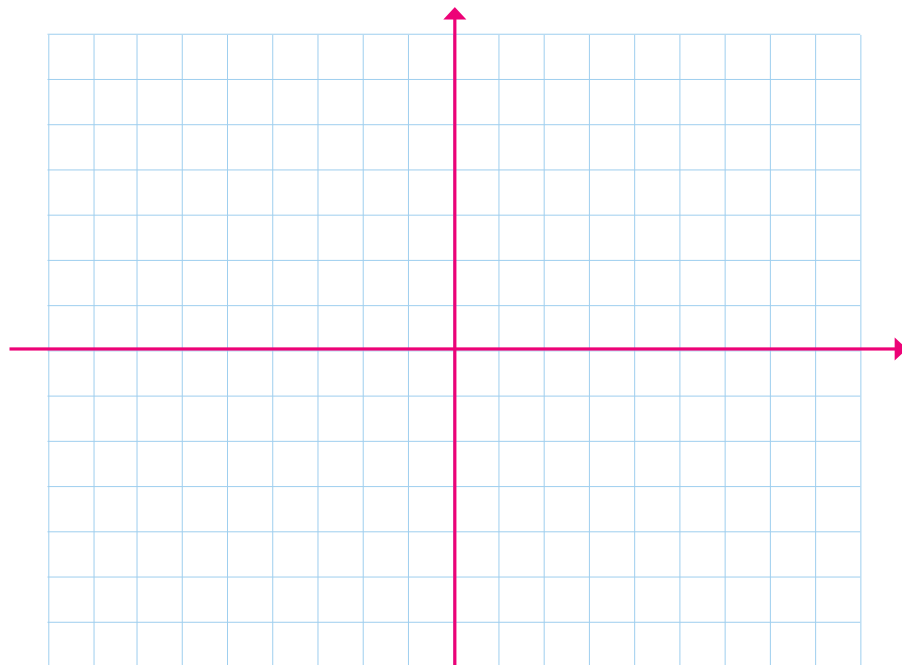




کار در کلاس

جدول زیر را کامل کنید و به کمک آن نمودار تابع با ضابطه $f(x) = 4 - x^2$ را رسم کنید.

x	...	-4	-3	-2	-1	0	1	3	4	...
y	...		9			0				...



کار در کلاس

نقاط رأس سهمی‌هایی را که معادله‌های آنها داده شده است، مشخص کنید.

الف) $y = x^2$

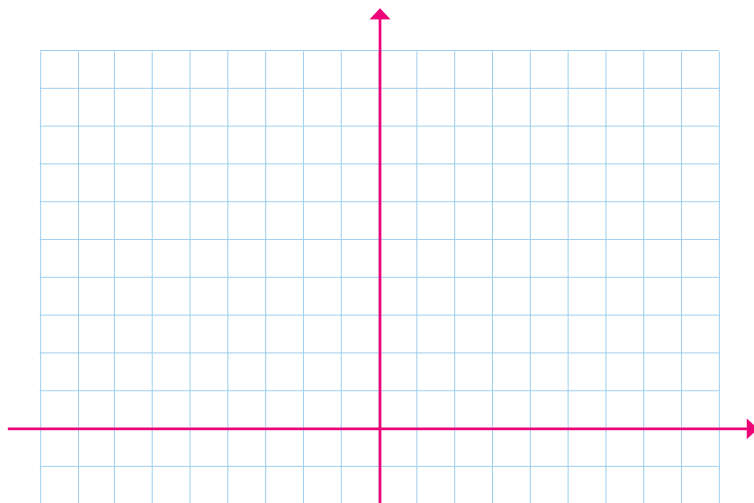
ب) $y = 4 - x^2$

پ) $y = 2x^2 - 4x + 1$

ت) $y = 2(x-1)^2 + 1$

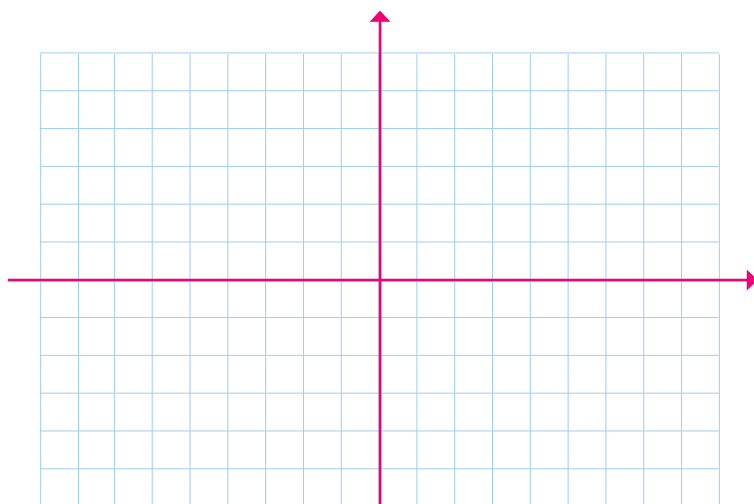


الف) رأس سهمی به معادله $y = 2(x-1)^2 + 1$ را مشخص کنید به کمک آن نمودار سهمی را رسم کنید.



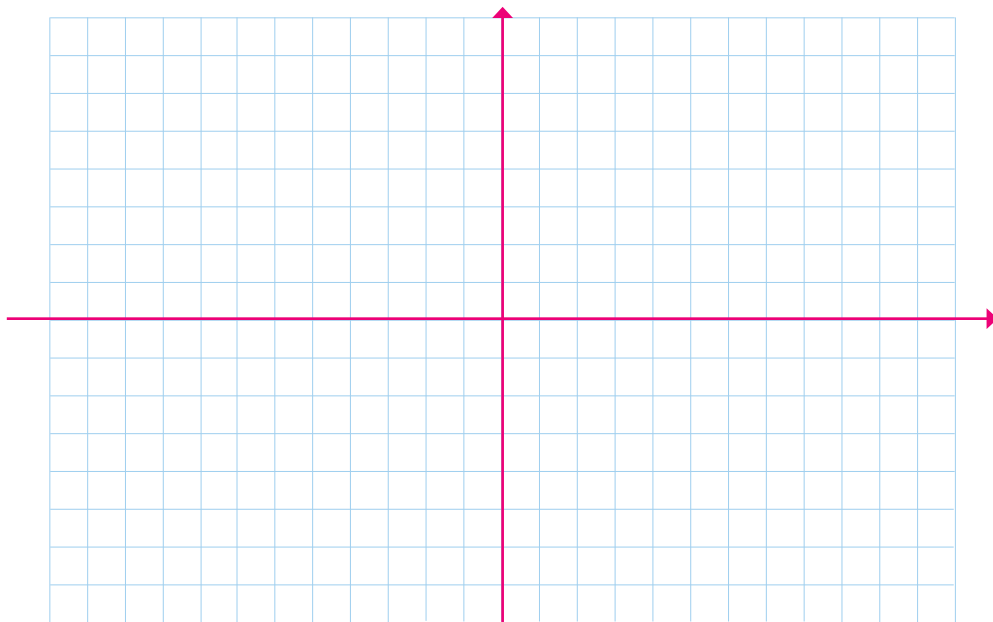
ب) جدول زیر را کامل کنید و به کمک آن نمودار سهمی به معادله $y = -\frac{1}{4}(x+2)^2 + 5$ را رسم کنید.

x	...	-5	-4	-3	-2	-1	0	1	...
y	...				0				...

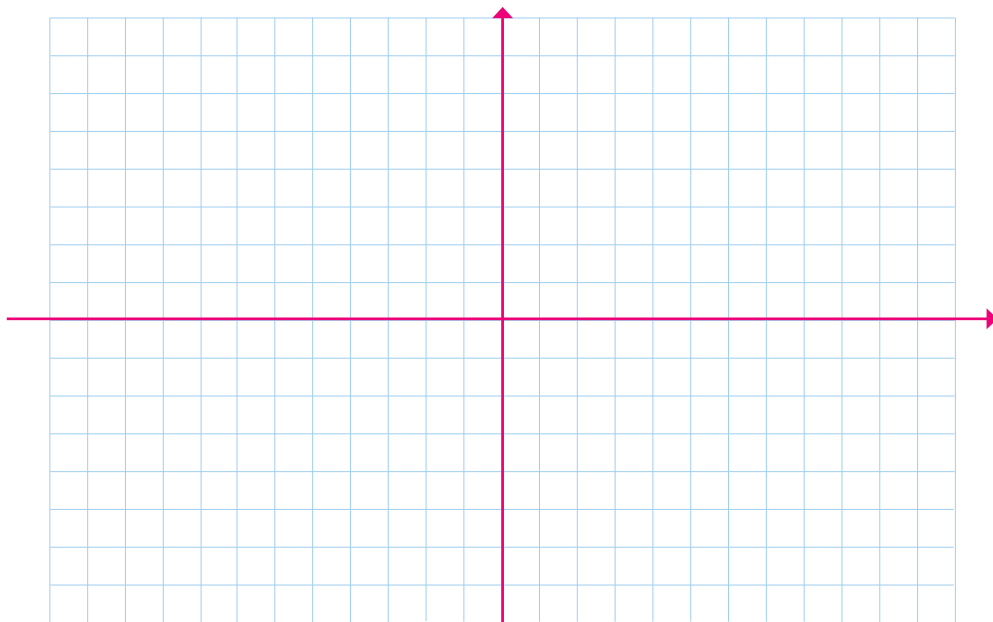




۱. نمودار توابع درجه دوم $y = -x^2 + 6x - 1$ و $y = 2x^2 + 4x + 1$ را رسم کنید.



۲. نمودار توابع $y = -x^2 + 6x - 1$ و $y = x^2 + 4x - 2$ را رسم کنید و در ادامه، مختصات برخورد این دو خم را مشخص کنید.

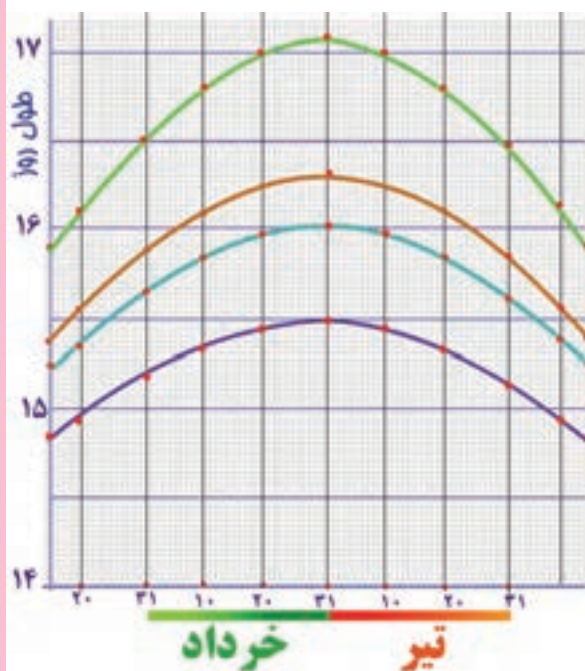




تمرین

۱. نمودار سهمی‌های به معادله‌های $y = x^2 - 2x$ و $y = -(x-1)^2 + 1$ و $y = x^2 + 4x + 1$ را رسم کنید.
 ۲. اگر تابع درآمد به صورت $y = -\frac{1}{4}x^2 + 30x$ و تابع هزینه به صورت $y = 18x + 40$ باشد، ماکسیم مقدار سود را مشخص کنید.
 ۳. محیط مستطیلی ۲۶ متر است. اگر اندازه یکی از اضلاع آن را با x و مساحت آن را با s نشان دهیم، ابتدا نمودار تابع مساحت را بر حسب x رسم کنید. سپس به کمک نمودار مشخص کنید به ازای چه مقداری از x مساحت مستطیل ماکسیم می‌شود.
 ۴. اگر $2x + a = 100$ باشد x و a را طوری بیابید که $y = xa$ ماکسیم شود.
 ۵. در یک تولیدی، نوعی لامپ، برای مصارف پزشکی تولید می‌شود. این تولیدی هریک از لامپ‌ها را می‌تواند به قیمت ۲۰۰ تومان بفروشد. اگر در هر روز x واحد لامپ تولید کند و بفروشد و تابع هزینه آن برابر $c(x) = x^2 + 40x + 100$ باشد:
- الف) تابع سود روزانه این تولیدی را بنویسید.
- ب) چند لامپ در روز تولید کند تا بیشترین سود را داشته باشد؟
- پ) بیشترین سود روزانه این کارگاه چقدر است؟

خواندنی



در روز ۳۱ خرداد، محور زمین در قطب شمال بیشترین انحراف به سمت خورشید را دارد. در این روز، خورشید در هنگام ظهر در بالاترین زاویه ممکن در آن محل قرار دارد. دقت کنید که روز ۳۱ خرداد طولانی‌ترین روز سال در نیم کره شمالی است. اما در عرض‌های بالاتر، این زمان بیشتر است. به عبارت دیگر هر چه به خط استوا نزدیک شویم، طول روز در ۳۱ خرداد عددی کوچک‌تر و هر چه از خط استوا دور شویم عدد بزرگ‌تری می‌باشد. نمودار مقابل نشان می‌دهد که طول روز در ایران در ایام خرداد ماه و تیرماه در شهرهای ایران چه اختلافی با هم دارند.



فصل چهارم

کار با داده‌های آماری

گردآوری داده‌ها

درس ۱

معیارهای گرایش به مرکز

درس ۲

معیارهای پراکندگی

درس ۳



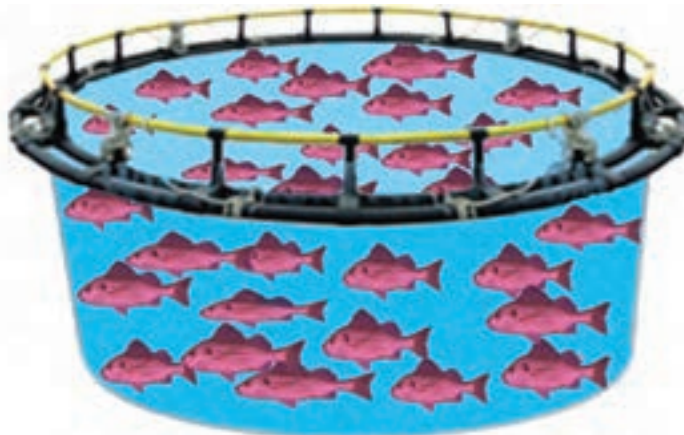
درس ۱

گردآوری داده‌ها

فعالیت



می‌خواهیم وزن ماهی‌های یک حوضچه پرورش ماهی را به منظور فروش آنها تخمین بزنیم. ابتدا از قسمت عمیق و در مرحله بعد، از قسمت کم عمق ۵ ماهی صید می‌کنیم.



انتخاب ۵ ماهی از قسمت عمیق



انتخاب ۵ ماهی از قسمت کم عمق



۱. آیا انتخاب ۵ ماهی از قسمت عمیق، تخمین خوبی از وزن ماهی‌های حوضچه به ما می‌دهد؟ یعنی آیا

می‌توان برای فروش آنها اعلام آمادگی کرد؟

۲. به نظر شما کدام تخمین بیش از مقدار واقعی است و کدام یک کمتر از مقدار واقعی؟

۳. اگر شما امکان صید ۵ ماهی داشتید، چگونه آنها را انتخاب می‌کردید تا منجر به تخمین بهتری از وزن ماهی‌های حوضچه می‌شد؟ فرض کنید ماهی‌ها، همانند شکل در حوضچه پخش شده‌اند و تحرک زیادی ندارند.

۴. اگر از نحوه پخش شدن ماهی‌ها اطلاعی نداشتیم، بهتر بود ۵ ماهی را چگونه انتخاب می‌کردیم؟ آیا انتخاب تصادفی چند بخش از حوضچه (مثلاً زمانی که آن را شطرنجی کرده‌ایم) به ما کمک می‌کند ماهی‌های انتخابی معرف بهتری از کل ماهی‌ها باشند؟

داده‌ها واقعیت‌هایی درباره یک چیزاند که در محاسبه، استنباط، یا برنامه‌ریزی به کار می‌روند. واحد آماری به هر یک از افراد یا چیزهایی می‌گویند که داده‌های مربوط به آنها در یک بررسی آماری گردآوری می‌شود. مجموعه کل واحدهای آماری را جامعه آماری می‌نامند. هر زیرمجموعه از جامعه آماری را که با روش مشخصی انتخاب شده باشد، یک نمونه می‌نامند. نمونه‌ای را که در آن، همه اعضای جامعه، شانس انتخاب یکسان در نمونه را داشته باشند نمونه تصادفی می‌نامند.

در مثال حوضچه ماهی، هر ماهی درون حوضچه یک واحد آماری است. به کل ماهی‌های حوضچه که عبارت است از مجموعه همه واحدهای آماری جامعه گفته می‌شود. اگر وزن تک‌تک ماهی‌ها را در اختیار داشته باشیم داده‌های جامعه را داریم. وزن نمونه ۵ ماهی از قسمت کم عمق معرف داده‌های یک نمونه پنج‌تایی است. اگر ۵ ماهی با یک روش تصادفی از حوضچه استخراج شود، عملاً یک نمونه تصادفی ۵ تایی از حوضچه در اختیار داریم.

خواندنی

تفاوت زیادی بین عدد در ریاضی و داده در آمار وجود دارد. به عبارت دیگر عدد ۵۰ یک مفهوم در ریاضی دارد و داده‌ای که مقدار آن ۵۰ است، علاوه بر مقدار آن حاوی اطلاعات زیادی است. به‌عنوان مثال این داده می‌تواند متوسط تلفات روزانه جاده‌های کشور در یکی از سال‌های اخیر باشد.

