

فصل اول - درس سوم چرخه‌ی آمار در حل مسائل

اهداف درس سوم

۱. یادآوری شاخص‌ها، نمودارها، اصطلاحات و مفاهیم آماری پایه‌های پیشین
۲. یادآوری و درک منطق انتخاب شاخص‌ها و نمودارهای آماری
۳. آشنایی با چرخه‌ی حل مسائل آماری به عنوان الگویی برای آماری فکر کردن
۴. درک روش به دست آوردن نتایج آماری و پرسیدن پرسش‌های تعیین‌کننده با توجه به بستر و زمینه‌ی مسئله
۵. پرورش نگاه نقادانه به مسائل و نتایج آماری و ارائه‌ی پیشنهاد مناسب

روش تدریس درس سوم

♦ در این درس می‌بایست تمرکز بر روی تقویت توانایی دانش‌آموزان در برقراری ارتباط بین مسئله و بستر آن در دنیای واقعی باشد. این مهم با تاکید کمتر بر روی محاسبات و ایجاد فضا و مجال بحث و گفتگو در کلاس میسر خواهد شد.

درس سوم با بیان مسئله‌ای آغاز می‌شود که تمامی گام‌های چرخه‌ی حل مسائل آماری در آن نهفته است. در واقع دانش‌آموزان در فعالیت صفحه‌ی ۲۸ تا ۳۰، پیش از معرفی چرخه‌ی حل مسائل آماری، گام‌های مختلف چرخه‌ی حل مسائل آماری را در یک مثال مشاهده می‌کنند. در این فعالیت زیر واژگان کلیدی هر گام خطی کشیده شده تا نشان داده شود در هر گام چه مواردی بیشتر مورد توجه است.

در صفحه‌ی ۲۹ در پرسش مربوط به گام دوم (طرح و برنامه‌ریزی) تاکید می‌شود که در هر مسئله واحد اندازه‌گیری متغیر مورد بررسی باید مشخص باشد. در پرسش‌نامه‌ی پرسش الف. صفحه ۲۸، اگر «واحد اندازه‌گیری مدت استفاده از تلفن هوشمند (دقیقه)» مشخص نباشد ممکن است پاسخ‌هایی با واحد ساعت یا پاسخ‌هایی کیفی (کم، متوسط، زیاد) دریافت کنیم که تحلیل داده‌ها را مشکل می‌کند.

در گام سوم این فعالیت در صفحه‌ی ۲۹ به حفظ محرمانگی اطلاعات شرکت‌کنندگان، به‌عنوان راهی برای گردآوری اطلاعات دقیق‌تر و حفظ حریم خصوصی افراد، اشاره می‌شود و دانش‌آموزان به این نتیجه می‌رسند

که اگر پرسش‌نامه‌ها بی‌نام باشد شرکت‌کنندگان اطمینان خاطر خواهند داشت که اطلاعاتشان محرمانه می‌ماند و والدین و کادر آموزشی و افرادی که داده‌ها را گردآوری می‌کنند از آن باخبر نخواهند شد. در این راستا به شرکت‌کنندگان توضیح داده می‌شود که اطلاعات گردآوری شده (مثلاً مدت زمان استفاده از تلفن هوشمند) به صورت کلی در قالب شاخص‌های آماری (مثلاً میانگین یا میانه و انحراف استاندارد یا دامنه‌ی میان چارکی) گزارش می‌شود و پاسخ تک‌تک افراد فاش نمی‌شود.

در ادامه در صفحه‌ی ۲۹ در گام چهارم (تحلیل داده‌ها)، به گزارش نتایج و یادآوری مفاهیمی که دانش‌آموزان پیش‌تر آموخته‌اند پرداخته می‌شود. در این فعالیت با توجه به مفهوم چارک‌ها، حدود ۲۵ درصد دانش‌آموزان بیش از ۱۰۰ دقیقه در روز با تلفن هوشمند کار می‌کنند. حضور در شبکه‌های اجتماعی بیشترین فراوانی و خواندن کتاب الکترونیک کمترین فراوانی را در میان فعالیت‌ها داشتند.

در این فعالیت صفحه‌ی ۲۹، چون شرکت‌کنندگان می‌توانند بیش از یک فعالیت را علامت بزنند، مجموع درصد فراوانی‌های نسبی بیش از ۱۰۰ درصد می‌شود. در این قسمت بهتر است توضیح داده شود که نمودار دایره‌ای گزینه‌ی مناسبی برای گزارش این نتایج نیست؛ چون در نمودار دایره‌ای مجموع زوایایی که به هر قسمت تخصیص داده می‌شود باید ۳۶۰ درجه شود.

در گام آخر (بحث و نتیجه‌گیری) صفحه‌ی ۳۰، به تعمیم‌پذیری نتایج پرداخته می‌شود و دانش‌آموزان می‌آموزند که نتایج حاصل فقط در مورد دانش‌آموزان پایه‌های دهم، یازدهم و دوازدهم همان مدرسه صادق است و نمی‌توانیم نتایج آن مدرسه را به مدارس دیگر تعمیم دهیم؛ چراکه از مدارس دیگر نمونه‌گیری نکرده‌ایم و اطلاعی در دست نداریم. همچنین این نتایج فقط در مورد میزان و چگونگی استفاده از تلفن هوشمند صادق است و قابل تعمیم به میزان و چگونگی استفاده از ابزارهایی مانند رایانه و تبلت و ... نیست.

در این گام همچنین به تفسیر نتایج می‌پردازیم و به این نتیجه می‌رسیم که حضور در شبکه‌های اجتماعی، تماشای فیلم یا گوش دادن به موسیقی و همچنین عکاسی بیشترین فعالیت‌هایی هستند که دانش‌آموزان با تلفن هوشمند خود انجام می‌دهند و نیمی از دانش‌آموزان بیش از ۷۰ دقیقه در روز با تلفن هوشمند خود کار می‌کنند و حدوداً ۲۵ درصد دانش‌آموزان کمتر از ۶۵ دقیقه و حدوداً ۲۵ درصد بیش از ۱۰۰ دقیقه با تلفن هوشمند خود کار می‌کنند. (در این قسمت به علت اینکه در حالت کلی اندازه‌ی نمونه (مثلاً زوج یا فرد بودن تعداد دانش‌آموزان مورد مطالعه در این فعالیت) می‌تواند بر تعداد و درصد دقیق داده‌هایی که بین چارک‌ها قرار می‌گیرند تاثیر بگذارد از کلمه‌ی «تقریباً» و «حدوداً» استفاده شده است.)

در کار در کلاس صفحه‌ی ۳۱ دانش‌آموزان به طور دقیق‌تر با اهداف گام‌های چرخه‌ی حل مسائل آماری آشنا می‌شوند.

