

آمار هنر پرورش و برداشش داده است. نمایش داده‌ها می‌تواند شبیه یک کوه و یا شبیه شاخه‌های درختان باشد. می‌توان برای ساختن یک پل از معیارهای گرایش به مرکز و یا معیارهای پراکندگی استفاده نمود.



## آمار توصیفی

۳

- ۱ توصیف و نمایش داده‌ها
- ۲ معیارهای گرایش به مرکز
- ۳ معیارهای پراکندگی





بعد از گردآوری داده‌ها، به تنظیم، رده‌بندی و خلاصه کردن آنها می‌پردازیم. به این منظور می‌توان از روش‌های زیر استفاده نمود:

الف) تنظیم و طبقه‌بندی داده‌ها در یک جدول به نام جدول فراوانی  
 ب) رسم کردن نمودارهای مختلف براساس مقادیر جدول فراوانی

## فعالیت



یک راننده تاکسی در یک روز، اسکناس‌های زیر را از مسافران دریافت می‌کند. او تصمیم دارد این اسکناس‌ها را در کیف خود دسته‌بندی کند. برای انجام این دسته‌بندی، می‌خواهد مراحل زیر را انجام دهد. شما او را کمک کنید تا این کار را انجام دهد.

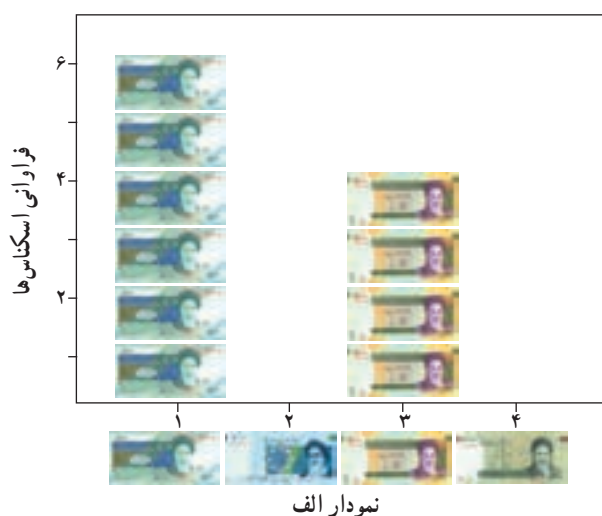


- ۱ ابتدا به هر نوع اسکناس عدد ۱ تا ۴ را بدهید و در ستون شماره وارد کنید.
- ۲ سپس به شمارش اسکناس‌ها بپردازید و تعداد تکرار هر اسکناس را در ستون سوم وارد کنید.
- ۳ در ادامه تعداد هر اسکناس را بر تعداد کل اسکناس‌ها تقسیم کنید و آن را در ستون چهارم قرار دهید.



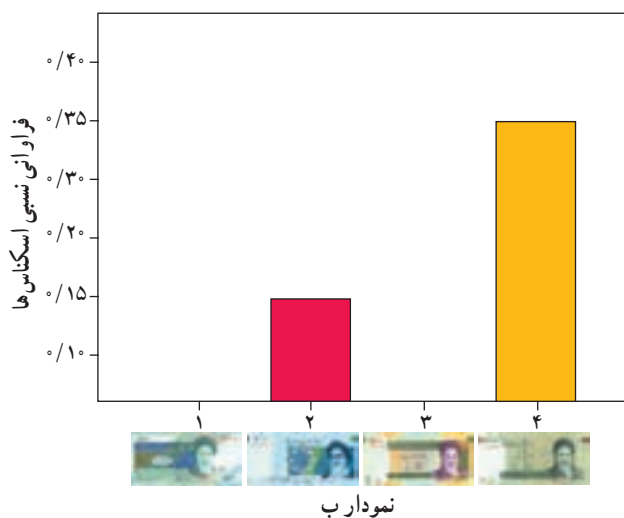
**۴** با توجه به اعداد موجود در جدول زیر، چند درصد اسکناس‌ها ۱۰ هزار تومانی، چند درصد ۱۰۰۰ تومانی، چند درصد ۵ هزار تومانی و چند درصد از اسکناس‌ها ۱۰ هزار تومانی است؟

انواع اسکناس‌ها	شماره	فراوانی یا تعداد تکرار هر اسکناس	فراوانی یا تعداد تکرار هر اسکناس تعداد کل اسکناس‌ها
	۱	۶	$\frac{۶}{۲۰}$
	۲		۰/۱۵
	۳	۴	
	۴		۰/۳۵
تعداد کل اسکناس‌ها		۲۰	



حال می‌خواهیم جدول بالا را به صورت سه نمودار الف، ب و پ نشان دهیم.

**۱** در نمودار الف، ابتدا دو محور مختصات رسم می‌کنیم که محور عمودی نشان‌دهندهٔ تعداد تکرار اسکناس‌ها، یا فراوانی اسکناس‌ها و محور افقی نشان‌دهندهٔ نوع اسکناس‌ها باشد. در این نمودار، اسکناس‌های ۱۰۰۰ و ۵ هزار تومانی روی هم قرار گرفته‌اند. شما هم اسکناس‌های ۲ هزار تومانی را به صورت  و اسکناس‌های ۱۰ هزار تومانی را به صورت  در نمودار قرار داده و آن را کامل کنید.

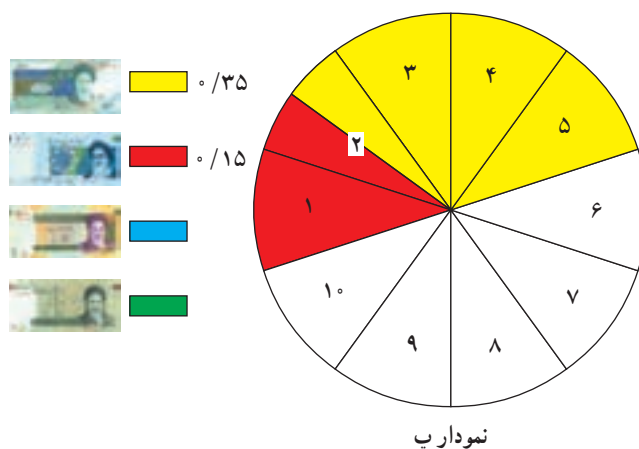


**۲** در نمودار ب، نیز دو محور مختصات رسم می‌کنیم که محور عمودی نشان‌دهندهٔ نسبت تکرار هر اسکناس به تعداد کل اسکناس‌ها یا فراوانی نسبی اسکناس‌ها و محور افقی نشان‌دهندهٔ نوع اسکناس‌ها است. با رسم مستطیل‌هایی برای فراوانی نسبی اسکناس‌های ۱۰۰۰ و ۵ هزار تومانی نمودار شکل ب را کامل کنید.



