

Lección 1: Formulación de preguntas estadísticas

La Estadística trata acerca del uso de datos para responder preguntas. En este módulo, vas a trabajar con datos y a crear resúmenes de la información obtenida siguiendo los cuatro pasos siguientes:

Paso 1: Formula una pregunta que se pueda responder con datos.

Paso 2: Determina un plan para recabar los datos.

Paso 3: Resume los datos con gráficos y resultados numéricos.

Paso 4: Responde la pregunta formulada en el paso 1 utilizando los datos y los resúmenes.

Se te guiará a lo largo de este proceso a medida que estudies estas lecciones. La primera lección trata acerca del primer paso: ¿qué es una pregunta estadística y qué significa que una pregunta se pueda responder con datos?

Trabajo en clase

Ejemplo 1: ¿Qué es una pregunta estadística?

Jerome, un estudiante de sexto grado de la escuela media Roosevelt, es un gran fan del béisbol. Le encanta coleccionar tarjetas de béisbol. Tiene tarjetas de jugadores actuales y de jugadores de temporadas de béisbol anteriores. Con el permiso de su maestro, Jerome llevó su colección de tarjetas de béisbol a la escuela. Cada tarjeta tiene una foto de un jugador de béisbol actual o retirado de las Grandes Ligas, junto con información sobre el jugador. Cuando Jerome sacó sus tarjetas para que los otros estudiantes las vieran, estos hicieron toda clase de preguntas sobre las tarjetas. Algunos preguntaron:

- ¿Cuántas tarjetas tiene Jerome en total?
- ¿Cuál es el coste normal de una tarjeta de la colección de Jerome?
- ¿Dónde obtuvo las tarjetas Jerome?

Ejercicios 1 a 5

1. Para cada una de las siguientes preguntas, determina si se trata de una pregunta estadística o no. Justifica tu respuesta.
 - a. ¿Quién es mi estrella de cine favorita?

 - b. ¿Cuáles son los colores favoritos de los estudiantes de sexto grado de mi escuela?

- c. ¿Durante cuántos años han tocado un instrumento los estudiantes de la banda u orquesta de mi escuela?
- d. ¿Cuál es la materia favorita de los estudiantes de sexto grado de mi escuela?
- e. ¿Cuántos hermanos y hermanas tiene mi mejor amigo/a?
2. Explica por qué cada una de las siguientes preguntas no es una pregunta estadística.
- a. ¿Cuántos años tengo?
- b. ¿Cuál es mi color favorito?
- c. ¿Cuántos años tiene el director de nuestra escuela?
3. Ronnie, un estudiante de sexto grado, quería averiguar si era el que vivía más lejos de la escuela. Escribe una pregunta estadística que ayudaría a Ronnie a encontrar la respuesta.
4. Escribe una pregunta estadística que se pueda responder recabando datos de los estudiantes de tu clase.
5. Cambia la siguiente pregunta para convertirla en una pregunta estadística: ¿cuántos años tiene mi maestro de matemáticas?

Ejemplo 2: Tipos de datos

Utilizamos dos tipos de datos para responder preguntas estadísticas: datos numéricos y datos categóricos. Si anotamos la antigüedad de 25 tarjetas de béisbol, tendríamos datos numéricos. Cada valor de un conjunto de datos numéricos es un número. Si anotamos el equipo de los jugadores que aparecen en 25 tarjetas de béisbol, tendrías datos categóricos. Aunque todavía tienes 25 valores de datos, los valores de datos no son números. Serían nombres de equipos, los cuales puedes considerar como categorías.

Ejercicios 6 y 7

6. Identifica cada uno de los siguientes conjuntos de datos como categóricos (C) o numéricos (N).
- Alturas de 20 estudiantes de sexto grado _____
 - Sabor de helado favorito de cada estudiante de sexto grado en un grupo de 10 estudiantes _____
 - Horas de sueño de 30 estudiantes de sexto grado durante la semana escolar _____
 - Tipo de bebida que bebe en el almuerzo cada estudiante de sexto grado en un grupo de 15 estudiantes _____
 - Color de ojos de cada estudiante de sexto grado en un grupo de 30 estudiantes _____
 - Cantidad de lápices en el escritorio de cada estudiante de sexto grado en un grupo de 15 estudiantes _____
7. Para cada una de las siguientes preguntas estadísticas, los estudiantes le pidieron a Jerome que identificara si los datos eran numéricos o categóricos. Explica tu respuesta y enumera cuatro valores de datos posibles.
- ¿Cuánta antigüedad tienen las tarjetas de la colección?
 - ¿Cuánto costaron las tarjetas de la colección?
 - ¿Dónde obtuviste las tarjetas?

Resumen de la lección

La Estadística trata acerca del uso de datos para responder preguntas. En este módulo vas a trabajar con datos y a crear resúmenes de la información obtenida siguiendo los cuatro pasos siguientes:

Paso 1: Formula una pregunta que se pueda responder con datos.

Paso 2: Determina un plan para recabar los datos.

Paso 3: Resume los datos con gráficos y resultados numéricos.

Paso 4: Responde la pregunta formulada en el paso 1 utilizando los datos y los resúmenes.

Conjunto de problemas

- Para cada una de las siguientes preguntas, determina si se trata de una pregunta estadística. Justifica tu respuesta.
 - ¿Cuántas letras hay en mi apellido?
 - ¿Cuántas letras hay en los apellidos de los estudiantes de mi clase de sexto grado?
 - ¿Cuáles son los colores de los zapatos que usan los estudiantes de mi escuela?
 - ¿Cuál es la cantidad máxima de pies que los vagones de las montañas rusas bajan durante un recorrido?
 - ¿Cuáles son las frecuencias cardíacas de los estudiantes de una clase de sexto grado?
 - ¿Cuántas horas suelen dormir por noche los estudiantes cuando tienen que ir a la escuela al día siguiente?
 - ¿Cuántas millas por galón recorren los automóviles compactos?
- Identifica cada uno de los siguientes conjuntos de datos como categóricos (C) o numéricos (N). Explica tu respuesta.
 - Envergadura de los brazos de 12 estudiantes de sexto grado.
 - Cantidad de idiomas que habla cada adulto en un grupo de 20 adultos.
 - Deporte favorito de cada persona en un grupo de 20 adultos.
 - Cantidad de mascotas que tiene cada estudiante de tercer grado en un grupo de 40 estudiantes.
 - Cantidad de horas por semana que un grupo de estudiantes de escuela media pasa leyendo un libro.
- Reescribe cada una de las siguientes preguntas como una pregunta estadística.
 - ¿Cuántas mascotas tiene tu maestro?
 - ¿Cuántos puntos anotó el equipo de fútbol de la escuela secundaria en su último partido?
 - ¿Cuántas páginas hay en nuestro libro de Matemáticas?
 - ¿Puedo hacer el pino (pararme sobre las manos)?
- Escribe una pregunta estadística que se podría responder recabando datos de los estudiantes de sexto grado de tu aula.

5. ¿Los datos que recabarías para responder esa pregunta son categóricos o numéricos? Explica tu respuesta.

Lección 2: Visualización de una distribución de datos

Trabajo en clase

Ejemplo 1: Frecuencia cardíaca

Mia, una estudiante de sexto grado de la escuela media Roosevelt, estaba pensando en unirse al equipo de atletismo de su escuela. Ella leyó que los atletas olímpicos tienen una frecuencia cardíaca en reposo menor que la mayoría de la gente. Se preguntó por su propia frecuencia cardíaca y cómo sería esta con respecto a la de otros estudiantes. Mia tenía interés en investigar la pregunta estadística: ¿cuáles son las frecuencias cardíacas de los estudiantes de mi clase de sexto grado?

Las frecuencias cardíacas se expresan como ppm (o pulsaciones por minuto). Mia sabía que su frecuencia cardíaca en reposo era de 80 pulsaciones por minuto. Le preguntó a su maestro si podía recabar las frecuencias cardíacas de los otros estudiantes de su clase. Con la ayuda de su maestro, los otros estudiantes de sexto grado de su clase averiguaron sus frecuencias cardíacas y se las dijeron a Mia. A continuación, se encuentran las frecuencias cardíacas (en pulsaciones por minuto) de los otros 22 estudiantes de la clase de Mia:

89 87 85 84 90 79 83 85 86 88 84 81 88 85 83 83 86 82 83 86 82 84

Ejercicios 1 a 10

1. ¿Cuál era la frecuencia cardíaca del estudiante con la frecuencia cardíaca más baja?
2. ¿Cuál era la frecuencia cardíaca del estudiante con la frecuencia cardíaca más alta?

3. ¿Cuántos estudiantes tenían una frecuencia cardíaca mayor que 86?
4. ¿Qué fracción de los estudiantes tenía una frecuencia cardíaca menor que 82?
5. ¿Cuál es la frecuencia cardíaca más común?
6. ¿Qué frecuencia cardíaca describe el centro de los datos?
7. ¿Qué frecuencias cardíacas son las más inusuales?
8. Si el maestro de Mia preguntara cuál es la frecuencia cardíaca más común de los estudiantes de sexto grado de la clase, ¿qué le dirías al maestro de Mia?
9. Agrega un punto para la frecuencia cardíaca de Mia en el diagrama de puntos del ejemplo 1.
10. ¿Qué diferencia hay entre la frecuencia cardíaca de Mia y la de los otros estudiantes de la clase?

Ejemplo 2: Ver la dispersión en diagramas de puntos

La clase de Mia recabó datos para responder muchas otras preguntas sobre su clase. Después de recabar los datos, dibujaron diagramas de puntos de sus hallazgos.

Aquí hay un diagrama de puntos que muestra los datos recabados para responder la pregunta: ¿cuántos libros de texto hay en los escritorios de los estudiantes de sexto grado?

Diagrama de puntos de la cantidad de libros de texto



Otro estudiante de la clase de Mia quería preguntar: ¿qué tan altos son los estudiantes de sexto grado de nuestra clase?

Este diagrama de puntos muestra la altura de los estudiantes de sexto grado de la clase de Mia.

Diagrama de puntos sobre la altura

**Ejercicios 11 a 14**

A continuación se enumeran cuatro preguntas estadísticas y cuatro diagramas de puntos diferentes de datos recabados para responder estas preguntas. Une cada pregunta estadística con el diagrama de puntos adecuado y explica cada elección.

Preguntas estadísticas:

11. ¿Cuáles son las edades de los estudiantes de cuarto grado de nuestra escuela?

12. ¿Cuáles son las alturas de los jugadores del equipo de baloncesto masculino de octavo grado?

13. ¿Cuántas horas de televisión ver los estudiantes de sexto grado de nuestra clase en una noche durante la semana escolar?

14. ¿Cuántos idiomas diferentes hablan los estudiantes de nuestra clase?

Diagrama de puntos A



Diagrama de puntos B

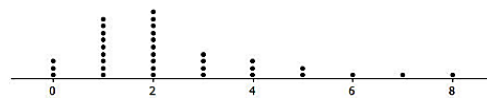


Diagrama de puntos C

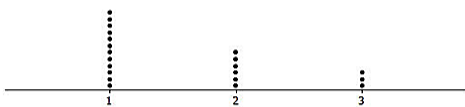
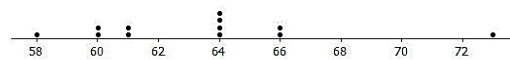


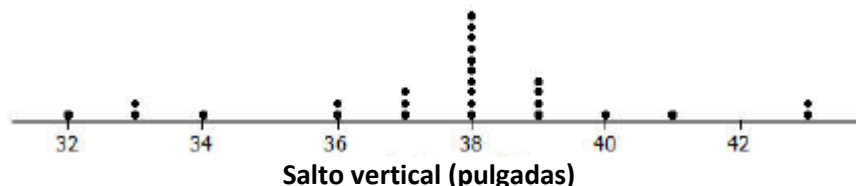
Diagrama de puntos D



Conjunto de problemas

- El siguiente diagrama de puntos muestra el salto vertical de algunos jugadores de la NBA. Un salto vertical es qué tan alto puede saltar un jugador sin tomar carrera.

Diagrama de puntos de salto vertical



- ¿Qué pregunta estadística piensas que se podría responder utilizando estos datos?
 - ¿Cuál fue el salto vertical más alto de un jugador?
 - ¿Cuál fue el salto vertical más bajo de un jugador?
 - ¿Cuál es el salto vertical más común?
 - ¿Cuántos jugadores realizaron el salto vertical más común?
 - ¿Cuántos jugadores saltaron más alto que 40 pulgadas?
 - Otro jugador de la NBA saltó 33 pulgadas. Agrega un punto para este jugador en el diagrama de puntos. ¿Qué diferencia hay entre este jugador y los otros jugadores?
- A continuación se enumeran dos preguntas estadísticas y dos diagramas de puntos diferentes de datos recabados para responder estas preguntas. Une cada pregunta estadística con su diagrama de puntos. Explica cada una de tus elecciones.

Preguntas estadísticas:

- ¿Cuál es la cantidad de peces (si hubiera) que los estudiantes de la clase tienen en un acuario en sus casas?
- ¿Cuántas personas viven en las casas de los estudiantes de sexto grado?

Diagrama de puntos A



Diagrama de puntos B



- Lee cada una de las siguientes preguntas estadísticas. Escribe una descripción acerca de cómo podría verse el diagrama de puntos de los datos recabados para responder la pregunta. Tu descripción debería incluir una descripción de la dispersión de los datos y del centro de los datos.
 - ¿Cuántas horas pasan en la escuela los estudiantes de sexto grado en un día escolar habitual?
 - ¿Cuál es la cantidad de videojuegos que tienen los estudiantes de sexto grado de nuestra clase?

Lección 3: Creación de un diagrama de puntos

Trabajo en clase

Ejemplo 1: Horas de sueño

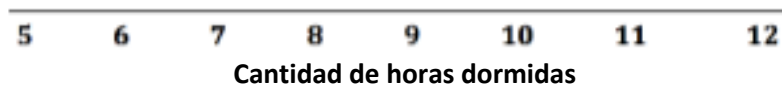
Robert, un estudiante de sexto grado de la escuela media Roosevelt, suele irse a dormir alrededor de las 10:00 p. m. y levantarse a las 6:00 a. m. para prepararse para ir a la escuela. Eso significa que duerme aproximadamente 8 horas en una noche durante la semana escolar. Él decidió investigar la pregunta estadística: ¿cuántas horas por noche suelen dormir los estudiantes de sexto grado durante la semana escolar?

Robert encuestó a 29 estudiantes de sexto grado y recabó los siguientes datos para responder la pregunta:

7 8 5 9 9 9 7 7 10 10 11 9 8 8 8 12 6 11 10 8 8 9 9 9 8 10 9 9 8

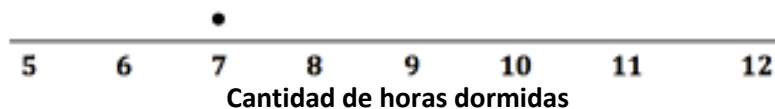
Robert decidió hacer un diagrama de puntos de los datos para poder responder su pregunta estadística. Primero, Robert dibujó una línea numérica y la rotuló del 5 al 12 para indicar las cantidades más baja y más alta de horas dormidas.

Diagrama de puntos de la cantidad de horas dormidas



Luego, colocó un punto encima del 7 para el primer dato que recabó. Continuó colocando puntos encima de los números hasta que cada uno estaba representado por un punto.

Diagrama de puntos de la cantidad de horas dormidas



Ejercicios 1 a 9

1. Completa el diagrama de puntos de Robert colocando un punto encima del número de la línea numérica para cada cantidad de horas dormidas. Si ya hay un punto encima de un número, agrega otro punto encima del punto que hay.

2. ¿Cuáles son las cantidades menor y mayor de horas dormidas en la encuesta de los estudiantes de sexto grado?
3. ¿Cuál es la cantidad más común de horas dormidas?
4. ¿Cuántas horas de sueño describen el centro de los datos?
5. Piensa cuántas horas sueles dormir en una noche durante la semana escolar. ¿Qué diferencia hay entre tu cantidad y la cantidad de horas de sueño que aparece en la encuesta que se hizo a los estudiantes de sexto grado?

Aquí están los datos de la cantidad de horas que duermen los estudiantes de sexto grado cuando no tienen que ir a la escuela al día siguiente:

7 8 10 11 5 6 12 13 13 7 9 8 10 12 11 12 8 9 10 11 10 12 11 11 11 12 11 11 10

6. Haz un diagrama de puntos de la cantidad de horas dormidas cuando no hay que ir a la escuela al día siguiente.
7. ¿Cuántas horas de sueño describen el centro de los datos cuando no hay que ir a la escuela al día siguiente?
8. ¿Cuáles son las cantidades menor y mayor de horas dormidas en la encuesta cuando no hay que ir a la escuela al día siguiente?

Ejercicios 10 a 15

10. Completa la columna de marcas de conteo de la tabla del ejemplo 2.
11. Para cada cantidad de horas, encuentra la cantidad total de marcas de conteo y colócala en la columna de frecuencias de la tabla del ejemplo 2.
12. Haz un diagrama de puntos de la cantidad de horas en las que se practica un deporte o se juega al aire libre.
13. ¿Qué cantidad de horas describe el centro de los datos?
14. ¿Cuántos estudiantes de sexto grado contestaron que pasan ocho o más horas por semana practicando un deporte o jugando al aire libre?
15. Los estudiantes de sexto grado querían responder la pregunta "¿cuántas horas por semana pasan los estudiantes de sexto grado practicando un deporte o jugando al aire libre?". ¿Cómo responderías la pregunta del estudiante de sexto grado utilizando la tabla de frecuencias y el diagrama de puntos?

Conjunto de problemas

1. Los siguientes datos representan la cantidad de goles marcados por un equipo de fútbol sala profesional en sus últimos 23 partidos.

8 16 10 9 11 11 10 15 16 11 15 13 8 9 11 9 8 11 16 15 10 9 12

- Haz un diagrama de puntos de la cantidad de goles marcados.
 - ¿Qué cantidad de goles describe el centro de los datos?
 - ¿Cuál es la menor y la mayor cantidad de goles marcados por el equipo?
 - De los 23 partidos jugados, el equipo perdió 10. Encierra en círculos los puntos del diagrama que piensas que representan los partidos que el equipo perdió. Explica tu respuesta.
2. Un estudiante de sexto grado tiró dos dados con números 21 veces. El estudiante sumó los dos números que salieron cada vez. Las siguientes son las sumas de las 21 veces que tiró los dos cubos:

9 2 4 6 5 7 8 11 9 4 6 5 7 7 8 8 7 5 7 6 6

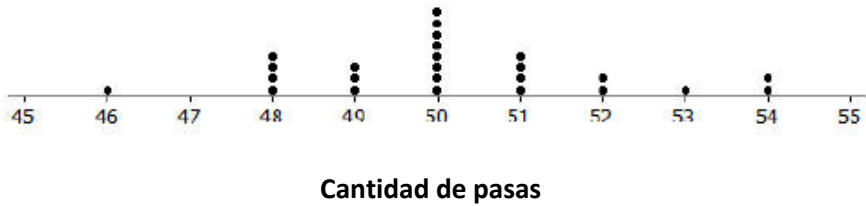
- a. Completa la tabla de frecuencias.

Suma que salió	Conteo	Frecuencia
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

- ¿Qué suma describe el centro de los datos?
- ¿Cuál fue la suma más común de los dados ?

3. El siguiente diagrama de puntos muestra la cantidad de pasas que hay en 25 cajas pequeñas de pasas.

Diagrama de puntos de la cantidad de pasas



- a. Completa la tabla de frecuencias.

Cantidad de pasas	Conteo	Frecuencia
46		
47		
48		
49		
50		
51		
52		
53		
54		

- b. Otro estudiante abrió una caja de pasas e informó que tenía 63 pasas. ¿Tenía este estudiante una caja de pasas del mismo tamaño? ¿Por qué sí o por qué no?

Lección 4: Creación de un histograma

Trabajo en clase

Ejemplo 1: Tabla de frecuencias con intervalos

Los equipos de baloncesto de niños y de niñas de la escuela media Roosevelt querían recaudar dinero para ayudar a comprar nuevos uniformes. Decidieron vender gorras con el logotipo de la escuela en la parte delantera a familiares y a otras personas interesadas. Para obtener el tamaño de gorra correcto, los estudiantes tuvieron que medir la circunferencia de la cabeza (distancia alrededor de la cabeza) de los adultos que querían pedir una gorra. El siguiente conjunto de datos representa las circunferencias de las cabezas de los adultos en milímetros (mm):

513, 525, 531, 533, 535, 535, 542, 543, 546, 549, 551, 552, 552, 553, 554, 555, 560, 561, 563, 563, 563, 565, 565, 568, 568, 571, 571, 574, 577, 580, 583, 583, 584, 585, 591, 595, 598, 603, 612, 618

Las gorras vienen en seis tamaños: XS, S, M, L, XL y XXL. Cada tamaño de gorra cubre una variedad de circunferencias de cabezas. El fabricante de gorras les dio a los estudiantes la siguiente tabla que muestra el intervalo de las circunferencias de cabezas para cada tamaño de gorra. El intervalo $510 - < 530$ representa las circunferencias de cabezas desde 510 a 530 (exclusive).

Tamaños de gorras	Intervalo de circunferencias de cabezas (mm)	Conteo	Frecuencia
XS	$510 - < 530$		
S	$530 - < 550$		
M	$550 - < 570$		
L	$570 - < 590$		
XL	$590 - < 610$		
XXL	$610 - < 630$		

Ejercicios 1 a 4

- ¿Qué tamaño de gorra necesitaría alguien con una circunferencia de cabeza de 570 mm?

2. Completa las columnas de conteos y de frecuencias de la tabla del ejemplo 1 para determinar la cantidad de cada tamaño de gorra que los estudiantes necesitan pedir para los adultos que querían pedir una gorra.
3. ¿En qué tamaño de gorra se centran los datos?
4. Describe cualquier patrón que observes en la columna de frecuencias.

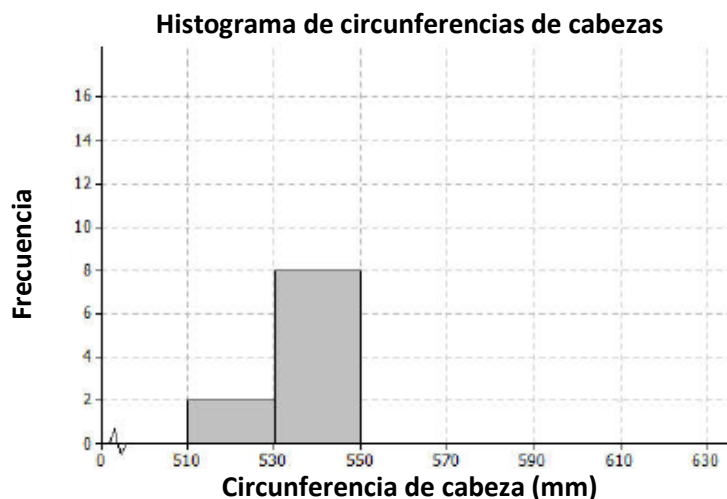
Ejemplo 2: Histograma

Un estudiante miró la columna de conteos y dijo que se parecía a un gráfico de barras colocado de lado. Un histograma es un gráfico que se parece a un gráfico de barras, excepto que el eje horizontal es una línea numérica que está marcada en intervalos iguales.

Para hacer un histograma:

- dibuja una línea horizontal y marca los intervalos;
- dibuja una línea vertical y rotúlala como "Frecuencia";
- marca el eje de Frecuencia con una escala que comience en 0 y aumente hasta un número que sea mayor que la frecuencia más grande de la tabla de frecuencias;
- para cada intervalo, dibuja una barra sobre el intervalo que tenga una altura igual a la frecuencia para ese intervalo.

A continuación, se dibujaron las primeras dos barras del histograma.



Ejercicios 5 a 9

5. Completa el histograma dibujando barras cuyas alturas sean las frecuencias de estos intervalos.

6. Según el histograma, describe el centro de las circunferencias de cabezas.

7. ¿Cómo cambiaría el histograma si sumaras las circunferencias de cabezas de 551 y 569 mm?

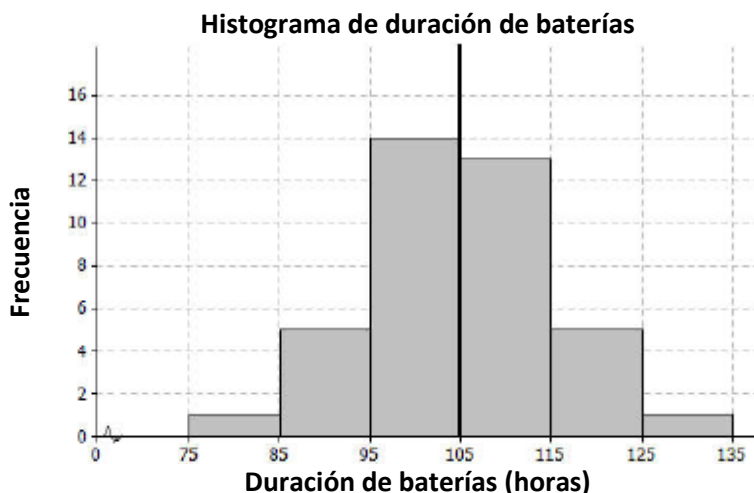
8. Como se proporcionaron los 40 valores de circunferencias de cabezas, podrías haber creado un diagrama de puntos a fin de mostrar los datos de las circunferencias de cabezas. ¿Qué información se pierde cuando se utiliza un histograma para representar una distribución de datos en lugar de un diagrama de puntos?

9. Supón que había 200 mediciones de circunferencias de cabezas en el conjunto de datos. Explica por qué preferirías resumir este conjunto de datos utilizando un histograma en lugar de un diagrama de puntos.

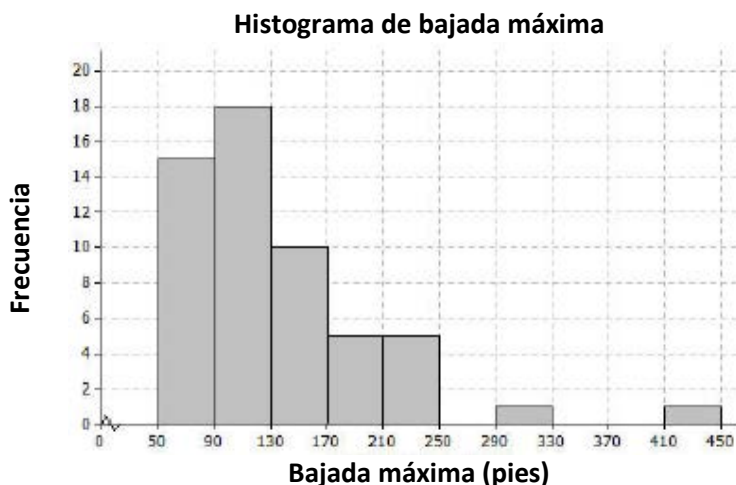
Ejemplo 3: Forma del histograma

Un histograma es útil para describir la forma de la distribución de datos. Es importante pensar acerca de la forma de una distribución de datos porque, según la forma que tenga, habrá diferentes maneras de describir características importantes de la distribución, tales como el centro y la variabilidad.