ترتیب گام	نام گام	توضیح گام
۲	طرح و برنامه‌ریزی	راهی برای رسیدن به پاسخ مسئله پیدا می‌کنیم. به نمونه‌گیری، شیوه اندازه‌گیری متغیر و چگونگی توصیف نتایج می‌اندیشیم.
۵	بحث و نتیجه‌گیری	به تفسیر نتایج به‌دست‌آمده می‌پردازیم و پاسخی برای پرسش اصلی پیدا می‌کنیم.
۱	بیان مسئله	مسئله‌ای را که در دنیای واقعی وجود دارد به صورت یک مسئله‌ی شفاف و دقیق آماری مطرح می‌کنیم.
۴	تحلیل داده‌ها	با استفاده از شاخص‌ها، نمودارها و مفاهیمی که آموخته‌ایم نتایج را متناسب با هدف‌های کارمان، نوع متغیرها و ویژگی‌های داده‌ها گزارش می‌کنیم.
۳	گردآوری داده‌ها	به گردآوری داده‌ها می‌پردازیم و تا حد ممکن از درستی آن‌ها مطمئن می‌شویم.

از صفحه‌ی ۳۱ به بعد دانش‌آموزان به صورت عمیق‌تر با هر گام چرخه آشنا می‌شوند. در تمامی گام‌ها ایجاد فضای گفتگو برای دانش‌آموزان بسیار اهمیت دارد. در گام اول (بیان مسئله) دانش‌آموزان تمرین می‌کنند تعرف دقیق و شفافی از مسئله داشته باشند و در مورد یک جامعه‌ی آماری محدود با متغیرهای مشخص مسئله را تعریف کنند. مثلاً در بخش «بحث و گفتگو» صفحه‌ی ۳۱ ممکن است دانش‌آموزان با بیان ارتباطی منطقی بگویند که نزدیکی به پارک‌ها یا ورزشگاه‌ها، جنسیت افراد و محدودیت اجتماعی برای انجام برخی ورزش‌ها، میزان آلودگی هوا در مکان مورد بررسی یا در فصل مورد بررسی می‌تواند بر میزان فعالیت بدنی تاثیر بگذارد؛ بدین ترتیب آن‌ها می‌توانند جامعه‌ی آماری مورد بررسی را محدود کنند. مثلاً مسئله‌ی دیگری را این‌گونه بیان کنند: «بررسی میزان فعالیت بدنی روزانه‌ی پسران پیش‌دبستانی دو مدرسه‌ی مشخص در روزهای تعطیل هفته در دی ماه». در بخش «بحث و گفتگو» هر نوع ارتباط منطقی که از نظر ایشان قابل توجیه باشد می‌بایست شنیده و تشویق شود.

یکی دیگر از نکات مهم برای رسیدن به پاسخ مناسب مسائل به کمک آمار این است که کارها به صورت گروهی انجام شود و از نظرات متخصصان استفاده شود. در مثال «بررسی میزان فعالیت بدنی» انتظار می‌رود دانش‌آموزان به مثال‌هایی مانند افراد زیر برای حضور در گروه اجراکننده اشاره کنند:

- متخصصان تربیت بدنی (به منظور طراحی پرسش‌نامه یا ابزاری برای اندازه‌گیری فعالیت بدنی)

- والدین (برای کمک به طراحی پرسش‌نامه یا ابزار اندازه‌گیری با توجه به شناخت آن‌ها از فعالیت فرزندان) (

- معلم‌های ورزش (چنانچه هدف اندازه‌گیری فعالیت بدنی در مدرسه باشد)

- کار درمان یا مددکار (چنانچه جامعه‌ی آماری مورد بررسی شامل افرادی با کم‌توانی جسمی باشد)

در گام «طرح و برنامه‌ریزی» دانش‌آموزان یاد می‌گیرند که به چگونگی اندازه‌گیری، نمونه‌گیری و گزارش نتایج توجه کنند و با نگاهی نقادانه این موارد را بررسی کنند. آن‌ها در این بخش می‌آموزند که هرچه پراکندگی متغیر مورد بررسی در جامعه بیشتر باشد به اندازه بزرگتری نیاز داریم. به همین دلیل در کار در کلاس صفحه‌ی ۳۳ پرسش ۱ قسمت الف. چون معدل دانش‌آموزان یک کلاس پراکنده‌تر از سن دانش‌آموزان یک کلاس است (در یک کلاس همه تقریباً هم‌سن هستند) برای بررسی معدل، اندازه‌ی نمونه‌ی بزرگتری نیاز داریم. به همین ترتیب در قسمت ب. چون حقوق کارکنان یک شرکت بزرگ خصوصی متنوع‌تر و پراکنده‌تر از حقوق معلم‌های یک مدرسه‌ی دولتی است برای بررسی حقوق کارکنان این شرکت به اندازه‌ی نمونه‌ی بزرگ‌تری احتیاج داریم.

در پرسش ۲ قسمت الف. اشکال پیشنهاد علی، قسمت ب. اشکال پیشنهاد امید و قسمت پ. نیز اشکال پیشنهاد علی است.

دو اشکال دیگر در مورد پیشنهاد علی می‌تواند موارد زیر باشد:

(۱) از همه‌ی افراد مورد بررسی باید همه‌ی پرسش‌ها پرسیده شود (مثلاً از کسی که در کتابخانه است نیز باید پرسیم چقدر به مسجد و تاتر و .. می‌رود).

(۲) باید به صورت تصادفی دانش‌آموزانی را انتخاب کنیم (نه فقط در مکان‌هایی که به ذهنمان می‌رسد).

در پرسش ۳ انتخاب افراد بر اساس رقم اول شماره تلفن منازل می‌تواند به انتخاب محله‌هایی بیانجامد که متعلق به طبقه‌ی اجتماعی-اقتصادی بالا یا پایینی هستند و بدین ترتیب تنوع خانوارها در نمونه کم می‌شود و تمامی قشرهای جامعه شانس حضور در نمونه نخواهند داشت.

گام سوم گردآوری و پاک‌سازی داده‌ها است. در این گام همچنین به سازمان‌دهی داده‌ها می‌پردازیم. یعنی تبدیل واحدهای اندازه‌گیری یا محاسبه‌ی یک متغیر جدید بر اساس اطلاعات گردآوری شده (مثلاً محاسبه‌ی شاخص توده‌ی بدنی با استفاده از قد و وزن) در این گام انجام می‌شود.

کار در کلاس صفحه‌ی ۳۴ به نکاتی در مورد پاک‌سازی داده‌ها می‌پردازد. در ادامه به پاسخ به پرسش‌های این کار در کلاس می‌پردازیم:

الف) در متغیر سن، کدام داده احتمالاً اشتباه تایپ شده است؟ چگونه می‌توانستیم با مرتب کردن داده‌ها در نرم‌افزار این داده را از میان تعداد بسیار زیادی سطر پیدا کنیم؟ ۵۱. اگر داده‌ها را با کمک نرم‌افزار به صورت نزولی (از بزرگ به کوچک) مرتب می‌کردیم ۵۱ در بالاترین سطر قرار می‌گرفت. (اگر یک بار هم داده‌ها را به صورت صعودی کنیم داده‌های بسیار کوچک به سرعت پیدا می‌شوند.)

ب) در داده‌های مربوط به کدام دانش آموز به دلیل جا افتادن کد جنسیت، داده‌های متغیرهای بعدی (ورزش مورد علاقه، متوسط ساعات مطالعه‌ی آزاد و نمره‌ی ادبیات) به سمت چپ جابه‌جا شده‌اند؟ ورزش مورد علاقه‌ی این دانش آموز چیست؛ نمره‌ی ادبیات او چند است؟ دانش‌آموز شماره‌ی ۱۹ (سطر ۷). بسکتبال. ۱۹.