۲ اگر راننده تاکسی بخواهد وضعیت تعداد اسکناس‌های خود را در یک هفته پیش‌بینی کند، کدام نمودار الف یا ب می‌تواند به او کمک کند؟

۴ برای رسم نمودار دایره‌ای ابتدا دایره را به ۱۰ قسمت مساوی تقسیم می‌کنیم که هر قسمت نشان دهنده ۱۰ درصد کل دایره است. سپس با استفاده از عدد مربوط به نسبت تکرار هر اسکناس به تعداد کل اسکناس‌ها یا فراوانی نسبی مربوط به اسکناس ۱۰ هزار تومانی در ستون چهارم، قسمت اول دایره و نصف قسمت دوم دایره رنگ قرمز شده است که معادل ۱۵ درصد کل دایره است و به طور مشابه برای اسکناس ۲ هزار تومانی سه قسمت دایره به علاوه نصف قسمت دوم رنگ زرد می‌شود که معادل ۳۵ درصد کل دایره است. برای اسکناس‌های ۱۰۰۰ و ۵ هزار تومانی دایره را رنگ آبی و سبز کنید.



**داده‌ها:** واقعیت‌هایی درباره یک شیء یا فردند که در محاسبه، برنامه‌ریزی و پیش‌بینی به کار می‌روند.  
**متغیر:** هر ویژگی از اشیا یا اشخاص، که در اعضای جامعه یکسان نیست و معمولاً از یک عضو به عضو دیگر تغییر می‌کند را **متغیر** می‌گویند و عددی که به آن ویژگی یک عضو نسبت داده می‌شود را **مقدار متغیر**، یا **مشاهده** می‌گویند.  
**فراوانی یک داده:** تعداد دفعاتی که هر داده مشاهده می‌شود را **فراوانی** آن داده می‌گویند.  
**فراوانی نسبی یک داده:** با تقسیم فراوانی هر داده به تعداد کل داده‌ها، **فراوانی نسبی** آن داده به دست می‌آید.  
 اگر فراوانی نسبی داده‌ها در ۱۰۰ ضرب شود، آن‌گاه درصد داده‌ها به دست می‌آید.

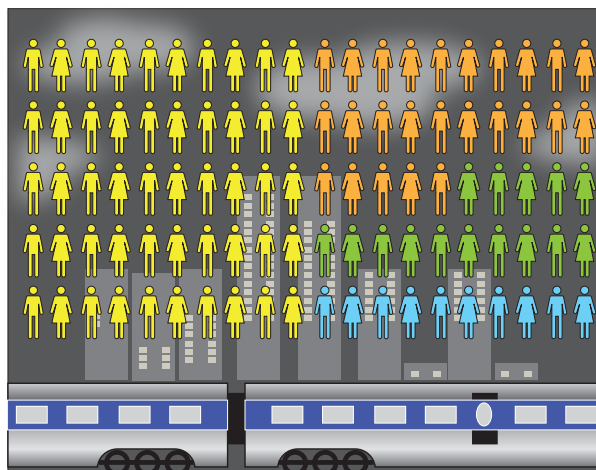
## کار در کلاس

در مورد اینکه مسافران یک قطار در طول سفر چگونه از وقت خود استفاده می‌کنند، تحقیقی صورت گرفته است و نتایج زیر به دست آمده است.

- در شکل ت، تعداد مسافران یک قطار به عنوان متغیر گسسته را ملاحظه می‌کنید.
- افرادی که با رنگ زرد مشخص شده‌اند، مسافرانی‌اند که در قطار استراحت می‌کنند.
- افرادی که با رنگ نارنجی مشخص شده‌اند، مسافرانی‌اند که در قطار با تلفن همراه خود بازی می‌کنند.
- افرادی که با رنگ سبز مشخص شده‌اند، مسافرانی‌اند که در قطار مطالعه می‌کنند.



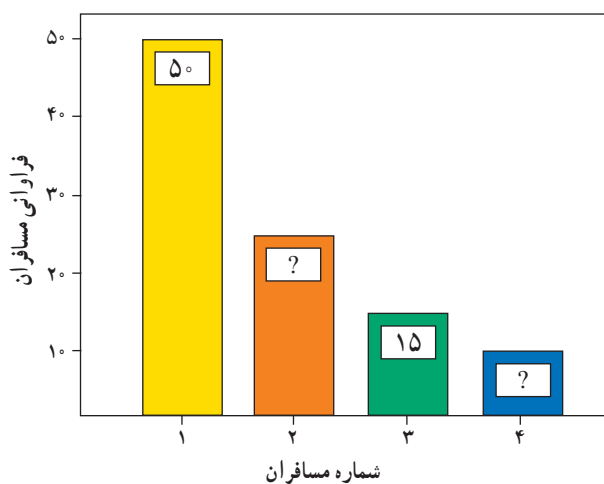
■ افرادی که بارنگ آبی مشخص شده‌اند، مسافرانی‌اند که در قطار غذا می‌خورند.



شکلت

جدول فراوانی مربوط به فراوانی تعداد مسافران را کامل کنید.

مسافران قطار	شماره مسافران	فراوانی مسافران	فراوانی نسبی مسافران
مسافرانی که استراحت می‌کنند	۱	۵۰	
مسافرانی که با تلفن همراه خود بازی می‌کنند	۲		۰/۲۵
مسافرانی که مطالعه می‌کنند	۳	۱۵	
مسافرانی که غذا می‌خورند	۴		۰/۱۰
تعداد کل مسافران		۱۰۰	



همچنین نمودار میله‌ای مربوط به فراوانی تعداد مسافران را کامل کنید.

فراوانی نسبی تعداد مسافران را براساس جدول کامل شده رسم کنید. نمودار دایره‌ای مربوط به فراوانی نسبی تعداد مسافران را رسم کنید.



سال هاست با مسئله آلودگی هوا آشنا هستیم و این مسئله به یکی از دغدغه‌های مهم تبدیل شده است.



