AD$ )	۴۷
واریانس	۴۷
انحراف معیار (انحراف استاندارد)	۴۹
تخصیص شپرد	۵۱
ضریب تغییرات	۵۱
نسبت تغییرات	۵۲
شاخص تغییرات کیفی	۵۳
گشتاور	۵۳
شاخص چولگی	۵۳
شاخص کشیدگی	۵۴
توزیع و منحنی نرمال (متقارن)	۵۵
سطوح زیر منحنی طبیعی استاندارد	۵۶
سوالات چهارگزینه‌ای فصل چهارم	۵۷

### فصل پنجم: آمار استنباطی (برآورد و آزمون فرضیه) و ضرایب همبستگی

برآورد	۶۱
خطای نمونه‌گیری	۶۲
محدوده اطمینان یا درجه اطمینان (d)	۶۲
اشتباه استاندارد میانگین	۶۲
نمره‌های استاندارد	۶۲

آزمون فرضیه

.....	۶۴
انواع خطا	۶۵
توان آزمون	۶۶
توصیف روابط بین متغیرها	۶۷
آزمون خی‌دو	۶۸
انواع ضرایب همبستگی	۷۰
ضریب همبستگی متغیرهای فاصله‌ای - نسبی	۷۷
کوواریانس	۷۹
تحلیل رگرسیون	۷۹
انواع آزمون‌های آماری	۸۲
آزمون تحلیل واریانس	۸۵
دامنه استودنت شده	۹۰
طریق خواندن جداول	۱۰۰
انواع تقسیم‌بندی ضرایب همبستگی	۱۰۳
سوالات چهارگزینه‌ای فصل پنجم	۱۰۴

### فصل ششم: تخمین آماری پارامترهای جامعه

برآورد	۱۱۷
خواص مطلوب برای آماره‌ها	۱۱۷
ویژگی‌های برآورد کننده	۱۱۸
تخمین فاصله‌ای	۱۱۹
تخمین فاصله‌ای تفاضل، میانگین دو جامعه	۱۲۲
تخمین فاصله‌ای نسبت موقعیت جامعه	۱۲۶
تخمین فاصله‌ای تفاضل نسبت موقعیت دو جامعه	۱۲۶
تخمین فاصله‌ای واریانس جامعه	۱۲۷
تعیین حجم نمونه برای برآورد میانگین جامعه	۱۲۹
تعیین حجم نمونه برای برآورد نسبت موقعیت جامعه	۱۳۰
سوالات چهارگزینه‌ای فصل ششم	۱۳۰

### فصل هفتم: احتمال

فاکتوریل	۱۳۷
مفاهیم مرتبط با احتمال	۱۳۷
نکات مهم بسط چند جمله‌ای	۱۴۱
امید ریاضی	۱۴۸
همبستگی	۱۵۱
توزیع‌های احتمالی گسسته	۱۵۳
سوالات چهارگزینه‌ای فصل هفتم	۱۵۵
فصل هشتم: نکات آماری مهم	
نکات آماری مهم	۱۶۳
سوالات چهارگزینه‌ای فصل هشتم	۱۷۳

### بخش دوم: روش تحقیق

#### فصل نهم: شناخت، علم، روش

شناخت	۱۸۷
دیدگاه‌های مربوط به واقعیت	۱۸۹
دیدگاه پسامدرن	۱۸۹
علم	۱۸۹
مبانی علم اجتماعی	۱۹۰
انواع علم	۱۹۰
رهیافت‌های روش شناختی	۱۹۱
انواع روش در علوم اجتماعی	۱۹۱
مروری بر برخی مکاتب فلسفه علم	۱۹۴
نظریه اجتماعی و نظریه‌پردازی	۱۹۵

## فصل چهاردهم: روش های تحقیق و تکنیک های جمع آوری

### داده ها در تحقیقات اجتماعی

روش پیمایش.....	۲۸۷
طراحی پرسشنامه.....	۲۸۸
انواع پرسش ها.....	۲۸۸
نظم و ترتیب پرسش ها.....	۲۹۱
روش های اجرای پیمایش.....	۲۹۲
انواع مصاحبه های تحقیقی.....	۲۹۳
نقاط قوت و ضعف روش پیمایش.....	۲۹۴
مطالعه میدانی و مشاهده.....	۲۹۴
اقسام مشاهده.....	۲۹۵
انواع نقش های مشاهده گر.....	۲۹۷
انواع مطالعه میدانی.....	۲۹۷
انواع آزمایش.....	۲۹۹
اجزای کلی آزمایش.....	۲۹۹
انواع طرح های آزمایش.....	۳۰۱
آزمایش گروه های کانونی.....	۳۰۵
روش های تحقیق غیرمخل.....	۳۰۶
تحلیل محتوا.....	۳۰۶
تحلیل آمارهای موجود.....	۳۰۸
تحلیل ثانویه.....	۳۰۹
اقسام دیگر تکنیک های گردآوری اطلاعات.....	۳۱۰
آشنایی با تحقیقات کیفی و نحوه انجام روش های آن.....	۳۱۱
نمونه گیری.....	۳۱۳
ارزیابی گزارش تحقیق کیفی.....	۳۱۹
نکات تکمیلی و مهم فصل چهاردهم.....	۳۲۰
سوالات چهارگزینه ای فصل چهاردهم.....	۳۲۳

### پاسخنامه ها

پاسخنامه فصل اول.....	۳۳۴
پاسخنامه فصل دوم.....	۳۳۴
پاسخنامه فصل سوم.....	۳۳۵
پاسخنامه فصل چهارم.....	۳۳۹
پاسخنامه فصل پنجم.....	۳۴۴
پاسخنامه فصل ششم.....	۳۵۲
پاسخنامه فصل هفتم.....	۳۵۸
پاسخنامه فصل هشتم.....	۳۶۵
پاسخنامه فصل نهم.....	۳۷۷
پاسخنامه فصل دهم.....	۳۷۹
پاسخنامه فصل یازدهم.....	۳۸۱
پاسخنامه فصل دوازدهم.....	۳۸۳
پاسخنامه فصل سیزدهم.....	۳۸۵
پاسخنامه فصل چهاردهم.....	۳۸۶

منابع.....	۳۹۲
------------	-----

## مسئله یابی در تحقیق علمی و پاسخ گویی به مسئله پژوهش..... ۱۹۸

نکات تکمیلی و مهم فصل نهم.....	۱۹۹
سوالات چهارگزینه ای فصل نهم.....	۲۰۱

### فصل دهم: مراحل و عناصر اساسی تحقیق

انواع طرح های تحقیق.....	۲۰۷
هدف های تحقیق.....	۲۰۸
انواع مختلف تحقیق.....	۲۰۹
واحد های تحلیل.....	۲۱۱
مراحل انجام تحقیق علمی.....	۲۱۲
مفاهیم.....	۲۱۵
متغیرها.....	۲۱۸
فرضیه تحقیق.....	۲۲۱
آشنایی با طراحی پژوهش های اجتماعی.....	۲۲۴
پرسش ها و اهداف پژوهش.....	۲۲۶
استراتژی های پاسخگویی به پرسش های پژوهش.....	۲۲۶
مفاهیم، نظریه ها، فرضیه ها و مدل ها.....	۲۲۸
روش های پاسخگویی به پرسش های پژوهش.....	۲۳۱
نکات تکمیلی و مهم فصل دهم.....	۲۳۴
سوالات چهارگزینه ای فصل دهم.....	۲۳۵

### فصل یازدهم: نمونه گیری

انواع نمونه گیری.....	۲۴۱
نمونه گیری احتمالی.....	۲۴۱
انواع نمونه گیری احتمالی.....	۲۴۲
نمونه گیری غیراحتمالی.....	۲۴۳
حجم نمونه.....	۲۴۵
خطای نمونه گیری.....	۲۴۷
نکات تکمیلی و مهم فصل یازدهم.....	۲۴۸
سوالات چهارگزینه ای فصل یازدهم.....	۲۴۸

### فصل دوازدهم: سطوح سنجش و طیف ها (مقیاس ها)

اندازه گیری.....	۲۵۳
مقیاس ها (طیف ها).....	۲۵۵
انواع طیف.....	۲۵۶
مقایسه زوجی.....	۲۵۶
طیف بوگاردوس یا طیف فاصله اجتماعی.....	۲۵۷
طیف تورستن.....	۲۶۰
طیف لیکرت.....	۲۶۲
طیف گاتمن.....	۲۶۵
طیف ازگود.....	۲۷۰

### فصل سیزدهم: معیارهای مقبولیت تحقیق

اعتبار.....	۲۷۹
انواع اعتبار.....	۲۸۰
نکات تکمیلی و مهم فصل سیزدهم.....	۲۸۴
سوالات چهارگزینه ای فصل سیزدهم.....	۲۸۵

# بخش اول: آمار

## مهم ترین عناوین بخش

- ❑ مفاهیم اساسی آمار
- ❑ آماده سازی، سازمان دهی و نمایش داده ها (نمودارها)
- ❑ نشان دادن مرکزیت داده ها (پارامترهای مکان مرکزی)
- ❑ پارامترهای پراکندگی
- ❑ آمار استنباطی (برآورد و آزمون فرضیه) و ضرایب همبستگی
- ❑ تخمین آماری پارامترهای جامعه
- ❑ احتمال
- ❑ نکات آماری مهم





# فصل اول

## مفاهیم اساسی آمار

### تعریف علم آمار

آمار علمی است که با استفاده از فنون و روش‌های علمی و ریاضی به جمع‌آوری، طبقه‌بندی و پیش‌بینی اطلاعات کمی و کیفی و نتیجه‌گیری و تعمیم آنها در جهت هدف معین می‌پردازد.