Un grupo de estudiantes quería averiguar cuánto duraba una marca determinada de baterías AA. El siguiente histograma muestra la distribución de datos de la duración (en horas) de algunas baterías AA. Mirando la forma del histograma, observa cómo los datos forman un "montículo" alrededor de un centro de aproximadamente 105. Describiríamos esta forma como forma de "montículo" o simétrica. Si tuviéramos que dibujar una línea por el centro, observa cómo cada lado del histograma es aproximadamente la misma imagen o una imagen espejo de la otra. Esto significa que el gráfico es aproximadamente simétrico.

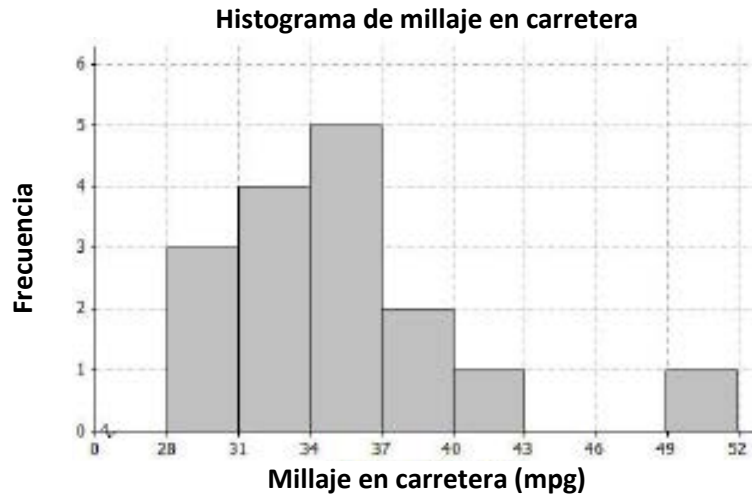


Otro grupo de estudiantes quería investigar la longitud de bajada máxima de las montañas rusas. El siguiente histograma muestra la bajada máxima (en pies) de un grupo seleccionado de montañas rusas. Este histograma tiene una forma asimétrica. La mayoría de los datos están en intervalos de 50 a 170 pies. Pero hay dos valores que son inusuales (o no típicos) cuando se los compara con los datos restantes. Estos valores son mucho más altos que la mayoría de los datos.



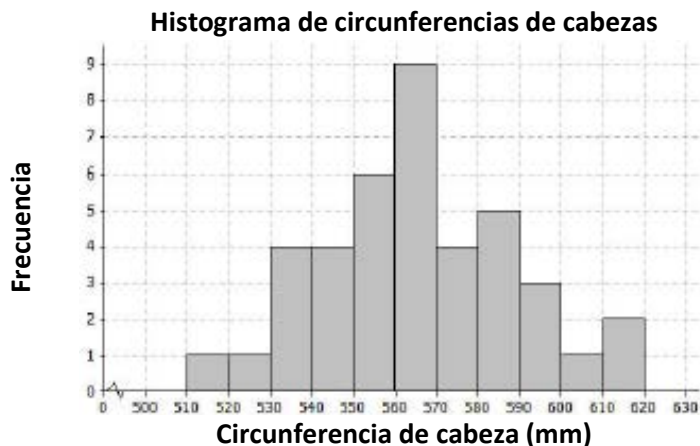
Ejercicios 10 a 12

10. El siguiente histograma muestra las millas por galón de diferentes automóviles compactos en carretera.



- Describe la forma del histograma como aproximadamente simétrica, asimétrica a la izquierda o asimétrica a la derecha.
 - Dibuja una línea vertical en el histograma para mostrar dónde estaría la cantidad típica de millas por galón de un automóvil compacto.
 - ¿Qué te indica la forma del histograma sobre las millas por galón de automóviles compactos?
11. Describe la forma del histograma de circunferencias de cabezas que completaste en el ejercicio 5 como aproximadamente simétrica, asimétrica a la izquierda o asimétrica a la derecha.

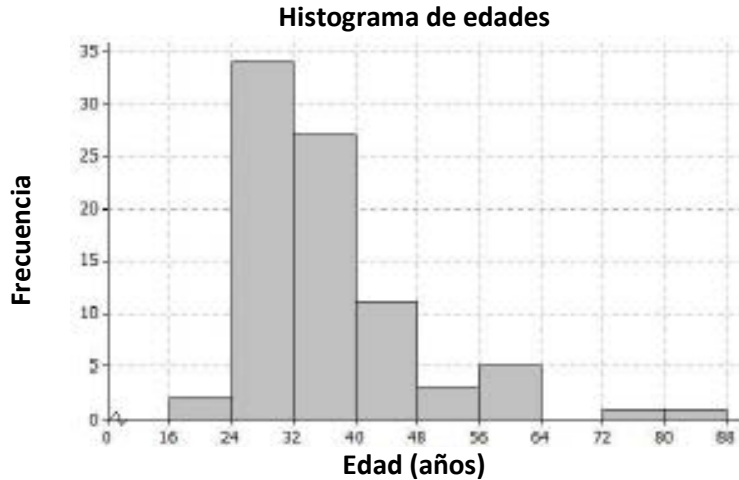
12. Otro estudiante decidió organizar los datos de las circunferencias de cabezas cambiando el ancho de cada intervalo para que sea 10 en lugar de 20. A continuación, se encuentra el histograma que hizo el estudiante.



- ¿Qué diferencia hay entre este histograma y el histograma de circunferencias de cabezas que completaste en el ejercicio 5?
- Describe la forma de este nuevo histograma como aproximadamente simétrica, asimétrica a la izquierda o asimétrica a la derecha.
- ¿Cuántas circunferencias de cabezas hay en el intervalo de 570 a 590 mm?
- ¿En qué intervalo se incluiría una circunferencia de cabeza de 571 mm? ¿En qué intervalo se incluiría una circunferencia de cabeza de 610 mm?

Conjunto de problemas

- El siguiente histograma muestra las edades de las actrices cuyas actuaciones han ganado la categoría de Mejor actriz principal en los Premios anuales de la Academia (es decir, los Óscar).



- ¿Qué intervalo de edades contiene a la mayor cantidad de actrices? ¿Cuántas actrices están representadas en ese intervalo?
 - Describe la forma del histograma.
 - ¿Qué te indica la forma sobre las edades de las actrices que ganaron el Óscar a la mejor actriz?
 - ¿Qué intervalo describe el centro de las edades de las actrices?
 - ¿En qué intervalo se incluiría la edad de 72?
- La siguiente tabla de frecuencias muestra la capacidad de asientos de los estadios de los equipos de baloncesto de la NBA.

Cantidad de asientos	Conteo	Frecuencia
17.000 - < 17.500		2
17.500 - < 18.000		1
18.000 - < 18.500	+++	6
18.500 - < 19.000	+++	5
19.000 - < 19.500	+++	5
19.500 - < 20.000	+++	5
20.000 - < 20.500		2
20.500 - < 21.000		2
21.000 - < 21.500		0
21.500 - < 22.000		0
22.000 - < 22.500		1

- Dibuja un histograma de la cantidad de asientos de los estadios de la NBA. Utiliza los histogramas que has visto a lo largo de esta lección para ayudarte a crear tu histograma.
- ¿Cuál es el ancho de cada intervalo? ¿Cómo lo sabes?

- c. Describe la forma del histograma.
 d. ¿Qué intervalo describe el centro de la cantidad de asientos?
3. Aquí se enumeran los gramos de carbohidratos de las hamburguesas de algunos restaurantes de comida rápida.

33 40 66 45 28 30 52 40 26 42
 42 44 33 44 45 32 45 45 52 24

- a. Completa la tabla de frecuencias con intervalos de un ancho de 5.

Cantidad de carbohidratos (gramos)	Conteo	Frecuencia
20 - < 25		
25 - < 30		
30 - < 35		
35 - < 40		
40 - < 45		
45 - < 50		
50 - < 55		
55 - < 60		
60 - < 65		
65 - < 70		

- b. Dibuja un histograma de los datos sobre los carbohidratos.
 c. Describe el centro y la forma del histograma.
 d. En la siguiente tabla de frecuencias, los intervalos están cambiados. Utilizando los datos de carbohidratos anteriores, completa la tabla de frecuencias con intervalos de un ancho de 10.

Cantidad de carbohidratos (gramos)	Conteo	Frecuencia
20 - < 30		
30 - < 40		
40 - < 50		
50 - < 60		
60 - < 70		

- e. Dibuja un histograma.
4. Utiliza los histogramas que creaste en las partes (b) y (e) del ejercicio 3 para responder las siguientes preguntas.
- a. ¿Por qué hay menos barras en el histograma de la parte (e) que en el histograma de la parte (b)?
 b. ¿Cambió la forma del histograma de la parte (e) con respecto a la forma del histograma de la parte (b)?
 c. ¿Tu estimación del centro cambió del histograma de la parte (b) al histograma de la parte (e)?

Lección 5: Descripción de una distribución representada en un histograma

Trabajo en clase

Ejemplo 1: Tabla de frecuencias relativas

En la Lección 4, investigamos las circunferencias de cabezas que recabaron los equipos de baloncesto de niños y de niñas. A continuación, se encuentra la tabla de frecuencias de las circunferencias de cabezas que midieron.

Tamaños de gorras	Intervalo de circunferencias de cabezas (mm)	Conteo	Frecuencia
XS	510 - < 530		2
S	530 - < 550	+++	8
M	550 - < 570	+++++	15
L	570 - < 590	+++	9
XL	590 - < 610		4
XXL	610 - < 630		2
			Total: 40

Isabel, una de las jugadoras de baloncesto, indicó que la mayoría de las gorras eran pequeñas, medianas o grandes. Para decidir si Isabel tenía razón, los jugadores agregaron una columna de frecuencia relativa a la tabla.

Frecuencia relativa es el valor de la frecuencia en un intervalo dividido por la cantidad total de valores de datos.

Ejercicios 1 a 4

1. Completa la columna de frecuencias relativas de la siguiente tabla.

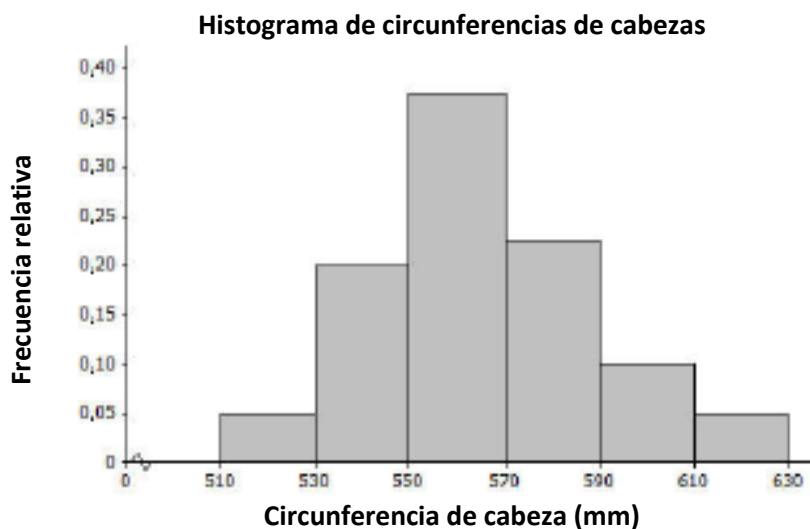
Tamaños de gorras	Intervalo de circunferencias de cabezas (mm)	Conteo	Frecuencia	Frecuencia relativa
XS	510 - < 530		2	$\frac{2}{40} = 0,05$
S	530 - < 550	+++	8	$\frac{8}{40} = 0,20$
M	550 - < 570	+++++	15	
L	570 - < 590	+++	9	
XL	590 - < 610		4	
XXL	610 - < 630		2	
			Total: 40	

2. ¿Cuál es el total de la columna de frecuencias relativas?
3. ¿Qué intervalo tiene la mayor frecuencia relativa? ¿Cuál es el valor?
4. ¿Qué porcentaje de las circunferencias de cabezas se encuentra entre 530 y 589 mm? Muestra cómo obtuviste la respuesta.

Ejemplo 2: Histograma de frecuencias relativas

Los jugadores decidieron crear un histograma utilizando las frecuencias relativas en lugar de las frecuencias. Observaron que las frecuencias relativas de la tabla variaban de casi 0 a casi 0.40. Dibujaron una línea numérica y marcaron los intervalos sobre esa línea. Luego, dibujaron la línea vertical y la rotularon "Frecuencia relativa". Agregaron una escala a esta línea a partir de 0, con aumentos cada 0.05, hasta llegar a 0.40.

Completaron el histograma dibujando las barras de manera que la altura de cada barra coincidiera con la frecuencia relativa de ese intervalo. Aquí está el histograma de frecuencias relativas terminado:



Ejercicios 5 y 6

- 5.
- Describe la forma del histograma de frecuencias relativas de las circunferencias de cabezas del ejemplo 2.
 - ¿Qué diferencia hay entre la forma de este histograma y el histograma de frecuencias que dibujaste en el ejercicio 5 de la Lección 4?
 - Isabel dijo que la mayoría de las gorras que se necesitaban pedir eran pequeñas, medianas y grandes. ¿Tenía razón? ¿Qué porcentaje de las gorras que se tienen que pedir son pequeñas, medianas o grandes?
6. Aquí está la tabla de frecuencias de la capacidad de asientos de los estadios de los equipos de baloncesto de la NBA.

Cantidad de asientos	Conteo	Frecuencia	Frecuencia relativa
17.000 - < 17.500		2	
17.500 - < 18.000		1	
18.000 - < 18.500	+++	6	
18.500 - < 19.000	+++	5	
19.000 - < 19.500	+++	5	
19.500 - < 20.000	+++	5	
20.000 - < 20.500		2	
20.500 - < 21.000		2	
21.000 - < 21.500		0	
21.500 - < 22.000		0	
22.000 - < 22.500		1	

- ¿Cuál es la cantidad total de estadios de la NBA?
- Completa la columna de frecuencia relativa. Redondea a la milésima más cercana.

- c. Crea un histograma de frecuencias relativas. Redondea a la milésima más cercana.
- d. Describe la forma del histograma de frecuencias relativas.
- e. ¿Qué porcentaje de los estadios tiene una capacidad de entre 18,500 y 19,999 asientos?
- f. ¿Qué diferencia hay entre este histograma de frecuencias relativas y el histograma de frecuencias que dibujaste en el problema 2 del Conjunto de problemas de la Lección 4?

Resumen de la lección

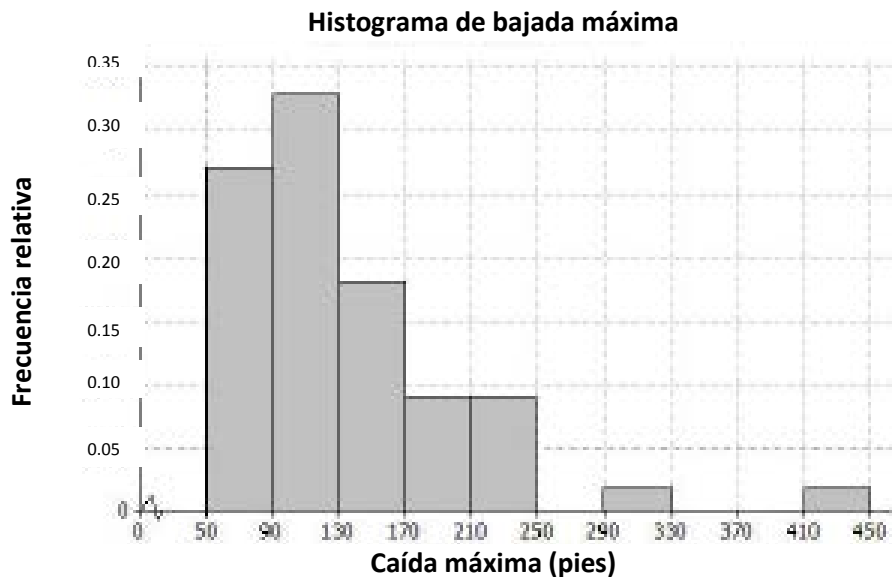
Un **histograma de frecuencias relativas** utiliza los mismos datos que un histograma de frecuencias, pero compara las frecuencias de cada intervalo con la cantidad total de elementos. Por ejemplo, si el primer intervalo contiene 8 del total de 32 elementos, la frecuencia relativa del primer intervalo es $\frac{8}{32} = \frac{1}{4} = 0.25$.

Resumen de la lección

Un **histograma de frecuencias relativas** utiliza los mismos datos que un histograma de frecuencias, pero compara las frecuencias de cada intervalo con la cantidad total de elementos. Por ejemplo, si el primer intervalo contiene 8 del total de 32 elementos, la frecuencia relativa del primer intervalo es $\frac{8}{32} = \frac{1}{4} = 0.25$.

Conjunto de problemas

1. A continuación, se encuentra un histograma de frecuencias relativas de la bajada máxima (en pies) de un grupo seleccionado de montañas rusas.



- a. Describe la forma del histograma de frecuencias relativas.

- b. ¿Qué te indica la forma sobre la bajada máxima (en pies) de las montañas rusas?
- c. Jerome dijo que más de la mitad de los datos se encuentran en el intervalo que va desde 50 hasta 130 pies. ¿Estás de acuerdo con Jerome? ¿Por qué sí o por qué no?

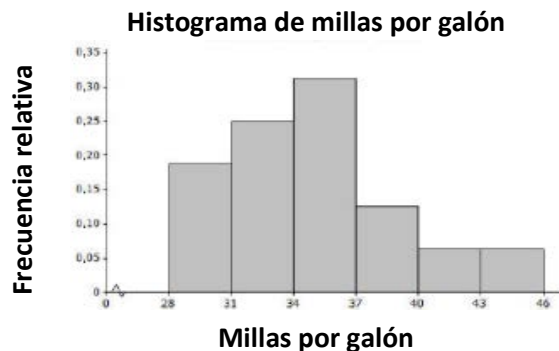
2. La siguiente tabla de frecuencias muestra la duración de las películas seleccionadas que se exhibieron en un cine local durante los últimos 6 meses.

Duración de películas (min)	Conteo	Frecuencia	Frecuencia relativa
80 - < 90		1	
90 - < 100		4	
100 - < 110	+++	7	
110 - < 120	+++	5	
120 - < 130	+++	7	
130 - < 140		3	
140 - < 150		1	

- Completa la columna de frecuencias relativas. Redondea a la milésima más cercana.
 - ¿Qué porcentaje de las duraciones de películas es mayor o igual que 130 minutos?
 - Dibuja un histograma de frecuencias relativas.
 - Describe la forma del histograma de frecuencias relativas.
 - ¿Qué te indica la forma sobre la duración de las películas?
3. La siguiente tabla muestra las millas por galón de diferentes automóviles compactos en carretera.

Millaje	Conteo	Frecuencia	Frecuencia relativa
128 - < 31		3	
31 - < 34		4	
34 - < 37	+++	5	
37 - < 40		2	
40 - < 43		1	
43 - < 46		0	
46 - < 49		0	
49 - < 52		1	

- ¿Cuál es la cantidad total de automóviles compactos?
- Completa la columna de frecuencias relativas. Redondea a la milésima más cercana.
- ¿Qué porcentaje de automóviles recorre entre 31 y 37 millas (no inclusive) por galón en carretera?
- Juan dibujó el histograma de frecuencias relativas de las millas por galón de los automóviles compactos, que se muestra a la derecha. ¿Estás de acuerdo con la manera en la que Juan dibujó el histograma? Explica tu respuesta.



Lección 6: Descripción del centro de una distribución mediante el uso de la media

Trabajo en clase

Ejemplo 1

Recuerda que en la Lección 3, Robert, un estudiante de sexto grado de la escuela media Roosevelt, investigó la cantidad de horas que dormían los estudiantes de sexto grado durante la semana escolar. Hoy debe hacer un breve informe sobre su investigación para la clase. Aquí está su informe.

"Encuesté a veintinueve estudiantes de sexto grado y les pregunté cuántas horas solían dormir por noche durante la semana escolar. Lo primero que tuve que hacer fue organizar los datos. Hice esto con un diagrama de puntos".

Diagrama de puntos de la cantidad de horas de sueño



Michelle es una compañera de clase de Robert. A ella le gustó su informe, pero opina de forma muy diferente sobre la manera de determinar el centro de la cantidad de horas de sueño. Su idea es emparejar los datos a fin de determinar un valor típico o central.

Ejercicios 1 a 6

Supón que Michelle les pregunta a diez de sus compañeros de clase la cantidad de horas que suelen dormir durante la semana escolar y que ellos responden (en horas): 8 10 8 8 11 11 9 8 10 7.

1. ¿Cómo piensas que Robert organizaría sus datos? ¿Cuál piensas que Robert diría que es el centro de estos diez puntos de datos? ¿Por qué?

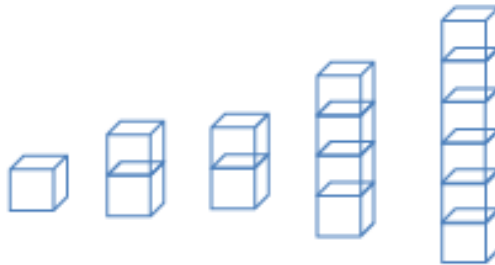
2. ¿Crees que su valor es una buena medida para utilizar para el centro del conjunto de datos de Michelle? ¿Por qué sí o por qué no?

El centro de Michelle se llama *media*. Ella averiguó que la cantidad total de horas de sueño de los diez estudiantes era de 90 horas. Tenía 90 cubos Unifix (cubos conectables) y le dio a cada uno de los diez estudiantes la cantidad de cubos que equivalía a la cantidad de horas de sueño que había dicho. Luego, le pidió a cada uno de los diez estudiantes que conectara sus cubos en una columna y que colocara su columna sobre una mesa para compararla. Después, hizo que compartieran los cubos entre ellos hasta que tuvieran la misma cantidad de cubos en sus columnas cuando hubieran terminado de compartirlos.

3. Haz diez columnas de cubos que representen la cantidad de horas de sueño de cada uno de los diez estudiantes. Con el método de Michelle, ¿cuántos cubos hay en cada una de las diez columnas cuando terminan de compartirlos?
4. Teniendo en cuenta que un cubo representa una hora de sueño, interpreta tu respuesta al ejercicio 3 en términos de cantidad de horas de sueño. ¿Qué representa esta cantidad de cubos en cada columna? ¿Cómo se llama este valor?
5. Supón que el estudiante que le dijo a Michelle que dormía 7 horas cambia su entrada de datos a 8 horas. ¿Qué genera ahora el procedimiento de Michelle para el centro del nuevo conjunto de datos? ¿Qué tuviste que hacer con el cubo adicional para que funcionara el procedimiento de Michelle?
6. Interpreta el procedimiento de distribución justa de Michelle a través de una fórmula matemática por medio de la que se averigüe el valor justo sin utilizar cubos. Asegúrate de que puedas explicar claramente cómo se relacionan entre sí el procedimiento de distribución justa y la fórmula matemática.

Ejemplo 2

Supón que Robert les preguntó a cinco estudiantes de sexto grado cuántas mascotas tenía cada uno. Sus respuestas fueron 2, 6, 2, 4 y 1. Robert mostró los datos con cubos de la siguiente manera:



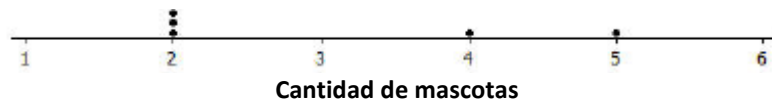
Observa que un estudiante tiene una mascota, dos estudiantes tienen dos mascotas cada uno, un estudiante tiene cuatro y un estudiante tiene seis. Robert también representó el conjunto de datos en el siguiente diagrama de puntos.

Diagrama de puntos de la cantidad de mascotas

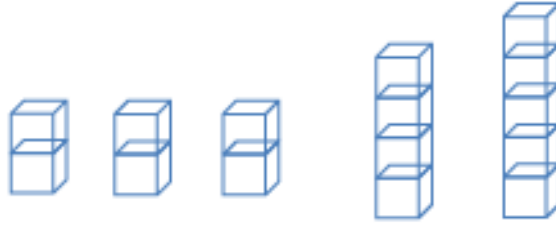


Robert quería ilustrar el método de distribución justa de Michelle utilizando diagramas de puntos. Dibujó el siguiente diagrama de puntos y dijo que este representaba el resultado que se obtenía cuando el estudiante que tenía seis mascotas compartía una de ellas con el estudiante que tenía una.

Diagrama de puntos de la cantidad de mascotas



Robert también representó los datos con cubos como se muestra a continuación.



Ejercicios 7 a 10

Ahora continúa distribuyendo las mascotas según los siguientes pasos.

7. Robert realiza una distribución justa al hacer que el estudiante que tiene cinco mascotas comparta una de ellas con uno de los estudiantes que tiene dos mascotas.

a. Dibuja la representación de cubos que muestre la distribución justa de Robert.

b. Dibuja el diagrama de puntos que muestre la distribución justa de Robert.

8. Robert realiza otra distribución justa al hacer que uno de los estudiantes que tiene cuatro mascotas comparta una de ellas con uno de los estudiantes que tiene dos mascotas.

a. Dibuja la representación de cubos que muestre la distribución justa de Robert.

- b. Dibuja el diagrama de puntos que muestre la distribución justa de Robert.
9. Robert realiza una última distribución justa al hacer que el estudiante que tiene cuatro mascotas comparta una de ellas con el estudiante que tiene dos mascotas.
- a. Dibuja la representación de cubos que muestre la última distribución justa de Robert.
- b. Dibuja la representación del diagrama de puntos que muestre la última distribución justa de Robert.
10. Explica con tus propias palabras por qué las representaciones finales en las que se utilizan cubos y un diagrama de puntos muestran que la cantidad media de mascotas de los cinco estudiantes es 3 mascotas.

Conjunto de problemas

1. Se llevó a cabo un juego en el que se lanzaban pelotas de tenis dentro de una canasta desde una determinada distancia. La cantidad de lanzamientos acertados de seis estudiantes fue 4, 1, 3, 2, 1 y 7.
 - a. Dibuja una representación de los datos utilizando cubos, en la que un cubo represente un lanzamiento acertados de una pelota de tenis dentro de la canasta.
 - b. Dibuja el conjunto de datos original utilizando un diagrama de puntos.

