

## مفاهیم و مبانی تحلیل داده‌ها و اطلاعات آماری در نظام آماری مدرن

اشکان شباک<sup>†\*</sup>، جواد حسین‌زاده<sup>‡</sup> و محمد شیری<sup>†</sup>

<sup>†</sup> پژوهشکده‌ی آمار

<sup>‡</sup> مرکز آمار ایران

چکیده. یکی از مهمترین اصول بنیادین آمار رسمی که همه سازمان‌های ملی، منطقه‌ای و بین‌المللی تولیدکننده آمار رسمی در سرتاسر جهان ملزم به رعایت آن بوده و در واقع برای ادامه حیات و فعالیت چنین نهادهایی ضروری است، اصل بی‌طرفی و رعایت اخلاق حرفه‌ای است. اطمینان طیف وسیع ذینفعان این سازمان‌ها و به ویژه کاربران آمار رسمی از استقلال حرفه‌ای سازمان‌های ملی آمار و به طور کلی تر نظام آمار رسمی، به‌عنوان یک سرمایه اجتماعی از مهمترین عوامل موفقیت این سازمان‌ها به‌عنوان نهادهای ملی و حاکمیتی، در تولید و انتشار آمارهای صحیح و معتبر به‌شمار می‌رود. با توجه به نقش مهم آمار رسمی در توسعه و پیشرفت کشورها، این اعتبار و سرمایه اجتماعی اهمیت فراوان و حیاتی دارد. از همین روی از دیر باز یکی از چالش‌های سازمان‌های تولیدکننده آمار رسمی به ویژه سازمان‌های ملی آمار، موضوع ارایه تحلیل و انجام پیش‌بینی بر اساس آمارهای تولیدی بوده است. بنابراین در گذشته سازمان‌های ملی آمار، به دلیل نگرانی از خدشه‌دار شدن استقلال حرفه‌ای و یا درگیر شدن در مسایل و حواشی ناشی از ارایه تحلیل‌ها، همواره از این مقوله دوری جسته و تحلیل را وظیفه خود نمی‌دانستند. با این وجود و با توجه به پیشرفت‌های فزاینده‌ای که در دانش و فناوری در زمینه‌های گوناگون در جوامع بشری رخ داده و منجر به دگرگونی و پیچیده‌تر شدن آن‌ها گشته، لزوم تغییر در رویکردها و روش‌های تولید آمار رسمی اجتناب‌ناپذیر شده است. از همین روی در سال‌های اخیر به ویژه از سال ۲۰۱۰ موضوع مدرن‌سازی سازمان‌های آماری و به تبع آن نظام‌های ملی آمار رسمی در ادبیات موضوع این مقوله و در بالاترین سطوح مرجع و بین‌المللی آمار

\* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۳۹۹/۴/۱۸، پذیرش: ۱۳۹۹/۱۲/۱۳.

رسمی، به طور جدی مطرح شده و دنبال می‌شود. یکی از موضوعاتی که در رویکرد مدرن‌سازی آمار رسمی مطرح است موضوع تحلیل آمارهای تولید و منتشر شده در نظام آمار رسمی و نقش سازمان‌های ملی آمار در این خصوص است. در این مقاله تلاش شده تا ضمن مطالعه این موضوع و بررسی بایدها و نبایدهای تحلیل داده‌ها و اطلاعات آماری توسط تولیدکنندگان آمار رسمی در یک نظام آماری مدرن، به بسترها و چگونگی انجام تحلیل در سازمان‌های ملی آمار به‌عنوان هماهنگ‌کننده و محور اصلی فعالیت‌های آمار رسمی در نظام‌های ملی آمار پرداخته شود.

واژه‌های کلیدی: داده‌ها و اطلاعات آماری، آمار رسمی، مدرن‌سازی نظام ملی آمار، هرم دانایی.

## ۱- مقدمه

اصل چهارم از اصول ده‌گانه بنیادین آمار با صراحت به وظیفه اساسی هر سازمان تولیدکننده آمار رسمی به ویژه سازمان ملی آمار در راستای جلوگیری از تفسیرهای اشتباه و استفاده نادرست از آمار رسمی تولید شده در نظام ملی آماری اشاره می‌کند [۴۸]. در دیدگاه سنتی بسیاری از کارشناسان بر این باور هستند که سازمان‌های ملی آمار و مراکز تولیدکننده آمار رسمی باید از هرگونه تحلیل و تفسیر آمارهای تولید شده، خودداری کرده و تنها وظیفه دارند به امر تولید و انتشار آمار رسمی بپردازند. دلیل این دیدگاه پرهیز از حواشی و چالش‌هایی بود که احتمال داشت در صورت ورود سازمان‌های ملی آمار به انجام تحلیل و تفسیر آمارها برای سازمان‌های مزبور ایجاد شود که در نتیجه ضمن بازداشتن مراکز آماری از انجام وظیفه اصلی خود که همانا تولید و انتشار آمار درست و بی‌طرفانه است، استقلال حرفه‌ای سازمان‌های ملی آمار که از دیگر اصول مهم و اساسی اصول بنیادین آمار است را نقض نماید. این نگاه که البته در جای خود درست است، موجب شده که برخی سازمان‌های ملی آمار به طور سنتی از ورود به موضوع تحلیل آمارهای تولیدی به شدت خودداری کرده و آن را خلاف حفظ استقلال حرفه‌ای خود بدانند. با این وجود اهمیت انجام اقدامات و فعالیت‌هایی برای ایجاد سهولت در فهم و درک کاربران و تسهیل شرایط استفاده آنان، برای مراکز تولیدکننده آمار رسمی کمتر از تولید

محصولات آماری نیست. این اهمیت آنچنان است که بخش آماری سازمان ملل مجموعه راهنمای ویژه‌ای را در چهار جلد تحت عنوان «معنی‌دار کردن داده‌ها»<sup>۱</sup> و برای درک و فهم بهتر کاربران از محصولات آماری سازمان‌های تولیدکننده آمار رسمی، به ویژه مراکز آماری، طی سال‌های ۲۰۰۹ تا ۲۰۱۴ منتشر کرده است. این مجموعه نتیجه کار فشرده کارگروه تخصصی بوده که برای ارتباطات و انتشارات آماری توسط کنفرانس آمارشناسان اروپایی<sup>۲</sup> و زیر نظر بخش اروپایی سازمان ملل (کمیسیون اقتصادی سازمان ملل برای اروپا)<sup>۳</sup> تشکیل شده بود.

بنابراین با توجه به اصل جلوگیری از تفاسیر نادرست آمار و در راستای تلاش برای درک و فهم بهتر کاربران، بخش آمار سازمان ملل تأکید دارد تا سازمان‌های تولیدکننده آمار رسمی در راستای تعهد و مسئولیت خود در جلوگیری از تفاسیر نادرست و بهره‌برداری مناسب کاربران آمار اقدام کرده، فراداده‌های مناسب به این منظور را با طراحی شیوه‌ها و در قالب فرایندهای مدون، به صورت نظام‌مند تولید نمایند [۴۷]. یکی از بهترین روش‌ها برای این کار علاوه بر تغییر در نحوه انتشار آمار و دادن آموزش‌های لازم به ذینفعان، به ویژه اصحاب رسانه، ارایه تحلیل‌های معین ساده و البته کاربرپسند بر روی داده‌ها و اطلاعات آماری تولید شده، در قالب فراداده و یا گزارش‌ها و نشریات تحلیلی مستقل است. در ۴۸<sup>۴</sup> آمین نشست کمیسیون آمار سازمان ملل نیز موضوع تحلیل داده‌ها، به‌عنوان یک فعالیت مهم در سازمان‌های ملی آمار برشمرده شده است. سایر سازمان‌های مرجع بین‌المللی آمار رسمی، مانند سازمان همکاری اقتصادی و توسعه<sup>۴</sup> (OECD) و یا اداره آمار کانادا، به اهمیت تحلیل داده‌ها و نقش آن در فعالیت‌های سازمان‌ها و مراکز آمار رسمی اشاره کرده‌اند.

از سوی دیگر در سال‌های اخیر، با توجه به پیشرفت دانش و فناوری، پیچیده‌تر شدن جوامع و نیازهای روزافزون آماری، یکی از موضوعات کلیدی که در مجامع و نشست‌های علمی و کارشناسی بین‌المللی آمار رسمی مطرح شده، موضوع مدرن‌سازی نظام آمار رسمی در کشورهای جهان است. در این نگاه نوین، در کنار تولید آمار رسمی، تحلیل

آمارها نیز از دیگر وظایف اصلی یک سازمان ملی آمار مدرن، به شمار رفته است. البته این تحلیل دارای تعریف خاص خود بوده و باید به گونه‌ای باشد که نه تنها اصول بی‌طرفی و استقلال حرفه‌ای را نقض نکند، بلکه منجر به خدشه‌دار شدن اصل محرمانگی اطلاعات نیز نشود [۲۹]. در همین راستا است که در نظام جامع مدیریت کیفیت آمارهای رسمی، تحلیل داده‌ها به‌عنوان یکی از اصول مهم چارچوب‌های کیفیت، هم به جهت ارزیابی کیفیت محصولات آماری و هم به جهت ایجاد وضوح<sup>۵</sup> (توضیح آمارهای تولید شده) به شمار می‌رود. بر همین اساس در مدل عمومی فرایند کسب و کار آماری، که از نوین‌ترین الگوهای کاری پیشنهاد شده توسط بخش آماری سازمان ملل برای سازمان‌های ملی آمار مدرن است، یک مرحله جداگانه به تحلیل اختصاص یافته است. در این الگو فاز تحلیل نه تنها برای ارزیابی کیفی اولیه محصول آماری و تأیید آن پیش از انتشار است، بلکه همچنین برای آماده کردن توضیحات و تفسیرهایی است که با عنوان فراداده، به درک و فهم بهتر کاربر از آمار تولیدی و جلوگیری از تفسیرهای نادرست کمک می‌کند [۲۱].

از آنچه گفته شد برمی‌آید که در یک نظام آماری مدرن، تحلیل داده‌ها و اطلاعات آماری تولید شده از آمارگیری‌ها، چه برای بررسی و ارزیابی کیفیت داده‌ها و اطلاعات آماری باشد و چه برای تسهیل استفاده کاربران و یا جلوگیری از تفسیرهای نادرست، در هر صورت وظیفه‌ای مهم برای سازمان ملی آمار و در راستای تکمیل فرایندهای تولید و انتشار آمارها است. بنابراین در یک رویکرد مدرن ضرورت دارد تا سازمان‌های ملی آمار هرچه سریع‌تر و دقیق‌تر به تحلیل آمارهای تولیدشده بپردازند. با این‌وجود شیوه و روش انجام این تحلیل‌ها و نیز چگونگی آن‌ها باید به گونه‌ای باشد تا موجب نقض بی‌طرفی و استقلال حرفه‌ای مراکز آماری نشود.

برای انجام هرچه بهتر تحلیل داده‌ها توسط سازمان‌های ملی آمار، لازم است نخست درک درستی از تعاریف اصلی و مبانی نظری تحلیل و انواع شیوه‌های آن داشت تا بتوان مشخص کرد که از یک سو با توجه به ویژگی‌های هر تحلیل و از سوی دیگر شرایط حرفه‌ای سازمان‌های آماری، کدام نوع تحلیل مناسب این سازمان‌ها خواهد بود. بنابراین

در ادامه این نوشتار تلاش می‌شود که در بخش دوم، ابتدا به تعاریف اولیه و مهم مورد استفاده در این مقوله پرداخته شود تا از این طریق امکان رسیدن به برداشت یکسان و فهم مشترک از مفاهیم اصلی بهتر فراهم شود. سپس در این مقاله به دسته‌بندی انواع تحلیل، از منظر این مطالعه پرداخته و ویژگی‌ها و ابعاد هر یک تشریح می‌شود. در بخش چهارم با بررسی منابع، مستندات و دستورالعمل‌های مراجع بین‌المللی، ضمن اثبات لزوم و دلیل انجام تحلیل توسط تولیدکنندگان آمار رسمی و سازمان‌های ملی آمار و جایگاه آن، به بررسی ویژگی‌ها و شرایط چنین تحلیلی در بستر و زمینه آمار رسمی اشاره خواهد شد.

## ۲- مفاهیم بنیادین

در این بخش برای آشنایی بیشتر خوانندگان، ضمن بررسی مفاهیم و تعاریف مهم و بنیادین در ادبیات موضوع، به مرور نوشتگان مرتبط نیز پرداخته می‌شود.

