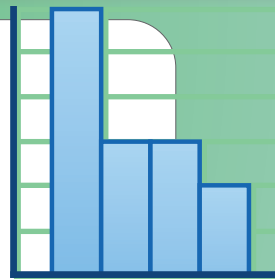


Unidad 8

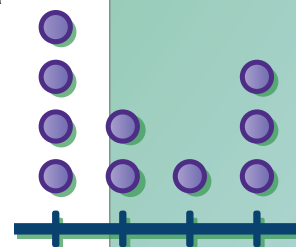
Descripción de datos



La estadística es la ciencia que estudia la recolección y el análisis de los datos. Es una de las áreas más relevantes de las matemáticas en la vida cotidiana. También es una herramienta de investigadores de muchos sectores, como la sociología, donde se usa para identificar y estudiar problemas de la sociedad. En todos los casos, es fundamental saber qué es típico para comprender qué no lo es.

Preguntas esenciales

- ¿Qué son las preguntas estadísticas y cómo se usan?
- ¿Cuáles son algunas formas de representar datos numéricos?
- ¿Cómo se mide la tendencia central de un conjunto de datos?
- ¿Cómo se mide la dispersión de un conjunto de datos?



Puedes usar encuestas y medidas para recopilar datos y responder preguntas sobre un tema. Estos datos pueden ser datos categóricos o datos numéricos. Los datos categóricos pueden clasificarse en categorías. Los datos numéricos son números, cantidades o medidas con los que pueden hacerse comparaciones útiles. Algunos datos que contienen números, como direcciones o fechas, son categóricos porque los números no son cantidades o medidas.

Estos son algunos ejemplos.

Datos categóricos

- Color favorito
- Alimentos que la gente come en el almuerzo
- Números de teléfono de personas

Datos numéricos

- Número de personas en cada aula
- Pesos de perros
- Edades de las personas en tu escuela

Prueba a hacer esto

Antwon hizo la siguiente pregunta a sus amigos: *¿Cuántas horas pasan en sus teléfonos por semana?*

Estos son los datos que recopiló.

12	24	25	0	40	0	28
----	----	----	---	----	---	----

- a** Antwon afirma que la mayoría de sus amigos utilizan el teléfono 3 o más horas al día. ¿Estás de acuerdo? Explica tu razonamiento.

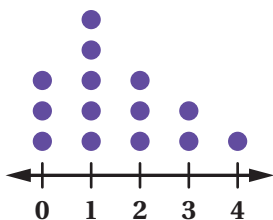
- b** Antwon también preguntó a sus amigos: *¿Tienes más de 2 mascotas?* ¿Esta pregunta producirá datos numéricos o datos categóricos? Explica tu razonamiento.

Una **pregunta estadística** es una pregunta que requiere más de un dato para responderla.

Aquí tienes un ejemplo:

- “¿Cuál aula de tu escuela tiene más libros?” es una pregunta estadística porque para responderla necesitas saber qué cantidad de libros hay en *cada* aula.
- “¿Cuántos libros hay en tu aula?” no es una pregunta estadística porque para responderla solo necesitas saber qué cantidad de libros hay en *un* aula.

Los datos que responden una pregunta estadística pueden organizarse en una lista o en un **diagrama de puntos**.

Lista	0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 3, 3, 4	Las listas permiten ver todos los datos y pueden utilizarse tanto con datos numéricos como categóricos.
Diagrama de puntos		Los diagramas de puntos son una representación visual de datos numéricos y permiten comparar varios conjuntos de datos.

Prueba a hacer esto

Este diagrama de puntos muestra el número de libros que 20 estudiantes de una clase de 6.º grado leyeron en un mes.



- ¿Qué nos dicen los 2 puntos trazados en el 4 sobre esta situación?
- Julián afirma que la mayoría de los estudiantes de la clase leyeron 6 libros o menos este mes. ¿Estás de acuerdo con él? Explica tu razonamiento.

Puedes utilizar diagramas de puntos para visualizar datos y comparar diferentes conjuntos de datos. Al crear un diagrama de puntos, es importante utilizar una escala coherente en la recta numérica y que el espacio vertical entre los puntos sea uniforme.

Aquí tienes un ejemplo de dos diagramas que muestran el mismo conjunto de datos, uno con errores y otro hecho con precisión.



Observaciones:

- La escala del diagrama de puntos con errores no es uniforme, pero la escala del diagrama de puntos preciso muestra uniformidad porque entre cada marca indicadora hay \$0.50.
- La altura de los puntos por encima de \$7.25 y \$8.00 debe ser la misma porque hay la misma cantidad de puntos encima de cada uno.

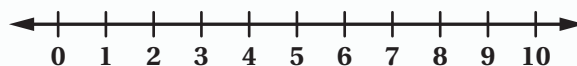
Prueba a hacer esto

Zola preguntó a 8 de sus amigos: *¿Cuántas horas trabajaste en la tienda de la escuela esta semana?*

Recopiló las respuestas en esta tabla.

8	0	2	5	5	2	1	5
---	---	---	---	---	---	---	---

Haz un diagrama de puntos con estos datos.

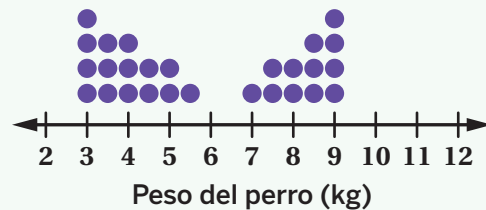


Se pueden usar el **centro**, la **dispersión** y la **forma** para describir la distribución de los datos en un diagrama de puntos. El centro es un número que representa un valor típico. La dispersión describe lo parecidos o diferentes que son los valores de una distribución, a menudo en relación con el centro.

Estos son algunos ejemplos.



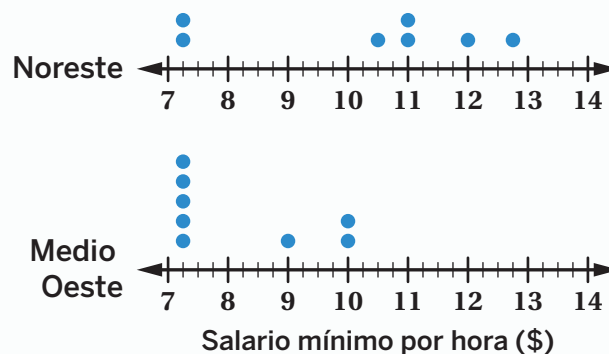
- Los datos son simétricos.
- No hay espacios vacíos entre los datos.
- El centro de los datos está entre 4 y 5 kilogramos.
- La mayoría de los datos están agrupados.
- Los datos tienen forma de montaña, con el mayor número de puntos en 4.5 kilogramos.