2. Encuentra la cantidad media de lanzamientos acertados de este conjunto de datos con el método de distribución justa de Michelle. Para cada paso, muestra la representación de cubos y el diagrama de puntos correspondiente. Explica cada paso con palabras en el contexto del problema. Puedes mover más de un lanzamiento acertado en un paso, pero asegúrate de que tu explicación sea clara. Debes mostrar dos o más pasos.

Paso descrito en palabras	Representación de distribución justa de cubos	Diagrama de puntos

3. La cantidad de bolsillos de las prendas que hoy utilizan cuatro estudiantes para ir a la escuela es 4, 1, 3 y 6. Paige realiza la siguiente representación de cubos a medida que lleva a cabo el proceso de distribución justa. Ayúdala a decidir cómo terminar el proceso de 3, 3, 3 y 5 cubos.



4. Supón que la cantidad media de pedacitos de chocolate en 30 galletas es de 14 pedacitos.
 - a. Interpreta la cantidad media de pedacitos de chocolate en términos de una distribución justa.
 - b. Describe la representación del diagrama de puntos de la media de la distribución justa de 14 pedacitos de chocolate en 30 galletas.

5. Supón que las siguientes son longitudes (en milímetros) de plántulas de rábano que crecen en idénticas condiciones durante tres días: 12 11 12 14 13 9 13 11 13 10 10 14 16 13 11.
 - a. Encuentra la longitud media de estas 15 plántulas de rábano.
 - b. Interpreta el valor de la parte (a) en términos de la longitud del centro de la distribución justa.

Lección 7: La media como punto de equilibrio

Trabajo en clase

En la Lección 3, Robert nos dio una interpretación informal del centro de un conjunto de datos. En la Lección 6, Michelle desarrolló una interpretación más formal del centro como la media de una distribución justa, un valor que toda persona del conjunto de datos tendría si todos tuvieran el mismo valor. En esta lección, Sabina nos mostrará cómo interpretar la media como punto de equilibrio.

Ejemplo 1: La media como punto de equilibrio

Sabina quiere saber cuánto tiempo les toma a los estudiantes llegar a la escuela. Les pregunta a dos estudiantes cuánto tiempo les toma hacerlo. A un estudiante le toma 1 minuto, y al otro estudiante le toma 11 minutos. Sabina representa estos datos sobre una regla, y coloca un centavo en 1 y otro en 11. Ella piensa que podría haber una conexión entre la media de dos puntos de datos y el lugar en donde alcanzan un equilibrio sobre una regla; cree que la media puede ser el punto de equilibrio. Sabina muestra el resultado en un diagrama de puntos.

Diagrama de puntos de la cantidad de minutos



Sabina decide mover el centavo de 1 a 4 y el otro centavo de 11 a 8 sobre la regla, y advierte que el movimiento de los dos centavos representa la misma distancia, pero en direcciones opuestas. Ella cree que si los puntos de datos se mueven la misma distancia, pero en direcciones opuestas, el punto de equilibrio sobre la regla no cambia. ¿Estás de acuerdo con Sabina?

Sabina sigue adelante moviendo el centavo de 4 a 6. Para mantener la regla equilibrada en 6, ¿cuánto debería mover Sabina el centavo desde 8 y en qué dirección?

Ejercicios 1 y 2

Ahora es tu turno para intentar equilibrar dos centavos sobre una regla.

1. Pega con cinta adhesiva un centavo en 2.5 sobre tu regla.
 - a. ¿Dónde se debería pegar el otro centavo para que la regla se equilibre en 6?

 - b. ¿Qué tan lejos de 6 se encuentra el centavo en 2.5? ¿Qué tan lejos de 6 se encuentra el otro centavo?

 - c. ¿Es 6 la media de las dos ubicaciones de los centavos?

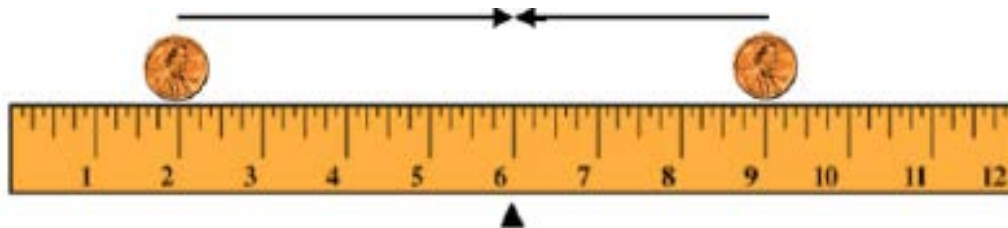
2. Mueve el centavo que está en 2.5 dos pulgadas hacia la derecha.
 - a. ¿Dónde colocarás el centavo?

 - b. ¿Qué tienes que hacer con el otro punto de datos (el otro centavo) para mantener el punto de equilibrio en 6?

 - c. ¿Cuál es la media de los dos puntos de datos nuevos? ¿Es este el mismo valor que el punto de equilibrio de la regla?

Ejemplo 2: Equilibrio de más de dos puntos

Sabina quiere saber qué ocurre si hay más de dos puntos de datos. Supón que hay tres estudiantes. Un estudiante vive a 2 minutos de la escuela, y otro vive a 9 minutos. Si el tiempo medio de los tres estudiantes es de 6 minutos, ella se pregunta cuánto le tomará al tercer estudiante llegar a la escuela. Utilizando lo que sabes sobre distancias con respecto a la media, ¿dónde debería colocarse el tercer centavo para que la media sea 6? Marca el diagrama y explica tu razonamiento.

**Ejercicios 3 a 6**

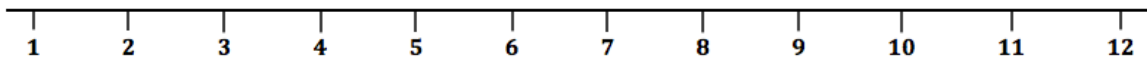
Imagina que estás equilibrando centavos sobre una regla.

3. Supón que colocas un centavo en 3, en 7 y en 8 sobre tu regla.
 - a. Haz un dibujo de la regla. ¿En qué valor piensas que se equilibrará? Marca el punto de equilibrio con el símbolo Δ .



- b. ¿Cuál es la media de 3, 7 y 8? ¿Tu regla se equilibra en la media?

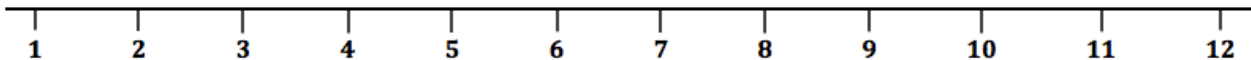
- c. Muestra la información de la parte (a) en un diagrama de puntos. Marca el punto de equilibrio con el símbolo Δ .



- d. ¿Cuáles son las desviaciones a cada lado del punto de equilibrio? ¿Cómo demuestra esto que la media es 6?

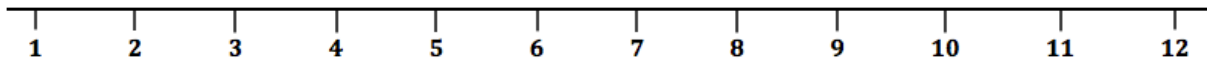
4. Ahora supón que colocas un centavo en 7 y en 9 sobre tu regla.

- a. Haz un diagrama de puntos que represente estos dos centavos.



- b. Estima dónde colocar un tercer centavo sobre tu regla para que esta se equilibre en 6, y marca el punto en el diagrama de puntos anterior. Marca el punto de equilibrio con el símbolo Δ .
- c. Explica por qué tu respuesta de la parte (b) es verdadera calculando las desviaciones de los puntos con respecto a 6. ¿Son iguales las desviaciones totales a cada lado de la media?

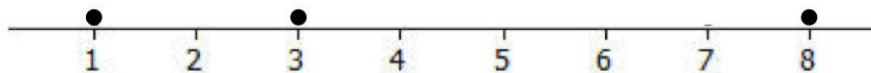
5. ¿Es verdadero el concepto de la media como punto de equilibrio si colocas varios centavos en un único lugar sobre la regla?
6. Supón que colocas dos centavos en 7 y un centavo en 9 sobre tu regla.
- Haz un diagrama de puntos que represente estos tres centavos.



- Estima dónde colocar un cuarto centavo sobre tu regla para que esta se equilibre en 6, y marca el punto en el diagrama de puntos anterior. Marca el punto de equilibrio con el símbolo Δ .
- Explica por qué tu respuesta de la parte (b) es verdadera calculando las desviaciones de los puntos con respecto a 6. ¿Son iguales las desviaciones totales a cada lado de la media?

Ejemplo 3: Averiguación de la media

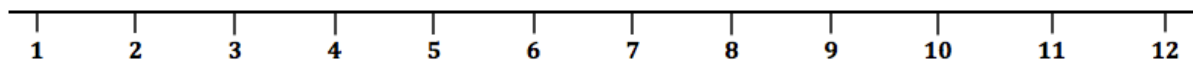
¿Qué ocurre si los datos de un diagrama de puntos fueran 1, 3 y 8? ¿Los datos se equilibrarían en 6? Si no es así, ¿cuál es el punto de equilibrio? ¿Por qué?



Ejercicio 7

Utiliza lo que has aprendido sobre la media para responder las siguientes preguntas.

7. Recuerda que, en la Lección 6, Michelle les preguntó a diez de sus compañeros de clase la cantidad de horas que suelen dormir durante la semana escolar. Sus respuestas (en horas) fueron 8, 10, 8, 8, 11, 11, 9, 8, 10 y 7.
 - a. Es difícil equilibrar diez monedas de un centavo. En lugar de utilizar realmente centavos y una regla, haz un diagrama de puntos que represente el conjunto de datos.



- b. Utiliza tu diagrama de puntos para encontrar el punto de equilibrio.

Conjunto de problemas

1. La cantidad de bolsillos de las prendas que hoy utilizan cuatro estudiantes para ir a la escuela es 4, 1, 3 y 4.
 - a. Realiza el proceso de distribución justa para encontrar la cantidad media de bolsillos de estos cuatro estudiantes. Dibuja las representaciones de cubos para cada paso del proceso.
 - b. Encuentra las desviaciones totales a cada lado de la media para demostrar que la media que encontraste en la parte (a) es correcta.

2. Los tiempos (redondeados al minuto más cercano) que tardaron seis compañeros de clase en correr una milla fueron 7, 9, 10, 11, 11 y 12 minutos.
 - a. Dibuja una representación del diagrama de puntos de los tiempos que tardaron en correr la milla.
 - b. Supón que Sabina cree que la media es de 11 minutos. ¿Tiene razón? Explica tu respuesta.
 - c. ¿Cuál es la media?

3. En el siguiente diagrama de puntos, se muestran los precios por galón de gasolina (en centavos) de cinco gasolineras de la ciudad en un día. Falta el precio de una sexta gasolinera, pero se informó que el precio medio de las seis gasolineras era de 380 centavos por galón. Utiliza el proceso de equilibrio para determinar el precio de un galón de gasolina en la sexta gasolinera.

Diagrama de puntos de precio (centavos por galón)



4. La cantidad de teléfonos (fijos y celulares) que tienen los miembros de cada familia en un grupo de nueve familias es 3, 5, 6, 6, 6, 6, 7, 7 y 8.
 - a. Utiliza la fórmula matemática para la media (determinar la suma de los puntos de datos y dividir por la cantidad de puntos de datos) a fin de encontrar la cantidad media de teléfonos que tienen estas nueve familias.
 - b. Haz un diagrama de puntos de los datos y verifica tu respuesta de la parte (a) utilizando el proceso de equilibrio.

Lección 8: Variabilidad en una distribución de datos

Trabajo en clase

Ejemplo 1: Comparación de dos distribuciones

La familia de Robert está planificando mudarse a la ciudad de Nueva York o a San Francisco. Robert tiene una prima en San Francisco a la que le preguntó si le gustaba vivir en un clima cálido. Ella le respondió que San Francisco no es muy cálido. Él se sorprendió y, como la temperatura era uno de los criterios que iba a utilizar para decidir dónde mudarse, decidió investigar las distribuciones de las temperaturas de la ciudad de Nueva York y de San Francisco. La siguiente tabla proporciona las temperaturas promedio (en grados Fahrenheit) de las dos ciudades por mes.

Ciudad	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Ciudad de Nueva York	39	42	50	61	71	81	85	84	76	65	55	47
San Francisco	57	60	62	63	64	67	67	68	70	69	63	58

Ejercicios 1 y 2

Utiliza la tabla anterior para responder lo siguiente:

- Calcula la temperatura media mensual por año de cada ciudad.
- Recuerda que Robert está intentando decidir a qué ciudad quiere mudarse. ¿Qué le aconsejarías de acuerdo con la comparación de todas las temperaturas medias mensuales por año de las dos ciudades?

Ejemplo 2: Comprensión de la variabilidad

Quizá Robert debería mirar qué tan dispersos están los datos de las temperaturas mensuales de la ciudad de Nueva York con respecto a su media y qué tanto lo están los de San Francisco con respecto a su media. Para comparar la variabilidad de las temperaturas mensuales entre las dos ciudades, puede ser útil mirar diagramas de puntos. A continuación, se encuentran los diagramas de puntos de las distribuciones de las temperaturas mensuales de la ciudad de Nueva York y de San Francisco.

Diagrama de puntos de la temperatura de la ciudad de Nueva York

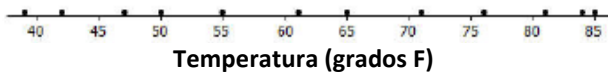
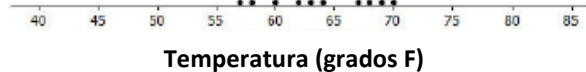


Diagrama de puntos de la temperatura de San Francisco

**Ejercicios 3 a 7**

Utiliza los diagramas de puntos anteriores para responder lo siguiente:

3. Marca la ubicación de la media en cada distribución con el símbolo de equilibrio Δ . ¿Qué diferencia hay entre las dos distribuciones según sus medias?
4. Describe la variabilidad de las temperaturas mensuales de la ciudad de Nueva York con respecto a la media de sus temperaturas.
5. Describe la variabilidad de las temperaturas mensuales de San Francisco con respecto a su media.

6. Compara la cantidad de variabilidad en las dos distribuciones. ¿La variabilidad es aproximadamente la misma o es diferente? Si es diferente, ¿qué distribución de temperaturas mensuales tiene más variabilidad? Explica tu respuesta.

7. Si Robert prefiere elegir la ciudad en donde las temperaturas varían menos de mes a mes, ¿qué ciudad debería elegir? Explica tu respuesta.

Ejemplo 3: Uso de la media y de la variabilidad en una distribución de datos

La media se utiliza para describir el valor típico para toda la distribución. Sabina le pregunta a Robert qué ciudad cree que tiene el mejor clima. ¿Cuál piensas que será la respuesta de Robert?

Sabina está confundida y le pide que le explique qué quiere decir con esta afirmación. ¿Cómo podría explicarle Robert a Sabina lo que quiere decir?

Ejercicios 8 a 14

Considera las siguientes dos distribuciones del tiempo que les toma a seis estudiantes llegar a la escuela por la mañana e ir a su casa desde la escuela por la tarde.

	Tiempo (minutos)					
Mañana	11	12	14	14	16	17
Tarde	6	10	13	18	18	19

8. Para visualizar las medias y la variabilidad, haz diagramas de puntos de cada una de las dos distribuciones.

Mañana**Tarde**

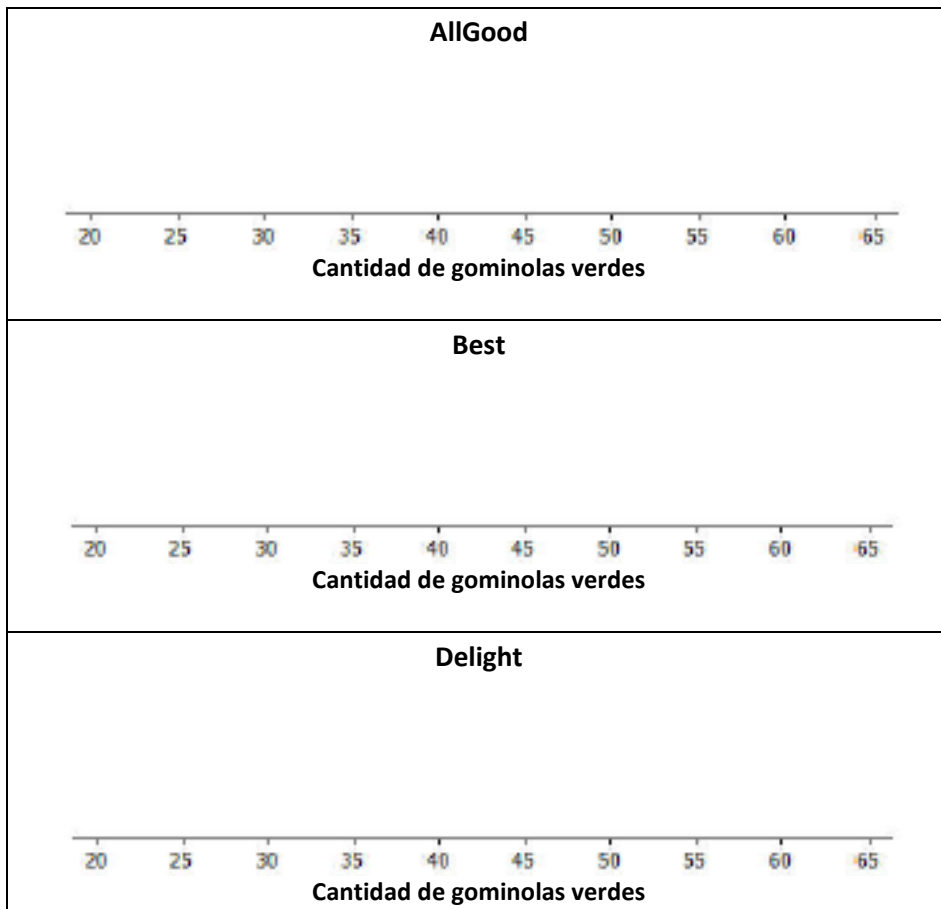
- ¿Cuál es el tiempo medio que les toma a estos seis estudiantes llegar desde su casa hasta la escuela por la mañana?
- ¿Cuál es el tiempo medio que les toma a estos seis estudiantes llegar desde la escuela hasta su casa por la tarde?
- ¿Para qué distribución la media proporciona un indicador más preciso de un valor típico? Explica tu respuesta.

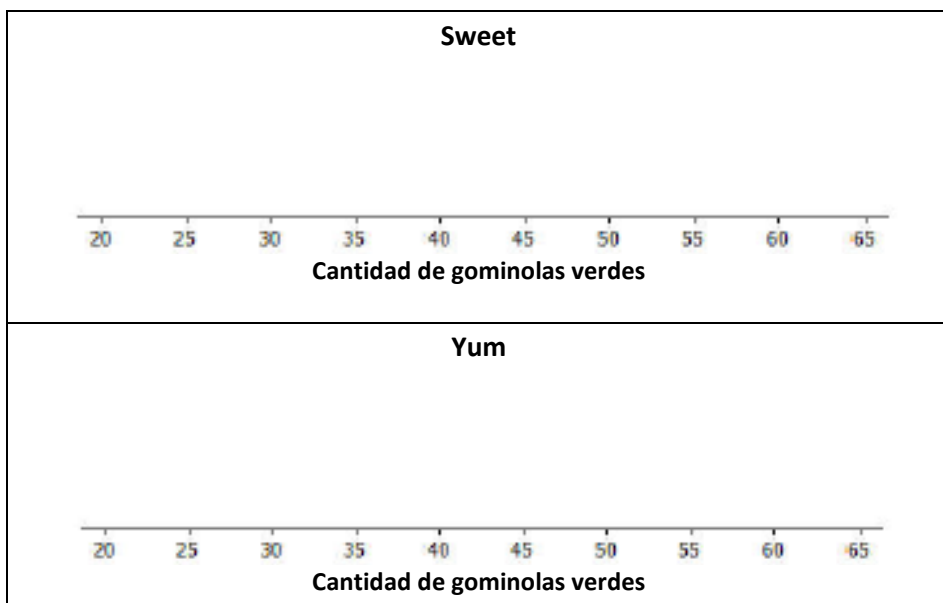
Las distribuciones se pueden ordenar según cuánto varíen los valores de los datos alrededor de sus medias.

Considera los siguientes datos sobre la cantidad de gominolas verdes que hay en siete bolsas de cinco fabricantes diferentes de dulces (AllGood, Best, Delight, Sweet y Yum). La media de cada distribución es de 42 gominolas verdes.

	Bolsa 1	Bolsa 2	Bolsa 3	Bolsa 4	Bolsa 5	Bolsa 6	Bolsa 7
AllGood	40	40	41	42	42	43	46
Best	22	31	36	42	48	53	62
Delight	26	36	40	43	47	50	52
Sweet	36	39	42	42	42	44	49
Yum	33	36	42	42	45	48	48

12. Haz un diagrama de puntos de la distribución de la cantidad de gominolas verdes para cada uno de los cinco fabricantes de dulces. Marca la ubicación de la media en cada distribución con el símbolo de equilibrio Δ .





13. Ordena los fabricantes de caramelos desde el que pienses que tenga la menor variabilidad hasta el que tenga la mayor variabilidad. Explica tu razonamiento para elegir el orden.
14. ¿Para qué compañía se consideraría que la media es un mejor indicador de un valor típico (según la menor variabilidad)?

Conjunto de problemas

- La cantidad de bolsillos de las prendas que utilizaron ayer siete estudiantes para ir a la escuela fue 4, 1, 3, 4, 2, 2 y 5. Hoy, esos siete estudiantes tienen cada uno tres bolsillos en sus prendas.
 - Haz un diagrama de puntos de las prendas que utilizaron ayer los estudiantes y otro diagrama de puntos de las que utilizaron hoy. Asegúrate de utilizar las mismas escalas. Muestra las medias usando el símbolo de equilibrio Δ .
 - Para cada distribución, encuentra la cantidad media de bolsillos que utilizan los siete estudiantes.
 - ¿Para qué distribución la cantidad media de bolsillos es un mejor indicador de lo que es típico? Explica tu respuesta.
- Se anotó la cantidad de minutos (redondeada al minuto más cercano) que le tomó a cada estudiante de un grupo de 5 estudiantes correr una determinada ruta pequeña a través del campo. Los datos resultantes fueron 9, 10, 11, 14 y 16 minutos. También se anotó la cantidad de minutos (redondeada al minuto más cercano) que les tomó a los cinco estudiantes correr una ruta diferente a través del campo. Se obtuvieron los siguientes datos: 6, 8, 12, 15 y 19 minutos.
 - Haz diagramas de puntos para las dos distribuciones del tiempo que tomó correr una ruta a través del campo. Asegúrate de utilizar la misma escala en ambos diagramas de puntos.
 - ¿Las distribuciones tienen la misma media? ¿Cuál es la media de cada diagrama de puntos?
 - ¿En qué distribución es la media un mejor indicador de la cantidad típica de tiempo que toma correr su ruta a través del campo? Explica tu respuesta.
- La siguiente tabla muestra los precios por galón de gasolina (en centavos) de cinco gasolineras de la ciudad que se anotaron el lunes, el miércoles y el viernes de una semana determinada.

Día	R&C	Gasolinera de Al	PB	Gasolinera de Sam	Gasolinera de Ann
Lunes	359	358	362	359	362
Miércoles	357	365	364	354	360
Viernes	350	350	360	370	370

- El precio medio por día de las cinco gasolineras es el mismo durante los tres días. Sin hacer ningún cálculo y simplemente mirando los precios del viernes, ¿cuál debe ser el precio medio?
- ¿En qué distribución diaria es la media un mejor indicador del precio típico por galón de las cinco gasolineras? Explica tu respuesta.

Lección 9: La desviación media absoluta (DMA) (Conocido también como MAD por sus siglas en inglés)

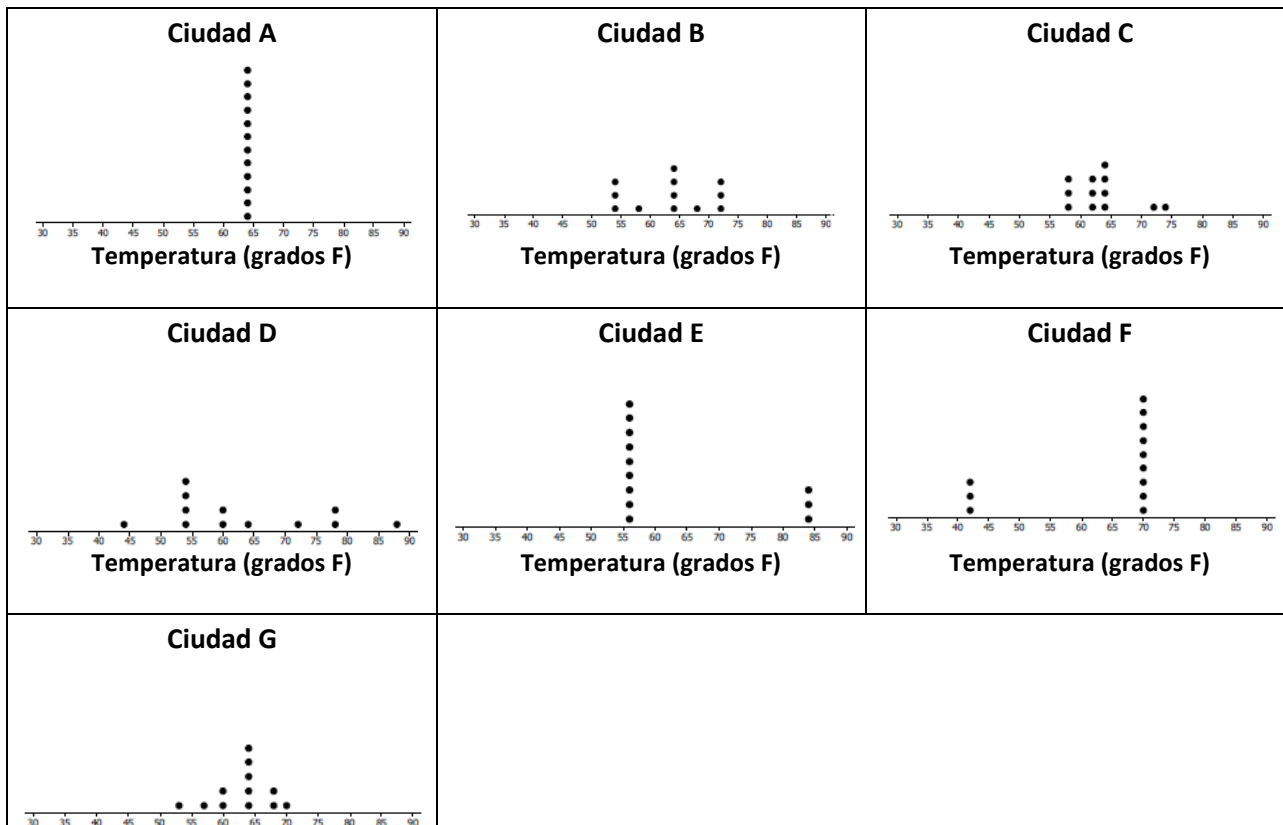
Trabajo en clase

Ejemplo 1: Variabilidad

En la Lección 8, Robert intentó decidir a cuál de las dos ciudades se mudaría según la comparación de sus temperaturas medias mensuales por año. Dado que la temperatura media mensual por año de la ciudad de Nueva York y la de San Francisco resultaron ser aproximadamente las mismas, decidió comparar las ciudades según la variabilidad de sus temperaturas mensuales con respecto a la media total. Miró las dos distribuciones y decidió que las temperaturas de la ciudad de Nueva York estaban más dispersas con respecto a su media que lo que lo estaban las temperaturas de San Francisco con respecto a su media.