### ۲-۱- داده‌ها<sup>۶</sup>

در فرهنگ فارسی عمید نوشته شده است: «داده عبارت است از اطلاعاتی است که برای یک کار آماری گردآوری می‌شوند». همچنین در واژه‌نامه مصوب فرهنگستان پارسی آورده شده است که داده‌ها اعداد و متون و مانند آن‌ها هستند که می‌توان آنها را به شکل خاص مرتب کرد تا امکان ذخیره‌سازی و پردازش آنها به وسیله رایانه فراهم شود». در فرهنگ انگلیسی وبستر نیز ترجمه واژه داده‌ها عبارت است از: «اطلاعات واقعی که به‌عنوان پایه‌ای برای استدلال، بحث و یا محاسبه استفاده می‌شوند». همانطور که مشاهده می‌شود این تعاریف همه به شکل عام هستند و گویا مفهوم داده و اطلاع را یکسان فرض کرده‌اند. برای دریافتن مفاهیم تخصصی‌تر و در پیوند با آمار رسمی، بهتر است نگاهی به واژه‌نامه‌های تخصصی اشاره شود. دانشنامه علوم آماری داده را حقیقتی<sup>۷</sup> می‌داند که بر اساس آن می‌توان نتیجه‌گیری کرد. این حقیقت می‌تواند عدد و یا هر چیز دیگری باشد [۳۱]. اداره آمار کانادا که مطالعات و مستندات آن به‌عنوان یک مرجع بین‌المللی در آمار

رسمی مطرح است بر مبنای همین تعریف، داده‌ها را در آمار رسمی این‌گونه تعریف می‌کند: «حقایق و اشکالی که بر اساس آن‌ها می‌توان نتیجه‌گیری کرد» [۱۴]. دوج و همکارانش در دانشنامه آمار، داده‌ها را نتایج مشاهده‌ها از یک جامعه و یا نمونه آماری تعریف کرده‌اند [۷]. بر اساس این تعریف، سازمان OECD نیز به‌عنوان یک سازمان مهم در آمار رسمی که استانداردها و دستورالعمل‌های آن به‌عنوان مرجع به کار می‌روند، داده‌ها را به این شکل تعریف کرده است: «ویژگی‌ها یا اطلاعاتی که معمولاً به صورت عددی بوده و از طریق مشاهده گردآوری می‌شوند» در توضیح این مفهوم آمده است که داده‌ها نمایش فیزیکی اطلاعات هستند که امکان ایجاد ارتباط، تفسیر و یا پردازش آن‌ها فراهم شده باشد. همچنین در ادامه تعریف توضیح داده شده است که در این مفهوم، داده‌های آماری<sup>۸</sup> داده‌هایی هستند که برای تولید آمار از منابع مختلف (آمارگیری و یا ثبت‌های اداری) به دست می‌آیند [۳۷].

بنا بر این و با توجه به تعریف‌های بالا، در این نوشتار داده‌ها را به صورت زیر تعریف می‌کنیم: «حقایقی به شکل عدد، تصویر، صوت، متن و مانند این‌ها که از نتایج مشاهدات بر روی پدیده‌ها، از جمله جامعه یا نمونه آماری، به دست آمده‌اند.»

## ۲-۲- اطلاعات

در فرهنگ عمید اطلاع در قالب اسم، به معنی آگاهی و در قالب اسم مصدر، به معنی آگاه‌شدن به کار می‌رود. در واژه‌نامه مصوب فرهنگستان نیز آنچه به دگرگونی وضعیت در نظام‌های طبیعی و ساختگی یا نشانه‌شناختی می‌انجامد، اطلاع نامیده شده است. واژه «اطلاعات» در متون علمی و از دیدگاه‌های گوناگون دارای تعریف‌های متفاوتی است [۳۵]. با این وجود و به طور عام، آنچه از پردازش، سازماندهی و ساختاردهی داده‌ها به‌منظور معنی‌دهی و قابل استفاده کردن آن‌ها حاصل می‌شود، اطلاعات تعریف شده است [۲۰]. در ارتباط با آمار رسمی، اداره آمار کانادا نیز «نتیجه حاصل از هرگونه عملیات ذخیره و ثبت، دسته‌بندی و سازماندهی، طبقه‌بندی و تفسیر بر روی داده‌ها که در قالب

چارچوب معینی انجام شود» را اطلاع نامیده است. در توضیح این واژه آمده است که اطلاعات حاصل شده از داده‌ها مطابق این تعریف، در واقع به داده‌ها معنی می‌بخشد [۱۴]. همچنین OECD نیز اطلاعات را شناخت از موجودیت‌هایی مانند حقایق، پیشامدها، فرایندها یا ایده‌ها تعریف کرده است که در یک زمینه خاص به آن‌ها معنی می‌بخشد [۳۷].

با توجه به موارد گفته شده، در این مقاله «نتایج حاصل از هرگونه عملی مانند دسته‌بندی، سازمان‌دهی، ساختاردهی، پردازش و تفسیر داده‌ها به‌منظور معنی‌دهی و قابل استفاده کردن آنها»، تعریف اطلاعات شناخته می‌شود. بدیهی است که اطلاعات با توجه به نوع داده‌ها و عملیاتی که بر روی آن‌ها انجام می‌شود می‌تواند کمی یا کیفی (عددی یا غیر عددی) باشد.

## ۲-۳- آمار

در واژه‌نامه انگلیسی آکسفورد مجموعه‌ای از اطلاعات که در قالب عدد نشان داده شده‌اند، آمار نامیده شده است [۳۹]. این در حالی است که در فرهنگ‌نامه انگلیسی وبستر، تعریف واژه آمار مجموعه‌ای از داده‌های کمی ذکر شده است. این نشان می‌دهد که در فرهنگ‌نامه‌های عمومی، جایی آمار را داده دانسته و در جایی دیگر مجموعه‌ای از اطلاعات فرض کرده است. در واقع به نظر می‌رسد در این نوع از واژه‌نامه‌ها به تمایز میان داده و اطلاع توجهی نشده است. شایان گفتن است که در هر دو واژه‌نامه تعریف دیگر آمار، علم گردآوری و تحلیل آمارها اشاره شده است و بر همین اساس فرهنگ فارسی عمید نیز آمار را در دو قالب «داده‌های عددی به دست آمده از بررسی اطلاعات درباره‌ی یک مجموعه» و «علم گردآوری و بررسی داده‌های عددی» تعریف کرده است. همچنین دانشنامه علوم آماری نیز فرای تعریف آمار به‌عنوان یک علم، آن را مجموعه داده‌های عددی تعریف کرده است [۳۱]. بنابراین و با توجه به این تعاریف، این پرسش پیش می‌آید که آمار از نظر ماهوی داده است یا اطلاع و در واقع وجه تمایز میان آن‌ها چیست؟

برای پاسخ به این پرسش، لازم است تخصصی‌تر به تعریف آمار نگاه شود. دانشنامه موجز آمار، «مجموعه‌ای از داده‌ها که در خصوص یک پدیده گردآوری شده و به صورت عددی نشان داده شده‌اند» را آمار نامیده است. واژه‌نامه OECD نیز آمار را داده‌های عددی مربوط به مجموع مشاهدات دانسته است [۳۷]. اگر چه در این تعریف‌های تخصصی مفهوم آمار بسیار نزدیک به مفهوم داده‌ها دانسته شده ولی با دقت بیشتر در عین حال می‌توان نتیجه گرفت که آمار به خودی خود یک داده خام نیست بلکه مجموعه‌ای از داده‌های خام اولیه (نه لزوماً عددی) که البته باید به صورت عددی بیان شوند، آمار را تشکیل می‌دهند. با توجه به تعریف اطلاعات، که پیشتر اشاره شد، این بدان معنی است که آمار بیش از آن که داده باشد یک اطلاع است. بر این اساس اداره آمار کانادا تعریف بسیار خوبی از آمار ارائه کرده است: «نوعی از اطلاعات است که از عملیات ریاضی بر روی داده‌ها حاصل می‌شود» [۱۴]. بنابراین می‌توان آمار به صورت «نوعی از اطلاع که به صورت عددی بیان شده است» تعریف کرد که در این مقاله ما این تعریف را مبنا قرار می‌دهیم.

## ۲-۴- اطلاعات آماری

با توجه به تعریف‌های ارائه شده در فوق می‌توان عبارت اطلاعات آماری را به‌عنوان یک تعبیر جدید به تعاریف فوق افزود. بر این اساس، اطلاعات آماری، اطلاعاتی حاصل از پردازش داده‌ها هستند که می‌تواند به صورت خروجی‌هایی مانند آمار (شامل شاخص و نماگر)، جدول آماری، نمودار آماری و نقشه آماری تولید شود. بنابراین در واقع محصولات آماری مراکز تولیدکننده آمار رسمی، معادل واژگان داده‌ها و اطلاعات آماری هستند. برای آشنایی بهتر با مفهوم و ضرورت تحلیل داده‌ها و انواع آن، چند تعریف پایه‌ای دیگر شامل «دانش، دانایی و بینش» نیاز است که به دنبال تعاریف گفته شده در بالا اضافه شود. بدین منظور و برای درک بهتر این مفاهیم باید ابتدا هرم دانایی تعریف شود.





شکل ۱- هرم دانایی (DIKW)

## ۲-۵-هرم دانایی

شکل ۱، هرم دانایی را نشان می‌دهد. در این هرم که به هرم DIKW نیز معروف است، دانش<sup>۹</sup> «انباشت هرچه بیشتر اطلاعات و ترکیب آن‌ها برای رسیدن به فهم و درک بهتر از پدیده‌های گوناگون مشاهده شده» تعریف می‌شود [۸]. در واقع دانش، دانستن این نکته است که اطلاعات به دست آمده «چه» هستند و «چگونه» استفاده می‌شوند [۲۵]. دانایی<sup>۱۰</sup> نیز به عنوان درک عمیق از پدیده‌ها و چرایی آن‌ها است [۳۸]. در واقع این مفهوم «توانایی قضاوت و تشخیص در خصوص جنبه‌های مختلف دانش کسب شده و به‌کارگیری آن است» [۴۲]. کیفیت این توانایی بستگی به درک و فهم ایجاد و انباشته شده از داده‌ها، اطلاعات و دانش، نسبت به پدیده‌ها و مشاهدات دارد. هانتر ویتنی در کتاب خود، هرم دانایی را گام‌های متوالی، برای افزایش ادراک<sup>۱۱</sup> از پدیده‌ها و مشاهدات، معرفی می‌کند که در نهایت به ایجاد بینش<sup>۱۲</sup> منتهی می‌شود [۴۹]. ادراک در واقع خود یک مرحله مستقل و یا مفهوم جداگانه از هرم دانایی نیست، بلکه پیش‌نیاز و یا ابزار پشتیبان برای گذار از هر مرحله به مرحله بعد در این هرم است. به بیان دیگر رسیدن از داده به دانایی با ادراک (به ترتیب درک الگوها، درک روابط و درک اصول<sup>۱۳</sup>) میسر خواهد بود. این ادراک

انباشته در نهایت با گذار از دانش و دانایی و عمل بر اساس آن‌ها، به شناخت کامل می‌رسد که بینش تعریف می‌شود [۱۱].

ویتنی [۴۸] بینش را عمیق‌ترین سطح از شناخت داده، اطلاع، دانش و دانایی تعریف می‌کند که تصویر واقعی و حکیمانه‌ای از پدیده‌ها و مشاهدات به دست می‌دهد. به بیان فلسفی، بینش ماهیت اساسی دانش و جوهر دانایی است.

فرهنگستان زبان و ادب پارسی بینش را اینگونه تعریف می‌کند: «درک عمیق از خود و روابط با دیگران که تجربه‌های پیشین را روشن یا فرد را در حل مسئله‌ای یاری می‌کند». در برخی متون بینش را بصیرت نیز نامیده‌اند که فرایندی است که به وسیله آن مسائل حل می‌شوند و مشخص‌کننده سازماندهی یا نوسازی است که به صورت عادت در فرد نهادیه شده و او را به فهم روابط مربوط به راه‌حل توانا می‌سازد. در بیان کلی اگر ادراک، شناختی به دست می‌دهد تا دانش و دانایی کسب شده از داده‌های ابتدایی طی فرایندی تجمیعی، به عمل و یا تصمیمی خردمندانه بیانجامد، بینش ایجاد و تقویت ظرفیت نهادینه شدن این فرایند و تبدیل آن به سبکی پیوسته و پایدار است. شکل ۲ این مفاهیم را در یک نما نشان می‌دهد.