- Hay un espacio vacío en los datos, entre 5.5 y 7 kilogramos.
- Los datos están dispersos en dos grupos, uno con un pico en 3 kilogramos y otro con un pico en 9 kilogramos.
- El centro de los datos está aproximadamente en 6 kilogramos.
- Los datos tienen forma de dos triángulos y la mayor cantidad de puntos está en 3 kilogramos y 9 kilogramos.

Prueba a hacer esto

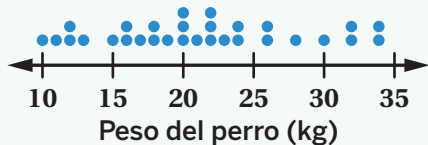
Estos son los diagramas de puntos que muestran los salarios mínimos de los estados de dos partes diferentes de los Estados Unidos.



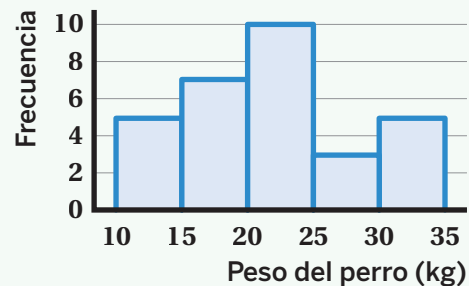
- ¿Qué diagrama de puntos tiene una dispersión mayor?
- ¿Qué nos dicen estos diagramas de puntos sobre los salarios mínimos en distintas partes de los Estados Unidos?

Puedes usar diagramas de puntos e **histogramas** para visualizar datos numéricos. Aquí tienes un ejemplo de un conjunto de datos de los pesos de 30 perros presentados en un diagrama de puntos y en un histograma.

Diagrama de puntos



Histograma



En un histograma, los valores de los datos se agrupan en intervalos que abarcan un conjunto de valores y cada **intervalo** tiene el mismo ancho. La altura de cada barra representa el número total de valores en ese conjunto, incluido el límite izquierdo (menor valor), pero excluyendo el límite derecho (mayor valor). Por ejemplo, la altura de la barra más alta, de 20 a 25, representa los pesos de 20 kilogramos a 25 kilogramos (sin incluir estos últimos).

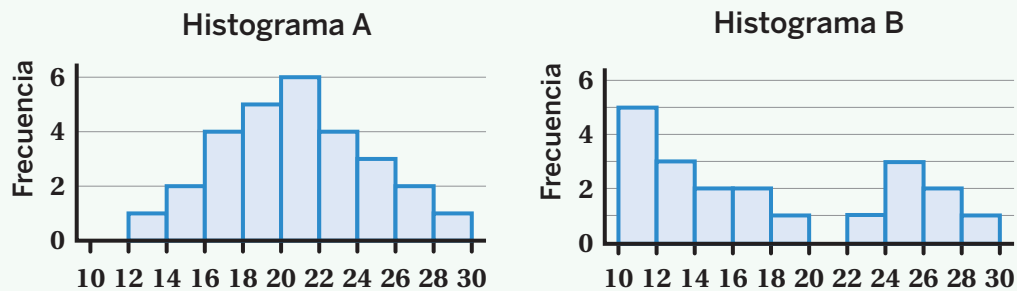
Prueba a hacer esto

Aquí tienes un histograma que muestra los salarios mínimos en los Estados Unidos en 2020.

- ¿Cuántos estados tienen salarios mínimos inferiores a \$10.00?
- Michigan tiene un salario mínimo de \$10.33. ¿En qué intervalo iría?
- Adriana afirma que este histograma representa los salarios mínimos de 4 estados diferentes. ¿Es correcta su afirmación? Explica tu razonamiento.



Puedes usar histogramas para comparar conjuntos de datos numéricos a partir de su forma, centro y dispersión. Aquí tienes dos histogramas con formas y características muy diferentes.



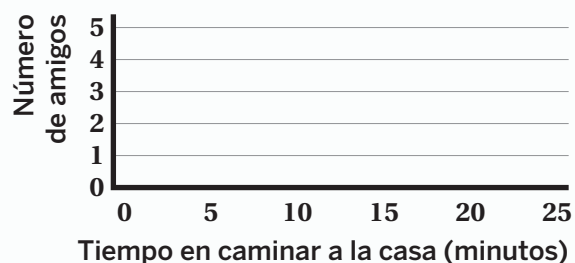
Forma	El histograma A tiene forma de montaña. Es simétrico, con un pico aproximadamente en 21.	El histograma B tiene dos agrupamientos y picos aproximadamente en 11 y 25. Hay un espacio vacío en los datos entre 20 y 22.
Centro	El centro está aproximadamente en 21.	El centro está aproximadamente en 17.
Dispersión	La mayoría de los valores están en el medio y hay aproximadamente la misma cantidad a cada lado del 21.	La mayoría de los valores están a la izquierda y disminuyen hacia la derecha.

Prueba a hacer esto

Marco anotó cuántos minutos tarda cada uno de sus amigos en volver caminando de la escuela a casa.

22	5.5	13	20	12
3	18	21.5	5	13

Haz un histograma de estos datos.



Una **estadística** es un número único que mide un aspecto de un conjunto de datos. Una forma de medir el centro de un conjunto de datos es determinar la **media**, o promedio, de todos los valores de los datos. Puedes entender la media como “una distribución equitativa”.

Por ejemplo, supongamos que este conjunto de datos representa la cantidad de litros de agua que hay en 5 botellas: 1, 4, 2, 3, 0. Para calcular la media, primero se suman todos los valores para determinar el total (10 litros) y luego se divide esa suma por la cantidad de valores (5 botellas). Este ejemplo puede representarse mediante la expresión $(1 + 4 + 2 + 3 + 0) \div 5$, o $10 \div 5$. Así que la cantidad media de agua en las 5 botellas es de 2 litros (por botella). La media es un número natural en este ejemplo, pero es posible que la media sea un número decimal.

