

نمودارها و تحلیل داده‌ها

فصل ۵- جلسه اول تا چهارم

مفاهیم: شاخص‌های هندسی، نمودار میله‌ای، نمودار مستطیلی، چندبر فراوانی، نمودار دایره‌ای و نمودار ساقه و برگ

مهارت‌ها: توانایی خلاصه کردن داده‌ها توسط نمودار میله‌ای، نمودار مستطیلی، چندبر فراوانی، نمودار دایره‌ای و نمودار ساقه و برگ.

اهداف:

- ساختن، خواندن و تحلیل جدول‌ها و نمودارها
- باور به قدرت روشهای آماری برای تصمیم‌گیری
- توانایی مقایسه ارجحیت‌های نمودارهای هندسی نسبت به یکدیگر

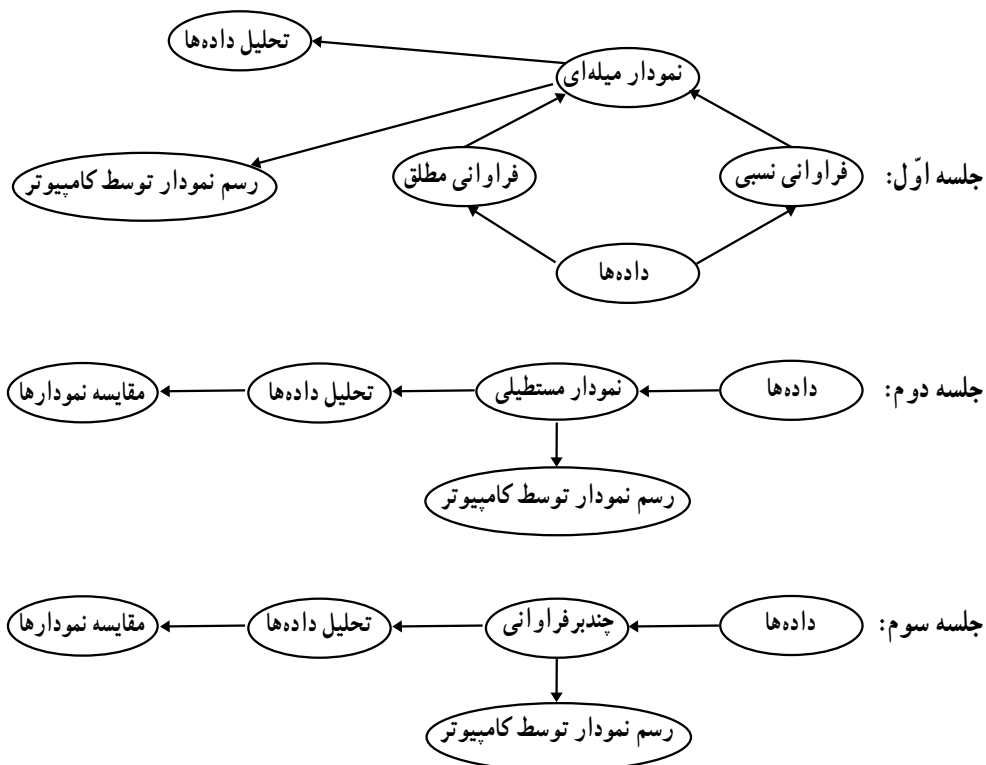
تأکید درس: مشخص کردن دامنه تغییرات داده‌ها، رسم نمودارهای ساده، خواندن داده‌ها و تحلیل آن‌ها به عنوان پاسخی برای سؤالات از پیش مشخص شده فعالیت‌های مهمی هستند اما جنبه‌های بسیار باریکی از آمار را تشکیل می‌دهند. در مقابل تأکید درسی آمار بر شرکت فعالانه دانش‌آموزان در تمام روند حل مسئله است: فرمول‌بندی سؤالات کلیدی، جمع‌آوری و نظم‌بخشی به داده‌ها، نمایش داده‌ها توسط نمودارها، جداول و خلاصه نمودن داده‌ها، تحلیل داده‌ها، پیشگویی، ارتباط برقرار کردن با دیگران و در دسترس قرار دادن اطلاعات به دست آمده به روشی که آنان را متقاعد نماید. این تمرین بخصوص برای تمام دانش‌آموزان لازم است، زیرا تفسیر کردن و توضیح دادن از توانایی‌هایی است که در سیستم آموزشی جدید بر آن تأکید شده است. به طور خلاصه، نباید موضوع درس باعث شود جنبه‌های مختلف کارآماری مستقل و جدا جدا تمرین شوند. سعی کنید دانش‌آموزان را با مسائلی درگیر کنید و در نهایت از نمودارها برای نمایش داده‌ها استفاده کنید و پس از آن به تحلیل داده‌ها بپردازید. توانایی نقد تمامی مراحل حل مسئله نیز مورد تأکید است. تا هنگامی که دانش‌آموزان توانایی قضاوت در مورد حل مسئله توسط دیگران را به دست نیاورده‌اند، قادر نخواهند بود خلاقانه به حل مسائل توسط ایده‌های آماری بپردازند. تا جایی که ممکن است داده‌ها را

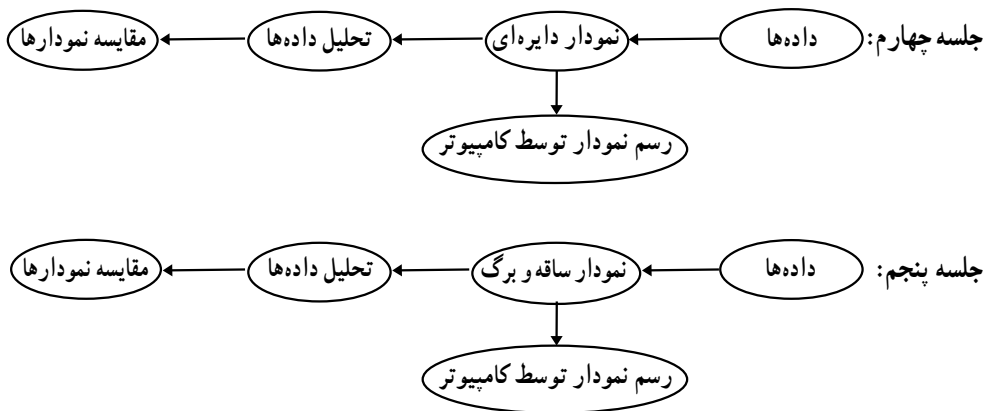
توسط چندین نمودار به نمایش بگذارید تا دانش آموزان بتوانند در مورد توانایی‌های هر روش نمایش با هم به گفتگو بپردازند.

بحث: سؤالاتی نظیر: چه الگوهایی در نمودار مشاهده می‌شوند؟ آیا داده‌ها اطلاعاتی غیرقابل انتظار در دسترس قرار می‌دهند؟ از روی اطلاعات چه پیشگویی‌هایی می‌توان کرد؟ تفاوت الگوهای مشاهده شده در نمودارهای مختلف را چگونه به هم مربوط می‌دانید؟ و ... سؤالاتی هستند که دانش آموزان را به سوی درونی شدن مهارت‌های آماریشان سوق می‌دهند و بر شاخص‌های هندسی تأکید دارند. سؤالاتی نظیر: چه داده‌هایی با تکرار بیش‌تر ظاهر می‌شوند؟ آیا روند رشدی در داده‌ها به چشم می‌خورد؟ تغییرات مهم در کجا رخ داده است؟ از روی داده‌ها چه پیشگویی‌هایی می‌توان کرد؟ چه مشکلاتی ممکن است بر سر راه تعمیم نتایج تحلیل‌ها سد به‌وجود بیاورد؟ چه داده‌های جدیدی ممکن است این تعمیم‌ها را صحت بگذارد؟ و ... سؤالاتی هستند که دانش آموزان را به سوی درونی شدن مهارت‌های آماریشان سوق می‌دهند ولی بر شاخص‌های عددی تأکید دارند.

مباحث ریاضی مربوط: جدول فراوانی، تحلیل داده‌ها، جامعه آماری.

نمودار مفاهیم:





نمودارها و تحلیل داده‌ها

درباره نمودارها بنظر می‌رسد ابهامی نباشد. جایی که توضیح می‌دهیم که دو سر نمودار چند بر فراوانی به محور افقی وصل شود تا مساحت زیر چند بر و مساحت نمودار مستطیلی برابر باشند اهمیت دارد. زیرا این مفهوم بعدها تبدیل به احتمال خواهد شد، که در این جا حرفی از آن به میان نیاورده‌ایم. شما هم در کلاس مطرح نکنید. نکته دیگر آنکه آنچه که در نمودارها مهم است ارتفاع نقاط از محور افقی نیست بلکه بلندی و یا کوتاهی آن‌ها نسبت به هم مقایسه می‌شود پس اگر شما نمودارها را با فراوانی مطلق و یا فراوانی نسبی رسم کنید تغییری در نتیجه‌ای که خواهید گرفت حاصل نخواهد شد. این وضع در نمودار دایره‌ای بخوبی مشهود است، زیرا هرگونه تغییر متناسب در داده‌ها شکل نمودار دایره‌ای را تغییر نمی‌دهد. (در سایر نمودارها ممکن است نمودار کمی بالاتر و یا پایین‌تر بیاید.) نمودار ساقه و برگ این حسن را دارد که در عین حال که از تمام داده‌ها استفاده می‌کند و ما می‌توانیم با مراجعه به نمودار داده‌ها را هم ببینیم، یک نمودار ستونی از داده‌ها فراهم می‌کند، علاوه از روی این نمودار به راحتی می‌توان میانه را هم حساب کرد. نکته‌ای که در کتاب درسی چندان واضح به آن پرداخته نشده است و اگر شما مایل باشید می‌توانید راجع به آن صحبت کنید چیزی است که در مثال زیر رخ می‌دهد.

۱۰۰۲۴

۱۶۶۷۸۹

۲۰۰۱۲۳۳۴

۲۵۵۵۶۶۶۶۶۷۷ ۱ ۶=۱۶

۳۲۲۲۳۳۴

۳۶۶۷۷

در این نمودار ۳۶ داده داریم و ملاحظه می‌کنید روی ساقه‌ها هر ساقه دوبار تکرار شده است برگ روی ساقه بالایی ارقام از ۰ تا ۴ را دارند و برگ روی ساقه پایینی ارقام از ۵ تا ۹ را دارند. اگر این کار را نمی‌کردیم نمودار زیر حاصل می‌شد.

۱۰۰ ۲ ۴ ۶۶۷۸۹

۲۰۰ ۱ ۲ ۳۳۴۵۵۵۶۶۶۶۷۷

۳ ۲۲۲۳۳۴۶۶۷۷

که برشهای روی ساقه‌ها بسیار طولانی می‌شوند. هرگاه که تعداد داده‌ها روی ساقه‌ها زیاد می‌شود می‌توان آن‌ها را به روشی که در بالا گفتیم به دو ساقه تقسیم کرد تا شکل بهتری از داده داشته باشیم.

اشاره کردیم که از نمودار، ساقه و برگ می‌توان برای محاسبه میانه استفاده کرد. برای این منظور در مثال بالا چون تعداد داده‌ها زوج است کافی است دو داده وسطی را پیدا کنیم، یعنی داده‌های هیجدهم و نوزدهم برای این منظور از ۰ داده‌های روی ساقه اول شروع به شمردن می‌کنیم تا به داده‌های هیجدهم و نوزدهم برسیم که عبارتند از ۵/۵ (دو پنج آخر قبل از ۶) پس میانه عبارت است از ۵/۵. نمودار ساقه و برگ برای داده‌هایی که از اعداد طبیعی درست شده باشند و در فاصله کمی پراکنده نشده باشند مناسب است.

توضیح فعالیت‌ها

بحث کلاسی ۱ صفحه ۸۰

نمودار بالا چپ به خاطر ضخیم بودن نوار مربوطه به محصول B گمراه‌کننده است زیرا برخی از مشاهده‌گران ممکن است به اشتباه به مقایسه مساحت‌ها بپردازند. نمودار پایین چپ به خاطر محور گرم شکر که عددگذاری نشده است گمراه‌کننده است. نمودار پایین راست به خاطر محور گرم شکر که از صفر عددگذاری نشده است گمراه‌کننده است مثلاً می‌تواند باعث این تصور شود که محصول B چندین برابر محصول A شکر دارد در صورتی که کمتر از دو برابر شکر دارد. برحسب ۱۰۰۰ نفر

بحث کلاسی ۲

برای رسم نمودار اطلاعات جدول بالا یعنی داده‌ها و فراوانی هر داده کافی است.



– در یکی از نمودارها میله‌ها نمایانگر تعداد دانش آموزان و دیگری نمایانگر درصد دانش آموزان است.

– تحلیل نمودار فراوانی نسبی ساده از نمودار فراوانی مطلق است. ولی هیچ تأثیری در تعبیر داده‌ها (مقایسه آن‌ها با هم) ندارد.

بحث کلاسی صفحه ۸۳

– طول قاعده نشانگر طول دسته است و ارتفاع نشانگر تعداد اعضای دسته (فراوانی دسته) تقسیم بر طول قاعده است به طوری که مساحت مستطیل برابر تعداد اعضای دسته بشود.