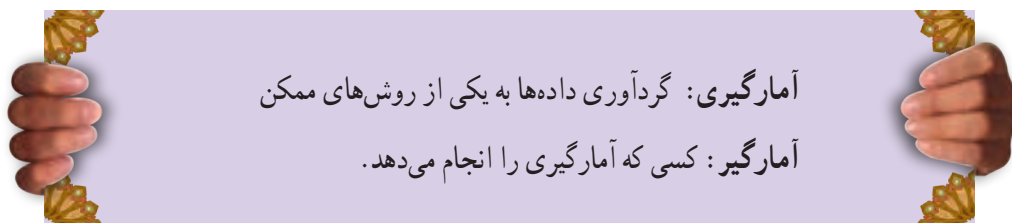


از قدیم گفته‌اند که
مشت نمونه خروار است،
ولی نه هر مشت. چرا؟
اگر نمونه تصادفی
انتخاب شود
آنگاه نمونه
معرف جامعه
خواهد بود.



فعالیت

می‌خواهیم طول قد دانش‌آموزان یک مدرسه را گردآوری کنیم. برای این منظور چه راهی پیشنهاد می‌کنید؟



آمارگیری: گردآوری داده‌ها به یکی از روش‌های ممکن
آمارگیر: کسی که آمارگیری را انجام می‌دهد.

اگر قرار شد آمارگیر باشیم، می‌توانیم جدولی به صورت زیر تکمیل کنیم.

مثالی از جدول طراحی شده برای ثبت داده‌ها

اندازه طول قد	چوب‌خط برای شمارش	تعداد دانش‌آموزان
کوتاه‌تر از ۱۴۰ سانتی‌متر		
۱۴۰-۱۴۹ سانتی‌متر		
۱۵۰-۱۵۹ سانتی‌متر		
۱۶۰-۱۶۹ سانتی‌متر		
۱۷۰ سانتی‌متر یا بلندتر		



چگونه مطمئن می‌شوید که دانش‌آموزی از قلم نیفتاده است؟ چه راهکاری برای این منظور پیشنهاد می‌کنید؟

آمارگیری زحمت زیادی برای آمارگیر دارد. آیا راه‌حل ساده‌تری برای انجام آن دارید؟ یکی از مرسوم‌ترین روش‌های آمارگیری، استفاده از پرسش‌نامه است. پرسش‌نامه شبیه همان جدولی است که هنگام ثبت‌نام در مدرسه، شما یا والدین، آن را تکمیل کرده‌اید. پرسش‌نامه را می‌توانند واحدهای جامعه یا نمونه تکمیل کنند.

مثالی از پرسش‌نامه طراحی شده

سلام، می‌خواهیم طول قد دانش آموزان مدرسه را آمارگیری کنیم.
لطفاً یکی از گزینه‌ها را انتخاب کنید.
طول قد شما چقدر است؟

- ☐ کوتاه‌تر از ۱۴۰ سانتی‌متر
- ☐ ۱۴۰-۱۴۹ سانتی‌متر
- ☐ ۱۵۰-۱۵۹ سانتی‌متر
- ☐ ۱۶۰-۱۶۹ سانتی‌متر
- ☐ ۱۷۰ سانتی‌متر یا بلندتر

خواندنی

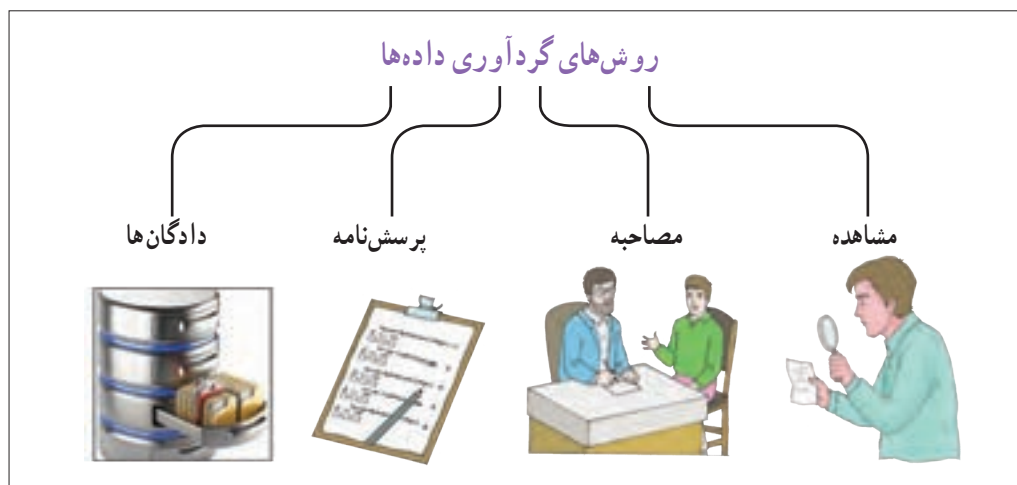
برای به انجام رساندن یک آمارگیری باید پاسخ سؤال‌های زیر را بدانید:

۱. می‌خواهید چه چیزی پیدا کنید؟ چرا؟
۲. چه داده‌هایی را باید گردآوری کنید؟ چگونه تصمیم می‌گیرید؟
۳. کدام راه، بهترین راه گردآوری داده‌ها است؟ چرا؟ آیا به ابزار خاصی نیاز دارید؟
۴. به چه مقدار داده نیاز دارید؟ چگونه تصمیم می‌گیرید؟
۵. داده‌های شما باید چقدر دقت داشته باشند؟ چرا؟
۶. داده‌هایتان را چگونه ثبت می‌کنید؟ چرا این روش را انتخاب کرده‌اید؟
۷. داده‌ها را چگونه ارائه می‌کنید؟ چرا؟
۸. آیا داده‌ها از الگوی خاصی پیروی می‌کنند؟ داده‌ها بیانگر چه هستند؟
۹. چه نتیجه‌گیری یا پیشگویی خاصی می‌توانید از داده‌ها ارائه کنید؟
۱۰. آیا نتایج، با آنچه انتظار داشتید، تطبیق دارد؟
۱۱. چگونه نتایج کار را ارائه می‌کنید؟ برای چه کسانی ارائه می‌کنید؟
۱۲. با توجه به نتایج به دست آمده، آیا می‌توانید سؤالات دیگری را نیز بررسی کنید، به نظر شما در مرحله گردآوری داده، به کدام یک از سؤالات فوق باید پاسخ داده شود؟



کار در کلاس

۱. چه راه دیگری برای آمارگیری طول قد دانش آموزان یک مدرسه پیشنهاد می کنید؟
۲. فرض کنید زمان لازم را برای گردآوری تمامی داده های دانش آموزان در اختیار نداشته باشید. اگر بخواهیم نمونه ای را انتخاب و آمارگیری کنیم، چه راهی پیشنهاد می کنید که نمونه به صورت تصادفی انتخاب شود؟



۱. **مشاهده:** گردآوری داده ها بدون نیاز به فرد پاسخ گو، مانند شمارش تعداد وسایل نقلیه عبوری از یک تقاطع در هر ساعت یا اندازه گیری وزن محصولات یک باغ میوه.





۲. پرسش نامه: مجموعه سؤالات از پیش تعیین شده که توسط تعدادی پاسخ دهنده تکمیل می شود. این روش مرسوم ترین ابزار گرفتن اطلاعات از مردم است. مرکز آمار ایران هر ۱۰ سال یکبار با استفاده از پرسش نامه اطلاعات تمامی خانوارهای ساکن در ایران را جمع آوری می کند. به این فرایند، سرشماری نفوس و مسکن می گوئیم.



۳. مصاحبه: معمولاً بین دو نفر صورت می گیرد، یکی مصاحبه گر (همان آمارگیر) و دیگری مصاحبه شونده یا پاسخ گو است. مثلاً اگر بخواهیم درباره مسائل فرهنگی کاهش شدآمد (ترافیک) پژوهش کنیم، مصاحبه از صاحب نظران راه حل مناسبی برای گردآوری داده هاست. این روش بیشتر زمانی استفاده می شود که آمارگیر اطلاع کافی از تمامی پاسخ های ممکن را ندارد.

۴. دادگان ها: شامل مجموعه ای از اطلاعات ذخیره شده اند. در بسیاری از موارد، داده ها را می توان از اطلاعاتی که قبلاً ذخیره شده اند، به دست آورد. اگر قرار است تحقیقی در مورد نمره های دروس ریاضی استان ها انجام شود، اطلاعات ثبتی اداره کل آموزش و پرورش راه گشا خواهد بود. از سوی دیگر به دلیل تولید داده ها به صورت خودکار، در بسیاری از مؤسسات و سامانه ها، استفاده از این روش برای گردآوری داده ها به سرعت رواج یافته است.

تمرین



کدام روش جمع آوری داده ها برای موارد زیر مناسب است؟ یک دلیل برای انتخاب خود ذکر کنید.

۱. میزان رضایت مشتریان بانک از نحوه برخورد و رسیدگی به درخواست های آنها.
۲. سن همه دانش آموزان مدرسه بر حسب ماه در پایه دهم.
۳. تعداد سرنشینان خودروهای سواری در یکی از محورهای خروجی شهر.



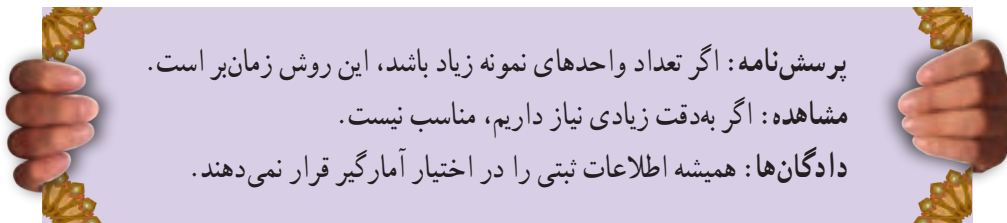
کار در کلاس



الف) کدام روش برای جمع‌آوری هر یک از داده‌ها مناسب است؟

۱. تعداد قلم‌های هر دانش‌آموز در یک کلاس.
۲. ساعات خواب دانش‌آموزان کلاس درس شما در شب گذشته.
۳. طول قد دانش‌آموزان یک کلاس.

ب) می‌خواهیم طول قد دانش‌آموزان یک کلاس یا مدرسه را به یکی از سه روش زیر آمارگیری کنیم. هریک از این روش‌ها محدودیت‌هایی دارند. چگونه می‌توان این محدودیت‌ها را از بین برد؟



پرسش‌نامه: اگر تعداد واحدهای نمونه زیاد باشد، این روش زمان‌بر است.
 مشاهده: اگر به‌دقت زیادی نیاز داریم، مناسب نیست.
 دادگان‌ها: همیشه اطلاعات ثبتی را در اختیار آمارگیر قرار نمی‌دهند.

آمارگیری را می‌توان به روش‌هایی بسیار سریع‌تر یا کم‌هزینه‌تر مانند آمارگیری پستی، تلفنی، اینترنتی یا پیامکی انجام داد. همچنین می‌توان با ابزاری نظیر گوگل فرم یک پرسش‌نامه طراحی کرد، و آن را به نشانی نمونه انتخابی ارسال کرد و نتایج را از گوگل فرم بازایی کنیم.

خواندنی



فعالیت

قرار است درباره‌ی افرادی که از کوه دنا بالا رفته‌اند، پژوهشی آماری انجام دهیم. واحدهای آماری این پژوهش، همه‌ی افرادی هستند که توانسته‌اند به قله برسند. هدف از این پژوهش می‌تواند فرهنگی یا علمی باشد. بسته به نوع پژوهش، یک یا چند ویژگی این افراد (مانند طول قد یا جنسیت) موردنیاز است. به هریک از این ویژگی‌ها که مورد پژوهش قرار می‌گیرد متغیر می‌گویند. سایر متغیرها می‌توانند مواردی مانند: سن، وزن، ملیت، میزان تحصیلات و درآمد باشند. متغیرهای مورد بررسی در یک پژوهش ممکن است کمی یا کیفی باشند.

متغیر: هر ویژگی از اشخاص یا اشیا که قرار است بررسی شود.
متغیر کمی: متغیرهایی هستند که مقادیر عددی می‌گیرند و برای آنها عملیات ریاضی از قبیل جمع، تفریق و معدل‌گیری قابل انجام است.
متغیر کیفی: متغیرهایی هستند که صرفاً برای دسته‌بندی افراد یا اشیا در گروه‌ها به کار می‌روند و لزوماً مقدار عددی نمی‌گیرند.

در مثال کوهنوردان دنا، سن، وزن، قد و درآمد یک کوهنورد متغیرهای کمی هستند. متغیرهای کیفی معمولاً از نوع مشاهدات غیر عددی‌اند و در مثال کوهنوردان دنا، جنسیت و ملیت را در بر می‌گیرند. به عنوان مثال جنسیت برای دسته‌بندی افراد به مرد و زن استفاده می‌شود.

پارامتر جامعه: یک مشخصه عددی است که توصیف کننده جنبه‌ای خاص از جامعه است و در صورتی که داده‌های کل جامعه در اختیار باشند قابل محاسبه است. مثلاً اگر داده‌های مربوط به تک‌تک کوهنوردان را داشته باشیم، یعنی به داده‌های جامعه دسترسی داریم. نسبت مردان در کل جامعه کوهنوردان، معرف یک پارامتر است.

اگر داده‌های بعضی از کوهنوردان را داشته باشیم؛ یعنی داده‌های نمونه را در اختیار داریم. نسبت مردان کوهنورد به این داده‌های نمونه‌ای را، آماره (مقدار آماره) گویند. آماره‌ها از یک نمونه به نمونه دیگر تغییر می‌کنند؛ این در حالی است که پارامترهای جامعه همیشه ثابت‌اند، چرا؟

در بسیاری از موارد، آمارگیری از کل جامعه امکان‌پذیر نیست. بنابراین علی‌رغم اینکه پارامتر دارای مقدار ثابتی است، این مقدار مجهول است و به همین دلیل از آماره‌ها برای تخمین پارامترها استفاده می‌کنند.

آماره نمونه: مشخصه‌ای عددی که توصیف کننده جنبه‌ای خاص از نمونه است و از داده‌های نمونه به دست می‌آید.

مثال: اداره کشاورزی استان خوزستان در حال ارزیابی هندوانه‌های آماده برداشت است. در این بررسی، هندوانه‌ها همان واحدهای آماری هستند. اگر پژوهشگران وزن هندوانه‌ها را مورد بررسی قرار دهند، متغیر، «وزن» آنهاست. وزن یک متغیر کمی است، زیرا با مقادیر عددی ارائه می‌شود. اگر وزن تک‌تک هندوانه‌های این زمین بررسی شود، سرشماری از جامعه انجام داده‌ایم (که امکان‌پذیر نیست). متوسط وزن تمامی هندوانه‌های قابل برداشت در این زمین، «پارامتر» است.

حال فرض کنیم پژوهشگران تصمیم دارند بر اساس معیار «مزه» هندوانه‌ها را مورد بررسی قرار دهند. در این حالت مزه هندوانه‌ها را می‌توان به سه دسته تقسیم کرد: بد، قابل قبول و خوب. حال که می‌خواهیم مزه هندوانه‌ها را امتحان کنیم، مطالعه به بخشی از کل هندوانه‌ها محدود می‌شود. در اینجا متغیر «مزه» متغیری کیفی است. از آنجا که نمی‌توانیم تمام هندوانه‌ها را مزه مزه کنیم، تنها بخشی از هندوانه‌ها مورد مطالعه قرار می‌گیرند؛ پس باید «نمونه» بگیریم. نسبت هندوانه‌های با مزه «خوب» در نمونه، یک «آماره» است.



محدودیت‌های سرشماری عبارت‌است از:

۱. هزینه زیاد
۲. زمان‌بر بودن
۳. خطای بیشتر در گردآوری داده‌ها
۴. امکان استفاده در بررسی‌های مخرب (واحدهای انتخاب شده را از بین می‌برند) ندارد.



کار در کلاس

یک شبکه تلویزیونی می‌خواهد نسبت دارندگان تلویزیون در شیراز را، که برنامه جدید این شبکه را حداقل یک‌بار در هفته تماشا می‌کنند، بداند. بدین منظور یک گروه ۱۰۰۰ نفری از دارندگان تلویزیون را در این شهر بررسی می‌کند.

الف) داده‌ها و متغیرهایی را که بررسی می‌شوند، مشخص کنید.

جواب: داده‌ها اطلاعات گروه ۱۰۰۰ نفری دارندگان تلویزیون در شیرازند، و متغیر، تماشای تلویزیون است که پاسخ آن «تماشا می‌کند» یا «تماشا نمی‌کند» افراد مورد بررسی است.

ب) آیا این داده‌ها یک نمونه‌اند؟ جامعه آماری کدام است؟

پ) متغیر کمی است یا کیفی؟

ت) چند متغیر کمی را که ممکن است در اینجا جالب باشد، مشخص کنید.

جواب: سن، درآمد،

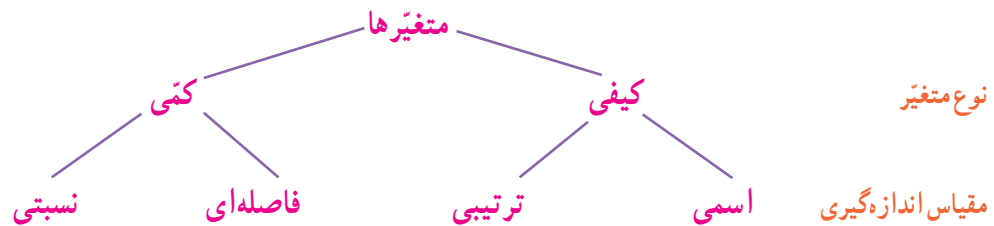
ث) نسبت افرادی در نمونه که برنامه جدید را تماشا می‌کنند، آماره است یا پارامتر؟ (تعداد اعضای مورد نظر تقسیم بر تعداد کل اعضای یک مجموعه را نسبت می‌گوییم.)

مقیاس‌های اندازه‌گیری

داده‌ها را به دو گروه کمی و کیفی تقسیم کردیم. از نگاهی دیگر، می‌توان متغیرهای داده‌ها را در چهار مقیاس اندازه‌گیری دسته‌بندی کرد. اندازه‌گیری در تعریف به معنی ایجاد تفکیک بین افراد یا اشیاء است. دو نوزاد دو قلو را نام‌گذاری می‌کنیم تا آنها را تفکیک کنیم، در واقع اندازه‌گیری کرده‌ایم. بسته به دقتی که این اندازه‌گیری صورت می‌گیرد آن را به چهار مقیاس اسمی، ترتیبی، فاصله‌ای و نسبتی تقسیم می‌کنیم. هدف بررسی این مقیاس‌ها بیانگر نوع محاسبه‌ای است که برای این داده‌ها مناسب است؛ نظیر: ترتیب، محاسبه اختلاف و نسبت گرفتن.