Prueba a hacer esto

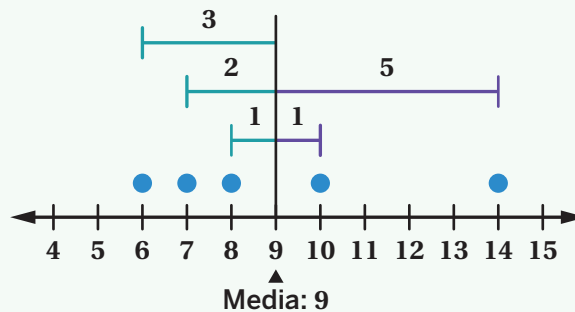
Seis amigos jugaron juntos en el salón recreativo Fly-Score.

Estos son los boletos que ganó cada amigo.

7	3	4	6	8	2
---	---	---	---	---	---

- Calcula la media del número de boletos para estos datos. Muestra tu razonamiento.
- ¿Qué nos indica la media en esta situación?

La distancia entre un punto de datos y la media se denomina **desviación absoluta**. Por ejemplo, el diagrama de puntos a continuación muestra un conjunto de datos con una media de 9. La desviación absoluta del punto en 14 es 5 porque el 14 está a 5 unidades del 9.

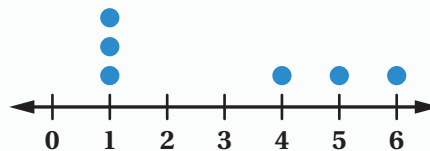


La suma de las desviaciones absolutas a la derecha de la media es igual a la suma de las desviaciones absolutas a la izquierda de la media. Se puede utilizar esto para verificar si un número es la media de un conjunto de datos. Por ejemplo, si 9 es la media correcta del conjunto de datos mostrado, la suma de las desviaciones absolutas a la izquierda del 9 debe ser igual a la suma del lado derecho.

- La suma de las desviaciones a la izquierda es $3 + 2 + 1 = 6$.
- La suma de las desviaciones a la derecha es $1 + 5 = 6$.
- Las sumas son iguales, así que 9 es la media correcta.

Prueba a hacer esto

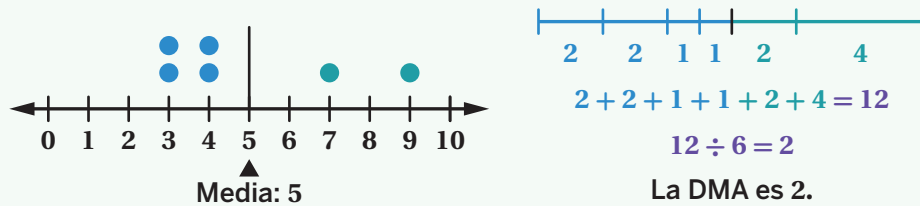
Ava hizo un diagrama de puntos para mostrar cuántas burbujas han reventado ella y sus amigos en dos segundos.



Amir dice que la media de estos datos es 2. ¿Está en lo correcto?

Utiliza las desviaciones absolutas para mostrar o explicar tu razonamiento.

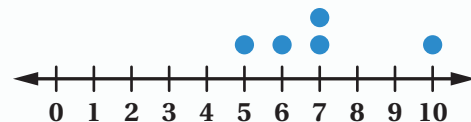
Puedes describir la dispersión de los valores de un conjunto de datos con un único número: la **desviación media absoluta (DMA)**. La DMA se calcula determinando la media de las desviaciones absolutas (es decir, la media de las distancias entre cada valor de los datos y la media).



La desviación media absoluta es un ejemplo de medida de dispersión. Una medida de dispersión es una forma de medir la uniformidad de los valores de un conjunto de datos. Cuanto menor es el valor de la DMA, menor es la dispersión de los puntos de datos con respecto a la media y mayor es la uniformidad de los datos. Cuanto mayor es la DMA, mayor es la dispersión de los puntos de datos con respecto a la media y menor es la uniformidad de los datos.

Prueba a hacer esto

Titus hizo un diagrama de puntos para mostrar el número de canastas que anotó en cada ronda del entrenamiento de baloncesto de la semana pasada.

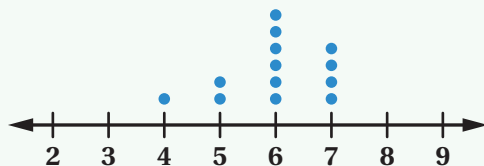


- Calcula la media de estos datos.
- Calcula la DMA de este conjunto de datos.

Al comparar dos conjuntos de datos, la media permite comparar el valor promedio de cada conjunto y la desviación media absoluta (DMA) permite comparar sus dispersiones.

Aquí tienes un ejemplo que muestra las edades de dos grupos de niños que empezaron a aprender a leer y que empezaron a aprender a andar en bicicleta.

Niños que aprendieron a leer



Media: 6 años **DMA:** ≈ 0.62 años

- En promedio, estos niños empezaron a aprender a leer alrededor de los 6 años, que es una edad superior a la edad media para aprender a andar en bicicleta.
- El valor de la DMA es menor, lo que significa que los datos están menos dispersos.

Niños que aprendieron a andar en bicicleta



Media: 5 años **DMA:** ≈ 1.86 años

- En promedio, estos niños empezaron a aprender a andar en bicicleta alrededor de los 5 años, que es una edad inferior a la media para aprender a leer.
- El valor de la DMA es mayor, lo que significa que los puntos de datos están más dispersos.

Prueba a hacer esto

Esta tabla muestra los salarios, en millones de dólares, de las 8 mujeres mejor pagadas de Hollywood en 2019. La media de los salarios de estas mujeres es de \$33.7 millones.

Calcula la DMA de los datos.

Usa la tabla para ayudarte a organizar tus ideas.

Salario (en millones)	Desviación absoluta
\$56	
\$44	
\$35	
\$34	
\$28	
\$25	
\$24	
\$23.5	

Se puede describir un conjunto de datos utilizando otra medida de tendencia central denominada **mediana**. La mediana es el valor de “en medio” en un conjunto de datos cuando los valores se ordenan de menor a mayor (o de mayor a menor). La mitad de los valores de los datos son menores o iguales que la mediana y la mitad de los valores de los datos son mayores o iguales que la mediana.

Para determinar la mediana a partir de una representación ordenada de datos, se puede repetir un proceso de eliminación de los pares de valores menor y mayor.