– خیر زیرا اطلاعات جزئی را در دسته‌بندی از دست داده‌ایم و نمی‌توانیم دقیقاً بگوییم چند نفر

در سن ۱۵ سالگی در کلاس اول دبیرستان مشغول به تحصیل هستند.

میلیون نفر $130 \approx (10 \times 12/3) + (4 \times 1/8)$ (میلیون $1/8 \times 4$)

– سؤال نمونه: به طور میانگین رشد سالانه جمعیت در کشور چقدر است؟

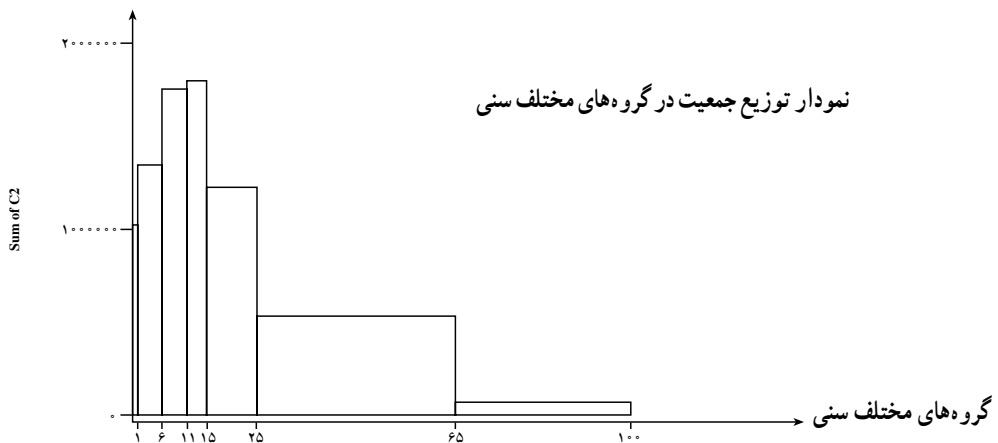
پاسخ تمرینات صفحات ۸۴، ۸۵ و ۸۶

۱- در حالت کلی این نمودار افزایش تولید را نشان نمی‌دهد. دلیل این گمراه‌کنندگی نیز نداشتن عنوان برای تعیین نوع فراوانی استفاده شده در رسم نمودار است. اگر نمودار براساس فراوانی مطلق رسم شده بود، بله ما شاهد افزایش تولید هستیم:

سال ۱۳۶۸، ۱۰۰ تا

سال ۱۳۶۹، ۱۵۰ تا

سال ۱۳۷۵، ۲۰۰ تا



ولی اگر براساس فراوانی تجمعی رسم شده باشد، خیر ما شاهد کاهش تولید در سال‌های ۱۳۶۹

و ۱۳۷۰ نسبت به ۱۳۶۸ هستیم:

سال ۱۳۶۸، ۱۰۰ تا

سال ۱۳۶۹، $150 - 100 = 50$ تا

سال ۱۳۷۰، $200 - 150 = 50$ تا

$$1 + 1 + 5 + 10 + 5 + 6 = 28$$

–۲

نفر واجد شرایط شرکت در دوره می‌باشند که با استفاده از نمودار فراوانی تجمعی داریم:

$$30 - 2 = 28$$

پس نمودار رسم شده براساس فراوانی تجمعی ما را سریع تر به پاسخ هدایت می کند.

۳- این نتیجه گیری صحیح نمی باشد. زیرا نمی دانیم که در هر سرعت، چند اتومبیل با آن سرعت حرکت کرده است. مثلاً ممکن است ۱۰ ماشین با سرعت بین ۱۱۰ تا ۱۲۰ کیلومتر بر ساعت حرکت کرده باشند که وجود ۱۰ تصادف در این حالت نشان دهنده ۱۰۰٪ خطر (۱۰۰٪) می باشد. تصور فوق به این دلیل ایجاد شده است که نمودار بر حسب فراوانی مطلق رسم شده است. برای تحلیل صحیح

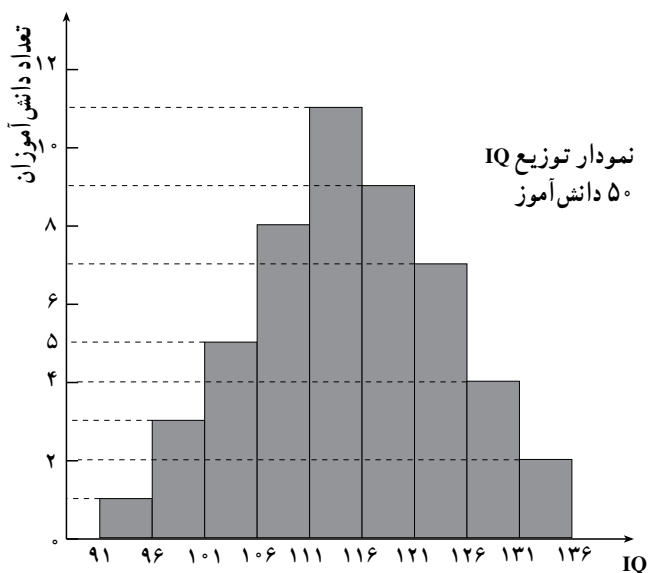
می بایست به جای فراوانی مطلق از تعداد تصادفات در هر دسته استفاده نمود.
تعداد ماشین های حرکت کرده با آن سرعت در هر دسته

با توجه به نداشتن این اطلاعات از دانش آموزان خواسته نشده تا نمودار را رسم کنند.

$$\frac{135-92}{9} \approx 4/8 \approx 5$$

۴-

حدود دسته ها	فراوانی
۹۱-۹۶	۱
۹۶-۱۰۱	۳
۱۰۱-۱۰۶	۵
۱۰۶-۱۱۱	۸
۱۱۱-۱۱۶	۱۱
۱۱۶-۱۲۱	۹
۱۲۱-۱۲۶	۷
۱۲۶-۱۳۱	۴
۱۳۱-۱۳۶	۲

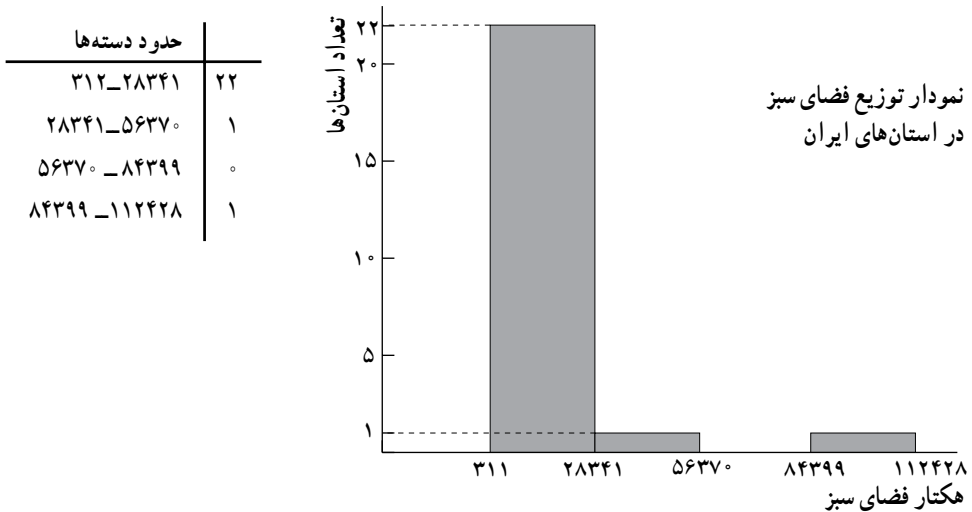


چون نمودار تقریباً متقارن است می توان گفت که جامعه تقریباً طبیعی و متعادل است.

۵- کمترین مقدار ۳۱۳ و بیشترین مقدار ۱۱۲۴۲۷ است. که هر مقیاسی انتخاب شود، نمی توان این اعداد را به خوبی نمایش داد.

— برای رسم نمودار مستطیلی باید داده ها را دسته بندی کنیم پس باید طول دسته ها یا تعداد دسته ها را بدانیم.

$$\text{طول دسته} = \frac{112427 - 313}{4} = 28028 / 5 \approx 28029$$



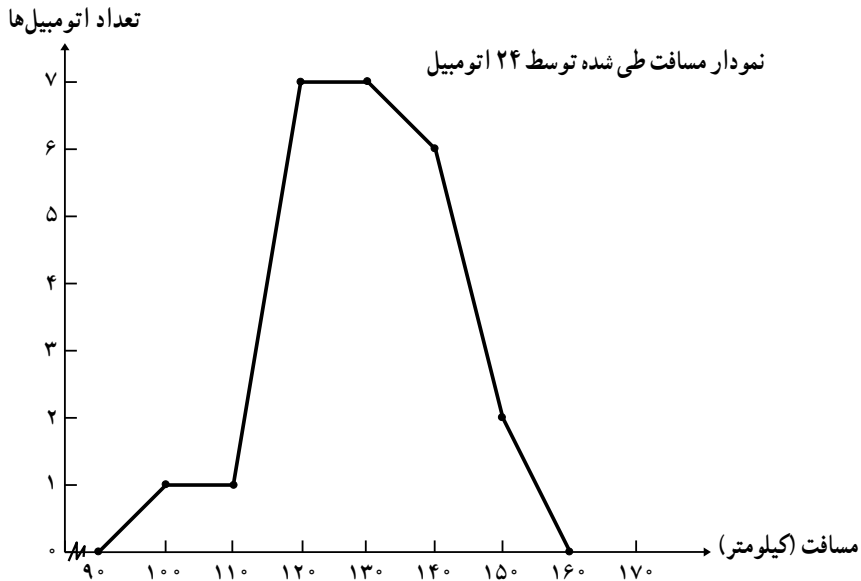
با توجه به این که دسته‌ای با فراوانی صفر به دست آمده است، فکر می‌کنید تغییر تعداد دسته‌ها (به طور مثال از ۴ به ۵)، کمکی به تفسیر اطلاعات می‌کند؟ آیا انتخاب ۴ دسته انتخاب مناسبی است؟ این سؤالات را در کلاس به بحث بگذارید.

پاسخ تمرینات صفحه ۹۱

$$\text{طول دسته} = \frac{153 - 98}{6} \approx 9 / 2 \approx 10 \quad \text{—۱}$$

حدود دسته‌ها	فراوانی	مرکز دسته
۹۵-۱۰۵	۱	۱۰۰
۱۰۵-۱۱۵	۱	۱۱۰
۱۱۵-۱۲۵	۷	۱۲۰
۱۲۵-۱۳۵	۷	۱۳۰
۱۳۵-۱۴۵	۶	۱۴۰
۱۴۵-۱۵۵	۲	۱۵۰

با توجه به آن که متغیر ما، پیوسته است برای نمایش متغیرات نمودار چند برفراوانی مناسب است.



۲- $3+2=5$ روز
چون سرعت باد متغیری پیوسته است و نمودار چندبر فراوانی پیوستگی را کاملاً واضح نشان می‌دهد مناسب‌تر است.

پاسخ تمرینات صفحات ۹۴ و ۹۵

۱- کمی گسسته یا کیفی

۲- خیر، زیرا با دو برابر کردن فراوانی‌ها، مجموع فراوانی‌ها نیز دو برابر می‌شود. لذا نسبت $\frac{\sum f_i}{\sum f_i}$ یعنی زاویه مرکزی ثابت می‌ماند.

۳- فراوانی داده‌ای که زاویه دو برابر دارد، دو برابر فراوانی داده‌های دیگر است.

۴- خیر، زیرا در نمودار دایره‌ای اندازه زاویه‌ها با هم مقایسه می‌شوند و در نتیجه ترتیب قرارگرفتن آن‌ها تأثیری در تفسیر داده‌ها ندارد.

گروه خونی	A	B	AB	O
فراوانی	۵	۶	۴	۳

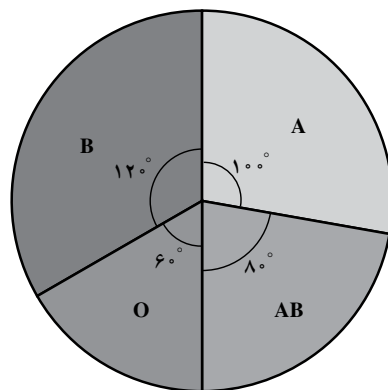
۵- پاسخ‌ها متفاوت است. به طور مثال

$$\alpha_A = \frac{5}{18} \times 360 = 100^\circ$$

$$\alpha_B = \frac{6}{18} \times 360 = 120^\circ$$

$$\alpha_{AB} = \frac{4}{18} \times 360 = 80^\circ$$

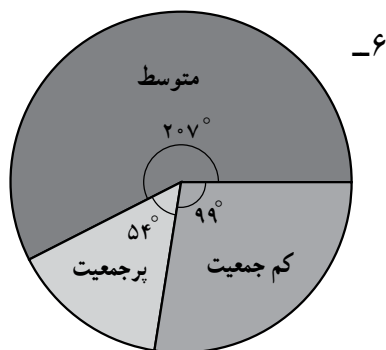
$$\alpha_O = \frac{3}{18} \times 360 = 60^\circ$$



$$\text{کم جمعیت: } \frac{11}{11+23+6} \times 360 = 99^\circ$$

$$\text{متوسط: } \frac{23}{4} \times 360 = 207^\circ$$

$$\text{پر جمعیت: } \frac{6}{4} \times 360 = 54^\circ$$



۷- خیر زیرا در نمودار دایره‌ای زاویه‌ها با هم مقایسه می‌شوند و با تغییر شعاع، زاویه‌ها تغییر نمی‌کنند.