<p>نسبتی: این مقیاس برای داده‌هایی است که قابل مرتب کردن هستند، اختلاف بین مقادیر داده‌ها، و نسبت مقادیر داده‌ها نیز بامعنا است. اغلب متغیرهای فیزیکی مانند نمره، وزن و قد دانش‌آموزان و متغیرهایی که با واژه تعداد شروع می‌شوند در این مقیاس اندازه‌گیری می‌شوند. در این مقیاس صفر به معنی نبود ویژگی در فرد یا شیء است.</p>	<p>فاصله‌ای: این مقیاس به دلیل استفاده از لوازم یا قواعد دقیق اندازه‌گیری ویژگی افراد یا اشیاء به دقت اندازه‌گیری می‌شود. به بیان دیگر مقیاس فاصله‌ای برای داده‌هایی است که قابل مرتب کردن هستند و همچنین، اختلاف بین مقادیر داده‌ها بامعناست. مانند درجه حرارت در شهرهای مختلف برحسب سلسیوس. مقادیری که به دو نفر یا دو شیء داده می‌شود صرفاً بیانگر فاصله بین آنهاست. در نتیجه صفر در این مقیاس قراردادی است. مثلاً اگر دمای یوشنهر ۲۰ و تهران ۱۰ درجه سلسیوس باشد نمی‌توان گفت دمای تهران دو برابر یوشنهر است ولی اختلاف دما ۱۰ می‌باشد.</p>	<p>ترتیبی: این مقیاس با استفاده از الفاظ، ضمن ایجاد تفکیک بین افراد و اشیاء، ارجحیت نیز قائل می‌شود. مقیاس ترتیبی برای متغیرهایی است که قابل مرتب کردن هستند؛ در عین حال محاسبه اختلاف بین مقادیر داده‌ها، یا امکان‌پذیر نیست و یا بی‌معناست، مانند رتبه دانش‌آموزان در یک کلاس. اگر رتبه‌های اول تا سوم معدل ۱۶، ۱۸، ۱۹، ۱۶، ۱۸، ۱۹ می‌دهیم کسب کرده باشند رتبه ۱، ۲، ۳ می‌دهیم و توجه نمی‌کنیم که اختلاف نمرات آنها چقدر است.</p>	<p>اسمی: این مقیاس برای متغیرهایی است که شامل نام‌ها، برحسب‌ها و گروه‌ها می‌شود. در اینجا هیچ معیاری که با آن بتوان داده‌ها را از کوچک به بزرگ مرتب کرد وجود ندارد. مانند گروه خونی انسان‌ها و شماره دانش‌آموزی. کدهای عددی در این مقیاس به‌واقع عدد نیستند بلکه صرفاً برای گروه‌بندی به کار می‌رود.</p>
--	--	--	---

اگر فقط یک متغیر از داده‌ها اندازه‌گیری شده باشد، به جای ذکر «مقیاس متغیر» از واژه «مقیاس داده‌ها» استفاده می‌کنند.



شعارهای برتر
سال جهانی آمار



بهترین برنامه‌ریزی
مستلزم بهترین آمار

با آمار بهتر بفهمیم،
بهتر تصمیم بگیریم

کار در کلاس



نوع متغیر داده‌های زیر را مشخص کنید:

- (الف) محسن، محمود، محمد و میثم همگی اسامی مذکر هستند.
(ب) در یک دبیرستان ۳۱۹ دانش‌آموز فارغ‌التحصیل وجود دارد. احمد رتبه بیست و پنجم، رضا رتبه نوزدهم، صادق رتبه دهم و جواد رتبه چهارم را کسب کرده است و می‌دانیم که رتبه یک، بالاترین است.
(پ) دمای بدن ماهی‌های قزل‌آلای رودخانه هراز (برحسب درجه سلسیوس).
(ت) طول ماهی‌های قزل‌آلا در رودخانه هراز.

تمرین



- داده‌های زیر مربوط به یک نماینده مجلس است. در هریک از سؤالات زیر نوع داده‌ها را مشخص کنید.
(الف) نام نماینده حسین ایرانی است.
(ب) این نماینده ۵۸ سال سن دارد.
(پ) سال‌هایی که این نماینده در مجلس انتخاب شده است، ۱۳۸۶، ۱۳۹۰ و ۱۳۹۴ است.
(ت) مجموع حقوق این نماینده در سال گذشته ۶۰۰,۰۰۰,۰۰۰ ریال بوده است.
(ث) این نماینده در حال بررسی لایحه پیشنهادی حفاظت از منابع آبی کشور است. گزینه‌های موردنظر:
حمایت کامل، حمایت، بی‌طرف، مخالف و کاملاً مخالف است.
(ج) وضعیت تأهل این نماینده: متأهل
(چ) می‌گویند این نماینده در رأی‌گیری لایحه مرتبط با آموزش عمومی، هفتمین نفری است که از آن حمایت کرده است.

آمار چیست؟

علوم تجربی نظیر کشاورزی و پزشکی نیازهای اساسی بشر را رفع می‌کنند. در این علوم یقین کامل برای حل مسئله وجود ندارد، بسیار کاربردی‌اند و مورد استفاده قرار می‌گیرد. یک پزشک معمولاً به صورت دقیق بیماری را تشخیص نمی‌دهد و داروی بیماری نیز به همین وضع دچار است. ولی پزشک بر اساس تجربه حکم به بیماری می‌دهد و دارو تجویز می‌کند و در اکثر مواقع نیز نتیجه می‌گیرد. از سوی دیگر در علوم ریاضی روابط به صورت صد درصد حاکم هستند و هیچ شک و شبهه‌ای به آن راه ندارد. وجود رابطه بین برخی پدیده‌ها در علوم تجربی باعث کشف حقایقی است که موجب پیشرفت آنها می‌شود. معمولاً روابط حاکم بر علوم تجربی را نمی‌توان به صورت ریاضی بیان کرد. علم آمار راهی برای بیان ریاضی چنین پدیده‌هایی است.



فعالیت

به نظر شما یک شهروند در زندگی روزمره خود از اطلاعات بیان شده در دو مثال زیر، چه استفاده‌ای می‌کند؟ این اطلاعات در رسانه‌ها منتشر شده‌اند؟

■ سلامت

- بیشترین آسیب‌دیدگی در منازل، افتادن یا زمین خوردن است.
- پنج درصد افراد به واکسن آنفولانزا واکنش شدید نشان می‌دهند.
- افراد سیگاری دو برابر دیگران در معرض سرطان قرار دارند.
- مصرف روزانه ۵ نوع میوه و سبزی بیشتر ویتامین‌های مورد نیاز بدن را تأمین می‌کند.

■ حمل و نقل

- اولین دلیل تماس با امداد خودرو فراموش کردن کلید داخل خودرو است.
- بیشتر راننده‌های مرد و زن یک کشور خارجی، در پشت چراغ قرمز به ترتیب با بینی خود کلنجار می‌روند و در آینه نگاه می‌کنند.
- راندن بین خطوط راهنمایی در اتوبان‌ها باعث کاهش ۳۰ درصدی شدآمد می‌شود.
- متوسط تعداد کشته‌های تصادفات حدود ۵۰ نفر در روز اعلام شده است.

چگونه این اطلاعات را به دست آورده‌اند؟ آیا تعریف دقیق کلمات پررنگ شده را حدس می‌زنید؟ جامعه، و پارامتر یا نمونه و آماره را در هریک از مثال‌ها مشخص کنید.

تصمیم‌گیری یکی از مهم‌ترین جنبه‌های زندگی ماست. ما بر اساس اطلاعاتمان و ارزش‌هایمان تصمیم‌گیری می‌کنیم. روش‌های آماری برای بررسی این اطلاعات به ما کمک می‌کنند. به علاوه، آمار در شرایطی که با عدم

قطعیت در تصمیم‌گیری روبه‌رویم، هم به یاری ما می‌آید. چگونه به اطلاعات گزارش‌شده در فعالیت قبل دست‌یافته‌اند؟ به‌عنوان مثال، اگر درصدد برآورد نسبت افرادی هستیم که به واکسن آنفلوآنزا واکنش شدید نشان می‌دهند، بدون تزریق بر روی همهٔ افرادی که مایل به انجام آن هستند، آمار روش‌های مناسبی را پیش رویمان می‌گذارد. روش‌های آماری ما را قادر می‌سازند تا با نگاه کردن به اطلاعات به دست آمده از مجموعهٔ کوچکی از افراد یا اقلام، برای گروه‌های بزرگ‌تری از این افراد یا اقلام تصمیم‌گیری کنیم. شیوهٔ تحلیل داده‌ها، در کنار قواعد استنباطی، موضوعات اصلی مطالعات آماری را تشکیل می‌دهند.

شعارهای برتر
سال جهانی آمار



آمار، جهت‌نمای
توسعهٔ پایدار

آمار، راهنمای ما
در تصمیم‌گیری
و برنامه‌ریزی صحیح

به مطالعهٔ نحوه گردآوری، سازمان‌دهی، تحلیل و تفسیر داده‌ها برای استخراج اطلاعات و تصمیم‌گیری، آمار گفته می‌شود.

گفتنی است که روش‌های آماری به‌تنهایی نمی‌توانند معجزه کنند؛ این روش‌ها می‌توانند به ما کمک کنند تا تصمیم‌هایی بگیریم؛ اما نه هر تصمیمی. به یاد داشته باشید که حتی یک روش آماری مناسب، نمی‌تواند دقیق‌تر یا صحیح‌تر از داده‌ها و حقایق اصلی باشد. در نهایت، نتایج آماری باید توسط فردی که نه تنها روش‌ها، بلکه موضوع مورد بحث را کامل درک کرده باشد، تفسیر شود.

تمرین



۱. فرق بین داده و متغیر چیست؟
۲. داده‌های در سطح اسمی، کمی هستند یا کیفی؟
۳. فرق بین آماره و پارامتر چیست؟
۴. در یک جامعه آماری، آیا ممکن است که یک پارامتر تغییر کند؟ اگر سه نمونه با اندازهٔ یکسان از یک جامعه داشته باشیم، آیا می‌توان سه مقدار متفاوت از یک آماره به دست آورد؟
۵. در یک مطالعه از ۱۲۶۱ مشتری غذاخوری‌های گیاه‌خوار، سؤال شده است که برای کدام وعدهٔ غذایی (ناهار یا شام) غذا سفارش داده‌اند؟
(الف) متغیر را مشخص کنید.
(ب) این متغیر کمی است یا کیفی؟
(پ) جامعهٔ آماری در اینجا چیست؟

۶. موضوعات زیر مرتبط است. متغیرهای آنها را در چهار مقیاس: اسمی، ترتیبی، فاصله‌ای و نسبی دسته‌بندی کنید.

الف) مدت زمان پاسخ‌گویی به سؤالات یک امتحان

ب) زمان اولین کلاس

پ) رشته تحصیلی

ت) مقیاس ارزیابی تحصیلی: ضعیف، معمولی و خوب

ث) نمره آخرین آزمون (از ۱۰۰ امتیاز)

ج) سن دانش‌آموز

لقمان و مرد پیاده

خواندنی

روزی لقمان در کنار چشمه‌ای نشسته بود. مردی که از آنجا می‌گذشت از لقمان پرسید: چند ساعت دیگر به ده بعدی خواهیم رسید. لقمان گفت: راه برو، آن مرد پنداشت که لقمان نشنیده است. دوباره سؤال کرد: مگر نشنیدی؟ پرسیدم: چند ساعت دیگر به ده بعدی خواهیم رسید؟ لقمان گفت: راه برو، آن مرد پنداشت که لقمان دیوانه است. برای همین راه خود را گرفت و رفت. زمانی که چند قدمی راه رفت، لقمان به بانگ بلند گفت: ای مرد، یک ساعت دیگر بدان ده خواهی رسید. مرد گفت: چرا اول نگفتی؟ لقمان گفت: چون راه رفتن تو را ندیده بودم، نمی‌دانستم تند می‌روی یا کند. حالا که دیدم دانستم که تو یک ساعت دیگر به ده خواهی رسید. در این داستان ساده و قدیمی تمام اصول آماری رعایت شده است. چرا؟ نکته ظریف این داستان این است که لقمان فقط می‌گوید، راه برو و توضیح دیگری نمی‌دهد. لقمان نمی‌گوید که می‌خواهم راه رفتن تو را ببینم تا از روی آن بگویم چه مدت طول می‌کشد تا به ده برسی، زیرا لقمان فکر می‌کند این اطلاع ممکن است در راه رفتن آن مرد اثر بگذارد و در نتیجه سرعتی که لقمان تخمین می‌زند، سرعت واقعی راه رفتن آن فرد نباشد و در نتیجه زمانی را که تخمین خواهد زد، مدت زمان دقیقی نباشد.

درس ۲

معیارهای گرایش به مرکز

فعالیت



امسال خیرین مدرسه چقدر به مدرسه کمک می کنند؟ مدیر مدرسه بر اساس اطلاعات سال های گذشته خود می گوید: معمولاً خیرین، به طور متوسط، ۱۰ درصد درآمد سالانه خود را به این امر اختصاص می دهند. فرض کنید درآمد ماهیانه حضار در انجمن خیریه این دبیرستان در سال جاری به ترتیب حروف الفبا به صورت زیر باشد:

درآمد (میلیون ریال)	نجمیه	سبحان	رسول	حسنا	جوانه	احمد	آرمان
	۴۰	۱۲	۲۸	۳۲	۳۰	۲۲	۲۵



پس برای پاسخ به سؤال طرح شده باید میانگین این اعداد را محاسبه کنیم. میانگین همان چیزی است که ما به آن معدل می گوئیم. برای محاسبه آن شما تمامی درآمدها را باهم جمع و بر تعداد افراد حاضر تقسیم می کنید. اگر تمامی درآمدها را باهم جمع کنید (۱۸۹)، و بر تعداد افراد تقسیم کنید (۷)، به عدد میانگین ۲۷ میلیون ریال در ماه می رسید. در نتیجه ۱۰ درصد درآمد سالانه برابر است با ۳۲/۴ است.

اگر n داده به صورت x_1, x_2, \dots, x_n داشته باشیم، میانگین آنها را با نماد

\bar{x} نشان می دهیم، که به صورت زیر تعریف می شود:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$



کار در کلاس

۱. میانگین داده‌های ۱۰، ۲۰، ۳۰، ۴۰، ۵۰ چقدر است؟
۲. اگر میانگین داده‌های ۱۰، x ، ۴۰، ۳۰ شود مقدار x چقدر است؟
۳. میانگین اعداد ۱، ۲، ۳، ۴، ۵ چقدر است؟
۴. میانگین اعداد ۲، ۴، ۶، ۸، ۱۰ چقدر است؟
۵. میانگین اعداد ۵، ۶، ۷، ۸، ۹ چقدر است؟
۶. میانگین اعداد ۵۰، ۶۰، ۷۰، ۸۰، ۹۰ چقدر است؟
۷. آیا می‌توانید چند قاعده کلی از تمرین‌های قبل درباره خواص میانگین ذکر کنید.

حال اگر یک میلیارد در با درآمد ماهیانه یک میلیارد ریال به انجمن خیریه دبیرستان ما بیاید، میانگین درآمد حضار چه تغییری می‌کند؟ (بیشتر مردم به او ثروتمند می‌گویند. آمارشناسان او را دورافتاده می‌نامند.) درآمد او میانگین را تا حدود ۱۴۸ میلیون ریال در ماه بالا می‌آورد (دقیقاً ۱۴۸/۶۲۵) و بر اساس شیوه تخمین گذشته، خیرین معادل ۱۷۸ میلیون ریال (دقیقاً ۱۷۸/۳۵) را به دبیرستان کمک خواهند کرد! که غیر واقعی به نظر می‌رسد یا امکان محقق شدن آن ضعیف است.

دورافتاده: مقداری متفاوت با سایر مقادیر داده‌هاست. معمولاً مقدار آن بسیار بزرگ‌تر یا بسیار کوچک‌تر از بقیه داده‌هاست.



در مثال خیریه، داده دورافتاده باعث اشتباه ما در تخمین متوسط داده‌ها شد. میانگین مرسوم‌ترین معیار گرایش به مرکز است که گاهی ممکن است ما را به اشتباه بیندازد، ولی می‌توان از معیار دیگری نیز برای بیان متوسط درآمد استفاده کرد. برای این منظور از میانه که داده وسطی داده‌های مرتب شده است، می‌توان استفاده کرد. پس از مرتب کردن داده‌ها، داده‌ای که تعداد داده‌های بعد از آن با تعداد داده‌های قبل از آن برابر است، میانه است. اگر تعداد داده‌ها زوج باشد، میانه برابر میانگین دو داده وسطی مرتب شده است. برای محاسبه میانه داده‌ها، قبل از ورود میلیارد، داده‌ها را از کوچک به بزرگ مرتب می‌کنیم. وسط ۷ عدد چهارمین عدد است، پس میانه داده‌ها عدد ۲۸ است که با میانگین داده‌ها تفاوت زیادی ندارد.

ردیف	۱	۲	۳	۴	۵	۶	۷
درآمد مرتب شده	۱۲	۲۲	۲۵	۲۸	۳۰	۳۲	۴۰

حال فرض کنید میلیارد در خیریه حضور دارد. میانه برای ۸ عدد وجود ندارد یا به عبارتی ۸ عدد، عدد وسط ندارند. پس بین داده چهارم و پنجم میانگین می‌گیریم. جایگاه چهارم درآمد ۲۸ و جایگاه پنجم آن ۳۰ و در نتیجه میانه جدید درآمد ۲۹ است.