آمار نه تنها در پژوهش‌های اجتماعی بلکه در سایر پژوهش‌های علمی نیز به مثابه وسیله است نه هدف. آمار شاخه‌ای از روش‌شناسی علمی است که با جمع‌آوری، طبقه‌بندی، توصیف و تفسیر داده‌هایی که از اجرای زمینه‌یابی‌ها و آزمایش‌ها به دست می‌آید، سر و کار دارد و هدف اساسی آن، توصیف و بیان استنباط‌هایی درباره ویژگی‌های عددی جوامع است.

به طور کلی آمار را می‌توان علم و عمل استخراج، بسط و توسعه دانش‌های تجربی انسانی با استفاده از روش‌های گردآوری، تنظیم، پرورش و تحلیل داده‌های تجربی دانست.

ویژگی منحصر به فرد آمار این است که نمودهای عددی بزرگ اجتماعی را در مرکز میدان دید خود می‌نهد و عمدتاً به پدیده‌هایی می‌پردازد که قابل شمارش، قابل اندازه‌گیری یا قابل کمیت‌یابی هستند.

### آمار توصیفی و آمار استنباطی

همان‌طور که از تعریف علم آمار معلوم می‌شود فعالیت آماری دو بخش را در برمی‌گیرد: الف - توصیف، ب - تبیین و برآورد (استنباط).

الف) جمع‌آوری، تلخیص، تنظیم و ارائه‌ی اطلاعات به صورت روشن و قابل درک و در صورت لزوم تعیین روابط موجود بین اطلاعات جمع‌آوری شده. این بخش از آمار را که بیشتر به مشخص کردن داده‌ها، تنظیم و ارائه آنها به صورت جدول‌بندی یا ترسیمی، محاسبه آمارها و تعیین ارتباط بین اطلاعات می‌پردازد، آمار توصیفی می‌نامند.

آمار توصیفی برای توصیف گروهی که مورد مطالعه‌ی محقق قرار گرفته‌اند، به کار می‌رود. هدف در آمار توصیفی، تقلیل، توصیف، تلخیص و نمایش روابط بین متغیرها است.

در واقع هنگامی که توده‌ای از اطلاعات کمی (مقداری) برای تفسیر گردآوری می‌شود، ابتدا لازم است آنها به صورتی که به روشنی قابل فهم و انتقال باشند سازمان بندی و خلاصه شوند که آمار توصیفی به همین منظور به کار برده می‌شود.

شاخص‌هایی که در آمار توصیفی به کار برده می‌شوند عبارتند از: مد، میانه، معدل و همبستگی. توجه شود که روش‌های آمار توصیفی همیشه برای تعیین و بیان ویژگی‌ها یا اطلاعاتی که به وسیله‌ی پژوهشگران جمع‌آوری شده‌اند به کار برده می‌شوند.

ب) چنانچه محقق قصد داشته باشد که با مطالعه بر روی نمونه‌ای منتخب از یک جامعه‌ی آماری، نتایج حاصله را به آن جامعه آماری تعمیم دهد، باید از روش‌های آمار استنباطی استفاده نماید. در بیشتر فعالیت‌های آماری جمع‌آوری، تنظیم و ارائه یافته‌ها و یا تعیین آمارها و روابط بین داده‌ها کفایت نمی‌کند، بلکه لازم است بر اساس این اطلاعات جمع‌آوری و تنظیم شده، تجزیه و تحلیل‌ها و استنباط‌هایی برای تبیین و تصمیم‌گیری صورت گیرد. این بخش از آمار که بر تحلیل، تفسیر و تعمیم نتایج حاصل از تنظیم و محاسبه مقدماتی آماری تأکید دارد، آمار استنباطی خوانده می‌شود.

در آمار استنباطی محقق به چیزی خارج از گروهی از موضوعات یا افرادی که در حقیقت مورد آزمون قرار می‌گیرند نظر دارد، و یا به عبارت دیگر، ویژگی‌های نمونه آماری را به جامعه آماری تعمیم می‌دهد. در آمار استنباطی فرض بر این است که مطالعه، یک مطالعه نمونه ای است. هدف عمده در آمار استنباطی، استنباط پارامتر از آماره است، و در واقع، در آمار استنباطی آماره‌ها را به پارامترها تعمیم می‌دهیم.



**نکته ۱:** در آمار توصیفی، آماره‌ها را به توصیف خصایص نمونه‌ها می‌پردازد و در آمار استنباطی مشخصات جامعه را از روی نمونه‌ها استنباط یا برآورد می‌کند.



**نکته ۲:** آمار توصیفی در خدمت آمار استنباطی است.

برای درک کامل اختلاف بین آمار توصیفی و آمار استنباطی نیاز به بحث در باره برخی از مفاهیم است.

### تعریف برخی از مفاهیم آماری



**جامعه آماری:** مجموعه‌ای از اشیاء، حوادث و افراد جامعه که در یک یا چند صفت مشترک باشند را یک جامعه آماری می‌گویند.  
**نمونه آماری:** مجموعه‌ای از افراد هستند که از جامعه آماری انتخاب می‌شوند. نمونه آماری باید معرف جامعه آماری باشد، یعنی باید ویژگی‌ها و خصوصیات جامعه آماری را در دل خود داشته باشد.  
**آماره:** آماره‌ها اندازه‌گیری‌هایی هستند که ویژگی‌های نمونه آماری را بیان می‌کنند و از آنها برای توصیف ویژگی‌های نمونه آماری استفاده می‌شود. معمولاً آماره‌ها با حروف لاتین نمایش داده می‌شوند.  
**پارامتر:** پارامترها اندازه‌گیری‌هایی‌اند که ویژگی‌های یک جامعه آماری را توصیف می‌کنند و معمولاً با حروف یونانی به نمایش در می‌آیند. پارامترها ثابت هستند اما آماره‌ها از نمونه‌ای به نمونه‌ی دیگر متفاوت می‌باشند.  
**نمونه‌گیری اریب و ناریب:** یک نمونه‌ی تصادفی یا ناریب (Unbiased Sample) نمونه‌ای است که تمامی افراد جامعه، شانس مساوی انتخاب شدن در آن را داشته باشند. در مقابل، نمونه اریب (Biased Sample) نمونه‌ای است که بعضی از افراد جامعه، شانس بیشتری برای انتخاب شدن پیدا می‌کنند.  
**قضیه‌ی حد مرکزی:** اگر تعداد دفعات نمونه‌گیری از یک جامعه به سمت بی نهایت میل کند، توزیع میانگین‌های نمونه به سمت نرمال میل خواهد نمود. اولین قدم در پژوهش، تعریف جامعه و سپس انتخاب نمونه است.  
**نمونه:** نمونه عبارت است از زیر مجموعه‌ای که از کل جامعه انتخاب می‌شود و معرف آن است.

### متغیر و انواع آن



اصطلاح **متغیر** به خصوصیتی اشاره دارد که به وسیله‌ی آن اعضای یک گروه یا مجموعه از هم متمایز می‌شوند. **متغیر کمی** است که می‌تواند از یک فرد به یک فرد دیگر و یا از یک مشاهده به یک مشاهده دیگر، مقادیر مختلفی را اختیار کند و به بیان دقیق‌تر، نمادی است که اعداد یا ارزش‌ها بدان انتساب داده می‌شود. به مجموعه‌ای از مقادیر یک **متغیر**، توزیع (Distribution) آن **متغیر** گفته می‌شود.

