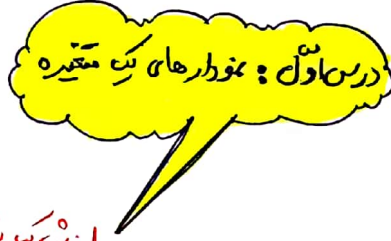


فصل پنجم: نمایش داده‌ها

- درس ۱ نمودارهای یک متغیره
- درس ۲ نمودارهای چند متغیره



ارزش یک تصویر به اندازه هزار کلمه است!

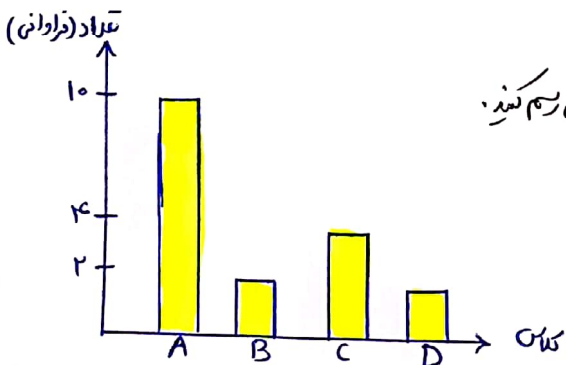
نمودار یک وسیله دیداری است که جامعه در آن قابل مشاهده است. هر نمودار با توجه به موضوعی که داده‌ها بر اساس آن جمع آوری شده‌اند و هم چنین نوع اطلاعات به دست آمده و کارایی دارد. نمایش داده‌ها روشی برای کمک به استخراج اطلاعات از داده‌ها با حس بینایی است. در سال‌های قبل با بعضی از نمودارها مثل نمودار میله‌ای، نمودار دایره‌ای و نمودار خط کُسته آشنا شده‌اید. در اینجا چنین یادآوریم نمودارهایی قبلی چون نمودار دایره‌ای، نمودار نقطه‌ای، جعبه‌ای و پراکنشی تعاریف، جابجایی و راداری را می‌آموزید. ابتدا نمودارهایی را بررسی می‌کنیم که همه آن‌ها فقط یک متغیر را بررسی می‌کنند.

۱. نمودار میله‌ای ۲. نمودار خط کُسته ۳. نمودار دایره‌ای ۴. نمودار نقطه‌ای ۵. نمودار جعبه‌ای

۱. نمودار میله‌ای

این نمودار از دو محور عمودی واقعی تشکیل شده است. روی محور افقی (x) نام متغیرها یا شماره آن‌ها را قرار می‌دهیم و روی محور عمودی (y) فراوانی آن‌ها را مشخص می‌کنیم.

- از این نمودار برای نمایش متغیرهای کیفی و متغیرهای کمی که مقدارش اعداد حسابی {۰، ۱، ۲، ۳، ۴، ۵، ...} هستند استفاده می‌شود.
- نمودار میله‌ای برای مقایسه فراوانی داده‌ها یا پیدا کردن داده‌ای که بیشترین و کمترین فراوانی را دارد، بسیار سودمند است.
- ترتیب قرار گرفتن میله‌ها در این نمودار اهمیت ندارد.



مثال. برای داده‌های جدول زیر، نمودار میله‌ای رسم کنید.

کلاس	A	B	C	D
تعداد	10	2	4	2

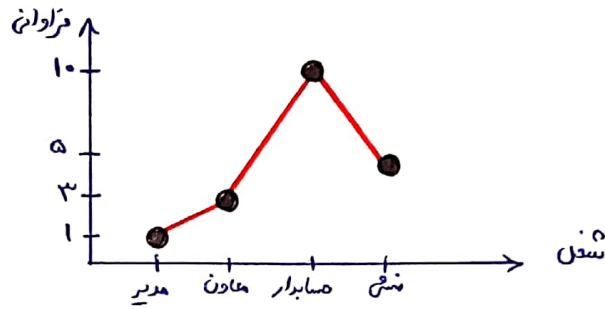
۲. نمودار خط سلسله

از این نمودار بیش تر برای نمایش متغیرها کنتی یا کنتی استفاده می کنند. در واقع نمودار از چندین نقطه تشکیل شده است که به طور متوالی با پاره خط هایی بهم وصل شده اند.

- نمودار خط سلسله برابر نمایش تغییرات مناسب است.