مد، داده‌ای است که بیشترین فراوانی را دارد. فروشندگان پوشاک از معیار گرایش به مرکز مد بسیار استفاده می‌کنند. آنها با آمارگیری‌های خود درمی‌یابند که چه نوع پوشاکی مورد پسند مصرف کنندگان است. از همان پوشاک برای فروش سفارش می‌دهند.

در رأی‌گیری‌ها، اساس تصمیم‌گیری مد است، چون موضوعی که بیشترین فراوانی را داشته باشد، انتخاب می‌شود. در انتخاب رئیس جمهور، نامزدی انتخاب می‌شود که بیشترین فراوانی (رأی) را داشته باشد. برای محاسبه مد فقط کافی است فراوانی داده‌ها را با هم مقایسه کنیم و داده با بیشترین فراوانی مد است. مد ممکن است منحصر به فرد نباشد.

بیشتر افراد معتقدند که عدد ۲۹ میلیون ریال گویای مقدار صحیح متوسط درآمد افراد است و عدد ۱۴۸ میلیون ریال کاملاً بی‌ربط است. پس ما به یک نتیجه می‌رسیم: اگر در داده‌هایمان، دور افتاده وجود داشت - دقیقاً مانند زمانی که یک میلیارد در هوس کار خیر می‌کند - باید از میانه استفاده کنیم.



تمرین

۱. میانه داده‌های ۱, ۱۴, ۸۶, ۲, ۶۸, ۹۹, ۱۰ چقدر است؟

۲. میانه داده‌های ۱, ۱۴, ۸۶, ۲, ۶۸, ۹۹, ۱۰ چقدر است؟

نکته اصلی در اینجا است: اگر در آمار در جایی به یک نتیجه شسته و رفته برخوردید، خیلی احتیاط

کنید. به دست آوردن «یک استنباط درست از داده‌ها تنها چیزی نیست که شما در یک بررسی آماری خواهان آن هستید؛ خواسته دیگر ما برنامه‌ریزی و تصمیم‌گیری است».

تصور کنید که به جای انجمن خیریه با یک بیمارستان سرو کار داریم و داده‌ها به جای درآمد، هزینه‌های جراحی هستند. فرض کنید به جای میلیارد، با یک بیمار روبه‌رویم که گرفتار بعضی از مشکلات بعد از عمل است و مجموع هزینه‌هایش بالغ بر ۲۵۰ میلیون ریال است.



بیمار	هزینه‌های جراحی میلیون ریال	هزینه‌های جراحی مرتب شده
الف	۲	۱
ب	۵	۲
پ	۴۵	۳
ت	۷	۴
ث	۳۵	۵
ج	۳۰	۶
چ	۵۰	۷
ح	۲۵۰	۸



میانگین موزون:

شما قطعاً شنیده‌اید که برخی از نمرات شما ضریب دارند. مثلاً اگر چهار بار از شما آزمون گرفته باشند، نمره آخرین آزمون را ممکن است در ۳ ضرب کنند، یعنی به آن ضریب ۳ بدهند. پس اگر نمرات شما به صورت زیر باشد:

آزمون اول	دوم	سوم	چهارم
۱۶	۱۵	۱۸	۱۷

نمره‌ای که برای این درس در کارنامه شما خواهد آمد میانگین ۶ عدد است یعنی نمره درس مذکور عبارت است از:

$$\frac{۱۶+۱۵+۱۸+۱۷+۱۷+۱۷}{۶} = ۱۶/۶۶$$

میانگین نمره‌ها بدون احتساب ضریب ۱۶/۵ می‌شود.

از میانگین موزون برای معدل دروسی که با واحدهای مختلف (تعداد ساعات متفاوت) در طول سال تحصیلی ارائه می‌شود نیز استفاده می‌شود.

میانگین هزینه‌های جراحی این بیمارستان برابر ۵۳ میلیون ریال و میانه آن برابر ۳۲/۵ میلیون ریال است. اگر شما مدیر این بیمارستان بودید، کدام عدد برایتان از همه مهم‌تر بود؟ عدد ۳۲/۵ میلیون نشان دهنده هزینه‌های معمول نگهداری یک بیمار است، اگر بخواهیم بودجه سال بعد را با فرض این عدد بنویسیم، به احتمال زیاد با کسر بودجه مواجه می‌شویم. با قدری تفکر به میانگین و میانه، دلیل اینکه افراد از بیمه تأمین اجتماعی استفاده می‌کنند، روشن می‌شود؛ (میانه هزینه‌های سالیانه درمان معمولاً بسیار کمتر از حق بیمه پرداختی است، اما با نگاه به میانگین سالانه هزینه‌ها، حساب می‌کنم که معامله خوبی کرده‌ام) و من همیشه کمربند ایمنی را می‌بندم، اگرچه میانه تعداد زخمی‌ها در هر سفر با خودرو صفر باشد.

این نتیجه را نیز به یاد داشته باشید: گاهی هیچ آمار درست یا غلطی وجود ندارد، و همه چیز به اینکه شما چرا می‌خواهید از آنها استفاده کنید، بستگی دارد.



کار در کلاس

با سه واژه: داده دورافتاده، میانه، و میانگین، ۵ جمله زیر را کامل کنید.

۱. آنچه اکثر مردم «حد وسط» می‌نامند، نزد آمارشناسان به معروف است. برای محاسبه، به داده‌های خود به عنوان فهرستی از اعداد نگاه کنید؛ همه اعداد را باهم جمع کنید و بر تعدادشان تقسیم کنید.

۲. در واقع نقطه وسطی فهرست اعداد مرتب شده است. نیمی از اعداد مقادیر بالاتر از و نیمی دیگر مقادیر پایین تر از قرار دارند.

۳. زمانی با مواجه هستیم که مشاهده‌ای داشته باشیم که از الگوی داده‌هایمان پیروی نکند.

۴. وقتی با مواجه هستید، معمولاً بازتاب بهتری از داده‌ها می‌دهد تا

۵. به طور کلی، برای برنامه‌ریزی و تصمیم‌گیری، بهتر از است.

شخصی دست راست خود را در یخچال و دست چپ خود را در فر کرده است. زمانی که از او درباره احساسش سؤال می‌شود پاسخ می‌دهد: «به طور متوسط خوب هستم». نتیجه اخلاقی این لطیفه این است که یک عدد به تنهایی، معمولاً مجموعه‌ای از داده‌ها را به خوبی توصیف نمی‌کند. بنابراین، اندیشه خوبی است که علاوه بر معیار گرایش به مرکز معیاری را نیز گزارش دهیم که نوسان و تغییر داده‌ها را هم بیان کند.

خواندنی

درس ۳

معیارهای پراکندگی

فعالیت



می‌خواهیم کلاس‌های دهم سه دبیرستان را بر اساس نتایج آزمون جامعی که هم‌زمان بین دانش‌آموزان ممتاز برگزار شده است، رتبه‌بندی کنیم. از هر دبیرستان ۱۰ نفر به تصادف انتخاب شده‌اند. نمرات آزمون جامع هر سه کلاس از ۱۲۰ نمره است.

داده‌های مربوط به سه کلاس دهم:

الف) {۶۵, ۷۵, ۷۳, ۵۰, ۶۰, ۶۴, ۶۹, ۶۲, ۶۷, ۸۵}

ب) {۸۵, ۷۹, ۵۷, ۳۹, ۴۵, ۷۱, ۶۷, ۸۷, ۹۱, ۴۹}

پ) {۴۳, ۵۱, ۵۳, ۱۱۰, ۵۰, ۴۸, ۸۷, ۶۹, ۶۸, ۹۱}



میانگین سه کلاس را محاسبه کنید. به نظر شما پراکندگی نمرات در کدام کلاس بیشترین و در کدام کلاس کمترین است؟ برای پاسخ به این سؤال، داده‌ها را روی سه محور موازی نمایش دهید. کدام دبیرستان نتایج بهتری با اطلاعات داده شده کسب کرده است؟ چرا؟ مثلاً اگر والدین برای ثبت نام فرزند بازیگوش خود در سال آینده بخواهند بر اساس این نتایج یکی از مدرسه‌ها را انتخاب کنند، مدرسه‌ای بهتر است که پراکندگی نمرات در آن کمتر است در مقابل برای فرزندی که به قصد شرکت در المپیاد می‌خواهد مدرسه را انتخاب کند، مدرسه‌ای بهتر است که پراکندگی نمرات در آن بیشتر باشد. چرا؟

یک معیار معمول سنجش میزان پراکندگی که نشان دهنده تغییرات داده‌های مورد مطالعه است، **انحراف معیار** است. انحراف معیار را با به‌کارگیری یک فرمول از داده‌ها به دست می‌آوریم (اختلاف هر عدد با میانگین داده‌ها را حساب کنید؛ به توان ۲ برسانید؛ میانگین اعداد حاصل را به دست آورید؛ و سپس جذر بگیرید.) توان دوم انحراف معیار را **واریانس** می‌نامند. داده‌ها در شکل زیر رسم شده‌اند. همچنین انحراف معیار نمرات هر سه کلاس محاسبه و در شکل نشان داده شده است.

اگر n را به صورت x_1, x_2, \dots, x_n داشته باشیم، انحراف معیار آنها را با نماد σ (سیگما) نشان می‌دهند، که به صورت زیر تعریف می‌شود:

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

کار در کلاس



۱. اگر داده‌ها با هم برابر باشند، انحراف معیار برابر با صفر خواهد شد؛ چرا؟
۲. انحراف معیار نمرات آزمون جامع سه کلاس را می‌توانید با تکمیل جدول زیر محاسبه کنید.

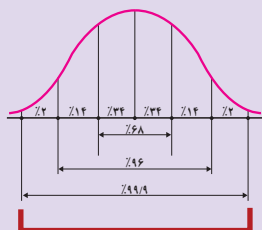
مشاهدات			انحراف مشاهدات از میانگین			انحراف مشاهدات از میانگین به توان ۲		

مشاهده می‌شود که انحراف معیار نه تنها نظر شما در خصوص کلاس با بیشترین پراکندگی نمرات را تأیید می‌کند، بلکه به شما امکان مقایسه کمی آنها را نیز می‌دهد. حال می‌توانید درستی پاسخ‌های خود را بررسی کنید.

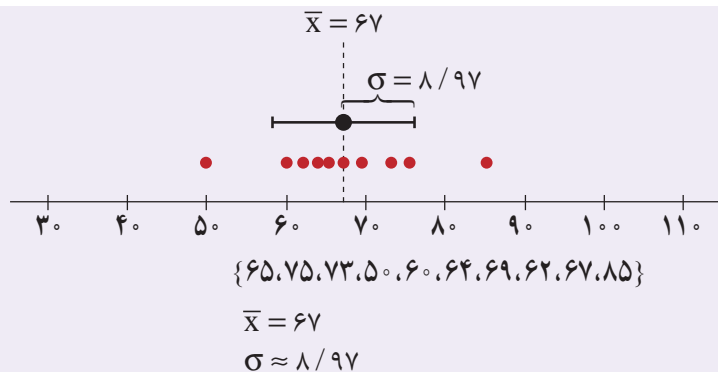


پراکندگی در خم بهنجار (منحنی نرمال)

اگر تعداد داده‌ها زیاد باشد، بافت‌نگاشت (هیستوگرام) آنها را می‌توان به صورت یک خم تصور کرد. اگر این خم به صورت یک «زنگ» باشد، به آن خم بهنجار گفته می‌شود. خداوند بسیاری از پدیده‌ها را بهنجار آفریده است. مثلاً طول قد یا وزن دانش‌آموزان بهنجار اند. یعنی بافت‌نگاشتی زنگی شکل دارند. خواص این خم کمک زیادی برای تصمیم‌گیری بر اساس داده‌ها یا کسب اطلاعات از آنها می‌کند. به یاد دارید که گفته شد اگر σ انحراف معیار داده‌ها باشد چند درصد داده‌ها به طور تقریبی بین میانگین که در واقع گرانگاه یا مرکز ثقل بافت‌نگاشت است و ضرابی از انحراف معیار قرار دارد. این اطلاعات را به طور مختصر در شکل زیر ملاحظه می‌کنید (مقادیر روی نمودار تقریبی هستند):



کلاس الف)

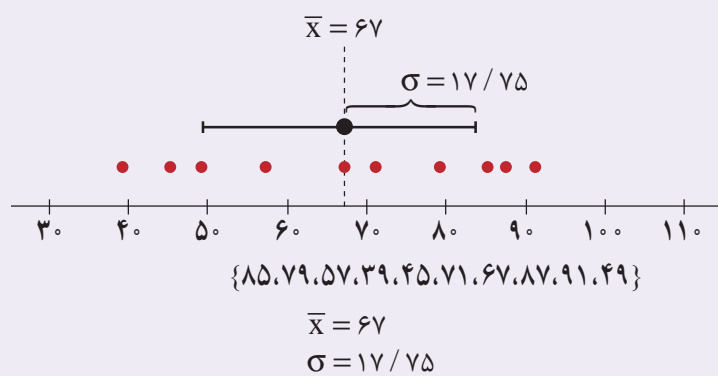


داده‌ها:

میانگین:

انحراف معیار:

کلاس ب)

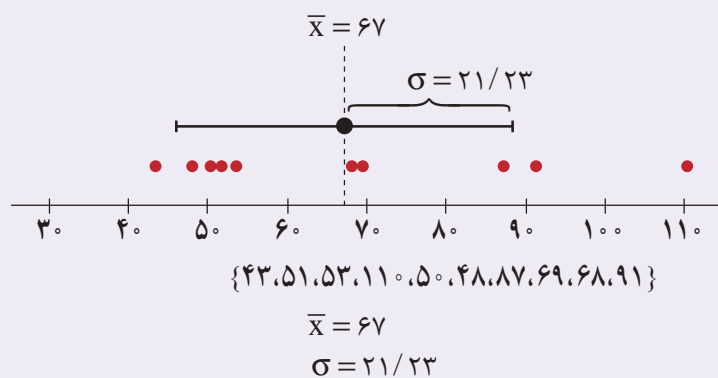


داده‌ها:

میانگین:

انحراف معیار:

کلاس ب)



داده‌ها:

میانگین:

انحراف معیار:

فعالیت



حتماً بر روی قوطی‌ها یا پاکت‌های مواد غذایی یا بهداشتی اعدادی به صورت زیر مشاهده کرده‌اید. به نظر شما هریک از این اعداد چه چیزی را نشان می‌دهند.

250 ± 10 گرم

عدد ۲۵۰ وزن خالص ماده غذایی است که قرار است به دست مصرف کننده برسد. ولی معمولاً این چنین نیست. کمی خطا اجتناب ناپذیر است. این خطا چقدر است و چگونه آن را محاسبه کنیم؟ اگر تعداد زیادی از محتوای پاکت‌ها یا قوطی‌های یک تولیدی را وزن و انحراف معیار آنها را محاسبه کنیم، دو برابر آن عدد ۱۰ است. یعنی عدد ۱۰ دو برابر انحراف معیار مواد غذایی در بسته‌بندی‌های مختلف است. به عبارت دقیق‌تر اگر ما ۱۰۰ قوطی حاوی این ماده غذایی را وزن کنیم، انتظار داریم حداقل ۹۶ تا از آنها وزنی بین ۲۴۰ تا ۲۶۰ گرم داشته باشند. درج چنین اعدادی بر روی کالاها از الزامات استاندارد هر کشوری است. چگونه به عدد ۹۶ رسیده‌اند؟

برای پی بردن به میزان تغییراتی که داریم، معمولاً با محاسبه مقدار انحراف از میانگین اطلاعات دقیق‌تری در خصوص میانگین و انحراف معیار به دست خواهیم آورد. برای این منظور از یک قانون مشهور در آمار استفاده می‌کنیم: تقریباً ۹۶ درصد از داده‌ها بین «دو انحراف معیار از میانگین» هستند. این جمله یعنی اگر بازه‌ای تعریف کنیم که ابتدای آن $\bar{x} = 25$ و انتهای آن $\bar{x} = 25$ باشد، تقریباً ۹۶ درصد داده‌ها در این فاصله قرار دارند این مانند آن است که بگوییم، ۴ درصد نمره‌ها بیش از دو برابر انحراف معیار از میانگین فاصله دارد. البته مشابه همین قانون در شکل‌های قبل برای یک برابر انحراف معیار نمایش داده شده است. در این حالت قانون قبل به «تقریباً ۶۸ درصد از مشاهدات بین یک انحراف معیار از میانگین هستند»، تغییر می‌کند. دو برابر انحراف از میانگین را روی همان شکل رسم کنید.



کار در کلاس

جدول زیر را تکمیل کنید.

$(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$	تقریباً ۹۹/۹ درصد از مشاهدات بین سه برابر انحراف معیار از میانگین قرار دارند
$(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$	تقریباً مشاهدات بین انحراف معیار از میانگین قرار دارند
.....	تقریباً مشاهدات بین یک برابر انحراف معیار از میانگین قرار دارند



فعالیت

می‌خواهیم با همان مثال خیریه درس قبل بحث را ادامه دهیم. اگر انحراف معیار مجموعه داده‌ها کوچک باشد، بدین معناست که درآمد همه افراد به هم نزدیک است؛ اگر انحراف معیار بزرگ باشد، بدین معناست که درآمد افراد آن انجمن بسیار متفاوت است. انحراف معیار درآمد اعضا به این صورت محاسبه می‌شود:

به کمک جدول زیر این کار را انجام دهید.