### انواع متغیر از حیث قابلیت عددپذیری



**متغیر کیفی:** متغیری است که اندازه‌های آن به صورت غیرعددی بیان می‌شود، نظیر جنسیت، رضایت شغلی، میزان مشارکت سیاسی و سطح سواد. متغیرهای کیفی متغیرهایی هستند که از یک رشته طبقات تشکیل می‌شوند که باید کامل و مانع‌الجمع باشند. یعنی مثلاً متغیر جنسیت که هم شامل زن و هم شامل مرد شود و یا متغیر مذهب که شامل کلیه مذاهب باشد. مثلاً اگر ما دین (مذهب) را به صورت مسلمان یا مسیحی بیاوریم در اینجا کامل نیست چون ممکن است فردی زرتشتی یا یهودی باشد. منظور از مانع‌الجمع بودن هم یعنی اینکه مثلاً اگر بگوییم مسیحی، مسلمان، شیعه، زرتشتی، اشتباه است چون دو طبقه آن یعنی شیعه و مسلمان با هم مانع‌الجمع نمی‌باشند و شیعه‌ها، مسلمان هستند. این متغیرها قابل تبدیل شدن به عدد نیستند و اگر چه به هر طبقه‌ای، شماره‌ای داده می‌شود ولی این شماره جنبه ریاضی ندارد و به معنای کدگذاری طبقات می‌باشد. مثلاً هم‌چنان که گفتیم، جنسیت یک متغیر کیفی است که به زن و مرد طبقه‌بندی می‌شود که ما می‌توانیم به زن کد ۱ و به مرد کد ۲ را بدهیم. در اینجا عددهای ۱ و ۲ جنبه‌ی ریاضی ندارند و صرفاً جهت کدگذاری به کار رفته‌اند. پس می‌توان گفت متغیرهای کیفی متغیرهایی هستند که پژوهش‌گر نمی‌تواند آن را اندازه بگیرد و با اعداد و ارقام ویژگی‌های آن را نشان دهد.  
**متغیر کمی:** متغیری است که اندازه‌های آن به صورت عددی بیان می‌شود: نظیر سن و بعد خانوار. این نوع متغیرها خود به دو دسته متغیرهای کمی پیوسته و متغیرهای کمی گسسته تقسیم می‌شوند.

**متغیر کمی پیوسته:** متغیری است که هر مقدار ممکن بین حدود معینی را به خود اختصاص می‌دهد، نظیر سن، وزن، درآمد. یعنی بین هر دو مقدار آن بتوان مقادیر جدیدی را در نظر گرفت. مثلاً بین ۲ کیلو و ۳ کیلو می‌توان مقادیر زیاد دیگری مثل ۲/۵ و ۲/۳۷ و... را در نظر گرفت پس وزن یک متغیر کمی پیوسته است. متغیر پیوسته اعشار می‌پذیرد. مثال دیگر، قد به عنوان یک متغیر کمی پیوسته است چون مثلاً بین ۱ متر و ۲ متر می‌توان اندازه‌های بی‌شمار دیگری مثل ۱/۵۰ و ۱/۶۲، ۱/۸۳ و غیره را بیان کرد.

**متغیر کمی گسسته:** متغیری است که فقط مقادیر معینی را می‌تواند به خود اختصاص دهد، یعنی نمی‌توان بین مقادیر آن عدد دیگری را آورد. نظیر تعداد رای دهندگان، بعد خانوار، دفعات ارتکاب جرم. مثلاً بعد خانوار به عنوان یک متغیر گسسته مطرح است. بعد یک خانوار نمی‌تواند به صورت ۴/۲۵ یا ۴/۵ مطرح شود یا مثلاً نمی‌توانیم بگوییم تعداد فرزندان یک خانواده ۲/۳ یا ۵/۴ و... است. پس بعد خانوار یک متغیر کمی گسسته است. متغیر گسسته اعشار نمی‌پذیرد.

## انواع متغیر از حیث تأثیر و تأثر

متغیرها از حیث این که تأثیرگذار باشند یا تأثیر پذیر، در سه وضعیت قرار می‌گیرند: متغیر مستقل، متغیر وابسته و متغیر همبسته.

**متغیر مستقل:** متغیری است که تأثیرگذار و پیش‌بینی‌کننده است.

**متغیر وابسته:** متغیری است که تأثیرپذیر و پیش‌بینی‌شونده است. در رابطه بین دو متغیر، متغیری را که تأثیر گذار است، مستقل و متغیری را که تأثیرپذیر است وابسته می‌نامند.

**متغیرهای همبسته:** دو متغیر وقتی همبسته هستند که رابطه تأثیر و تأثر با یکدیگر داشته باشند و نتوان بین آنها علت و معلول را تشخیص داد.

**متغیر میانجی:** اگر متغیر X از طریق Z بر Y تأثیر گذارد، آنگاه متغیر Z را متغیر میانجی یا مداخله‌گر می‌نامیم. متغیر میانجی بیانگر مکانیسم تأثیر X بر Y است. این متغیر را گاه متغیر واسطه می‌نامند.

**متغیر کاذب:** متغیر A برای Y کاذب است وقتی که هر دوی آنها تحت تأثیر متغیر X باشند. به عنوان مثال اندازه کف دست برای استعدادهای کلامی کودکان یک متغیر کاذب است. چون هر دو تحت تأثیر متغیر سن قرار دارند. باید توجه داشت که بین متغیر کاذب و متغیر وابسته نمی‌توان ربط علی معقولی پیدا نمود.

**متغیر منکوب یا سرکوبگر (متغیر بازدارنده - پنهان کننده):** هرگاه عدم کنترل یک متغیر (مثلاً Z) موجب شود رابطه‌ی بین دو متغیر (مثلاً X و Y) ضعیف شود، آن متغیر (Z) را سرکوبگر می‌نامیم.

**متغیر کنترل:** هر متغیری که تأثیرش بر روابط بین متغیرهای پژوهش از طریق روش‌های آماری یا نمونه‌گیری کنترل شده باشد، متغیر کنترل نامیده می‌شود. به عنوان مثال اگر عملکرد تحصیلی دانش‌آموزان دختر و پسر را یک‌بار در مدارس دولتی و یک‌بار در مدارس غیرانتفاعی با یکدیگر مقایسه کنیم، تأثیر نوع مدرسه را بر عملکرد تحصیلی، کنترل نموده‌ایم.

امروزه با روش‌های پیشرفته آماری نظیر ضرایب همبستگی تک‌یک‌ی و نیمه تک‌یک‌ی و نیز تحلیل کوواریانس می‌توان تأثیر متغیرهای ناخواسته را کنترل نمود. لازم به ذکر است که آنچه که به محقق کمک می‌کند تا نوع یک متغیر (مثلاً مستقل یا وابسته بودن یا کنترل یا کاذب بودن و...) را تشخیص دهد آمار نیست، بلکه نظریه‌ها و روش تحقیق است.

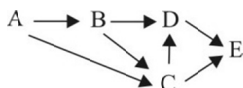
## تحلیل مسیر

برای آزمون‌های مدل‌های علی به کار می‌رود و مستلزم تنظیم مدل است که به صورت نمودار نشان داده می‌شود یعنی از طریق نمودار مشخص می‌کنیم کدام نوع متغیر بر متغیر دیگر تأثیر دارد.

در تحلیل مسیر با دو نوع متغیر سروکار داریم یعنی با توجه به مدل، دو نوع متغیر داریم: الف) متغیرهای درونی، ب) متغیرهای بیرونی.

**متغیرهای درونی:** متغیرهایی هستند که تغییرات آن‌ها از طریق متغیرهای درونی و بیرونی تبیین می‌شود و به‌عنوان متغیرهای رابطی هستند بین متغیر مستقل و وابسته.

**متغیرهای بیرونی:** متغیرهایی هستند که از طریق متغیرهای درونی بر متغیر وابسته تأثیر می‌گذارند. مثلاً در مدل روبه‌رو متغیرهای B، D و C و متغیرهای درونی هستند و متغیر A متغیر بیرونی می‌باشد.



متغیر E که همان متغیر وابسته می‌باشد از نظر زمانی آخرین متغیر می‌باشد.