شغل	مدیر	معاون	حسابدار	منشی
فراوانی	۱	۳	۱۰	۵

مثال: نمودار خط سلسله جدول مقابل را رسم کنید.



۳. نمودار دایره ای

یکی از پرکاربردترین نمودارها در آمار، نمودار دایره ای است. از آن برای نمایش متغیرها کنتی استفاده می شود. در این روش دایره را با توجه به زاویه مرکزی و یا با توجه به درصد مساحت، به قسمت های کوچک تر تقسیم می کنند.

$$\theta = \frac{f}{n} \times 360^\circ$$

فراوانی داده (f) →
 زاویه مربوط به هر قسمت (θ) ←
 تعداد کل داده ها (n) →

- شعاع دایره هیچ اطلاعاتی به ما نمی دهد.

- زمانی که درصد را گزارش می دهیم، بهتر است از نمودار دایره ای استفاده کنیم تا بتوانیم مقایسه انجام دهیم.

مثال: گروه خوبی دانش آموزان یک کلاس به شرح زیر است. نمودار دایره ای آن را رسم کنید.

گروه خوبی	A	B	AB	o
فراوانی	۱۷	۱۴	۱	۲

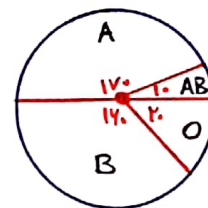
تعداد کل داده ها $n = 17 + 14 + 1 + 2 = 34$

$$\theta_A = \frac{17}{34} \times 360^\circ = 180^\circ$$

$$\theta_B = \frac{14}{34} \times 360^\circ = 147^\circ$$

$$\theta_{AB} = \frac{1}{34} \times 360^\circ = 10^\circ$$

$$\theta_o = \frac{2}{34} \times 360^\circ = 20^\circ$$

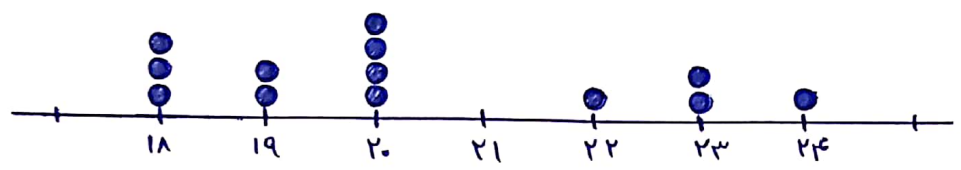


۴. نمودار نقطه ای

از این نمودار برای نمایش مقی‌ها، کتفی و یا متغیرهای کتفی با مقادیر حسابی استفاده می‌شود. برای این کار داده‌ها را که همان اعداد حسابی هستند، با توجه به کمترین و بیشترین داده موجود، روی محور افقی مرتب‌سیم و روی هر کدام از آن‌ها به تعداد فراوانی داده موجود، نقطه می‌نشانیم.

مثال. سن کنکورهای یک کتابخانه به صورت زیر است. نمودار نقطه‌ای مربوط به آن را رسم کنید.

۱۸، ۱۹، ۱۸، ۲۰، ۲۲، ۲۰، ۲۰، ۱۹، ۱۸، ۲۳، ۲۳، ۲۴، ۲۰.

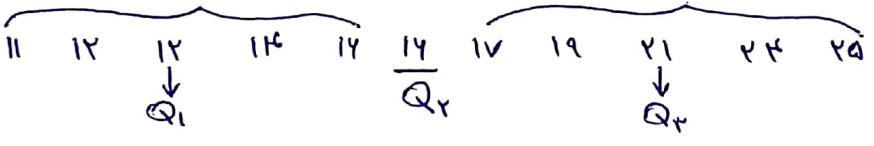


حد اقل سن ۱۸ و حد اکثر آن ۲۴ است. ملاحظه می‌کنید که با ۲۱ هیچ نقطه‌ای نیست یعنی هیچ فرد کنکوری از این کتابخانه ۲۱ سال ندارد.

مثال. در نمودار نقطه‌ای زیر، مقدار میانگین و دامنه میان چارگی را تعیین کنید.