شکل ۲- فرایند دستیابی به بینش

در مقوله مدیریت داده‌ها، برای رسیدن به بینش درباره پدیده‌ای که در ارتباط با آن داده‌ها گردآوری شده‌اند، دو مفهوم گذشته‌نگاری<sup>۱۴</sup> و آینده‌نگاری<sup>۱۵</sup> نیز وجود دارند که لازم است در این مقاله به اختصار تعریف شوند. در واقع به بیان دقیق‌تر، مفهوم بینش به اجزای

«بینش از گذشته»، «بینش از حال» و «بینش از آینده» بخش می‌شود. بینش از حال که معمولاً همان بینش به طور عام نامیده می‌شود، در واقع درک و فهم عمیق از محیط کنونی و آنچه در حال حاضر رخ می‌دهد، است. بینش از گذشته را گذشته‌نگاری و بینش از آینده را آینده‌نگاری می‌نامند. تعریف بینش گذشته‌نگارانه همان بصیرتی است که پس از وقوع رویدادی به درک این که چه رخ داده کمک می‌کند و بینش آینده‌نگارانه نیز با درک آنچه که اتفاق خواهد افتاد به نگاشت آینده کمک می‌کند. در واقع یک بینش کامل، بینشی است که با درک و فهم گذشته و حال، به ساخت آینده مطلوب می‌پردازد [۴۴]. شایان گفتن است که گذشته‌نگاری با گذشته‌نگری<sup>۱۶</sup> متفاوت است. در واقع منظور از گذشته‌نگری کشف وقایع و رویدادهای گذشته و نگرش در آن‌ها برای کسب دانش و رسیدن به دانش است. در حالی که همان‌گونه که گفته شد، گذشته‌نگاری مطالعه همه جانبه گذشته و بررسی روابط میان رخدادها و تحلیل آن‌ها، برای رسیدن به درک و شناخت عمیق از آنچه در گذشته رخ داده، می‌باشد. همین معنی برای تفاوت میان آینده‌نگاری و آینده‌نگری<sup>۱۷</sup> نیز وجود دارد. آینده‌نگری کسب دانش و رسیدن به مرحله‌ای از دانایی است که بتوان بر مبنای آن به پیش‌بینی و کشف آینده پرداخت، حال آن‌که آینده‌نگاری از جنس بینش است و با درک و فهم خلاقانه و فعالانه از آنچه روی خواهد داد به نگاشت آینده محتمل و البته مطلوب می‌پردازد. برای مطالعه بیشتر به مراجع [۴۱] و [۳۴]، مراجعه کنید.

بنابر آن‌چه تا کنون گفته شد، برای درک و فهم عمیق از پدیده‌ها و مشاهدات باید مسیری طولانی از گردآوری داده‌ها تا رسیدن به اقدامی درست، بر مبنای تصمیم خردمندانه، پیموده شود. این مسیر در شکل ۳ به خوبی نشان داده است.



شکل ۳- مسیر تبدیل داده به اقدام درست

در شکل ۳، نقش تحلیل بر روی داده‌های مزبور بسیار کلیدی و البته حیاتی است. در واقع برای کاربردی کردن داده‌ها و اطلاعات آماری و تبدیل آن‌ها به تصمیم و به تبع آن اقدام، وجود تحلیل ضروری است. در واقع بدون تحلیل، داده‌ها و اطلاعات از همان آغاز هیچ معنی و مفهومی نخواهند داشت و پیمودن راه رسیدن از داده به بینش و گذار از هرم دانایی امکان‌پذیر نیست. برای روشن شدن بهتر موضوع به یک مثال توجه کنید. فرض کنید گزارشی ارائه شده که در ماه گذشته ۱۴۰۰۰ پیامک ارسال شده است. این داده‌ها و اطلاعات به تنهایی ارزش چندانی ندارند. اما اگر بر اساس یک روش تحلیلی، داده‌ها کاویده شده و گزارش جدیدی حاصل شود که ضمن دسته‌بندی پیامک‌ها به تفکیک مخاطب و قصد ارسال، مشخص نماید که به طور متوسط ۳/۷ پیامک به ازای هر کاربر و به چه منظور ارسال شده است. حال فرض کنید تعداد پیامک‌های ماه گذشته را بتوان با ۱۲ ماه گذشته، با همان دسته‌بندی و به‌طور مشابه مقایسه و روند ارسال را به تفکیک هر مخاطب و منظور از ارسال مطالعه کرد. با توجه به این‌که هزینه هر پیامک معلوم است، این همان ارزش افزوده‌ای است که فرایند تحلیل به داده‌های اولیه اهدا کرده است.

در ادبیات موضوع علوم داده، عبارت چرخه حیات داده، یک مفهوم کاملاً تعریف شده و عبارت است از یک سلسله مراحل مشخص که به ترتیب از گردآوری یا ایجاد داده تا ذخیره و احتمالاً حذف آن به ترتیب و به‌منظور مدیریت بهتر داده‌ها برای استفاده و استفاده‌های مجدد انجام می‌شود [۱۸]. البته تعداد این مراحل و جزئیات آن‌ها با توجه به بستر تعریف و محیط مورد استفاده تا حدودی تفاوت کند اما کلیت آن مشابه است و در همه آن‌ها مرحله تحلیل داده‌ها، به‌عنوان یکی از مراحل اصلی وجود دارد [۲۸]. ویتنی [۴۸] مراحل چرخه حیات داده‌ها را در شکل ۴ معرفی می‌کند.



شکل ۴- نمودار چرخه حیات داده‌ها (هانترویتنی - ۲۰۱۲)

در پایان این بخش و پیش از پرداختن به انواع تحلیل و توضیح ویژگی‌ها و ابعاد آن‌ها، با ارایه تعریف نظری تحلیل داده‌ها تلاش می‌شود تا این مفهوم در ارتباط با داده‌ها و اطلاعات به خوبی تبیین شود.

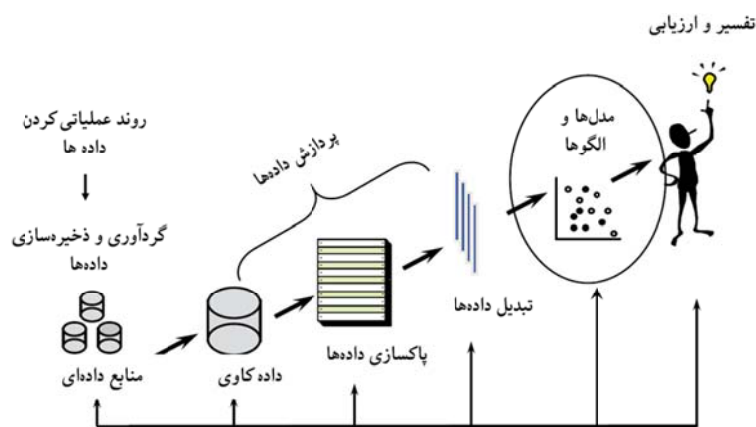
## ۲-۶- تحلیل داده‌ها

واژه تحلیل خود به تنهایی، واژه بسیار پرکاربردی است، اما در عین حال نمی‌توان معنی دقیق و یکتایی برای آن پیدا کرد و چون در هر بستری متناسب با مفاهیم آن تعریف می‌شود رسیدن به مفهوم مناسب برای آن دشوار است. البته از یک زاویه معنی لغوی آن همان موشکافی و باز کردن اجزای مسایل است. در واژه‌نامه انگلیسی آکسفورد، تحلیل «مطالعه دقیق یا بررسی چیزی به منظور درک بیشتر در مورد آن»، نامیده شده است. در فرهنگ‌نامه انگلیسی وبستر نیز تحلیل «بررسی دقیق هر چیزی پیچیده به منظور درک ماهیت آن و یا تعیین ویژگی‌های اصلی آن» تعریف شده است. در ادبیات موضوعی علم داده‌ها، برخی تحلیل داده‌ها را تبدیل داده به اطلاعات تعبیر کرده‌اند که مرحله‌ای ضروری، در ادامه پردازش و جدول‌بندی داده‌های گردآوری شده بوده و باعث تبدیل داده‌ها به

اطلاعاتی می‌شود که قابل فهم و روشن باشند. این اطلاعات به دست آمده بر پایه تحلیل، می‌تواند مانند چراغی به جامعه نور افشانی کرده و موجب توسعه اجتماعی شود [۱۱]. در دانشنامه موجز آمار تحلیل داده‌ها، «مجموعه اقداماتی برای استخراج اطلاعات از داده‌ها» تعریف شده است [۱۹].

بارت باسنس تحلیل را یک اصطلاح که اغلب به‌عنوان جایگزینی با علم داده، داده‌کاوی، کشف دانش و مانند آن استفاده می‌شود، بیان می‌کند. از دید او تحلیل، استخراج الگوهای صحیح و یا مدل‌های ریاضی درست از میان داده‌ها و اطلاعات آماری است و روش‌های آماری، یادگیری ماشین، درخت تصمیم، الگوریتم ژنتیک، هوش مصنوعی و مانند این‌ها ابزاری برای انجام تحلیل هستند. بنابراین تحلیل داده‌ها یک مهارت فناورانه است که باید توسط افراد متخصص آموزش دیده و آشنا با علم داده و داده‌کاوی به‌طور حرفه‌ای و مبتنی بر نرم‌افزار و فناوری اطلاعات انجام شود. شکل ۵ به طور مفهومی تحلیل و کاربرد آن را به خوبی نشان می‌دهد [۷]. ایکسیا و گانگ در سال ۲۰۱۵، تحلیل داده‌ها را «فرآیندی برای بازرسی، پیرایش، تبدیل و مدل‌سازی داده‌ها»<sup>۱۸</sup>، تعریف کرده‌اند که با هدف «کشف و استخراج اطلاعات مفید»، «نتیجه‌گیری آگاهی بخش» و «حمایت از تصمیم‌سازی» انجام می‌شود. از نظر ایشان تحلیل داده‌ها دارای چندین جنبه و رویکرد، شامل فنون‌های متعدد با نام‌های مختلف است که در زمینه‌های گوناگون علمی از قبیل مدیریت کسب و کار و سایر زمینه‌های اجتماعی و اقتصادی کاربرد داشته و نقش مهمی در تصمیم‌گیری‌های مؤثر دارد [۵۰].

سازمان OECD نیز تحلیل را «فرایند تبدیل داده‌های خام به اطلاعات قابل استفاده تعریف کرده است» [۳۷]. در توضیح این معنی آمده است که تحلیل در واقع ارزش‌افزوده‌ای است که به آمارهای تولید شده، در قالب محصول آماری مستقل (شامل گزارش‌ها و یا مستندهای تحلیلی)، اضافه می‌شود.



شکل ۵- مدل مفهومی فرایند تحلیل داده‌ها (باسنس - ۱۴۰۲)

اداره آمار کانادا نیز در همین راستا تحلیل را روندی تعریف می‌کند که «از طریق بررسی و تفسیر داده‌ها به پاسخ سؤالاتی در خصوص حقیقت و یا پدیده‌ای که داده‌ها در خصوص آن گردآوری شده‌اند، می‌انجامد. در توضیح این مفهوم آمده است که مراحل اصلی در فرایند تحلیل شناسایی موضوعات، تعیین مناسبت و دسترس‌پذیری داده‌ها، تعیین روش و ابزار مناسب برای تحلیل و در پایان ارزیابی، خلاصه کردن و اطلاع دادن نتایج تحلیل است [۱۲]. در برخی متون تحلیل را ابزار و راه رسیدن به بینش تعریف کرده‌اند که تصمیم و اقدام مبتنی بر بینش به واسطه آن انجام خواهد شد [۳۲]. آلت سیمونز می‌گوید تحلیل، کشف الگوها و روندهای درون داده‌ها است [۳]. بنابراین بهره‌برداری از داده‌ها و اقدام صحیح بر اساس آن‌ها و بهبود فرایند تصمیم‌گیری، نیازمند داشتن ادراک و بینش کافی است که پیش‌نیاز رسیدن به این درجه از ادراک و بینش، همانطور که پیش‌تر نیز گفته شد، گذار از هرم دانایی بوده که تحلیل ابزار مؤثر این گذار است.

در یکی از مطالعات انجام شده برای طراحی آمارگیری‌های نمونه‌ای خانواری، توسط گروه کارشناسی بخش آمار سازمان ملل، تحلیل را یک فعالیت ضروری در آمار رسمی دانسته که فرایند تبدیل داده‌های گردآوری شده در آمارگیری‌ها را به اطلاعات آماری تکمیل ساخته و از آرایه یک جدول ساده تا روش‌های پیچیده‌تر استنباط آماری گسترش می‌یابد [۲۶].