Estos son algunos ejemplos.

~~0~~ ~~1~~ ~~1~~ 2 ~~2~~ ~~4~~ ~~5~~

Número impar de valores

Una vez eliminados todos los pares, solo queda un valor en medio, que es la mediana.

Mediana: 2

~~0~~ ~~1~~ ~~1~~ 1 2 ~~2~~ ~~4~~ ~~5~~

Número par de valores

Una vez eliminados todos los pares, quedan dos valores.

Su media es la mediana.

$$(1 + 2) \div 2 = 1.5$$

Mediana: 1.5

Prueba a hacer esto

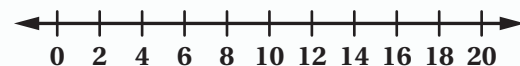
Jayla lanzó su carrito de juguete 7 veces y registró cada distancia en pulgadas.

Distancias del carrito de Jayla:

[5, 6, 17, 19, 9, 20, 18]

¿Cuál es la distancia mediana que recorrió el carrito?

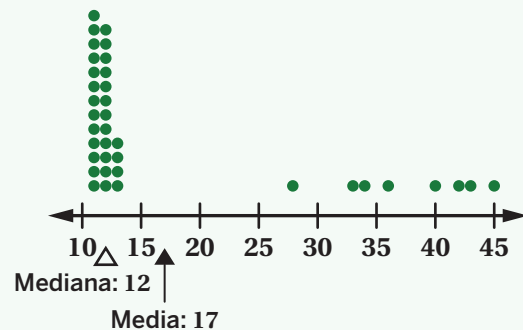
Desarrolla un diagrama de puntos si te ayuda a pensar.



La media y la mediana pueden ser útiles para obtener información diversa sobre un conjunto de datos. Según la forma de los datos y la situación, una de las medidas puede ser la más apropiada.

Aquí hay datos de la fiesta de cumpleaños de un niño de 12 años. La mediana de este conjunto de datos es 12. La media de este conjunto de datos es 17. En esta situación, la mediana es una mejor medida para representar las edades de los invitados a la fiesta porque la mayoría de los invitados tiene entre 11 y 13 años. La media se aleja de donde se encuentra la mayoría de los datos porque las edades mayores se suman a las menores.

Edades de los invitados a la fiesta (años)



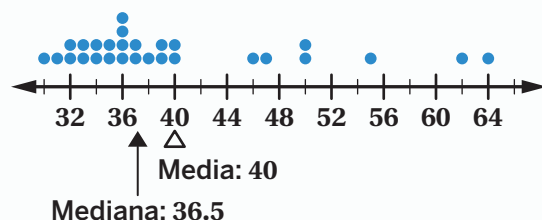
Cuando la mayoría de los datos están en un mismo lugar y hay unos pocos puntos de datos alejados del grupo (como en el diagrama de puntos que se muestra), es probable que la media y la mediana estén muy separadas.

Prueba a hacer esto

Este es un diagrama de puntos que muestra las edades de los maestros de una escuela.

- a** ¿Es la media o la mediana una mejor medida de tendencia central para representar las edades de los maestros?

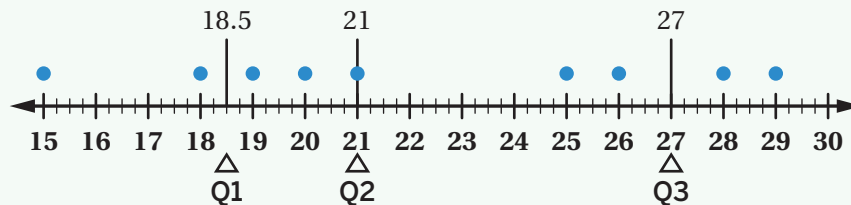
Explica tu razonamiento.



- b** ¿Por qué crees que la media es mayor que la mediana?
- c** ¿Cuándo es probable que la media y la mediana estén muy separadas?

Puedes describir la mitad central de un conjunto de datos dividiéndolo en cuatro secciones iguales denominadas **cuartiles**.

El valor de los cuartiles puede determinarse dividiendo todo el conjunto de datos a la mitad y luego volviendo a dividir las mitades. La mitad central son todos los puntos de datos que se encuentran entre Q1 y Q3. Hay representaciones útiles, como los diagramas de puntos, para identificar los cuartiles y describir conjuntos de datos.



El *primer cuartil (Q1)* es la mediana de la mitad inferior del conjunto de datos.

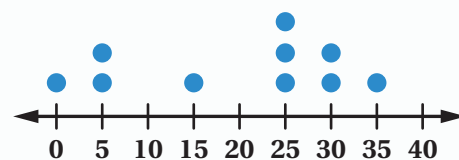
El *segundo cuartil (Q2)* es la mediana de todo el conjunto de datos.

El *tercer cuartil (Q3)* es la mediana de la mitad superior del conjunto de datos.

Prueba a hacer esto

Aki hizo un diagrama de puntos para registrar cuántas ardillas persiguió su perro durante 10 paseos distintos.

Determina los valores de Q1, Q2 (mediana) y Q3.



Q1	
Q2 (mediana)	
Q3	

Puedes hacer un **diagrama de caja** para visualizar un conjunto de datos. Aunque un diagrama de caja muestra los mismos datos que un diagrama de puntos, nos proporciona información adicional sobre los datos. En lugar de mostrar todos los puntos de datos, un diagrama de caja separa los datos en cuartiles.

Podemos utilizar diagramas de caja para describir la dispersión de los datos de dos maneras.

- El **rango** representa la diferencia entre los valores máximo y mínimo de un conjunto de datos. Describe la dispersión general de los datos.

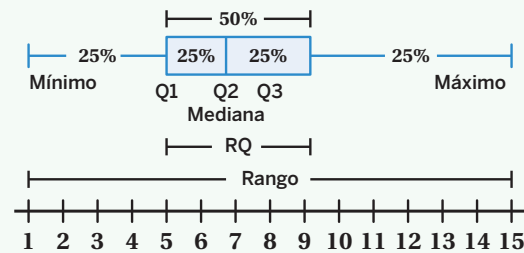
$$\text{Rango: } 15 - 1 = 14$$

- El **rango intercuartílico (RQ)** representa el rango del 50% central de los datos (entre Q3 y Q1). Describe la dispersión del centro de los datos.

$$\text{RQ: } 9.25 - 6.75 = 2.5$$

Los diagramas de caja no muestran cuántos puntos de datos hay en cada conjunto, ni el valor de ningún dato individual, excepto el mínimo y el máximo.