Ejercicios 1 a 3

Las siguientes distribuciones de temperaturas de otras siete ciudades tienen todas una temperatura media mensual de aproximadamente 63 grados Fahrenheit. No tienen la misma variabilidad. Considera los siguientes diagramas de puntos de las temperaturas medias mensuales de las siete ciudades en grados Fahrenheit.



Temperatura (grados F)

1. ¿Qué distribución tiene la menor variabilidad de temperaturas con respecto a su media de 63 grados Fahrenheit? Explica tu respuesta.
2. ¿Qué distribución(es) parece(n) tener la mayor variabilidad de temperaturas con respecto a la media de 63 grados Fahrenheit? Explica tu respuesta.
3. Ordena las siete distribuciones desde la que tenga menor variabilidad hasta la que tenga mayor variabilidad. Explica por qué enumeraste las distribuciones en el orden que elegiste.

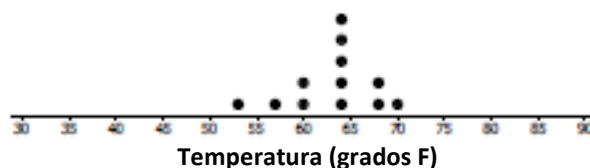
Ejemplo 2: Medición de la variabilidad

Con solo mirar las distribuciones, hay diferentes órdenes de variabilidad que parecen tener algo de sentido. A Sabina le interesa desarrollar una fórmula que genere un número que mida la variabilidad en una distribución de datos. Luego, aplicaría la fórmula a cada conjunto de datos y ordenaría las distribuciones de la menor a la mayor. Utilizar desviaciones para desarrollar una fórmula que mida la variabilidad es una buena idea para tener en cuenta.

Ejercicios 4 a 6

A continuación, se muestra el diagrama de puntos de las temperaturas mensuales de la ciudad G. Utiliza el diagrama de puntos y la temperatura media mensual de 63 grados Fahrenheit para responder las siguientes preguntas.

Ciudad G

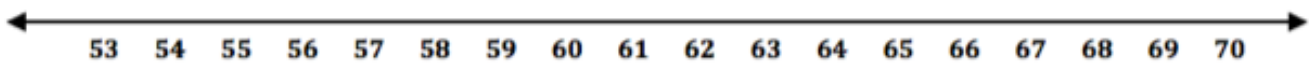


4. Completa la siguiente tabla de las desviaciones de las temperaturas de la ciudad G.

Temp.	Distancia	Desviación
53	53 - 63	10 a la izquierda
57		
60		
60		
64		
64		
64		
64		
64		
64		
68		
68		
70		
Suma		

5. ¿Cuál es la desviación total a la izquierda de la media? ¿Cuál es la desviación total a la derecha de la media?
6. Otra manera de representar en un gráfico las desviaciones es escribirlas en una línea numérica como se indica a continuación. Marca la media en la línea numérica. Coloca las desviaciones de la media en la línea numérica. ¿Cuál es la desviación total a la izquierda de la media? ¿Cuál es la desviación total a la derecha de la media?

Temperaturas de la ciudad G



Ejemplo 3: Averiguación de la desviación media absoluta (DMA) (Conocida también como MAD)

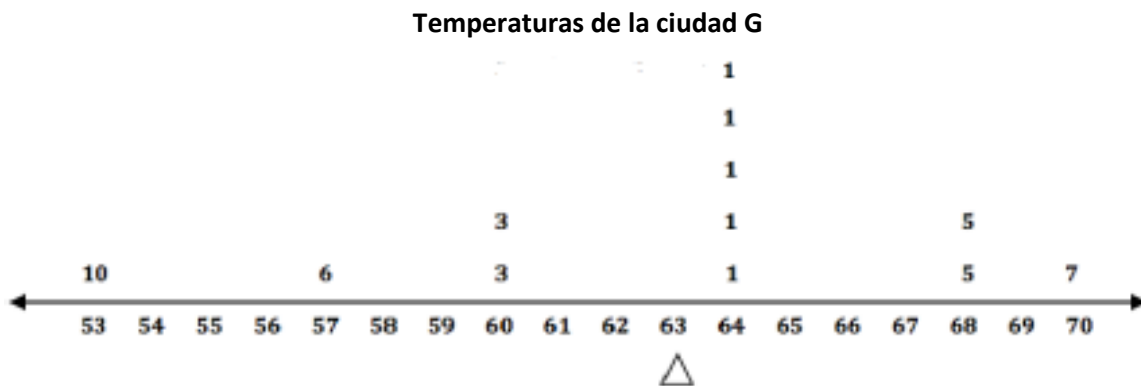
Con la interpretación del equilibrio de la media, las desviaciones totales a cada lado de la media siempre serán iguales. Sabina observa que hemos estado contando de manera consistente las distancias a la izquierda y a la derecha de la media, que es el punto de equilibrio. Se pregunta cómo puede utilizar las direcciones izquierda y derecha a fin de que esto la ayude a desarrollar una fórmula para encontrar la variabilidad.

Ejercicios 7 y 8

- 7. Una sugerencia posible para ayudar a Sabina es tomar el valor absoluto de las desviaciones.
 - a. Completa la siguiente tabla.

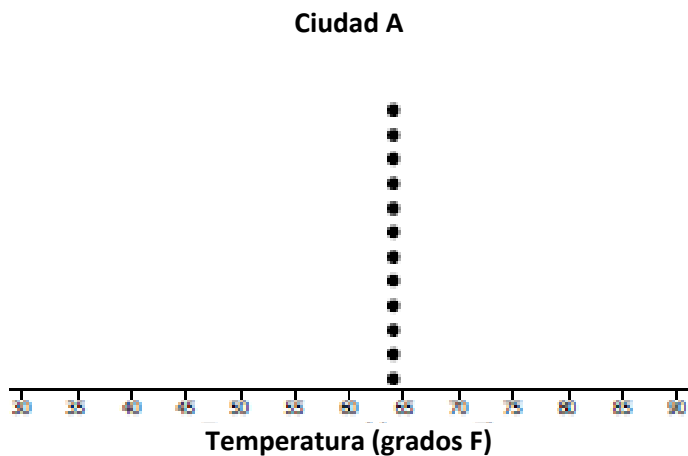
Temp.	Desviación de la media	Desviación representada como un número con signo	Valor absoluto (desviaciones absolutas)
53	10 a la izquierda	- 10	10
57			
60			
60			
64			
64			
64			
64			
64			
68			
68			
70			

- b. Teniendo en cuenta el siguiente gráfico, ¿cuál es la suma de las desviaciones absolutas?



- c. Sabina sugiere que la media de las desviaciones absolutas podría ser una medida de la variabilidad en un conjunto de datos. Su valor es la distancia promedio a la que se encuentran todos los valores de datos con respecto a la temperatura media mensual. Se llama desviación media absoluta y se denota con las letras DMA. Encuentra la DMA de este conjunto de datos de las temperaturas de la ciudad G. Redondea a la décima más cercana.
- d. Encuentra la DMA de las distribuciones de temperatura de cada una de las siete ciudades y utiliza los valores para ordenar las distribuciones desde la que tenga menor variabilidad hasta la que tenga mayor variabilidad. Recuerda que la media de cada conjunto de datos es 63 grados Fahrenheit. Mirando solamente las distribuciones, ¿coincide la lista que hiciste en el ejercicio 2 con la lista que hiciste ordenando los valores de la DMA?
- e. ¿Cuál de las siguientes es una interpretación correcta de la DMA?
- Las temperaturas mensuales de la ciudad G se encuentran todas dentro de los 3.7 grados con respecto a la media aproximada de 63 grados.
 - Las temperaturas mensuales de la ciudad G se encuentran, como promedio, a 3.7 grados con respecto a la temperatura media aproximada de 63 grados.
 - Todas las temperaturas mensuales de la ciudad G difieren de la temperatura media aproximada de 63 grados en 3.7 grados.

8. A continuación se encuentra el diagrama de puntos de las temperaturas de la ciudad A.



- a. ¿Cuánta variabilidad hay en las temperaturas de la ciudad A? ¿Por qué?
- b. ¿Concuerda la DMA con tu respuesta de la parte (a)?

Resumen de la lección

En esta lección, se desarrolló una fórmula que mide la cantidad de variabilidad en una distribución de datos.

- La desviación absoluta de un punto de datos representa qué tan lejos de la media se encuentra ese punto de datos.
- La desviación media absoluta (DMA) se calcula averiguando la media de las desviaciones absolutas de la distribución.
- El valor de la DMA es la distancia promedio a la que se encuentran todos los valores de datos con respecto a la media.
- Una DMA pequeña indica que la distribución tiene muy poca variabilidad.
- Una DMA grande indica que los puntos de datos están dispersos lejos de la media.

Conjunto de problemas

1. Supón que el diagrama de puntos de la izquierda muestra la cantidad de goles que anotó un equipo de fútbol de niños en los seis partidos que van de esta temporada, y que el diagrama de puntos de la derecha muestra la cantidad de goles que anotó un equipo de fútbol de niñas en los seis partidos que van de esta temporada. La media de ambos equipos es 3.

Diagrama de puntos de la cantidad de goles anotados por el equipo de los niños



Cantidad de goles anotados

Diagrama de puntos de la cantidad de goles anotados por el equipo de las niñas



Cantidad de goles anotados

- a. Antes de realizar cálculos, ¿qué diagrama de puntos tiene la mayor DMA? Explica cómo lo sabes.
- b. Utiliza las siguientes tablas para encontrar la DMA de la cantidad de goles de cada distribución. Redondea tus cálculos a la centésima más cercana.

Equipo de los niños	
Cantidad de goles	Desviaciones absolutas
0	
0	
3	
3	
5	
7	
Suma	

Equipo de las niñas	
Cantidad de goles	Desviaciones absolutas
2	
2	
3	
3	
3	
5	
Suma	

- c. Según los valores calculados de la DMA, ¿para qué distribución es la media un mejor indicador de un valor típico? Explica tu respuesta.

2. Recuerda el problema de Robert en el que tenía que decidir si mudarse a la ciudad de Nueva York o a San Francisco. La siguiente es la tabla de temperaturas (en grados Fahrenheit) y las desviaciones absolutas de la distribución correspondiente a la ciudad de Nueva York:

Mes	Temperatura promedio de la ciudad de Nueva York											
	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Temperatura	39	42	50	61	71	81	85	84	76	65	55	47
Desviación absoluta	24	21	13	2	8	18	22	21	13	2	8	16

- a. Las desviaciones absolutas de las temperaturas mensuales se muestran en la tabla anterior. Utiliza esta información para calcular la DMA. Explica con palabras la DMA de este problema.
- b. Completa la siguiente tabla y luego utiliza los valores para calcular la DMA de la distribución de datos de San Francisco.

Mes	Temperatura promedio de San Francisco											
	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Temperatura	57	60	62	63	64	67	67	68	70	69	63	58
Desviaciones absolutas												

- c. Comparando los valores de la DMA de la ciudad de Nueva York y de San Francisco, ¿qué ciudad elegiría Robert para mudarse si le interesa tener mucha variabilidad en las temperaturas mensuales? Explica utilizando la DMA.

3. Ten en cuenta los siguientes datos sobre la cantidad de gominolas verdes que hay en siete bolsas de muestra de cinco fabricantes diferentes de dulces (Awesome, Delight, Finest, Sweeties y YumYum). Observa que la media de cada distribución es de 42 gominolas.

	Bolsa 1	Bolsa 2	Bolsa 3	Bolsa 4	Bolsa 5	Bolsa 6	Bolsa 7
Awesome	40	40	41	42	42	43	46
Delight	22	31	36	42	48	53	62
Finest	26	36	40	43	47	50	52
Sweeties	36	39	42	42	42	44	49
YumYum	33	36	42	42	45	48	48

- a. Completa la siguiente tabla de desviaciones absolutas de la cantidad de gominolas verdes con respecto a la cantidad media de estos dulces que hay en las siete bolsas para cada fabricante de dulces.

	Bolsa 1	Bolsa 2	Bolsa 3	Bolsa 4	Bolsa 5	Bolsa 6	Bolsa 7
Awesome	2	2	1	0	0	1	4
Delight	20	11	6				
Finest	16						
Sweeties							
YumYum							

- b. En función de lo que aprendiste sobre la DMA, ¿qué fabricante crees que tendrá la menor DMA?
 Calcula la DMA del fabricante que seleccionaste.

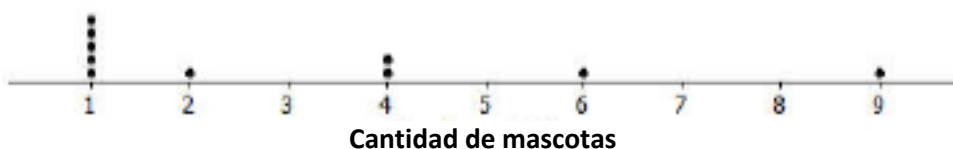
	Bolsa 1	Bolsa 2	Bolsa 3	Bolsa 4	Bolsa 5	Bolsa 6	Bolsa 7	SUMA	DMA
Awesome									
Delight									
Finest									
Sweeties									
YumYum									

Lección 10: Descripción de distribuciones mediante el uso de la media y la desviación media absoluta (DMA) (Conocida también como MAD)

Trabajo en clase

Ejemplo 1: Descripción de distribuciones

En la Lección 9, Sabina desarrolló la desviación media absoluta (DMA) como un número que medía la variabilidad en una distribución de datos. Utilizar la media y la DMA con un diagrama de puntos te permite describir el centro, la dispersión y la forma de una distribución de datos. Por ejemplo, supón que los datos sobre la cantidad de mascotas de diez estudiantes se muestran en el siguiente diagrama de puntos.



Hay varias maneras de describir la distribución de datos. La cantidad media de mascotas de estos estudiantes es tres, lo cual es una medida del centro. Hay variabilidad en la cantidad de mascotas que tienen los estudiantes, lo que promedia 2.2 mascotas con respecto a la media (la DMA). La forma de la distribución es pesada a la izquierda y se aligera hacia la derecha.

Ejercicios 1 a 4

1. Supón que los pesos de siete mochilas de estudiantes de la escuela intermedia se proporcionan a continuación.
 - a. Completa la siguiente tabla.

Estudiante	Alan	Beth	Char	Damon	Elisha	Fred	Georgia
Peso (libras)	18	18	18	18	18	18	18
Desviaciones							
Desviaciones absolutas							

- b. Haz un diagrama de puntos de estos datos, y calcula la media y la DMA.

c. Describe esta distribución de los pesos de las mochilas a través del análisis del centro, de la dispersión y de la forma.

2. Supón que el peso de la mochila de Elisha es de 17 libras en lugar de 18.

a. Haz un diagrama de puntos de la nueva distribución.

b. Sin hacer ningún cálculo, ¿cómo se ve afectada la media por el peso más liviano? ¿La nueva media sería la misma, más pequeña o más grande?

c. Sin hacer ningún cálculo, ¿cómo se ve afectada la DMA por el peso más liviano? ¿La nueva DMA sería la misma, más pequeña o más grande?

3. Supón que, además de que el peso de la mochila de Elisha cambió de 18 a 17 libras, el peso de la mochila de Fred cambió de 18 a 19 libras.

a. Haz un diagrama de puntos de la nueva distribución.

- b. Sin hacer ningún cálculo, ¿qué diferencia hay entre la nueva media y la media original?
- c. Sin hacer ningún cálculo, ¿la DMA de la nueva distribución sería la misma, más pequeña o más grande que la DMA original?
- d. Sin hacer ningún cálculo, ¿qué diferencia hay entre la DMA de la nueva distribución y la del ejercicio 2?

4. Supón que los pesos de las mochilas de siete estudiantes de segundo grado fueran los siguientes:

Estudiante	Alice	Bob	Carol	Damon	Ed	Felipe	Gale
Peso (libras)	5	5	5	5	5	5	5

- a. ¿En qué se parecen la distribución de los pesos de las mochilas de los estudiantes de segundo grado y la distribución original de los estudiantes de la escuela media que se proporciona en el ejercicio 1?
- b. ¿En qué se diferencian las distribuciones?

Ejemplo 2: Uso de la DMA (Conocida también como MAD por sus siglas en inglés)

La toma de decisiones a través de la comparación de distribuciones es una función importante de Estadística. Recuerda que Robert está intentando decidir si mudarse a la ciudad de Nueva York o a San Francisco según la temperatura. Comparar el centro, la dispersión y la forma de las dos distribuciones de temperatura podría ayudarlo a decidirse.

Diagrama de puntos de la temperatura de la ciudad de Nueva York

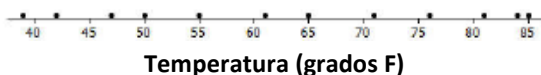
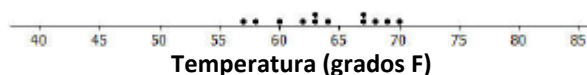


Diagrama de puntos de la temperatura de San Francisco



A partir de los diagramas de puntos, Robert vio que las temperaturas mensuales de la ciudad de Nueva York estaban dispersas casi en forma pareja desde aproximadamente 40 grados hasta más de 80 grados, pero que en San Francisco las temperaturas mensuales no variaban tanto. Se sorprendió al ver que la temperatura media era casi la misma en ambas ciudades. La DMA de 14 grados de la ciudad de Nueva York le indicaba que, en promedio, la temperatura de un mes estaba a 14 grados de la media de 63 grados Fahrenheit. Eso representa mucha variabilidad, lo que es congruente con el diagrama de puntos. Por otro lado, la DMA de San Francisco le indicaba que las temperaturas mensuales de San Francisco diferían, en promedio, solo 3.5 grados con respecto a la media de 64 grados Fahrenheit. Entonces, la media no ayuda mucho a Robert para tomar una decisión, pero la DMA y el diagrama de puntos son útiles.

¿Qué ciudad debería elegir si le encanta el clima cálido y realmente le desagrada el clima frío?

Ejercicios 5 a 7

5. Robert quiere comparar las temperaturas de las ciudades B y C.

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Ciudad B	54	54	58	63	63	68	72	72	72	63	63	54
Ciudad C	54	44	54	61	63	72	78	85	78	59	54	54

- a. Haz un diagrama de puntos de las temperaturas mensuales de cada una de las ciudades.

- b. Verifica que la temperatura media mensual de cada distribución sea de 63 grados Fahrenheit.
- c. Encuentra la DMA de cada ciudad. Interpreta las dos DMA en palabras y compara sus valores.
6. ¿Cómo describirías las diferencias en las formas de las distribuciones de las temperaturas mensuales de las dos ciudades?
7. Supón que Robert tenía que decidir entre las ciudades D, E y F.

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Media	DMA
Ciudad D	54	44	54	59	63	72	78	87	78	59	54	54	63	10.5
Ciudad E	56	56	56	56	56	84	84	84	56	56	56	56	63	10.5
Ciudad F	42	42	70	70	70	70	70	70	70	70	70	42	63	10.5

- a. Haz diagramas de puntos para cada distribución.

- b. Interpreta la DMA de las distribuciones. ¿Qué significa esto en cuanto a la variabilidad?
- c. ¿Cómo decidirá Robert a qué ciudad se debería mudar? Enumera las posibles razones que podría tener Robert para elegir cada ciudad.

Resumen de la lección

Una distribución de datos se puede describir en términos de su centro, dispersión y forma.

- El centro se puede medir con la media.
- La dispersión se puede medir con la desviación media absoluta (DMA).
- Un diagrama de puntos muestra la forma de la distribución.

Conjunto de problemas

1. Haz un diagrama de puntos de los tiempos durante los que cinco estudiantes estudiaron para un examen si el tiempo medio durante el que estudiaron fue de dos horas y la DMA fue de cero horas.
2. Supón que los tiempos durante los que cinco estudiantes estudiaron para un examen son los siguientes:

Estudiante	Aria	Ben	Chloe	Dellan	Emma
Tiempo (horas)	1.5	2	2	2.5	2

Michelle dijo que la DMA de este conjunto de datos es 0 porque el diagrama de puntos se equilibra aproximadamente en 2. Sin hacer ningún cálculo, ¿estás de acuerdo con Michelle? ¿Por qué sí o por qué no?

3. Supón que la cantidad de mensajes de texto que reciben ocho estudiantes un día común es la siguiente:

Estudiante	1	2	3	4	5	6	7	8
Cantidad	42	56	35	70	56	50	65	50

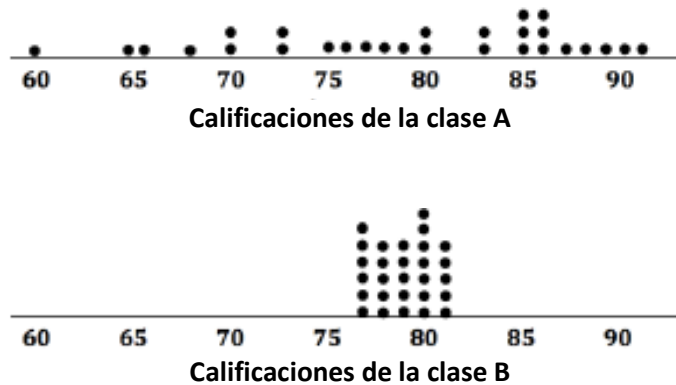
- a. Haz un diagrama de puntos de la cantidad de mensajes de texto que reciben estos ocho estudiantes un día común.
- b. Encuentra la cantidad media de mensajes de texto que reciben estos ocho estudiantes un día común.
- c. Encuentra la DMA de la cantidad de mensajes de texto y explica su significado utilizando las palabras de este problema.
- d. Describe la forma de esta distribución de datos.
- e. Supón que, en el conjunto de datos original, el estudiante 3 recibe cinco mensajes de texto adicionales por día y que el estudiante 4 recibe cinco mensajes de texto menos por día.
 - i. Sin hacer ningún cálculo, ¿la media del nuevo conjunto de datos permanece igual, aumenta o disminuye en comparación con la media original? Explica tu razonamiento.
 - ii. Sin hacer ningún cálculo, ¿la DMA del nuevo conjunto de datos permanece igual, aumenta o disminuye en comparación con la DMA original? Explica tu razonamiento.

Lección 11: Descripción de distribuciones mediante el uso de la media y la desviación media absoluta (DMA) (Conocido también como MAD)

Trabajo en clase

Ejemplo 1: Comparación de distribuciones con la misma media

En la Lección 10, una distribución de datos se caracterizaba principalmente por su centro (media) y su variabilidad (DMA, conocida también como MAD). La manera en la que estas medidas nos ayudan a tomar una decisión suele depender del contexto de la situación. Por ejemplo, supón que dos clases de estudiantes realizaron el mismo examen y que sus calificaciones (en una escala de 100 puntos) se muestran en los siguientes diagramas de puntos. La calificación media de cada distribución es de 79 puntos. ¿Preferirías estar en la clase A o en la clase B si tuvieras una calificación de 79?



Ejercicios 1 a 6

1. Mirando los diagramas de puntos, ¿qué clase tiene la mayor DMA? Explica sin calcular realmente la DMA.
2. Si Liz tuvo una de las calificaciones más altas de su clase, ¿en qué clase preferiría estar? Explica tu razonamiento.
3. Si Logan obtuvo una calificación por debajo del promedio, ¿en qué clase preferiría estar? Explica tu razonamiento.

Tu hermano menor te pide que cambies la batería de su auto de control remoto favorito. El auto está hecho de manera tal que es difícil cambiar la batería. Tu investigación sobre la duración (en horas) de dos marcas diferentes de baterías (A y B) muestra los siguientes datos de 20 baterías de cada marca:

A	12	14	14	15	16	17	17	18	19	20	21	21	23	23	24	24	24	25	26	27
B	18	18	19	19	19	19	19	19	20	20	20	20	20	21	21	21	21	22	22	22

- Para ayudarte a decidir qué batería comprar, comienza haciendo un diagrama de puntos de cada marca.
- Encuentra la duración media de las baterías de cada marca y compáralas.
- Mirando la variabilidad de cada conjunto de datos que se muestra en su diagrama de puntos, escribe una razón por la que elegirías la marca A. ¿Cuál sería una razón por la que elegirías la marca B? Explica tu razonamiento.

Ejemplo 2: Comparación de distribuciones con medias diferentes

Has estado comparando distribuciones que tienen la misma media, pero una variabilidad diferente. Como has visto, decidir si es mejor una variabilidad grande o pequeña depende del contexto y de lo que se pide. Por ejemplo, en el ejercicio 2, Liz prefirió estar en la distribución con más variabilidad porque tenía una de las calificaciones más altas de la clase. Entonces, su calificación hubiera sido más alta si hubiera estado en la clase A en lugar de la clase B. Logan, por otro lado, prefirió la clase con menos variabilidad (es decir, la clase B), ya que su calificación estaba por debajo del promedio.

Si dos distribuciones de datos tienen diferentes medias, ¿qué papel desempeña una medida de variabilidad en la toma de decisiones?

Ejercicios 7 a 9

Supón que quisieras responder la siguiente pregunta: ¿los grillos de campo predicen mejor la temperatura atmosférica que las cigarras? Ambas especies de insectos emiten sonidos frotándose las alas delanteras.

Los siguientes datos representan la cantidad de sonidos (por minuto) que emite cada insecto en un grupo de 10 insectos. Todos los datos se tomaron en la misma noche y al mismo tiempo.

Insecto	1	2	3	4	5	6	7	8	9	10
Grillos	35	32	35	37	34	34	38	35	36	34
Cigarras	66	62	61	64	63	62	68	64	66	64

7. Haz diagramas de puntos de estas dos distribuciones de datos utilizando la misma escala, de 30 a 70. Visualmente, ¿a qué conclusiones puedes llegar a partir de los diagramas de puntos?