بنا بر این می‌توان گفت که به دست آوردن درک عمیق و بینش از داده‌های گردآوری شده به منظور تصمیم‌گیری بهتر، جمع‌بندی مناسب و جامعی برای تبیین مفهوم تحلیل داده‌ها بوده که مورد پذیرش عام است. به عبارت دیگر، تجزیه و تحلیل فرایندی علمی، جهت تبدیل داده به درک عمیق و بینش و سپس تصمیم‌گیری کارآمد و متعاقب آن اقدامات بهینه و مناسب است. حال با توجه به همه مفاهیم و ادبیات گفته شده در بالا می‌توان تعریف مناسب در ادبیات آمار رسمی که مورد نظر این مقاله است را برای مفهوم تحلیل داده‌ها ارایه داد: «تحلیل داده‌ها فرایندی است که با بررسی و پردازش داده‌ها، به استخراج الگوهای صحیح از میان آن‌ها پرداخته و آن‌ها را تبدیل به ابزار کاربردی و مؤثر برای اقدام و عمل مبتنی بر درک و بینش می‌کند.»

حال با دانستن مفهوم دقیق تحلیل داده‌ها و جنبه‌های آن، در ادامه مقاله به معرفی انواع تحلیل پرداخته می‌شود تا با شناخت بهتری بتوان به موضوع تحلیل داده‌ها و اطلاعات آماری در آمار رسمی ورود کرد تا انجام تحلیل و ارایه گزارش‌های آن، ضمن خدشه‌دار نکردن بی‌طرفی سازمان‌های ملی آمار و پرهیز از هرگونه سوگیری، مطابق وظایف و رسالت یک سازمان مدرن تولیدکننده آمار رسمی باشد.

### ۳- انواع تحلیل داده‌ها

تحلیل داده‌ها، به طور کلی از ابعاد گوناگون قابل دسته‌بندی هستند. به عنوان مثال از نظر نوع داده، به دو دسته کلی تحلیل کمی و کیفی<sup>۱۹</sup> تقسیم می‌شوند. تحلیل کیفی بر روی داده‌هایی انجام می‌شود که به صورت غیر ساختاریافته هستند. داده‌های غیر ساختاریافته داده‌هایی هستند که یک منبع از پیش تعیین شده مانند یک بانک داده و یا مدل تعریف‌شده‌ای ندارند و از منابع مختلف (مانند مصاحبه، اسناد متنی، وبلاگ‌ها، نظرسنجی، تصویر، فیلم) که در واقع در دسترس بوده‌اند به دست آمده‌اند و بنابراین الگو و ویژگی مشخصی ندارند. همین‌طور داده‌های کیفی و رسته‌ای که بیشتر برای اندازه‌گیری و توصیف ویژگی‌های مختلف مربوط به پدیده‌ها و اشیا به کار می‌روند. در این نوع تحلیل



تلاش می‌شود که برای حجم وسیعی از چنین داده‌هایی توضیح و تفسیرهای مناسب، بر اساس درک و استنباط تحلیل‌گر از پدیده مورد مطالعه، ارائه شود [۴]. بنا بر این تحلیل کیفی، درصد جمع‌آوری داده و ایجاد یک شناخت عمیق از دلایل و انگیزه‌ها برای وقوع یک رویداد خاص است. در تحلیل کمی، داده‌ها ساختاریافته هستند و معمولاً از نوع عددی بوده و بر اساس منابع مشخص و تعریف شده با ویژگی‌های یکسان و معلوم تولید و یا گردآوری شده‌اند. تحلیلی کمی به بررسی نظام‌مند پدیده‌ها و اشیا از طریق روش‌های معمول ریاضی و محاسباتی می‌پردازد [۱۷].

از نظر مخاطب و هدف نیز معمولاً تحلیل‌ها به دو نوع بنیادین و کاربردی<sup>۲۰</sup> دسته‌بندی می‌شوند. تحلیل بنیادین در پژوهش‌هایی از همین نوع که به پژوهش‌های نظری<sup>۲۱</sup> نیز معروف هستند به‌کار می‌رود و هدف از انجام آن شکافتن مرزهای دانش و ایجاد نظریه‌های علمی از طریق تحلیل است. مخاطب این نوع تحلیل تنها مجامع علمی و پژوهشی هستند. تحلیل کاربردی نیز همانطور که از نام آن هویدا است، تحلیلی است که هدف آن بررسی پدیده‌ها برای پاسخ عملی به پرسش‌های مطرح در خصوص آن پدیده و پیدا کردن راه‌های اجرایی برای حل و توضیح مسایل است. در این نوع تحلیل تلاش می‌شود تا نیازهای کاربران پاسخ داده شود. طبیعی است این نوع تحلیل، طیف وسیعی از کارشناسان و خبرگان را به‌عنوان مخاطب پوشش می‌دهد و در هر دو جنبه کمی و کیفی، بسته به نوع داده‌ها، استفاده می‌شود.

تحلیل‌های کاربردی بر اساس روش‌های به‌کار گرفته شده و اهداف تحلیل داده‌ها، به انواع متعددی دسته‌بندی می‌شوند. در متون مختلف تعداد این دسته‌ها متفاوت است. به‌عنوان مثال جفری لیک تحلیل‌های کاربردی را در شش گروه اصلی تحلیل‌های «توصیفی، اکتشافی، استنباطی، پیشگویانه، علی و مکانیزمی»<sup>۲۲</sup> تقسیم کرده است [۳۳]. از سوی دیگر موسسه گارتنر، تحلیل‌های کاربردی را در چهار دسته کلی خلاصه کرده است که عبارت‌اند از: «توصیفی، اکتشافی، پیشگویانه و تجویزی» [۲۲]. این نوع از تقسیم‌بندی در بسیاری از منابع دیگر نیز تکرار شده است [۳۶] و [۹]. داونپورت و همکارانش نیز

تحلیل‌های کاربردی را به سه دسته کلی‌تر تحلیل‌های «توصیفی، پیشگویانه و تجویزی تقسیم کرده‌اند»<sup>۲۳</sup> [۱۷].

شایان توجه است که در مطالعه پیش‌روی با توجه به طیف گسترده کاربران تحلیل کاربردی، و نقش آن در کسب و کار آمار رسمی، به این نوع از تحلیل پرداخته می‌شود. شایان گفتن است که دسته‌بندی دقیق انواع تحلیل کاربردی در دسته‌های جداگانه دشوار است و در بسیاری موارد هم‌پوشانی بالایی در میان دسته‌بندی‌های تحلیل داده‌ها وجود دارد. با این وجود سعی شده است تا با مطالعه منابع مختلف و بررسی ادبیات علمی موضوع بتوان یک جمع‌بندی جامع و کاربردی در این زمینه ارائه داد. بر این اساس در این مقاله، در مجموع تحلیل‌های کاربردی در چهار دسته‌ی کلی زیر دسته‌بندی می‌شوند.

### ۳-۱- تحلیل توصیفی

این تحلیل ساده‌ترین عملیات تحلیلی بر روی داده‌ها و اولین گام در جهت معنی‌دار کردن آن‌ها شامل جمع‌آوری، سازماندهی، گروه‌بندی و به تصویر کشاندن داده‌ها است. در واقع در این نوع تحلیل، یک گزارش‌گیری اولیه از داده‌ها صورت می‌پذیرد. هدف اصلی این تحلیل پاسخ سؤالاتی از قبیل «چه رویدادی رخ داده است؟» و «چه رویدادی دارد در حال رخ دادن است؟» با استفاده از داده‌های گذشته و خلاصه سازی آن‌ها است.

### ۳-۲- تحلیل تشخیصی

در مقایسه با تحلیل توصیفی، تحلیل تشخیصی<sup>۲۴</sup> به جای تمرکز بر آن‌چه رویداده است، به دنبال تشخیص دلیل اصلی رخ دادن آن رویداد است. به طور کلی این تجزیه و تحلیل‌ها به جای نتیجه رویدادها، به فرایندها و علل آن‌ها نگاه می‌کنند. هدف اصلی پاسخ به پرسش‌هایی از قبیل «چرا این رویداد رخ داده است؟» و یا «چرا این رویداد در حال رخ دادن است؟» است. برای پاسخ به این سؤالات، این سطح از تحلیل در مقایسه با تحلیل توصیفی، سطح پیشرفته‌تری را طلب می‌کند که پس از توصیف اولیه داده‌ها، حال به

مشاهده و کشف «روندها»، «پیوندهای میان داده‌ها»، «متغیرها» و «الگوهای پنهان و آشکار درون داده‌ها» می‌پردازد. سپس در این نوع تحلیل، به توضیح و تفسیر این «روندها، الگوها و روابط مشاهده یا کشف شده» اقدام و در نهایت استنباط در خصوص آن‌ها انجام می‌شود. یک جنبه مهم و قابل توجه در مقایسه تحلیل تشخیصی نسبت به تحلیل توصیفی، ایجاد اطلاعات دقیق‌تر از داده‌ها است. به بیان دیگر اطلاعات ایجاد شده در تحلیل توصیفی نیز، در این مرحله مورد استفاده قرار می‌گیرند. بنابر آن‌چه گفته شد به‌طور کلی می‌توان تحلیل تشخیصی را به سه نوع تحلیل جزئی‌تر «تحلیل اکتشافی، تحلیل توضیحی<sup>۲۵</sup>، و تحلیل استنباطی» تقسیم‌بندی کرده است.

### ۳-۲-۱- تحلیل اکتشافی

در این نوع که اولین قدم برای تشخیص علت رخ دادن رویداد بر پایه داده‌ها است ابتدا به داده‌ها و اطلاعات کسب شده از تحلیل توصیفی نگاه دقیق‌تری انداخته می‌شود. بیشترین هدف این تحلیل شناسایی ساختار، متغیرها و پیوندهای پنهان و ناشناخته میان داده‌ها، مانند استقلال و همبستگی، است. در واقع در این مرحله دقیق‌تر مشخص می‌شود که چه رویدادی رخ داده و یا در حال رخ دادن است. تشخیص روندها و الگوهای میان داده‌ها نیز بخشی از همان شناخت ساختار داده‌ها است.

### ۳-۲-۲- تحلیل توضیحی

هدف اصلی این مرحله از تحلیل تشخیصی، توضیح روابط و الگوهای مشاهده و یا شناسایی شده میان داده‌ها در تحلیل اکتشافی است. در واقع حال پس از شناسایی دقیق یک رویداد در مرحله قبل، به چگونگی رخ دادن آن پرداخته می‌شود. در این مرحله است که با توجه به روابط و الگوهای مشاهده و شناسایی شده در تحلیل اکتشافی، فرضیه‌های مناسبی طرح می‌شود. طبیعی است که در این مرحله تحلیل‌گر برای طرح فرضیه‌های

مناسب باید نسبت به داده‌های جمع‌آوری شده از نظر موضوعی و علمی، تخصص و تجربه کافی داشته باشد.