Mín.	Q1	Mediana	Q3	Máx.
1	5	6.75	9.25	15



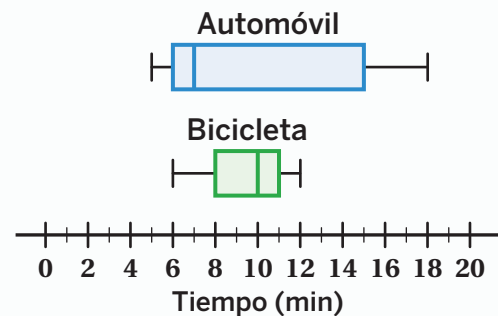
Prueba a hacer esto

Keya hizo diagramas de caja para mostrar cuánto tarda en llegar a la escuela en automóvil y en bicicleta.

- a** Para cada diagrama de caja, determina la mediana, el RQ y el rango.

- b** Si Keya quiere una forma más predecible de viajar, ¿debería ir a la escuela en automóvil o en bicicleta?

Explica tu razonamiento.



	Automóvil	Bicicleta
Mediana		
RQ		
Rango		

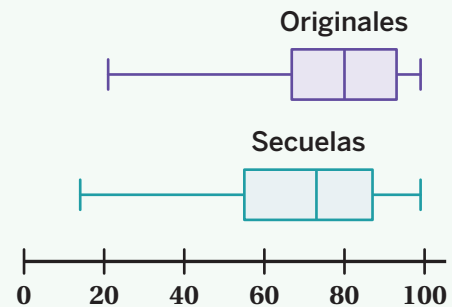
Puedes usar diagramas de caja para hacer muchas afirmaciones diferentes sobre conjuntos de datos.

Por ejemplo, estos diagramas de caja muestran datos de puntuaciones de películas originales y secuelas.

Puedes utilizar el RQ o la mediana para hacer afirmaciones sobre información específica, así como sobre tendencias generales. Por ejemplo:

- El RQ de las originales es 20 y el de las secuelas es 30. Esto nos indica que el 50% central de los datos está más disperso en el caso de las secuelas.
- La puntuación mediana de las originales es mayor que la puntuación mediana de las secuelas. Esto nos indica que las películas originales suelen tener puntuaciones mayores que las secuelas.

Reseñas de películas (puntuación de 0–100)



Originales

Mediana: 80
RQ: 20

Secuelas

Mediana: 75
RQ: 30

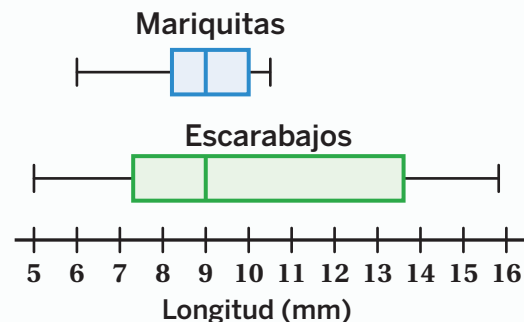
Prueba a hacer esto

Estos diagramas de caja representan los datos de longitud de una colección de mariquitas y de una colección de escarabajos.

- a** Completa la tabla para estimar la mediana y el RQ de cada diagrama de caja.

- b** Compara el RQ de cada colección de insectos.

¿Qué te dicen estas cifras sobre la longitud de las mariquitas y los escarabajos?



	Mediana	RQ
Mariquitas		
Escarabajos		

Puedes responder preguntas estadísticas utilizando estadísticas y representaciones de datos. Aquí tienes datos sobre cuántas horas pasaron viendo películas estos estudiantes y maestros el mes pasado.

Estudiantes				
0	0	5	10	10
20	25	60	75	100

Maestros				
0	15	15	20	20
30	30	35	35	40

Para afirmar algo sobre estos estudiantes y maestros y su hábito de ver películas, podrías hacer representaciones de datos, como diagramas de caja, o calcular medidas de tendencia central (media o mediana) y de dispersión (DMA o RQ).

Estudiantes	Maestros
Media: 30.5 DMA: 28.7 Mediana: 15 RQ: 55	Media: 24 DMA: 10 Mediana: 25 RQ: 20

A partir de estos datos, se podría afirmar que estos maestros vieron más películas porque su mediana es mayor. También se podría afirmar que estos estudiantes ven más películas porque su media es mayor.

Prueba a hacer esto

Durante un periodo de dos semanas, Jada y Mai anotaron el número de problemas matemáticos que recibieron cada día de clase.

Mai piensa que en general tiene más problemas matemáticos que Jada.

¿Es correcto el razonamiento de Mai?
Explica tu respuesta.

Mai

2	15	20	0	5
25	1	0	5	17

Jada

4	0	23	7	5
0	4	10	8	5

Lección 1

- a *Las respuestas pueden variar.* No estoy de acuerdo con Antwon. Tiene datos sobre cuántas horas usan sus amigos sus teléfonos durante toda la semana, pero no puede saber cuántas horas usan sus teléfonos por día. Debería hacer otra pregunta.
- b *Datos categóricos. Las explicaciones pueden variar.* Aunque Antwon pregunta por un número concreto de mascotas, 2, las respuestas serán “Sí” o “No”, que son categóricas.

Lección 2

- a Los 2 puntos nos indican que hay 2 estudiantes en la clase que leyeron 4 libros este mes.
- b Sí. *Las respuestas pueden variar.* Estoy de acuerdo con Julián porque 12 de los 20 estudiantes leen 6 libros o menos y eso es más de la mitad de los estudiantes.

Lección 3



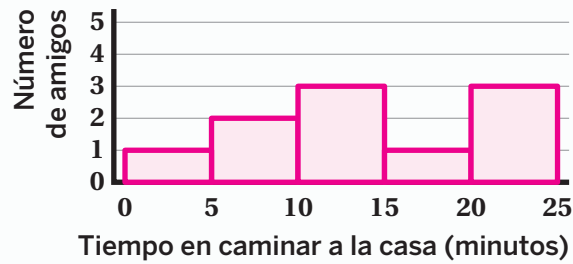
Lección 4

- a Noreste
- b *Las respuestas pueden variar.* Estos diagramas de puntos nos indican que los salarios mínimos de los estados del noreste son más diferentes o están más alejados entre sí. En el Medio Oeste, los salarios mínimos de los estados son más parecidos.