شاخص کیفیت هوا<sup>۱</sup> (AQI)، متغیری پیوسته برای بیان کیفیت روزانه هواست. شاخص کیفیت هوا، برای شش آلاینده اصلی هوا شامل کربن مونواکسید، اوزون، گوگرد دی‌اکسید، نیتروژن دی‌اکسید و میزان ذرات معلق در هوا سنجیده می‌شود.

آلاینده	تأثیر بهداشتی	منبع انتشار
	کاهش عملکرد ریه و افزایش علائم تنفسی مانند سرفه، تنگی نفس، تشدید آسم و سایر بیماری‌های ریوی، افزایش استفاده از دارو، مراجعات و پذیرش بیمارستانی، اورژانس و مرگ و میر زودرس	این آلاینده ثانویه در اثر واکنش شیمیایی ترکیبات آلی فزار و اکسیدهای نیتروژن در حضور نور خورشید تولید می‌شود.
	مواجهه کوتاه مدت با این آلاینده می‌تواند منجر به تشدید علائم بیماری‌های قلبی ریوی و علائم تنفسی، افزایش نیاز به استفاده از دارو و پذیرش بیمارستانی گردد. مواجهه طولانی مدت عامل مرگ و میر زودرس و تشدید بیماری‌های قلبی و ریوی است.	ذرات معلق در اثر انتشار مستقیم و یا واکنش‌های شیمیایی ایجاد می‌شوند. عمده‌ترین منابع انتشار این آلاینده شامل احتراق سوخت (مانند سوزاندن زغال سنگ، چوب و سوخت دیزل)، فرایندهای صنعتی، کشاورزی و انتشار از جاده، خودروها (اگزوز، لنت، لاستیک و...) می‌باشند.
	تشدید بیماری‌های ریوی، افزایش مراجعات و پذیرش بیمارستانی، اورژانس و افزایش آسیب‌پذیری و استعداد ابتلا به عفونت‌های ریوی	احتراق سوخت (از وسایل نقلیه، واحدهای تولید برق، صنایع، بویلرها و همچنین سوزاندن چوب)
	کاهش اکسیژن‌رسانی به بافت‌ها و اندام‌های مختلف بدن، تشدید بیماری‌های قلبی و درد قفسه سینه، افزایش مراجعات و پذیرش بیمارستانی	احتراق سوخت (به خصوص در وسایل نقلیه موتوری)
	تشدید آسم و افزایش علائم تنفسی، کمک به شکل‌گیری و تشدید علائم و اثرات بیماری‌های ریوی	احتراق سوخت (به‌ویژه سوخت‌های با گوگرد بالا)، فرایندهای تولید برق و صنایع، منابع طبیعی مانند آتشفشان

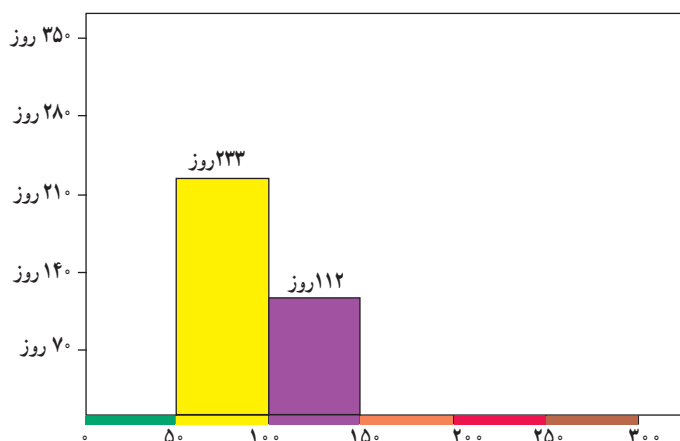
۱- Air Quality Index



اطلاعات تکمیلی و داده‌های مربوط به شاخص آلودگی هوا در سایت شرکت کنترل کیفیت هوا<sup>۱</sup> قابل دسترسی است. میزان شاخص کیفیت هوا در شهر تهران برای تمام روزهای سال ۱۳۹۳ در جدول زیر گزارش شده است. این جدول را کامل کنید:

وضعیت هوا	شاخص کیفیت هوا	فراوانی	فراوانی نسبی
پاک	$0 \leq AQI \leq 50$	۱۶	
سالم	$50 < AQI \leq 100$	۲۳۳	
ناسالم برای گروه‌های حساس	$100 < AQI \leq 150$	۱۱۲	
ناسالم	$150 < AQI \leq 200$	۴	
بسیار ناسالم	$200 < AQI \leq 250$	۱۰	
خطرناک	$250 < AQI \leq 300$	۰	
تعداد کل روزهای یک سال		۳۶۵	

نمودار مربوط به فراوانی تعداد روزها براساس وضعیت آلودگی هوا را کامل کنید؟



- نمودار فراوانی نسبی تعداد روزها را براساس وضعیت آلودگی هوا رسم کنید.
- چند درصد از روزهای سال، هوا سالم بوده است؟
- چند درصد روزهای سال، هوا ناسالم و بسیار ناسالم بوده است؟
- کدام نمودار، در پاسخ دادن به سؤالات، ما را بهتر راهنمایی می‌کند؟

<sup>۱</sup> <http://air.tehran.ir>

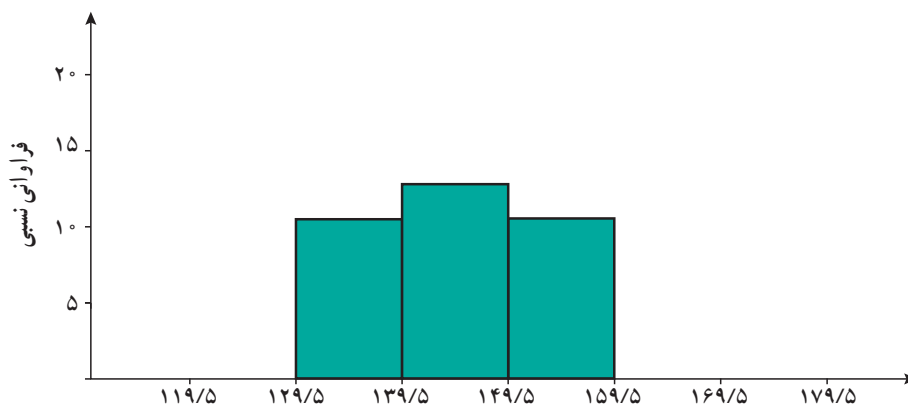




جدول فراوانی زیر مربوط به قد  $50$  (H) دانش‌آموز پایه یازدهم است. جاهای خالی جدول زیر را کامل کنید.