پ) در متغیر «ورزش مورد علاقه» کدام عدد به اشتباه تایپ شده است؟ عدد ۲۲ برای دانش‌آموز شماره‌ی ۴۵.

ت) در متغیر «متوسط ساعات‌های مطالعه‌ی آزاد در هفته» کدام عدد با الگوی داده‌های دیگر تفاوت دارد؟ آیا می‌توانیم این عدد را بدون بررسی حذف یا اصلاح کنیم؟ ۴۴. خیر؛ داده‌های دورافتاده لزوماً داده‌های اشتباه نیستند و صحت آن‌ها باید مورد بررسی قرار گیرد. (مثلاً در مورد بررسی حقوق شهروندان یک شهر ممکن است چند داده‌ی دورافتاده بزرگ (نجومی) مشاهده شود. اما این داده‌ها لزوماً اشتباه نیستند و نباید کنار گذاشته شوند.)

ث) در متغیر «نمره‌ی ادبیات» داده‌های مربوط به کدام دانش‌آموز با واحد اندازه‌گیری متفاوتی ثبت شده است؟ چگونه می‌توانستیم از این اتفاق جلوگیری کنیم؟ نمره‌ی دانش‌آموز شماره‌ی ۴۵ به صورت کیفی ثبت شده است. پیش از گردآوری داده‌ها لازم بود تاکید شود که تمامی دبیران باید آزمون ۲۰ نمره‌ای برگزار کنند.

در بخش بحث و گفتگو صفحه‌ی ۳۵ پیشنهادهای دانش‌آموزان در مورد اینکه اگر دانش‌آموزی در روز برگزاری آزمون ادبیات غایب باشد مورد بحث قرار می‌گیرد؛ یک پیشنهاد می‌تواند این باشد که میانگین نمره‌های آزمون‌های قبلی او را به جای نمره‌ی او بگذاریم. در اینجا می‌توان درباره‌ی صفر گذاشتن برای این دانش‌آموز و تاثیر آن بر میانگین کلاس، حذف کردن تمامی اطلاعات این دانش‌آموز و یا قرار دادن میانگین کلاس به جای میانگین این دانش‌آموز بحث کرد.

در ادامه‌ی بخش بحث و گفتگوی صفحه‌ی ۳۵ لازم است تاکید شود که هرگز نمی‌توانیم اطمینان کامل داشته باشیم که داده‌ها به طور کامل پاک‌سازی و اصلاح شده‌اند. مثلاً ممکن است سن یک فرد به جای ۱۴، ۱۳ ثبت شده باشد. اما از آنجا که ۱۳ نیز در دامنه‌ی داده‌های مورد بررسی قرار دارد با صعودی و نزولی کردن داده‌های این متغیر نمی‌توانیم به این اشتباه پی ببریم.

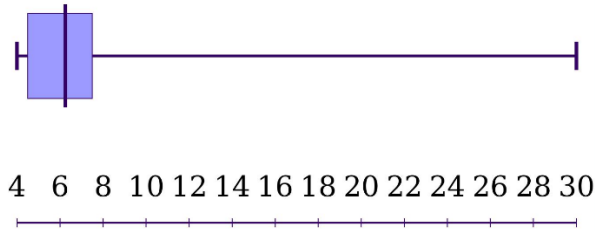
یکی از اهداف مهم طراحی این کار در کلاس این است که دانش‌آموزان یاد بگیرند در بسیاری از نرم‌افزارها «هر سطر شامل اطلاعات یک واحد آماری و هر ستون نشان‌دهنده‌ی یک متغیر است»

در گام چهارم (تحلیل داده‌ها) به گزارش نتایج با معیارها و نمودارهای مناسب پرداخته می‌شود و یادآوری می‌شود که میانگین معیار گرایش به مرکز مناسبی برای داده‌هایی که مقادیر دورافتاده دارند نیست. چرا که این معیار تحت تاثیر داده‌های دورافتاده قرار می‌گیرد (تاکید بر اینکه میانگین مرکز ثقل داده‌ها است و به سمت داده‌ای دورافتاده می‌رود). در این موارد انحراف استاندارد (انحراف معیار) نیز معیار مناسبی نیست چرا که با استفاده از واریانس به دست می‌آید که واریانس نیز از جنس میانگین است (میانگین مجذور فاصله‌ها تا میانگین).

در کار در کلاس صفحه‌ی ۳۵ به این مطلب پرداخته شده است.

متغیر	داده‌ها	معیارهای گرایش به مرکزی		معیارهای پراکندگی		
		\bar{x}	s_x	IQR	SD	دامنه
سن (سال)	۱۷ ۱۸ ۱۶ ۱۷ ۱۵ ۱۹ ۱۵ ۱۶ ۱۷ ۱۸ ۱۵ ۱۹	۱۶/۸۳	۱۷	$18-15/5=2/5$	۱/۳۴	$19-15=4$
میزان تماشای تلویزیون	۴ ۲۹ ۴ ۶ ۷ ۶/۵ ۷ ۸ ۶ ۳۰ ۴ ۵	۹/۷۱	۶/۲۵	$7/5-4/5=3$	۸/۵۵	$30-4=26$

با توجه به توزیع داده‌ها برای متغیر «سن» میانگین و انحراف استاندارد (انحراف معیار) معیارهای مناسبی هستند چون داده‌ی دورافتاده مشاهده نمی‌کنیم؛ اما برای «میزان تماشای تلویزیون» میانه و دامنه‌ی میان چارکی معیارهای گرایش به مرکز و پراکندگی مناسبی هستند؛ زیرا چند داده‌ی دورافتاده وجود دارد. داده‌های ۲۹ و ۳۰ فاصله‌ی زیادی از میانگین دارند و مجذور این فاصله‌ها باعث بزرگ شدن واریانس و به تبع آن بزرگ شدن انحراف استاندارد (انحراف معیار) شده است. در واقع میانگین (۹/۷۱) نماینده‌ی مناسبی برای داده‌های «میزان تماشای تلویزیون» نیست و به سمت داده‌های دورافتاده رفته است. در قسمت ث. کارد در کلاس صفحه‌ی ۳۶ اشاره می‌شود که اگر میانگین و انحراف استاندارد معیارهای مناسبی برای توصیف داده‌ها باشند می‌توان از نموداری که با این دو معیار رسم می‌شود داده‌ها را خلاصه و توصیف کرد. در این کار در کلاس فقط برای متغیر سن می‌توان از این نمودار استفاده کرد که مقادیر شکل الف منطبق با داده‌های متغیر سن است. در این کار در کلاس همچنین تاکید می‌شود که اگر داده‌ی دورافتاده داشته باشیم نمودار حاصل از میانگین و انحراف استاندارد (انحراف معیار) نمودار مناسبی نیست و به همین دلیل است که برای متغیر کمی مطمئن‌ترین نمودار، نمودار جعبه‌ای است.

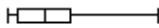


نمودار جعبه‌ای به خوبی نشان می‌دهد که داده‌ها در کدام قسمت پراکنده‌تر و کجا فشرده‌ترند. در نمودار جعبه‌ای میزان تماشای تلویزیون ۷۵ درصد داده‌ها در بازه‌ی ۴ تا ۷/۵ به صورت فشرده قرار گرفته‌اند و ۲۵ درصد داده‌ها به صورت پراکنده‌تر در بازه‌ی ۷/۵ تا ۳۰ قرار دارند.