Lección 5

- a 32 estados
Explicación: Una estrategia consiste en sumar el número de estados del intervalo 6–8, que es 21, y el número de estados del intervalo 8–10, que es 11.
- b Intervalo 10–12
- c No. *Las explicaciones pueden variar.* Cada uno de los 4 rectángulos no representa un estado, sino intervalos que organizan los datos. La altura de cada rectángulo indica cuántos estados hay en cada intervalo.

Lección 6



Lección 7

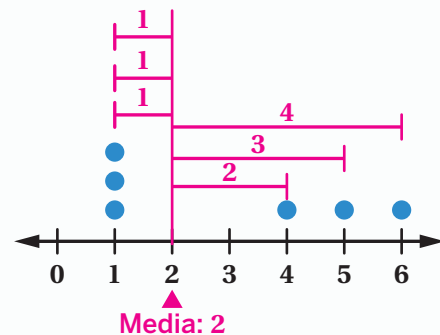
- a) 5 boletos. *Los trabajos pueden variar.* $\frac{7+3+4+6+8+2}{6} = 5$ (o equivalente)
- b) *Las respuestas pueden variar.* En esta situación, la media nos dice cuántos boletos recibiría cada amigo si se distribuyeran todos los boletos equitativamente.

Lección 8

No. *Las explicaciones pueden variar.* La afirmación de Ava no es correcta porque la suma de las desviaciones a la izquierda de 2 no es igual a la suma de las desviaciones a la derecha de 2.

Izquierda: $1 + 1 + 1 = 3$

Derecha: $2 + 3 + 4 = 9$



Lección 9

- a) 7 canastas
- b) 1.2 canastas

Explicación: Esta es una estrategia para calcular la DMA. Halla la distancia entre cada punto de datos y la media, así como la suma de todas las distancias. Luego divide la suma por el número de puntos de datos. $\frac{2+1+0+0+3}{5} = 1.2$

Lección 10

La DMA es aproximadamente \$8.6 millones.

Salario (en millones)	Desviación absoluta
\$56	22.3
\$44	10.3
\$35	1.3
\$34	0.3
\$28	5.7
\$25	8.7
\$24	9.7
\$23.5	10.2

Lección 11

17 pulgadas

Lección 12

- a** Mediana. *Las explicaciones pueden variar.* La mediana, 36.5, es una mejor medida de tendencia central para representar las edades de los maestros que la media, 40, porque la mayoría de los maestros son treintañeros.
- b** La media es mayor que la mediana porque todas esas edades mayores, como 64, se suman a las edades menores y eso hace que la media sea mayor. La mediana es simplemente el número de en medio, por lo que no importa que algunas de las edades sean muy grandes.
- c** Cuando la mayoría de los datos están en un mismo lugar y hay unos pocos puntos de datos alejados del grupo, es probable que la media y la mediana estén muy separadas.

Lección 13

Q1	5
Q2 (mediana)	25
Q3	30

Lección 14

a		Automóvil	Bicicleta
	Mediana	7 min	10 min
	RQ	9 min	3 min
	Rango	13 min	6 min

- b En bicicleta. *Las explicaciones pueden variar.* Keya debería elegir la bicicleta porque el rango de tiempos en automóvil es mayor que el rango de tiempos en bicicleta.

Lección 15

- a *Las estimaciones pueden variar.*

	Mediana	RQ
Mariquitas	9	2
Escarabajos	9	6.5

- b *Las respuestas pueden variar.* Las medianas de las dos colecciones son iguales, pero el RQ de las mariquitas es mucho menor. Esto nos dice que la longitud típica de una mariquita es similar a la longitud típica de un escarabajo, pero las mariquitas son más parecidas entre sí en sus longitudes que los escarabajos.

Lección 16

Las respuestas pueden variar.

- No, porque las medianas de ambos conjuntos de datos son iguales: 5 problemas matemáticos.
- Sí, porque Jada tiene muchos valores más bajos, por lo que su media, 6.6 problemas, es inferior a la media de Mai, 9 problemas.

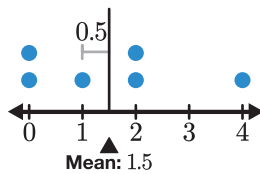
Grade 6 Unit 8 Glossary/6.º grado Unidad 8 Glosario

English

absolute deviation

The distance between a value and the mean of a data set.

If a class's mean number of pets is 1.5, then the absolute deviation of a student who has 1 pet is 0.5.



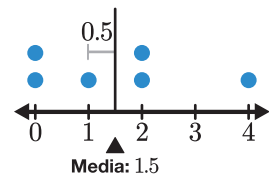
A

Español

desviación absoluta

La distancia entre un valor y la media de un conjunto de datos.

Si la cantidad media de mascotas de una clase es 1.5, entonces la desviación absoluta de un estudiante que tiene 1 mascota es 0.5.

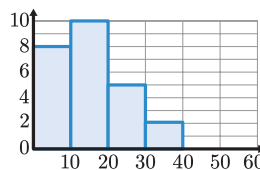


B

bin (of a histogram)

The intervals used to group data values in a histogram.

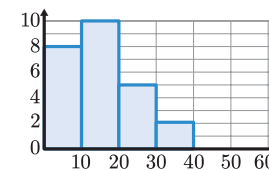
For example, this histogram shows bins of 10 units each.



intervalo (de un histograma)

Los intervalos que se usan para agrupar los valores de datos en un histograma.

Por ejemplo, este histograma muestra intervalos de 10 unidades cada uno.



box plot A way to visualize numerical data sets. The data is divided into four sections using five values: the minimum, Q1, Q2 (or the median), Q3, and the maximum.

A box is drawn between Q1 and Q3. The line inside the box represents the median.

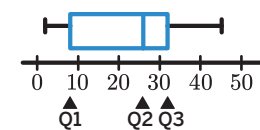
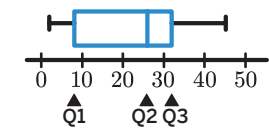


diagrama de caja

Una forma de visualizar conjuntos de datos numéricos.