8. Calcula la media y la DMA de cada distribución.

9. La temperatura exterior T se puede predecir contando la cantidad de sonidos que emiten estos insectos.
- a. En el caso de los grillos, T se encuentra sumando 40 a su cantidad media de sonidos por minuto. ¿Qué valor de T predicen los grillos?
- b. En el caso de las cigarras, T se encuentra sumando 161 a su cantidad media de sonidos por minuto y luego dividiendo la suma por 3. ¿Qué valor de T predicen las cigarras?
- c. La temperatura era de 75 grados cuando se registraron estos datos, por lo que utilizar la media de cada conjunto de datos proporcionó una predicción precisa de la temperatura. Si fueras a usar la cantidad de sonidos de un único grillo o de una única cigarra para predecir la temperatura, ¿usarías un grillo o una cigarra? Explica cómo la variabilidad en las distribuciones de la cantidad de sonidos desempeñó un papel importante en tu decisión.

Conjunto de problemas

1. Dos clases realizaron el mismo examen de Matemáticas. Las medidas de resumen de las dos clases son las siguientes:

	Media	DMA
Clase A	78	2
Clase B	78	10

- Supón que obtuviste la calificación más alta de tu clase. ¿Tu calificación hubiera sido más alta si hubieras estado en la clase A o en la clase B? Explica tu razonamiento.
 - Supón que tu calificación estuvo por debajo de la calificación media. ¿En qué clase preferirías haber estado? Explica tu razonamiento.
2. Ocho plantas de tomate de dos variedades (LoveEm y Wonderful) crecen en las mismas condiciones. A continuación, se muestran la cantidad de tomates que produce cada planta de cada variedad:

Planta	1	2	3	4	5	6	7	8
LoveEm	27	29	27	28	31	27	28	27
Wonderful	31	20	25	50	32	25	22	51

- Haz diagramas de puntos para ayudarte a decidir qué variedad es más productiva.
- Calcula la cantidad media de tomates que produce cada variedad. ¿Cuál produce más tomates de promedio?
- Si quieres poder predecir con precisión la cantidad de tomates que va a producir una planta, ¿qué variedad elegirías: la que tiene la DMA más pequeña o la que tiene la DMA más grande? Explica tu razonamiento.
- Calcula la DMA de cada variedad de planta.

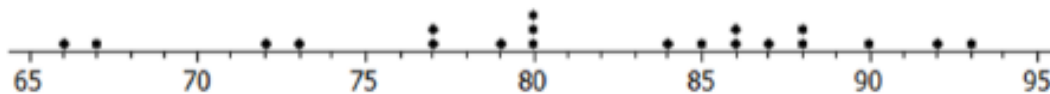
Lección 12: Descripción del centro de una distribución mediante el uso de la mediana

¿Cómo resumimos una distribución de datos? ¿Qué nos proporciona una buena descripción de los datos? Los siguientes ejercicios nos ayudan a comprender cómo un resumen numérico responde a estas preguntas.

Trabajo en clase

Ejemplo 1: La mediana - Un número típico

Supón que un restaurante de una cadena (restaurante A) promociona que la cantidad típica de papas fritas que hay en una bolsa grande es de 82 papas fritas. El gráfico muestra la cantidad de papas fritas que hay en muestras seleccionadas de bolsas grandes del restaurante A.



Cantidad de papas fritas que hay en una bolsa grande (restaurante A)

Algunas veces, es útil saber qué punto separa una distribución de datos en dos partes iguales, donde una parte representa la mitad más grande de los valores de datos y la otra parte representa la mitad más pequeña de los valores de datos. Este punto se llama *mediana*. Cuando los datos se ordenan de menor a mayor, habrá la misma cantidad de valores por encima y por debajo de ese punto.

Ejercicios 1 a 3

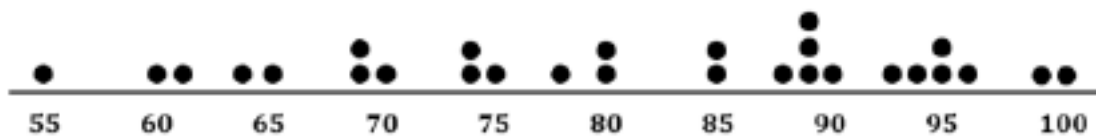
- Acabas de comprar una bolsa grande de papas fritas del restaurante. ¿Crees que tienes 82 papas fritas? ¿Por qué sí o por qué no?
- ¿Cuántas bolsas había en la muestra?

3. ¿Cuál(es) de las siguientes afirmaciones parecería(n) ser verdadera(s) según los datos? Explica tu razonamiento.
- La mitad de las bolsas contenía más de 82 papas fritas.
 - La mitad de las bolsas contenía menos de 82 papas fritas.
 - Más de la mitad de las bolsas contenía más de 82 papas fritas.
 - Más de la mitad de las bolsas contenía menos de 82 papas fritas.
 - Si obtuviste una bolsa de papas fritas al azar, podrías tener exactamente 93 papas fritas.

Ejemplo 2

Examina el siguiente diagrama de puntos.

Notas en un examen de Ciencias



- ¿Cuántos valores de datos se representan en el diagrama de puntos anterior?
- ¿Cuántos valores de datos debería haber por encima de la mediana? ¿Cuántos por debajo de la mediana? Explica tu respuesta.
- ¿Cuál es la mediana de los datos presentados en el diagrama de puntos?

- d. ¿Qué nos indica esta información acerca de los datos?

Ejemplo 3

Utiliza la información del diagrama de puntos del ejemplo 2.

- a. ¿Qué porcentaje de estudiantes obtuvo una calificación mayor que la mediana? ¿Y menor que la mediana?
- b. Supón que el maestro calificó exámenes de manera incorrecta y que el estudiante que obtuvo 65, en realidad, obtuvo 71. ¿Cambiaría la mediana? ¿Por qué sí o por qué no?
- c. Supón que el estudiante que obtuvo 65, en realidad, obtuvo 89. ¿Cambiaría la mediana? ¿Por qué sí o por qué no?

Ejemplo 4

Supón que estabas intentando convencer a tu familia de que necesitabas un par de tenis nuevas. Después de hablar con tus amigos, argumentaste que la mitad de ellos tenía más de cuatro pares de tenis y que tú solo tenías dos pares. Da otro ejemplo de cuándo querrías saber que un valor de datos es un punto medio. Explica tu razonamiento.

Ejercicios 4 y 5

4. El dueño de la cadena decidió comprobar la cantidad de papas fritas de otro restaurante de la cadena. Aquí están los datos del restaurante B: 82, 83, 83, 79, 85, 82, 78, 76, 76, 75, 78, 74, 70, 60, 82, 82, 83, 83 y 83.
- ¿Cuántas bolsas de papas fritas se contaron?
 - Sallee afirma que la mediana es 75, ya que ve que 75 es el número que se encuentra en el medio del conjunto de datos enumerados anteriormente. Ella cree que la mitad de las bolsas tenían menos de 75 papas fritas. ¿Crees que cambiaría de opinión si los datos se graficaran en un diagrama de puntos? ¿Por qué sí o por qué no?
 - Jake dijo que la mediana era 83. ¿Qué le dirías a Jake?
 - Betse afirmó que la mediana se encontraba a la mitad entre 60 y 85 o 72.5. ¿Crees que tiene razón? ¿Por qué sí o por qué no?
 - Chris pensó que la mediana era 82. ¿Estás de acuerdo? ¿Por qué sí o por qué no?

5. Calcula la media y compárala con la mediana. ¿Qué observas acerca de los dos valores? Si tanto la media como la mediana son medidas del centro, ¿por qué piensas que una de ellas es menor que la otra?

Ejercicios 6 a 8: Averiguación de medianas a partir de tablas de frecuencias

6. En un tercer restaurante (restaurante C), se realizó un conteo de una muestra de bolsas de papas fritas y se encontraron los siguientes resultados.

Cantidad de papas fritas	Frecuencia
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	

- a. ¿Cuántas bolsas de papas fritas se contaron?
- b. ¿Cuál es la mediana de la cantidad de papas fritas de la muestra de bolsas de este restaurante? Describe cómo encontraste tu respuesta.

7. Robere decidió dividir los datos en cuatro partes y encontró la mediana de todo el conjunto.
 - a. Enumera los 13 valores de la mitad inferior. Encuentra la mediana de estos 13 valores.

 - b. Enumera los 13 valores de la mitad superior. Encuentra la mediana de estos 13 valores.

8. ¿Cuál de los tres restaurantes parece tener más probabilidades de tener realmente 82 papas fritas en una bolsa típica? Explica tu razonamiento.

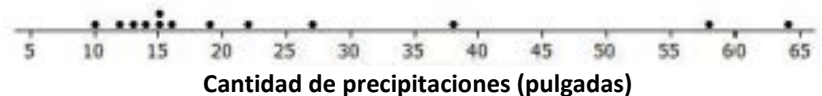
Resumen de la lección

La **mediana** es el punto medio de un conjunto de datos ordenados. Esta separa los datos en dos partes con la misma cantidad de valores por debajo del punto como por encima de ese punto. Para encontrar una mediana, primero tienes que ordenar los datos. En caso de haber una cantidad par de valores de datos, encuentra el promedio de los dos números del medio. En caso de haber una cantidad impar de valores de datos, utiliza el valor del medio.

Conjunto de problemas

- La cantidad de precipitaciones en los estados occidentales de los Estados Unidos se proporciona en la tabla, así como en el diagrama de puntos.

Estado	Cantidad de precipitaciones (pulgadas)
Washington	38.4
Oregón	27.4
California	22.2
Montana	15.3
Idaho	18.9
Wyoming	12.9
Nevada	9.5
Utah	12.2
Colorado	15.9
Arizona	13.6
Nuevo México	14.6
Arkansas	58.3
Hawái	63.7



Fuente: <http://www.currentresults.com/Weather/US/average-annual-state-precipitation.php>

- ¿Cómo varían las cantidades en los estados?
- Encuentra la mediana. ¿Qué te indica la mediana sobre la cantidad de precipitaciones?
- Utiliza la mediana y el rango para describir las precipitaciones mensuales promedio en los estados occidentales de los Estados Unidos.
- ¿Cuál crees que sería una mejor descripción de la cantidad típica de precipitaciones: la media o la mediana? Explica tu razonamiento.

2. Identifica lo siguiente como verdadero o falso. Si una afirmación es falsa, da un ejemplo que muestre por qué lo es.
- La mediana siempre es igual a uno de los valores del conjunto de datos.
 - La mediana es el punto medio entre el menor y el mayor valor de un conjunto de datos.
 - Como máximo, la mitad de los valores de un conjunto de datos tiene valores menores que la mediana.
 - En un conjunto de datos con 25 valores diferentes, si cambias los dos valores más bajos de un conjunto de datos a valores menores, la mediana no cambiará.
 - Si sumas 10 a cada elemento de un conjunto de datos, la mediana no cambiará.
3. Forma un conjunto de datos de manera que lo siguiente sea verdadero:
- El grupo tiene 11 valores diferentes, y la mediana es 5.
 - El grupo tiene 10 valores, y la mediana es 25.
 - El grupo tiene 7 valores, y la mediana es igual al menor valor.
4. El diagrama de puntos muestra la cantidad de teléfonos fijos que una muestra de personas tiene en sus casas.



- ¿Cuántas personas había en la muestra?
- ¿Por qué crees que tres personas no tienen teléfonos fijos en sus casas?
- Encuentra la mediana de la cantidad de teléfonos de las personas de la muestra.
- Utiliza la mediana y el rango (máximo-mínimo) para describir la distribución de la cantidad de teléfonos.

5. A continuación, se proporcionan los salarios de Los Angeles Lakers para la temporada de baloncesto 2012-2013.

Jugador	Salario (\$)
Kobe Bryant	\$27,849,149
Dwight Howard	\$19,536,360
Pau Gasol	\$19,000,000
Steve Nash	\$8,700,000
Metta World Peace	\$7,258,960
Steve Blake	\$4,000,000
Jordan Hill	\$3,563,600
Chris Duhon	\$3,500,000
Jodie Meeks	\$1,500,000
Earl Clark	\$1,240,000
Devin Ebanks	\$1,054,389
Darius Morris	\$962,195
Antawn Jamison	\$854,389
Robert Sacre	\$473,604
Darius Johnson-Odom	\$203,371

Fuente: www.basketball-reference.com/contracts/LAL.html

- Con solo mirar los datos, ¿qué adviertes acerca de los salarios?
 - Encuentra la mediana de los salarios y explica qué te indica sobre ellos.
 - Encuentra la mediana de la mitad inferior de los salarios y la mediana de la mitad superior de los salarios.
 - Encuentra el ancho de cada uno de los siguientes intervalos. ¿Qué adviertes acerca del tamaño de los anchos de los intervalos y qué te indica eso sobre los salarios?
 - Del salario mínimo a la mediana de la mitad inferior:
 - De la mediana de la mitad inferior a la mediana de todo el conjunto:
 - De la mediana de todo el conjunto a la mediana de la mitad superior:
 - De la mediana de la mitad superior al salario más alto:
6. Utiliza la tabla de salarios anterior para responder lo siguiente:
- Si tuvieras que encontrar el salario medio, ¿qué diferencia crees que este tendría con la mediana? Explica tu razonamiento.
 - ¿Qué medida crees que daría una mejor idea de un salario típico de los Lakers: la media o la mediana? Explica tu razonamiento.

Lección 13: Descripción de la variabilidad mediante el uso del rango intercuartílico (R. I.C.) (También conocido por sus siglas en inglés- IQR)

Trabajo en clase

La mediana se utilizó para describir el valor típico de nuestros datos de la Lección 12. Claramente, no todos los datos se describen con este valor. ¿Cómo encontramos una descripción de la manera en la que varían los datos? ¿Cuál es una buena manera de indicar cómo varían los datos cuando utilizamos una mediana como nuestro valor típico? Estas preguntas se desarrollan en los siguientes ejercicios.

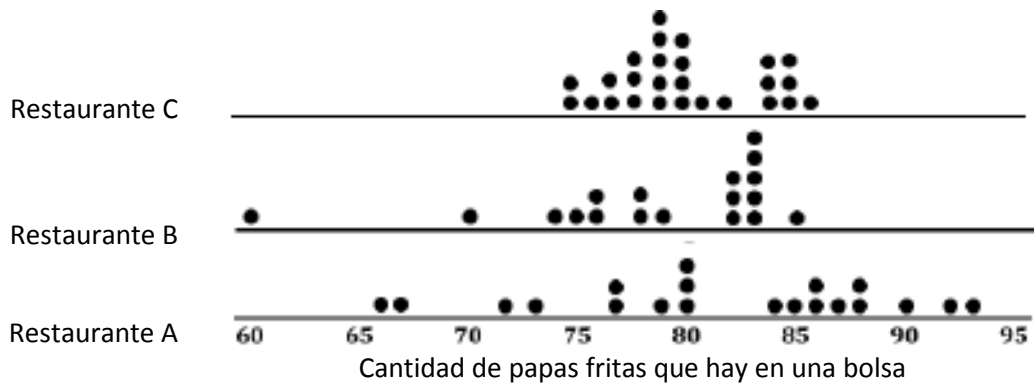
Ejercicios 1 a 4: Más papas fritas

1. En la Lección 12, pensaste acerca de la afirmación que hizo un restaurante de una cadena de que la cantidad típica de papas fritas que había en una bolsa grande era de 82 papas fritas. Luego, miraste los datos sobre la cantidad de papas fritas que había en una bolsa de tres de los restaurantes.
 - a. ¿Cómo crees que se recabaron los datos y qué problemas podrían haber surgido al hacerlo?

 - b. ¿Qué escenario(s) daría(n) conteos que podrían no ser representativos de las bolsas típicas?

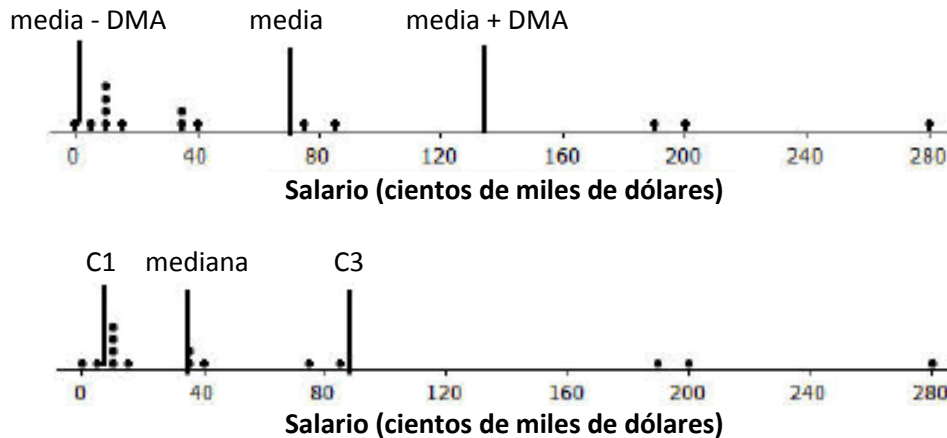
2. En el ejercicio 7 de la Lección 12, contrastaste la mediana de la mitad superior y la mediana de la mitad inferior de los conteos de cada uno de los tres restaurantes. Estas son las cantidades que encontraste: 87.5 y 77 en el restaurante A, 82 y 79 en el restaurante B, y 84 y 78 en el restaurante C. La diferencia entre las medianas de las dos mitades se llama *rango intercuartílico* o R. I.C (IQR)
 - a. ¿Cuál es el R. I.C. de cada uno de los tres restaurantes?

- b. ¿Qué restaurante tenía el R. I.C. más pequeño y qué te indica eso?
- c. ¿Aproximadamente qué fracción de los conteos se encontraría entre los cuartiles? Explica tu razonamiento.
3. Las medianas de la mitad inferior y superior de un conjunto de datos se llaman *cuartiles*. La mediana de la mitad superior de los datos se llama *cuartil superior*, y la mediana de la mitad inferior de los datos se llama *cuartil inferior*. ¿Tienen sentido estos nombres? ¿Por qué sí o por qué no?
4. a. Marca los cuartiles de cada restaurante en los siguientes gráficos.



Ejercicio 5: ¿Cuándo deberías utilizar el R. I.C.?

5. ¿Cuándo deberías utilizar el R. I.C.? Los datos de los salarios de 2012 del equipo de baloncesto de los Lakers se muestran en los siguientes dos diagramas. (Consulta el problema 5 del Conjunto de problemas de la Lección 12).



- a. Los datos se proporcionan en cientos de miles de dólares. ¿Qué representaría un salario de 4 millones de dólares?
- b. Las líneas verticales del diagrama superior muestran la media, y la media más y menos la DMA. El diagrama inferior muestra la mediana y el R. I.C. ¿Qué intervalo da una mejor idea de los salarios típicos? Explica tu razonamiento.

Ejercicio 6: Por tu cuenta con R. I.C.

6. Crea tres contextos diferentes en los que un conjunto de datos recabados en relación con esos contextos pudiera tener un R. I.C. de 20. Define una mediana para cada contexto. Sé específico acerca de la manera en la que se podrían haber recabado los datos y de las unidades involucradas. Prepárate para describir qué significan la mediana y el R. I.C. en cada caso.

a.

b.

c.

Resumen de la lección

Para encontrar el R. I.C., conocido también por sus siglas en inglés IQR, ordenas los datos, encuentras su mediana, y luego encuentras la media de la mitad inferior de los datos (cuartil inferior) y la mediana de la mitad superior de los datos (cuartil superior). El R. I.C. es la diferencia entre el cuartil superior y el cuartil inferior, lo que representa la longitud del intervalo que incluye la mitad intermedia de los datos (dado que la mediana y los dos cuartiles dividen los datos en cuatro secciones, con aproximadamente $\frac{1}{4}$ de los datos en cada sección). Dos de las secciones se encuentran entre los cuartiles, de manera que el intervalo entre los cuartiles contendría aproximadamente el 50% de los datos.

Conjunto de problemas

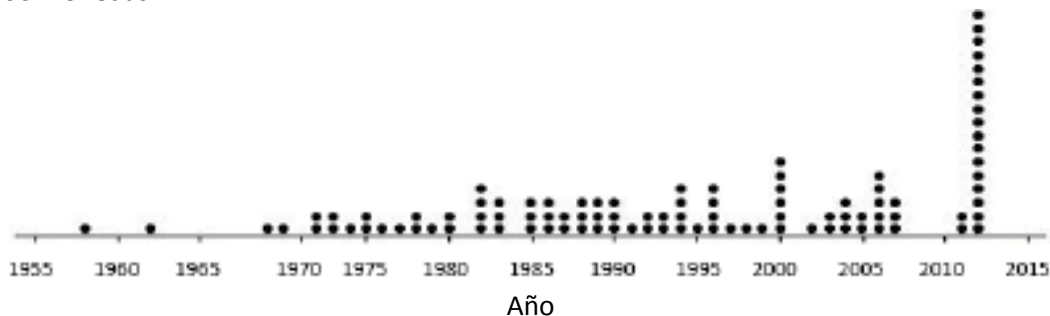
1. En la siguiente tabla, se proporcionan las temperaturas máximas mensuales promedio (en grados Fahrenheit) de St. Louis y de San Francisco.

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
St. Louis	40	45	55	67	77	85	89	88	81	69	56	43
San Francisco	57	60	62	63	64	67	67	68	70	69	63	57

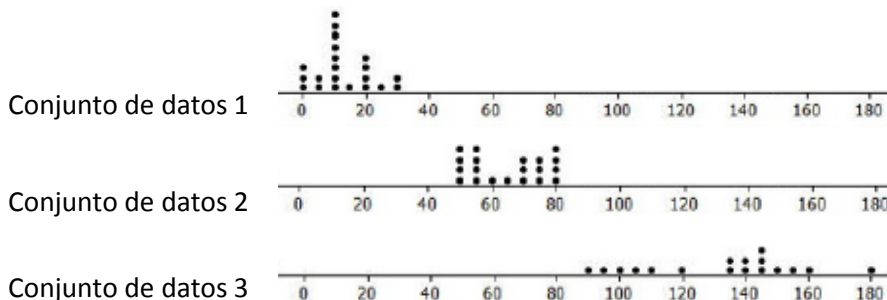
Fuente de datos: www.weather.com/weather/wxclimatology/monthly/graph/USCA0987
www.weather.com/weather/wxclimatology/monthly/graph/USMO0787

- ¿Cómo crees que se podrían haber recabado los datos?
- ¿Crees que sería posible que $\frac{1}{4}$ de las temperaturas de julio de St. Louis sea de 95 °F o más? ¿Por qué sí o por qué no?
- Haz una predicción acerca de qué diferencia hay entre los tamaños de los R. I.C. de las temperaturas de cada ciudad. Explica tu razonamiento.
- Encuentra el R. I.C. de la temperatura máxima mensual promedio de cada ciudad. ¿Qué diferencia hay entre los resultados y tu conjetura?

2. El siguiente diagrama muestra los años en los que se fabricó cada moneda de un centavo en un grupo de 100 monedas.



- ¿Qué te indica la pila de 17 puntos de 2012 (que representa 17 monedas de un centavo) sobre la antigüedad de las monedas de un centavo de 2014?
 - Aquí tienes información sobre la muestra de monedas de un centavo: El año medio en el que se fabricaron es 1994; el primer año en el que se fabricaron monedas de un centavo fue 1958; las monedas más nuevas se fabricaron en 2012; el C1 es 1984; la mediana es 1994; el C3 es 2006; y la DMA es 11.5 años. (Recuerda que C1 es el cuartil inferior y C3 es el cuartil superior.) Utiliza la información para indicar los años en los que se fabricó la mitad intermedia de las monedas de un centavo.
3. Crea un conjunto de datos con, al menos, 6 elementos de manera que tenga las siguientes propiedades:
- Un R. I.C. pequeño y un rango grande (máximo-mínimo).
 - Un R. I.C. igual al rango.
 - El cuartil inferior igual a la mediana.
4. Ordena los siguientes tres conjuntos de datos según el valor del R. I.C.



5. Aquí están los conteos de las papas fritas de cada una de las bolsas del restaurante A:
- 80, 72, 77, 80, 90, 85, 93, 79, 84, 73, 87, 67, 80, 86, 92, 88, 86, 88, 66 y 77.
- Supón que se pasó por alto una bolsa de papas fritas de la muestra y que esa bolsa contenía solo 50 papas fritas. ¿Cambiaría el R. I.C.? Explica tu razonamiento.
 - ¿Agregar otro valor de datos siempre cambia el R. I.C.? Da un ejemplo para justificar tu respuesta.

Lección 14: Resumen de una distribución mediante el uso de un diagrama de caja

Trabajo en clase

Un diagrama de caja es un gráfico que se utiliza para resumir una distribución de datos. ¿Qué nos indica el diagrama de caja sobre la distribución de datos? ¿Cómo indica el diagrama de caja la variabilidad de la distribución de datos?

Ejemplo 1: Tiempo para llegar a la escuela

¿Cuál es la cantidad típica de tiempo que le toma a una persona de tu clase llegar a la escuela? La cantidad de tiempo que toma llegar a la escuela por la mañana varía según la persona de tu clase. Tómame un minuto para responder las siguientes preguntas. Tu clase utilizará esta información para crear un diagrama de puntos.

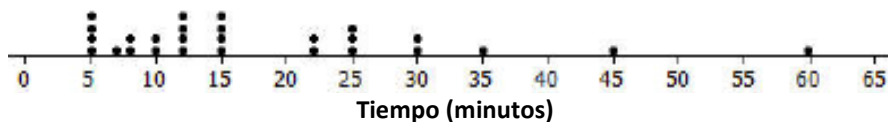
Escribe en una nota autoadhesiva tu nombre y una estimación de la cantidad de minutos que te tomó llegar a la escuela hoy.

¿Cuáles fueron algunas de las cosas en las que tuviste que pensar cuando hiciste tu estimación?

Ejercicios 1 a 4

Aquí se muestra un diagrama de puntos de las estimaciones de los tiempos que les tomó a los estudiantes de la clase del señor S llegar a la escuela una mañana.

Clase del señor S



1. Coloca una línea en el diagrama de puntos que parezca separar los tiempos más cortos y los tiempos más largos.
2. Coloca otra línea en el diagrama que separe a aquellos que parecen vivir muy cerca de la escuela y otra línea que marque a aquellos a los que les tomó mucho tiempo llegar a la escuela.

- Tu diagrama debería estar dividido en cuatro secciones. Anota la cantidad de valores de cada una de las cuatro secciones.
- Comparte tu diagrama de puntos marcado con algunos de tus compañeros. Compara cómo dividió cada uno el diagrama en cuatro secciones.