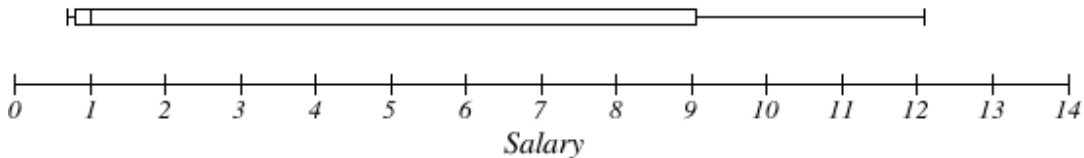
در کار در کلاس صفحه‌ی ۳۷ تاکید می‌شود که برای متغیرهای کمی هم گزارش معیار گرایش به مرکز و هم گزارش معیار پراکندگی لازم است. در این کار در کلاس همچنین اهمیت انتخاب درست معیار گرایش به مرکز نشان داده شده‌است.

در بخش الف. دو نمودار جعبه‌ای مربوط به دو شرکت بر روی یک محور رسم می‌شود تا دانش‌آموزان بتوانند به خوبی آن‌ها را با هم مقایسه کنند.

Company 1



Company 2



حقوق تقریباً ۵۰٪ کارکنان شرکت دوم کمتر از کم‌ترین حقوق کارکنان شرکت اول است. (درست است؛ زیرا میان‌ه‌ی حقوق کارکنان شرکت دوم ۱ میلیون تومان و برابر با کمینه‌ی حقوق کارکنان شرکت اول است.)

اگر نرگس سطح تحصیلات و تجربه‌ی متوسطی داشته باشد و این دو شرکت افراد را بر اساس شایستگی‌هایشان انتخاب کنند، شرکت اول برای او مناسب است. (درست است؛ زیرا یک کارمند متوسط در شرکت یک وضعیت بهتری نسبت به شرکت دو خواهد داشت؛ زیرا میان‌ه‌ی این شرکت بیشتر از میان‌ه‌ی شرکت دو است.)

✘ در شرکت دوم، میانگین معیار گرایش به مرکز مناسبی است. (نادرست است؛ زیرا داده‌ی دورافتاده داریم و توزیع داده‌ها متقارن نیست بنابراین میان‌معیار گرایش به مرکز مناسبی است.)

نکته‌ی مهم این‌که همان‌طور که مریم با یک پرسش تعیین‌کننده و به‌جا اشاره کرده است اگر معیار گرایش به مرکز نیز گزارش می‌شد حتی با انحراف استاندارد داده‌های شرکت دوم (بیش از ۴) نیز می‌توانستیم بفهمیم که در این شرکت حقوق افراد بسیار پراکنده توزیع شده است.

در کار در کلاس دوم صفحه‌ی ۳۷ دانش‌آموزان باید به این نکته توجه کنند که برای متغیرهای کیفی هم گزارش تعداد و هم گزارش درصد اهمیت دارد و گزارش ناقص نتایج می‌تواند گمراه‌کننده باشد. در این کار در کلاس «قبولی در ۱۵ دانشگاه برتر کشور» (قبول شدن / قبول نشدن) یک متغیر کیفی است و مدرسه‌ی اول که قبلاً تعداد قبولی‌هایش ۱ نفر بوده و با تغییر رویکرد آموزشی ۳ نفر شده است ۲۰۰٪ افزایش داشته و مدرسه‌ی دوم که تعداد قبولی‌هایش قبلاً ۷۰ نفر بوده و با تغییر رویکرد آموزشی ۸۴ نفر شده است ۲۰٪ افزایش داشته است. در این مثال دانش‌آموزان متوجه می‌شوند که به درصد اکتفا نکنند و حتماً در مورد این نوع متغیرها همیشه پرسش‌های تعیین‌کننده‌ای در مورد تعداد بپرسند.

در کار در کلاس صفحه‌ی ۳۹ بر یکپارچگی چرخه‌ی حل مسائل با آمار و اهمیت کارگروهی و استفاده از نظر افراد متخصص و بومی در تمامی گام‌های چرخه تاکید می‌شود.

◆ نکته‌ی مهم در این بخش این است که دانش‌آموزان بدانند برای نقد و بررسی نتایجی که روزانه در رسانه‌ها و دنیای واقعی با آن برخورد می‌کنند نه تنها به بخش نتایج که به تمامی گام‌هایی که می‌تواند نتایج را تحت تاثیر قرار دهد می‌بایست نگاهی پرسشگر و نقادانه داشته باشند.

در پرسش ۱ این کار در کلاس نشان داده می‌شود که مشورت با افراد بومی چگونه می‌توانست به ما کمک کند بدانیم بیشتر ساکنین این منطقه چند ساله هستند و چه نیازهایی دارند.

در این کار در کلاس در پرسش ۲ قسمت الف. در گام اول مسئله به صورت دقیق و شفاف تعریف نشده است. این مسئله می‌تواند تمامی گام‌های دیگر را تحت تاثیر قرار دهد. در قسمت ب. در گام طرح و برنامه‌ریزی واحد اندازه‌گیری تعیین نشده است و داده‌ها با واحدهای مختلفی گردآوری شده‌اند. اگر واحدها قابل تبدیل به یکدیگر باشند می‌توان در گام سازمان‌دهی داده‌ها مشکل را حل کرد اما گاهی اگر در گام طرح و برنامه‌ریزی در مورد شکل اندازه‌گیری تصمیم‌گیری نشده باشد، اطلاعات بسیاری از دست می‌رود؛ مثلاً وقتی که برخی داده‌ها به صورت کیفی و برخی کمی گردآوری شده باشند. در قسمت پ. در گام داده‌ها، داده‌ها به درستی پاک‌سازی نشده‌اند؛ دورافتاده بودن داده‌ها لزوم به معنای اشتباه بودن آن‌ها نیست

و گاهی اتفاقاً اطلاعات بسیار مهمی در آن‌ها نهفته است. حذف داده‌های دور افتاده نتایج و تفسیر نتایج را کاملاً تحت تاثیر قرار می‌دهد. در قسمت ت. مسئله برای تمامی گروه توضیح داده نشده است و گروه به صورت یکپارچه در جریان اهداف اصلی مطالعه نیست و این اشکال سمت و سوی اجرای کار را به صورت کلی از اهداف دور می‌کند. در قسمت ث. گروه اجراکننده می‌خواهند نتایج را طبق سلیقه یا صلاحدید خود گزارش کنند. مثلاً اگر میانه معیار گرایش به مرکز بهتری است آن‌ها بنا به مصلحت می‌خواهند میانگین گزارش کنند. این کار تفسیر نتایج و شکل نتیجه‌گیری و درنهایت تصمیمات حاصل از نتایج را کاملاً تحت تاثیر قرار می‌دهد. در قسمت ج. متخصصین آمار بدون دانش و شناخت زمینه و بستر موضوع مورد بررسی به تحلیل نتایج می‌پردازند. بدین ترتیب گام بحث و نتیجه‌گیری که نیازمند این شناخت است و به ارائه‌ی ایده‌ها و پیشنهادهای می‌انجامد به خوبی انجام نمی‌شود.