Los datos se dividen en cuatro secciones utilizando cinco valores: el mínimo, Q1, Q2 (o la mediana), Q3 y el máximo. Se dibuja una caja entre Q1 y Q3. La línea dentro de la caja representa la mediana.



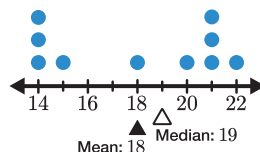
C

categorical data Data that can be sorted into categories instead of counted. Categorical data usually have values that are represented by words instead of numbers.

“What kind of pet do you have?” is a question that would result in categorical data.

center A single value that can be used to summarize the typical value in a data set.

Mean and median are measures of center.

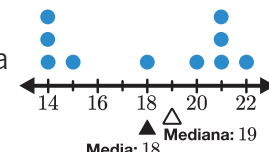


datos categóricos Datos que pueden clasificarse en categorías en lugar de contarse. Los datos categóricos suelen tener valores que se representan mediante palabras en lugar de números.

“¿Qué tipo de mascota tienes?” es una pregunta que produciría datos categóricos.

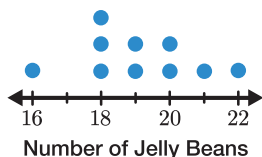
centro Un solo valor que puede utilizarse para resumir el valor típico de un conjunto de datos.

La media y la mediana son medidas de tendencia central, o sea, medidas de centro.



English

dot plot A way to visualize numerical data sets, where each data point is represented by a dot on a number line. Data points with the same value are stacked on top of each other. A dot plot is sometimes called a line plot.



For example, this dot plot shows that 3 students guessed that there were 18 jelly beans in a jar.

Español

diagrama de puntos Una forma de visualizar conjuntos de datos numéricos, en la que cada punto de datos se representa mediante un punto en una recta numérica. Los puntos de datos con el mismo valor se apilan unos sobre otros. Un diagrama de puntos a veces se conoce como gráfica de puntos.



Por ejemplo, este diagrama de puntos muestra que 3 estudiantes estimaron que había 18 frijolitos de jalea en un tarro.

F

frequency table

A table that shows the number of times each value or category occurs in a data set.

Favorite Color	Frequency
Red	5
Blue	3
Pink	4

For example, this frequency table shows that 5 students selected red as their favorite color.

tabla de frecuencia

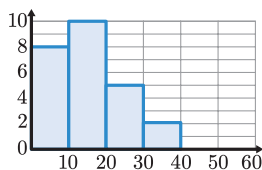
Una tabla que muestra el número de veces que aparece cada valor o categoría en un conjunto de datos.

Color favorito	Frecuencia
Rojo	5
Azul	3
Rosado	4

Por ejemplo, esta tabla de frecuencia muestra que 5 estudiantes seleccionaron el rojo como su color favorito.

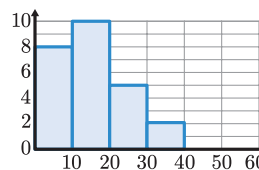
H

histogram A way to visualize numerical data where the data is grouped into bins represented by rectangles. The height of each rectangle shows how many values are in that bin.



For example, this histogram shows that there are 8 data values from 0 up to (but not including) 10.

histograma Una forma de visualizar datos numéricos en la que los datos se agrupan en intervalos que se representan con rectángulos. La altura de cada rectángulo muestra cuántos valores hay en ese intervalo.



Por ejemplo, este histograma muestra que hay 8 valores de datos del 0 al 10 (sin incluirlo).

Grade 6 Unit 8 Glossary/6.º grado Unidad 8 Glosario

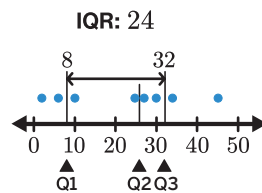
English

Español

I

interquartile range (IQR)

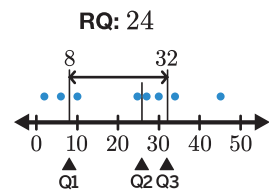
A measure of spread. The IQR is calculated as the distance from Q1 to Q3, or the width of the box in a box plot.



For example, the IQR of this data set is $32 - 8 = 24$.

rango intercuartílico (RQ)

Una medida de dispersión. El RQ se calcula como la distancia de Q1 a Q3, o el ancho de la caja en un diagrama de caja.



Por ejemplo, el RQ de este conjunto de datos es $32 - 8 = 24$.

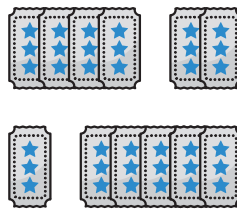
M

maximum The greatest value in a data set.

máximo El mayor valor en un conjunto de datos.

mean (average)

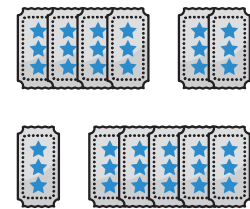
A measure of center. If you equally distribute a set of items into different groups, the mean is the number of items in each group. It is also the balance point of a dot plot. To calculate the mean, you can add the values of all the data points, then divide by the number of data points.



The mean in this example is 3 tickets because that is the number each person would get if the tickets were distributed equally among four people.

media (promedio)

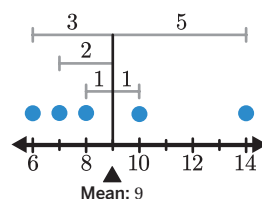
Una medida de tendencia central. Si se distribuye equitativamente un conjunto de elementos en diferentes grupos, la media es la cantidad de elementos en cada grupo. También es el punto de equilibrio de un diagrama de puntos. Para calcular la media, se pueden sumar los valores de todos los puntos de datos y luego dividir por la cantidad de puntos de datos.



La media en este ejemplo es 3 boletos porque es la cantidad que obtendría cada persona si los boletos se distribuyeran equitativamente entre cuatro personas.

mean absolute deviation (MAD)

One way to measure how spread out a data set is from the mean. Sometimes we call this the MAD. To find the MAD, add the distances between each data value and the mean. Then divide by the number of data values in the set.

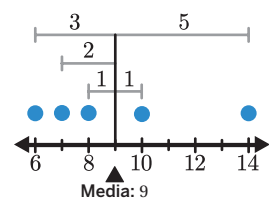


For example, the MAD of this data set is 2.4 because $3 + 2 + 1 + 1 + 5 = 12$ and $\frac{12}{5} = 2.4$.

desviación media absoluta (DMA)

Una forma de medir cuán disperso está un conjunto de datos con respecto de la media.