		*				*									
	*	*		*		*	*		*		*		*	*	
۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲	۲۳	۲۴	۲۵

با توجه به هموار مرتب می‌شویم که داده‌ها به صورت زیر هستند:



$IQR = Q_3 - Q_1 = 21 - 12 = 9$ دامنه میان چارگی

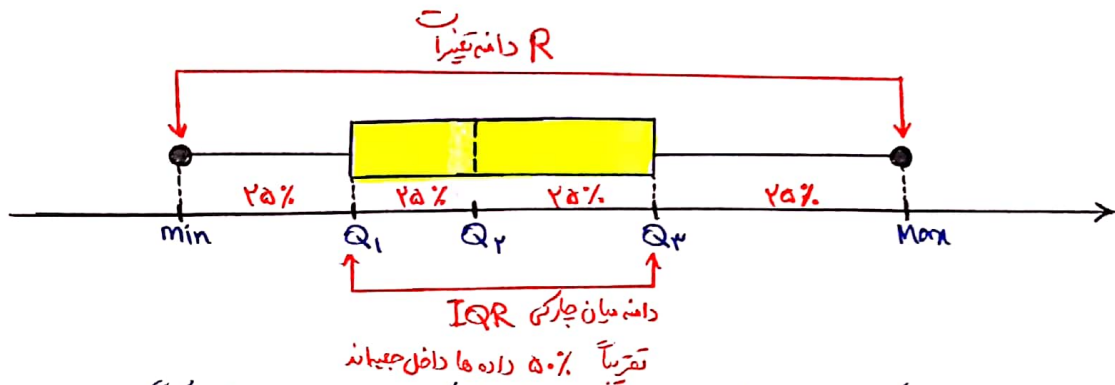
$Q_2 = 14$ میانگین

۵. نمودار جعبه‌ای

این نمودار روشی سودمند برای نمایش دامنه‌ها و چارک‌ها را داده‌هاست. این نمودار علاوه بر مقاسم داده‌ها با هم، مقدار بزرگترین داده‌ها

(دوره یا نزدیک داده‌ها نسبت به هم) را نیز نشان می‌دهد. برای رسم نمودار جعبه‌ای به ۵ مقدار زیر نیاز داریم:

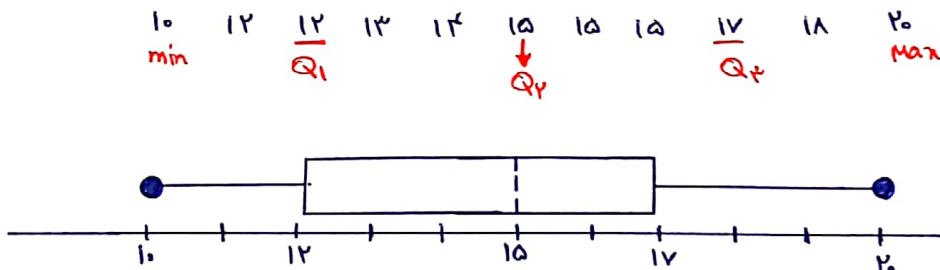
- ۱. کوچکترین داده min
- ۲. چارک اول Q_1
- ۳. میانگین یا چارک دوم Q_2
- ۴. چارک سوم Q_3
- ۵. بزرگترین داده Max



- این نمودار به دلیل اینکه به سرعت امکان بررسی هم زمان شاخص مرکزی (میان) و شاخص پراکنش (دامنه میان چارگی) را امکان پذیر می کند، می تواند وسیله خوبی برای مقایسه داده ها باشد.
- نمودار جعبه ای، کل داده ها را به ۴ قسمت مساوی از نظر تعداد تقسیم می کند که به هر کدام از آن ها چارک می گویند.
- میان (Q2) همیشه نزدیکاً در وسط جعبه قرار ندارد. اگر میان در سمت راست جعبه باشد نشان می دهد که تراکم داده ها بین Q2 و Q3 بیش از تراکم داده ها بین Q1 و Q2 است. یعنی در سمت راست جعبه داده ها به هم نزدیک ترند.

مثال • به داده های آماری ۱۰، ۱۲، ۱۲، ۱۳، ۱۴، ۱۵، ۱۵، ۱۴، ۱۳، ۱۲، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۱۶، ۱۷، ۱۸، ۱۹، ۲۰ درج کنید.