حل تمرین درس سوم

۱- به نظر می‌رسد کودکانی که زبان مادری‌شان فارسی نیست در مقطع دبستان مشکلات بیشتری برای یادگیری مفاهیم درسی دارند. با انتخاب یک جامعه‌ی آماری محدود (از نظر پایه‌ی تحصیلی، جنسیت، منطقه، زبان مادری و ..) و انتخاب متغیر تصادفی مناسب این مسأله را به صورت دقیق و شفاف بیان کنید. پاسخ: بررسی معدل نیم‌سال اول دانش‌آموزان دختر کرد زبان پایه‌ی اول دبستان مدرسه‌ای در سنج (مقایسه‌ی آن با مدرسه‌ای مشابه در شهری مشابه از نظر اجتماعی-اقتصادی با گویش فارسی یا مقایسه‌ی آن با متوسط معدل دانش‌آموزان دختر پایه‌ی اول کل کشور) (هر پاسخ دیگری که با متغیری قابل اندازه‌گیری و جامعه‌ی آماری مشخص به حل این مسئله برسد قابل قبول است.)

۲- با توجه به تاثیر شیوه‌های مختلف حمل‌ونقل بر محیط‌زیست می‌خواهیم در یک کلاس ۲۸ نفره با انتخاب تصادفی ۹ نفر از دانش‌آموزان، از آن‌ها بپرسیم در هفته‌ی گذشته بیشتر به کدام یک از شیوه‌های زیر به مدرسه آمده‌اند؟

● پیاده یا با دوچرخه ○ با وسایل حمل و نقل عمومی ● با سرویس مدرسه یا خودرو شخصی
الف) در این مطالعه جامعه‌ی آماری، نمونه‌ی آماری، اندازه‌ی جامعه و اندازه‌ی نمونه را مشخص کنید.
پاسخ: جامعه‌ی آماری: تمامی دانش‌آموزان این کلاس؛ نمونه‌ی آماری: ۹ دانش‌آموزی که به تصادف انتخاب شده‌اند؛ اندازه‌ی جامعه: ۲۸؛ اندازه‌ی نمونه: ۹.

ب) با چه روشی می‌توان داده‌های این مطالعه را گردآوری کرد؟ چه مشکلاتی ممکن است در جمع‌آوری این داده‌ها رخ دهد؟

پاسخ: می‌توانیم با طراحی یک پرسش‌نامه داده‌ها را گردآوری کنیم. ممکن است برخی دانش‌آموزان پیاده به مدرسه بیایند ولی با اتوبوس برگردند. قبل از گردآوری داده‌ها باید به این مسائل فکر کنیم.

پ) متغیر تصادفی مورد بررسی این مطالعه چیست؟ نوع متغیر و مقیاس اندازه‌گیری را مشخص کنید.
پاسخ: شیوه‌ی حمل و نقل؛ با توجه به اینکه شیوه‌های حمل و نقل در این مسئله بر اساس تاثیر آن‌ها بر محیط زیست طبقه‌بندی شده‌اند (از سازگارترین با محیط زیست تا زیان‌بارترین برای محیط زیست) این متغیر کیفی ترتیبی است.

ت) چه نمودارها و چه آماره‌هایی برای گزارش نتایج این مطالعه مناسبند؟
پاسخ: می‌توانیم از نمودار میله‌ای یا دایره‌ای و از تعداد و درصد برای گزارش نتایج استفاده کنیم.

ج) آیا می‌توانیم این نتایج را به شیوه‌ی رفت و آمد دانش‌آموزان این کلاس به کل سال تحصیلی تعمیم دهیم (مثلاً به فصل‌های مختلف یا زمان برگزاری آزمون‌های پایان نیم‌سال)؟ توضیح دهید.

پاسخ: خیر؛ زیرا ممکن است شیوه‌ی رفت و آمد دانش‌آموزان در فصل‌های مختلف یا زمان برگزاری آزمون‌های پایان نیم‌سال این‌طور نباشد و تغییر کند.

ث) اگر ۹ نفر دیگر از این کلاس را به تصادف انتخاب می‌کردیم، آیا لزوماً نتایج مشابهی حاصل می‌شد؟ توضیح دهید.

پاسخ: خیر؛ ممکن است نتایج متفاوتی حاصل شود. (لازم است در اینجا به این نکته که در پایه‌ی دهم آمده بود اشاره شود که پارامتر جامعه (تعداد و درصد در کل کلاس) مقدار ثابتی است ولی آماره‌ی نمونه (تعداد و درصد در افرادی که در نمونه قرار گرفته‌اند) متأثر از نمونه‌گیری است و تغییر می‌کند.)

۳- در هر یک از موارد زیر علت عدم تناسب جامعه‌ی آماری با نمونه‌های انتخاب شده را توضیح دهید. برای بهبود نمونه‌گیری چه پیشنهادی دارید؟ انتخاب نادرست نمونه در هر یک از موارد چگونه بر نتایج تاثیر می‌گذارد؟

الف) مسأله: بررسی میزان رضایت شغلی در بین کارمندان یک کارگاه شبانه‌روزی

نمونه: انتخاب تصادفی تعدادی از مدیران ارشد شیفت روز

پاسخ: جامعه‌ی آماری تمامی کارمندان در تمامی شیفت‌ها هستند. اما در نمونه فقط مدیران ارشد شیفت روز شانس حضور دارند. این نمونه‌گیری نتایج را تحت تاثیر قرار می‌دهد و ممکن است میزان رضایت این افراد نسبت

به کارمندان ساده‌ی شیفت شب بالاتر باشد و نتایج گمراه‌کننده شود. باید از لیست تمامی کارمندان همه‌ی شیفت‌ها به صورت تصادفی افرادی را انتخاب کنیم.

(ب) مسأله: نظرسنجی از مادران یک شهر در مورد میزان رضایت‌شان از برنامه‌های کودک تلویزیون نمونه: انتخاب تصادفی برخی منازل از تمامی مناطق آن شهر و مراجعه بین ساعت ۸ تا ۱۰ صبح

پاسخ: در این شیوه‌ی نمونه‌گیری مادران شاغل شانس حضور ندارند و این مشکل می‌تواند نتایج را تحت تأثیر قرار دهد. گردآوری داده‌ها را می‌توان به عصرها یا ساعات مختلفی از شبانه‌روز موکول کرد. (پیشنهادهای دانش‌آموزان می‌بایست مورد بحث و بررسی قرار گیرد و توسط دیگر دانش‌آموزان نقد شود. هر پیشنهادی که به تمامی افراد جامعه شانس حضور برابر دهد قابل قبول است.)

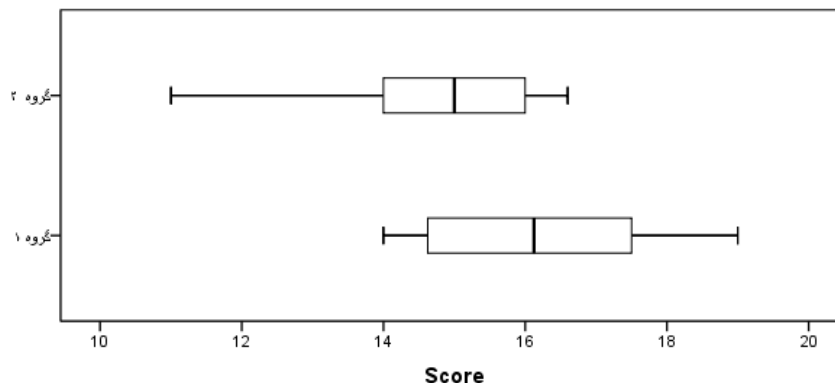
(پ) مسأله: نظرسنجی از دانش‌آموزان پایه‌ی دوازدهم منطقه‌ی دو بوشهر درباره‌ی اینکه به صورت متوسط چند ساعت در روز را به انجام تکالیف مدرسه اختصاص می‌دهند.
نمونه: دختران پایه‌ی دوازدهم یک مدرسه از این منطقه که به صورت داوطلبانه در این مطالعه مشارکت کرده‌اند.