انحراف مشاهدات از میانگین به توان ۲		انحراف مشاهدات از میانگین		مشاهدات	
بعد از ورود میلیارد	قبل از ورود میلیارد	بعد از ورود میلیارد	قبل از ورود میلیارد	بعد از ورود میلیارد	قبل از ورود میلیارد
				۴۰	۴۰
				۱۲	۱۲
				۲۸	۲۸
				۳۲	۳۲
				۳۰	۳۰
				۲۲	۲۲
				۲۵	۲۵
				۱۰۰۰	

زمانی که میلیارد در وارد انجمن خیریه می‌شود، انحراف معیار داده‌ها برابر است با
این دو مقدار باهم اختلاف زیادی دارند. چرا و چه باید کرد؟ زمانی را به خاطر بیاورید که ما معیار گرایش به مرکز را برای این داده‌ها محاسبه کردیم. ما برای رفع آن مشکل به جای میانگین، میانه را به کار بردیم. مشابه همان کار را می‌توانیم انجام دهیم. معیار پراکندگی که تعریف می‌شود، انحراف معیار نیست. ما آن را دامنه میان چارکی، که با IQR نشان می‌دهیم، می‌نامیم. «میانه» جایی در وسط داده‌هاست؛ به‌طور مشابه، چارک‌ها هم یک چهارم و سه چهارم مشاهدات هستند. به عبارت ساده‌تر اگر برای داده‌های مرتب شده قبل از میانه، یک میانه حساب کنیم، همان چارک اول است و به همین صورت میانه داده‌های بعد از میانه، چارک سوم خواهد بود. برای سادگی از نمادهای Q_1 ، Q_2 و Q_3 به ترتیب برای چارک اول، میانه (چارک دوم)، و چارک سوم استفاده می‌شود.

جدول زیر را برای مثال خیرین مدرسه تکمیل کنید.

Q_1 چارک اول	Q_2 میانه	Q_3 چارک سوم	$IQR = Q_3 - Q_1$ دامنه میان چارکی	σ انحراف معیار

این اعداد امکان مشاهده چند چیز را به شما می‌دهند: میلیارد در هم‌ریختگی میانگین درآمد و انحراف معیار بوده است. اما میانه و دامنه میان چارکی همچنان ثابت ماندند. این یکی از دلایلی است که ما می‌گوییم، اگر داده دور افتاده داشته باشیم، از میانه (و در نتیجه دامنه میان چارکی) به جای میانگین استفاده کنید، در این صورت نتیجه بهتری از داده‌ها به دست می‌آورد. زمانی که تعداد داده‌ها زیاد باشد، به نحو دیگری می‌توان از تفاوت معیارهای گرایش به مرکز و پراکندگی نتیجه بگیرید که داده دور افتاده وجود داشته است.

شعارهای برتر
سال جهانی آمار



آمارهای دقیق
پشتوانه توسعه پایدار

آمار، ابزاری علمی
در توسعه مبتنی بر
دانشی

دلیل دیگر استفاده از میانه و دامنه میان چارکی تفسیر ساده آنها در مقابل تفسیر میانگین و انحراف معیار است. اگر نتایج مطالعات مربوط به سرطان را مطالعه کنیم، اولین چیزی که در گزارش‌ها مشاهده می‌کنیم، مشخصات عمومی بیماران مورد مطالعه است: آنها چند سال سن دارند؟ نسبت مردان به زنان چقدر است؟ چند نفر بیمارشان زود تشخیص داده شده و در مراحل ابتدایی است و چند نفر بیمارشان عود کرده است؟ اگر در گزارش، میانگین و انحراف معیار را ذکر کنیم، هر خواننده‌ای می‌تواند از روی اطلاعات داده شده با توجه به دو قانون ذکر شده نحوه توزیع سنین بیماران را به دست آورد. اما نکته در اینجاست که، آنها این کار را نخواهند کرد. به ندرت می‌توانید یک پزشک متخصص سرطان بسیار پرمشغله را در حال فکر کردن به این مسئله ببینید که، «خوب میانگین $64/3$ و انحراف معیار $9/8$ است؛ 68 درصد از بیماران در فاصله یک برابر انحراف معیار از میانگین هستند؛ این یعنی، $9/8 \pm 64/3$ ، که می‌شود، یک لحظه صبر کنید، ماشین حساب کجاست؟» شما تنها می‌توانید با نگاهی سریع به میانه و دامنه میان چارکی تصور خوبی نسبت به توزیع داده‌های پیش رویتان به دست آورید. به عبارت دیگر، میانه و دامنه میان چارکی در توصیف مجموعه داده‌ها بسیار مفیدند و این دقیقاً همان کاری است که ما از آنها انتظار داریم انجام دهند:

تمام آمارهایی که در اینجا ذکر شد (میانگین‌ها، میانه‌ها، انحراف‌های معیار، دامنه‌های میان چارکی) تحت عنوان آمار توصیفی شناخته می‌شوند.



کار در کلاس

جملات زیر را کامل کنید:

۱. میانگین‌ها و میانه‌ها برای توصیف مجموعه داده‌ها مفیدند. و انواعی از معیارهای گرایش به مرکزی هستند.
۲. شما معمولاً نه تنها معدل یک مجموعه را می‌خواهید، بلکه میزان تغییرات حوالی آن نقطه را هم نیاز دارید که آن معیار است.
۳. معیار پراکندگی که معمولاً با میانگین بیان می‌شود، است.
۴. معیار پراکندگی که معمولاً با میانه بیان می‌شود، نام دارد.
۵. و اطلاعات سریعی درباره داده‌ها بدون نیاز به هرگونه محاسبه می‌دهند.
۶. آماری که برای توصیف یک مجموعه داده، میانگین‌ها و میانه‌ها، انحرافات معیار و دامنه‌های میان چارکی به کار می‌رود، دارد.
۷. 50 درصد داده‌ها قبل از و 50 درصد داده‌ها بعد از قرار دارند.
۸. 75 درصد داده‌ها قبل از یا بعد از قرار دارند.
۹. 25 درصد داده‌ها قبل از یا بعد از قرار دارند.
۱۰. 50 درصد داده‌ها بین و قرار دارند.

۱۱. تقریباً ۹۶ درصد مشاهدات در فاصله از میانگین هستند.

۱۲. تقریباً ۶۸ درصد مشاهدات در فاصله از میانگین هستند.



تمرین

۱. یک نمونه ۲ تایی از بین اعداد ۱ تا ۶ انتخاب کنید. اگر بخواهید این نمونه حتماً تصادفی باشد، چه راهی را پیشنهاد می‌کنید. اگر بخواهید اعضای انتخابی این نمونه تصادفی، تکراری نباشند، چگونه این کار را انجام می‌دهید؟

۲. سؤال ۱ را برای اعداد ۱ تا ۳۶ تکرار کنید. (راهنمایی: می‌توانید از پرتاب دو تاس به عنوان روشی برای انتخاب اعداد ۱ تا ۳۶ استفاده کنید)

۳. مجموعه افراد فامیل درجه اول و درجه دوم خود را بنویسید و آنها را شماره‌گذاری کنید. سپس یک نمونه ۴ تایی از این جامعه انتخاب کنید. موضوع‌های زیر را در نظر بگیرید. درباره بهترین روش جمع‌آوری داده برای این فرضیه‌ها تصمیم بگیرید:

الف) بیشتر مردم فکر می‌کنند «حداکثر سرعت در اتوبان‌ها باید تعیین شود».

ب) آبی، رنگ مورد علاقه بیشتر مردم برای ماشین است.

پ) در زمان مطالعه، گوش دادن به موسیقی کلاسیک به یادگیری کمک می‌کند.

ت) بیشتر تصادفات اتومبیل‌ها را رانندگان با سن کمتر از ۲۵ سال مرتکب می‌شوند.

ث) رژیم گرفتن، باعث کاهش هوش می‌شود.

۴. دو موضوع برای هر یک از روش‌های جمع‌آوری داده بیان کنید.

الف) از طریق مصاحبه ب) از طریق مشاهده پ) از طریق دادگان ث) از طریق پرسش‌نامه

۵. کدام یک از نمونه‌گیری‌های زیر، یک نمونه‌گیری تصادفی است. در هر یک واحدهای آماری، جامعه و نمونه را مشخص کنید.

الف) با تمام پلیس‌های یک پاسگاه پلیس برای پیدا کردن نظر آنها راجع به تخلفات مصاحبه شد.

ب) با بچه‌هایی که وارد یک پارک بازی می‌شدند، پنج در میان مصاحبه شد تا وسیله بازی مورد علاقه کودکان بررسی شود.

پ) برای بررسی PH شامپوهای تولیدی یک کارخانه، شامپوها را صد در میان مورد آزمایش قرار دادیم.

۶. برای تحقیقات و بررسی‌های زیر جامعه را مشخص کرده و روش نمونه‌گیری پیشنهاد کنید.

الف) پیدا کردن درصد بیکاران استان مازندران. ب) مردان کدام شامپو را بیشتر ترجیح می‌دهند؟

۷. می‌خواهیم مدت زمانی را که دانش‌آموزان کلاس شما در طول یک هفته صرف مطالعه کتاب‌های غیردرسی می‌کنند، آمارگیری کنیم.

الف) در این آمارگیری جامعه را مشخص کنید.

شعارهای برتر
سال جهانی آمار



آمار دقیق،

برنامه‌ریزی بهتر،

آینده روشن

آمار سیمای واقعیت،

ابزار مدیریت

- (ب) یک روش نمونه گیری برای انتخاب نمونه معرفی کنید.
- (پ) از چه روشی برای گردآوری اطلاعات استفاده می کنید؟
- (ت) اندازه (تعداد) این جامعه چقدر است؟
- (ث) اندازه نمونه مورد بررسی چقدر است؟ آیا این تعداد برای بررسی مورد نظر مناسب است؟ در صورت مناسب نبودن، روش نمونه گیری و نمونه خود را اصلاح کنید.
- (ج) متغیر مورد مطالعه در این مسئله چیست؟ توضیح دهید.
- (چ) متغیر مورد مطالعه از چه نوعی است؟
۸. نوع هر یک از متغیرهای زیر را مشخص و بهترین مقیاس اندازه گیری آنها را ذکر کنید.
- (الف) رنگ اتومبیل های موجود در یک نمایشگاه اتومبیل (ب) درجه حرارت کلاس شما در روزهای سال
- (پ) گنجایش آب یک باری مخزنی (تانکر) (ت) تعداد شکایات رسیده شده به یک پاسگاه پلیس
- (ث) درآمد دانشجویان شاغل به کار (ج) وضعیت تأهل کارمندان یک شرکت
- (چ) سن دانشجویان شرکت کننده در یک دوره هنری
۹. میزان پرداخت حقوق در یک شرکت خصوصی در جدول زیر آمده است :

سِمَت	مدیر شرکت	معاون	حسابدار	بازاریاب	۲ نفر فروشنده	منشی	مستخدم
حقوق میلیون ریال	۱۰۰	۶۰	۳۰	۲۰	۲۰	۱۲	۸

- مدیر شرکت در آگهی دعوت به همکاری در روزنامه اعلام می دارد که میانگین پرداخت حقوق در این شرکت بیشتر از ریال است. آیا این رقم می تواند دلیلی برای پرداخت حقوق بالاتر در این شرکت باشد؟ چرا؟ کدام یک از معیارهای گرایش به مرکز برای نشان دادن وضع پرداختی کارکنان این شرکت مناسب تر است؟
۱۰. نمرات درس فیزیک دانش آموزی در طول سال برابر است با:

۱۹	۱۸	۷	۱۸	۱۹	۱۷
----	----	---	----	----	----

- (الف) میانه و میانگین را برای نمرات این دانش آموز حساب کنید.
- (ب) کدام یک از شاخص های فوق، بیانگر وضع این دانش آموز در درس فیزیک است؟
- (پ) اگر معلم درس فیزیک این دانش آموز، برای جبران نمره ۷، امکان امتحان مجدد را به او بدهد، برای اینکه میانگین وی در این درس بیشتر از ۱۸ شود، او در این امتحان چه نمره ای باید کسب کند؟
۱۱. دو دسته داده زیر قیمت کالایی را در دو بازار جداگانه برحسب هزار ریال نشان می دهد.

بازار الف	۱۰	۱۱	۱۰	۹	۱۲	۱۱	۱۰	۱۲	۹	۱۳	۸
بازار ب	۱۰	۹	۱۱	۱۰	۱۰	۱۱	۹	۱۰	۸	۱۳	۱۰

- (الف) در کدام دسته پراکندگی بیشتر است؟
- (ب) دامنه تغییرات را محاسبه کنید.
- (پ) آیا دامنه تغییرات با پراکندگی مشاهده شده در «الف» همخوانی دارد؟
- (ت) ترجیح می دهید از کدام بازار خرید کنید؟ چرا؟
- (ث) اگر داده ها را در اختیار نداشته باشید، فقط به صرف داشتن دامنه تغییرات می توانید تصمیم گیری کنید؟

فصل پنجم

نمایش داده‌ها

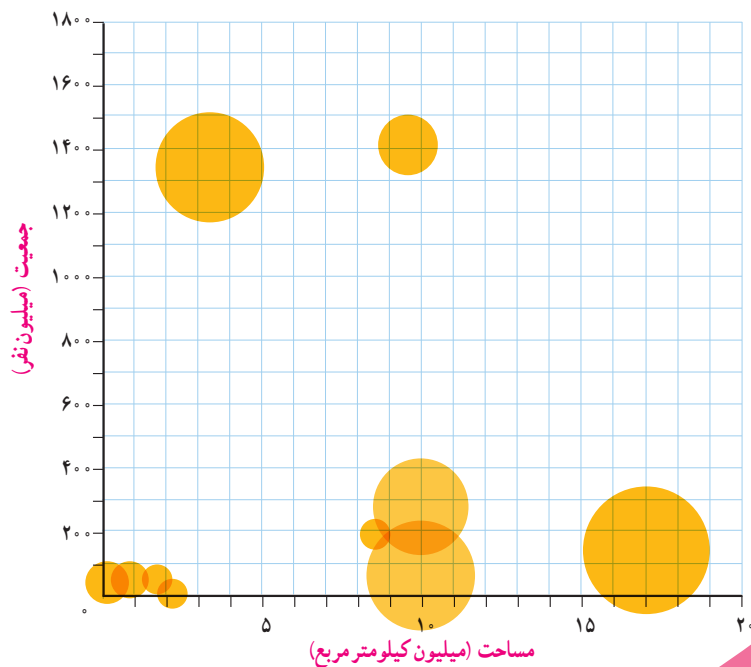
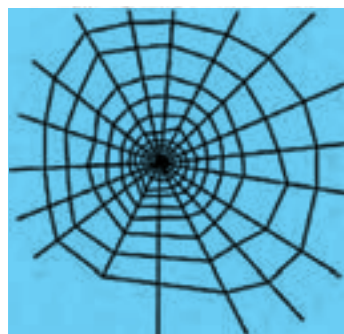


نمودارهای یک متغیره

درس ۱

نمودارهای چندمتغیره

درس ۲



درس ۱

نمودارهای یک متغیره



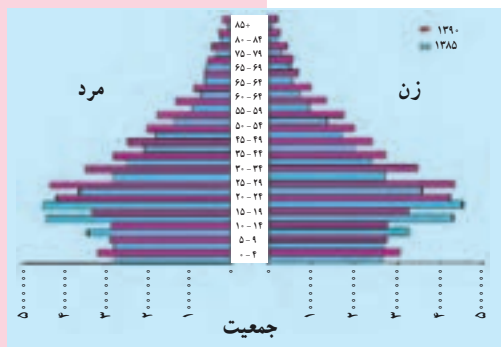
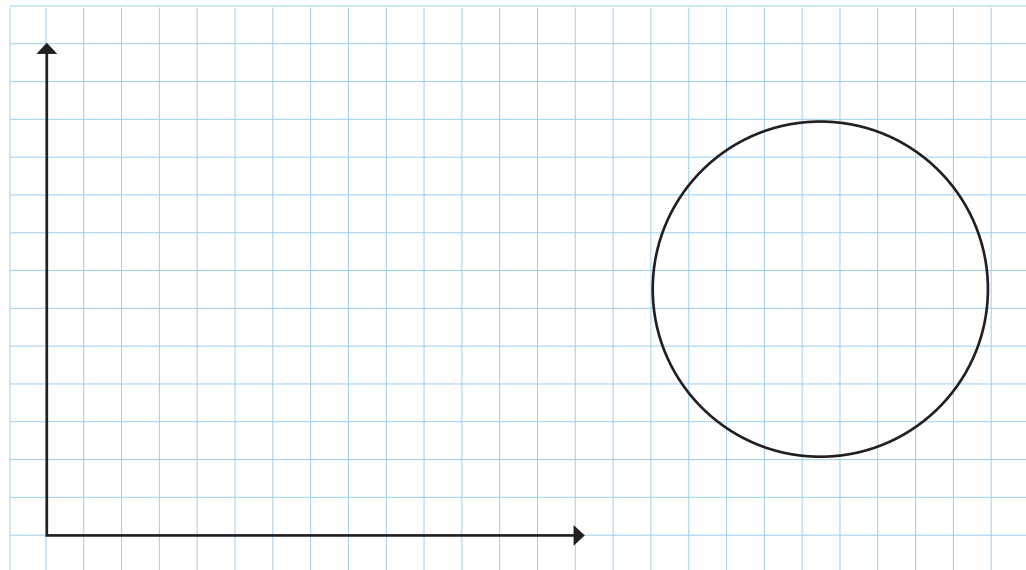
فعالیت

در کلاس قرار شد هر دانش‌آموز با توجه به داده‌های موجود در رسانه‌ها، درباره یک موضوع اجتماعی یا فرهنگی گزارشی در قالب یک شکل ارائه کند. مثلاً حسن در مورد موضوع علل آسیب‌دیدگی ۲۰۰ نفر در منازل اطلاعاتی کسب کرده بود. او گزارش خود را در یک شکل خلاصه کرده است. او با اطلاعاتی که در اختیار داشته است، ابتکار جالبی به خرج داده و نمودار زیر را رسم کرده است. از این نمودار چه اطلاعاتی می‌توان کسب کرد؟ به عنوان مثال، بیشترین آسیب‌دیدگی در منازل بر اثر افتادن یا زمین خوردن است.



معلم از موضوع انتخابی و نمایش ترتیبی مکعب‌ها برحسب طول آنها رضایت کامل داشت؛ ولی از دانش‌آموزان درخواست کرد که اگر می‌توانند پیشنهادی برای بهتر شدن آن ارائه کنند. آیا شما قبل از دیدن نظر‌ها می‌توانید چند پیشنهاد برای بهبود نمودار ارائه کنید؟

نمودار میله‌ای انواع آسیب‌دیدگی را، برحسب تعداد رسم کنید. مراحل رسم نمودار دایره‌ای را به یاد بیاورید و آن را در چند گام خلاصه کنید. نمودار دایره‌ای درصد آسیب‌دیدگی‌ها را نیز رسم کنید.