دسته‌ها	قد دانش‌آموزان	فراوانی	فراوانی نسبی	کرانه‌ها
۱۲۰-۱۲۹	$120 \leq H < 130$	۳		[۱۱۹/۵-۱۲۹/۵]
۱۳۰-۱۳۹	$130 \leq H < 140$	۱۰		[۱۲۹/۵-۱۳۹/۵]
۱۴۰-۱۴۹	$140 \leq H < 150$	۱۳		[۱۳۹/۵-۱۴۹/۵]
۱۵۰-۱۵۹	$150 \leq H < 160$	۱۱		[۱۴۹/۵-۱۵۹/۵]
۱۶۰-۱۶۹	$160 \leq H < 170$	۱۰		[۱۵۹/۵-۱۶۹/۵]
۱۷۰-۱۷۹	$170 \leq H < 180$	۳		[۱۶۹/۵-۱۷۹/۵]
	مجموع	۵۰		

بر اساس اعداد جدول، نمودارهای بافت نگاشت<sup>۱</sup> مربوط به فراوانی نسبی قد دانش‌آموزان را کامل کنید.



قد چند درصد از دانش‌آموزان بین  $160$  تا  $170$  سانتی‌متر است؟ همچنین قد چند درصد از دانش‌آموزان بین  $120$  تا  $140$  سانتی‌متر است؟



۱ داده‌های زیر، مسافتی را که ۲۰ راننده از مکان‌های مختلف برای رسیدن به مقصد A طی می‌کنند نشان می‌دهد. این داده‌ها، در جدول زیر گردآوری شده است. جدول را کامل کرده و نمودار بافت نگاشت مربوطه را رسم کنید.

دسته‌ها	کیلومترهایی که توسط راننده طی شده است	فراوانی	فراوانی نسبی
۶-۱۰	از ۵/۵ کیلومتر تا ۱۰/۵ کیلومتر	۱	
۱۱-۱۵	از ۱۰/۵ کیلومتر تا ۱۵/۵ کیلومتر	۲	
۱۶-۲۰	از ۱۵/۵ کیلومتر تا ۲۰/۵ کیلومتر	۳	
۲۱-۲۵	از ۲۰/۵ کیلومتر تا ۲۵/۵ کیلومتر	۵	
۲۶-۳۰	از ۲۵/۵ کیلومتر تا ۳۰/۵ کیلومتر	۴	
۳۱-۳۵	از ۳۰/۵ کیلومتر تا ۳۵/۵ کیلومتر	۳	
۳۶-۴۰	از ۳۵/۵ کیلومتر تا ۴۰/۵ کیلومتر	۲	
	مجموع	۲۰	

۲ رنگ چشم ۱۲۸ فرد به شرح زیر است: ۶۴ نفر قهوه‌ای، ۲۳ نفر آبی، ۳۶ نفر سبز و ۵ نفر سایر رنگ‌هاست. چه نمودارهایی می‌توان برای این اعداد رسم کرد. آن نمودار را رسم کنید؟

نمودار میله‌ای ☐ نمودار دایره‌ای ☐ هر دو ☐

۳ جملات زیر را کامل کنید:

الف) برای متغیرهای پیوسته از نمودار..... استفاده می‌شود.

ب) برای متغیرهای گسسته از نمودارهای ..... و ..... استفاده می‌شود.

پ) برای متغیرهای کیفی از نمودارهای ..... و ..... استفاده می‌شود.

۴ گروه خونی ۵۰ دانش‌آموز پایه یازدهم به صورت زیر گردآوری شده‌اند:

الف) جدول فراوانی مربوط به گروه خونی این افراد را رسم کنید. ب) نمودار میله‌ای مربوط به فراوانی و فراوانی نسبی و همچنین نمودار دایره‌ای مربوط به این افراد را رسم کنید؟ پ) چند درصد افراد، دارای گروه خونی O هستند؟



O	O	A	A	O
B	O	B	A	O
AB	B	A	B	AB
O	O	A	A	O
AB	O	A	B	A
O	A	A	O	A
O	A	O	AB	A
O	B	A	A	O
O	O	O	A	O
O	A	O	A	O



۵ اگر فراوانی نسبی مربوط به گروه خونی O،  $\frac{4}{100}$  باشد و مجموع فراوانی‌های همه گروه‌های خونی برابر  $20^\circ$  در نظر گرفته شود. فراوانی گروه خونی O چه عددی است؟

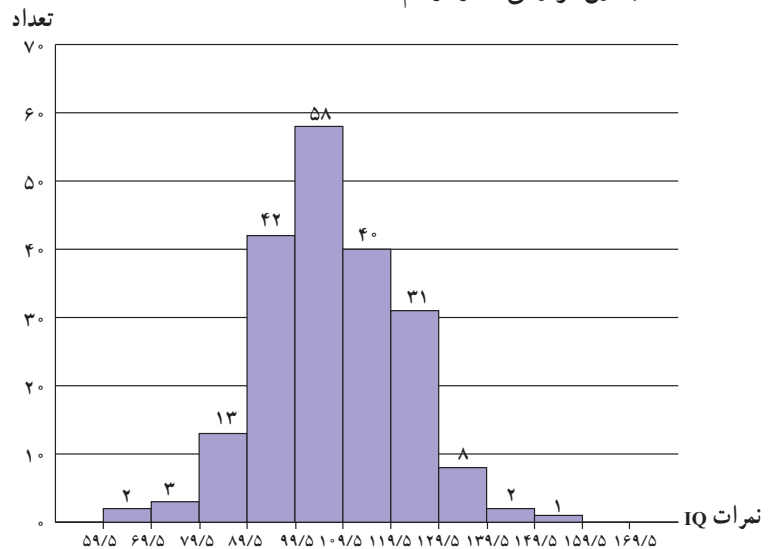
۶ نمودار بافت نگاشت نمرات IQ کودکان یک مهدکودک به صورت زیر رسم شده است. با توجه به این نمودار، به سؤالات زیر پاسخ دهید؟

(الف) تعداد کل کودکان که نمره IQ آنها، مورد بررسی قرار گرفته است، چند نفر است؟

(ب) نمره IQ در کدام رده بیشترین و در کدام رده کمترین فراوانی را دارد؟

(پ) چند درصد کودکان دارای نمره IQ بین  $140^\circ$  تا  $160^\circ$  هستند؟

(ت) جدول فراوانی آن را رسم کنید؟



۷ جدول فراوانی و نمودارهای مناسب مربوط به تعداد حروف بیت شعر زیر را به دست آورید؟

کیست این پنهان مرا در جان و تن      کز زبان من همی گوید سخن



## تاریخچه علم آمار و علم احتمال

علم آمار تاریخچه‌ای بسیار طولانی دارد؛ منشأ ظهور آمار به صورت توصیف اطلاعات را می‌توان سرشماری‌هایی که حدود ۴۰۰۰ سال قبل از میلاد مسیح توسط بابلی‌ها و مصری‌ها و بعداً توسط امپراتوری‌های روم و ایران درباره اطلاعات مربوط به زاد و ولد و دارائی‌های افراد جامعه زیر سلطه خود انجام می‌گرفته، به حساب آورد.