پاسخ: فقط دختران مورد بررسی قرار گرفته‌اند. داوطلبان مشارکت ممکن است دانش‌آموزانی باشند که تکالیف خود را مرتب انجام می‌دهند و این نوع نمونه‌گیری می‌تواند ما را از نتایج واقعی منحرف کند. می‌توان لیستی از دانش‌آموزان (دختران و پسران) مدارس این منطقه تهیه کرد و به صورت تصادفی افرادی را انتخاب کرد.

۴- «مطالعات دانشمندان علوم اعصاب نشان می‌دهد که مغز انسان برای انجام هم‌زمان چند فعالیت ساخته نشده است و هر بار که ما گمان می‌کنیم مشغول انجام دادن چند کار به صورت هم‌زمان هستیم در واقع مغز ما دارد با سرعت از کاری به کار دیگر می‌پرد. مغز انسان تشنه‌ی خبرها و اتفاقات جدید است. وقتی مشغول مطالعه و یادگیری هستیم و ناگهان صدای دریافت یک پیام را از تلفن خود می‌شنویم مغز ما، که بی‌تابانه منتظر کشف آن اطلاع جدید است، دیگر روی یک کار متمرکز نیست. این مطالعات نشان می‌دهد که مثلاً اگر دانش‌آموزی هنگام مطالعه و یادگیری، هم‌زمان تلویزیون تماشا کند یا گفتگویی را در یک شبکه‌ی اجتماعی دنبال کند، مطالب درسی به‌جای اینکه به قسمت حقایق و ایده‌ها برود به قسمت مهارت‌ها فرستاده می‌شود.»

سعید و سهراب پس از خواندن این مطلب در نشریه‌ی مدرسه تصمیم گرفتند نمره‌ی عملکرد تحصیلی ۲۵ دانش‌آموز پایه‌ی دوازدهم مدرسه را که هنگام مطالعه، تلفن همراه خود را در حالت بی‌صدا قرار می‌دهند

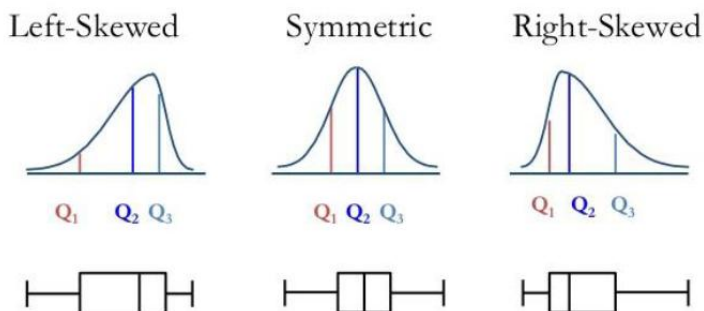
(گروه یک) با ۲۵ دانش‌آموز پایه‌ی دوازدهم دیگر که این کار را نمی‌کنند (گروه دو) مقایسه کنند. نتایج این مطالعه به این شرح است:



الف) از مقایسه‌ی میانه‌ی گروه یک با چارک سوم گروه دو چه نتیجه‌ای می‌گیرید؟
 پاسخ: نیمی از دانش‌آموزان گروه یک عملکرد بهتری نسبت به ۷۵ درصد دانش‌آموزان گروه دو داشته‌اند.
 ب) پراکندگی دو گروه را با استفاده از دامنه‌ی تغییرات و دامنه‌ی میان چارکی مقایسه کنید.
 پاسخ: دامنه‌ی تغییرات گروه دو بیشتر از گروه یک و دامنه‌ی میان چارکی گروه یک بیشتر از گروه دو است.

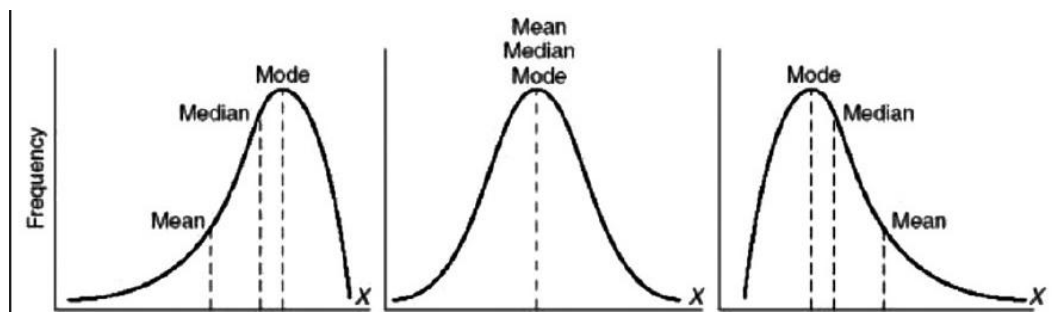
پ) از مقایسه‌ی کمینه‌ی گروه یک با چارک اول گروه دوم چه نتیجه‌ای می‌گیرید؟
 پاسخ: تمامی دانش‌آموزان گروه یک عملکرد بهتری نسبت به ۲۵ درصد دانش‌آموزان گروه دو دارند.

ت) در کدام گروه گزارش میانگین و انحراف معیار می‌تواند گمراه‌کننده باشد؟
 پاسخ: به نظر می‌رسد گروه دو داده یا داده‌های دورافتاده‌ی کوچکی دارد (کشیدگی یک چهارم اول نمودار). به همین دلیل گزارش میانگین و انحراف استاندارد (معیار) برای این گروه مناسب نیست. در این قسمت برای درک بهتر دانش‌آموزان می‌توان به شکل‌های زیر و چگونگی تاثیر توزیع داده‌ها بر نمودار جعبه‌ای اشاره کرد.



ث) در کدام گروه مقدار میانگین و میانه به هم نزدیک‌تر هستند؟

پاسخ: به نظر می‌رسد در گروه یک توزیع داده‌ها متقارن‌تر است و در نتیجه میانگین و میانه نزدیک‌تر هستند. در این قسمت برای درک بهتر دانش‌آموزان می‌توان به شکل‌های زیر و چگونگی تاثیر توزیع داده‌ها بر معیارهای گرایش به مرکز اشاره کرد.