۸- الف :

$$\text{چربی: } \frac{3}{36} = \frac{1}{12}$$

$$\text{کربوهیدرات: } \frac{20}{36} = \frac{5}{9}$$

$$36 \times \frac{6}{36} = 6 \text{ گرم} \quad \text{ب:}$$

۹- بدون محاسبه: سطح مربوط به نبردن بیشتر از بردن است. با محاسبات زیر نیز می‌توان آن

را بررسی نمود.

$$6 \times 20 = 120 \Rightarrow \frac{120}{360} \times 100 = 33\%$$

امکان جایزه بردن :

$$1 - 33\% = 67\%$$

امکان جایزه نبردن

پس امکان نبردن بیشتر است.

$$3 \times \left(\frac{2^\circ}{36^\circ} \times 100\right) \approx 16/6\%$$

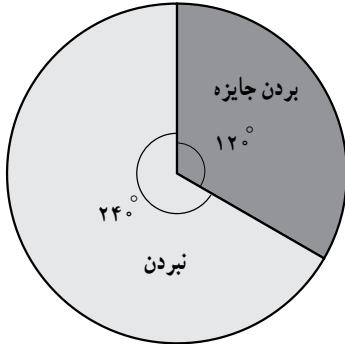
– امکان بردن رادیو

$$1 \times \left(\frac{2^\circ}{36^\circ} \times 100\right) \approx 5/6\%$$

– امکان بردن تلویزیون

$$2 \times \left(\frac{2^\circ}{36^\circ} \times 100\right) \approx 11/11\%$$

– امکان بردن جایزه سفر به مشهد



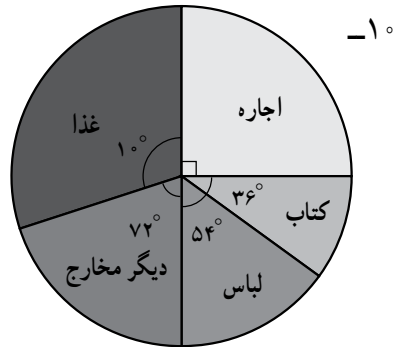
$$\text{غذا} : \frac{3^\circ}{100^\circ} \times 360 = 10.8^\circ$$

$$\text{اجاره} : \frac{25}{100} \times 360 = 90^\circ$$

$$\text{لباس} : \frac{15}{100} \times 360 = 54^\circ$$

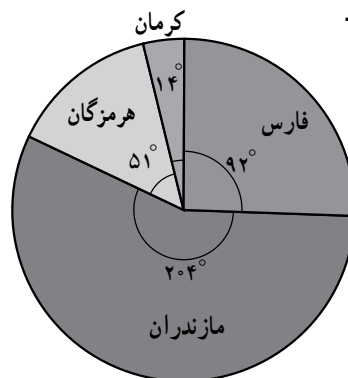
$$\text{کتاب} : \frac{1^\circ}{100} \times 36 = 36^\circ$$

$$\text{دیگر مخارج} : \frac{2^\circ}{100} \times 36 = 72^\circ$$



–۱۰

	میزان تولید	فراوانی نسبی
مازندران	۱۳۷۳۶۸۰	$\approx 0/566 \approx 20.4^\circ$
فارس	۶۱۷۶۳۰	$\approx 0/255 \approx 92^\circ$
کرمان	۹۴۸۷۶	$\approx 0/0.39 \approx 14^\circ$
هرمزگان	۳۴۱۲۳۹	$\approx 0/141 \approx 51^\circ$
	۲۴۲۷۴۲۵	



–۱۱

پاسخ تمرینات صفحات ۱۰۳ - ۱۰۰

۱- ۱۰، ۱۳، ۱۳، ۱۴، ۲۰، ۲۲، ۲۴، ۲۸، ۲۸، ۳۲ و ۱۰

۲- در ساقه شماره ۱: ۲ و ۳ و ۴ و ۵

در ساقه شماره ۳: ۴ و ۵ و ۶ و ۷ و ۸ و ۹

۳- ۱۹۹۰۰ فوت یا ۵۹۷۰۰۰ سانتی متر

۲۲- (بدون در نظر گرفتن ۱۰۰۰۰ فوت)

۵- آتششان به ارتفاع ۴۳۰۰ فوت

۴-

۵۷

۱۰۱۳۷۹

۲ ۱۲۳۶۶۹

۳ ۲۳۷

۴ ۳۸

۵ ۰

۶ ۵

کلید نمودار

۳ ۲ = ۳۲

۵- الف:

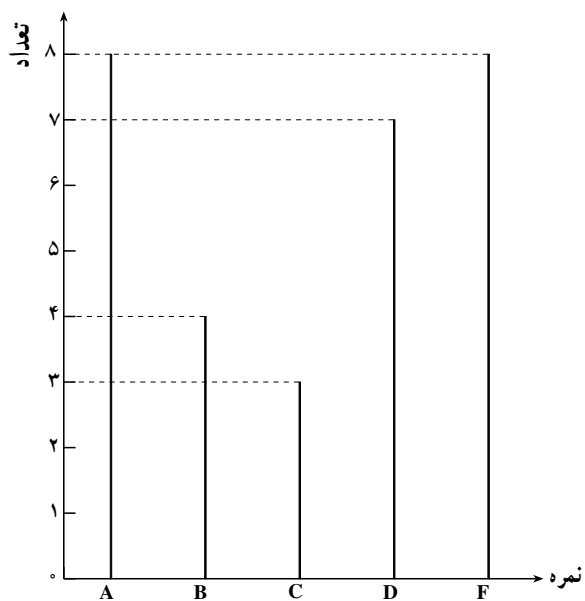
کمترین نمره در کلاس دختران: ۲۹

کمترین نمره در کلاس پسران: ۳۰

ب:

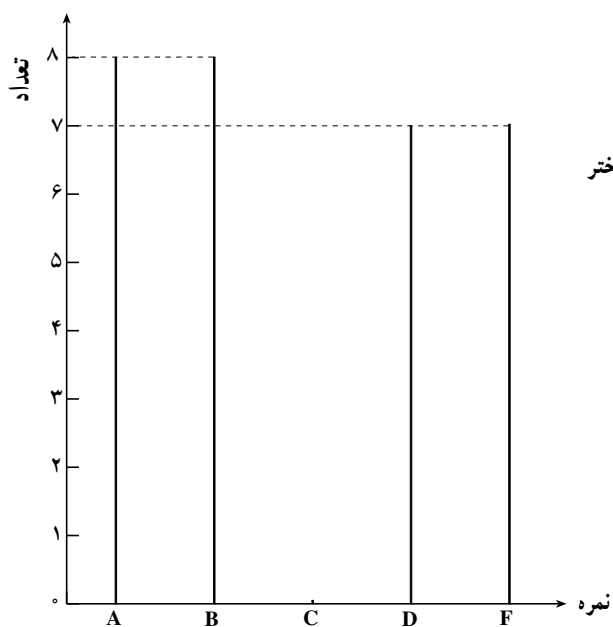
۲۳ نفر در کلاس دختران نمره بیشتر یا مساوی ۶۰ کسب کرده‌اند.

۲۲ نفر در کلاس پسران نمره بیشتر یا مساوی ۶۰ کسب کرده‌اند.



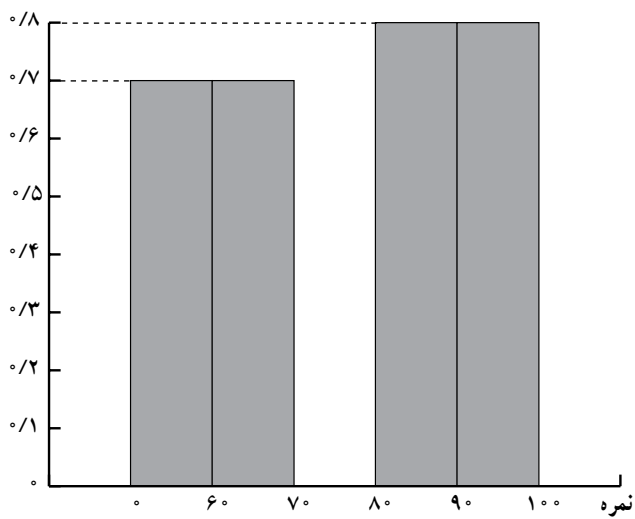
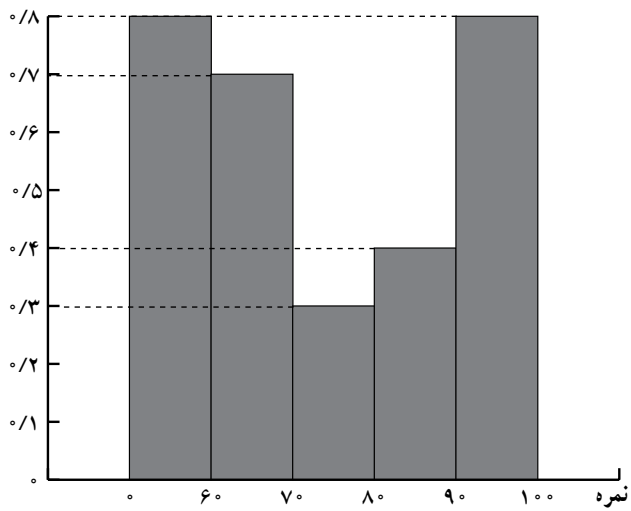
نمودار نمرات دانش آموزان پسر

نمرات	تعداد
A	۸
B	۴
C	۳
D	۷
F	۸



نمودار نمرات دانش آموزان دختر

نمرات	تعداد
A	۸
B	۸
C	۰
D	۷
F	۷



$$\text{درصد قبولی دختران} : \frac{22}{30} \times 100$$

$$\text{درصد قبولی پسران} : \frac{23}{30} \times 100$$

در نتیجه درصد قبولی دختران بیشتر از درصد قبولی پسران است.

برگ کلاس ب	ساقه	برگ کلاس الف	۶- الف
۲	۱۱	۶۷۹	
۹۸۸۶۲	۱۲	۵۹	
۳۱	۱۳	۰۱۳۴۵۷۷۷	کلید نمودار
۷۶۴۴۳۲۰	۱۴	۰۲۴۶۷۷	سانتی متر = ۱۱۶ = ۶ ۱۱
۹۸۷۶۵۴	۱۵	۱	
۹۶۵۴	۱۶	۰۳۸۹	

ب- کلاس ب با ۱۱۲ سانتی متر

ج- ۱۶۹ سانتی متر در هر دو کلاس

د- قد هیچکدام از دانش آموزان کمتر از ۱۱۰ سانتی متر نیست.

ه- ۲۱ نفر از کلاس الف و ۲۴ نفر از کلاس ب

شاخصهای مرکزی

فصل ۶- جلسه اول

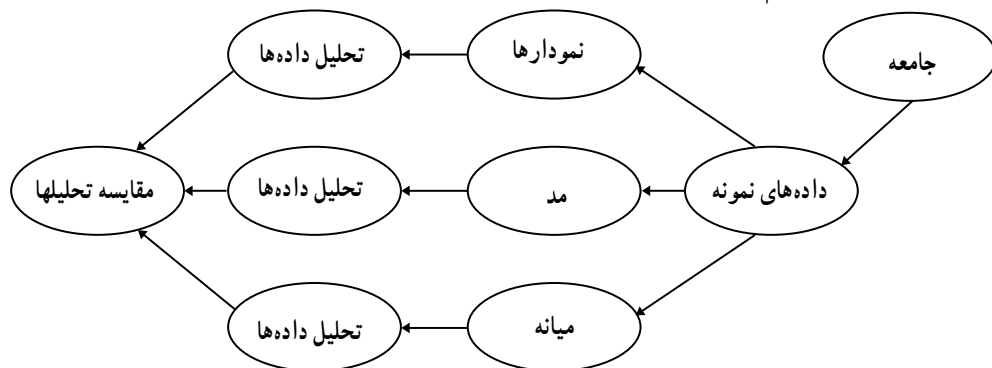
مفاهیم: شاخص مرکزی، مد، میانه، جامعه چند مدی
مهارت‌ها: مرتب کردن داده‌ها، تشخیص اعتبار شاخص‌ها، محاسبه میانه
اهداف:

- تشخیص جوامع یک‌دست و غیر یک‌دست
 - استفاده از شاخصهای عددی برای تحلیل داده‌ها
 - تشخیص اعتبار شاخص‌ها با توجه به مسئله مورد تحقیق
- تأکید درس: مقایسه میانه و مد و تحلیل اطلاعات در مورد وضع نسبی آنها مورد تأکید است چون اگر دانش‌آموزان بتوانند دو مفهوم را به خوبی مقایسه کنند، تا حد خوبی آن‌ها را می‌شناسند. به‌علاوه کم‌کم از شاخص‌های هندسی به سمت درک شاخصهای عددی میل می‌کنند زیرا با مقایسه وضع نسبی شاخصهای عددی سعی می‌شود تعبیری هندسی از این شاخص‌ها معرفی شود. تشخیص جوامع یک‌دست و غیر یک‌دست از یکدیگر مورد تأکید است. جداول فراوانی چند کوهانه نماینده جوامع غیر یک‌دست هستند. سعی کنید مثالهایی از جوامع غیر هم‌دست را در اختیار دانش‌آموزان قرار دهید و از آن‌ها بخواهید سعی کنند با استفاده از داده‌ها و جدول فراوانی جامعه را به بخشهای یک‌دست تقسیم کنند. هم‌چنین بر تحلیل داده‌ها توسط شاخص‌های مرکزی و بر کاربرد آن در مورد تصمیم‌گیری و پیشگویی تأکید شده است. خوب است نمونه‌هایی از تصمیم‌گیری‌هایی که براساس شاخص‌های مرکزی صورت می‌گیرد توسط دانش‌آموزان از وسایل ارتباط جمعی استخراج شود.
- بحث: اولین قدم پس از معرفی شاخص‌های مرکزی این است که دانش‌آموزان برای هر یک از آن‌ها مصداق‌هایی در زندگی روزمره بیابند. آشنایی با الگوریتم‌هایی برای مرتب کردن داده‌ها درک بهتری از اینکه محاسبه میانه چه حجمی از محاسبات را دربردارد، در اختیار دانش‌آموزان قرار می‌دهد. از دانش‌آموزان بخواهید توانایی تحلیل توسط شاخص‌های مرکزی را با شاخص‌های هندسی مقایسه کنند و در کلاس به بحث بگذارند. خوب است دانش‌آموزان سعی کنند خودشان شاخص‌هایی

مرکزی برای خلاصه کردن اطلاعات معرفی کنند. این فعالیت باعث می‌شود در تحلیل میانه و مد و مقایسه آن‌ها تجربه بیشتری پیدا کنند.

مباحث ریاضی مربوط: جامعه‌آمار، شاخص‌های عددی

نمودار مفاهیم:



فصل ۶ — جلسه دوم

مفاهیم: کمترین مقدار داده، بیشترین مقدار داده، چارک اول، چارک سوم
 مهارت‌ها: رسم و تحلیل نمودار جعبه‌ای، مقایسه نمودار جعبه‌ای و جدول فراوانی
 اهداف:

- تحلیل نمودار جعبه‌ای و مقایسه توانایی‌های آن با جدول فراوانی
- استخراج اطلاعاتی درباره توزیع داده‌ها با استفاده از شاخص‌های مرکزی
- درک تأثیر تغییرات داده‌ها بر شاخص‌های مرکزی

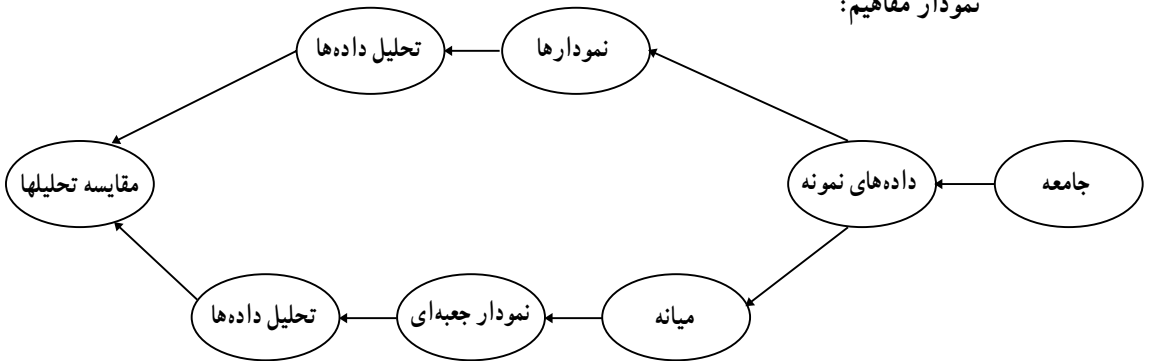
تأکید درس: نمودار جعبه‌ای به نوعی شاخص‌های عددی و هندسی را ترکیب می‌کند. درک و بررسی مفاهیم توزیع داده‌ها مورد تأکید است. این که با داشتن میانه، و در دست داشتن نمودار جعبه‌ای هر کدام تا چه حد توزیع داده‌ها را مشخص می‌کنند باید مورد بحث قرار گیرد. مثلاً داده‌هایی مثال زده شوند که دارای نمودار جعبه‌ای یکسانی باشند ولی دارای توزیع متفاوتی باشند. اینکه تغییرات داده‌ها چه تأثیری بر شاخص‌های مرکزی می‌گذارد باید مورد تأکید باشد. این بررسی موجب درک بهتر ضعف‌های نمودار جعبه‌ای برای تحلیل داده‌ها می‌شود. آشنایی مقدماتی با مفاهیم مربوط به پراکنندگی داده‌ها با توجه به نمودار جعبه‌ای شهود هندسی خوبی در اختیار دانش‌آموزان قرار می‌دهد. از دانش‌آموزان بخواهید نمودار جعبه‌ای و تحلیل‌های حاصل از آن را با نمودارهای مستطیلی، ساقه و برگ، دایره و غیره مقایسه کنند. این مقایسه باید در چارچوب حل مسئله با انجام کامل همه مراحل صورت بگیرد. مقایسه نمودارهای جعبه‌ای مختلف برای تحلیل داده‌ها مورد تأکید است. مثلاً کارکرد

دو کلاس مختلف را توسط نمودارهای جعبه‌ای آن‌ها مقایسه کنید.

بحث: اگر تعداد داده‌ها زیاد باشد مرتب کردن داده‌ها کار دشواری است. به همین دلیل همیشه نمودار جعبه‌ای برای تحلیل کاربرد ندارد. استفاده از نرم افزارها برای مرتب کردن داده‌ها توصیه می‌شود. دانش‌آموزان باید بتوانند تحلیل‌های نادرست از نمودار جعبه‌ای را ابطال کنند. دانش‌آموزان را تشویق کنید خودشان شاخص‌های مرکزی و هندسی را ترکیب کنند و نموداری برای نمایش آن معرفی کنند. تحلیل شاخص‌های مرکزی با این تمرین برای دانش‌آموزان آسان‌تر می‌شود. اگر جامعه همگن نباشد چه اشکالی در تحلیل نمودار جعبه‌ای پیدا می‌شود؟ این سؤال را در کلاس به بحث بگذارید.

مباحث ریاضی مربوط: نمودارها، تحلیل داده‌ها، جامعه آماری

نمودار مفاهیم:



فصل ۶- جلسه سوم

مفاهیم: میانگین، میانگین وزنی، جدول فراوانی

مهارت‌ها: محاسبه میانگین، مقایسه میانگین، میانه و مد

اهداف:

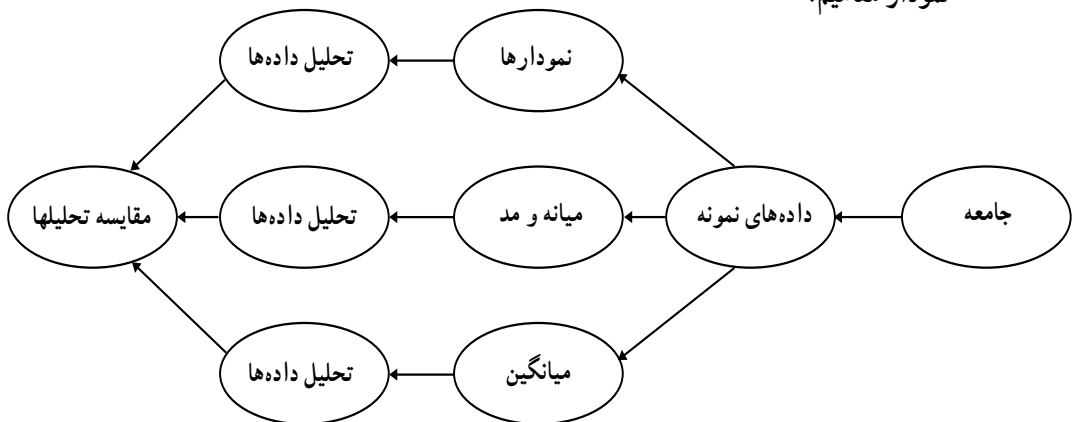
- کاربرد شاخص‌های مرکزی در زندگی روزمره
- استفاده از شاخص‌های مرکزی برای ارزشیابی
- درک حساسیت متفاوت شاخص‌های مرکزی نسبت به تغییرات داده‌ها

تأکید درس: دانش‌آموزان باید بتوانند شاخص مرکزی مناسب مسئله را انتخاب کنند. خوب است دانش‌آموزان سعی کنند مصداق‌های شاخص میانگین درس در زندگی روزمره را بشناسند. نقد معدل به عنوان ارزشیابی دانش‌آموزان را در کلاس به بحث بگذارید و از آن‌ها بخواهید شاخص‌های بهتری برای ارزشیابی کار دانش‌آموزان معرفی کنند. چگونگی استفاده از میانگین برای حل مسائل آماری مورد تأکید است. ارتباط بین دانش‌آموزان نقش مهمی در حل مسائل آماری دارد. در بسیاری

اوقات نمایش‌های مختلف داده‌ها پیام‌های کاملاً متفاوتی می‌دهند. بررسی و تحقیق بیشتر ممکن است بعضی از این پیام‌ها را بی‌اعتبار تشخیص دهد. روند قضاوت و تحلیل دانش‌آموزان تنها با ارتباطات آن‌ها رشد می‌کند. همچنین این ارتباطات باعث می‌شود دانش‌آموزان در مورد مفاهیم آماری شهود پیدا کنند و این به ارزشیابی تحلیلی‌های آماری کمک می‌کند. مقایسه میانگین، میانه و مد و وضعیت نسبی آن‌ها نیز مورد تأکید است. این مقایسه به درک بهتر کاربرد هر کدام از این شاخص‌ها کمک می‌کند. نقد تحلیل داده‌ها در اینجا نیز مورد تأکید است.

بحث: حساسیت شاخص‌های مرکزی را نسبت به تغییرات داده‌ها مقایسه کنید. این مقایسه حساسیت بیشتر میانگین نسبت به میانه را در برابر تغییرات داده‌ها به خوبی به نمایش می‌گذارد. مفهوم میانگین کاربردهای بسیار وسیعی در زندگی روزمره دارد. از کلاس بخواهید که در مورد این کاربردها بحث کنند و یا راهبردهایی برای زندگی بهتر با استفاده از مفهوم میانگین ارائه دهند. منحنی نرمال و اینکه در اکثر اندازه‌گیری‌های طبیعی این منحنی ظاهر می‌شود مطرح شود. همچنین بحث شود که چه وزن‌هایی به دروس مختلف برای ارزشیابی صحیح هنگام محاسبه معدل باید نسبت داده شود.

مباحث ریاضی مربوط: شاخص‌های مرکزی، اندازه‌گیری، تحلیل داده‌ها، خلاصه کردن داده‌ها نمودار مفاهیم:



شاخصهای مرکزی

در تشکیل جدولهای فراوانی گفتیم که با در نظر گرفتن مرکز دسته‌ها به تعداد فراوانی‌های نظیر آن دسته‌ها داده‌های متفرق و متفاوت را تبدیل به چند مقدار که همان مراکز دسته‌ها هستند می‌کنیم و به این وسیله با حذف اختلاف‌های جزئی، حجم وسیعی از داده‌ها و اطلاعات جزئی را به تعداد کمی از مقادیر تبدیل می‌کنیم. دیدیم که می‌توانیم برحسب وضعیت هریک از این دسته‌ها به آن‌ها برچسب

بزنیم. مثلاً به کسانی که بهره هوشی آنها از ۱۲۰ بیشتر است بگوییم دانش‌آموزان باهوش و دسته دانش‌آموزان ضعیف به کسانی اطلاق شود که نمره‌ای کمتر از ده آورده‌اند و غیره. یکی از هدفها در آمار این است که تا جایی که ممکن است داده‌ها را کم کرد بی‌آنکه اطلاعات زیادی از دست برود. فایده‌ای که این عمل دارد در این است که می‌توان از میان انبوه داده‌ها شاخصی استخراج کنیم که جامعه را به ما بشناساند. یک دسته از این شاخص‌ها، شاخص‌هایی هستند که مرکزیت داده‌ها را مشخص می‌کنند. معلوم می‌کنند که داده‌ها در کجا جمع شده‌اند. مهمترین این شاخص‌ها میانگین است. البته همانطوری که در کتاب درسی توضیح داده شده ممکن است در مسائلی میانه و یا مد اهمیت بیشتری داشته باشد. ولی حسن میانگین در این است که اولاً تمام داده‌ها مستقیماً در آن دخیل هستند و ثانیاً ویژگیهای خوب ریاضی دارد و قضایای مهمی درباره آن ثابت شده است. به این علت است که میانگین جایگاه ویژه‌ای در بین شاخص‌های مرکزی پیدا کرده است. ذیلاً قوانین مربوط به میانگین را بدون اثبات می‌نویسیم:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{X_1 + \dots + X_n}{n} \quad (\text{تعریف میانگین})$$

$$\overline{aX + b} = a\bar{X} + b$$

$$\sum_{i=1}^n (X_i - \bar{X}) = 0$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

اگر صلاح دانستید به دانش‌آموزان تمرین‌هایی نظری مانند تمرین‌های زیر بدهید.