## مقیاس اندازه‌گیری



**اندازه‌گیری:** اندازه‌گیری عبارت است از نسبت عددی دادن به یک صفت یا رویداد براساس یک قانون معین. متغیرها از حیث مقیاس اندازه‌گیری در يك سطح نیستند و برای سنجش هر متغیری لازم است که مقیاس متناسب با آن را انتخاب نمود. مقیاس‌ها از مهمترین ابزار پژوهش به شمار می‌آیند پس در هر سنجش ناچار باید:

- ۱- ویژگی مورد آزمون را به کمیّت تبدیل کرد؛
  - ۲- درجات آن را با اعداد یا علائم خاصی مشخص ساخت؛
  - ۳- در رده‌بندی خاصی با نام مقیاس اندازه‌گیری جای داد.
- برای سنجش و اندازه‌گیری، چهار مقیاس را می‌توان نام برد:

### ۱- مقیاس اسمی یا طبقه‌ای



ابتدایی‌ترین مقیاس، اسمی است. در این مقیاس به تعیین طبقه‌هایی پرداخته می‌شود که افراد، اشیاء یا رویدادها را می‌توان در آنها جایگزین کرد. البته باید این طبقه‌ها ناسازگار باشند.

متغیر اسمی، متغیری است که فقط از جنبه کیفی مورد نظر است. این مقیاس برای نشان دادن وجود یا عدم وجود يك صفت در نمونه می‌باشد. در این مقیاس اعدادی که به مقولات يك متغیر اختصاص داده می‌شود ارزش حقیقی ندارند، یعنی بیانگر بزرگی یا کوچکی مقولات نیستند و می‌توان آنها را جابه‌جا نمود.

مقیاس اسمی مطلبی درباره کم بودن یا زیاد بودن بیان نمی‌کند و دارای این ویژگی مهم است که فقط دارای طبقه‌های مشخص و متمایزی است که جنبه کیفی دارند و تنها رابطه موجود بین آنها تفاوت آنها با یکدیگر است. بسیاری از این مقیاس‌ها، دو ارزشی هستند. برای مثال جنسیت که دارای دو ارزش مذکر و مؤنث است دارای مقیاس اسمی است. در مقیاس اسمی نمی‌توانیم بگوییم یک طبقه بزرگتر از طبقه دیگری است و یا کوچک‌تر از دیگری است بلکه فقط می‌توانیم بگوییم که دو طبقه از یکدیگر متفاوت هستند.

متغیر اسمی یک متغیر ذاتی است یعنی خصیصه‌ای که روز به روز عوض نمی‌شود. اعدادی که در این سطح به کار برده می‌شوند، معنای عددی و کمی ندارند. مناسب‌ترین آماره در این سطح، کاربرد موارد نما در این مقیاس است. نمونه انواع متغیر اسمی عبارتند از: جنسیت، ملیت، شغل، شماره پیراهن بازیکنان، شماره پلاک خانه، شماره تلفن، شماره کلاس (در این چهار مورد اخیر عدد و شماره معنای عددی ندارد) و . . .

### ۲- مقیاس رتبه‌ای یا ترتیبی



متغیرهای رتبه‌ای متغیرهایی‌اند که ذاتاً رتبه‌بندی شده‌اند. در این سطح محقق فقط به دسته بندی داده‌ها نمی‌پردازد بلکه آزاد خواهد بود تا آنها را با توجه به درجه‌ای اهمیت مرتب نماید.

در مقیاس ترتیبی نیز همانند مقیاس اسمی، هدف، طبقه‌بندی و تشخیص طبقات از یکدیگر است. با این تفاوت که ترتیب طبقه‌ها هم مورد توجه قرار می‌گیرد. در این مقیاس از اعداد به منظور درجه‌بندی و رتبه‌بندی استفاده می‌شود. در این سطح اعدادی که به مقولات تعلق می‌گیرند باید از نظم و ترتیب خاصی برخوردار باشند، به عنوان مثال عددی که به مقوله متوسط تعلق می‌گیرد حتماً باید بین اعدادی باشد که به مقوله‌های کم و زیاد تعلق گرفته است.

مقیاس رتبه‌ای به منظور مرتب کردن افراد یا اشیاء از بیشترین میزان مورد اندازه‌گیری تا کمترین میزان آن به کار برده می‌شود. در اندازه‌گیری ترتیبی می‌توانیم تعیین کنیم که يك مشاهده (Case) بزرگ‌تر از، مساوی با، یا کوچک‌تر از بقیه مشاهدات است ولی نمی‌توانیم فاصله مشاهدات را مشخص نماییم و نیز ابتدا و انتهای آنها مشخص نیست. نظیر رضایت شغلی (خیلی راضی، راضی، متوسط، ناراضی، خیلی ناراضی).

در این مقیاس به داده‌ها بر اساس کوچک بودن یا بزرگ بودن آنها رتبه داده شده و داده‌ها با یکدیگر مقایسه می‌گردد. مثلاً مشخص می‌کنیم که یکی از دیگری بزرگ‌تر و کوچک‌تر و . . . است. به عنوان مثال تحصیلات، در مقیاس رتبه‌ای است.

در این مقیاس فاصله بین دو طبقه را نمی‌توان یکسان فرض کرد. در واقع اعداد اختصاص یافته به مقولات فاقد ارزش عددی هستند.

مثال‌های زیر در زمره متغیر ترتیبی به حساب می‌آیند: (۱۵ تا ۲۰، ۱۰ تا ۱۵، ۵ تا ۱۰، زیر ۵ سال)، (بالای ۲۰ سال، ۱۵ تا ۲۰ سال، ۱۰ تا ۱۵، ۵ تا ۱۰، زیر ۵ سال)، (بالای ۲۰ سال، ۱۵ تا ۲۰ سال، ۱۰ تا ۱۵، ۵ تا ۱۰).

استفاده از چهار عمل ریاضی ( $\times$ ،  $\div$ ،  $+$ ،  $-$ ) در مقیاس‌های اسمی و ترتیبی امکان‌پذیر نیست.

### ۳- مقیاس فاصله‌ای

یکی از مشکلات مقیاس ترتیبی این است که فاصله بین مقادیر در آن نسبی است. این مشکل در متغیر فاصله‌ای وجود ندارد، زیرا در آن فاصله بین طبقات یا مقادیر یکسان است. مقیاس فاصله‌ای، مقیاسی مدرج و با درجات مساوی است. هنگامی که نتایج یک تحقیق را بر روی چنین مقیاسی وارد می‌کنیم، هم امکان رده‌بندی آنها در دو جهت (بالا و پایین) وجود دارد و هم مقایسه داده‌ها امکانپذیر است (چون درجات مساوی هستند). این مقیاس دارای صفر اختیاری یا قراردادی است. قراردادی بودن بدین معنی که ما خود قرارداد کرده‌ایم و برای آن یک صفر تعیین کرده‌ایم و اینکه صفر آن طبیعی نیست. مانند نقطه صفر دماسنج که ما خود نقطه‌ای را برای صفر قرارداد کرده‌ایم و صفر را برای آن تعیین نموده‌ایم. به عبارت دیگر، در مقیاس فاصله‌ای نقطه «صفر واقعی» وجود ندارد. یعنی اگر کسی در آزمون هوش نمره صفر بگیرد به این معنی نیست که او اصلاً هوش ندارد. در این مقیاس می‌توانیم اعداد را از هم کم کنیم و تفاوت بین آنها را به دست آوریم. داده‌ها همه دارای فاصله‌های مساوی هستند. با وجودی که صفر واقعی وجود ندارد، گاهی به علت سهولت و سادگی صفر قراردادی منظور می‌کنند. اندازه‌گیری درجه حرارت بر حسب فارنهایت مثالی از یک مقیاس فاصله‌ای است که نقطه صفر مطلق ندارد، بلکه دارای صفر قراردادی است و به همین دلیل عملیات ضرب و تقسیم در مورد آن انجام نمی‌گیرد و تنها می‌توان عملیات جمع و تفریق را بکار برد. سن، بهره‌هوشی، نمره کنکور، توزیع نمرات Z، هوش، عزت نفس و افسردگی در مقیاس فاصله‌ای سنجیده می‌شوند.