A veces se denomina DMA. Para calcular la DMA, se suman las distancias entre cada valor y la media. Luego se divide por la cantidad de puntos de datos (valores) del conjunto.



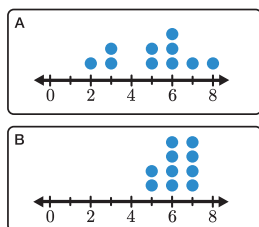
Por ejemplo, la DMA de este conjunto de datos es 2.4 porque $3 + 2 + 1 + 1 + 5 = 12$ y $\frac{12}{5} = 2.4$.

English

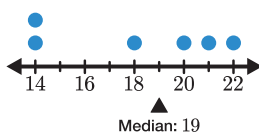
measure of center A single number that summarizes all of the values in a data set. It is usually a typical value for a data set. Mean and median are measures of center.

measure of spread

A single number that describes the spread of a data set. Range, interquartile range, and mean absolute deviation are measures of spread.



median A measure of center. It is the middle value of a data set when the values are in numerical order. When there is an even number of data points, the median is the average of the two middle values.



The median of this data set is 19 because it is the average of the two middle values: 18 and 20.

minimum The least value in a data set.

numerical data Data that includes values that are numbers and can be measured and meaningfully compared.

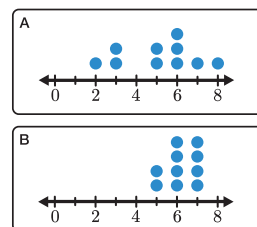
“How many pets do you have?” is a question that would result in numerical data.

Español

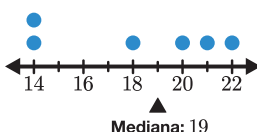
medida de tendencia central, medida de centro Un solo número que resume todos los valores de un conjunto de datos. Suele ser un valor típico de un conjunto de datos. La media y la mediana son medidas de tendencia central.

medida de dispersión

Un solo número que describe la dispersión de un conjunto de datos. El rango, el rango intercuartílico y la desviación media absoluta son medidas de dispersión.



mediana Una medida de tendencia central. Es el valor del medio de un conjunto de datos cuando los valores se clasifican en orden numérico. Cuando hay un número par de puntos de datos, la mediana es el promedio de los dos números centrales.



La mediana de este conjunto de datos es 19 porque es el promedio de los dos valores centrales: 18 y 20.

mínimo El menor valor en un conjunto de datos.

datos numéricos Datos que incluyen valores numéricos que pueden medirse y compararse de forma significativa.

“¿Cuántas mascotas tienes?” es una pregunta que produciría datos numéricos.

Grade 6 Unit 8 Glossary/6.º grado Unidad 8 Glosario

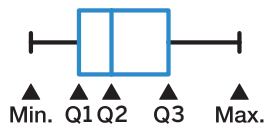
English

quartile Quartiles divide an ordered data set into four equal sections.

Quartile 1 (Q1) is the median of the lower half of the data.

Quartile 2 (Q2) is the median.

Quartile 3 (Q3) is the median of the upper half of the data.



Q

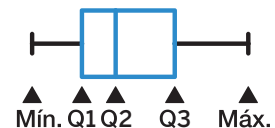
Español

cuartil Los cuartiles dividen un conjunto de datos ordenado en cuatro secciones iguales.

El cuartil 1 (Q1) es la mediana de la mitad inferior de los datos.

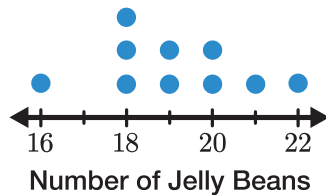
El cuartil 2 (Q2) es la mediana.

El cuartil 3 (Q3) es la mediana de la mitad superior de los datos.



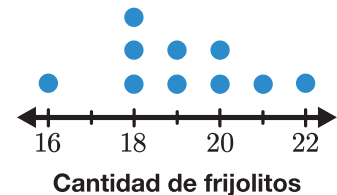
R

range A measure of spread. It is the difference between the maximum and minimum values in a data set.



For example, the range of this data set is 6 jelly beans because $22 - 16 = 6$.

rango Una medida de dispersión. Es la diferencia entre los valores máximo y mínimo de un conjunto de datos.



Por ejemplo, el rango de este conjunto de datos es 6 frijolitos de jalea porque $22 - 16 = 6$.

S

shape A description of the distribution (or pattern) of the data within a data set. Descriptions of the shape of a distribution can include: symmetry, skew, clusters, and outliers.

spread A description of how alike or different the values in a distribution are, often in relation to the center. The spread also describes the variability of a distribution.

For example, Dot Plot A has a larger spread than Dot Plot B.

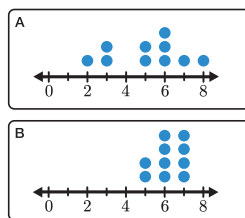
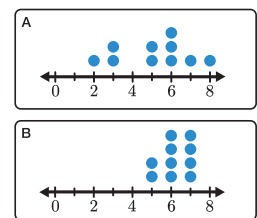


figura Una descripción de la distribución (o patrón) de los datos en un conjunto de datos. Las descripciones de la forma de una distribución pueden incluir: simetría, asimetría, agrupaciones y valores atípicos.

dispersión Una descripción de las semejanzas o diferencias de los valores en una distribución, a menudo en relación con el centro. La dispersión también describe la variabilidad de una distribución.

Por ejemplo, el diagrama de puntos A tiene una dispersión mayor que el diagrama de puntos B.



English

statistic A single number that measures something about a data set.

Examples of statistics include: mean, median, MAD, and IQR.

statistical question A question that requires more than one piece of data to answer.

Here are some examples of statistical questions:
What is the most popular band at your school?
When do students in your class typically eat dinner?

Español

dato estadístico, estadística Un número único que mide algo de un conjunto de datos.

Ejemplos de datos estadísticos: media, mediana, DMA y RQ.

pregunta estadística Una pregunta que requiere más de un dato para responder.

Estos son algunos ejemplos de preguntas estadísticas:
¿Cuál es la banda más popular en tu escuela?
¿Cuándo suelen cenar los estudiantes de tu clase?