### ۳-۲-۳- تحلیل استنباطی

در این نوع تحلیل با استفاده از روش‌های علمی و به کمک برآورد پارامترها و آزمودن فرضیه‌های مطرح شده در تحلیل توضیحی، به استخراج نتیجه و استنباط نهایی در خصوص چرایی رخ دادن رویدادهای مورد نظر پرداخته می‌شود. به بیان دیگر هدف اصلی این مرحله، یافتن قطعی علت رویداد رخ داده و یا در حال رخ دادن است. همانطور که پیش‌تر نیز گفته شد در بسیاری از متون علمی، معمولاً سه تحلیل اکتشافی، توضیحی و یا استنباطی را در یک عنوان کلی تحلیل تشخیصی و یا به ویژه تحلیل‌های اکتشافی و توضیحی را ادغام و در قالب یک عنوان در نظر گرفته می‌گیرند [۲۲]. با توجه به منابع مختلف در این زمینه تقسیم‌بندی که در بالا گفته شد، در واقع کاملترین تقسیم‌بندی است که می‌توان در این خصوص انجام داد [۲۷] و [۳۰]. البته مرزبندی میان آن‌ها سخت است. به همین دلیل و برای روشن شدن بهتر موضوع به مثال زیر توجه کنید. فرض کنید پس از انجام تحلیل اکتشافی بر روی یک مجموعه داده مشخص شده است که میزان قراردادهای فروش بسته شده توسط نمایندگی‌های یک شرکت در کل کشور، برای استان‌های پر جمعیت بیشتر است. همچنین این میزان با تعداد پرسنل بخش قراردادهای نیز همبستگی بالایی دارد که البته تعداد پرسنل با افزایش جمعیت هر استان، افزایش می‌یابد. حال سؤالی که مطرح می‌شود این است که آیا افزایش میزان قرار دادهای، تنها به علت افزایش جمعیت استان‌ها است، که خود به خود باعث افزایش تعداد پرسنل بخش قراردادهای شده و یا اینکه ممکن است دلیل دیگری نیز وجود داشته باشد. این آغاز مرحله تحلیل توضیحی است. در این مرحله تحلیل‌گر، با شناختی که هم از بعد نظری و هم تجربی نسبت به موضوع دارد، می‌داند که قاعداً باید در مجموع، سرانه پرسنل نسبت به جمعیت استان‌ها مقدار ثابتی باشد، که با یک عملیات ساده بر روی داده‌ها این موضوع

ثابت می‌شود و لذا نباید دلیل افزایش قراردادها، افزایش جمعیت باشد و ممکن است متغیرهای دیگری مانند تجربه و تحصیلات پرسنل، که البته می‌تواند در استان‌های با جمعیت بالا ظهور بیشتری پیدا کنند، مؤثر باشند. پس در این مرحله دو فرضیه اصلی «آیا افزایش جمعیت استان‌ها در افزایش میزان قراردادها نقش دارد؟» و «آیا تجربه و تحصیلات پرسنل بخش قرار داده‌ها در افزایش میزان قراردادها نقش دارد؟» مطرح می‌شوند که در مرحله تحلیل استنباطی با روش‌های علمی آماری مورد آزمون قرار می‌گیرند.

### ۳-۳- تحلیل پیشگویانه

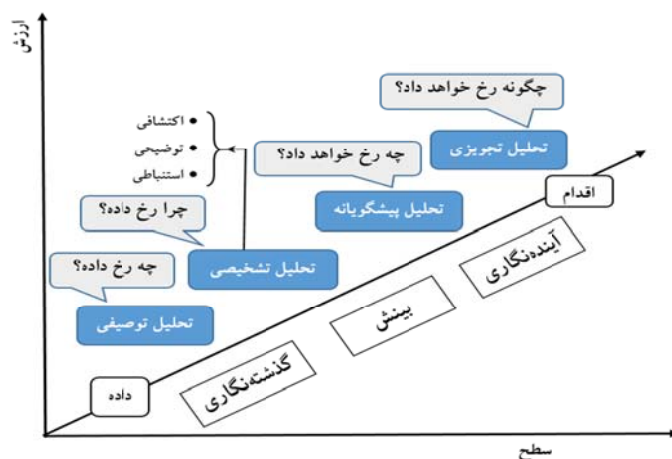
در این تحلیل به پیش‌بینی رخ داده‌های آینده با استفاده از داده‌ها و اطلاعات گذشته و موجود پرداخته می‌شود. هدف اصلی تحلیل پیشگویانه پاسخ به سؤالاتی از قبیل «در آینده چه رویدادی رخ خواهد داد؟». در این مرحله تحلیل‌گر، از تشریح صرف ویژگی‌هایی داده و روابط بین متغیرها (شاخص‌هایی که می‌توانند دارای محدوده‌ای از مقادیر باشند) پا را فراتر نهاده و از داده گذشته به منظور پیش‌گویی و پیش‌بینی<sup>۲۶</sup> آینده استفاده می‌نماید. بنابراین پس از توصیف و کشف ارتباط میان متغیرها و انجام استنباط‌های اولیه در ادامه، احتمال وقوع یک پدیده و یا رویداد را پیش‌بینی می‌کند. روش‌های استنباط پیشرفته آماری، مدل‌سازی و رگرسیون و سری‌های زمانی از روش‌های مرسوم تحلیل پیش‌گویانه هستند.

### ۳-۴- تحلیل تجویزی

اگر چه در تحلیل پیشگویانه از ارتباط کشف شده میان داده‌ها و متغیرها برای پیش‌گویی و پیش‌بینی استفاده می‌شود، اما نمی‌تواند رابطه علت و معلولی را میان داده‌ها نشان دهد و در واقع روابط علی و معلولی نیز نمی‌توانند مبنای تحلیل پیشگویانه قرار داد. به عنوان مثال، در مدل‌سازی میان پاسخ بیمار به یک دارو، اگر چه میان بهبود بیمار و مصرف دارو

رابطه وجود دارد اما هرگز نمی‌توان اثر روانی مصرف کردن دارو را بر بیمار (فارغ از نوع دارو و میزان آن) به تنهایی با یک مدل رگرسیونی یا سایر روش‌های معمول در تحلیل پیش‌گویانه سنجید. مثال دیگر برای توضیح بیشتر موضوع این است که هرگز نمی‌توان با تحلیل پیش‌گویانه بررسی کرد علت نخوردن گوشت در میان کسانی که به پویش اجتماعی مصرف نکردن گوشت پیوسته‌اند، به دلیل پیوستن به پویش است یا این‌که ایشان از ابتدا از خوردن گوشت بیزاز بوده‌اند. در واقع جنس این نوع تحلیل‌ها، «تحلیل علی» و در مراحل پیشرفته‌تر «تحلیل‌های مکانیزی» است که از روش‌های طرح آزمایش و تشکیل گروه‌های مشاهده و کنترل برای کشف روابط علی و معلولی بهره می‌گیرند [۱۶]. هدف اصلی از این نوع تحلیل‌ها پاسخ به این سؤال است که «تغییر یک متغیر چه تأثیری ایجاد می‌کند و یا دقیقاً موجب چه تغییری در سایر متغیرها می‌شود؟» [۳۳] در این مقاله با پیروی از مراجع [۲۲] و [۳۶]، این دو نوع تحلیل را ادغام کرده و فراتر از اهداف این دو تحلیل، آخرین و پیشرفته‌ترین مرحله تحلیل، یعنی تحلیل تجویزی معرفی می‌شود. در این تحلیل هدف پاسخ به این پرسش است که «چگونه و با چه عواملی می‌توان موجب رخ دادن رویدادی شد؟». بنابراین تحلیل تجویزی، شامل روش‌هایی نظیر طراحی تجربی، بهینه‌سازی و حتی موارد بیشتر است. در این تحلیل، همانند نسخه‌ای که یک پزشک تجویز می‌کند، مجموعه‌ای از اقدامات لازم پیشنهاد می‌شود. طراحی تجربی در تلاش است تا به سؤالاتی در خصوص چرایی رخ داده‌ها از طریق برخی آزمایش‌ها پاسخ دهد. بهینه‌سازی، یکی دیگر از روش‌های تحلیل تجویزی است. در این روش، هدف شناسایی سطح ایده‌آل یک متغیر به خصوص در ارتباط با عناصر دیگر است. به‌عنوان نمونه تعیین قیمت آینده یک محصول با در نظر گرفتن میزان سودآوری آن و یا تعیین سطح موجودی کالا در انبار می‌تواند از طریق تحلیل تجویزی رخ دهد.

شکل ۶ همه انواع تحلیل کاربردی داده‌ها، ابعاد و روابط میان آنان را در یک نما نشان می‌دهد.



شکل ۶- انواع تحلیل داده‌های کاربردی و ابعاد آنها

#### ۴- نقش و جایگاه تحلیل داده‌ها در فرایند تولید آمار رسمی

همانطور که در مقدمه این مقاله گفته شد، تحلیل داده‌ها و آمار تولید شده، حداقل برای جلوگیری از تفسیرهای نادرست، از فعالیتهای مهم و ضروری هر سازمان تولیدکننده آمار رسمی، به ویژه در نظام آماری مدرن به شمار می‌رود. البته این تحلیل باید طوری باشد که سایر اصول آمار رسمی به ویژه بی‌طرفی و استقلال حرفه‌ای سازمان آماری را خدشه‌دار نکند. در این بخش از مقاله تلاش می‌شود با مرور دقیق‌تر مستندات و دستورالعمل‌های بین‌المللی در این خصوص، با دقت بیشتری این موضوع بررسی شود. شایان گفتن است که با توجه به تعریف تحلیل داده‌ها در بخش ۲ و دیگر تعاریف ارائه شده در آن بخش، می‌توان در آمار رسمی از عبارت دقیق‌تر تحلیل داده‌ها و اطلاعات آماری نیز به‌عنوان معادل تحلیل داده‌ها استفاده کرد. گسترش چشمگیر و رشد سریع فناوری اطلاعات به ویژه در دو دهه اخیر و پیدایش ابزارهای نرم‌افزاری و سخت‌افزاری باعث شده که انجام تحلیل بر روی داده‌ها و اطلاعات آماری به سرعت و سهولت بیشتری امکان‌پذیر باشد.

نتیجه مطالعه‌ای که توسط یک گروه کارشناسی بخش آمار سازمان ملل در سال ۲۰۰۳ انجام شد، ضمن تصریح بر لزوم وجود گروه‌های سازمان یافته و متمرکز برای انجام تحلیل

آمار تولید شده در سازمان‌های آماری، تحلیل توصیفی و تشخیصی را بهترین نوع تحلیل در این نوع سازمان‌ها می‌داند. در این مطالعه پیشنهاد شده است که تحلیل داده‌های آماری به دست آمده از آمارگیری‌ها، به‌عنوان اطلاعات آماری به همراه نتایج آمارگیری توسط سازمان انجام‌دهنده‌ی آمارگیری منتشر شود [۲۶].

در بیانیه پایانی اولین نشست بین‌المللی داده سازمان ملل، که در سال ۲۰۱۷ در شهر کیپ‌تون آفریقای جنوبی برگزار شده و به‌عنوان سند رسمی در ۴۸ امین نشست کمیسیون آمار نیز با عنوان «برنامه جهانی اقدام عملی کیپ‌تون برای توسعه پایدار داده‌ها» پذیرفته شد، تصریح شده که تحلیل داده‌ها و اطلاعات آماری گردآوری شده، برای رسیدن به اهداف توسعه پایدار در سال ۲۰۳۰، از اقدامات مهم تولیدکنندگان آمار رسمی است. در فهرست اقدامات کلیدی بر شمرده شده در این سند، به تلاش برای تسهیل و ترویج بکارگیری هرچه بیشتر فناوری‌های نوین در گردآوری، پردازش، انتشار و البته تحلیل داده‌ها به‌عنوان یک جریان اصلی در فعالیت‌های آماری<sup>۲۷</sup> اشاره و تأکید شده است [۲۴].

سازمان OECD در گزارش همکاری توسعه سال ۲۰۱۷، با اشاره به نقش و وظایف نظام ملی آمار رسمی در عصر انقلاب داده‌ای<sup>۲۸</sup>، اشاره می‌کند که «قدرت داده‌ها، تنها زمانی پدیدار می‌شود که کاربران آن‌ها را بفهمند و آن هنگامی است که داده‌ها به خوبی توضیح، تفسیر و تحلیل شوند. این درک و فهم ناشی از توضیح روشن داده‌ها که به استفاده هرچه بهتر داده‌ها می‌انجامد، همان تعبیر مطلوب از سواد آماری<sup>۲۹</sup> است» [۶].

اداره آمار کانادا به سازمان‌های تولیدکننده آمار رسمی توصیه می‌کند تا با انجام تحلیل بر روی داده‌های تولید شده، پیش از انتشار نهایی آن‌ها، به ارزیابی کیفیت و در صورت نیاز به تصحیح، تعدیل و یا ادیت داده‌ها بپردازند. اهمیت این امر تا آن‌جا است که موضوع تحلیل داده‌ها به‌عنوان یکی از وظایف مهم اداره آمار کانادا، در قانون آمار این کشور تصریح شده است [۴۰]. نه تنها در قانون آمار کانادا در قانون آماری بسیاری از کشورهای توسعه یافته، تحلیل داده‌ها به‌عنوان یک وظیفه مهم سازمان‌های آماری به ویژه سازمان ملی آمار بر شمرده شده است [۱]. در برنامه ملی راهبردی توسعه آمار رسمی کشورهای



اروپایی از جمله اتریش، برای سال ۲۰۲۰ تحلیل داده‌ها به‌عنوان یکی از ۵ رکن اصلی موضوع این برنامه‌ها، در بخش موضوعی «تبدیل داده به اطلاعات» قرار گرفته است [۵]. در چارچوب کیفیت آمارهای کانادا، تحلیل داده‌ها و اطلاعات آماری به‌عنوان یک ابزار مهم برای ارزیابی شاخص‌های مهم کیفیت در حوزه برون‌داده‌های آماری به صورت نظام‌مند و ساختار یافته استفاده می‌شود [۱۳]. در عین حال نتایج این تحلیل‌ها برای اصلاح و بهبود فرایندهایی مانند «انتخاب منابع داده‌ای و ارزیابی آن‌ها، طراحی یا بازطراحی آمارگیری‌ها، درک درست داده‌ها و اطلاعات آماری تولید شده، برنامه‌ریزی برای انجام فعالیت‌های جدید آماری و سایر فرایندهای اصلی در تولید و انتشار آمارها»، ضروری است.