### ۴- مقیاس نسبی (کسری)

در این سطح فاصله‌ها در تمام طول مقیاس پیوستگی دارد و همچنین مبدأ سنجش در آن صفر واقعی یا صفر مطلق است. به عبارتی دقیق‌تر نمره صفر در این سطح به این معناست که فرد از آن متغیر هیچی ندارد. در این مقیاس مبدأ سنجش همیشه صفر مطلق است. در واقع مقیاس نسبی از تمام ویژگی‌های انواع مختلف مقیاس‌های پیشین برخوردار است؛ دارای بیش از یک طبقه، دارای ترتیب، مبین طول دقیق فواصل و دارای نقطه صفر حقیقی است. به عنوان مثال اگر میزان درآمد صفر باشد به این معناست که فرد اصلاً درآمد ندارد. در این مقیاس، تمام چهار عمل اصلی ریاضی را می‌توان اعمال کرد. مقیاس نسبی قابل تبدیل به دیگر مقیاس‌هاست. وزن، طول قد، جمعیت یک روستا، و میزان درآمد نمونه‌هایی از متغیرهای واقع در مقیاس نسبی هستند.

**نکته ۳:** وجه مشترک سطوح اندازه‌گیری چهارگانه اسمی، رتبه‌ای، فاصله‌ای و نسبی، در طبقه‌بندی آنها است.

- ۲) مقیاس اسمی و ترتیبی (رتبه‌ای) در سطح کیفی‌اند و مقیاس فاصله‌ای و نسبی در سطح کمی‌اند.
- ۳) از مقیاس نسبی به طرف مقیاس اسمی مقیاس‌ها قابل تبدیل به یکدیگرند.
- ۴) مقیاس اسمی قابل تبدیل به هیچ مقیاسی نیست.
- ۵) گرایش‌های مرکزی این مقیاس‌ها به ترتیب عبارتند از:
  - الف: مقیاس اسمی: مد یا نما (Mode)
  - ب: مقیاس ترتیبی: میانه (Median)
  - ج: مقیاس فاصله‌ای: میانگین (Mean)
  - د: مقیاس نسبی: میانگین (Mean)

### کارکردهای آمار توصیفی

آمار توصیفی کارکردهای مختلفی دارد از جمله:

- ۱- سازمان‌دهی داده‌ها
- ۲- نمایش داده‌ها
- ۳- نشان دادن مرکزیت داده‌ها (مانند میانگین، میانه، نما و ...)
- ۴- نشان دادن پراکندگی داده‌ها (مانند واریانس، انحراف معیار و ...)
- ۵- نشان دادن شاخص‌های تقارن و نرمال بودن داده‌ها
- ۶- توصیف روابط بین داده‌ها.



سؤالات چهارگزینه‌ای کنکور سراسری فصل اول

- ۱- توصیف خلاصه‌ای از یک متغیر در یک نمونه چه نامیده می‌شود؟  
 (۱) آماره (۲) پارامتر (۳) فاصله اطمینان (۴) متغیر  
 (سال ۸۵)
- ۲- سطوح سنجش متغیرهای قومیت، تعداد برادر و مدرک تحصیلی به ترتیب کدامند؟  
 (۱) رتبه‌ای - نسبی - اسمی (۲) اسمی - نسبی - رتبه‌ای  
 (۳) اسمی - فاصله‌ای - نسبی (۴) رتبه‌ای - فاصله‌ای - رتبه‌ای  
 (سال ۸۸)
- ۳- متغیر نوع شغل در چه سطحی سنجیده می‌شود؟  
 (۱) اسمی (۲) نسبی (۳) فاصله‌ای (۴) رتبه‌ای  
 (سال ۸۸)



سؤالات چهارگزینه‌ای تألیفی فصل اول

- ۴- کدام مقیاس اندازه‌گیری صرفاً به تعیین ط بقاتی می‌پردازد که افراد، اشیاء یا حوادث را می‌توان در آنها قرار داد؟  
 (۱) اسمی (۲) ترتیبی (۳) فاصله‌ای (۴) نسبی
- ۵- در کدام مقیاس اندازه‌گیری، تمام عملیات آماری و ریاضی را می‌توان انجام داد؟  
 (۱) ترتیبی (۲) فاصله‌ای (۳) نسبی (۴) اسمی
- ۶- مقیاس اندازه‌گیری دو متغیر «وزن» و «بهره هوشی» در فرضیه «وزن با میزان بهره هوشی کودکان پایه اول دبستان رابطه دارد» به ترتیب کدام است؟  
 (۱) فاصله‌ای - فاصله‌ای (۲) نسبی - نسبی (۳) نسبی - فاصله‌ای (۴) فاصله‌ای - نسبی
- ۷- موارد استفاده میانه بیشتر در کدام یک از مقیاس‌های اندازه‌گیری است؟  
 (۱) اسمی (۲) رتبه‌ای (۳) فاصله‌ای (۴) نسبی
- ۸- کدام یک از متغیرهای زیر گسسته است؟  
 (۱) طول عمر مفید یک رادیو (۲) نمرات آزمون آمار دانش‌آموزان  
 (۳) فواصل بین واحدهای آموزشی یک شهر (۴) میزان زاد و ولد در شهر تهران
- ۹- اگر در یک مطالعه جنسیت یکی از متغیرها باشد، این متغیر چه نوع متغیری است؟  
 (۱) وابسته (۲) پیوسته (۳) کمی (۴) ناپیوسته
- ۱۰- کدام یک از گزینه‌های زیر متغیر گسسته است؟  
 (۱) تعداد صندلی‌ها (۲) قد مردم (۳) نمره پیشرفت تحصیلی (۴) وزن دانشجویان
- ۱۱- مناسب‌ترین مورد استفاده از نما زمانی است که مقیاس اندازه‌گیری به کار رفته مقیاس ..... باشد.  
 (۱) اسمی (۲) ترتیبی (۳) فاصله‌ای (۴) نسبی
- ۱۲- با کدام روش آماری می‌توان افزایش تولید ناخالص ملی ایران را در سه سال گذشته نمایش داد؟  
 (۱) استنباطی غیر مشروط (۲) استنباطی مشروط (۳) توصیفی (۴) توصیفی استنباطی
- ۱۳- هدف از کاربرد آمار استنباطی چیست؟  
 (۱) تعمیم یافته‌ها به جامعه‌ی مرجع (۲) مشخص نمودن میزان خطای شاخص‌ها  
 (۳) تفسیر آماره‌های محاسبه شده (۴) محاسبه شاخص‌ها یا آماره‌ها

## فصل دوم

# آماده‌سازی، سازمان‌دهی و نمایش داده‌ها (نمودارها)

### آماده‌سازی داده‌ها

داده‌ها به هر روشی که گردآوری شده باشند داده‌های خام‌اند. اولین مرحله در تحلیل داده‌ها آماده کردن داده‌ها برای انواع تحلیل‌های مورد نظر است. اگر تحقیق، مطالعه میدانی است نوع اصلی تحلیل شامل تنظیم یادداشت‌ها برای بررسی مسأله اصلی تحقیق و سپس ارائه آنها به صورت مقاله است. ای بسا در این نوع تحلیل داده‌های کمیت‌پذیر وجود نداشته باشد.

در اکثر روش‌های دیگر برای تحلیل‌های کمی و (احتمالاً) آزمون‌های آماری باید اطلاعات گردآوری شده اولیه (داده‌های خام) را به معادل عددی تبدیل کرد. این قسمت به نحوه تبدیل داده‌های خام بالقوه کمیت‌پذیر به اعدادی اختصاص دارد که نشان‌دهنده دامنه معنای - تغییر - داده‌های خام است. این کار برای تهیه جداول و نمودارها و شکل‌هایی که داده‌ها را به‌طور دقیقی توصیف می‌کنند و برای تبیین نحوه تأیید فرضیه‌ها براساس داده‌ها ضروری است. اگر در اندیشه آزمون آماری فرضیه‌های خود هستید باید از سنجش‌هایی برای تعیین میزان روابط استفاده کنید. چنانچه داده‌های تحقیق از نمونه احتمالی به‌دست آمده باشد می‌توان با آزمون‌های آماری دیگری به استنباط‌هایی در مورد جمعیتی که نمونه از آن برگرفته، نایل آمد.

آماده کردن داده‌ها در چهار مرحله صورت می‌گیرد. مرحله اول که داده‌ها به صورت عدد **کدگذاری** می‌شوند، مرحله **کدگذاری** است. مرحله دوم که اطلاعات کدگذاری شده برای تبدیل از شکل مکتوب به شکل کامپیوتری آماده می‌شوند، مرحله **تبدیل** است. مرحله سوم که شامل فرایند عملی ورود داده‌های کدگذاری شده به کامپیوتر می‌شود، مرحله **انتقال به کامپیوتر** است.