۱- میانگین داده‌های x_1, \dots, x_n برابر a و میانگین داده‌های y_1, \dots, y_n برابر b است.

میانگین $x_1, \dots, x_n, y_1, \dots, y_n$ چقدر است؟

مثلاً میانگین نمرات درس آمار و مدل‌سازی در کلاس ۳۵ نفری A برابر ۱۶ و در کلاس ۴۰ نفری B برابر ۱۵ است. اگر دانش‌آموزان این دو کلاس را روی هم در نظر بگیریم، میانگین نمره درس آمار و مدل‌سازی این $75 = (35 + 40)$ دانش‌آموز چقدر خواهد بود؟

۲- پس از محاسبه میانگین داده‌های x_1, \dots, x_4 معلوم شده است که داده‌ی x_1 باید برابر

x_4 باشد، در این صورت میانگین چه مقداری خواهد بود؟

مثلاً، میانگین نمرات درس آمار و مدل‌سازی کلاس ۳۵ نفری A برابر ۱۶ است، بعداً معلوم می‌شود که نمره دانش‌آموزی بجای آن که ۱۶ باشد ۱۴ منظور شده است. میانگین نمره کلاس چقدر خواهد بود؟

۳- این تمرین یکی از ویژگیهای میانگین است. اگر میانگین داده‌های x_1, \dots, x_n برابر a باشد، میانگین داده‌های x_1, \dots, x_n, x_{n+1} چقدر است؟

حال برسید x_{n+1} چقدر باشد تا میانگین این دو دسته داده تغییر نکند.

حل: می‌دانیم

$$a = \frac{x_1 + \dots + x_n}{n} \Rightarrow x_1 + \dots + x_n = na$$

$$b = \frac{x_1 + \dots + x_n + x_{n+1}}{n+1} \quad \text{می‌خواهیم}$$

را حساب کنیم، داریم:

$$b = \frac{na + x_{n+1}}{n+1} = \frac{n}{n+1}a + \frac{x_{n+1}}{n+1}$$

مثلاً میانگین ۳۵ داده برابر ۱۸ است، بعداً متوجه می‌شویم یک یک داده که اندازه آن برابر ۱۷ است از قلم افتاده، حال میانگین چقدر می‌شود. در این مثال داریم:

$$a = 18, \quad n = 35, \quad x_{n+1} = 17$$

پس

$$b = \frac{35}{36} \times 18 + \frac{17}{36}$$

اگر بخواهیم a و b مساوی باشند باید داشته باشیم

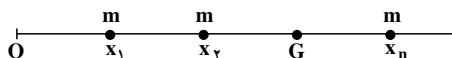
$$\frac{n}{n+1}a + \frac{x_{n+1}}{n+1} = a$$

از اینجا خواهیم داشت:

$$na + x_{n+1} = na + a$$

پس $x_{n+1} = a$. یعنی اگر مقداری که به داده‌ها اضافه می‌شود برابر میانگین باشد میانگین داده‌های جدید تغییر نمی‌کند.

یکی از تعبیرهایی که از میانگین می‌شود عبارت است از مرکز ثقل چند گلوله به وزنهای مساوی که در نقاط x_1, x_2, \dots, x_n قرار داده شده باشند.

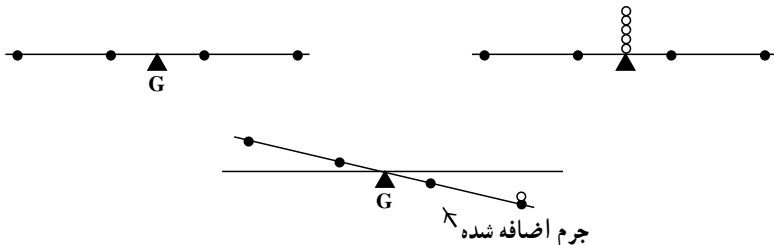


مرکز ثقل این گلوله‌ها را در نقطه G فرض کنیم بنابر قوانین فیزیک داریم:

$$OG \times (m + m + \dots + m) = mx_1 + mx_2 + \dots + mx_n$$

از اینجا $OG = \bar{x}$ به دست می‌آید.

مرکز ثقل نقطه‌ای است که اگر دستگاه را در آن نقطه روی پایه‌ای قرار دهیم دارای تعادل خواهد بود. حال اگر وزنه دیگری به جایی دیگر از میله اضافه کنیم تعادل آن به هم خواهد خورد (مرکز ثقل دستگاه عوض خواهد شد). اگر مرکز این جرم جدید را روی مرکز ثقل قرار دهیم در این صورت تعادل محفوظ خواهد ماند.



تمرین: از دانش‌آموزان بخواهید با استفاده از جدول فراوانی میانگین را حساب کنند و کم‌کم آنها را به سمت دستور زیر

$$\bar{x} = \sum_{i=1}^n \varphi_i x_i$$

که در آن φ_i فراوانی نسبی است، هدایت کنید.
$$\left[\bar{x} = \frac{1}{n} \sum f_i x_i = \sum \frac{f_i}{n} x_i = \sum \varphi_i x_i \right]$$

همانطوری که می‌دانید در یک جدول توزیع احتمال اگر متغیر تصادفی X مقادیر x_1, x_2, \dots

را با احتمالهای $\varphi_1, \varphi_2, \dots$ اختیار کند آنگاه امید ریاضی برابر است با

$$E(X) = \sum_{i=1}^n \varphi_i \cdot x_i$$

البته شما در این مورد با دانش‌آموزان صحبت نکنید. همین که توجه شود این مطالب علاوه بر استفاده‌های مفهومی از میانگین، استفاده‌های دیگری بعداً خواهد داشت اهمیت آنها را روشن می‌کند.

برابری $\sum (x_i - \bar{x}) = 0$ گاهی ممکن است سؤال برانگیز شود. معمولاً دانش‌آموزان \bar{x} را از

دستور $\frac{1}{n} \sum f_i x_i$ حساب می‌کنند ولی در رابطه اخیر مرکز دسته‌ها را قرار می‌دهند و ملاحظه می‌کنند

که حاصل برابر صفر نیست. این قبیل دانش‌آموزان فراموش کرده‌اند که باید مرکز دسته‌ها را به تعداد فراوانی دسته‌ها تکرار کنند یعنی باید مجموعه مقادیر زیر را حساب کنند.

$$f_1 \begin{cases} x_1 - \bar{x} \\ x_1 - \bar{x} \\ \vdots \\ x_1 - \bar{x} \end{cases}, \quad f_2 \begin{cases} x_2 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_2 - \bar{x} \end{cases}, \quad \dots$$

که در این صورت خواهیم داشت :

$$\sum f_i (x_i - \bar{x}) = 0$$

تمرین: به دانش‌آموزان بگویید که دیدیم میانگین مقداری است که مجموع انحرافات از آن برابر صفر است، یعنی $\sum (x_i - \bar{x}) = 0$. حال اگر مقداری مانند a چنان باشد که مجموع انحرافات از a برابر صفر باشد آیا a باید برابر میانگین باشد؟ قبل از آنکه پاسخ سؤال را بدهید اجازه دهید دانش‌آموزان توجیه‌های خود را ارائه کنند. از دانش‌آموزان کم‌حرف و خجول هم بپرسید تا حقی از آن‌ها هم ضایع نشود و به این ترتیب در گفتگوهای کلاس شرکت کنند. البته با مثالهایی خواهید دید که اگر a برابر میانگین نباشد مجموع انحرافات برابر صفر نخواهد شد. حال آن را به صورت زیر ثابت کنید:

$$\begin{aligned} \sum (x_i - a) &= (x_1 - a) + (x_2 - a) + \dots + (x_n - a) \\ &= x_1 + \dots + x_n - na \\ a &= \frac{x_1 + \dots + x_n}{n} = \bar{x} \end{aligned}$$

یعنی

پس شرط لازم و کافی برای آن که مجموع انحرافات داده‌ها از مقداری مانند a برابر صفر شود آن است که a برابر میانگین داده‌ها باشد.

اگر در کلاس علوم تجربی تدریس می‌کنید می‌توانید از مثالهای جالب مربوط به رشته آنها استفاده کنید، مثلاً از دانش‌آموزان بخواهید محیط تنه درختان چنار محله خود را اندازه بگیرند (به فاصله یک متر از سطح زمین) و جدول فراوانی و میانگین این محیط‌ها را حساب کنند. آیا این محیط‌ها می‌تواند مبنایی برای تشخیص سن درخت باشد؟ (به این سؤال باید به کمک معلم زیست‌شناسی پاسخ دهند) این مسأله را می‌توانند برای درختان مختلف تکرار کنند. می‌توانند تعداد افراد خانواده‌های هم‌کلاسی‌های خود را بپرسند و جدول مربوطه را تشکیل و میانگین تعداد نفرات را حساب کنند.

از جمله مسائل جالبی که از آمار در ادبیات استفاده می‌شود عبارت است از میانگین طول جملاتی که یک نویسنده به کار می‌برد. می‌توانید از دانش‌آموزان بخواهید، طول جملات کتاب آمار و مدل‌سازی را بررسی کنند طول جمله عبارت است از تعداد کلماتی که در آن جمله به کار رفته است. از این کتاب یک نمونه‌گیری تصادفی کنند. طول جملات را بشمارند، جدول فراوانی را تشکیل و میانگین را حساب کنند. این کار را با کتاب فارسی انجام دهند و قضاوت کنند در کدام یک از این دو کتاب جملات کوتاه‌تری به کار رفته است؟

یکی دیگر از شاخص‌های مرکزی مد است. مد داده‌ای است که بیشترین تکرار را دارد. روش محاسبه آن همان است که در کتاب توضیح داده شده است. ممکن است مانند میانگین سؤال

شود چگونه می‌توان مد را از جدول فراوانی حساب کرد. ما این روش را در کتاب درسی نگفته‌ایم ولی در اینجا به ذکر آن می‌پردازیم. ضرورتی ندارد که این مطلب در کلاس گفته شود. روش محاسبه مد را در جدول فراوانی زیر توضیح می‌دهیم.

مراحل یافتن مد عبارت است از

دسته‌ها	فراوانی‌ها
۸-۱۰	۳
۱۰-۱۲	۵
۱۲-۱۴	۹
۱۴-۱۶	۶
۱۶-۱۸	۴
جمع	۲۷

۱- یافتن دسته مد. دسته‌ای است که بیشترین

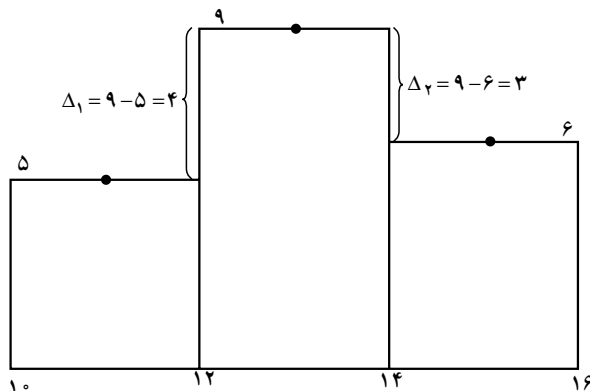
فراوانی را دارد. در مثال مورد بحث دسته مد دسته سوم است که فراوانی آن ۹ است که از سایر فراوانی‌ها بیشتر است.

۲- حد پایین دسته مد را مشخص کنید. در مثال

مورد بحث این حد عبارت است از ۱۲.

۳- نمودار مستطیلی جدول را رسم می‌کنیم، فقط

سه دسته‌ای را در نظر می‌گیریم که دسته مد در وسط قرار می‌گیرد.



۴- حال مد به دسته‌ای نزدیکتر است که فراوانی آن بیشتر باشد، در شکل مد به دسته ۱۴-۱۶

نزدیکتر است، از این رو دسته مد را به نسبت عکس تفاضل فراوانی‌های دسته مد از دو دسته مجاور تقسیم می‌کنیم. اگر این اختلاف فراوانی‌ها را با Δ_1 و Δ_2 نشان دهیم و اگر Δ_1 از Δ_2 بیشتر باشد، آن‌گاه مد از رابطه تقریبی زیر به دست می‌آید:

$$\hat{X} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} C$$

که در آن L حد پایین دسته وسط و C طول دسته است.

در مثال مورد بحث مد عبارت است از :

$$\begin{aligned}\hat{X} &= 12 + \frac{4}{4+2} \times 2 \\ &= 12 + \frac{8}{6} \approx 13/33\end{aligned}$$

تمرین: متنی از یک نویسنده را در اختیار دانش‌آموزان قرار دهید و از آن‌ها بخواهید که تعیین کنند این نویسنده از کدام کلمه بیش از همه استفاده کرده است، در واقع مد کلمات او را بیابند. اخیراً نرم افزارهای آماری درباره قرآن تهیه شده است که با مراجعه به آنها می‌توان مد کلماتی را که در هر سوره و یا حتی در قرآن به کار رفته است بیابند.