در قرن ۱۴ میلادی برای محاسبه نرخ بیمه، جمع‌آوری اطلاعات درباره تولد و وفات و حوادث رایج شد. در اواسط قرن ۱۷ مطالعات آماری به صورت توصیفی انجام می‌گرفت. مثلاً گرونت با مطالعه تعداد متولدین کشف نمود که تعداد پسرها کمی از تعداد دخترها بیشتر است، اما سال‌های اول زندگی تعداد بیشتری از پسرها فوت می‌کنند. استفاده از علم احتمال در آمار، در اواخر قرن ۱۷ شروع شد. در این مورد می‌توان به مطالعات مندل در مورد قانون وراثت اشاره کرد.

دامنه علم آمار در اوایل قرن ۱۹ شامل جمع‌آوری و تحلیل داده‌ها می‌شد. لژاندر در ۱۸۰۵ «روش کمترین مربعات» را برای اولین بار شرح داد.

آمار مدرن در اواخر قرن ۱۹ و اوایل قرن ۲۰ پدید آمده است. گالتون و کارل پیرسون آمار را وارد چارچوب دقیق ریاضیات کردند و فیشر در ابداع روش‌های مختلف «استنباط آماری» از جمله «آزمون فرض» قدم‌های مهمی برداشت.

علم احتمال عمری کوتاه‌تر از علم آمار دارد :

در اواسط قرن ۱۶ اولین کتاب احتمال توسط کاردانو با عنوان «بازی‌های شانس» نوشته شد.

در اواسط قرن ۱۷ پاسکال و فرما اولین کسانی بودند که مطالعه احتمال را به طور علمی شروع نمودند. البته آنها کتابی در این مورد ننوشتند بلکه در مکاتبات خود به آنالیز ترکیبی و مسائل مربوط به علم احتمال پرداختند. هویگنس کتابی در مورد احتمال نوشت که از نظر تحلیل علمی در سطح بسیار بالاتری از کتاب کاردانو قرار داشت.

یاکوب برنولی و دموآور در قرن ۱۸ کار را ادامه دادند. در قرن ۱۸ و ابتدای قرن ۱۹ علم احتمال در دانش‌های طبیعی و صنعت به طور جدی کاربرد پیدا کرد. در این دوره نخستین قضیه‌های علم احتمال یعنی قضایای لاپلاس، پواسون، لژاندر و گاوس ثابت شد.

در نیمه دوم قرن ۱۹ دانشمندان روسی تأثیر زیادی در پیشرفت علم احتمال داشتند؛ چییشف و شاگردانش، مسئله‌های لیاپونوف و مارکوف، از مسئله‌های کلی علم احتمال را حل کردند و قضایای برنولی و لاپلاس را تعمیم دادند.

در آغاز قرن ۲۰ متخصصان کارهای قبلی را منظم نموده و ساختمان اصول موضوعه احتمال را بنا نمودند. در این دوره دانشمندان زیادی روی علم احتمال کار کردند اما درخشان‌ترین نام در این عرصه کولموگروف روسی است که اصول موضوع احتمال را در کتابی به نام مبانی علم احتمال در آلمان منتشر کرد.

با توجه به اینکه شکل‌گیری علم احتمال در اروپای قرن شانزدهم برای بررسی بازی‌های شانس بوده است، بسیاری از مسائل احتمال هنوز هم به زبان بازی و شرط‌بندی بیان می‌شود، در حالی که امروزه علم احتمال در بسیاری از مسائل مهندسی، پزشکی، اقتصادی، سیاسی، علوم انسانی و... مورد استفاده قرار می‌گیرد.



### الف) میانگین داده‌ها

#### فعالیت



در یک باغ، برای تعیین میزان محصولات گردو، چهار نوع درخت گردو وجود دارد که میزان محصولات انواع گردوها برحسب تعداد به شرح زیر است :

نوع گردو	گردوی نوع اول	گردوی نوع دوم	گردوی نوع سوم	گردوی نوع چهارم
میزان محصول گردو (تعداد)	۵۰۰۰	۲۵۰۰	۳۵۰۰۰	۱۰۰۰

الف) میانگین تعداد گردوی تولید شده برای این چهار نوع درخت چه تعداد است؟  
حال اگر علاوه بر داشتن اطلاعات میزان تولید گردو برای هر نوع درخت گردو، تعداد درخت‌های باغ مطابق جدول زیر مشخص شده باشند :

نوع	گردوی نوع اول	گردوی نوع دوم	گردوی نوع سوم	گردوی نوع چهارم
میزان محصول گردو برای هر درخت (تعداد)	۵۰۰۰	۲۵۰۰	۳۵۰۰۰	۱۰۰۰
تعداد درخت‌ها	۱۰	۵	۷	۳



ب) آیا می‌توان میانگین تعداد گردوی تولید شده در قسمت (الف) را در این حالت به عنوان میانگین گردوی تولید شده برای این چهار نوع درخت گردو در نظر گرفت؟  
پ) میانگین گردوی تولید شده در این حالت، به چه صورت است؟

**مجموع داده‌ها:** اگر  $n$  داده  $x_1, x_2, \dots, x_n$  داشته باشیم، مجموع آن داده‌ها را با نماد سیگما ( $\Sigma$ ) نمایش

می‌دهیم و داریم:

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$$

و عبارت  $\sum_{i=1}^n x_i$ ، سیگمای  $i$  از ۱ تا  $n$ ،  $x_i$  می‌خوانیم.