جمع‌بندی پیشنهادها براساس آموخته‌های ما در سال‌های گذشته، به صورت زیر است:

- رسم مکعب مستطیل به جای مستطیل خوب نیست، چون ممکن است اطلاعات دقیق را به بیننده انتقال ندهد.
- برای متغیرهای اسمی نمودار میله‌ای بهتر از بافت‌نگاشت (هیستوگرام) است.
- زمانی که درصد را گزارش می‌کنیم، بهتر است از نمودار دایره‌ای استفاده کنیم تا بهتر بتوانیم مقایسه انجام دهیم. البته به شرطی که بیشتر از ۶ مقدار نداشته باشیم.

مقایسه هرم سنی جمعیت ایران
سال‌های ۱۳۸۵ و ۱۳۹۰

کار در کلاس



در کلاس درس خود متوسط مصرف روزانه انواع میوه‌ها یا سبزی‌ها را به یکی از روش‌های فصل قبل جمع‌آوری کنید و آنها را با یک نمودار مناسب نمایش دهید. مثلاً اگر در روز شنبه ۳، یکشنبه ۶، دوشنبه ۵، سه‌شنبه ۴، چهارشنبه ۳، و پنجشنبه ۷ و جمعه ۸ نوع میوه یا سبزی خورده باشید، میانگین آنها برابر ۵/۱۴ است که آن را ۵ در نظر می‌گیریم.

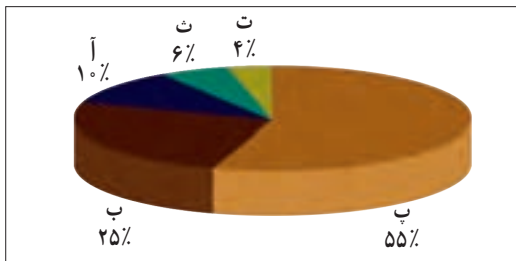
نمایش داده‌ها

نمایش داده‌ها، روشی برای کمک به استخراج اطلاعات از داده‌ها با حس بینایی است.

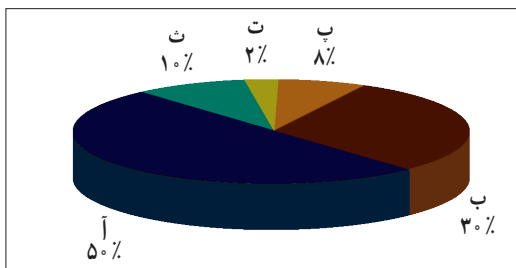


فعالیت

درصد تعداد کاربران اینترنت در پنج کشور



درصد کاربران اینترنت نسبت به کل جمعیت در پنج کشور

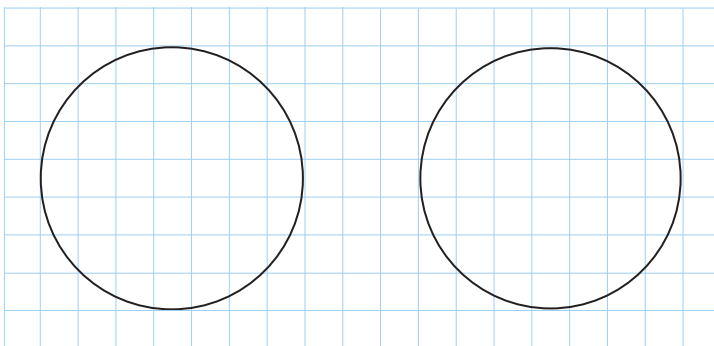


از نمودار دایره‌ای مقابل، برای نشان دادن وضعیت استفاده از اینترنت در پنج کشور استفاده شده است. کدام کشور دسترسی بیشتری به اینترنت دارد؟ آیا اطلاعات این نمودار برای پاسخ به سؤال قبل کافی است؟ چه اطلاعاتی از این نمودار کسب می‌کنید؟

اگر عامل جمعیت را در نظر بگیریم، واقعیت میزان توسعه اینترنت و میزان استفاده از آن در کشورهای مختلف را می‌توان به صورت روبه‌رو نشان داد.

آیا برداشت شما از نمودار قبل تغییر کرد؟

به نظر شما رسم نمودار دایره‌ای به صورت دو بعدی که در سال‌های گذشته با آن آشنا شده‌اید، بهتر است یا سه بعدی؟ مانند دو نمودار قبلی.



انتقال اطلاعات از طریق نمایش داده‌ها بسیار جذاب‌تر و سریع‌تر از ارائه همان اطلاعات در جدول‌هاست. نمودارها باید به گونه‌ای رسم شوند که از آنها سوء برداشت نشود و بی‌طرفی را حفظ کنند. می‌گویند ارزش یک تصویر به اندازه هزار کلمه است.

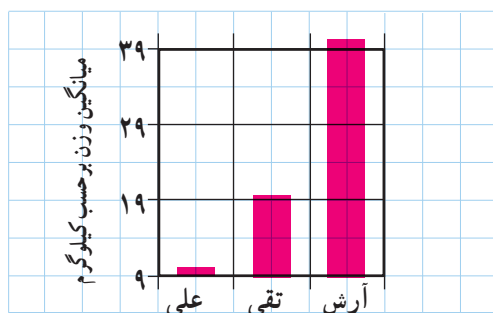
خواندنی



کار در کلاس

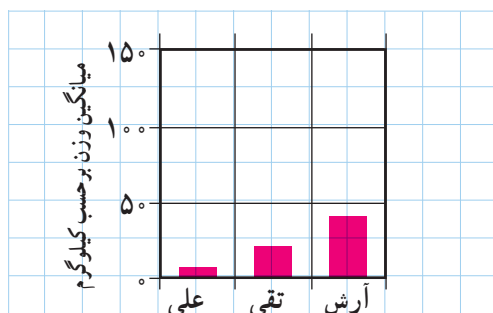
بازی با نمودارها

۳. اگر نقطه شروع محور عرض‌ها را از صفر به عدد ۹ تغییر دهیم، چه تغییری در نتیجه حاصل می‌شود؟



در واقع اعداد همان اعداد قبلی‌اند، اما محور عرض‌ها تغییر کرده است. اکنون این‌طور به نظر می‌رسد که کدو تنبل‌های می‌رسد

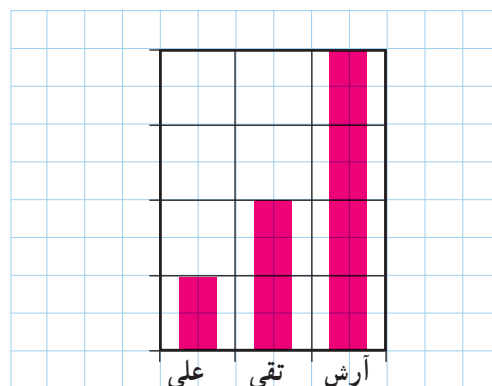
۴. اینک اگر کسی بخواهد دیگران را متقاعد کند که همه کدو تنبل‌ها حدوداً به یک اندازه‌اند، چه کاری می‌تواند بکند؟ به این نمودار نگاه کنید:



اعداد تغییری نکرده‌اند؛ اما محور عمودی دوباره تغییر کرده است. چه تغییری کرده است؟

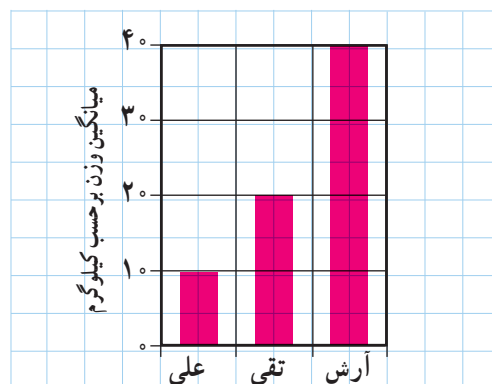
به کارگیری غلط نمودارها، راه ساده‌ای برای گمراه کردن افراد است. افراد مایل‌اند به نمودارها به عنوان روشی سریع برای ارزیابی مجموعه‌ای از اعداد بنگرند. اما مراقب باشید که فریب نخورید.

۱. بیایید از کدو تنبل‌های باغ‌های آرش، تقی و علی استفاده کنیم. اولین نمودار به صورت زیر است:

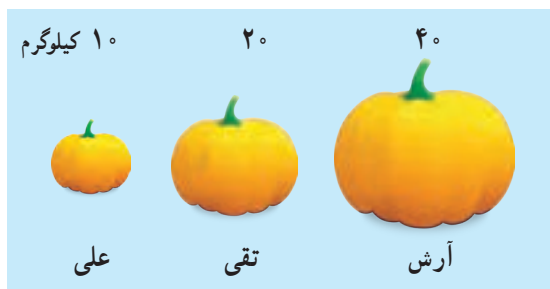


این نمودار چه چیزی را نمایش می‌دهد؟ وزن، حجم، پهنای ارتفاع کدو تنبل‌ها را؟

۲. نمودار زیر چه تفاوتی با نمودار بالا دارد؟



این نمودار به چه اطلاعات بیشتری اشاره می‌کند؟



گاهی برای نمایش داده‌ها از یک تصویر استفاده می‌شود، مانند شکل روبه‌رو.

این تصویر، داده‌ها را از شکل طبیعی خارج می‌کند. برای نمایش تفاوت میانگین وزن‌ها، این تصویر ارتفاع هر کدو تنبل را، برای نشان دادن وزنش، تغییر می‌دهد. ارتفاع

کدو تنبل تقی (با وزن ۲۰ کیلوگرم) دو برابر ارتفاع کدو تنبل علی (۱۰ کیلوگرم) است؛ ارتفاع کدو تنبل آرش (۴۰ کیلوگرم) دو برابر ارتفاع کدو تنبل تقی و چهار برابر ارتفاع کدو تنبل علی است. آیا این شکل اطلاعات مورد نظر را به درستی به بیننده منتقل می‌کند؟ (راهنمایی: فرمول‌های محاسبه و مساحت را به یاد آورید: مساحت مستطیل = طول × عرض و مساحت دایره πr^2)

شما چگونه این شکل را رسم می‌کنید تا اطلاعات موجود را بهتر نمایش دهد؟
به نظر شما آیا رنگ در انتقال نظر تان به بیننده تأثیر دارد؟

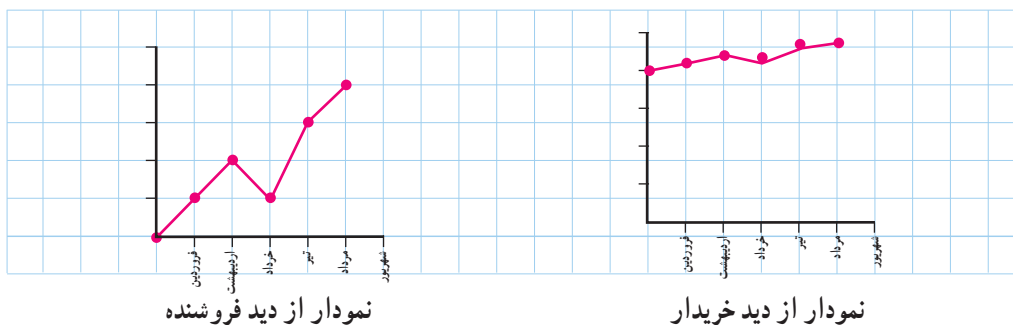


تمرین

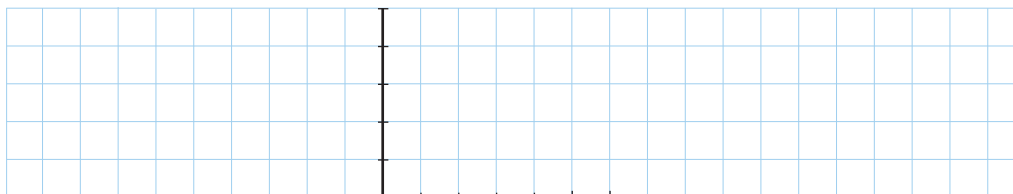
۱. سود خالص یک شرکت خدماتی در شش ماه نخست سال بر حسب میلیارد ریال به صورت زیر است:

فروردین	اردیبهشت	خرداد	تیر	مرداد	شهریور
۲/۰	۲/۱	۲/۲	۲/۱	۲/۳	۲/۴

خریدار و فروشنده سهام این شرکت، نمودارهای زیر را رسم کرده‌اند. اعداد روی محورها را مشخص کنید:



به نظر شما کدام یک منطقی‌تر است؟ آیا می‌توانید نموداری بهتر از این دو نمودار رسم کنید؟



۲. اگر درصد یا فراوانی متغیرهایی که نمودارهای آنها را رسم می‌کنیم، نزدیک به هم باشند، آیا نمودار میله‌ای یا دایره‌ای برای مقایسه مناسب‌تر است؟

۳. رسم نمودارهای میله‌ای و دایره‌ای برای داده‌های کمی مناسب‌تر است یا برای داده‌های کیفی؟



فعالیت

مربی گروه (تیم) بسکتبال شهر می‌خواهد بر اساس نتایج بازی‌های قبلی دو بازیکن، نسبت به حضور یکی از آنها در بازی بعدی تصمیم بگیرد. امتیازهای کسب‌شده توسط این دو بازیکن به صورت زیر است.

۱۱ بازی بازیکن الف:

۱	۱	۱	۳	۲	۱	۲	تعداد بازی‌ها
۳۰	۱۳	۱۱	۱۰	۷	۶	۳	امتیاز کسب شده

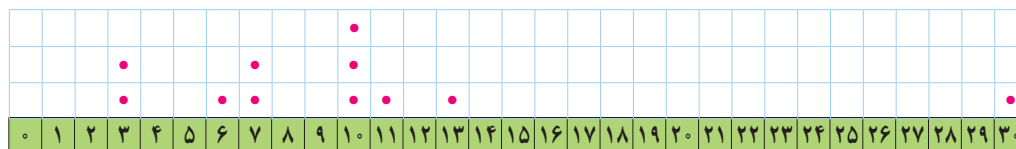
۱۰ بازی بازیکن ب:

۱	۱	۲	۲	۲	۱	۱	تعداد بازی‌ها
۱۳	۱۲	۱۱	۱۰	۹	۸	۷	امتیاز کسب شده

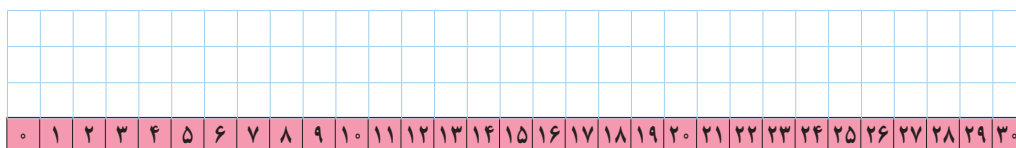
می‌توان داده‌های مرتب شده را به صورت زیر نشان داد:

۳	۳	۶	۷	۷	۱۰	۱۰	۱۰	۱۱	۱۳	۳۰	بازیکن الف
۷	۸	۹	۹	۱۰	۱۰	۱۱	۱۱	۱۲	۱۳		بازیکن ب

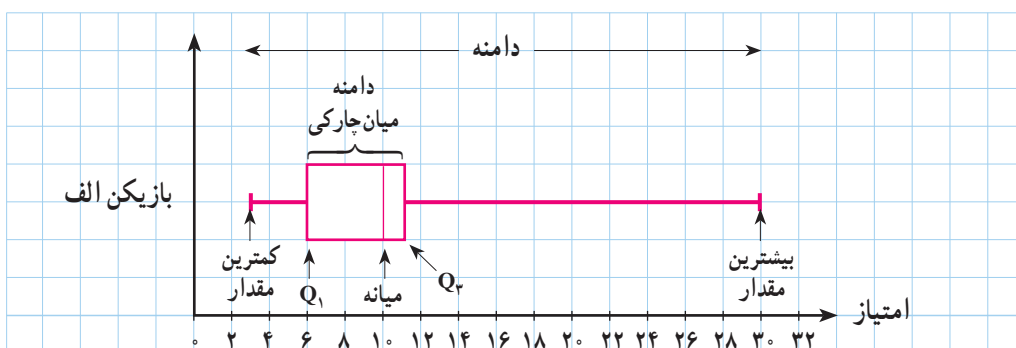
اگر هر یک از اعداد را به صورت نقطه بالای یک محور علامت بزنیم، به نمودار حاصل نمودار نقطه‌ای داده‌ها می‌گویند. نمودار نقطه‌ای امتیازهای بازیکن الف را می‌توان به صورت زیر نمایش داد.



نمودار نقطه‌ای داده‌ها را برای بازیکن ب بر روی محور نمایش دهید.