در برخی از مسائل میانگین نمی‌تواند مبنای تصمیم‌گیری قرار گیرد. مثلاً در مسأله انتخابات میانگین اصلاً موضوعیت ندارد. آنچه که به انتخاب کاندیدایی منجر می‌شود، مد است. یعنی فردی که بیشترین رأی را آورده است انتخاب می‌شود.

همانطوری که در کتاب درسی اشاره شد، فروشندگان مطابق فصل و یا رویدادهای روز اجناس مغازه خود را تنظیم می‌کنند. کالایی را در حیطة کاری خود عرضه می‌کنند، که بیشتر مورد نیاز مردم است. یعنی آنها روی مد برنامه‌ریزی می‌کنند. وقتی فردی می‌خواهد مغازه‌ای باز کند، به دنبال مکانی می‌گردد مشتری بیشتری داشته باشد، یعنی به مد مراجعین توجه می‌کند.

مد ویژگیهای خاص میانگین را ندارد مع الوصف برخی از آن‌ها را در زیر فهرست می‌کنیم. اگر مایل بودید می‌توانید به عنوان تمرین به دانش‌آموزان بدهید.

الف - مد هم مانند میانگین از قانون خطی بودن پیروی می‌کند. یعنی اگر مد داده‌های x_1, \dots, x_n

را با \hat{X} و مد داده‌های $ax_1 + b, \dots, ax_n + b$ را با $\hat{ax} + b$ نشان دهیم، آن‌گاه

$$\hat{ax} + b = a\hat{x} + b$$

دلیل این برابری هم واضح است زیرا اگر مقدار x بیش از سایر مقادیر تکرار شده باشد آن‌گاه مقدار $ax + b$ هم بیش از سایر مقادیر تکرار خواهد شد.

ب - همانطوری که در کتاب درسی توضیح داده شد، مد با تغییر داده‌ها ممکن است ثابت بماند، مثلاً دو دسته داده زیر دارای مدهای برابر هستند.

۱۰، ۱۱، ۱۲، ۱۲، ۱۲، ۱۳، ۱۵

۴، ۶، ۷، ۱۲، ۱۲، ۱۲، ۲۰، ۳۰

بنابراین مد در همه موارد به مقدار تک‌تک داده‌ها پاسخ نمی‌دهد ولی در میانگین مقدار تک‌تک داده‌ها مؤثر بودند.

ج - نمی‌توان دربارهٔ مد قضایا و روابط خوب و وسیعی که درباره میانگین وجود دارد، ثابت کرد. بنابراین نظریهٔ مربوط به مد به گستردگی نظریهٔ مربوط به میانگین نیست.

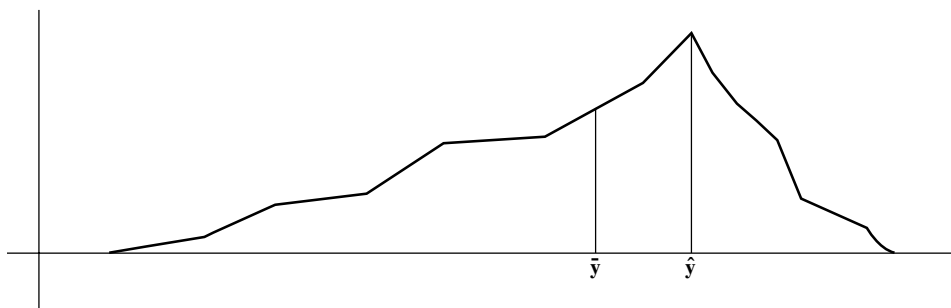
د - اگر فراوانی مقادیر به یک نسبت کم و یا زیاد مثلاً نصف یا دو برابر شوند، مد تغییر نخواهد کرد.

از دانش‌آموزان بخواهید مد قد، و مد وزن آن‌ها و مد نمره درس آمار و مدل‌سازی را حساب کنند. اگر مد کمتر از میانگین باشد چه نتیجه‌ای از آن حاصل می‌شود. مقایسهٔ این پارامترها سبب می‌شود که شکل کلی جامعه را بتوان مجسم کرد.

اگر مد کمتر از میانگین باشد، چند بر فراوانی به صورت زیر خواهد بود.

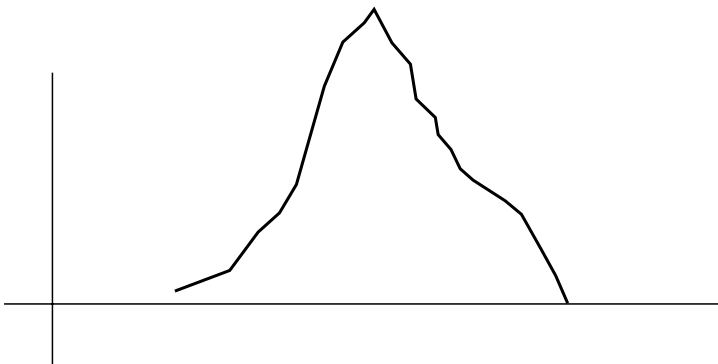
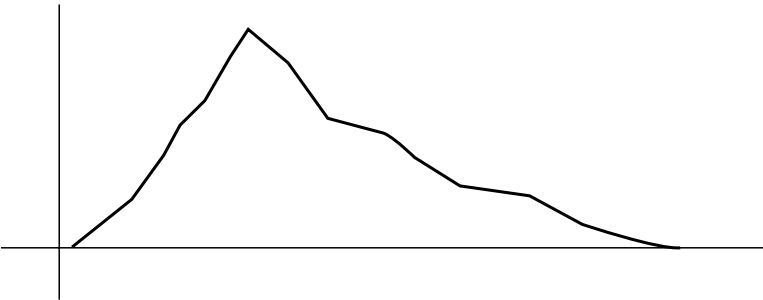
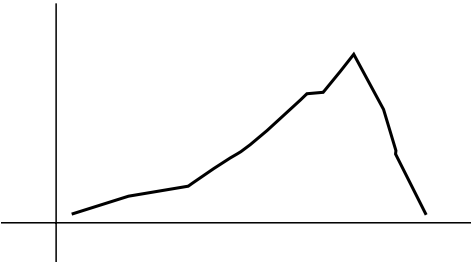
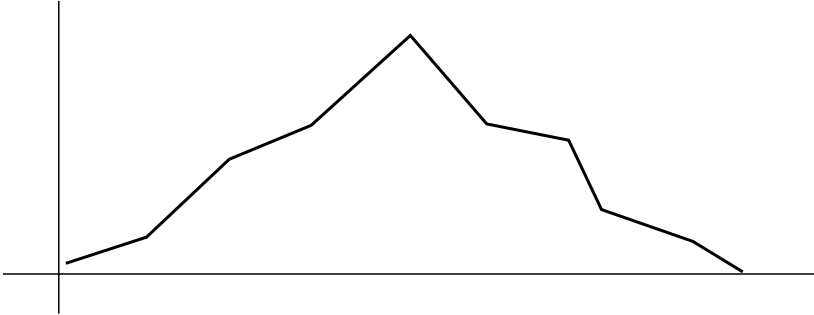


چون بیشترین فراوانی‌ها در سمت چپ میانگین هستند، برای حفظ تعادل لازم است میانگین به مقادیر کمتر نزدیکتر باشد و در نتیجه در سمت راست میانگین نمودار کشیده‌تر است و مقادیر بسیار بزرگتر با فراوانی‌های کم در سمت راست میانگین دیده می‌شود. اگر نمره درس در یک کلاس به این صورت باشد، معنای آن چیست؟ این شکل نشان می‌دهد در کلاس چند نمرهٔ خیلی بالا وجود دارد و بیشتر دانش‌آموزان زیر میانگین هستند. اما اگر نمودار چند بر فراوانی به صورت زیر باشد.

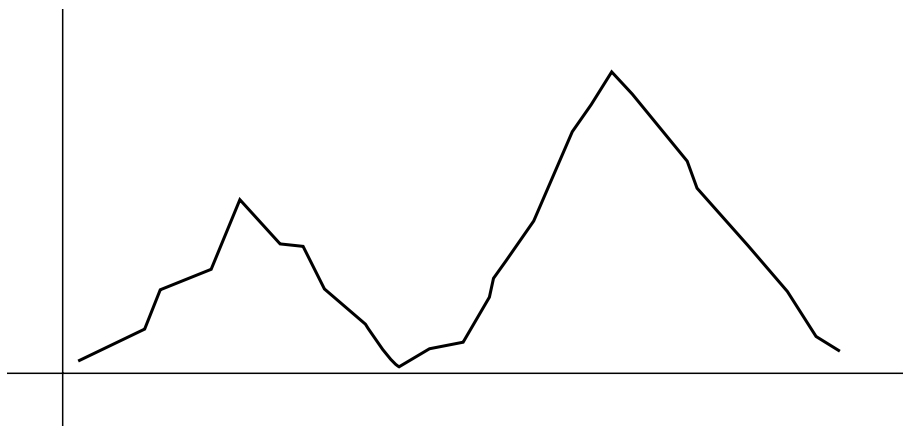


نتیجه می‌شود که چند نمره خیلی پایین در کلاس وجود دارد و بیشتر دانش‌آموزان قوی هستند. در چنین شرایطی باید فکری برای آن دانش‌آموزان خیلی قوی و یا برای این دانش‌آموزان

خیلی ضعیف کرد. با برگزار کردن کلاسهای فوق برنامه، مطالعه آزاد از هدر رفتن انرژی دانش آموزان قوی می توان جلوگیری کرد و همچنین از عقب ماندن دانش آموزان ضعیف هم کاست. شما به عنوان معلم مایلید که وضع نمره کلاس شما به کدام صورت زیر باشد.



اگر نمودار نمره کلاس شما به صورت زیر باشد چه نتیجه‌ای می‌گیرید؟



میانه

گفتیم میانه داده‌ای است که جامعه را نصف می‌کند. این جمله بسیار عام و کلی است. ممکن است بگوییم میانه داده‌ای است که نصف داده‌ها از آن بیشتر و نصف داده‌ها از آن کمتر باشند. اگر چه این جملات مفهوم را می‌رساند ولی دقیق نیستند. تعریف دقیق میانه عبارت است از مقداری که تعداد داده‌های بعد از آن برابر تعداد داده‌های قبل از آن باشند. از دانش‌آموزان بخواهید اشکالی را که در عبارات قبلی وجود دارد روشن کنند، تا در این صورت اشکال عبارتی را که به کار می‌برند بدانند. برای یافتن میانه لازم است داده‌ها از کوچک و بزرگ و یا از بزرگ به کوچک مرتب شوند. در این صورت داده‌ای که در وسط قرار می‌گیرد (در حالتی که تعداد داده‌ها فرد باشند) و یا میانگین دو داده‌ای که در وسط قرار می‌گیرند (در حالتی که تعداد داده‌ها زوج باشند) برابر میانه است. میانه هم مانند مد دارای ویژگی خطی است

$$a\bar{x} + b = a\bar{x} + b$$

همچنین اگر فراوانی‌ها به یک نسبت کم و یا زیاد شوند، میانه تغییر نمی‌کند. میانه هم مانند مد در همه موارد به تغییرات داده‌ها پاسخ نمی‌دهد. درباره میانه هم قضایا و روابط وسیع مانند میانگین وجود ندارد.

میانه و تعمیم‌هایی از آن مانند چهارکها، دهکها و صدکها بیشتر در مسائل اقتصادی استفاده می‌شوند. میانه جامعه را نصف می‌کند، دهکها و صدکها به ترتیب جامعه را به ده و صد قسمت مساوی تقسیم می‌کنند. مثلاً می‌خواهند بدانند چه درصد از افراد جامعه از یک دهم درآمد ملی بهره می‌برند، چه تعدادی از دو دهم و الی آخر و از این‌جا درباره توزیع عادلانه ثروت قضاوت می‌کنند. ما وارد بحثهای چارکها و غیره نمی‌شویم و فقط درباره میانه صحبت می‌کنیم. مثلاً اگر بخواهیم بدانیم نمره

نصف دانش‌آموزان کلاس از چه نمره‌ای کمتر است از میانه استفاده می‌کنیم. اگر این میانه درباره نمرات کلاس عدد بزرگی باشد مثلاً ۱۷، در این صورت نتیجه می‌گیریم که رقابت در این کلاس بسیار شدید است زیرا نمره نصف دانش‌آموزان در فاصله‌ای به طول ۳ (از ۱۷ تا ۲۰) متمرکز شده است، اما اگر میانه ۱۳ باشد، نتیجه می‌گیریم که رقابت چندان شدید نیست. به کمک دستورهایی می‌توان میانه را از روی جدول فراوانی هم حساب کرد مراحل آن چنین است.

۱- ابتدا دسته میانه را پیدا می‌کنیم، این دسته، دسته‌ای است که مجموع فراوانی‌های آن دسته و دسته‌های ماقبل آن برای اولین بار از $\frac{n}{4}$ تجاوز می‌کند. مجموع فوق‌الذکر را فراوانی تجمعی دسته می‌گوییم.

و ما در متن صحبتی از آن به میان نیاوردیم.