آخرین محله، مرحله **تصحیح داده‌های وارد شده به کامپیوتر** است. در این مرحله باید صحت داده‌هایی را که وارد کامپیوتر شده‌اند واریسی کرد، چه بعید نیست در هر مرحله از پردازش داده‌ها اشتباهاتی پیش آید.

### تشکیل نظام کدگذاری

#### اصول کلی

کدگذاری داده‌ها متضمن کمی کردن داده‌های خام برای تحلیل است. برای تشکیل نظام کدگذاری باید از اصولی تبعیت کرد. اولین اصل این است که **کدگذاری باید دقیقاً تعریف شده، خالی از ابهام باشد**. این امر مستلزم اتخاذ بهترین شیوه کدگذاری برای پاسخ پرسش‌ها (متغیرها) است به گونه‌ای که معنای مفهومی که گویای متغیر تحقیق است یکسان باقی بماند. کدگذاری را می‌توان براساس معیارهای نظری و تجربی انجام داد.

معیارهای نظری برای آزمون این است که مقیاس (یا رشته‌ای از مقیاس‌های) صفاتی که انواع تغییرات متغیری را تشکیل می‌دهند به صورت رشته‌ای از کدهای عددی درآمده باشد که دقیقاً مبین دامنه مقادیر متغیر مورد نظرند. چنین معیاری آزمون اعتبار متغیر است. معیارهای تجربی آزمون این است که نتایج سنجش متغیر تا چه حد گویای نتایج منطقی و قابل انتظار است. به عبارت دیگر این معیارها به تشخیص این که پاسخگویان تا چه حد

پرسش‌ها را فهمیده و به معنای پرسش شما پی برده‌اند کمک می‌کند. مسأله اصلی در اینجا این است که پاسخگویان به‌طور استوار و ثابتی به سوالات پاسخ داده باشند؛ یعنی بتوان پیش‌بینی کرد که چنانچه بار دیگر همان سوالات از آنها پرسیده شود همان پاسخ‌ها را می‌دهند. این آزمون همانا **آزمون پایایی** است. تفاوت معیارهای نظری و تجربی در **گذرداری** در این است که معیارهای نظری را می‌توان قبل از پردازش داده‌ها به کار برد اما معیارهای تجربی را فقط بعد از گردآوری داده‌ها و الگوی خاص شواهد موجود در پاسخ‌ها که راهنمای نحوه **گذرداری** است می‌توان تعیین کرد. از جنبه نظری کدها را باید به گونه‌ای طرح کرد که گویای دامنه معنای در نظر گرفته شده برای متغیرها باشد. این دامنه معنا باید تمام تغییرات ممکن را دربر گیرد.

از جنبه تجربی، **گذرداری** داده‌ها باید دربرگیرنده تمام پاسخ‌های مختلفی باشد که پاسخگویان ارائه داده‌اند و نیز برای هنگامی است که برآیند انواع خاصی از پاسخ‌ها را ترکیب کنید. پس دومین اصل کلی **گذرداری** این است که **حتی‌الامکان معنای واقعی پاسخ‌ها و کل دامنه تغییرات داده‌ها را گذرداری کرد.**

مثال بارز **گذرداری** تجربی، **گذرداری** سوالات باز است. گو این‌که در پرسش‌های باز می‌توان پاسخ‌های احتمالی را پیشاپیش در نظر گرفت اما **گذرداری** نهایی منوط به گردآوری داده‌هاست. این نوع **گذرداری**، **گذرداری** بعد از داده‌هاست و بر نتایج تجربی به‌دست‌آمده استوار است. سومین قاعده کلی این است که **نظام گذرداری را حتی‌الامکان در مرحله طرح ابزار گردآوری داده‌ها طرح‌ریزی کرد.** اگر ابزار گردآوری داده‌ها پرسش‌نامه باشد و بتوان آن را **گذرداری** کرد معیارهای نظری **گذرداری** را باید دقیقاً مراعات کرد. چهارمین قاعده این است که **نظام گذرداری باید به تقلیل دفعات پردازش داده‌ها بکشد**، چه در هر بار پردازش داده‌ها اشتباهاتی رخ خواهد داد. به همین دلیل نظام **گذرداری** باید روشن و دقیق باشد.

پنجمین قاعده **گذرداری** این است که **برای تثبیت نظام گذرداری و تمام تصمیماتی که در مورد گذرداری اتخاذ شده است باید کدنامه‌ای تهیه کرد.** این کدنامه نه تنها برای راهنمای **گذرداری** لازم است، بلکه در جریان تحلیل هم برای یادآوری چیزی که متغیر می‌سجد لازم است.

#### انواع استراتژی‌های گذرداری

پرسش‌ها معمولاً پنج گونه‌اند. گونه اول متضمن رشته‌ای از پاسخ‌های ترتیبی است مانند موافقت کامل، موافقت، مردد، مخالفت، مخالفت کامل (پرسش‌های نگرشی غالباً این‌گونه‌اند). گونه دوم دارای پاسخ‌های دوشقی است، مانند بله یا خیر و نیز پاسخ‌هایی که دال بر داشتن یا نداشتن صفتی است. پرسش «تنبیس بازی می‌کنید؟ بله / خیر» با پرسش «به کدامیک از این ورزش‌ها می‌پردازید؟ تنیس، شنا، بکس و ...» هم‌سنگ است. در پرسش اخیر هریک از ورزش‌ها را می‌توان مبنای یک متغیر دوشقی قرار داد مانند ۱ = تنیس بازی می‌کنم و ۰ = تنیس بازی نمی‌کنم. پرسش‌های گونه سوم، پرسش‌های مطلق بوده دارای رشته‌ای از پاسخ‌های بدون ترتیب‌اند (مثلاً مذهب - کاتولیک، پروتستان، یهودی - که هیچ نظم و ترتیب منطقی ندارد). گونه چهارم پرسش‌هایی است که پاسخ آنها مستقیماً با اعداد داده می‌شود (درآمد، سن، تعداد فرزندان). گونه پنجم پرسش‌های باز است که پاسخ‌های آنها براساس محتوای پاسخ‌ها یک به یک **گذرداری** شده، سپس می‌توان آنها را از نو به طبقات عددی تبدیل کرد.

#### شکل نهایی کدها

شکل نهایی کدهای مکتوب همانی است که برای انتقال داده‌ها به کامپیوتر مورد استفاده قرار می‌گیرد. باز هم باید تأکید کرد که تحدید دفعات پردازش داده‌ها تأثیر بسزایی در تقلیل خطاها دارد. به همین دلیل محققان سعی می‌کنند حتی‌الامکان داده‌ها را مستقیماً از پرسش‌نامه به کامپیوتر منتقل کنند. از آنجا که این کار در پاره‌ای موارد امکان‌پذیر نیست باید به روش‌های دیگری تمسک جست:

**گذرداری حاشیه‌ای در پرسش‌نامه.** در این روش اطلاعات **گذرداری** در حاشیه پرسش‌نامه درج می‌شود به گونه‌ای که پردازش‌کننده داده‌ها فقط باید پاسخ را به محلی که مختص ستون کامپیوتری است منتقل کند (در پاره‌ای پرسش‌نامه‌ها محل پاسخ‌ها را در همان حاشیه قرار می‌دهند). کلاً **گذرداری** حاشیه‌ای فرایند **گذرداری** را ساده کرده، همسانی آن را بالا می‌برد.

**استفاده از برگه گذرداری.** هرگاه قرار باشد داده‌ها روی کارت ای بی ام (IBM) که دارای ۸۰ ستون برای ذخیره اطلاعات است پانچ شود بهتر است داده‌ها را از منبع اصلی (پرسش‌نامه) به برگه‌های مخصوص ۸۰ ستونه‌ای که عین کارت‌های ای بی ام است منتقل کرد. بدین ترتیب پانچ داده‌ها ولو آن‌که پانچ‌کن با تحقیق آشنا نباشد بسیار ساده خواهد شد. با این وصف فراموش نکنید که همین انتقال داده‌ها ممکن است باعث خطاهایی بشود که قبلاً وجود نداشته‌اند. به همین دلیل محققان ترجیح می‌دهند داده‌ها مستقیماً از پرسش‌نامه وارد کامپیوتر گردد.