Ejercicios 5 a 7: Tiempo para llegar a la escuela

El maestro le pidió a la clase que hiciera una representación que resumiera los tiempos que tardaron los estudiantes de la clase del señor S en llegar a la escuela y que mostrara la manera en la que estaban dispersos. Tim decidió deshacerse de los puntos y solo utilizar un dibujo de las divisiones que hizo de los tiempos más cortos y de los tiempos más largos. Colocó un recuadro alrededor de las dos secciones intermedias.

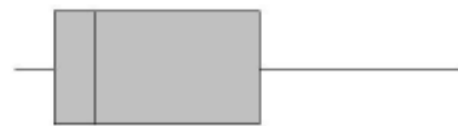
Tanya pensó que eso era una buena idea e hizo un dibujo de la manera en la que había dividido los tiempos. Aquí están sus dibujos.

Dibujo de Tim



Tiempo (minutos)

Dibujo de Tanya



Tiempo (minutos)

- ¿Qué te indican los dibujos sobre la cantidad de tiempo que tardan los estudiantes en llegar a la escuela?
- ¿Qué no te indican los dibujos sobre la cantidad de tiempo que tardan los estudiantes en llegar a la escuela?

7. ¿Qué diferencia hay entre los dos dibujos?

Ejemplo 2: Elaboración de un diagrama de caja

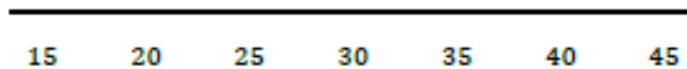
El señor S sugirió que, para estar seguros de que todos tuvieran la misma imagen, los estadísticos desarrollaron un procedimiento estándar para hacer las marcas de corte de las secciones.

El señor S escribió lo siguiente en la pizarra:

Para hacer un diagrama de caja:

- Encuentra la mediana de todos los datos.
- Encuentra el C1 (la mediana de la mitad inferior de los datos) y el C3 (la mediana de la mitad superior de los datos). Dibuja un recuadro que vaya desde el C1 hasta el C3 y que esté dividido en dos mitades por la mediana.
- Dibuja un segmento que una el valor mínimo con el recuadro y otro que una el valor máximo con el recuadro.

Ahora, utiliza la línea numérica proporcionada para hacer un diagrama de caja de los siguientes datos.
20, 21, 25, 31, 35, 38, 40, 42, 44



El resumen de cinco números es el siguiente:

Mín. =

C1 =

Mediana =

C3 =

Máx. =

Resumen de la lección

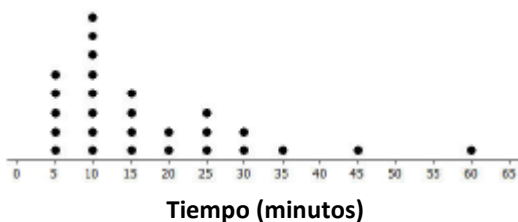
Aprendiste cómo hacer un diagrama de caja haciendo lo siguiente:

- Averiguaste la mediana de todo el conjunto de datos.
- Averiguaste el C1 (la mediana de la mitad inferior de los datos) y el C3 (la mediana de la mitad superior de los datos).
- Dibujaste un recuadro que iba desde el C1 hasta el C3 y que estaba dividido en dos mitades por la mediana.
- Dibujaste un segmento que unía el valor mínimo con el recuadro y otro que unía el valor máximo con el recuadro.

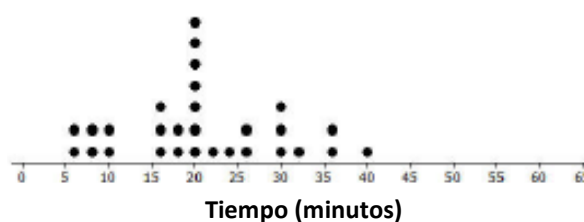
Conjunto de problemas

1. A continuación, se muestran los diagramas de puntos de la cantidad de tiempo que les toma llegar a la escuela a los estudiantes de las clases del señor S y de la señora J.

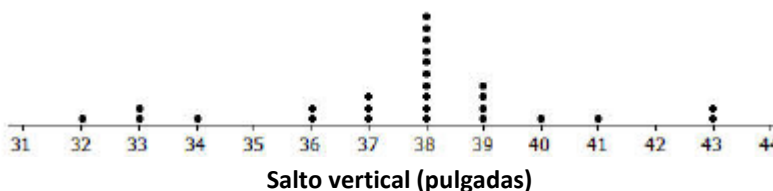
Clase del señor S



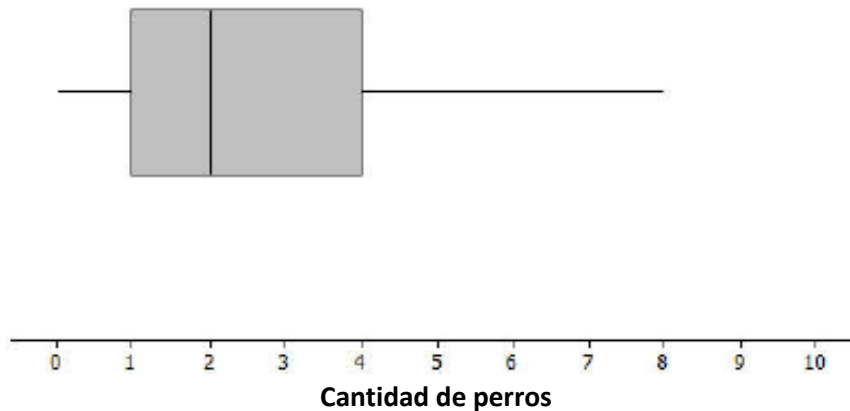
Clase de la señora J



- a. Haz un diagrama de caja de los tiempos de cada clase.
 - b. ¿Qué es lo que puedes ver en el diagrama de puntos que no puedes ver en el diagrama de caja? ¿Qué es lo más fácil de ver en el diagrama de caja que en el diagrama de puntos?
2. El siguiente diagrama de puntos muestra el salto vertical de algunos jugadores de la NBA. Un salto vertical es qué tan alto puede saltar un jugador sin tomar carrera. Haz un diagrama de caja de las alturas de los saltos verticales de los jugadores de la NBA por encima del diagrama de puntos.



3. Las temperaturas medias diarias en grados Fahrenheit del mes de febrero en una ciudad determinada son las siguientes:
- 4, 11, 14, 15, 17, 20, 30, 23, 20, 35, 35, 31, 34, 23, 15, 19, 39, 22, 15, 15, 19, 39, 22, 23, 29, 26, 29, 29
- Haz un diagrama de caja de las temperaturas.
 - Haz una predicción sobre la parte de los Estados Unidos en donde piensas que se podría encontrar la ciudad. Explica tu razonamiento.
 - Describe la distribución de datos de la temperatura. Incluye una descripción del centro y de la dispersión.
4. El siguiente diagrama muestra los resultados de una encuesta de hogares sobre la cantidad de perros que tienen. Identifica los siguientes enunciados como verdaderos o falsos. Explica tu razonamiento en cada caso.



- La cantidad máxima de perros por casa es de 8.
- Al menos $\frac{1}{2}$ de las casas tiene 2 o más perros.
- Todas las casas tienen perros.
- La mitad de las casas encuestadas tiene entre 2 y 4 perros.
- La mayoría de las casas encuestadas no tiene perros.

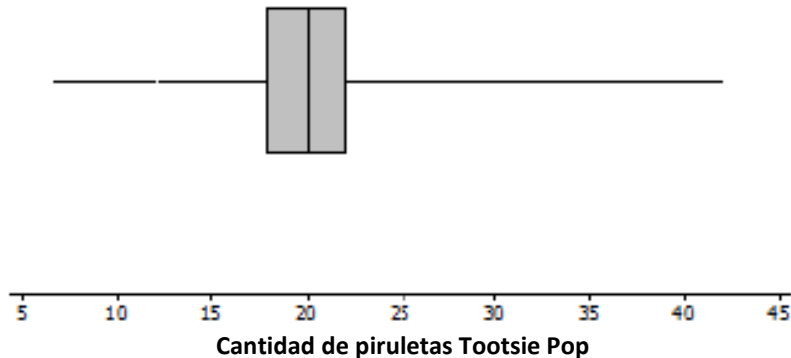
Lección 15: Más práctica con diagramas de caja

Trabajo en clase

Introduces tu mano en una jarra con piruletas Tootsie Pop. ¿Cuántas piruletas Tootsie Pop piensas que podrían caber en una mano? ¿Crees que esta cantidad es mayor o menor que la de otros estudiantes? ¿La cantidad que podría caber es una cantidad típica de piruletas Tootsie Pop? Esta lección analiza estas preguntas.

Ejemplo 1: Piruletas Tootsie Pop

Se les pidió a noventa y cuatro personas que agarraran la mayor cantidad de piruletas Tootsie Pop que pudieran. Aquí se muestra un diagrama de caja de estos datos. ¿Estás sorprendido?



Ejercicios 1 a 5

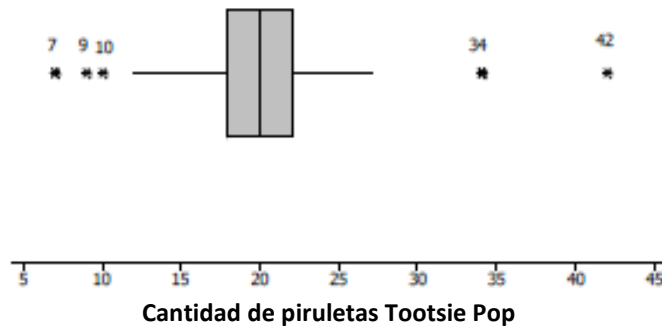
1. ¿Qué podría explicar la variabilidad en la cantidad de piruletas Tootsie Pop que esas 94 personas podían agarrar?
2. Estima los valores del resumen de cinco números del diagrama de caja.

3. Describe cómo el diagrama de caja te puede ayudar a comprender la diferencia en la cantidad de piruletas Tootsie Pop que las personas podrían agarrar.

4. Esta es la descripción de Jayne acerca de lo que ve en el diagrama de caja. ¿Estás de acuerdo o en desacuerdo con su descripción? Explica tu razonamiento.

"Una persona podía agarrar exactamente 42 piruletas Tootsie Pop. La cantidad de piruletas Tootsie Pop que las personas podían agarrar era realmente diferente y se dispersaba de forma casi homogénea entre 7 y 42. Aproximadamente la mitad de las personas podía agarrar más de 20 piruletas Tootsie Pop".

5. Este es un diagrama de caja diferente de los mismos datos sobre la cantidad de piruletas Tootsie Pop que podían agarrar 94 personas.



- a. ¿Por qué supones que los cinco valores son puntos separados y están rotulados?
- b. ¿Conocer estos valores de datos cambia algo acerca de tus respuestas de los ejercicios 1 a 4 anteriores?

Ejercicios 6 a 10: Velocidades máximas

En las siguientes tablas, se proporcionan las velocidades máximas de aves y de animales terrestres seleccionados.

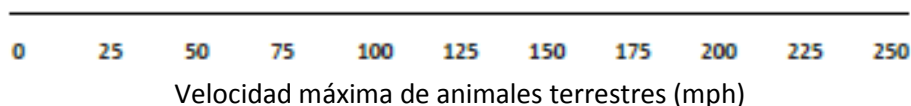
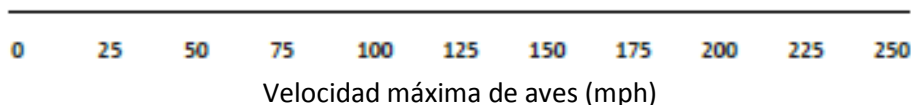
Ave	Velocidad (mph)
Halcón peregrino	242
Vencejo	120
Vencejo de cola espinosa	106
Vencejo mongol	105
Alcotán europeo	100
Paloma	100
Fregata	95
Ganso espolonado	88
Serreta mediana	80
Pato de lomo blanco	72
Colibrí de Ana	61.06
Avestruz	60

Animal terrestre	Velocidad (mph)
Guepardo	75
Murciélago cola de ratón (en vuelo)	60
Antílope americano	55
León	50
Ñu	50
Liebre	44
Perro salvaje africano	44
Canguro	45
Caballo	43.97
Gacela de Thomson	43
Galgo	43
Coyote	40
Ciervo mula	35
Oso pardo	30
Gato	30
Elefante	25
Cerdo	9

Fuentes de datos: *Natural History Magazine*, marzo de 1974, derechos de autor de 1974; The American Museum of Natural History; y James G. Doherty, curador general de The Wildlife Conservation Society; <http://www.thetravelalmanac.com/lists/animalspeed.htm>; http://en.wikipedia.org/wiki/Fastest_animals

6. Al mirar las velocidades, ¿qué te parece más interesante?
7. ¿Cuáles parecen tener la mayor variabilidad de velocidad: las aves o los animales terrestres? Explica tu razonamiento.

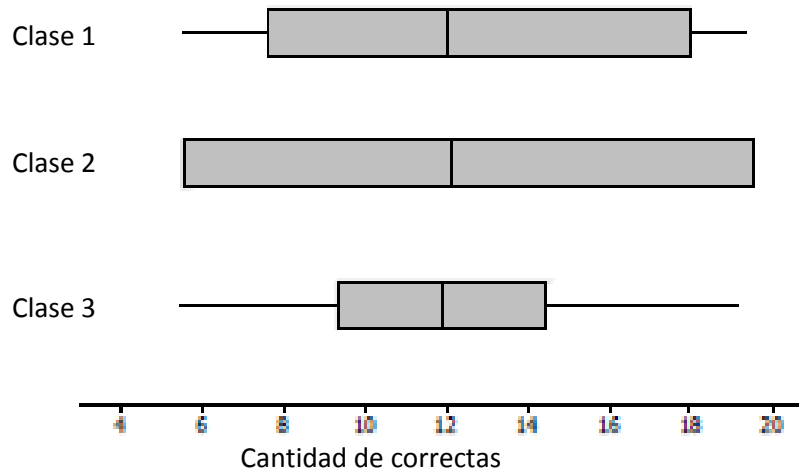
8. Encuentra el resumen de cinco números de las velocidades de cada conjunto de datos. ¿Qué te indican los resúmenes de cinco números sobre la distribución de las velocidades de cada conjunto de datos?
9. Utiliza los resúmenes de cinco números para hacer un diagrama de caja de cada uno de los dos conjuntos de datos.



10. Escribe varias oraciones para contarle a alguien sobre las velocidades de las aves y de los animales terrestres.

Ejercicios 11 a 15: ¿Qué es igual y qué es diferente?

Ten en cuenta los siguientes diagramas de caja que muestran la cantidad de preguntas (de un cuestionario de 20 preguntas) que respondieron correctamente diferentes estudiantes de tres clases distintas.

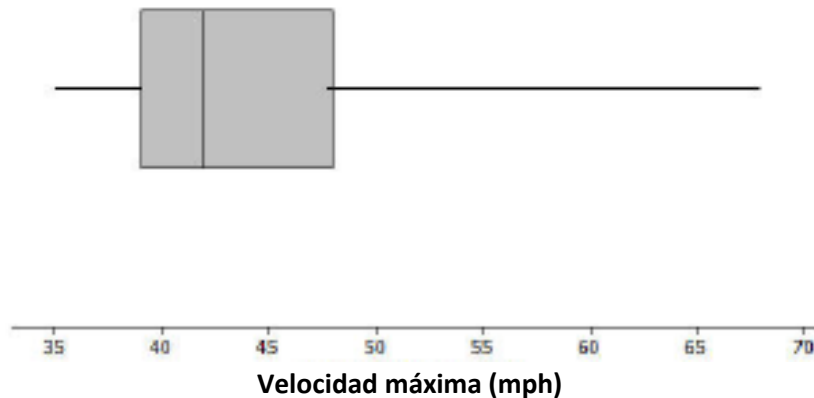


11. Describe la variabilidad de las calificaciones de las tres clases.
- 12.
- Estima el rango intercuartílico de cada uno de los tres conjuntos de calificaciones.
 - ¿Qué fracción de estudiantes representa el rango intercuartílico?
 - ¿Qué te indica el valor del R. I.C. sobre la manera en la que están distribuidas las calificaciones?

13. ¿Qué clase crees que lo hizo mejor? Asegúrate de utilizar los datos de los diagramas de caja para justificar tu respuesta.
- 14.
- Encuentra el R. I.C. de los tres conjuntos de datos de los dos primeros ejemplos: velocidad máxima de las aves, velocidad máxima de los animales terrestres y cantidad de piruletas Tootsie Pop.
 - ¿Qué conjunto de datos tenía el porcentaje más alto de valores de datos entre el cuartil inferior y el cuartil superior? Explica tu razonamiento.
15. El maestro les pidió a los estudiantes que dibujaran un diagrama de caja con un valor mínimo en 34 y un valor máximo en 64, que tuviera un rango intercuartílico de 10. Jeremy dijo que no podía dibujar solo uno porque no sabía dónde colocar el recuadro en la línea numérica. ¿Estás de acuerdo con Jeremy? ¿Por qué sí o por qué no?

Conjunto de problemas

1. El siguiente diagrama de caja resume las velocidades máximas de determinados tipos de peces.



- Estima el resumen de cinco números del diagrama de caja.
 - El pez más rápido es el pez vela, que se desplaza a 68 mph, seguido del pez espada, que lo hace a 50 mph. ¿Qué te indica esto sobre la dispersión de las velocidades de los peces en el cuarto superior del diagrama?
 - Utiliza el resumen de cinco números y el R. I.C. para describir las velocidades de los peces.
2. Supón que el rango intercuartílico de la cantidad de horas que los estudiantes pasaron jugando a los videojuegos durante la semana escolar fue de 10. ¿Qué piensas sobre cada una de las siguientes afirmaciones? Explica tu razonamiento.
- Aproximadamente la mitad de los estudiantes jugó a los videojuegos durante 10 horas en una semana escolar.
 - Todos los estudiantes jugaron a los videojuegos al menos 10 horas durante la semana escolar.
 - Aproximadamente la mitad de la clase podría haber jugado a los videojuegos entre 10 y 20 horas por semana, o entre 15 y 25 horas.
3. Supón que sabes lo siguiente de un conjunto de datos: el valor mínimo es 130, el cuartil inferior es 142, el R. I.C. es 30, la mitad de los datos son inferiores a 168, y el valor máximo es 195.
- Piensa en un contexto en el que estos números pudieran tener sentido.
 - Haz un diagrama de caja.
 - ¿Hay más valores de datos por encima o por debajo de la mediana? Explica tu razonamiento.

4. En la siguiente tabla, se proporcionan las velocidades de los perros más rápidos.

Raza	Velocidad (mph)
Galgo	45
Perro salvaje africano	44
Galgo persa	43
Whippet	36
Basenji	35
Pastor alemán	32
Braco húngaro	32
Dóberman	30

Raza	Velocidad (mph)
Lobero irlandés	30
Dálmata	30
Border Collie	30
Husky de Alaska	28
Schnauzer gigante	28
Jack Russell terrier	25
Pastor ganadero australiano	20

Fuente de datos: <http://www.vetstreet.com/our-pet-experts/meet-eight-of-the-fastest-dogs-on-the-planet>;
<http://canidaepetfood.blogspot.com/2012/08/which-dog-breeds-are-fastest.html>

- Encuentra el resumen de cinco números de este conjunto de datos y utilízalo para crear un diagrama de caja de las velocidades.
- ¿Por qué la mediana no está en el centro del recuadro?
- Escribe algunas oraciones en las que le cuentes a tu amigo acerca de la velocidad de los perros más rápidos.

Lección 16: Comprensión de diagramas de caja

Trabajo en clase

Ejercicio 1: Presidentes de la Corte Suprema

1. La Corte Suprema es el tribunal superior de los Estados Unidos y toma decisiones que afectan a todo el país. Se designa su Presidente, quien desempeñará el cargo de juez de la Corte de por vida a menos que renuncie o se enferme. Algunas personas creen que esto le da mucho tiempo para estar en la Corte Suprema. El primer Presidente de la Corte se designó en 1789.

La tabla muestra los años que desempeñaron su cargo cada uno de estos Presidentes de la Corte Suprema hasta 2013:

Nombre	Años	Designado en
John Jay	6	1789
John Rutledge	1	1795
Oliver Ellsworth	4	1796
John Marshall	34	1801
Roger Brooke Taney	28	1836
Salmon P. Chase	9	1864
Morrison R. Waite	14	1874
Melville W. Fuller	22	1888
Edward D. White	11	1910
William Howard Taft	9	1921
Charles Evens Hughes	11	1930
Harlan Friske Stone	5	1941
Fred M. Vinson	7	1946
Earl Warren	16	1953
Warren E. Burger	17	1969
William H. Rehnquist	19	1986
John G. Roberts	8	2005

Fuente de datos: http://en.wikipedia.org/wiki/List_of_Justices_of_the_Supreme_Court_of_the_United_States

Utiliza la tabla para responder lo siguiente:

- a. ¿Qué Presidente de la Corte prestó servicio durante más tiempo y cuál lo hizo durante menos tiempo? ¿Cuántos años prestó servicio cada uno de estos Presidentes de la Corte?

- b. ¿Cuál es la mediana de la cantidad de años durante los cuales estos Presidentes han prestado servicio en la Corte Suprema? Explica cómo encontraste la mediana y qué significa en términos de los datos.
- c. Haz un diagrama de caja de los años durante los cuales los jueces prestaron servicio. Describe la forma de la distribución, y la manera en que se relacionan la mediana y el R. I.C. con el diagrama de caja.
- d. ¿La mediana se encuentra a la mitad entre la menor y la mayor cantidad de años de servicio? ¿Por qué sí o por qué no?

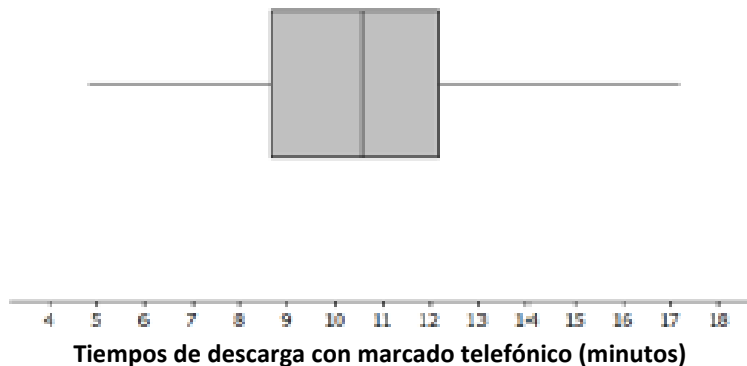
Ejercicios 2 y 3: Descarga de canciones

2. Una compañía de banda ancha registró el tiempo que le tomó descargar 232 canciones de cuatro minutos con una conexión de marcado telefónico. El siguiente diagrama de puntos muestra sus resultados.



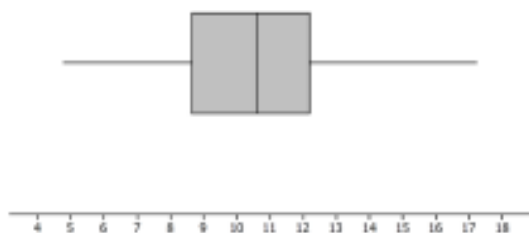
- a. ¿Qué puedes observar sobre los tiempos de descarga a partir del diagrama de puntos?
- b. ¿Es fácil indicar si 12.5 minutos se encuentra o no en el cuartil superior de los tiempos de descarga?

- c. A continuación, se muestra el diagrama de caja de los datos. Ahora, responde las partes (a) y (b) anteriores utilizando el diagrama de caja.

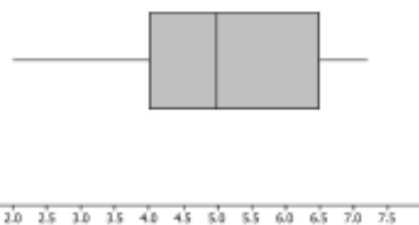


- d. ¿Cuáles son las ventajas de utilizar un diagrama de caja para mostrar un conjunto grande de datos? ¿Y las desventajas?

3. Molly presentó los siguientes diagramas para argumentar que usar una conexión de marcado telefónico sería mejor que usar una conexión de banda ancha. Ella afirmó que la conexión de marcado telefónico parecía tener menos variabilidad alrededor de la mediana, aunque el rango total parecía ser aproximadamente igual para los tiempos de descarga utilizando banda ancha. ¿Qué dirías?



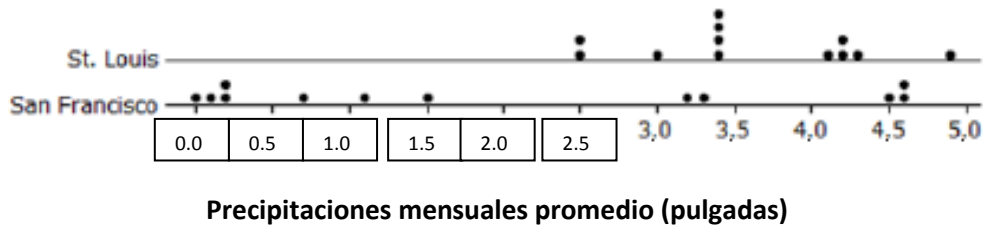
Tiempos de descarga con marcado telefónico (minutos)



Tiempos de descarga con banda ancha (segundos)

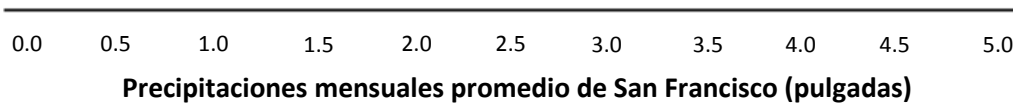
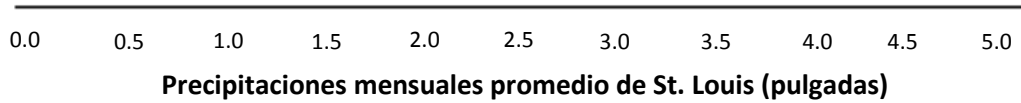
Ejercicios 4 y 5: Precipitaciones de lluvia

4. Se utilizaron los datos del promedio de precipitaciones de lluvia de cada uno de los doce meses del año para preparar los dos diagramas de puntos siguientes.



- ¿Cuántos puntos de datos hay en cada diagrama de puntos? ¿Qué representa cada punto de datos?
- Haz una conjetura acerca de qué ciudad tiene la mayor variabilidad en la cantidad mensual promedio de precipitaciones y cómo se reflejaría esto en el R. I.C. de los datos de ambas ciudades.
- Según los diagramas de puntos, ¿cuáles son los valores aproximados de los rangos intercuartílicos (R. I.C.) de la cantidad de precipitaciones mensuales promedio en pulgadas para cada ciudad? Utiliza cada R. I.C. para comparar las ciudades.
- En una lección anterior, las temperaturas mensuales promedio se redondearon al grado Fahrenheit más cercano. ¿Tendría sentido redondear la cantidad de precipitaciones a la pulgada más cercana? ¿Por qué sí o por qué no?