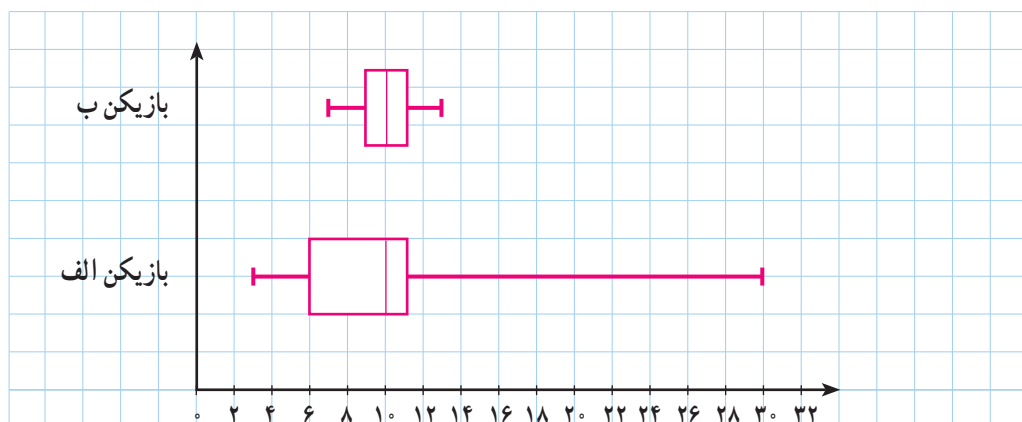
به نظر شما کدام بازیکن بهتر است؟ البته سؤال کلی است و اگر به شما بگویم این بازی آخر است یا گروه مقابل تیمی است که تاکنون امتیاز بیشتری نسبت به گروه شهر کسب کرده، ممکن است پاسخ سؤال عوض شود. در فصل گذشته برای تصمیم درست‌تر، از معیارهای گرایش به مرکز و پراکندگی به طور هم‌زمان استفاده می‌شد. پس خالی از لطف نیست که معیارهای گرایش به مرکز و پراکندگی مجموعه‌های گوناگون از داده‌ها را به شکل تصویری مورد مقایسه قرار دهیم. امتیازهای بازیکن الف را در نظر بگیرید. برای این منظور کمترین مقدار، چارک اول، میانه، چارک سوم و بیشترین مقدار را محاسبه کنید و روی یک محور نمایش دهید. برای مشخص کردن حدود دامنه میان‌چارکی، یک جعبه به عرض دلخواه رسم می‌کنیم. برای مشخص کردن دامنه دو خط، از دو طرف جعبه به کمترین مقدار و بیشترین مقدار داده‌ها وصل می‌کنیم. با مشخص کردن میانه‌روی جعبه، نمودار جدیدی ارائه کرده‌ایم به نام نمودار جعبه‌ای. برای این منظور خطی عمودی بکشید تا میانه مشخص شود.



این نمودار دامنه، دامنه میان‌چارکی و میانه مجموعه داده‌ها را به طور هم‌زمان نشان می‌دهد. بیش از یک مجموعه داده را می‌توان در یک نمودار نشان داد. این بدان معناست که این روش برای مقایسه داده‌ها بسیار عالی است.

اگر داده‌ها در برگرفته داده دور افتاده بودند، دامنه وسیع‌تر می‌بود. در نمودار جعبه‌ای، طول سبیل‌ها با توجه به مرزهای بالا و پایین افزایش پیدا می‌یابد. با نگاه کردن به سبیل‌های نمودار جعبه‌ای، می‌توانید به نامتقارن بودن داده‌ها پی ببرید.

نمودار جعبه‌ای دو بسکتبالیست در اینجا نشان داده شده است. دامنه این امتیازات را مقایسه کنید. اگر مجبور بودید از بین این دو بازیکن، یک نفر را انتخاب کنید، کدام را انتخاب می‌کردید؟ اکنون بهتر می‌توانید به سؤال قبل در وضعیت‌های مختلف گروه مقابل یا حساسیت بازی تصمیم‌گیری کنید.



بازیکن ب دامنه نسبتاً کوچکی دارد، ولی میانه هر دو برابر با ۱۰ است. بازیکن الف دامنه امتیازات بزرگی دارد. گاهی این بازیکن، امتیازاتی بسیار بیشتر از بازیکن ب می‌آورد، و گاهی هم بسیار کمتر. بازیکن ب ثبات بیشتری دارد و معمولاً امتیازاتش از بازیکن الف بیشتر است (میانه‌ها و دامنه میان‌چارکی را با هم مقایسه کنید)، پس بهتر است بازیکن ب را انتخاب کنیم.

نمودار جعبه‌ای، روشی سودمند برای نمایش دامنه‌ها و چارک‌های داده‌هاست. یک جعبه، مکان قرارگیری چارک‌ها و دامنه میان‌چارکی را نشان می‌دهد و سبیل‌ها کمترین و بیشترین داده را نشان می‌دهند. در یک نمودار جعبه‌ای بیش از یک مجموعه داده را می‌توان نشان داد؛ پس این نمودار برای مقایسه مناسب است.

دامنه میان‌چارکی به نظر مفید می‌رسد. این دامنه در مورد بازیکنانی که گاه‌گاهی امتیازات بسیار پایینی کسب می‌کنند چگونه؟ اگر بازیکنی در روز مسابقه، بد بازی کند، برای ما به قیمت از دست دادن قهرمانی در تیمگان (لیگ) تمام می‌شود. مطمئن نیستیم که دامنه یا دامنه میان‌چارکی برای ما مشخص می‌کند که کدام بازیکن واقعاً ثبات بیشتری دارد.



از آنجا که دو خط دوطرف جعبه شبیه سبیل گریه است. بعضی مواقع به آن نمودار جعبه و سبیل گفته می‌شود.

مربی نباید فقط دامنه امتیازات بازیکنان را مقایسه کند. او به راهی نیاز دارد تا به طور دقیق از روی مقادیر موجود محاسبه کند که کدام بازیکن در روز مسابقه ثبات بیشتری دارد. به عبارت دیگر، او باید بازیکنی را پیدا کند، امتیازاتش کمترین تغییرات را داشته باشد.

مشکل دامنه و دامنه میان چارکی این است که فقط تفاوت بین مقادیر کم و زیاد را به شما می گویند؛ اما نمی گویند که چند وقت به چند وقت بازیکنان این امتیازات کم و زیاد را در مقابل امتیازات نزدیک به میانه کسب می کنند. این برای مربی مهم است.



کار در کلاس



برای مجموعه داده های زیر نمودار جعبه ای بکشید.

$\frac{4}{8}, \frac{2}{5}, \frac{4}{1}, \frac{1}{25}, \frac{1}{5}, \frac{2}{5}, \frac{3}{1}, \frac{3}{2}, \frac{4}{25}, \frac{4}{75}, \frac{4}{95}, \frac{5}{1}$

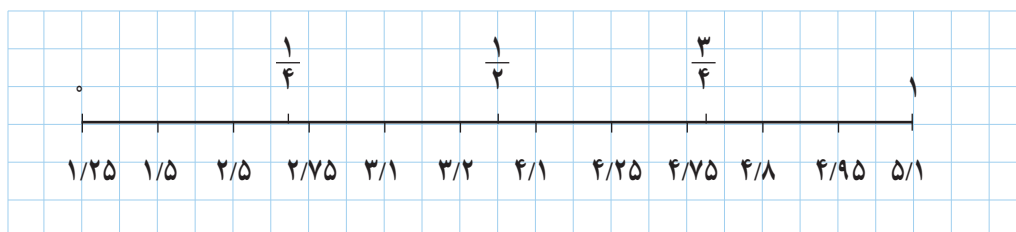
راهنمایی:

گام اول: کمترین مقدار و بیشترین مقدار را مشخص کنید.

از آنجا که اعداد به ترتیب چیده شده اند، اولین عدد مقدار است و آخرین عدد مقدار.

گام دوم: چارک ها را مشخص کنید.

در این مجموعه ۱۲ عدد وجود دارد. می توانیم از شکل زیر یا از فرمول برای تعیین چارک ها استفاده کنیم.

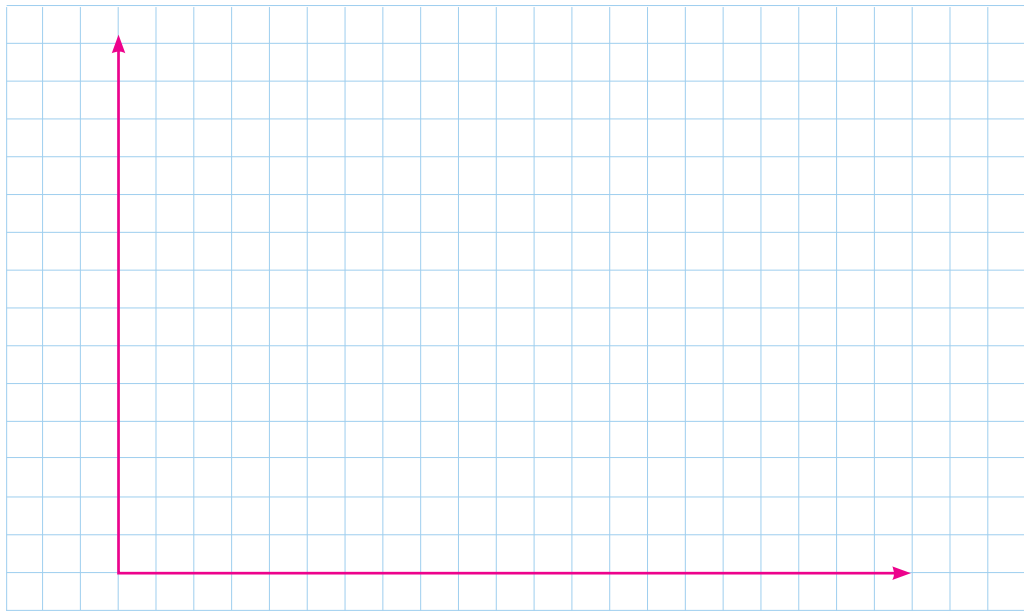


با نگاه به شکل بالا در می یابیم که میانه بین مقادیر و است، بنابراین، مقدار میانه می شود:

چارک اول، بین مقادیر و قرار می گیرد. بنابراین، مقدار اولین چارک می شود:

چارک سوم بین مقادیر و قرار می گیرد. بنابراین، مقدار سومین چارک می شود:

گام سوم: نمودار جعبه‌ای رسم کنید.



نمودارهای میله‌ای افقی درست شبیه نمودارهای میله‌ای عمودی‌اند، با این فرق که محورها چرخیده‌اند. در نمودارهای میله‌ای افقی، رده‌ها روی محور عمودی و درصد یا فراوانی روی محور افقی نشان داده می‌شود.

نمودارهای میله‌ای عمودی مرسوم‌ترند، اما نمودارهای میله‌ای افقی در صورتی که اسم رده‌ها طولانی باشند، مفید واقع می‌شوند. آنها برای نمایش اسامی رده‌ها، بدون نیاز به عمودی نوشتن، فضای زیادی را در اختیارمان می‌گذارند.

تمرین



۱. نمودارهای میله‌ای فراوانی یا درصدها را نشان می‌دهند. چه زمانی باید از فراوانی‌ها و چه زمانی از درصدها استفاده کرد؟
۲. نمودارهای میله‌ای افقی درست شبیه نمودارهای میله‌ای عمودی هستند. با این تفاوت که محورها چرخیده‌اند. نمودارهای میله‌ای عمودی مرسوم‌تر هستند. به نظر شما رسم نمودارهای میله‌ای افقی چه زمانی مفید است؟
۳. سن بازیکنان تیم ملی فوتبال یک کشور به شرح زیر است :

۲۷	۲۴	۲۶	۲۶	۲۹	۱۹	۳۱	۱۸	۲۳	۲۲	۲۵	۲۶	۲۷	۲۳	۲۹	۲۵	۲۵	۳۳	۳۱	۲۱	۲۶	۲۵
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

الف) نمودار نقطه‌ای رسم کنید و مقادیر میانگین، مد و میانه سن بازیکنان این تیم روی محور افقی نشان دهید.

ب) نمودار جعبه‌ای داده‌ها را رسم کنید.

پ) تعداد بازیکنانی که سن آنها بیشتر از میانگین است، بیشتر است یا تعداد بازیکنانی که سن آنها از میانگین کمتر است؟

ت) تعداد بازیکنانی که سن آنها بالاتر از میانه است بیشتر است یا تعداد بازیکنانی که سن آنها از میانه کمتر است؟ میانه و میانگین را در این بررسی مقایسه کنید. چه نتیجه‌ای می‌گیرید؟

ث) چه تعداد از بازیکنان سن آنها بین چارک اول و سوم قرار دارد؟ آیا بدون محاسبه چارک‌ها می‌توانستید به این سؤال پاسخ دهید؟

درس ۲

نمودارهای چندمتغیره



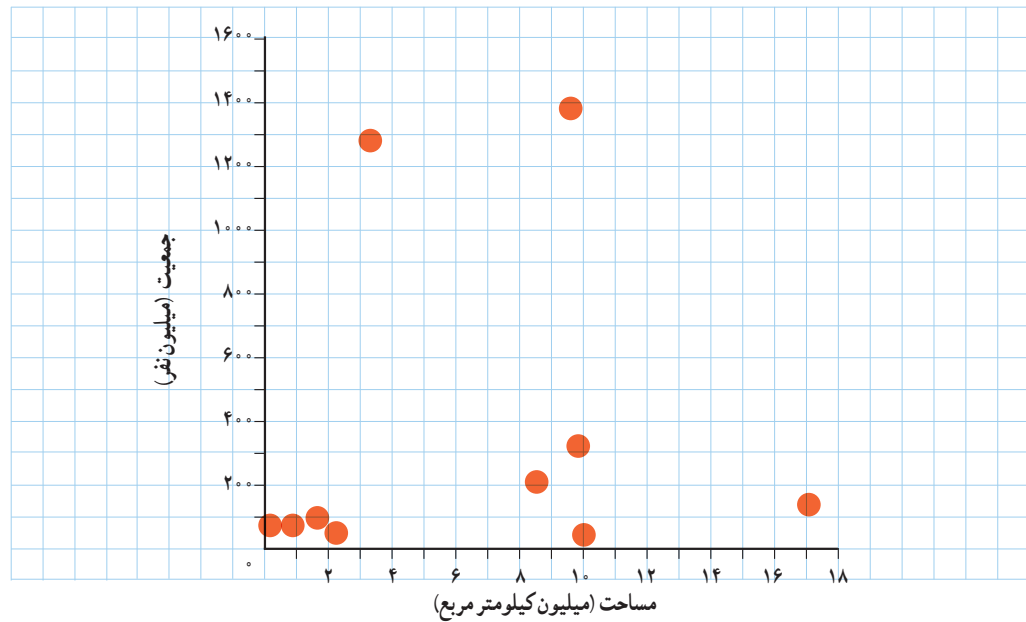
فعالیت

پوریا علاقه زیادی به جغرافیا دارد. او فهرستی از مقادیر سه متغیر مربوط به ۱۰ کشور را تهیه کرده است:

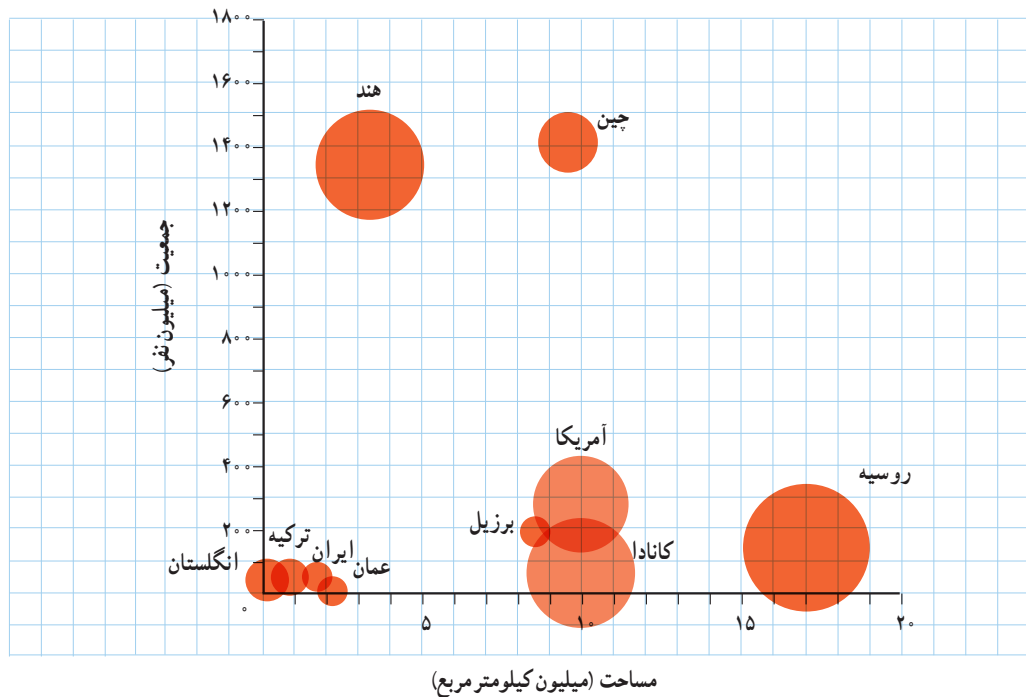
نام کشور	جمعیت (نفر)	مساحت (کیلومتر مربع)	آبها (درصد)
۱ ایران	۷۹,۲۰۰,۰۰۰	۱,۶۴۸,۱۹۵	۰/۷۰
۲ آمریکا	۳۲۲,۳۶۹,۳۱۹	۹,۸۵۷,۳۰۶	۷/۱۰
۳ انگلستان	۶۴,۷۱۶,۰۰۰	۲۴۲,۴۹۵	۱/۳۴
۴ برزیل	۲۰۵,۳۳۸,۰۰۰	۸,۵۱۵,۷۶۷	۰/۶۵
۵ ترکیه	۷۹,۴۶۳,۶۶۳	۸۱۴,۵۷۸	۱/۳۰
۶ چین	۱,۳۷۶,۰۴۹,۰۰۰	۹,۵۹۶,۹۶۱	۲/۸۰
۷ روسیه	۱۴۴,۱۹۲,۴۵۰	۱۷,۰۹۸,۲۴۲	۱۳/۰۰
۸ عمان	۳۰,۷۷۰,۳۷۵	۲,۱۴۹,۶۹۰	۰/۷۰
۹ کانادا	۳۶,۰۴۸,۵۲۱	۹,۹۸۴,۶۷۰	۸/۹۲
۱۰ هند	۱,۲۷۶,۲۶۷,۰۰۰	۳,۲۸۷,۲۶۳	۹/۶۰

پوریا می‌داند که برای هر کدام از متغیرهای مربوط به کشورها؛ یعنی مساحت و جمعیت و درصد آبها، می‌تواند نموداری میله‌ای رسم کند. برای هر جفت از این متغیرها هم می‌تواند پراکنش‌نگاشت رسم کند (مثلاً نمودار جمعیت در مقابل مساحت)؛ اما این سؤال برایش مطرح است که آیا راهی وجود دارد تا هر سه متغیر مربوط به این کشورها را به‌طور هم‌زمان فقط در یک نمودار نشان دهد؟

پراکنش‌نگاشت جمعیت در مقابل مساحت کشورها



پوریا همین‌طور که به پراکنش‌نگاشت جمعیت در مقابل مساحت کشورها نگاه می‌کرد، ناگهان به این فکر افتاد که اندازه نقطه‌های نمودار را متناسب با درصد آب‌های کشورها رسم کند. پس نموداری به شکل زیر کشید:



به این ترتیب، او موفق شد سه متغیر مربوط به کشورها را در یک نمودار، نمایش دهد. نموداری که پوریا رسم کرد، به نام نمودار حبابی معروف است.