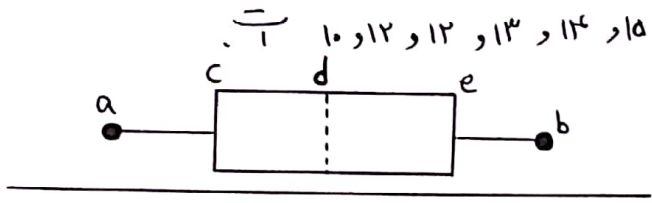
- نمودار جعبه ای مربوط به این داده ها را رسم کنید.
- دامنه میان چارگی را به دست آورید.



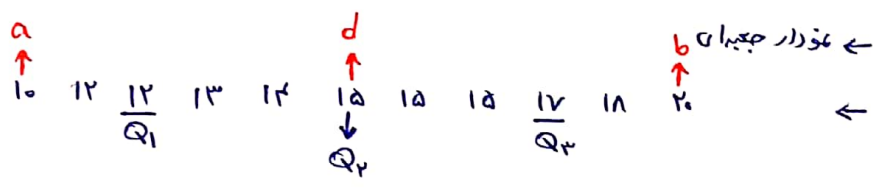
$$IQR = Q_3 - Q_1 = 17 - 12 = 5$$

۴ و ۱۰ و ۵ و ۲۵ و ۳۰ و ۱۷ و ۱۹ و ۲۸ و ۳۹ و ۳ و ۴۰

مثال • برای داده های مقابل نمودار جعبه ای رسم کنید.
با عنوان کردن



مثال. سن زیر مربوط به داده‌های ۱۰، ۱۲، ۱۲، ۱۳، ۱۴، ۱۵، ۱۵، ۱۷، ۱۸، ۲۰. این نمودار چه نام دارد؟
 - حاصل عبارت $\frac{a+b}{d}$ را بدست آورید.
 - مقدار IQR را بیابید.
 - چند درصد داده‌ها بیش از میانگین اول هستند؟

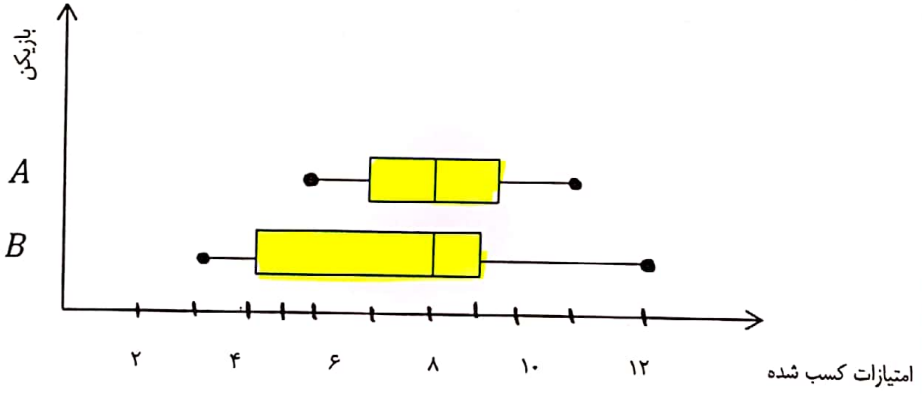


$$\frac{a+b}{d} = \frac{10+20}{15} = \frac{30}{15} = 2$$

$$IQR = Q_3 - Q_1 = 17 - 12 = 5$$

تقریباً ۷۵٪ داده‌ها از بزرگ‌تر اول هستند.

مثال. اگر یک مربی ورزشی بخواهد از دو بازیکن A و B یکی را انتخاب کند. با توجه به نمودار زیر که نشان دهنده امتیازها و کسب شده توسط این دو نفر است، مربی کدام بازیکن را ترجیح می‌دهد؟ چرا؟



با توجه به نمودار داریم:

A: min = 4 $Q_1 = 7$ $Q_2 = 8$ $Q_3 = 10$ Max = 11

B: min = 3 $Q_1 = 4$ $Q_2 = 8$ $Q_3 = 9$ Max = 12

$$\left. \begin{array}{l} R = \text{Max} - \text{min} = 11 - 4 = 7 \\ IQR = Q_3 - Q_1 = 10 - 7 = 3 \end{array} \right\} \text{بازیکن A}$$

$$\left. \begin{array}{l} R = \text{Max} - \text{min} = 12 - 3 = 9 \\ IQR = Q_3 - Q_1 = 9 - 4 = 5 \end{array} \right\} \text{بازیکن B}$$

چون R و IQR برای بازیکن A کم است پس بازیکن A کم تر بوده و مربی این بازیکن را ترجیح می‌دهد.