**میانگین یا متوسط داده‌ها:** میانگین یا متوسط داده‌ها را با نماد  $\bar{x}$  نشان می‌دهیم و آن را به صورت زیر تعریف

می‌کنیم:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

**میانگین موزون داده‌ها:** اگر  $n$ ، داده  $x_1, x_2, \dots, x_n$  داشته باشیم به طوری که هر یک از این داده‌ها دارای

تعداد تکرار  $w_1, w_2, \dots, w_n$  هستند که به هر یک از آنها وزن داده متناظر با آن می‌گوییم. میانگین موزون داده‌ها را

با نماد  $\bar{x}_w$  نشان می‌دهیم و آن را به صورت زیر تعریف می‌کنیم:

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n}$$

## کار در کلاس

دانش‌آموزی در کنکور سراسری شرکت می‌کند و نتیجه کارنامه‌آزمون آن به

شرح زیر است:



مواد امتحانی	ریاضیات	فیزیک	شیمی	زبان انگلیسی	ادبیات و زبان فارسی	دین و زندگی
درصد	۷۱	۶۵	۸۰	۵۲	۹۵	۱۰۰
ضریب درس	۴	۳	۱	۱	۴	۳



الف) متوسط درصد مواد امتحانی این دانش آموز بدون احتساب ضرایب مواد امتحانی چه عددی است؟  
 ب) متوسط درصد مواد امتحانی این دانش آموز با احتساب ضرایب مواد امتحانی را کامل کنید؟

$$\bar{x}_w = \frac{\sum_{i=1}^6 w_i x_i}{\sum_{i=1}^6 w_i} = \frac{4 \times 71 + \dots + \dots + \dots + \dots + 3 \times 10}{4 + \dots + \dots + \dots + \dots + 3}$$

پ) کدام متوسط، مناسب است؟

## ب) میانه داده‌ها

### فعالیت



شکل الف

در شکل الف افرادی را به ترتیب قد، در یک صف مرتب کرده‌اند و داده‌های مربوط به اندازه قد آنها (برحسب سانتی متر)، به صورت روبه‌رو می‌باشد.  
 در شکل الف در بین پنج فرد، کدام فرد از نظر قد در وسط صف قرار گرفته است؟



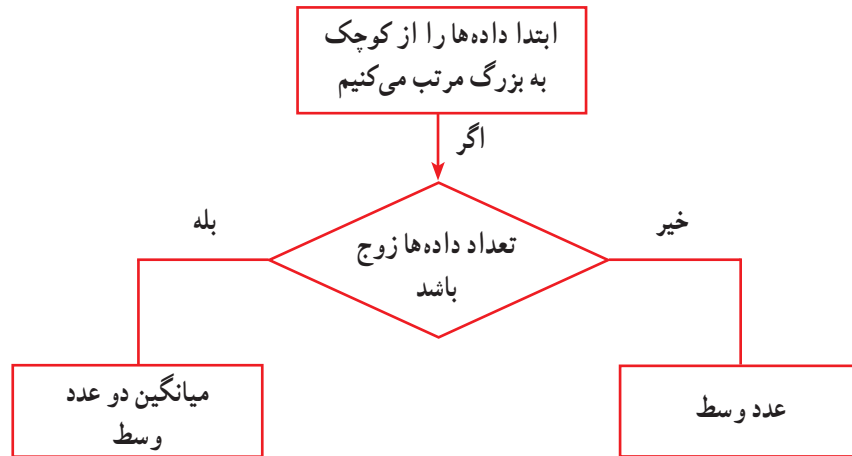
شکل ب

حال به شکل ب توجه کنید. در بین شش فرد، کدام فرد در وسط صف قرار دارد؟  
 همان‌طور که مشاهده می‌شود، به راحتی نمی‌توانید عدد وسط در این حالت را پیدا کنید. برای به دست آوردن عدد وسط در این حالت مراحل زیر را انجام دهید:  
 الف) دو فردی که در جایگاه وسط صف قرار گرفته‌اند را پیدا کنید.  
 ب) میانگین این دو عدد را به عنوان عدد وسط قد این افراد به دست آورید.

میانه، چارک اول و چارک سوم: عدد وسط مجموعه‌ای از داده‌ها را که از کوچک به بزرگ مرتب شده باشند میانه داده‌ها می‌گوییم و آن را با  $Q_2$  نشان می‌دهیم. میانه یک چهارم اول داده‌های مرتب شده را چارک اول داده‌ها می‌گوییم و آن را با  $Q_1$  نشان می‌دهیم. همچنین میانه سه چهارم داده‌های مرتب شده را چارک سوم می‌گوییم و آن را با  $Q_3$  نشان می‌دهیم.



## نحوه به دست آوردن میانه داده‌ها



## کاردر کلاس



در یک شعبه بانک تراکنش‌های مالی بسیاری در یک روز انجام می‌گردد. یک تراکنش مالی ممکن است انتقال مبلغی از حساب پس‌انداز یک مشتری به حساب جاری مشتری دیگری در یک بانک باشد. این تراکنش را می‌توان به دو عملیات تقسیم کرد: بدهکار کردن حساب پس‌انداز یک مشتری به اندازه مبلغ مورد نظر و طلبکار کردن حساب جاری مشتری دیگر به اندازه همان مبلغ است.

الف) فرض کنید تراکنش‌های مالی در بازه زمانی ۸ تا ۹ صبح یک شعبه بانک (به میلیون تومان) به شرح زیر گردآوری شود.

۲۵    ۱۲    ۱۰    ۸/۷    ۱۰

- میانه، چارک اول و سوم مربوط به تراکنش‌های مالی بر اساس داده‌های جمع‌آوری شده را مشخص کنید.
- ب) حال فرض کنید تراکنش‌های مالی دیگری در بازه زمانی ۹ تا ۱۰ صبح در همان شعبه بانک (به میلیون تومان) به شرح زیر گردآوری شود.

۳۴    ۳۲    ۲۰    ۸۱/۷    ۳۰    ۷۰

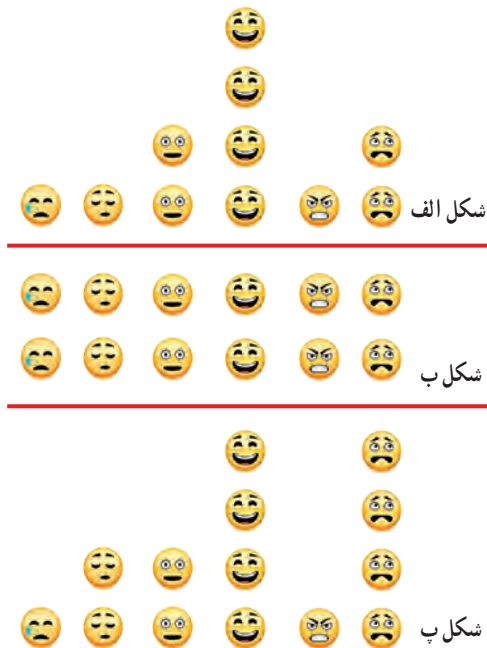
- در این حالت نیز میانه، چارک اول و سوم مربوط به تراکنش‌های مالی بر اساس داده‌های جمع‌آوری شده را مشخص کنید.