در راهنمای کیفیت اداره آمار کانادا نیز تأکید می‌شود که در همه گزارش‌ها و انتشارات رسمی محصولات سازمان‌های آماری، باید تحلیل داده‌ها و اطلاعات آماری مربوط نیز وجود داشته باشد. همچنین این راهنما اشاره می‌کند که محصول نهایی برخی آمارگیری‌ها که به دلیل اصل محرمانگی اطلاعات امکان انتشار نتایج آن‌ها وجود ندارد، می‌تواند گزارش تحلیل داده‌ها و اطلاعات آمارگیری‌های مزبور باشد. این راهنما مشخص می‌کند که تحلیل داده‌ها و اطلاعات آماری در یک سازمان تولیدکننده آمار رسمی باید چه شرایط و مراحل داشته باشد [۱۲]. با توجه به اهمیت تحلیل در سازمان ملی آمار، در اداره آمار کانادا یک واحد تخصصی مشخص برای این فعالیت با نام «واحد مطالعات تحلیلی» معین شده است.<sup>۳۰</sup>

نهاد مشارکت برای توسعه آماری در قرن ۲۱<sup>۳۱</sup>، که به پاریس ۲۱ معروف شده، از جمله سازمان‌های مرجع آماری است که توسط بخش «آمار سازمان ملل، کمیسیون اروپا، صندوق بین‌المللی پول و بانک جهانی» برای بهبود فعالیت‌های نظام‌های ملی آمار در سال ۱۹۹۹ تأسیس شده است. این نهاد به ویژه در تهیه و طراحی راهبردهای توسعه نظام آماری و برنامه‌های ملی آمار و تبیین نقش و وظایف سازمان‌های ملی آمار کارآمد در یک نظام آماری فعالیت دارد. در اغلب طراحی‌هایی که با مشارکت این نهاد صورت می‌پذیرد

تحلیل داده‌های آماری از جمله فعالیت‌هایی است که برای یک سازمان ملی آمار، نه تنها برای داده‌های تولید شده آن سازمان بلکه برای همه آمارهای رسمی تولید شده در نظام ملی آمار، به‌عنوان وظیفه تعریف می‌کند. به‌عنوان مثال سازمان ملی آمار تایلند در تغییر ساختار تشکیلات سازمانی که با مشارکت پاریس ۲۱ در سال ۲۰۰۲ انجام داده، به‌طور مشخص یک اداره مستقل برای انجام تحلیل‌های آماری ایجاد کرده است [۲].

یکی از مهمترین و معتبرترین منابعی که در حال حاضر در زمینه ساختار سازمانی و چگونگی عملکرد صحیح یک سازمان ملی آمار رسمی وجود دارد، ویرایش سوم کتابچه راهنمای ساختار و عملکرد سازمان آماری، کمیسیون آماری سازمان ملل است. در این مجموعه تحلیل داده‌ها و اطلاعات آماری به‌عنوان یکی از چند وظیفه مهم و کلیدی سازمان آماری برشمرده شده است. این کتاب راهنما ارایه تحلیل به همراه داده‌ها و اطلاعات آماری تولید شده را مکمل فرایند انتشار در سازمان‌های آماری می‌داند. در واقع این تحلیل‌ها نوعی فراداده برای فهم بیشتر داده‌ها و اطلاعات آماری منتشر شده هستند. البته سازمان‌های تولیدکننده آمار رسمی باید توجه داشته باشند که تحلیل منتشر شده توسط آن‌ها نباید به‌گونه قضاوت و یا ارزیابی عملکرد نهادهای در پیوند با داده‌ها و اطلاعات آماری تولید شده باشد. بلکه ارایه تحلیل‌ها در سازمان‌های آمار رسمی باید به صورت کاملاً بی‌طرفانه و با حفظ اصول استقلال حرفه‌ای، تنها برای درک بهتر کاربران از آمار انتشار یافته و آشنایی با درجه کیفیت و صحت آن‌ها باشد. وجود یک واحد تخصصی در سازمان ملی آمار برای انجام تحلیل‌های تخصصی و ارایه آن همراه با آمار منتشر شده در این کتاب به شدت توصیه می‌شود. این تحلیل‌ها به ویژه برای رسانه‌ها برای اطلاع‌رسانی داده‌ها و اطلاعات آماری و هدایت تحلیل‌های بیشتری که توسط رسانه‌ها و سایر عوامل خارج از سازمان آماری انجام خواهد شد، بسیار سودمند است. ضمن آن که این تحلیل‌ها می‌توانند در داخل فرایند تولید و انتشار آمار قرار گرفته و با بررسی و کنترل‌هایی از قبیل روند داده‌ها و یا سری‌های زمانی، اشتباه‌های پیش از انتشار و حتی پس از انتشار را اصلاح نمایند.

موضوع مدرن‌سازی نظام‌های آمار رسمی یکی از موضوعات اصلی و کلیدی حوزه آمار رسمی و ادبیات علمی آن طی سال‌های اخیر بوده و هست. مدرن‌سازی نظام آماری، مفهومی چند بعدی است که تولید و انتشار آمار و همه عوامل و فعالیت‌هایی را که به طور مستقیم و یا غیر مستقیم (در سطوح ملی، منطقه‌ای و بین‌المللی) در آن پویا هستند را به صورت یک کسب و کار و در قالب «صنعت آمار رسمی» در نظر می‌گیرد. تدوین و ارایه مجموعه‌ای از چارچوب‌های استاندارد کاری، برای همه فعالیت‌های حوزه تولید و انتشار آمار رسمی، از اقدامات مهمی است که در راستای نوین‌سازی آمار رسمی صورت پذیرفته است. یکی از مهمترین این چارچوب‌ها که سلسله‌ای از مراحل و فرایندها را برای تولید و انتشار آمارها معرفی می‌کند، «مدل عمومی فرایند کسب و کار آماری»<sup>۳۲</sup> است. در این چارچوب که بستر اصلی مدرن‌سازی فرایند تولید و انتشار آمارهای رسمی و مدیریت کیفیت آن‌ها را برای مراکز آمار رسمی فراهم می‌سازد، «تحلیل»، یکی از مراحل اصلی و برجسته تعریف شده است و پس از پایان مراحل «طراحی و تولید آمار» و پیش از انتشار داده‌ها و اطلاعات آماری انجام می‌شود. بنابر GSBPM، تحلیل در نظام آماری مدرن شامل اضافه کردن توضیحات و نکات فنی به آمار برای استفاده بهتر و درک بهتر کاربران است. همچنین بررسی دقیق داده‌ها و اطلاعات آماری تولید شده برای اطمینان از مطابقت آن‌ها شاخص‌های کیفیت و نیازهای کاربران پیش از انتشار در این مرحله انجام می‌شود. در مدل مزبور به روشنی تأکید شده که انجام تحلیل با روش‌های مختلف توصیفی و استنباط آماری برای توضیح و تفسیر صحیح از داده‌ها و اطلاعات آماری تولید شده و نیز اطمینان از اعمال شدن روش‌های فنی برای جلوگیری از افشای اطلاعات فردی در هنگام انتشار، یکی از وظایف مهم یک سازمان آماری استاندارد در نظام آماری مدرن است. [۲۱].

هچنین در طبقه‌بندی فعالیت‌های آماری، نسخه ۲۰۰۹، موضع تحلیل داده‌ها در حوزه اصلی چهارم از مجموعه حوزه‌های فعالیت پنج‌گانه این طبقه‌بندی استاندارد بین‌المللی، قرار گرفته است. در توضیحات این طبقه‌بندی اشاره شده که تحلیل یکی از فعالیت‌های

اصلی در زمینه آمار رسمی است که برای ارزیابی کیفیت، توضیح آمارها و روندها، ساختن شاخص‌های ترکیبی و انجام ادیت و ویرایش داده‌ها با اقداماتی مانند «تعدیل‌های فصلی، برون‌یابی»<sup>۳۳</sup> به‌کار می‌رود [۴۵].

بدیعی و همکاران به نقش ارایه تحلیل، همراه با سایر محصولات یک سازمان تولیدکننده آمار، در افزایش سواد آماری جامعه مربوط اشاره کرده‌اند. آن‌ها سازمان ملی آمار را در عصر انفجار اطلاعات، تنها زمانی مؤثر می‌دانند که نه تنها به تولید و انتشار آمار مورد نیاز جامعه بپردازد بلکه بتواند با ارایه تحلیل و توضیح‌های مناسب، درک و فهم مردم را از آمار منتشر شده بالا ببرد. ایشان در مطالعه خود نشان داده‌اند در نظام‌های آماری که تحلیل‌های مناسب همراه با آمار منتشر می‌شود سواد آماری در مقایسه با کشورهای دیگر به صورت بسیار معنی‌داری بیشتر است [۶]. از سوی دیگر با توجه به این‌که در سند اقدام عملی کیپ‌تون برای توسعه پایدار داده‌ها، ارایه و انجام برنامه‌هایی برای افزایش سواد آماری در جامعه، یکی از وظایف مهم سازمان‌های ملی آمار برشمرده شده است، لذا انجام و ارایه تحلیل برای داده‌ها و اطلاعات آماری منتشر شده، اقدامی در راستای انجام این وظیفه مهم است [۲۴].

در آخرین دستورالعمل‌ها و استانداردهای سازمان‌های بین‌المللی مرجع آماری، موضوع تحلیل داده‌ها و اطلاعات آماری به‌عنوان وظیفه آنچنان برای سازمان‌های ملی آمار مهم شده است که توصیه شده که در تدوین قوانین مدرن برای نظام‌های ملی آمار، موضوع تحلیل داده‌ها و اطلاعات آماری توسط تولیدکننده‌های آمار رسمی به صورت یک تکلیف مطرح بیان شود. در الگوی عمومی تدوین قوانین آمار که توسط سازمان ملل در سال ۲۰۱۶ منتشر شده، فرایند تولید آمار رسمی را شامل مراحل گردآوری، پردازش، تحلیل و ذخیره‌سازی داده‌ها را تعریف کرده است. همچنین در تعریف واژه انتشار آمار رسمی مشخصاً در دسترس عموم قرار دادن تحلیل داده‌ها را نیز، جزیی از فرایند انتشار آمار رسمی دانسته است [۴۶]. همچنین در اصلاحیه همین الگو که در سال ۲۰۱۸، توسط

سازمان ملل منتشر شده، بر لزوم شمول تحلیل داده‌ها در قوانین آمار تأکید شده است [۱۵].

همچنین در روزگار انقلاب داده‌ها و پیشرفت چشمگیر و سریع فناوری و در زمانی که مدرن‌سازی کسب و کار آماری مورد توجه و تأکید سازمان‌های مرجع آمار رسمی قرار گرفته، موضوع «مه‌داده‌ها» به‌عنوان یک منبع داده‌ای نوین، نمی‌تواند به‌عنوان یک چالش اصلی در نظام آمار رسمی مطرح نباشد. به همین دلیل است که اخیراً موضوع تحلیل مه‌داده‌ها برای استفاده به‌عنوان یک منبع مهم تولید آمار رسمی، مورد توجه و تأکید سازمان‌های آمار مرجع بین‌المللی قرار گرفته است که این امر خود دلیل دیگری است که مقوله تحلیل در نظام آمار رسمی مدرن روزبه‌روز اهمیت بیشتری می‌یابد [۴۳]. دکتر عبدالرحمن حسن رئیس وقت سازمان ملی آمار کشور مالزی (از سال ۲۰۱۱ تا ۲۰۱۷) با توجه به تغییرات روزافزون جوامع و نیازهای کاربران آمار رسمی، انجام تحلیل را از وظایف اصلی یک سازمان آماری در نظام آماری مدرن دانسته تأکید می‌کند لازم است سازمان‌های ملی آمار باید تمرکز بیشتری روی آن داشته باشند [۲۳].

بدین ترتیب و با توجه به مستندات اشاره شده واضح است که به ویژه در یک نظام آماری مدرن وظیفه مهم و حتی تکلیف قانونی یک تولیدکننده آمار رسمی، انجام و ارزیابی تحلیل داده‌ها و اطلاعات آماری از آمار منتشر شده است.