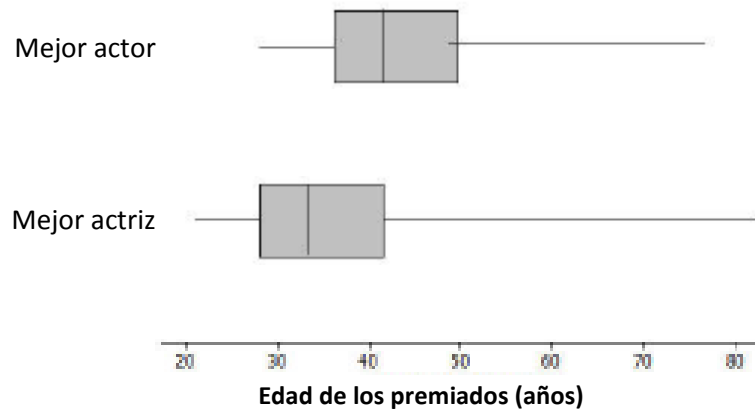
5. Utiliza los datos del ejercicio 4 para responder lo siguiente.
- a. Haz un diagrama de caja de la cantidad de precipitaciones de cada ciudad.



- b. Compara el porcentaje de los meses que tengan por encima de 2 pulgadas de precipitaciones de las dos ciudades. Explica tu razonamiento.
- c. ¿Qué diferencia hay entre los cuartos superiores de precipitaciones mensuales promedio de las dos ciudades?
- d. Describe los intervalos que contengan el 25% más pequeño de las cantidades de precipitaciones mensuales promedio de cada ciudad.
- e. Piensa en los diagramas de puntos y en los diagramas de caja. ¿Qué representación crees que te ayudará más para comprender cómo varían los datos?

Conjunto de problemas

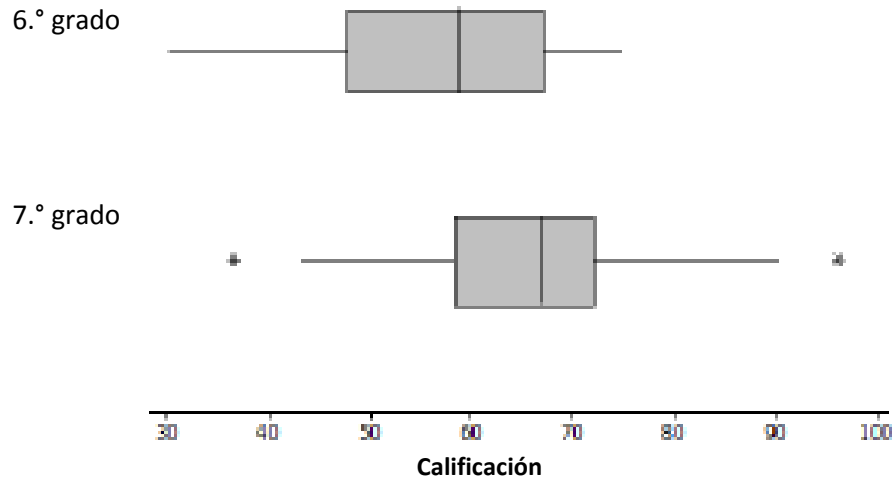
1. Los siguientes diagramas de caja resumen las edades que tenían los ganadores premiados en las categorías de mejor actriz principal y mejor actor principal de los premios de la Academia.



Fuente de datos: http://en.wikipedia.org/wiki/List_of_Best_Actor_winners_by_age_at_win
http://en.wikipedia.org/wiki/List_of_Best_Actress_winners_by_age_at_win

- ¿Crees que es más difícil para una mujer mayor ganar un premio de la Academia a la mejor actriz que para un hombre mayor ganar el premio al mejor actor? ¿Por qué sí o por qué no?
- La mujer de más edad que ganó un premio de la Academia fue Jessica Tandy en 1990 por *Paseando a Miss Daisy*. El actor de más edad fue Henry Fonda por *En el estanque dorado* en 1982. ¿Qué edad tenían cuando ganaron el premio? ¿Cómo lo sabes? ¿Eran mucho más mayores que la mayoría de los otros ganadores?
- El actor ganador de 2013 fue Daniel Day-Lewis por *Lincoln*. Tenía 55 años en ese momento. ¿Qué puedes decir sobre el porcentaje de ganadores masculinos que eran mayores que Daniel Day-Lewis cuando ganaron los óscaros?
- Utiliza la información que puedes ver en los diagramas de caja para escribir un párrafo que respalde o refute la afirmación de que menos actrices mayores ganan premios de la Academia en comparación con actores mayores.

2. En los siguientes diagramas de caja, se resumen las calificaciones que obtuvieron los estudiantes de sexto y de séptimo grado en un examen sobre los polígonos y sus características.



- ¿En qué grado los estudiantes obtuvieron mejores calificaciones? Explica cómo lo sabes.
 - ¿Por qué crees que dos de los valores de datos de séptimo grado no son parte de los segmentos?
 - ¿Qué diferencia hay entre las medianas de las calificaciones de los dos grados? ¿Te sorprende esto? ¿Por qué sí o por qué no?
 - ¿Qué diferencia hay entre los R. I.C. de los dos grados?
3. Una fórmula del R. I.C. se podría escribir como $C3 - C1 = R. I.C.$ Supón que sabes el R. I.C. y el C1. ¿Cómo podrías encontrar el C3?
4. Ten en cuenta esta afirmación: "Históricamente, la duración promedio del servicio de un Presidente de la Corte era de menos de 15 años; sin embargo, desde 1970, la duración promedio aumentó". Utiliza los datos proporcionados en el ejercicio 1 para responder las siguientes preguntas.
- ¿Estás de acuerdo o en desacuerdo con la afirmación? Explica tu razonamiento.
 - ¿Cambiaría tu respuesta si utilizaras la mediana de la cantidad de años en lugar de la media?

Lección 17: Desarrollo de un proyecto estadístico

Trabajo en clase

Desafío de exploración

Revisión de preguntas estadísticas

Las preguntas estadísticas que investigaste en este módulo incluyen estas preguntas:

- ¿Cuántas horas suelen dormir los estudiantes de sexto grado durante la semana escolar?
- ¿Cuál es la cantidad típica de libros que lee un estudiante de sexto grado durante 6 meses?
- ¿Cuál es la frecuencia cardíaca típica de un estudiante de una clase de sexto grado?
- ¿Cuántas horas pasa típicamente un estudiante de sexto grado practicando un deporte o jugando al aire libre?
- ¿Cuáles son las circunferencias de las cabezas de los adultos interesados en comprar gorras de béisbol?
- ¿Cuál es la duración de las baterías de una determinada marca?
- ¿Cuántas mascotas tienen los estudiantes?
- ¿Cuánto tiempo le toma a un estudiante llegar a la escuela?
- ¿Cuál es una temperatura diaria típica de la ciudad de Nueva York?
- ¿Cuál es el peso típico de una mochila de los estudiantes de una determinada escuela?
- ¿Cuál es la cantidad típica de papas fritas que hay en un pedido grande de un restaurante de comida rápida?
- ¿Cuál es la cantidad típica de minutos que pasa un estudiante haciendo la tarea cada día?
- ¿Cuál es la altura típica de un salto vertical de un jugador de la NBA?

¿Qué tienen estas preguntas en común?

¿Por qué varias de estas preguntas incluyen la palabra "típico" o "típica"?

Recuerda de la primera lección de este módulo, que una pregunta estadística es una pregunta que se responde con datos que sabes que variarán.

Repasemos los pasos de una investigación estadística.

Paso 1: Formula una pregunta que se pueda responder con datos.

Paso 2: Recaba los datos adecuados.

Paso 3: Resume los datos con gráficos y con resúmenes numéricos.

Paso 4: Responde la pregunta formulada en el paso 1 con resúmenes numéricos y con gráficos.

El primer paso es formular una pregunta estadística. Selecciona una de las preguntas y escríbela en la siguiente Plantilla de revisión de estudio estadístico.

El segundo paso es recabar los datos. En todas estas investigaciones, se te proporcionaron datos. ¿Cómo crees que se recabaron los datos para la pregunta que seleccionaste en el paso 1? Escribe tu respuesta del paso 2 en el siguiente resumen.

El tercer paso está relacionado con las diferentes maneras en las que resumes los datos. Enumera las diferentes maneras en las que resumiste los datos del paso 3.

Plantilla de revisión de estudio estadístico

Paso 1: Formula una pregunta estadística.

Paso 2: Recaba los datos.

Paso 3: Resume los datos.

Paso 4: Responde la pregunta.

Ahora es tu turno para responder una pregunta estadística a partir de los datos que recabaste. Antes de recabar datos, estudia posibles preguntas estadísticas. Para cada pregunta, indica los datos que recabarías y resumirías para responderla. También, indica cómo planeas recabar los datos.

Piensa en preguntas que se podrían responder con datos recabados de los integrantes de tu clase o de tu escuela, o con datos que se podrían recabar de sitios web reconocidos (por ejemplo, la Asociación de Estadística de Estados Unidos y el proyecto "Censo en la escuela"). Tu maestro necesitará aprobar tanto tu pregunta como tu plan para recabar los datos antes de que lo hagas.

Como clase, estudien las posibilidades de una investigación estadística. Anota algunas de las ideas que se compartieron en tu clase utilizando la siguiente tabla.

Posibles preguntas estadísticas	¿Qué datos se recabarían y cómo?

Después de conversar acerca de varias de las posibilidades anteriores de un proyecto estadístico, prepara una pregunta estadística y un plan para recabar datos para presentárselo a tu maestro. Después de que tu maestro apruebe tu pregunta y tu plan para recabar datos, comienza a obtenerlos. Organiza con atención tus datos a medida que comienzas a desarrollar los resúmenes para responder tu pregunta estadística. En las próximas lecciones, se te pedirá que comiences a hacer un cartel o un bosquejo de una presentación para compartir con tu maestro y con otros integrantes de tu clase.

Resumen de la lección

Un estudio estadístico implica un proceso de investigación de cuatro pasos:

- Formular preguntas que se puedan responder con datos.
- Diseñar un plan para recabar los datos adecuados y, luego, utilizar el plan para hacerlo.
- Analizar los datos.
- Interpretar los resultados y llegar a conclusiones válidas a partir de los datos de la pregunta formulada.

Conjunto de problemas

Tu maestro indicará los pasos que se espera que completes en los próximos días para desarrollar este proyecto. Ten en cuenta que el primer paso para desarrollar tu proyecto es una pregunta estadística. Con una de las preguntas estadísticas formuladas en esta lección o con una nueva pregunta desarrollada en esta lección, organiza tu pregunta, y planifica cómo recabar y resumir los datos. Completa el proceso según las instrucciones de tu maestro.

Lección 18: Conexión de representaciones gráficas y resúmenes numéricos

Trabajo en clase

Puede ser difícil comprender un conjunto de datos con solo mirar los datos sin procesar. Imagina que el proyecto de Joaquin trata sobre el peso típico de los osos. Él encontró un artículo sobre osos que le proporcionó una lista desordenada de los pesos de 250 osos, sin resúmenes numéricos ni gráficos de estos datos. Sería muy difícil llegar rápidamente a conclusiones. Joaquin decidió diseñar su proyecto utilizando estos datos.

Imagina que el artículo te proporciona esta afirmación: "La mediana del peso de los osos estudiados era de 305 libras". Esto es útil pero, a veces, incluso los resúmenes numéricos por sí solos no pueden transmitir por completo los aspectos interesantes de una distribución de datos. Por lo general, los lectores quieren tener un resumen conciso y útil de la información que sea tanto numérico como visual.

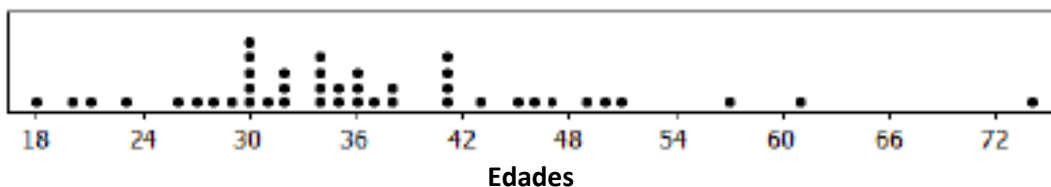
Ejemplo 1: Información de resumen a partir de gráficos

Este es un conjunto de datos de las edades (en años) de 43 participantes de una carrera local reciente de 5 kilómetros.

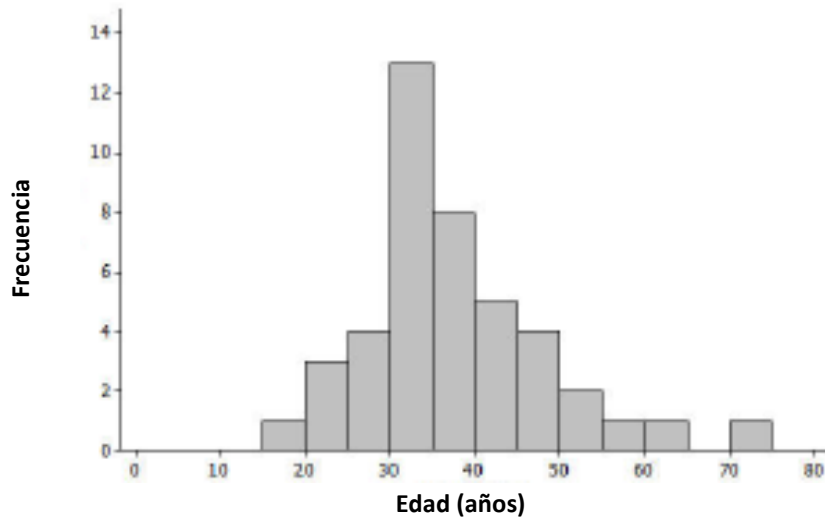
20	30	30	35	36	34	38	46
45	18	43	23	47	27	21	30
32	32	31	32	36	74	41	41
51	61	50	34	34	34	35	28
57	26	29	49	41	36	37	41
38	30	30					

Estas son algunas estadísticas de resumen, un histograma y un diagrama de puntos de los datos:

Mínimo = 18; C1 = 30; Mediana = 35; C3 = 41; Máximo = 74; Media = 36.81; DMA = 8.1



Histograma de edades de los participantes

**Ejercicios 1 a 7**

1. Según el histograma, ¿describirías la forma de la distribución de datos como aproximadamente simétrica o como asimétrica? ¿Hubieras llegado a la misma conclusión mirando al diagrama de puntos?
2. ¿Es más fácil ver la forma de la distribución de datos del histograma o del diagrama de puntos?
3. ¿Qué es lo que puedes ver en el diagrama de puntos que no es tan fácil ver en el histograma?

4. ¿El diagrama de puntos y el histograma parecen estar centrados aproximadamente en el mismo lugar?

5. ¿Tanto el diagrama de puntos como el histograma ofrecen información sobre la variabilidad en la distribución de edades?

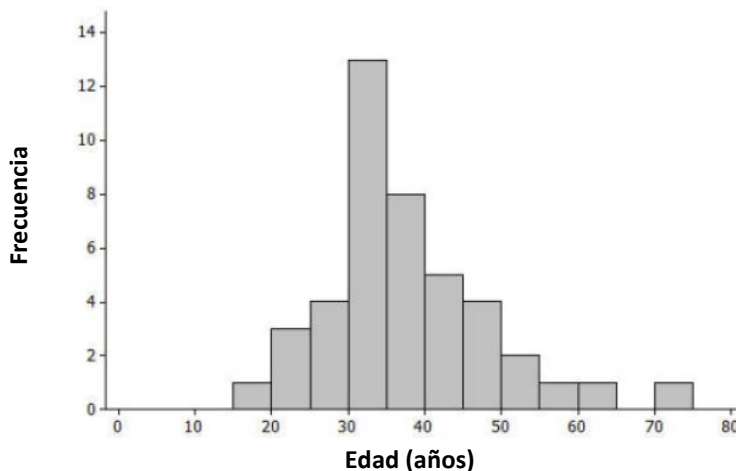
6. Si no tuvieras el conjunto de datos original, y solo tuvieras el diagrama de puntos y el histograma, ¿podrías encontrar el valor de la mediana de la edad a partir del diagrama de puntos?

7. Explica por qué podrías solamente estimar el valor de la mediana si solo tuvieras un histograma de los datos.

Ejercicios 8 a 13: Gráficos y resúmenes numéricos

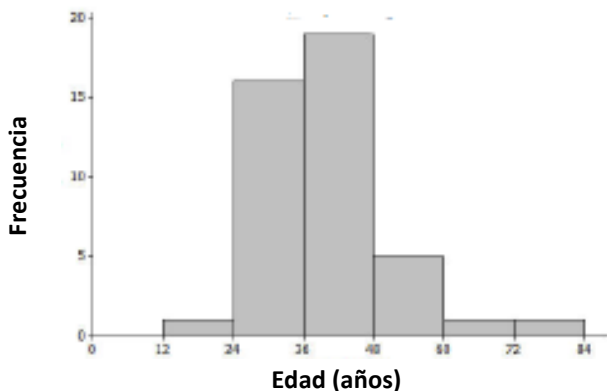
8. Supón que se escribió un artículo en el periódico sobre la carrera y que en él se mostraba el histograma de las edades del ejemplo 1. El escritor afirmó: "La carrera atrajo a más corredores mayores este año, la mediana de la edad fue de 45". Explica cómo sabríamos que esta afirmación es incorrecta con tan solo mirar el histograma.

Histograma de edades de los participantes de una carrera de 5 kilómetros

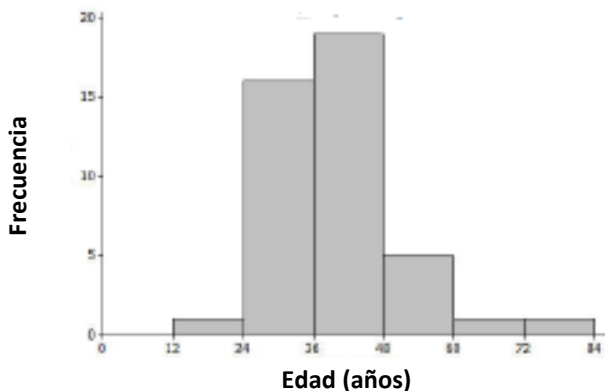


9. Uno de los siguientes histogramas es otro histograma válido de las edades de los corredores. Selecciona el histograma correcto y explica cómo determinaste qué gráfico es válido (y cuál es incorrecto) según las medidas de resumen y el diagrama de puntos.

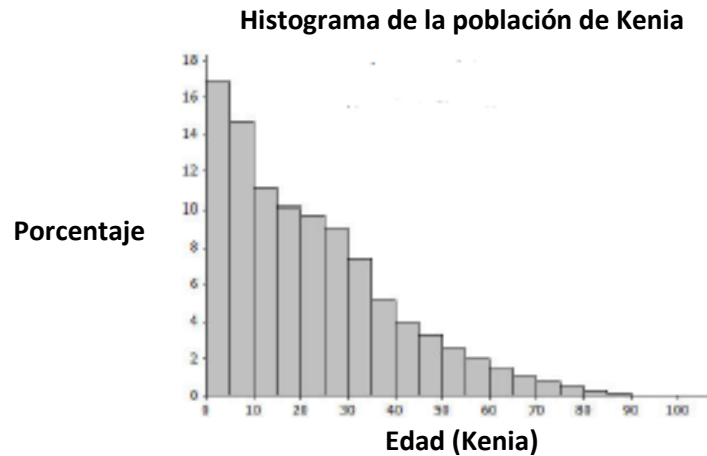
Histograma de edades de los participantes de una carrera de 5 kilómetros



Histograma de edades de los participantes de una carrera de 5 kilómetros



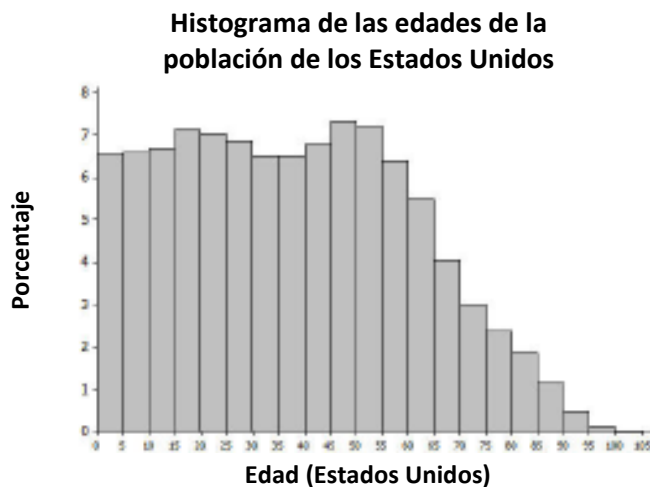
10. El siguiente histograma representa la distribución de las edades de la población de Kenia en 2010.



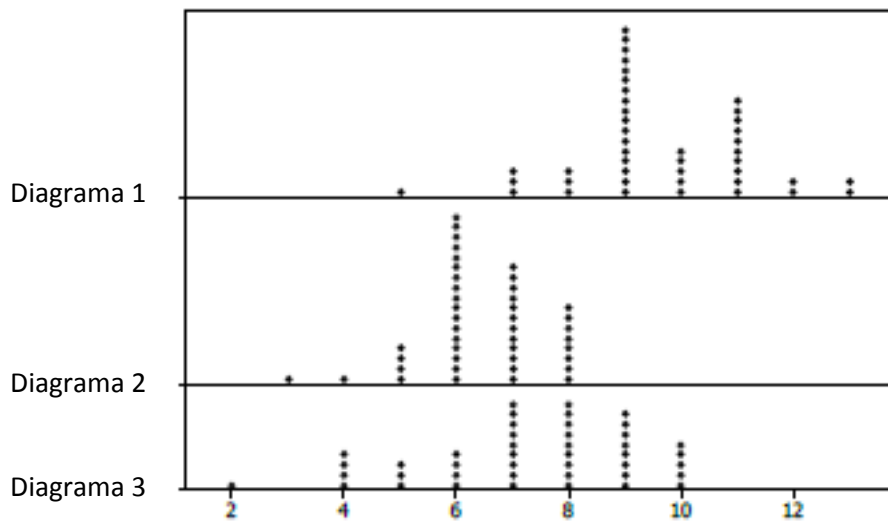
a. ¿Cómo sabemos, a partir del gráfico anterior, que el primer cuartil (C1) de esta distribución de edades está entre los 5 y los 9 años?

b. Alguien cree que la mediana de la edad está cerca de 30. Explica cómo el gráfico respalda esta creencia o explica por qué el gráfico no lo hace.

11. El siguiente histograma representa la distribución de las edades de la población de los Estados Unidos en 2010. Según el histograma, ¿cuál de los siguientes rangos crees que incluye la mediana de la edad en los Estados Unidos: 20–29, 30–39 o 40–49? ¿Por qué?



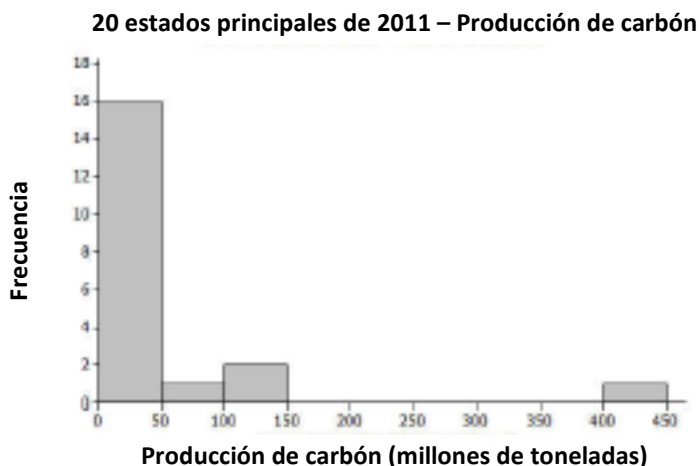
12. Utiliza los histogramas de los ejercicios 11 y 12 para responder lo siguiente:
- ¿La distribución de edades de qué país (Kenia o Estados Unidos) tiene un tercer cuartil en los 50? ¿Cómo lo determinaste?
 - Si alguien creyera que la edad promedio de una persona que vive en los Estados Unidos es mayor que la edad promedio de una persona que vive en Kenia, ¿cómo podrías justificar esa afirmación mediante la comparación de los histogramas?
13. Une los siguientes conjuntos de medidas de resumen con el diagrama de puntos correspondiente. Solo **UN** diagrama de puntos coincide con cada conjunto de medidas de resumen. Explica por qué seleccionaste el diagrama de puntos o por qué los otros diagramas de puntos no representarían las medidas de resumen. Nota: se utiliza la misma escala en cada diagrama de puntos.



- Mediana = 8 y R. I.C. = 3 Gráfico _____
- Media = 9.6 y DMA = 1.28 Gráfico _____
- Mediana = 6 y Rango = 5 Gráfico _____

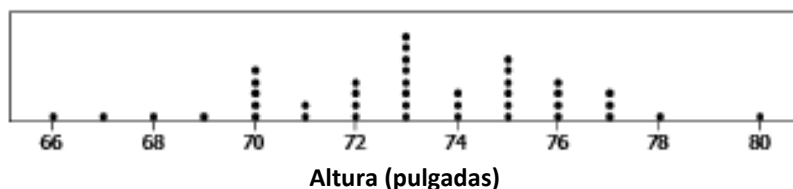
Conjunto de problemas

1. El siguiente histograma muestra la cantidad de carbón producido (por estado) en los 20 estados que más produjeron carbón en 2011. Muchos de estos estados produjeron menos de 50 millones de toneladas de carbón, pero un estado produjo más de 400 millones de toneladas (Wyoming). Según el histograma, ¿cuál de los tres conjuntos de medidas de resumen podría coincidir con el gráfico? Para cada opción que elimines, enumera al menos un motivo para hacerlo.



Fuente: datos de la producción de carbón por estado en los Estados Unidos de acuerdo con lo que informa la Asociación Nacional de Minería en http://www.nma.org/pdf/c_production_state_rank.pdf, sitio web al que se accedió el 5 de mayo de 2013.