## دستکاری داده‌ها پیش از تحلیل

### دستکاری داده‌های ناقص

در همه برنامه‌های نرم‌افزاری، روش‌هایی برای دستکاری و جابه‌جایی داده‌های ناقص وجود دارد. باید به کامپیوتر اعلام کرد که در چه مواردی داده‌های یک متغیر را ناقص به‌شمار آورد (یا با علامت مشخص کند) تا بتوان آنها را از تحلیل حذف کرد. البته اگر موردی (مثلاً پرسش‌نامه‌ای) معدودی متغیر ناقص داشته باشد به‌طور کامل از تحقیق حذف نمی‌شود، بل از کامپیوتر خواسته می‌شود متغیرهایی (پرسش‌هایی) که جواب درستی دارند باقی مانده، متغیرهایی که بی‌پاسخ مانده‌اند به‌عنوان متغیر ناقص مشخص شوند. در زبان برنامه نرم‌افزار به این تفکیک، **حذف جزئی داده‌های ناقص** اطلاق می‌شود. اما اگر موردی وجود داشته باشد که داده‌های همه متغیرهایش ناقص باشد به‌طور کامل از تحلیل حذف می‌شود که بدان **حذف کامل داده‌های ناقص** اطلاق می‌شود.

معمولاً با کد ۹ یا ۹۹ یا ۹۹۹ مقادیر ناقص را مشخص می‌کنند. بعد از تعیین مقادیر ناقص آنها را حذف می‌کنند، هر چند تعداد موردی ناقص هر متغیر در جداول ثبت می‌شود.

### گذگذاری مجدد متغیرها

یکی از رایج‌ترین دستکاری‌ها در تعیین متغیرهای تحلیل، **گذگذاری مجدد** آنهاست. عموماً هدف گذگذاری مجدد تقلیل تعداد طبقات متغیر به تعداد معقول‌تری جهت تحلیل عددی است. وقتی برآیند رابطه دو یا چند متغیر را بررسی کنید احتمالاً باید تعداد طبقات متغیرها را تقلیل دهید تا معلوم شود اثر متغیرها بر یکدیگر قوی است یا ضعیف. اما با گذگذاری مجدد، طبقات اولیه متغیر دگرگون خواهد شد. برای اجتناب از این اختلاط معمولاً بهتر است از متغیرهای اولیه متغیرهای جدیدی درست کرد تا هم طبقات اولیه متغیر حفظ شود و هم از متغیری با یک رشته طبقات جدید (مجدداً گذگذاری شده) برخوردار شد.

برای گذگذاری مجدد متغیر دو قاعده کلی وجود دارد. یکی **ملاحظات نظری** است؛ یعنی متغیر مجدداً گذگذاری شده به نحو بهتری معرف مفهوم متغیر تحقیق است (مثلاً در تحقیقی بر آن می‌شویم تمام غیرسفیدپوستان را در یک طبقه جای دهیم چه فقط در پی تمایز این دو هستیم). قاعده دیگر **ملاحظات تجربی** است؛ یعنی برحسب پاسخ‌های ارائه شده متغیر را مجدد طبقه‌بندی می‌کنیم (مثلاً بعد از بررسی پاسخ‌ها می‌بینیم که تعداد هر طبقه اقلیت کمتر از آنی است که بتوان به تنهایی آن را تحلیل کرد در نتیجه همه اقلیت‌ها را در یک طبقه غیرسفیدپوست ادغام می‌کنیم).

### سازماندهی داده‌ها

با توجه به اینکه اطلاعات به‌دست آمده از یک تحقیق غالباً زیاد و به صورت نامنظم هستند، هر نوع نتیجه‌گیری و تعبیر در مورد آنها مشکل است. بنابراین برای اینکه بتوانیم داده‌های به دست آمده را تحلیل کنیم باید آنها را به صورت یک نظم منطقی درآوریم تا به صورت معنی‌دار قابل تعبیر باشند. در این راستا اولین گام، تشکیل جدول فراوانی می‌باشد که در آن مراحل طبقه‌بندی و محاسبه داده‌ها نشان داده می‌شود. پس توزیع فراوانی عبارت است از سازمان دادن به داده‌ها یا مشاهدات به صورت طبقات، همراه با فراوانی هر طبقه. هدف کلی جداول توزیع فراوانی عبارت است از: درک بهتر داده‌های جمع‌آوری شده، کمی ساختن واقعیت مورد مطالعه، ارائه تصویری نسبتاً دقیق از واقعیت مورد مطالعه، انعکاس واقعیت تجربی.

سازمان‌دهی داده‌ها معمولاً بر اساس جداول توزیع فراوانی صورت می‌گیرد. معمولاً داده‌های مربوط به یک متغیر به علت طبیعت ظاهری نامنظم خود گویای مطلبی درباره‌ی جامعه نیستند. برای آنکه بتوان به آنها نظم بهتری داد آنها را معمولاً در قالب جدول تنظیم می‌کنند. جداول ابزارهایی برای مهار پراکندگی داده‌ها هستند. یک جدول فراوانی معمولاً از ستون‌های زیر تشکیل می‌شود:

- ۱- **طبقات یا دسته‌ها:** اگر چه گاهی اوقات ستون نخست جدول می‌تواند از اعداد منفرد تشکیل شود اما اکثر اوقات دسته‌های اعداد تشکیل‌دهنده آن هستند. به طور قراردادی تعداد طبقات را بین ۱۰ تا ۲۰ انتخاب می‌نماییم. قاعده کلی این است که اگر تعداد نمره‌ها نسبتاً کم باشد (کمتر از ۱۰۰)، تعداد طبقات را نزدیک به ۱۰، ولی اگر توزیع شامل تعداد زیادی نمره است، برای مثال بیش از ۵۰۰، آنگاه طبقات را نزدیک به ۲۰ انتخاب می‌نماییم. در گام دوم فاصله بزرگ‌ترین و کوچک‌ترین عدد توزیع را محاسبه و سپس بر تعداد طبقات تقسیم می‌نماییم تا فاصله طبقات مشخص شود.
- ۲- **نماینده‌ی طبقات (حد میانی یا طبقه وسط طبقات):** اگر کران پایین و بالای هر طبقه را جمع و تقسیم بر ۲ نماییم نماینده‌ی طبقه به دست می‌آید.
- ۳- **فراوانی مطلق:** این ستون از شمارش تعداد مقادیر مربوط به هر طبقه حاصل می‌شود. جمع این ستون باید با تعداد کل نمونه برابر باشد.

- ۴- **فراوانی نسبی:** این ستون از تقسیم فراوانی مطلق هر طبقه بر جمع کل فراوانی‌ها به دست می‌آید، جمع این ستون باید ۱ شود.  
 ۵- **فراوانی درصدی:** این ستون از حاصل ضرب فراوانی نسبی در عدد ۱۰۰ به دست می‌آید. جمع کل این ستون باید ۱۰۰ شود.  
 ۶- **فراوانی تجمعی درصدی:** این ستون از جمع فراوانی‌های درصدی هر طبقه با فراوانی‌های درصدی طبقات ماقبل حاصل می‌شود.

طبقات	نماینده‌ی طبقات	فراوانی مطلق	فراوانی نسبی	فراوانی درصدی	فراوانی تجمعی درصدی
۱۰ تا ۲۰ هزار تومان	۱۵	۱۰	%۵	۵	۵
۲۰ تا ۳۰ هزار تومان	۲۵	۵۰	%۲۵	۲۵	۳۰
۳۰ تا ۴۰ هزار تومان	۳۵	۷۰	%۳۵	۳۵	۶۵
۴۰ تا ۵۰ هزار تومان	۴۵	۶۰	%۳۰	۳۰	۹۵
۵۰ تا ۶۰ هزار تومان	۵۵	۱۰	%۵	۵	۱۰۰
جمع	۱۷۵	۲۰۰	۱	۱۰۰	

**خطای گروه‌بندی:** اختلاف بین شاخص آماری محاسبه شده با استفاده از توزیع فراوانی طبقه‌بندی شده و توزیع فراوانی طبقه‌بندی نشده، خطای گروه‌بندی نام دارد.

زمانی که داده‌های یک جامعه آماری گسترده باشند، از جداول طبقاتی استفاده می‌کنیم که البته در این حالت طبقه‌بندی داده‌ها به لحاظ اقتصادی مقرون به صرفه است اما توزیع داده‌ها از نظر دقت، کمتر می‌شود.