۲- با استفاده از دستور زیر میانه را حساب می‌کنیم.

$$\tilde{Y} = L + C \frac{\frac{n}{4} - F}{f}$$

که در این دستور n تعداد داده‌ها، L حد پایین دسته میانه، F فراوانی تجمعی دسته ماقبل دسته میانه، C طول دسته‌ها و f فراوانی مطلق دسته میانه است.

مثال: در جدول روبرو میانه را بیابید:

در این مثال $n = 20$

پس $\frac{n}{4} = 5$ و اولین دسته‌ای که فراوانی تجمعی

آن از ۵ تجاوز می‌کند دسته ۸-۱۰ است پس

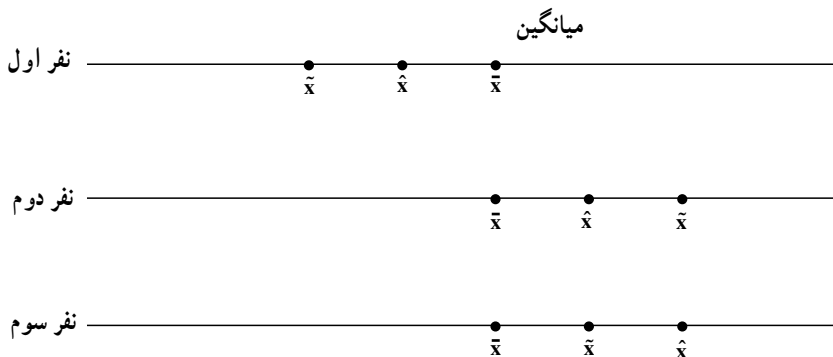
$L = 8$ ، $C = 2$ ، $F = 6$ و $f = 14$ بنابراین

$$\tilde{X} = 8 + 2 \frac{5 - 6}{14} = 8 + \frac{4}{7} \approx 8.57$$

البته زمانی این دستور و دستور مد بسیار با ارزش بودند، زیرا در آن روزگار به علت نبودن وسایل سریع محاسبه می‌بایستی از دستورهایی که تا حدودی کارها را ساده کند استفاده نمایند. در روزگار ما این قبیل محاسبات زاید است، زیرا در رایانه‌ها به آسانی شاخص‌های موردنظر را محاسبه می‌کنند. به جای آن که فرمول اخیر را به دانش‌آموزان ارائه کنید، آن‌ها را تشویق و راهنمایی کنید تا از رایانه‌ها استفاده کنند.

میانه معمولاً در سمتی از میانگین قرار دارد که مد در آن سمت است. میانه و مد به هم

نزدیک‌اند. ما برای این جمله اخیر اثبات ریاضی نداریم. می‌توانید کارهایی که دانش‌آموزان انجام داده‌اند نمودار را باهم مقایسه و محل قرار گرفتن میانه، مد و میانگین را باهم مقایسه کنید، آنچه که در بالا گفته شد در بیشتر نمودارها رخ خواهد داد. از دانش‌آموزان بخواهید تا روی یک صفحه کاغذ میانگین، مد و میانه را برای مسائل خود رسم کنند تا شکلی به صورت زیر حاصل شود.



میانگین‌ها را زیر هم قرار دهند، بعد از محل قرار گرفتن میانه و مد راجع به آن جامعه قضاوت کنند. مسائل دانش‌آموزان ممکن است متفاوت باشند، آنچه که برای ما در این مسأله مهم است ترتیب قرار گرفتن این مد شاخص مرکزی است. اگر میانگین، میانه و مد باهم برابر باشند جامعه را متقارن می‌گوییم.

توضیح فعالیت‌ها

فعالیت ۱ صفحه ۱۲۰

ابتدا داده‌ها را مرتب می‌کنیم

۲۳، ۲۵، ۲۷، ۳۱، ۳۲، ۳۴، ۴۱، ۴۳، ۱۰، ۱۲، ۱۴، ۱۵، ۱۶، ۱۸، ۱۹، ۲۳، ۲۵، ۲۷، ۳۱، ۳۲

است.

— اعدادی که در نیمه قبل از میانه هستند عبارتند از: ۱۰، ۱۲، ۱۴، ۱۵، ۱۶، ۱۸، ۱۹

— میانه داده‌های نیمه اول برابر ۱۵ است.

— اعدادی که در نیمه بعد از میانه هستند عبارتند از: ۲۵، ۲۷، ۳۱، ۳۲، ۳۴، ۴۱، ۴۳

— میانه داده‌های نیمه آخر برابر ۳۲ است.

فعالیت ۲ صفحه ۱۲۱

- میانه برابر ۹ است.
- بین ۱ و ۸ یک چهارم داده‌ها قرار گرفته‌اند.
- چارک سوم ۱۱ است و بین ۱۱ و ۱۲ یک چهارم داده‌ها قرار داده‌اند.
- تمرکز داده‌های درون جعبه در سمت چپ بیشتر از سمت راست است.
- چارک اول داده‌ها در دنباله سمت چپ پراکنده‌ترند تا چارک آخر داده‌ها در دنباله سمت راست.

فعالیت ۳ صفحه ۱۲۲

- کلاس B بیشترین دامنه تغییرات را دارد.
- میانه نمرات کلاس C از همه بیشتر است و میانه نمرات کلاس B از همه کمتر است.
- در کلاس B هم داده‌های درون جعبه و هم کل داده‌ها بیش از سایر کلاس‌ها پراکنده‌اند.
- کلاس B، کلاس A و کلاس C (از راست به چپ).

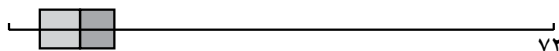
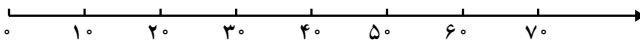
فعالیت ۴ صفحه ۱۲۳

- ابتدا داده‌های مربوط به زنان را که قبلاً مرتب شده است در نظر می‌گیریم. تعداد زنان ۵۰ نفر است.

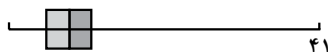
نفرات ۲۵ ام و ۲۶ ام هر دو ۸ روز بستری شده‌اند. پس میانه ۸ است. و میانه نیمه اول ۴ و میانه نیمه دوم ۱۴ است.

- داده‌های مربوط به مردان هم قبلاً مرتب شده است. تعداد مردان ۱۵۲ نفر است. نفرات ۱۷۶ ام و ۱۷۷ ام هر دو ۸ روز بستری شده‌اند. پس میانه ۸ است و میانه نیمه اول ۵ و میانه نیمه دوم ۱۱ است. تعداد کل بیماران ۲۰۲ نفر است.

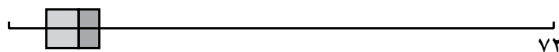
نفرات ۱۰۱ ام و ۱۰۲ ام هر دو ۹ روز بستری شده‌اند. پس میانه ۹ است و میانه نیمه اول ۵ و میانه نیمه دوم ۱۲ است.



نمودار جعبه‌ای مدت بستری شدن زنان



نمودار جعبه‌ای مدت بستری شدن مردان



نمودار جعبه‌ای مدت بستری شدن زنان و مردان

مدت بستری مردان و زنان تقریباً مساوی است ولی پراکندگی مدت بستری زنان بیشتر است.

فعالیت صفحه ۱۲۷

$$\text{میانگین} = \frac{\text{مجموع داده‌ها}}{\text{تعداد داده‌ها}} = \frac{۴۵۸۸}{۳۰} = ۱۵۲/۹۳ \approx ۱۵۳$$

– به مجموع ۵×۳۰ اضافه می‌شود پس به میانگین $۵ \times \frac{۳۰}{۳۰} = ۵$ اضافه می‌شود.

– از مجموع ۱۲×۳۰ کم می‌شود پس از میانگین $۱۲ \times \frac{۳۰}{۳۰} = ۱۲$ کم می‌شود.

– a به میانگین اضافه می‌شود.

– مجموع ۲ برابر می‌شود و لذا میانگین هم ۲ برابر می‌شود.

– مجموع $۰/۲$ برابر می‌شود و لذا میانگین هم $۰/۲$ برابر می‌شود.

– میانگین a برابر می‌شود.

– میانگین -a برابر می‌شود.

– الف – کفایت میانگین ۱۱۲، ۰، ۴۲، ۷، ۱۷، ۲۸ را به دست آوریم که برابر است با $۳۴/۳$

و با ۱۳۰۰ جمع کنیم.

– ب – کفایت میانگین ۸، ۱۰، ۳، ۴، ۱، ۲ را به دست آوریم که برابر است با $۴/۶۶۶۶$ و در

هزار ضرب کرده با ۳۸ جمع نماییم.

پاسخ تمرینات فصل

پاسخ تمرینات صفحه ۱۱۹

۱- ۱۵، ۱۲، ۶، ۵، ۵، ۵، ۴، ۰، ۰

–۱

مد: ۵

$$\text{میانگین: } ۵ = \frac{۵+۵}{۲}$$

۲- هم میانگین و هم مد دو برابر می‌شود.

۳- مؤلفه x نقطهٔ ماکزیم نمودار، نمایانگر مد می‌باشد.

۴- جامعه همگن نبوده و به دو زیرمجموعه قابل تفکیک است که در هر یک، مد قابل تشخیص

است.

پاسخ تمرینات صفحه ۱۲۴

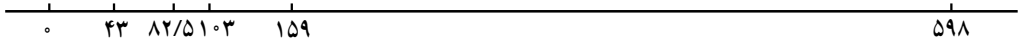
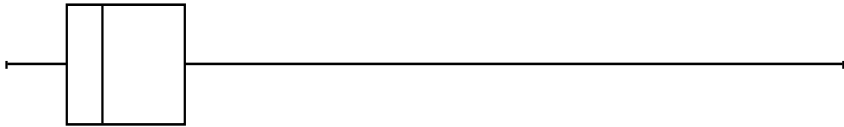
۱- نمودار ساقه و برگ برای توصیف این داده‌ها مناسب نیست زیرا پراکندگی داده‌ها خیلی زیاد است. به‌طور مثال از ۲۱۴ به ۲۴۳ و سپس به ۳۲۹، ۳۸۰، ۴۰۳ و ۵۱۱ رسیده است.

میانه : ۱۰۳

$$\text{چارک اول} : \frac{۸۲+۸۳}{۲} = ۸۲/۵$$

$$\text{چارک سوم} : \frac{۱۵۶+۱۶۲}{۲} = ۱۵۹$$

بیشترین مقدار : ۵۹۸
کم‌ترین مقدار : ۴۳



۲- حقوق تقریباً ۵۰ درصد کارمندان شرکت ۲ بیش‌تر از ماکزیمم حقوق کارمندان شرکت ۱ است صحیح است و حقوق ۵۰٪ کارمندان شرکت ۱ کمتر از ۱۰۰ هزار تومان دریافت می‌کنند. برخی از کارکنان شرکت ۲ نیز کمتر از ۱۰۰ هزار تومان دریافت می‌کنند پس ادعای دوم صحیح نیست. از چنین نمودارهایی برای مقایسه عملکردها در گروه‌های مختلف و نیز تشخیص پراکندگی داده‌ها در هر دسته می‌توان استفاده نمود.

پاسخ تمرینات صفحات ۱۴۰ - ۱۳۶

۱- الف - $۲۰۰۰+۳۹$ ، $۲۰۰۰+۷۳$ ، $۲۰۰۰+۱۶$ ، $۲۰۰۰+۲۰۲$ ،

$۲۰۰۰+۹۶$ ، $۲۰۰۰+۲۰$

$$\Rightarrow \bar{X} = ۲۰۰۰ + \frac{۳۹+۷۳+۱۶+۲۰۲+۹۶+۲۰}{۶}$$

$$= ۲۰۰۰ + ۷۵ \approx ۲۰۷۴ / ۳۳$$

۲۰۱۶ ، ۲۰۲۰ ، ۲۰۳۹ ، ۲۰۷۳ ، ۲۰۹۶ ، ۲۲۰۶

میانه : $\frac{۲۰۳۹+۲۰۷۳}{۲} = ۲۰۵۶$

ندارد : مد

$$\bar{X} = 20000 + \frac{39 + 73 + 16 + 22 + 96 + 20}{6} \approx 2044 / 33 \quad \text{ب -}$$

$$\text{میانۀ : } \frac{2022 + 2039}{2} = 2030 / 5$$

ندارد : مد

ج - میانگین

د - میانگین زیرا این تفاوت همیشه مستقیماً در محاسبات وارد می‌شود. برخی موارد تغییر یک داده، حتی میانۀ را نیز تغییر نمی‌دهد ولی میانگین با تغییر یک داده همواره تغییر می‌یابد.

$$\bar{X} \approx 13 / 73 \quad \text{۲ -}$$

مد : ۱۸

میانۀ : ۱۶

به دلیل وجود داده‌های پرت چون ۰، ۱ و ۳ که بر میانگین تأثیری زیادی می‌گذارند، میانگین شاخص خوبی نیست. در این مثال میانۀ نمایان‌گر بهتری از وضع کلاس می‌باشد.