ج) نتایج این مطالعه را به چه جامعه‌ای می‌توان تعمیم داد؟
 پاسخ: دانش‌آموزان پایه‌ی دوازدهم مدرسه‌ی سعید و سهراب

۵- در یک نظرسنجی از ۱۵ دانش‌آموز علوم انسانی پایه‌ی یازدهم یک مدرسه پرسیده‌ایم به نظر آن‌ها چند ساعت آموزشی در ماه برای برگزاری جلسات نقد کتاب یا فیلم مناسب است. داده‌های جمع‌آوری شده به این شرح است:

۱۲ ۱۴ ۲۰ ۲۲ ۱۸ ۲۰ ۰ ۴ ۶ ۱۳ ۱۲ ۱۰ ۱۲ ۱۰ ۱۶

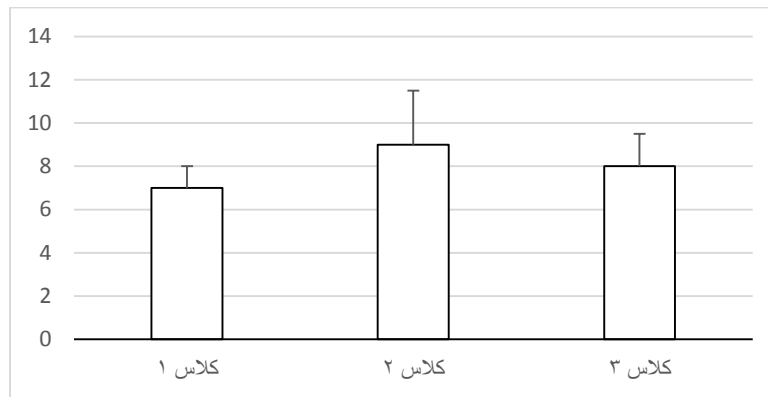
الف) این داده‌ها را طوری تغییر دهید که میانگین و انحراف معیار بیشتر شود اما میانه، چارک اول و چارک سوم تغییر نکند.

پاسخ: پس از مرتب کردن داده‌ها متوجه می‌شویم که مثلاً اگر به جای بزرگ‌ترین داده (۲۲)، عددی مانند ۳۲ بگذاریم میانگین و انحراف استاندارد بزرگ‌تر می‌شوند ولی میانه و چارک‌ها تغییر نمی‌کنند.

ب) فرض کنید می‌خواهیم دو نفر دیگر به این نمونه اضافه کنیم. داده‌های این دو نفر را طوری انتخاب کنید که میانگین و میانه تغییر نکند.

پاسخ: میانگین این داده‌ها $12/6$ است. با توجه به مفهوم مرکز ثقل برای میانگین، اگر دو عدد با فاصله‌های برابر (یکی کمتر و دیگری بیشتر از میانگین) به داده‌ها اضافه شود میانگین تغییر نخواهد کرد. اگر بخواهیم میانه نیز تغییر نکند نباید عددی بین میانگین ($12/6$) و میانه (۱۲) قرار گیرد. مثلاً دو عدد $11/6$ و $13/6$ مناسبند.

۶- با توجه به تاثیر کمیت و کیفیت خواب بر یادگیری، ندا و آوا در مطالعه‌ای با بررسی تمامی دانش‌آموزان سه کلاس پایه‌ی دهم میانگین ساعت خواب این دانش‌آموزان را در هفته‌ی گذشته بر حسب ساعت ثبت کردند. آن‌ها پس از گردآوری داده‌ها و بررسی درستی آن، نتایج را به صورت نمودارهای زیر نمایش دادند. (در این نمودارها بلندی مستطیل نشان‌دهنده‌ی میانگین و میله‌ی خطا نشان‌دهنده‌ی انحراف معیار است.)



الف) چه عواملی ممکن است هنگام گردآوری داده‌ها نتایج را از واقعیت دور کند؟ (هنگامی که از دانش‌آموزان راجع به میانگین ساعت خوابشان در هفته‌ی گذشته می‌پرسیم آیا آنها می‌توانند به درستی به یاد آورند که هر شب چند ساعت خوابیده‌اند تا بتوانند میانگین را به درستی اعلام کنند؟ برای حل این مشکل چه راهکاری پیشنهاد می‌کنید؟)

پاسخ: فراموش کردن ساعت‌های خواب می‌تواند بر نتایج تاثیر بگذارد. بهتر است پرسش‌نامه‌ای تهیه کنیم و از آن‌ها بخواهیم در طول هفته هر شب میزان خواب خود را یادداشت کنند.

ب) نوع متغیر و مقیاس اندازه‌گیری مورد بررسی در این مطالعه را مشخص کنید.

پاسخ: کمی، نسبی

ت) با استفاده از اطلاعات نمودارها، نتایج را به صورت تقریبی در جدول زیر بنویسید.

	کلاس ۱	کلاس ۲	کلاس ۳
میانگین	۷	۹	۸
انحراف معیار	۱	۲/۵	۱/۵

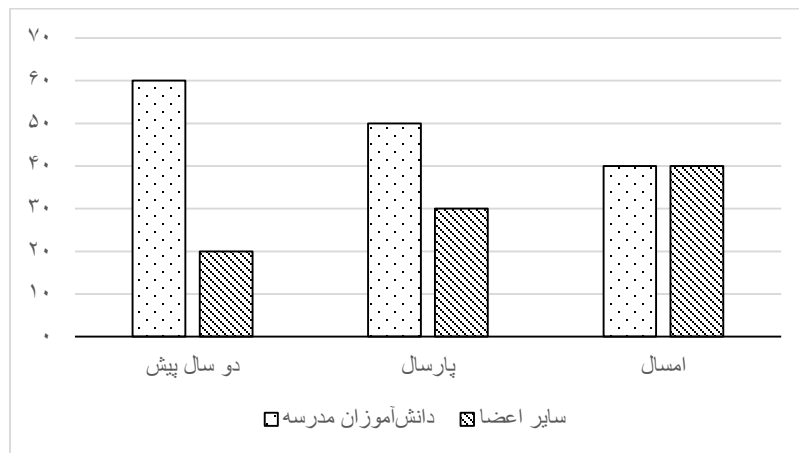
ث) چه عواملی می‌تواند نتایج این مطالعه را تحت تاثیر قرار دهد؟ چگونه می‌توانیم با بیان مسأله‌ای جدید پاسخ دقیق‌تری برای این مسأله پیدا کنیم؟

پاسخ: ممکن است یکی از کلاس‌ها در زمان انجام این مطالعه آزمونی داشته که میزان خواب دانش‌آموزان آن کلاس را تحت تاثیر قرار دهد. بهتر است از این نظر هماهنگی بین کلاس‌ها برقرار باشد یا اینکه مطالعه در زمان طولانی‌تری (چند هفته) انجام شود.

ج) چه کسانی می‌توانند در اجرای بهتر این مطالعه به ما کمک کنند؟ چگونه؟

پاسخ: متخصصین اختلالات خواب، روانشناسان، والدین و مشاوران آموزشی می‌توانند در اندازه‌گیری و طراحی یک پرسش‌نامه‌ی مناسب و اجرا به ما کمک کنند.

۷- مدرسه‌ای به منظور ترویج فرهنگ کتاب‌خوانی از دو سال پیش به غیر از دانش‌آموزان مدرسه، از خارج از مدرسه نیز عضو می‌پذیرد. نمودار میله‌ای زیر نشان‌دهنده‌ی روند تغییرات فراوانی اعضای جدید کتابخانه در دو گروه مذکور است.