داده‌های متغیر را می‌توان به دو صورت طبقه‌بندی کرد:

- ۱- **طبقه‌بندی پیوسته:** اگر زمانی داده‌ها به صورت پیوسته و نسبی (اعشارپذیر) بود، آنها را به صورت پیوسته طبقه‌بندی می‌کنیم.  
 ۲- **طبقه‌بندی گسسته:** طبقه‌بندی گسسته برای داده‌های غیراعشاری استفاده می‌شود. داده‌هایی که از طریق شمارش به دست می‌آید (هر عددی شمارش پذیر باشد گسسته است مانند ماشین، آدم).

طبقه‌بندی پیوسته	طبقه‌بندی گسسته
۰-۵	۰-۵
۵-۱۰	۶-۱۱
۱۰-۱۵	۱۲-۱۷
۱۵-۲۰	۱۸-۲۳

در حالتی که داده‌ها به صورت گسسته طبقه‌بندی می‌شوند، امکان دارد داده‌هایی باشند که در هیچ یک از طبقات جای نگیرند برای مثال در طبقه‌بندی گسسته زیر:

- ۲-۴  
 ۵-۷  
 ۸-۱۰  
 ۱۱-۱۳

اگر در داده‌های ما عدد ۴/۵ وجود داشته باشد، آن را در هیچ یک از طبقات نمی‌توان جای داد. به همین دلیل باید طبقات را به صورت پیوسته طبقه‌بندی کرد. برای اینکه طبقه‌بندی از حالت پیوسته به گسسته تبدیل شود باید حدود واقعی اعداد را به دست آوریم.

$$\text{حدود واقعی عدد } ۱۹ \leftarrow ۱۸/۵ - ۱۹/۵$$

$$\text{حدود واقعی عدد } ۱۹/۵ \leftarrow ۱۹/۴۵ - ۱۹/۵۵$$

$$\text{حدود واقعی عدد } ۱۲/۲۵ \leftarrow ۱۲/۲۴۵ - ۱۲/۲۵۵$$

در مثال بالا طبقه‌بندی پیوسته با احتساب حدود واقعی اعداد، اینگونه می‌شود:

- ۱/۵ - ۴/۵  
 ۴/۵ - ۷/۵  
 ۷/۵ - ۱۰/۵  
 ۱۰/۵ - ۱۳/۵

**مراحل طبقه‌بندی:**

- ۱- مرتب کردن داده‌ها از کوچک به بزرگ
- ۲- به دست آوردن دامنه تغییرات (R):

کمترین عدد - بیشترین عدد = دامنه تغییرات (R) برای طبقه‌بندی پیوسته
۱+ کمترین عدد - بیشترین عدد = دامنه تغییرات (R) برای طبقه‌بندی گسسته

تعیین تعداد طبقات؛ تعداد طبقات به نحوه‌ی پراکندگی داده‌ها و خواست محقق بستگی دارد و معمولاً بین ۵ تا ۱۵ طبقه توصیه می‌شود. تعیین فاصله طبقات.

$$I = \frac{R(\text{دامنه طبقات})}{K(\text{تعداد طبقات})}$$

برای تعیین فاصله طبقات باید دامنه تغییرات را تقسیم بر تعداد طبقات کرد:

تعیین حد میانی یا نماینده طبقه: معدل دو حد بالا و پایین را مرکز یا حد وسط آن طبقه می‌نامند و از طریق فرمول روبرو به دست می‌آید:

$$X = \frac{\text{حد بالا} + \text{حد پایین}}{2}$$

$$\frac{1/5 + 4/5}{2} = 3$$

مثلاً در طبقه ۴/۵-۱/۵ حد میانی طبقه عدد ۳ است:

**مثال ۱-** در زیر نمرات درس جامعه‌شناسی یک کلاس آورده شده است. توزیع فراوانی این نمرات را با فاصله طبقاتی مناسب به دست آورید.

	۱۸	۱۵	۲۰	۱۲	۱۴	۱۸/۵	۱۹	۱۳	۱۷
	۱۲/۲۵	۱۶	۱۹/۵	۱۵/۷۵	۱۱	۹	۲۰	۱۸/۲۵	۱۶/۷۵

۱+ پایین‌ترین نمره - بالاترین نمره = R (دامنه تغییرات)

$$R = 20 - 9 + 1 = 12$$

تعداد طبقات = ۴

$$\text{فاصله طبقات} = i = \frac{12}{4} = 3$$

**توزیع فراوانی تراکمی:** فراوانی تراکمی هر طبقه مساوی است با مجموع فراوانی‌های طبقه‌های پایین‌تر و طبقه مورد نظر به عنوان مثال فراوانی تراکمی طبقه (۹ - ۱۱) همان عدد ۲، فراوانی تراکمی طبقه‌ی (۱۲ - ۱۴) عدد ۶، فراوانی تراکمی طبقه‌ی (۱۵ - ۱۷) عدد ۱۱، فراوانی تراکمی طبقه‌ی (۱۸ - ۲۰) عدد ۱۸ یعنی همان N است.

جدول توزیع فراوانی نمرات درس جامعه‌شناسی گروهی از دانش‌آموزان

طبقه نمرات	فراوانی
۱۸-۲۰	۷
۱۵-۱۷	۵
۱۲-۱۴	۴
۹-۱۱	۲
N = 18	

نکته ۴: فراوانی تراکمی درصدی از این فرمول به دست می‌آید: برای این کار فراوانی تراکمی هر طبقه  $\frac{CF}{N} \times 100$  را بر N (تعداد کل اعداد) تقسیم و آنگاه ضربدر ۱۰۰ می‌کنیم.

## نمایش داده‌ها (نمودارها)

یکی دیگر از راه‌های ارائه یافته‌ها ترسیم نمودار است. نمودار نوعی تصویر دیداری از یک مجموعه نمره یا جدول توزیع فراوانی به دست می‌دهد. هدف نمودار، ساده‌سازی است. نمودار به تفهیم مطلب کمک کرده و توجه را معطوف به ماهیت اساسی داده‌ها می‌کند. نمودارها انواع مختلفی دارند که هر کدام بر حسب شرایط خاصی طراحی می‌شوند. مزیت نمودارها نسبت به جداول، ضریب بالای سرعت انتقال مفاهیم و یافته‌ها به مخاطبان و نیز ارائه تصویر روشن‌تری از داده‌های جمع‌آوری شده است.

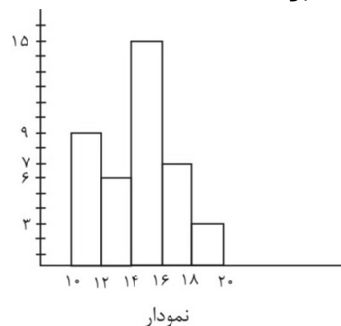
### ۱- نمودار هیستوگرام

هیستوگرام نموداری است که در آن، فراوانی‌ها به صورت سطوح عمودی نشان داده می‌شوند. این نمودار ویژه **متغیرهای کمی (فاصله‌ای و نسبی)** و همچنین **داده‌های پیوسته** است. این نمودار از یک سری ستون‌های به هم چسبیده تشکیل می‌شود، که در آن هر ستون نشان‌دهنده طبقه‌ای از اعداد است؛ عرض هر ستون برابر فاصله طبقه و ارتفاع آن مساوی فراوانی همان طبقه است. اتصال ستون‌ها موجب می‌شود که این نمودار برای نمایش داده‌های ناشی از اجرای **متغیرهای پیوسته** وسیله مناسبی باشد. برای ترسیم این نمودار از حد بالا و پایین طبقات استفاده می‌شود.

**مثال ۲- اگر نمره ریاضی ۲۰ نفره دانشجوی به شکل زیر دسته‌بندی شده باشد:**

فراوانی مطلق	طبقات نمره
۹	۱۰ تا ۱۲
۶	۱۴ تا ۱۶
۱۵	۱۶ تا ۱۸
۷	۱۸ تا ۲۰
۳	۲۰ تا ۲۲

نمودار هیستوگرام آن به شکل زیر خواهد بود:



### ۲- نمودار ستونی یا میله‌ای

این نمودار شبیه به نمودار هیستوگرام است با این تفاوت که در این نمودار ستون‌ها مجزا از یکدیگر هستند. هر میله یا ستون نشان‌دهنده یک طبقه یا یک رده از داده‌هاست. این نمودار را می‌توان برای نمایش **داده‌های ناپیوسته با مقیاس‌های اسمی** که مقوله‌های جدایی از یکدیگر هستند، به آسانی مورد استفاده قرار داد. این نمودار ویژه **متغیرهای اسمی و رتبه‌ای** است. در این نمودار باید داده‌ها **گسسته** باشند.