## پ) مد یا نما داده‌ها

### فعالیت

به تصاویر روبه‌رو توجه کنید. در شکل (الف)، (ب) و (پ) یک سری از حالت‌های صورتک را مشاهده می‌کنید. تعداد این حالت‌ها را در شکل (الف)، (ب) و (پ) در جدول زیر کامل کنید.



شماره صورتک‌ها	انواع صورتک‌ها	شکل الف	شکل ب	شکل پ
۱		۴	۲	۴
۲			۲	
۳		۲		
۴			۲	
۵				۲
۶				۱

- در شکل الف کدام صورتک بیشتر از همه تکرار شده است؟
- در شکل ب کدام صورتک بیشتر از همه تکرار شده است؟
- در شکل پ کدام صورتک بیشتر از همه تکرار شده است؟

**مد یا نما داده‌ها:** داده‌ای که بیشترین فراوانی را داشته باشد، مد یا نما داده‌ها نام دارد. اگر در داده‌هایی، همه داده‌ها یک فراوانی داشته باشند، آن‌گاه این داده‌ها مد ندارند. اگر در داده‌هایی، دو داده بیشترین فراوانی را داشته باشند، آن‌گاه این داده‌ها دو مد دارند.

### کار در کلاس

در یک مسابقه پرتاب دارت، سه نفر شرکت کرده‌اند. بر اساس ۱۰ پرتابی که آنها انجام داده‌اند، امتیازهای زیر به دست آمده است:



- مد نفر اول چه عددی است؟
- مد نفر دوم چه عددی است؟
- مد نفر سوم چه عددی است؟

۸	۸	۹	۱۰	۹	۵	۷	۱۰	۹	۱۰	نفر اول
۷	۴	۵	۳	۲	۱	۶	۸	۹	۱۰	نفر دوم
۷	۴	۵	۹	۱۰	۱۰	۷	۹	۹	۹	نفر سوم



## میانگین، میانه و مد داده‌ها، کدام معیار را انتخاب می‌کنید؟

### کاردر کلاس



دو کارخانه تولید لامپ را در نظر بگیرید. کارخانه (الف)، لامپ‌های کم مصرف و کارخانه (ب)، لامپ‌های پر مصرف تولید می‌کند. مدیر این دو کارخانه می‌خواهد در مورد طول عمر لامپ‌های تولیدی کارخانه‌هایشان تحقیقی انجام دهد.

بر اساس داده‌های سال‌های گذشته در کارخانه (الف) و (ب)، طول عمر پنج لامپ برحسب ماه ثبت شده است و نتایج را به صورت زیر جمع‌آوری می‌نماید.

پنجم	چهارم	سوم	دوم	اول	لامپ انتخاب شده
۱۶	۱۵	۱۴	۱۵	۱۷	طول عمر لامپ تولید شده در کارخانه (الف)
۱۳	۱۶	۰	۱۵	۰	طول عمر لامپ تولید شده در کارخانه (ب)

■ آیا میانگین طول عمر لامپ‌های تولید شده در کارخانه (الف)، معیار گرایش به مرکز خوبی برای طول عمر لامپ‌های تولید شده کارخانه (الف) است؟

■ به دلیل وجود لامپ‌های تولید شده با طول عمر صفر در کارخانه (ب) آیا بازهم میانگین طول عمر لامپ‌های تولید شده در کارخانه (ب)، معیار گرایش به مرکز خوبی برای طول عمر لامپ‌های تولید شده است؟ چه معیار گرایش به مرکزی مناسب است؟

■ مدیر کارخانه براساس فروش سال گذشته، متوجه شده است که لامپ‌های کم مصرف با نور سفید در منازل مردم رایج شده است. اگر او بخواهد برای امسال لامپ‌های کم مصرف با نور سفید تولید کند، کدام معیار گرایش به مرکز، برای تعداد این لامپ‌های تولیدی به او کمک می‌کند؟

**داده دورافتاده:** مشاهده‌ای که تفاوت بسیار زیادی با سایر مشاهدات مجموعه داده‌ها داشته باشد، میانگین داده‌ها را تحت تأثیر قرار داده در حالی که تأثیری بر میانه و مد داده‌ها ندارد. در فعالیت مربوط به تعداد لامپ‌های تولیدی کارخانه (ب)، عدد صفر داده دورافتاده است.

در تفسیر و تحلیل مسائل آماری، در نظر گرفتن تنها یک شاخص گرایش به مرکز کافی نیست. می‌بایست هر سه معیار میانگین، میانه و مد محاسبه شود و براساس هدف مورد بررسی، معیار مناسب انتخاب و از آن برای انجام تفسیر، قضاوت و پیش‌بینی مورد استفاده قرار گیرد.



## تمرین

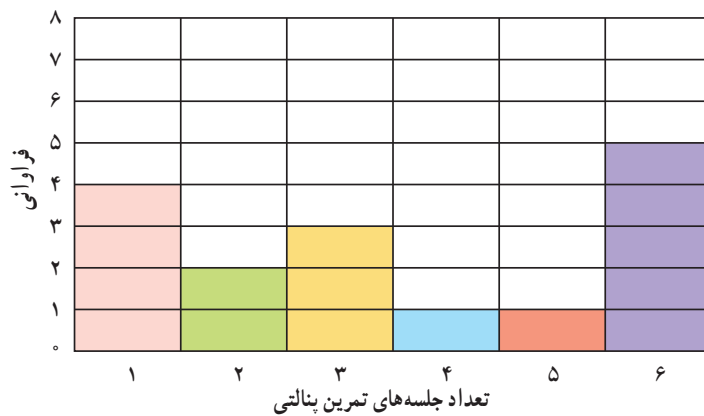
۱ تعداد حمله‌های یک تیم فوتبال در شش مسابقه گذشته به صورت ۴۳، ۴۲، ۴۵، ۴۴، ۴۵، ۴۸ است. میانگین تعداد حملات این تیم در شش بازی گذشته را به دست آورید؟