الف) سال گذشته چند نفر از خارج مدرسه عضو کتابخانه شده‌اند؟

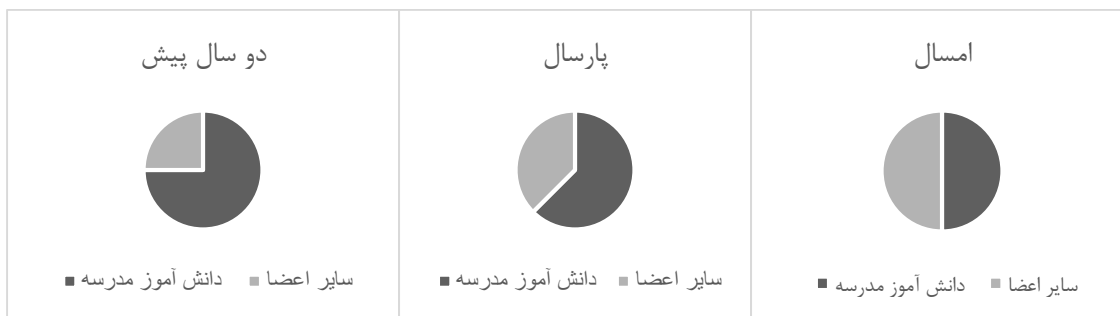
پاسخ: ۳۰ نفر

ب) با توجه به روند این تغییرات، پیش‌بینی می‌کنید سال آینده چه تعداد از دانش‌آموزان این مدرسه عضو کتابخانه شوند؟

پاسخ: ۳۰ نفر

ت) با استفاده از نمودارهای میله‌ای داده شده، برای هریک از این سه سال یک نمودار دایره‌ای جداگانه رسم کنید.

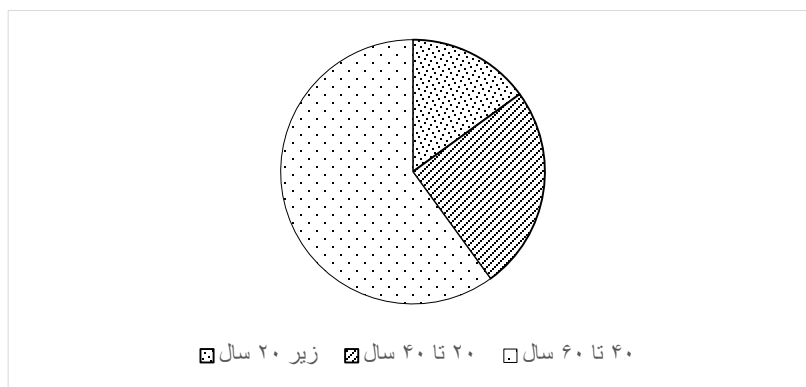
پاسخ:



ث) به نظر شما چه عواملی موجب این روند تغییرات شده‌است؟

پاسخ: هر پاسخی از دانش‌آموزان که بتواند نتایج را توجیه کند پذیرفته است. مثلاً شلوغی فضای کتابخانه با افرادی ناآشنا، تمایل کتابخانه به ثبت‌نام افراد خارج از مدرسه به دلیل حق عضویت بیشتر یا کمک‌های آن‌ها به کتابخانه و یا اینکه به مرور زمان با افزایش اعضای خارج مدرسه کتاب‌ها از سلايق دانش‌آموزان دور شده است.

۸- نتیجه‌ی مطالعه‌ی فراوانی بازدیدکنندگان از یک سالن نمایش را در گروه‌های مختلف سنی در نمودار دایره‌ای زیر نمایش داده است. (در این نمودار ۵۴ درجه مربوطه به گروه سنی زیر ۲۰ سال، ۹۰ درجه مربوط به گروه سنی ۲۰ تا ۴۰ سال و مابقی مربوط به گروه سنی ۴۰ تا ۶۰ سال است)



الف) با توجه به این نمودار، چند درصد افراد بازدیدکننده بین ۴۰ تا ۶۰ سال سن دارند؟
پاسخ: ۶۰ درصد

ب) مسئول فروش بلیت این سالن پس از مشاهده نتایج، آن را غیرواقعی دانست و گفت: «بیشتر بازدیدکنندگان این سالن کمتر از ۴۰ سال دارند.» از موارد زیر کدام می‌توانند نتایج این مطالعه را بدین شکل از واقعیت دور کرده باشد؟ توضیح دهید.

- بسیاری از افراد کمتر از چهل سال تمایلی به پاسخ‌گویی به سؤالات پرسشگر نداشته‌اند.
- گردآوری داده‌ها در ساعات اداری انجام شده است و بازنشستگان بیشتری در نمونه قرار گرفته‌اند.

☒ هنگام گردآوری داده‌ها، دانش‌آموزان بسیاری از طرف مدرسه برای بازدید حضور داشته‌اند.

۹- جدول زیر نشان دهنده‌ی تعداد تصادف خودروها و سرعت حرکت آنها در زمان تصادف است.

سرعت (کیلومتر در ساعت)	۱۰	۲۰	۳۰	۴۰	۵۰	۶۰	۷۰	۸۰	۹۰	۱۰۰	۱۱۰	۱۲۰
تعداد خودروهای تصادف کرده	۴۰	۱۳۰	۲۱۰	۲۵۰	۳۵۰	۲۴۰	۱۵۰	۷۰	۴۰	۲۰	۲۰	۱۰

همانطور که دیده می‌شود، تعداد تصادف خودروهایی که سرعتشان بیش از ۹۰ کیلومتر در ساعت بوده، کمتر است. پس: «هرچه سریع‌تر بروید، مطمئن‌تر و امن‌تر است.» نتیجه‌گیری بالا چه اشکالی دارد؟ چرا این تصور ایجاد شده است؟ پاسخ: این نتیجه‌گیری صحیح نیست. زیرا نمی‌دانیم که در هر سرعت چند خودرو با آن سرعت حرکت کرده است. مثلاً اگر ۱۰ خودرو با سرعت بین ۱۱۰ تا ۱۲۰ حرکت کرده باشد، این نتایج نشان می‌دهد که صد درصد آنها تصادف داشته‌اند. برای تحلیل صحیح داده‌ها نباید فراوانی گزارش شود بلکه باید نسبت تعداد تصادف‌های هر دسته به تعداد خودروهایی که با سرعت آن دسته حرکت می‌کرده‌اند گزارش شود. (این اطلاعات در این مسئله موجود نیست.)

پروژه

انجام پروژه توسط دانش‌آموزان به صورت گروه‌های چند نفره از اهمیت ویژه‌ای برخوردار است و دانش‌آموزان مفاهیم و نکاتی را که در این درس آموخته‌اند در قالب انجام پروژه به خوبی درک خواهند کرد. انجام پروژه گروهی با موضوعی که دانش‌آموزان به آن علاقمند هستند «توانایی نقد و بررسی نتایج مطالعه‌ی دیگر گروه‌ها» را در آنها تقویت می‌کند. دانش‌آموزان با کار گروهی می‌آموزند که لزوماً پاسخ یکتایی برای حل هر مسئله وجود ندارد و می‌توان به شیوه‌های مختلف به راهی برای حل مسئله فکر کرد. انجام پروژه با هدایت معلم به دانش‌آموزان کمک می‌کند در تمامی گام‌های چرخه‌ی حل مسئله به اهداف نخستین کار بیاورند و به محدودیت‌های انجام پروژه واقف باشند.