۳ - خیر. به دلیل وجود اعدادی چون ۳۰۰,۰۰۰ و ۲۰۰,۰۰۰ میانگین بالا رفته است. پس میانگین یعنی ۱۰۰,۰۰۰ تومان شاخص مناسبی برای تعیین وضعیت نمی‌باشد. در این مثال :

۳۰۰۰۰۰ ، ۲۰۰۰۰۰ ، ۸۵۰۰۰ ، ۶۰۰۰۰ ، ۶۰۰۰۰ ، ۶۰۰۰۰ ، ۴۰۰۰۰ ، ۳۵۰۰۰

میانۀ : $\frac{60000 + 60000}{2}$ یعنی ۶۰۰۰۰ تومان است که معرف خوبی برای نشان دادن وضعیت

این شرکت است.

$$\bar{X} = \frac{3 + 8 + 5 + 9 + 3 + 7 + 8 + a + 3 + 4}{10} = \frac{50 + a}{10} = 5 / 5 \quad \text{۴ - الف}$$

$$\Rightarrow 50 + a = 55 \Rightarrow a = 5$$

ب - ۳، ۳، ۳، ۴، ۵، ۵، ۷، ۸، ۸، ۹

$$\text{میانۀ : } \frac{5 + 5}{2} = 5$$

۵ - a, a+1, a+2, a+3, a+4, a+5, a+6

میانۀ برابر است با a+3

$$a + \left(\frac{0 + 1 + 2 + 3 + 4 + 5 + 6}{7} \right) = a + 3 \quad \text{میانگین :}$$

$$(a + 3) - (a + 3) = 0 \quad \text{پس}$$

$$7, 17, 18, 18/5, 19, 19 \quad \text{۶- الف-}$$

$$\frac{7+17+18+18/5+19+19}{6} \approx 16/42 \quad \text{میانگین:}$$

$$\frac{18+18/5}{2} = 18/25 \quad \text{میانه:}$$

$$19 \quad \text{مد:}$$

ب- میانه

$$\frac{x+17+18+18/5+19+19}{6} = 18/5 \quad \text{ج-}$$

$$\Rightarrow x = 19/5$$

$$\text{①} \quad \text{②} \quad \text{③} \quad \text{④} \quad \text{⑤} \quad \text{۷-}$$

در ابتدا در نظر می‌گیریم که x بتواند یکی از وضعیتهای ①، ②، ③، ④ و ⑤ را اختیار کند. هریک از این وضعیتها را بررسی می‌کنیم. اگر x در یکی از وضعیتهای ① یا ⑤ قرار گیرد میانه به ترتیب برابر 7° یا 8° خواهد بود. با توجه به فرض مسئله که مد با میانه برابر است پس مد نیز باید 7° یا 8° باشد که با فرض اولیه ما تناقض خواهد داشت. پس x نمی‌تواند در وضعیتهای ① و ⑤ قرار داشته باشد.

اگر x در وضعیت ② قرار داشته باشد با فرض برابری مد و میانه، x باید برابر 7° باشد، که با محاسبه میانگین در این حالت، این فرض نیز رد می‌شود.

اگر x در وضعیت ③ باشد با فرض برابری مد و میانه، x می‌تواند یا 7° یا 8° باشد. $x = 7^\circ$ در بالا رد شد. اگر $x = 8^\circ$ باشد، میانگین نیز برابر 8° است که برای این مسئله قابل قبول است.

اگر x در وضعیت ④ باشد نیز برابری مد و میانه، $x = 8^\circ$ را حاصل می‌شود که قابل قبول است.

روش دیگر: چون تعداد داده‌ها فرد است، میانه برابر یکی از داده‌ها است.

با مرتب کردن داده‌ها به این نتیجه می‌رسیم که میانه یا 7° است یا 8° ، و با فرض برابری میانه و مد، x یا باید 7° باشد و یا 8° . با محاسبه میانگین در هر حالت به نتیجه $x = 8^\circ$ دست پیدا می‌کنیم.

۸- الف

تعداد محصول در هر بسته	فراوانی
۱	۵
۲	۵
۳	۳
۴	۶
۶	۳

$$\frac{۵ \times ۱ + ۵ \times ۲ + ۳ \times ۳ + ۶ \times ۴ + ۳ \times ۶}{۲۲} = ۳$$

میانگین:

$$\frac{۳ + ۳}{۲} = ۳$$

میانه:

۴

مد:

ب- مد و میانه با هم، مناسب‌ترین است. اگر میانه را نیز در نظر بگیریم می‌بینیم که ۵۰ درصد مردم بسته‌بندی‌های با تعداد سه تا بیشتر را می‌خواهند. در این میان بیشترین افراد خواهان بسته‌بندی‌های ۴ تایی هستند. پس دسته‌بندی ۴ تایی می‌تواند مطلوب کارخانه باشد.

۹-

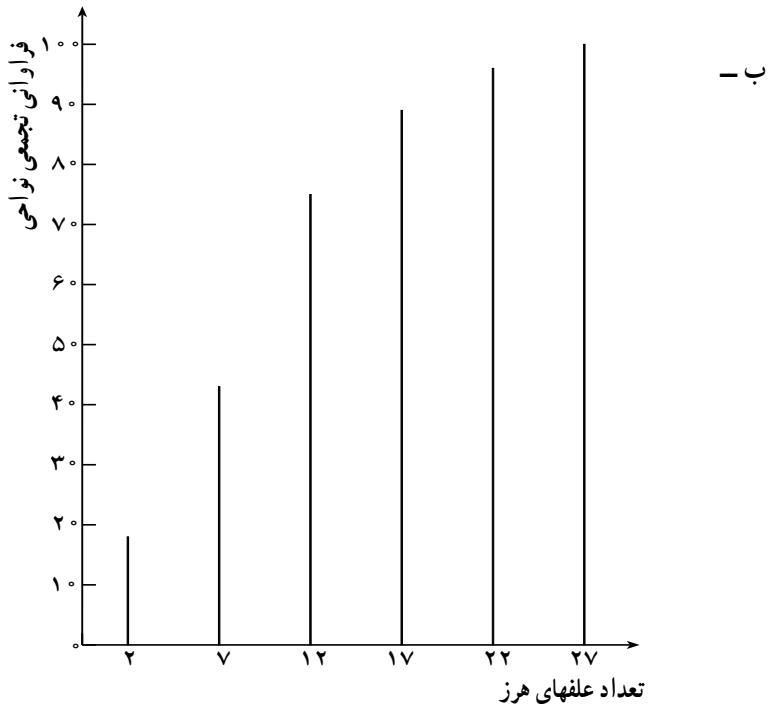
	نمره	تعداد واحد
ریاضیات	B	۳
فیزیک	A	۳
ورزش	C	۲
تاریخ	B	۲
ادبیات	C	۳
آزمایشگاه	A	۱

$$\bar{X}_w = \frac{۳ \times ۳ + ۳ \times ۴ + ۲ \times ۲ + ۲ \times ۳ + ۳ \times ۲ + ۱ \times ۴}{۱۴}$$

$$\approx ۲/۹۳$$

۱۰- الف

تعداد علف‌های هرز	فراوانی نواحی	فراوانی تجمعی	مرکز دسته
۰-۴	۱۸	۱۸	۲
۵-۹	۲۵	۴۳	۷
۱۰-۱۴	۳۲	۷۵	۱۲
۱۵-۱۹	۱۴	۸۹	۱۷
۲۰-۲۴	۷	۹۶	۲۲
۲۵-۲۹	۴	۱۰۰	۲۷



نمودار تعداد علفهای هرز در ۱۰۰ ناحیه مورد بررسی

$$\bar{X} = \frac{۲ \times ۱۸ + ۷ \times ۲۵ + ۱۲ \times ۳۲ + ۱۷ \times ۱۴ + ۲۲ \times ۷ + ۲۷ \times ۴}{۱۰۰} = ۱۰/۹۵ \text{ - ج}$$

$$۱۰/۹۵ \times (۸۰۰۰ \times ۱۰۰) = ۸۷۶۰۰۰۰ \text{ - د}$$

$$\text{میانگین} = \frac{۵۶۱}{۲۲} = ۲۵/۵ \text{ - ۱۱}$$

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

مد: ۲۵ و ۲۶

$$\text{میانه: } \frac{۲۵ + ۲۶}{۲} = ۲۵/۵$$

تعداد بازیکنانی که سن آنها بیشتر از میانگین است: ۱۱

تعداد بازیکنانی که سن آنها کمتر از میانگین است: ۱۱

تعداد بازیکنانی که سن آنها بیشتر از میانه است: ۱۱

تعداد بازیکنانی که سن آنها کمتر از میانه است: ۱۱

- نتایج به دست آمده نشان می دهد که جامعه ما تقریباً متقارن است.

$$\bar{X} = \frac{1 \times 12 + 2 \times 6 + 3 \times 4 + 4 \times 2 + 6 \times 1}{25} = 2 \quad -12$$

– مد ۱ سرنشین است.

۱۳– خیر. زیرا پراکندگی داده‌ها با توجه به تعداد داده‌ها زیاد است.

– خیر. زیرا پراکندگی زیاد است و داده‌های پرت چون ۵ و ۲۰۰ مشاهده می‌شوند.

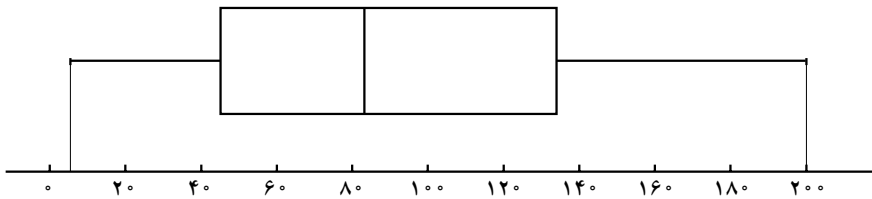
۵، ۱۰، ۱۲، ۱۵، ۱۵، ۱۷، ۱۹، ۲۵، ۲۵، ۲۵، ۳۰، ۳۵، ۴۵، ۴۷، ۵۵، ۶۶،
 ۶۶، ۶۸، ۶۸، ۶۸، ۷۵، ۷۵، ۷۵، ۷۷، ۷۸، ۸۵، ۸۸، ۸۸، ۹۳، ۹۵، ۹۸، ۹۸، ۱۰۱،
 ۱۱۰، ۱۱۵، ۱۲۰، ۱۲۰، ۱۳۵، ۱۳۵، ۱۴۰، ۱۵۰، ۱۶۰، ۱۶۵، ۱۷۵، ۱۷۷، ۱۸۰، ۱۸۰،
 ۱۹۰، ۱۹۵، ۲۰۰

۲۰۰ = بیشترین داده ۵ = کمترین داده

$$\text{میانۀ} = \frac{78 + 85}{2} = 81.5$$

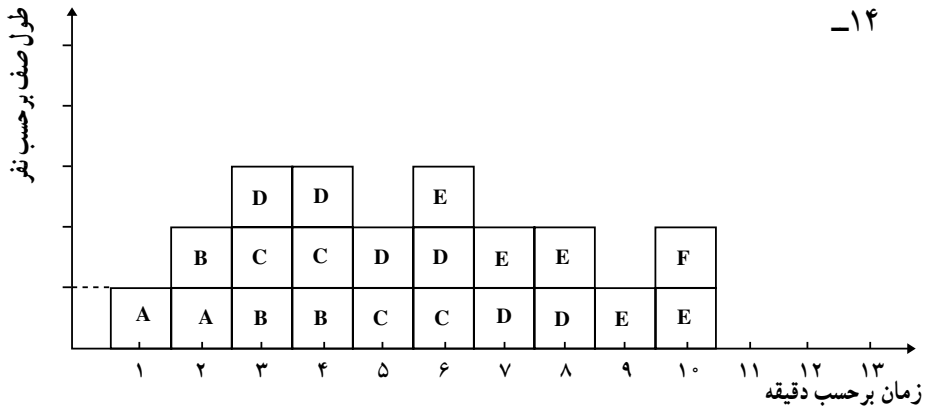
$$Q_1 = 45$$

$$Q_3 = 135$$



تراکم مصرف در فاصله ۴۵ تا ۸۰ کیلووات ساعت دیده می‌شود.

-14



اولین پیشنهاد این است که به جای فواصل زمانی ۱۰ دقیقه‌ای، عمل نمونه‌گیری را برای

فواصل زمانی پنج دقیقه‌ای انجام داده و الگو را رسم نماید. چرا که وجود دقایقی چون ۷، ۸، ۹ که مشتری جدید به صندوق مراجعه ننموده موجب می‌شود تا میانگین طول صف کاهش پیدا کند. پیشنهاد دوم این است که عمل نمونه‌گیری را روی زمان در فواصل زمانی پنج دقیقه‌ای در ساعات مختلف کار فروشگاه به‌طور تصادفی انجام داده تا متوجه شویم آیا الگوی به‌دست آمده برای تمام ساعت‌های کار مناسب است یا خیر. در صورت وجود الگوی مشترک، متناسب با الگوی به‌دست آمده تعداد صندوق‌های مشغول به کار را تعیین نموده در صورت عدم وجود الگوی مشترک تعداد صندوق‌ها متناسب با الگوی زمانی (به‌طور مثال صبح‌ها، بعدازظهرها، شب‌ها) در زمان‌های مختلف تعیین شود.