با این وجود، همانطور که در مقدمه این مقاله نیز اشاره شده است، حفظ بی‌طرفی و استقلال حرفه‌ای نیز از دیگر وظایف مهم و تکالیف قانونی هر سازمان ملی آمار و تولیدکننده آمار رسمی است که هرگز نباید کوچکترین خدشه‌ای به آن وارد شود. بنابراین پرسش این است که با توجه به اصل حفظ بی‌طرفی و استقلال از سویی و لزوم ارزیابی تحلیل از آمار تولید شده توسط تولیدکنندگان آمار رسمی از سوی دیگر، چند و چون و ویژگی‌های تحلیلی که باید توسط مراکز آمار رسمی انجام و ارزیابی شود، چیست و چه ابعاد و مشخصاتی را در بر می‌گیرد؟ برای پاسخ دادن به این پرسش، ضمن مرور و بررسی دقیق دستوالعمل‌ها و توصیه‌های مراجع بین‌المللی در خصوص تحلیل داده‌ها و اطلاعات

آماري در نظام آمار رسمي در ميابيم که تحليل داده‌ها و اطلاعات آماری در یک نظام آماری مدرن باید به طور کلی برای «کاربران» ویژگی‌های عمده زیر را دارا باشد:

- تسهیل در درک و فهم کاربران از داده‌ها و اطلاعات آماری انتشار یافته با ایجاد حس و بینش به ایشان و معنی بخشی به داده‌ها و اطلاعات آماری
- ایجاد شناخت و بینش کافی برای کاربران از روابط میان داده‌ها و اطلاعات آماری انتشار یافته و ساختارهای درونی آنها
- ایجاد اطمینان از صحت و دقت داده‌ها و اطلاعات آماری تولید و منتشر شده و توصیف و توضیح بیشتر آنها
- پاسخ‌گویی به شبهات مطرح شده و تفاسیر اشتباه در خصوص داده‌ها و اطلاعات آماری انتشار یافته
- پرهیز از هرگونه قضاوت و یا ارزیابی در خصوص عملکرد نهادها و واحدهای اجرایی

تحلیل‌هایی که دارای ویژگی‌های فوق باشند ضمن جلوگیری از تفاسیر نادرست از داده‌ها و اطلاعات آماری منتشر شده با تولید فراداده‌های مناسب، به شدت موجب افزایش سرمایه اجتماعی نظام ملی آمار و همه ارکان و اجزای آن خواهند شد. در این راستا تحلیل‌های مورد نظر به‌عنوان مثال باید بتوانند دلایل تغییرهای معنی‌دار در نرخ‌ها و نسبت‌ها را به خوبی تشریح و تبیین نموده و به مقایسه روندها، شبیه‌سازی شرایط و حتی تجربه و وضعیت دیگر کشورها بپردازند. سهم متغیرها در شاخص‌ها، به ویژه شاخص‌های کلیدی را بیان کنند و به توضیح تغییرات اجزا و سهم آنها در محاسبه شاخص و متغیر نهایی بپردازند. تحلیل داده‌ها و اطلاعات آماری باید بتوانند یک شاخص را به اجزای کوچک‌تر شکسته و سهم اجزا از کل و ارتباط آنها با هم را تبیین کند. ارتباط، وابستگی و در عین حال سازگاری شاخص‌ها و متغیرهای مختلف با هم را بررسی کرده و به توضیح آنها از منظر نظری و تئوریک موضوع بپردازند.

همچنین تحلیل‌های مزبور از دید تولیدکنندگان آمار، علاوه بر موارد بالا، باید دارای ویژگی‌های عمده زیر نیز باشند:

- تولید اطلاعات لازم برای انجام ارزیابی‌های کیفیت و ایجاد شناخت کافی برای تأیید محصولات نهایی آماری پیش از انتشار آن‌ها
- ایجاد بینش کافی برای اجرای فرایندهای اصلاح خطاهای انتشار و انجام رویه‌های بازنگری با تولید اطلاعات لازم
- ایجاد بینش کافی برای تهیه و اجرای روش‌های جلوگیری از افشای اطلاعات و حفظ محرمانگی با تولید اطلاعات لازم
- ایجاد اطلاعات لازم برای انجام تعدیل‌ها، ادیت‌ها (به ویژه کلان ادیت‌ها<sup>۳۴</sup>)

با توجه به ویژگی‌های برشمرده شده برای کاربران و تولیدکنندگان و با نگاه به انواع تحلیل که در بخش ۳ این مقاله گفته شد، می‌توان نتیجه گرفت که تحلیلی که مناسب یک سازمان ملی آمار است تحلیلی گذشته‌نگارانه است چرا که تحلیل‌های آینده‌نگارانه ممکن در دل خود به نوعی قضاوت و یا ارزیابی را تداعی کند. بر این اساس تحلیل‌های «توصیفی» و «تشخیصی» می‌توانند ویژگی‌های فوق را شامل شده و انتظارات را برآورده نمایند. در واقع تحلیل مطلوب برای نظام آماری که ضمن حفظ بی‌طرفی و پرهیز از هرگونه قضاوت می‌تواند با توصیف، کشف، توضیح و استنباط در خصوص داده‌ها و اطلاعات آماری با رویکردی گذشته‌نگارانه و با روشن کردن آنچه رخ داده و چرایی آن، با پاسخ به پرسش‌ها و شباهت کاربران، به درک و فهم ایشان کمک کرده موجب ایجاد بینش کافی در ایشان شده و در عین حال آمار و ارقام آرایه شده را معنی و مفهوم بخشد، همان ترکیب تحلیل توصیفی و تشخیصی و استفاده از روش‌ها و فنون آن‌ها است.

البته بدیهی است با توجه به زمینه تحلیل، هر کدام از انواع تحلیل شامل «توصیفی، اکتشافی، توضیحی و استنباطی»، بخشی از آن‌ها یا همه آن‌ها، می‌تواند مورد استفاده قرار گیرد.

## ۵- نتیجه‌گیری و پیشنهادات

جوامع امروزی، جوامعی با آگاهی بالا و در معرض اطلاعات عمومی فراوان هستند. بنابراین آرایه آمارهای کلان به تنهایی نمی‌تواند نیازهای متنوع، افزایش‌یافته و پیچیده‌تر شده این کاربران را، به ویژه در سطوح خردتر، رفع کند. هدف از تولید آمار رسمی ایجاد امکان و بستر مناسب برای تصمیم‌گیری و سیاست‌گذاری درست است و بنابر تعاریف و مراجعی که در این مقاله ارائه شد، تحلیل فرایندی است که این امکان را برای داده‌ها و اطلاعات آماری فراهم می‌سازد تا تبدیل به ابزار مناسب تصمیم‌گیری، سیاست‌گذاری و اقدام مبتنی بر بینش شوند. بنابراین و طبق مطالعاتی که در این مقاله انجام شده، نتیجه‌گیری می‌شود که در یک نظام آماری مدرن، نه تنها تحلیل داده‌ها یکی از فعالیت‌های مهم در فرایند تولید آمار رسمی است بلکه وظیفه قانونی و ضروری سازمان‌های تولیدکننده آمار رسمی، به ویژه سازمان‌های ملی آمار، به‌شمار می‌رود که باید به صورت نظام‌مند و نهادینه در ساختار نظام ملی آمار صورت پذیرد. بنابر نتایج مطالعات این مقاله، تحلیل داده‌ها در یک سازمان ملی آمار و بلکه در یک نظام آماری مدرن را می‌توان به صورت کلی، در یک کلام و به طور خلاصه به دلیل «ایجاد شناخت و بینش لازم در کاربران، تامین‌کنندگان و تولیدکنندگان داده‌ها و اطلاعات آماری»، ضروری دانست. این بینش هم از جنبه گذشته‌نگارانه و هم از جنبه آینده‌نگارانه، در به‌کارگیری و استفاده از داده‌ها تأثیر حیاتی دارد و ارزش افزوده شایان توجهی به داده‌ها و آمار تولید شده اهدا می‌کند. در همین راستا می‌توان ضرورت انجام تحلیل داده‌ها و اطلاعات آماری در نظام آماری و توسط سازمان‌های ملی آمار را به صورت جزئی‌تر در دلایل زیر نیز تعبیر کرد:

- جلوگیری از برداشت‌ها و تفسیرهای نادرست؛
- ایجاد وضوح در داده‌ها و اطلاعات آماری تولید شده برای بکارگیری بهتر از آن‌ها؛



- ایجاد بینش صحیح از داده‌ها و اطلاعات آماری تولید شده در نظام‌های آماری برای ذینفعان (به ویژه کاربران، تأمین‌کنندگان و تولیدکنندگان داده‌ها)؛
- افزایش سواد آماری و ظرفیت‌سازی برای تولید و استفاده از داده‌ها و اطلاعات آماری؛
- ایجاد امکان ارزیابی بهتر کیفیت آمارهای تولید شده؛
- ایجاد امکان ارزیابی بهتر فرایندهای تولید آمار و منابع داده‌ای و فراهم شدن زمینه انجام بهبودها و بازنگری‌های دقیق‌تر و صحیح‌تر؛
- ایجاد امکان برای انجام تعدیل‌ها و ادیت‌های فنی بهتر بر داده‌ها و اطلاعات آماری پیش از انتشار و اصلاح خطاها بعد از انتشار؛
- ایجاد امکان بیشتر برای یافتن راه‌های مناسب جلوگیری از افشای اطلاعات و حفظ محرمانگی.

البته با توجه به لزوم بی‌طرفی و رعایت اخلاق حرفه‌ای توسط سازمان‌های تولیدکننده آمار رسمی به ویژه سازمان‌های ملی آمار که از اصول بنیادین آمارهای رسمی به‌شمار می‌رود، انجام تحلیل بر داده‌ها و اطلاعات آماری توسط این سازمان‌ها باید دارای شرایط و ویژگی‌هایی باشد. با توجه به انواع تحلیل و تعاریفی که در این مقاله بررسی و ارایه شده‌اند، نتیجه‌گیری شده است که تنها تحلیل‌های توصیفی و تشخیصی، از انواع تحلیل می‌توانند در سازمان‌های آماری مورد استفاده قرار گیرند. با توجه به نقشی که این دو نوع تحلیل مذکور و ترکیب آن‌ها در تبیین مفهوم و معنی داده‌ها و اطلاعات آماری ایفا می‌کنند می‌توان ترکیب هر دو را (تحلیل توصیفی و همه انواع تحلیل تشخیصی) تحلیل تبیینی نامید و به ادبیات موضوع اضافه کرد. این نوع تحلیل نه تنها باید در همه گزارش‌ها و نشریات منتشر شده در سازمان‌های ملی آمار به همراه داده‌ها و اطلاعات آماری تولید شده منتشر شود، بلکه خود به‌عنوان یک محصول مستقل (اطلاع آماری) و یک فراداده مهم به شمار رفته و در تقویم انتشار مراکز آماری قرار گیرد. از همین روی سازمان‌های ملی آمار نیاز است تا به همراه ارایه آمار، تحلیل‌های روش‌شناسانه خود را به صورت نظام‌مند و

ساختار یافته ارائه کنند. بنابراین یکی از چالش‌هایی که سازمان‌های تولیدکننده آمار رسمی، به ویژه سازمان‌های ملی آمار، با آن مواجه هستند ایجاد توانایی تحلیل بر پایه روش‌های علمی مانند برآوردیابی، داده‌کاوی و مانند آن‌ها و البته ظرفیت‌سازی برای انجام تحلیل است. برای این کار، همانطور که در این مقاله اشاره شد، مراکز ملی آمار باید ضمن ایجاد واحدهای مشخص ثابت و رسمی در ساختار سازمانی و تشکیلاتی خود، به صورت هدفمند کارشناسان متخصص و آشنا با روش‌های اشاره شده را آموزش و پرورش دهند تا به امر انجام تحلیل، با ویژگی‌ها و به‌گونه گفته شده، مبادرت ورزیده و در انجام تخصصی آن متمرکز باشند. البته بدیهی است وجود یک واحد مشخص با کارشناسان متخصص تحلیل و آشنا به علوم داده، به این معنی نیست که انجام تحلیل بر روی داده‌ها و اطلاعات آماری تولید شده صرفاً توسط ایشان صورت پذیرد، بلکه طبق اصول فعالیت بر مبنای فرایندگرایی، متولی فرایند تحلیل هستند و ضمن تهیه و ارائه تحلیل‌های مزبور به‌عنوان وظیفه اصلی و سازمانی خود، با همکاری سایر کارشناسان واحدهای تولید، به تهیه و ارائه تحلیل‌های تبیینی مورد نظر به صورت حرفه‌ای خواهند پرداخت. برای انجام این فعالیت‌ها به نحو مطلوب و کارآمد، بازطراحی ساختارهای تشکیلاتی سازمان‌های ملی آمار بر اساس مفاهیم و مؤلفه‌های مدرن‌سازی آمار رسمی ضروری است.