هر نقطه یا دایره در نمودار حبابی، در واقع یک سه‌تایی مرتب به صورت (V_1, V_2, V_3) را مشخص می‌کند، که در آن داریم:

- V_1 : موقعیت نقطه روی محور افقی،
- V_2 : موقعیت نقطه روی محور عمودی،
- V_3 : اندازه نقطه (مساحت دایره).

نمودارهای حبابی را می‌توان گونه خاصی از پراکنش نگاشت دانست که برای نمایش هم‌زمان سه متغیر عددی به کار می‌روند و در آنها به جای نقطه از دایره‌های توپُر استفاده می‌شود.

سیستم بینایی انسان به‌طور طبیعی اندازه یک دایره را بر اساس مساحت آن در نظر می‌گیرد. مساحت دایره بر خلاف قطر یا محیط آن، متناسب با شعاع دایره نیست؛ بلکه متناسب با توان دوم شعاع است. پس اگر شعاع دایره‌ها را دقیقاً متناسب با مقادیر متغیر سوم در نظر بگیریم، اختلاف ظاهری اندازه دایره‌ها غیر واقعی و گمراه‌کننده خواهد بود. برای برطرف کردن این مشکل باید شعاع دایره‌ها را متناسب با جذر مقادیر متغیر سوم در نظر گرفت. توجه: متغیر سوم در نمودارهای حبابی نباید دارای مقادیر منفی یا صفر باشد.



کار در کلاس

قد و وزن و نمره درس‌های ریاضی و ادبیات فارسی دانش‌آموزان کلاس را مشابه آنچه در فعالیت دیدید، در یک جدول یادداشت کنید.

برای هر کدام از دسته‌های سه‌تایی متغیرهای زیر، نمودار حبابی رسم کنید:

(الف) قد، وزن، نمره ریاضی
(ب) قد، وزن، نمره ادبیات فارسی
(پ) قد، نمره ریاضی، نمره ادبیات فارسی

سپس در هر نمودار، جای متغیرها را چنان عوض کنید؛ به طوری که دایره‌ها هربار، نشان‌دهنده یکی از متغیرها باشند.



تمرین

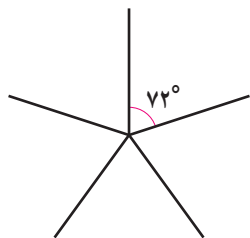
۱. مقدار متغیر سوم در نمودار حبابی، متناسب با کدام یک از موارد زیر است؟
 - (الف) شعاع دایره‌ها
 - (ب) قطر دایره‌ها
 - (پ) محیط دایره‌ها
 - (ت) مساحت دایره‌ها
۲. نمودارهای حبابی برای نمایش چند متغیر عددی در یک نمودار به کار می‌روند؟
 - (الف) یک متغیر
 - (ب) دو متغیر
 - (پ) سه متغیر
 - (ت) محدودیتی ندارد



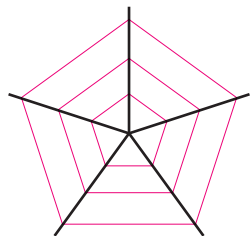
اطلاعات اضافی درباره واحدهای آماری (علاوه بر سه متغیر اصلی) را می‌توان با استفاده از رنگ‌ها و طرح‌ها وارد نمودار حبابی کرد. مثلاً در نمودار مربوط به درصد آب‌های کشورها می‌توان قاره‌های مختلف را با رنگ‌های متفاوت دایره‌ها مشخص کرد. اگر علاوه بر آن بخواهیم زبان رسمی کشورها را هم در نمودار مشخص کنیم، چه راهی پیشنهاد می‌کنید؟



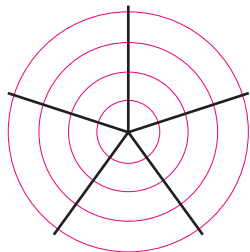
نمودار راداری، روشی برای نمایش داده‌های چندمتغیره در قالب نموداری دویعدی است، که در آن سه متغیر کمی یا بیشتر بر روی محورهای نشان داده می‌شوند که نقطه شروع همه آنها یکی است. محل قرارگیری و زاویه محورها نسبت به یکدیگر، هیچ اطلاعات خاصی را بیان نمی‌کند.



نمودار راداری از چند خط (به‌طور دقیق‌تر: نیم‌خط) به‌شکل پره‌های چرخ دوچرخه تشکیل می‌شود که در نقطه‌ای مرکزی به یکدیگر اتصال دارند و با زاویه‌های یکسان بین هم، گرداگرد آن نقطه قرار گرفته‌اند. هر کدام از این خطوط، در واقع یک محور و نشان‌دهنده یک متغیر است؛ پس می‌توانند درجه‌بندی شوند. به این خطوط، اصطلاحاً شعاع‌های نمودار راداری می‌گویند.



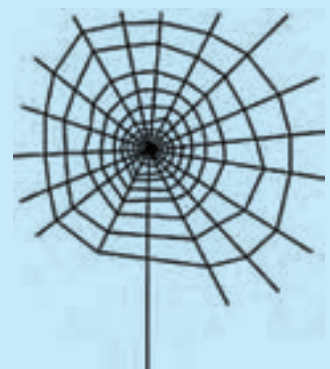
اگر درجه‌بندی شعاع‌ها به‌گونه‌ای انجام گیرد که فاصله بین نشانک‌های متوالی بر روی همه شعاع‌ها یکسان باشد، می‌توان خطوط راهنمای مقادیر را هم رسم کرد؛ به‌شرطی که باعث شلوغی نمودار نشود.

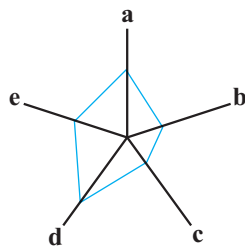


گاهی این خطوط راهنما را به‌جای راست‌خط به‌صورت خط خم و دایره رسم می‌کنند. در این حالت، نمودار واقعاً شبیه صفحه نمایش رادار می‌شود.

خواندنی

از آنجا که نمودار راداری گاهی شبیه تار عنکبوت است، عده‌ای به آن «نمودار تار عنکبوتی» هم می‌گویند. اگر خطوط راهنما رسم نشوند، این نمودار ممکن است شبیه یک ستاره شود. به همین دلیل، آن را «نمودار ستاره‌ای» هم می‌نامند. البته نام‌های دیگری هم برای این نمودار وجود دارد.





نسبت مقدار یک متغیر، برای یک مشاهده به پیشینه آن متغیر به ازای همه مشاهده‌ها، اندازه آن مقدار روی شعاع مربوط را مشخص می‌کند. این کار را برای همه متغیرها و بر روی همه شعاع‌ها انجام می‌دهیم. سپس نقاط مشخص شده روی شعاع‌های مجاور به ازای هر مشاهده را به هم وصل می‌کنیم. به این ترتیب، نمودار راداری برای یک مشاهده به دست خواهد آمد. اگر نمودار مربوط به دو یا چند مشاهده را روی هم بیندازیم، می‌توانیم مشاهده‌ها را با هم مقایسه کنیم.

نمودار راداری به ما می‌گوید:

- کدام مشاهده‌ها شبیه به یکدیگرند؟
 - آیا داده دورافتاده‌ای وجود دارد؟
 - مقدار کدام متغیر برای یک مشاهده نسبت به متغیرهای دیگر بیشتر یا کمتر است؟
 - مقدار کدام متغیر برای یک مشاهده نسبت به مشاهده‌های دیگر بیشتر یا کمتر است؟
- از این نمودار در ورزش برای نشان دادن میزان قدرت و ضعف بازیکنان نسبت به یکدیگر استفاده می‌شود.



یادآوری: منظور از مشاهده یا نقطه داده‌ای در آمار، مجموعه‌ای است شامل یک یا چند مقدار اندازه‌گیری شده برای یک عضو جامعه‌ی آماری (= واحد آماری).

مثلاً اگر واحدهای آماری، دانش‌آموزان کلاس باشند و سن و قد و وزن و گروه خونی آنان را ثبت کنیم، مجموعه مقادیر این متغیرها برای هر کدام از دانش‌آموزان، یک «مشاهده» است.

◆ **مثال:** نامزدهای نهایی دریافت جایزه «توپ طلا» در سال ۲۰۱۳ میلادی سه بازیکن A و B و C بودند که در نهایت، جایزه به بازیکن A تعلق گرفت.

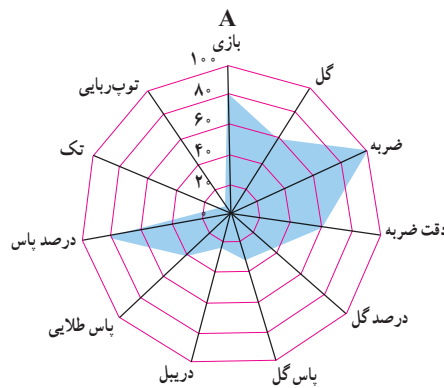
برای اینکه ببینیم آیا این انتخاب عادلانه بوده است یا نه، داده‌های گردآوری شده از این سه بازیکن در طول فصل ۲۰۱۳-۲۰۱۲ را بررسی می‌کنیم. (همه داده‌ها، به جز تعداد بازی، به صورت میانگین مقادیر اندازه‌گیری شده در طول هر ۹۰ دقیقه محاسبه شده‌اند.)

متغیر	A	B	C	پیشینه
تعداد بازی	۳۰/۱	۲۹/۱	۲۳/۵	۳۸
گل زده (غیر از پنالتی)	۰/۹۳	۱/۴۴	۰/۴۳	۱/۶
ضربه به سمت دروازه	۷/۷۸	۵/۶	۲/۷۷	۸
دقت ضربه (درصد)	۴۴/۶	۴۹/۱	۴۷/۷	۷۵
ضربه‌های گل شده (درصد)	۱۴/۵	۲۸	۱۵/۳	۴۰
پاس گل	۰/۳۳	۰/۴۱	۰/۶	۱
دریبل موفق	۱/۸۹	۴/۱۹	۵/۲	۷
پاس طلایی (مهم)	۲/۰۵	۱/۵۸	۳/۶۶	۵
پاس‌های موفق (درصد)	۷۶/۸	۸۵	۸۷/۸	۹۵
تک (تکل)	۰/۴۶	۰/۶۵	۱/۱۵	۶
توپ‌ربایی	۰/۳	۰/۳۴	۰/۸۹	۶

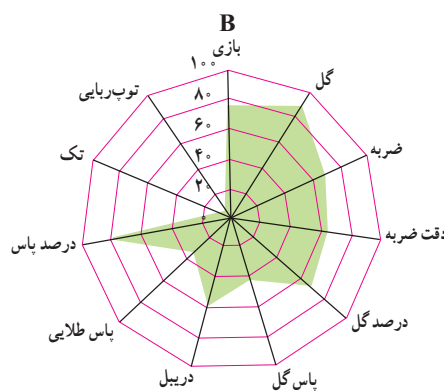
توپ طلای فدراسیون بین‌المللی فوتبال (فیفا)، جایزه‌ای است که هر سال توسط فیفا و مجله فرانس فوتبال به بهترین بازیکن فوتبال مرد داده می‌شود. انتخاب بهترین بازیکن بر اساس رأی‌گیری از نمایندگان رسانه‌های بین‌المللی و سرمربیان و کاپیتان‌های تیم‌های ملی فوتبال صورت می‌گیرد.

اکنون نمودار را دارای مربوط به هر سه بازیکن را رسم می‌کنیم.

ملاحظه می‌شود که تعداد ضربه‌های بازیکن A به سمت دروازه، تقریباً به حد نهایی نمودار رسیده است. یعنی تعداد ضربه‌های او به سمت دروازه، بیشتر از هر بازیکن دیگری در جهان بوده است؛ اما از سوی دیگر، تعداد دریبل‌های موفق او به‌ویژه نسبت به دو نامزد دیگر، خیلی کم است.

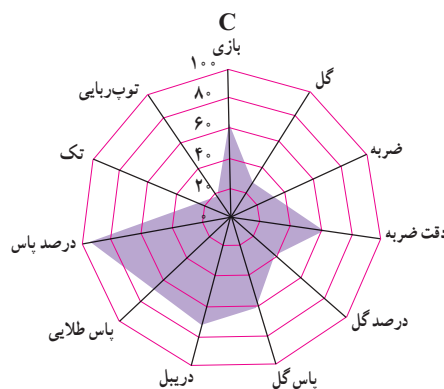


تعداد گل‌های زده بازیکن B در جریان بازی، نزدیک به حد نهایی نمودار شده و از این لحاظ، او جزء برترین‌های جهان بوده است. درصد ضربه‌هایی هم که او به سمت دروازه زده و گل شده (درصد گل)، بسیار زیاد است. به‌طور کلی، ناحیه مربوط به گل زدن در نمودار B، وسعت زیادی دارد.

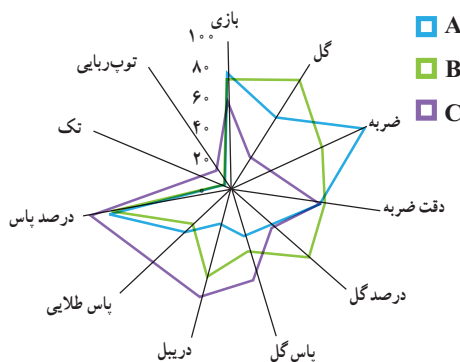


نمودار C به‌شکل جذابی کاملاً متفاوت با نمودار دو نامزد دیگر است. علت این موضوع، نقش متفاوت او در زمین است.

نمودارهای A و B شباهت زیادی به هم داشت. آنان از بزرگ‌ترین گلزنان اند؛ اما بازیکن C یک «گل‌ساز» بزرگ است. با اینکه آمار تعداد گل‌های او برای یک هافبک کناری، بسیار خوب است و درصد ضربه‌های گل شده‌اش نیز عالی است، درخشش اصلی او در آمار پاس و دریبل است. ببینید که ناحیه مربوط به پاس و دریبل در نمودارش چه وسعت زیادی دارد.



اگر داده‌های مربوط به هر سه نفر را در قالب یک نمودار رسم کنیم، شاید مقایسه بین آنان آسان‌تر شود. اکنون آیا می‌توانید بگویید که کدام بازیکن، شایستگی بیشتری برای دریافت توپ طلای سال ۲۰۱۳ میلادی داشت؟

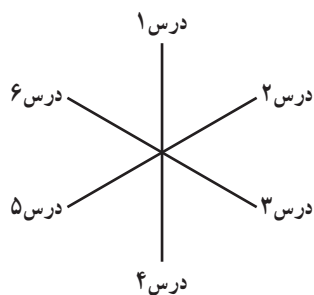




کار در کلاس

آموزگارتان شش درس را انتخاب و برای شما به صورت زیر فهرست می کند :

- ۱- ۲-
۳- ۴-
۵- ۶-



الف) نمودار راداری نمره های خود را به صورت مقابل رسم کنید :

ب) نمودار خود را با یکی دیگر از دانش آموزان ادغام کنید و با کمک یکدیگر، نمودار جدیدی رسم کنید.

پ) به نظر شما نمره های کدامتان بهتر است؟



تمرین

۱. نمودار راداری برای نمایش داده های چند متغیر کمی به طور هم زمان به کار می رود؟
۲. زاویه بین شعاع های مجاور در نمودار راداری، چه چیزی را نشان می دهد؟
۳. نمودار راداری چه چیزی به ما می گوید؟
۴. کاربرد نمودار راداری در ورزش چیست؟
۵. اگر زاویه بین دو شعاع مجاور در نمودار راداری، 40° درجه باشد، چند متغیر در نمودار حضور دارد؟
۶. داده های زیر را که مربوط به شاخص های سلامت است، در قالب یک نمودار راداری نمایش دهید :

متغیر	ایران	پاکستان	ترکیه	بیشینه
X_1 عمر مورد انتظار در بدو تولد (سال)	۷۵/۵	۶۶/۴	۷۵/۸	۸۴
X_2 نسبت متخصصان سلامت (به ازای هر ۱۰ هزار نفر جمعیت)	۲۳/۰	۱۴/۰	۴۱/۱	۲۴۵
X_3 نسبت ولادت ها به وسیله متخصصان سلامت (درصد)	۹۶	۵۲	۹۷	۱۰۰
X_4 نسبت جمعیت دارای دسترسی به شبکه فاضلاب	۹۰	۶۴	۹۵	۱۰۰
X_5 شاخص آمادگی اجرای مقررات بین المللی سلامت	۸۵	۴۳	۷۸	۱۰۰

(داده ها برگرفته از گزارش آماری سازمان جهانی بهداشت، سال ۲۰۱۶ میلادی)



* منابع

۱. آشفته (۱۳۹۵)، ترندها و سواد آماری، خانه آمار اصفهان.
۲. ج. بهبودیان (۱۳۹۳)، آمارو احتمال مقدماتی، انتشارات آستان قدس.
۳. گروه مؤلفان (۱۳۹۵)، آمارو مدل سازی، وزارت آموزش و پرورش.
4. C. H. Brase and C. P. Brase (2012). Under-standable Statistics, Brooks/Cole.
5. D. Griffiths (2009). Head First Statistics, O'Reilly Media.
6. S. McKillup (2004). Statistics Explained: An Introductory Guide for Life Scientists, Cambridge Univ. Press.
7. A. J. Vickers (2009). What is a p-value anyway? 34 Stories to Help You Actually Understand Statistics, Pearson.
8. Written by Volunteers (2014). Everything Maths: Grade 10 and 11 Mathematical Literacy, Siyavula.