۲ بالاترین دما در هریک از روزهای هفته گذشته اندازه‌گیری شده و نتایج زیر به دست آمده است. معدل یا میانگین دما در هفته گذشته چه عددی است؟  
۵۵، ۲۷، ۲۹، ۳۲، ۲۸، ۳۱، ۲۹

۳ میانه و مد هریک از داده‌های زیر را به دست آورید؟

الف) ۹، ۹، ۹، ۹، ۸، ۶۰، ۵۰، ۴۰، ۲۴، ۳۰۰ (ب)  
ب) ۱۰، ۳، ۸، ۱۵  
ت) ۴، ۶، ۹، ۱۲، ۵  
ث) ۲۳، ۱۲، ۱۲، ۲۳ (ج)  
ج) ۷، ۴، ۱۳، ۷

۴ نمودار زیر، نمودار میله‌ای مربوط به تعداد ضربات پنالتی گل شده یک بازیکن در شش جلسه تمرین پنالتی است. با توجه به نمودار، میانگین، میانه و مد تعداد ضربات گل شده را به دست آورید؟



۵ در جدول زیر، نمرات درس ریاضی ۱۰ دانش‌آموز گردآوری شده و میانگین نمرات داده شده است. علامت‌های سؤال چه اعدادی اند؟

۱۷/۵	۱۹	۱۷	۱۶	۲۰	نمرات درس ریاضی
۱۶	۱۵	۱۸	؟	۱۸	
میانگین نمرات = ۱۵/۶۵					
مد نمرات = ؟					



۶ داده‌های زیر مدت زمان مطالعه یک دانش‌آموز را در روزهای هفته نشان می‌دهد.

روزهای هفته	شنبه	یکشنبه	دوشنبه	سه شنبه	چهارشنبه	پنجشنبه	جمعه
مدت زمان مطالعه (ساعت)	۲	۱/۵	۲/۵	۱/۵	۲	۳	۳



این دانش آموز به طور میانگین چند ساعت در روز، در هفته گذشته مطالعه کرده است؟

۷ یک شرکت بیمه برای تعیین حق بیمه شخص ثالث در سال آینده، نمونه‌ای از خسارت‌های پرداخت شده امسال را جمع‌آوری نموده است. میانگین خسارت‌های پرداخت شده برابر ۸۵ میلیون ریال به دست آمده است در صورتی که میانه و مد آن برای این خسارت‌های پرداخت شده برابر ۴۲/۲ میلیون ریال و عدد ۹۰ میلیون ریال می‌باشد. به نظر شما مدیر شرکت، کدام معیار گرایش به مرکز را به منظور تعیین حق بیمه در سال آینده در نظر بگیرد تا اینکه این شرکت ضرر نکند؟

۸ دانش آموزی در کنکور سراسری شرکت می‌کند و نتیجه کارنامه آزمون آن به شرح زیر است :

مواد امتحانی	ریاضیات	فیزیک	شیمی	زبان انگلیسی	ادبیات و زبان فارسی	دین و زندگی
درصد	۵۳	؟	۶۷	۳۴	۸۰	۶۷
ضریب درس	۴	۳	۱	۱	۴	۳

اگر معدل موزون درصد این دانش آموز ۷۳ باشد، درس فیزیک را چند درصد زده است؟

۹ میانگین ۵ داده آماری ۱۷ است، اگر دو عدد ۱۷ و ۱۱ را به داده‌های قبلی اضافه کنیم، میانگین جدید چه عددی خواهد شد؟


$$\begin{pmatrix} ۴۷۰ & ۵۸۰ \\ ۶۹۰ & ۶۹۰ \end{pmatrix}$$


$$\begin{pmatrix} ۴۷۰ & ۴۷۰ \\ ۵۸۰ & ۶۹۰ \end{pmatrix}$$

۱۰ دو دانش آموز، جدول‌های چهارخانه‌ای را به صورت روبه‌رو رنگ‌آمیزی کرده‌اند، بر اساس جدول مربوط به طیف رنگ‌ها، جدول عددی این دو شکل به صورت روبه‌رو نشان داده شده است :

حال جدول عددی مربوط به این دو شکل را ابتدا با هم جمع و سپس هریک از اعضای جدول عددی را به عدد ۲ تقسیم می‌کنیم. جدول عددی حاصل را به دست آورده و شکل مورد نظر را با توجه به جدول طیف رنگ‌ها، به دست آورید. آیا این شکل میانگین دو شکل بالا است؟

رنگ‌ها	طیف رنگ‌ها
	۴۵۰ تا ۴۹۵
	۴۹۵ تا ۵۷۰
	۵۷۰ تا ۵۹۰
	۵۷۰ تا ۶۲۰
	۶۲۰ تا ۷۵۰

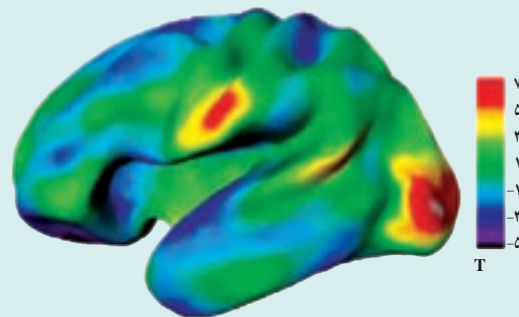
برای پاسخ به این سؤال، کاربرد علم آمار در علوم شناختی و مغز را مطالعه کنید. عدد مربوط به طیف رنگ‌ها در جدول موجود در حاشیه نشان داده شده است.



## کاربرد علم آمار در علوم شناختی و مغز

### تفسیر تصاویر مغزی

در شکل زیر با میانگین تصویر مغزی ۱۰ فرد آشنا می‌شویم.



این تصویر از شکل زیر، برشی از یک تصویر مغزی است. رنگ قرمز نشان دهنده ناحیه‌ای است که در آن فشار خون بالایی وجود دارد. این ناحیه که مشکوک به وجود تومور است با استفاده از علم آمار شناسایی و محل تومور حدس زده می‌شود. به عنوان مثال، نقطه  $a$  به عنوان نقطه‌ای شناخته می‌شود که با احتمال بالایی محل قرارگرفتن تومور است، ولی نقاط  $b$  و  $c$  به رغم داشتن فشار خون بالا، محل تومور نیستند.