تحلیل تبیینی، که در واقع همان ترکیب تحلیل‌های توصیفی و تشخیصی است، با توجه به رویکرد گذشته‌نگارانه خود ضمن پاسخ به شبهات و ابهامات کاربران و جلوگیری از تفاسیر نادرست، با ورود نکردن به روابط علی و معلولی داده‌ها و پیشگویی بر اساس آن‌ها، طبیعتاً منجر به سوگیری و یا قضاوت نخواهد شد. با این وجود و با توجه به نقش و جایگاه داده‌ها و اطلاعات آماری در تصمیم‌گیری‌ها و سیاست‌گذاری‌ها، انجام تحلیل‌های «پیش‌گویانه و تجویزی» که منجر به ایجاد بینش آینده‌نگارانه در کاربران و سایر ذینفعان نظام آمار رسمی می‌شود، نیز نباید به طور کلی از نظر مراکز آماری دور بماند. انجام این نوع تحلیل‌ها می‌تواند در مراکز علمی و پژوهشی به صورت مستقل، اما با مشارکت و همکاری سازمان‌های آماری به صورت هدفمند انجام شده و نتایج آن مورد استفاده

سازمان‌های ملی آمار نیز قرار گیرد. شایان گفتن است این نوع تحلیل‌ها در هر حال توسط مراکز پژوهشی و پژوهشگران انجام خواهد شد. آنچه بر اساس نتایج این مطالعه می‌توان پیشنهاد داد، هدایت و راهبری این نوع تحلیل‌ها (تحلیل‌های پیشگویانه و تجویزی) توسط سازمان‌های ملی آمار است تا با مدل‌ها و پیش‌بینی‌هایی که انجام می‌شود، ضمن کمک فراوان به کاربران آمار رسمی در سطوح کلان تصمیم‌گیری، نتایج تحلیل‌ها بدون این‌که بی‌طرفی و استقلال حرفه‌ای مراکز آمار رسمی را خدشه‌دار کرده و آن‌ها را از وظایف اصلی خود بازدارد، آن‌ها را در فرایند تولید آمار، ارزیابی کیفیت محصولات، بازنگری‌های فرایندی و تعیین و به‌روزرسانی منابع داده‌ای یاری رساند.

### توضیحات

1. Making Data Meaningful
2. Conference of European Statisticians (CES)
3. The United Nations Economic Commission for Europe (UNECE)
4. Organization for Economic Co-operation and Development (OECD)
5. clarity
۶. برابر نهاده داده‌ها DATA است و مفرد آن DATUM است که واژه‌ای لاتین است. با این وجود معمولاً در ادبیات آماری DATA هم به مفهوم جمع یعنی داده‌ها و هم مفرد یعنی داده، استفاده می‌شود.
7. Fact
8. Statistical data
9. Knowledge
10. Wisdom
11. Understanding
12. Insight
13. Understanding relations and patterns and principles
14. Hindsight
15. Foresight
16. Retrospective

17. Forecasting
18. Inspecting, Cleansing, Transforming & Modeling data
19. Quantitative and Qualitative Analysis
20. Basic (Fundamental) and Applied analysis
21. Theoretical Researches
22. Descriptive, Exploratory, Inferential, Predictive, Casual, and Mechanistic Analysis
23. Descriptive, Predictive and Perspective Analysis
24. Diagnostic Analysis
25. Explanatory analysis
26. Diagnostic Analysis
27. Explanatory analysis
28. Mainstream statistical activities
29. Data revolution
30. Statistical literacy
31. Analytical Studies, Methodology and Statistical Infrastructure  
Field\Methodology Branch\Business Survey Methods Division\Time Series Research and Analysis Centre
32. The Partnership in Statistics for Development in the 21st Century (PARIS21)
33. The Generic Statistical Business Process Model (GSBPM)
34. Extrapolation
35. Macro Edit

### مرجع‌ها

- [1] ABS, The Australian Bureau of Statistics Act. (1975). The Australian Bureau of Statistics, Canberra.

- [2] ADB/PARIS21 (2002). Joint High Level Forum. Country Paper; Statistical Activities in Thailand. in Statistical Capacity Building for ASEAN Countries. EDSA Shangri-la Hote
- [3] Alt-Simmons, R. (2015). Agile by Design: An Implementation Guide to Analytic Lifecycle Management. John Wiley & Sons.
- [4] Attride-Stirling, J. (2001). Thematic networks: an analytic tool for qualitative research. *Qualitative research*, **1**, 385-405.
- [5] Austria, S.(2019). STRATEGY 2020, Vienna.
- [6] Badiee, S., Jütting, J., Appel, D., Klein, T. and Swanson, E. (2017). The role of national statistical systems in the data revolution, in Development Co-operation Report 2017: Data for Development. OECD Publishing, Paris.
- [7] Baesens, B. (2014). Analytics in a big data world: The essential guide to data science and its applications. John Wiley & Sons
- [8] Baskarada, S., and Koronios, A. (2013). Data, information, knowledge, wisdom (DIKW): a semiotic theoretical and empirical exploration of the hierarchy and its quality dimension. *Australasian Journal of Information Systems*, **18**.
- [9] Bekker, A. (2019). (4) Types of Data Analytics to Improve Decision-Making. @ScienceSoft, Inc.: Dallas, USA.
- [10] Bellinger, G., Castro, D., and Mills, A. (2006). Data, information, knowledge, and wisdom (2004). Available at: [www. systems-thinking. org/dikw/dikw. htm](http://www.systems-thinking.org/dikw/dikw.htm) (accessed: 5 February 2006)
- [11] Bihani, P. and Patil, S. (2014). A comparative study of data analysis techniques. *International journal of emerging trends & technology in computer science*, **3**, 95-101.

- [12] Canada, S., Statistics Canada quality guidelines (2009). Authority of the Minister responsible for Statistics Canada, © Minister of Indus: Ottawa.
- [13] Canada, S., Statistics Canada's Quality Assurance Framework (2017). Authority of the Minister responsible for Statistics Canada, © Minister of Industry.
- [14] Canada, S., Statistics: Power from Data (2011). Authority of the Minister responsible for Statistics Canada, © Minister of Industry: Ottawa.
- [15] CES, Guidance on common elements of statistical legislation (2018). United Nations Economic Commission for Europe (UNECE): Geneva.
- [16] Constantine, N.A. (2012). Regression analysis and causal inference: cause for concern?. Perspectives on sexual and reproductive health, 134-137.
- [17] Davenport, T.H., and Kim, J. (2013). Keeping up with the quants: Your guide to understanding and using analytics. Harvard Business Review Press.
- [18] Demchenko, Y., Grosso, P., De Laat, C., and Membrey, P. (2013). Addressing big data issues in scientific data infrastructure. In 2013 International Conference on Collaboration Technologies and Systems (CTS). IEEE.
- [19] Dodge, Y. (2008). The concise encyclopedia of statistics. Springer Science & Business Media.
- [20] Floridi, L. (2005). Is semantic information meaningful data?. Philosophy and phenomenological research, 70, 351-370.

- [21] Generic Statistical Business Process Model (GSBPM), Version 5.1, (2019). HLG-MOS, Conference of European Statisticians, UNECE: Geneva.
- [22] Hagerty, J. (2017). Planning Guide for Data and Analytics. Gartner. Published: 13 October 2016. Gartner, Inc: Stamford, CT USA.
- [23] Hasan, A.R. (2012). Challenges Faced by the National Statistics Office Department of Statistics, Malaysia, in The 13th East Asian Statistical Conference. Tokyo, Japan.
- [24] High-level Group for Partnership Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development, Cape Town Global Action Plan for Sustainable Development Data (2017). First UN World Data Forum: Cape Town, South Africa.
- [25] Hjørland, B. (2017). Encyclopedia of knowledge organization. Knowledge Organization, **44**, 436-464.
- [26] Jambwa, M.M. (2003). Data processing, analysis, and dissemination. in Expert Group Meeting to Review the Draft Handbook on Designing of Household Sample Surveys. United Nations Statistics Division, New York.
- [27] Julia Kokina, C.P.A. and Stephen Kozlowski, C.P.A. (2016). The next frontier in data analytics. Journal of Accountancy, **222**, 58.
- [28] Khatri, V. and Brown, C.V. (2010). Designing data governance. Communications of the ACM, **53**, 148-152.
- [29] Kiregyera, B. (2015). The Emerging Data Revolution in Africa, Strengthening the Statistics, Policy and Decision-making Chain. SUN MeDIA Stellenbosch @ SUN PRESS.

- [30] Knaflic, C.N. (2015). *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons.
- [31] Kotz, S., Read, C.B., Balakrishnan, N. and Vidakovic, B. (2005). *Encyclopedia of Statistical Sciences, 16 Volume Set, 2nd Edition*. Vol. 3. New York, John Wiley & Sons. 1531.
- [32] LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., and Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT sloan management review*, **52**, 21-32.
- [33] Leek, J. (2013). Six types of analyses every data science should know. *Data scientist Insights blog*.
- [34] MacKay, R.B., and McKiernan, P. (2004). The role of hindsight in foresight: refining strategic reasoning. *Futures*, **36**, 161-179.
- [35] Madden, A.D. (2000). A definition of information. In *Aslib Proceedings*. MCB UP Ltd.
- [36] Maydon, T. (2017). *The 4 Types Of Data Analytics*. 2017, @Principa, Inc, Transaction Capital Group. South Africa, Cape Town
- [37] OECD. Publishing (2007). *OECD glossary of statistical terms*. Organisation for Economic Co-operation and Development.
- [38] Pantzar, E. (2000). Knowledge and wisdom in the information society. *Foresight*, **2**, 230-236.
- [39] Press, O.U. (2016). *Oxford Advanced Learner's Dictionary*. Obtenido de <http://www.oxfordlearnersdictionaries.com/us>.
- [40] R.S.C. 1985 c. S-19, Statistics Act, in Current to June 20, 2019, Last amended on December 12, 2017. 2017, the Minister of Justice Canada.
- [41] Roese, N.J., and Vohs, K.D. (2012). Hindsight bias. *Perspectives on psychological science*, **7**, 411-426.



- [42] Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*, **33**, 163–180.
- [43] Tam, S. M., and Clarke, F. (2015). Big data, official statistics and some initiatives by the Australian Bureau of Statistics. *International Statistical Review*, **83**, 436–448.
- [44] Thomson, R., and Holland, J. (2003). Hindsight, foresight and insight: The challenges of longitudinal qualitative research. *International Journal of Social Research Methodology*, **6**, 233–244.
- [45] UNECE, Classification of Statistical Activities (2009). United Nations Economic Commission for Europe (UNECE), Geneva.
- [46] UNECE, Generic Law on Official Statistics, for Eastern Europe, Caucasus and Central Asia. (2016), United Nations Publication, Geneva.
- [47] UNSD, United Nations Statistical Division (2015). United Nations Fundamental Principles of Official Statistics, Implementation guidelines, United Nations Statistics Division, New York.
- [48] UNSD, United Nations Statistical Division (2014). Fundamental Principles of Official Statistics U.N.S. Division, Economic and Social Council, New York.
- [49] Whitney, H. (2012). Data insights: new ways to visualize and make sense of data Newness.
- [50] Xia, B.S., and Gong, P. (2014). Review of business intelligence through data analysis. *Benchmarking: An International Journal*, **21**, 300–311.

## اشکان شباک

دکترای آمار

تهران، خیابان فاطمی، خیابان باباطاهر، خیابان سرتیپ فکوری، ساختمان ۱۴۵، پژوهشکده‌ی آمار  
رایانشانی: a.shabbak@gmail.com

## جواد حسین زاده

دکترای اقتصاد

تهران، خیابان فاطمی، خیابان رهی معیری، شماره ۱، مرکز آمار ایران  
رایانشانی: j\_hoseinzadeh@sci.org.ir

## محمد شیری

دکترای جمعیت‌شناسی

تهران، خیابان فاطمی، خیابان باباطاهر، خیابان سرتیپ فکوری، ساختمان ۱۴۵، پژوهشکده‌ی آمار  
رایانشانی: mshiri@src.ac.ir