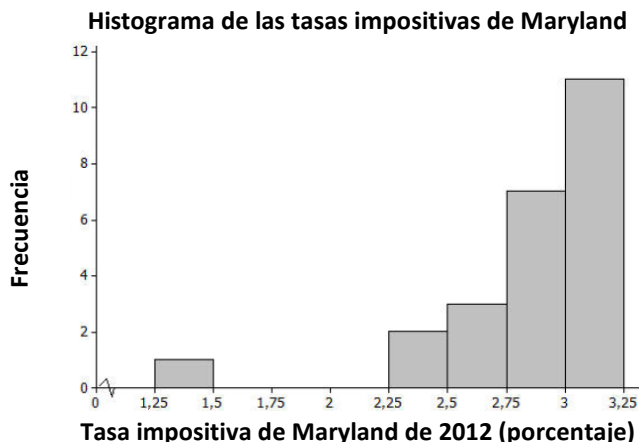
- Mínimo = 1; C1 = 12; Mediana = 36; C3 = 57; Máximo = 410; Media = 33; DMA = 2.76
 - Mínimo = 2; C1 = 13.5; Mediana = 27.5; C3 = 44; Máximo = 439; Media = 54.6; DMA = 52.36
 - Mínimo = 10; C1 = 37.5; Mediana = 62; C3 = 105; Máximo = 439; Media = 54.6; DMA = 52.36
2. En el siguiente diagrama de puntos, se muestran las alturas (redondeadas a la pulgada más cercana) de los 41 miembros del equipo masculino de natación y de sato de la Universidad de Texas de 2012-2013.



Fuente: <http://www.texassports.com/sports/m-swim/mtt/tex-m-swim-mtt.html>, sitio web al que se accedió el 30 de abril de 2013.

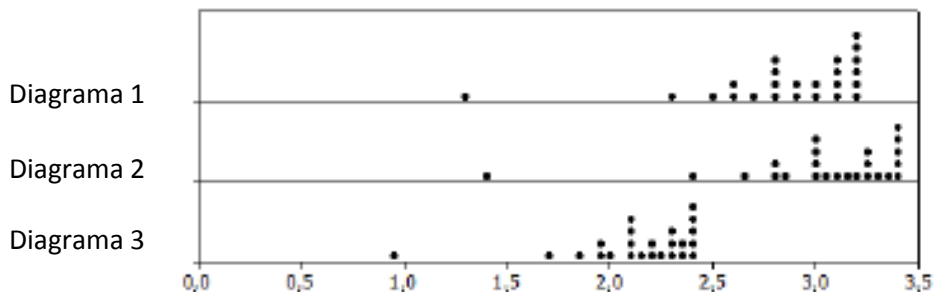
- Utiliza el diagrama de puntos para determinar el resumen de cinco números (mínimo, cuartil inferior, mediana, cuartil superior y máximo) del conjunto de datos.
- De acuerdo con este diagrama de puntos, haz un histograma de las alturas utilizando los siguientes intervalos: $66 \leq 68$ pulgadas, $68 \leq 70$ pulgadas y así sucesivamente.

3. Según el sitio web del Interventor de Maryland, "en los 23 condados de Maryland y en la ciudad de Baltimore se grava un impuesto local sobre la renta [...] Los funcionarios locales fijan las tasas, que varían entre 1.25% y 3.20% para el año fiscal en curso (2012)". A continuación, aparece un histograma de las 24 tasas impositivas (en porcentajes).



Fuente: <http://taxes.marylandtaxes.com>, sitio web al que se accedió el 5 de mayo de 2013.

¿Cuál de los tres diagramas de puntos coincide con el histograma de "Tasas impositivas de Maryland de 2012" anterior? Explica cómo determinaste el diagrama de puntos correcto.



4. Para cada uno de los siguientes cinco conjuntos de medidas de resumen, indica si el conjunto podría coincidir con el histograma de "Tasas impositivas de Maryland de 2012" anterior. Para cada conjunto de medidas de resumen que eliminaste, explica por qué eliminaste esa opción.
- Media = 1.01; DMA = 5.4
 - Mediana = 2.93; R. I. = 0.45
 - Media = 3.5; DMA = 1.1
 - Mediana = 3.10; R. I. = 2.15
 - Mínimo = 1.25; Máximo = 3.20

Lección 19: Comparación de distribuciones de datos

Trabajo en clase

Como has visto en lecciones anteriores, puede ser difícil comprender un conjunto de datos con solo mirar los datos sin procesar. Por lo general, los lectores quieren tener un resumen conciso y útil.

Esto se vuelve extremadamente importante cuando las distribuciones de datos se comparan entre sí. Si bien un lector puede estar interesado en saber si un oso polar macho adulto de Alaska es más grande que un oso pardo macho adulto típico de Columbia Británica, también sería útil poder comparar la variabilidad y la forma de las distribuciones de estos dos grupos de osos. Con los gráficos de resumen de las dos distribuciones colocadas de manera contigua, puedes acceder a las características de una distribución y compararlas con las de otra distribución más fácilmente.

A estas alturas, deberías haber terminado de recabar los datos para tu pregunta estadística. Esta lección proporcionará representaciones gráficas de distribuciones de datos que son parte de los resúmenes que se espera ver en tu proyecto.

Ejemplo 1: Comparación de grupos utilizando diagramas de caja

Recuerda que un *diagrama de caja* es una representación visual de un resumen de cinco números. Está dibujado con una referencia detallada a una línea numérica, por lo que la diferencia entre dos valores cualesquiera del resumen de cinco números se representa visualmente como una distancia. Por ejemplo, el recuadro de un diagrama de caja se dibuja de manera que el ancho del recuadro represente el R. I.C. Los bigotes (las líneas que se extienden desde el recuadro) se dibujan de manera tal que la distancia desde el extremo de un bigote hasta el extremo del otro bigote represente el rango. Si dos diagramas de caja (que representan distribuciones diferentes) se dibujaran de manera contigua utilizando la misma escala, se podrían comparar rápidamente los R. I.C. y los rangos de las dos distribuciones, y también obtener una idea de los valores del resumen de cinco números de cada distribución.

Este es un conjunto de datos de las edades de 43 participantes de una carrera local de 5 kilómetros (que se muestra en una lección anterior).

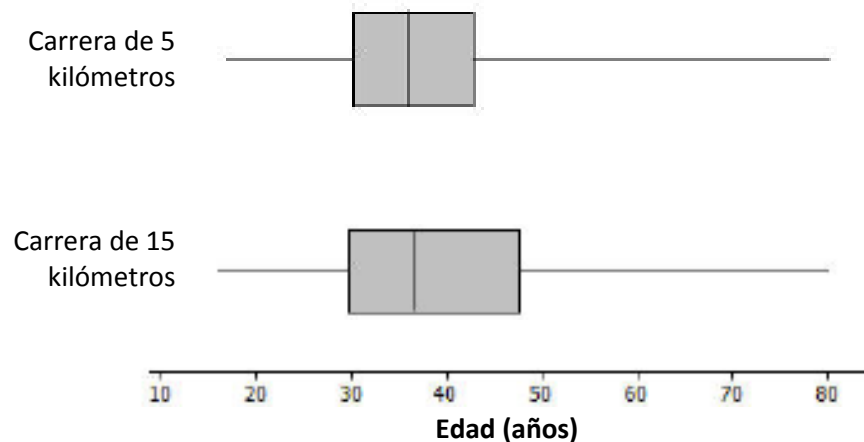
20	30	30	35	36	34	38	46
45	18	43	23	47	27	21	30
32	32	31	32	36	74	41	41
51	61	50	34	34	34	35	28
57	26	29	49	41	36	37	41
38	30	30					

Este es el resumen de cinco números de los datos: Mínimo = 18; C1 = 30; Mediana = 35; C3 = 41; Máximo = 74.

Más adelante ese año, también se llevó a cabo una carrera de 15 kilómetros en la misma ciudad. A continuación, aparecen las edades de los 55 participantes de esa carrera.

47	19	30	30	36	37	35	39
19	49	47	16	45	22	50	27
19	20	30	32	32	31	32	37
22	81	43	43	54	66	53	35
22	35	35	36	28	61	26	29
38	52	43	37	38	43	39	30
58	30	48	49	54	56	58	

¿La carrera más larga parece atraer a corredores diferentes en términos de edad? Estos son diagramas de caja contiguos que pueden ayudarte a responder esa pregunta. Los diagramas de caja contiguos son dos o más diagramas de caja dibujados utilizando la misma escala.

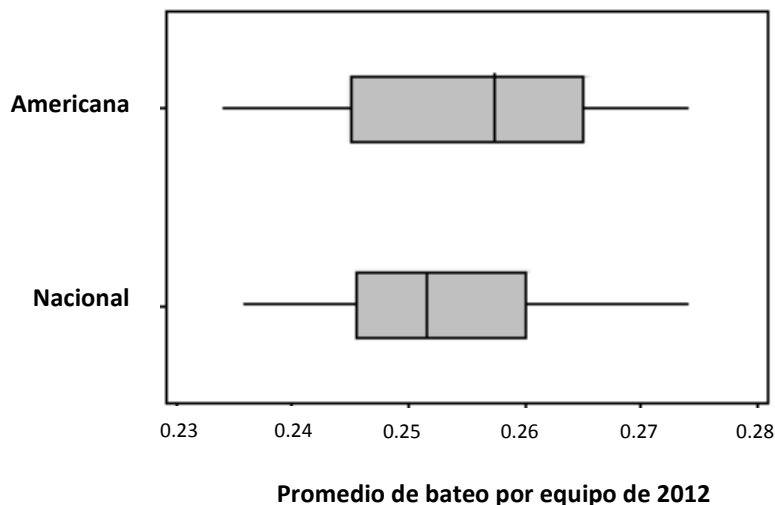


Ejercicios 1 a 6

1. Según los diagramas de caja contiguos, estima el resumen de cinco números del conjunto de datos de la carrera de 15 kilómetros.
2. ¿Los dos conjuntos de datos tienen la misma mediana? Si no es así, ¿qué carrera tuvo la mediana de edad más alta?
3. ¿Los dos conjuntos de datos tienen el mismo R. I.C.? Si no es así, ¿qué distribución tiene la mayor dispersión en el 50% central?
4. ¿Qué carrera tuvo el rango total de edades más pequeño? ¿Cuál crees que es el rango de edades de la carrera de 15 kilómetros?
5. ¿En qué carrera se encontraba el participante de más edad? ¿Qué edad tenía este participante aproximadamente?
6. Ahora, ten en cuenta solo el 25% más joven de los participantes de la carrera de 15 kilómetros. ¿Qué edad tenía el corredor más joven de este grupo? ¿Qué edad tenía el corredor de más edad de este grupo? ¿Qué diferencia hay con la carrera de 5 kilómetros?

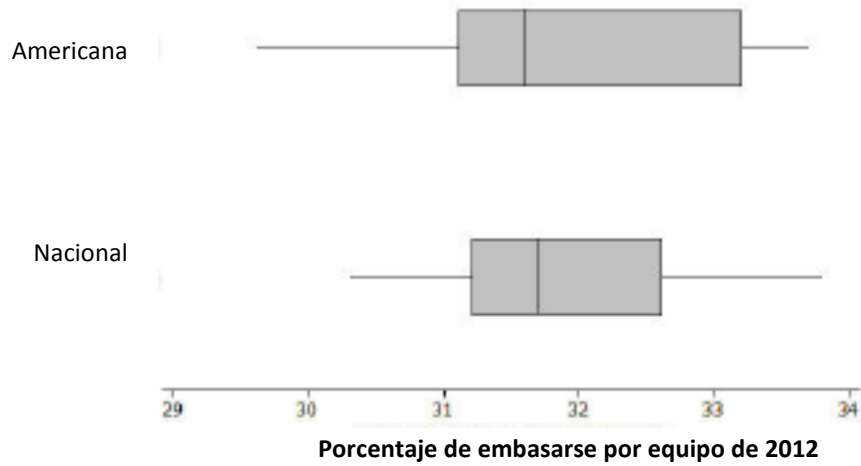
Ejercicios 7 a 12: Comparación de diagramas de caja

En 2012, las Grandes Ligas de Béisbol estaban conformadas por dos ligas: una Liga Americana de 14 equipos y una Liga Nacional de 16 equipos. Se cree que los equipos de la Liga Americana, generalmente, tendrían valores más altos de determinadas estadísticas ofensivas, como el promedio de bateo y el porcentaje de embasarse. (Los equipos quieren tener valores altos en estas estadísticas). Utiliza los siguientes diagramas de caja contiguos para investigar estas afirmaciones. (Fuente: <http://mlb.mlb.com/stats/sortable.jsp>, sitio web al que se accedió el 13 de mayo de 2013).



- ¿El promedio de bateo más alto por equipo de la Liga Americana era muy diferente al promedio de la Liga Nacional? ¿Aproximadamente qué tan grande era la diferencia y qué liga tenía el valor máximo más alto?
- ¿El rango de promedios de bateo por equipo de la Liga Americana era muy diferente o solo un poco diferente con respecto al de la Liga Nacional?
- ¿Qué liga tenía la mediana más alta del promedio de bateo por equipo? Dada la escala del gráfico y el rango de los conjuntos de datos, ¿la diferencia entre los valores de las medianas de las dos ligas parece ser pequeña o grande? Explica por qué piensas que es pequeña o grande.

10. De acuerdo con los siguientes diagramas de caja del porcentaje de embasarse, ¿qué tres valores de resumen (del resumen de cinco números) parecen ser iguales o prácticamente iguales para ambas ligas?



11. ¿El conjunto de datos de qué liga parece tener menos variabilidad? Explica tu respuesta.
12. Responde la afirmación original: "Se cree que los equipos de la Liga Americana, generalmente, tendrían valores más altos de [...] porcentaje de embasarse". Teniendo en cuenta los gráficos anteriores, ¿estás de acuerdo o en desacuerdo? Explica.

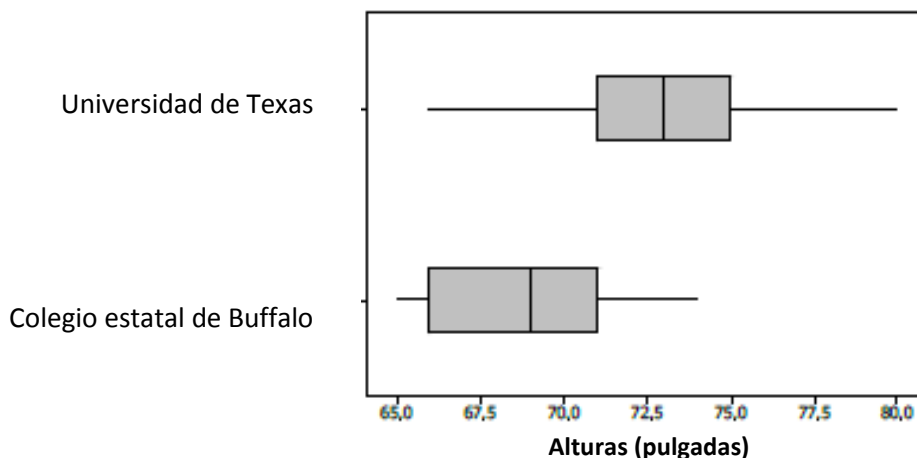
Resumen de la lección

Al comparar la distribución de una variable cuantitativa de dos o más grupos distintos, es útil mostrar gráficos de las distribuciones de los grupos de manera contigua utilizando la misma escala. Por lo general, puedes observar, cuantificar y describir las similitudes y las diferencias de las distribuciones de los grupos con mayor facilidad.

Conjunto de problemas

- Los programas escolares de atletismo se separan en divisiones según el tamaño de la escuela, las becas disponibles de atletismo y otros factores. Una investigadora quiere saber si los integrantes de los programas de natación y de salto de la División I (escuelas que ofrecen becas de atletismo y tienden a tener una gran cantidad de inscritos) suelen ser más altos que los nadadores y saltadores de los programas de la División III (escuelas que no ofrecen becas de atletismo y tienden a tener una menor cantidad de inscritos). Para comenzar con la investigación, la investigadora hace diagramas de caja contiguos de las alturas (en pulgadas) de los integrantes del equipo masculino de natación y de salto de la Universidad de Texas de 2012-2013 (un programa de la División I) y de las alturas (en pulgadas) de los integrantes del equipo masculino de natación y de salto del Colegio Estatal de Buffalo (un programa de la División III).

Fuente: <http://www.texasports.com/sports/m-swim/mtt/tex-m-swim-mtt.html>, sitio web al que se accedió el 30 de abril de 2013, donde se enumeraban todas las alturas de los 41 integrantes, y <http://www.buffalostateathletics.com/roster.aspx?path=mswim&>, sitio web al que se accedió el 15 de mayo de 2013, donde solo se enumeraban 11 integrantes y 10 alturas.

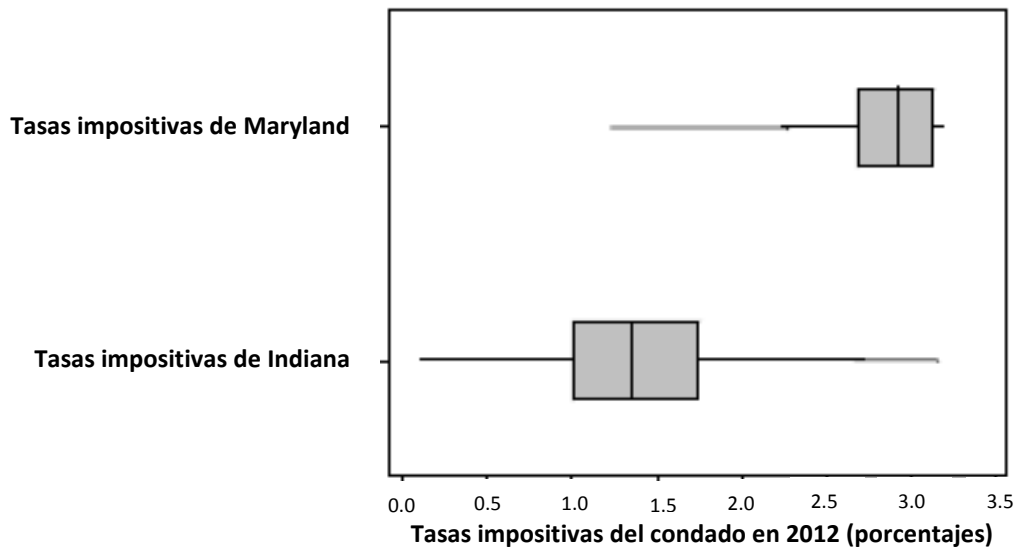


- ¿Qué conjunto de datos tiene el rango más pequeño?
- Verdadero o falso: un integrante del equipo de la Universidad de Texas con una altura igual a la mediana sería más alto que un miembro del equipo del Colegio Estatal de Buffalo con una altura igual a la mediana.
- Para ser exhaustiva, la investigadora examinará muchos otros programas escolares de deportes para investigar más en profundidad su afirmación de que los integrantes de los programas de natación y de salto de la División I suelen ser más altos que los nadadores y los saltadores de la División III. Sin embargo, dado el gráfico anterior, en su etapa inicial de investigación, ¿crees que

su afirmación podría ser válida? Justifica cuidadosamente tu respuesta utilizando medidas comparativas de resumen o atributos gráficos.

2. Diferentes estados utilizan distintos métodos para determinar el impuesto sobre la renta de una persona. Sin embargo, tanto Maryland como Indiana tienen sistemas en los que una persona paga una tasa impositiva sobre la renta diferente según el condado en el que viva. A continuación, se muestran los diagramas de caja que resumen las 24 tasas impositivas diferentes del condado que se aplican en los 23 condados de Maryland y en la ciudad de Baltimore (la que se considera como un condado en este caso) y las tasas impositivas para residentes de 91 condados de Indiana en 2012.

Fuente: http://taxes.marylandtaxes.com/Individual_Taxes/Individual_Tax_Types/Income_Tax/Tax_Information/Tax_Rates/Local_and_County_Tax_Rates.shtml, sitio web al que se accedió el 5 de mayo de 2013, y www.in.gov/dor/files/12-county-rates.pdf, sitio web al que se accedió el 16 de mayo de 2013.

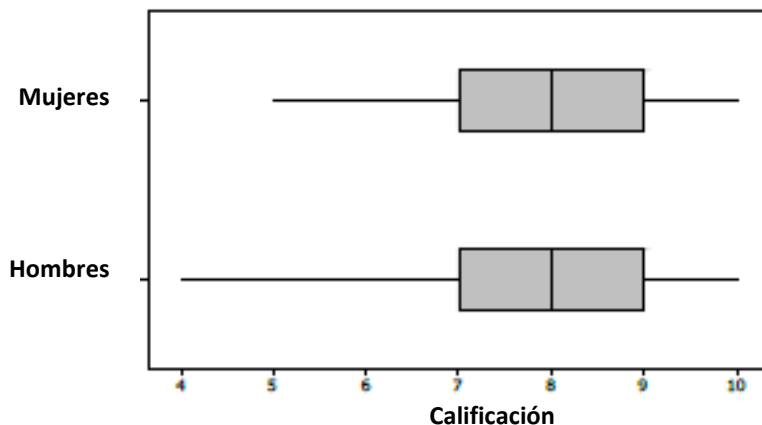


- Verdadero o falso: al menos una tasa impositiva sobre la renta del condado de Indiana es más alta que la mediana de la tasa impositiva sobre la renta del condado de Maryland. Explica cómo lo sabes.
- Verdadero o falso: las 24 tasas impositivas sobre la renta del condado de Maryland tienen menos variabilidad que las 91 tasas impositivas sobre la renta del condado de Indiana. Explica cómo lo sabes.
- ¿Qué estado parece tener tasas impositivas sobre la renta del condado típicamente menores? Explica tu respuesta.

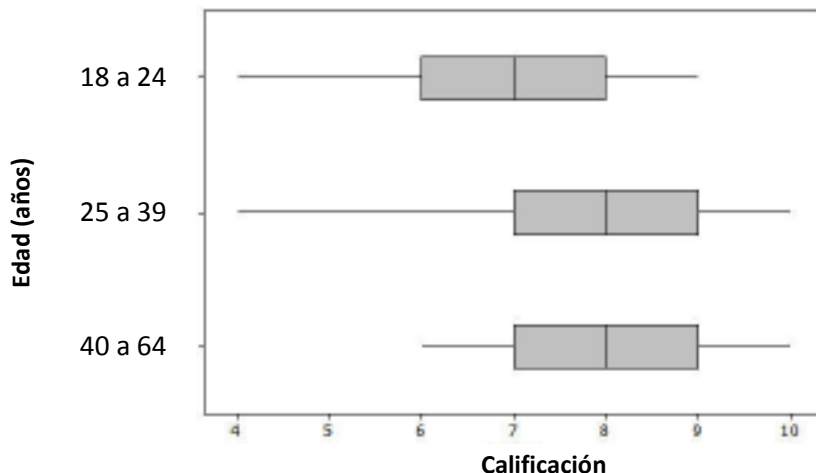
3. Varios estudios de cine confían mucho en los datos sobre clientes de los mercados de prueba a fin de determinar cómo se promocionará y distribuirá una película. Recientemente, se mostraron los avances o pre-estrenos de una película que se estrenará pronto ante 300 personas. Se le pidió a cada persona que calificara la película en una escala del 0 al 10, en la que 10 representa "la mejor película que he visto" y 0 representa "la peor película que he visto".

A continuación, hay algunos diagramas de caja contiguos que resumen las calificaciones según determinadas características demográficas.

Para 150 mujeres y 150 hombres:



Para 3 grupos distintos de edades:



- Por lo general, ¿parece que los hombres y las mujeres calificaron la película de manera similar o de manera muy diferente? Escribe algunas oraciones en las que expliques tu respuesta utilizando información comparativa del gráfico sobre el centro y la dispersión.
- En general, parece que la película típicamente recibió mejores calificaciones por parte de los integrantes mayores del grupo. Escribe algunas oraciones utilizando medidas comparativas del centro y de la dispersión o atributos gráficos para justificar esta afirmación.

Lección 20: Descripción de centro, variabilidad y forma de una distribución de datos a partir de una representación gráfica

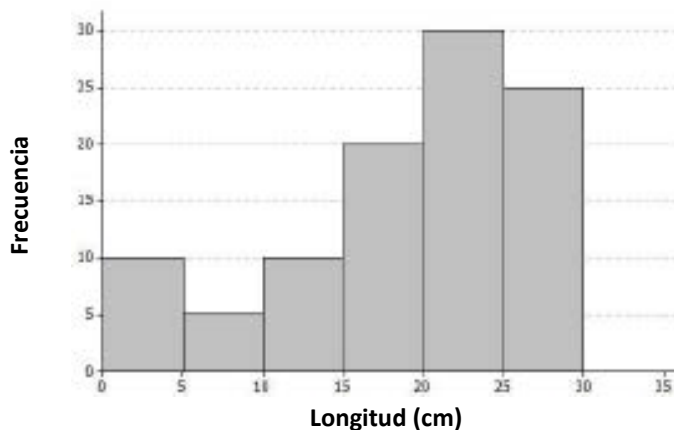
Trabajo en clase

La perca flavescens de los Grandes Lagos es un pez que vive en los cinco Grandes Lagos y en muchos otros lagos de las regiones oriental y superior de los Grandes Lagos de los Estados Unidos y de Canadá. Ambos países participan activamente de actividades que tienen el fin de mantener una población saludable de percas en estos lagos.

Ejemplo 1: La perca flavescens de los Grandes Lagos

Los científicos recabaron datos de muchas muestras de percas flavescens porque estaban preocupados por su supervivencia. ¿Qué datos crees que los investigadores podrían querer recabar acerca de la perca?

Los científicos capturaron percas flavescens de un lago de esta región. Registraron datos sobre cada pez y luego los devolvieron al lago. Ten en cuenta el siguiente histograma de datos sobre la longitud (en centímetros) de una muestra de percas flavescens.



Ejercicios 1 a 11

Los científicos se preocupaban por la supervivencia de la perca flavescens a medida que estudiaban el histograma.

1. ¿Qué pregunta estadística se podría responder según esta distribución de datos? ¿Cómo crees que los científicos recabaron estos datos?

2. Utiliza el histograma para completar la siguiente tabla:

Longitud del pez en centímetros (cm)	Cantidad de peces
$0 \leq 5$ cm	
$5 \leq 10$ cm	
$10 \leq 15$ cm	
$15 \leq 20$ cm	
$20 \leq 25$ cm	
$25 \leq 30$ cm	

3. La longitud de cada pez se midió y se registró antes de devolverlo al lago. ¿Cuántas percas flavescens se midieron en esta muestra?
4. ¿Describirías la distribución de las longitudes de los peces de la muestra como una distribución de datos asimétrica o simétrica? Explica tu respuesta.

5. ¿Qué porcentaje de peces de la muestra medían menos de 10 centímetros de largo?

6. Si el pez más pequeño de esta muestra medía 2 centímetros de largo, ¿cuál es tu estimación de un intervalo de longitudes que contendría las longitudes del 25% más corto de los peces? Explica cómo obtuviste tu respuesta.

7. Si la longitud de la perca flavescens más grande era de 29 centímetros, ¿cuál es tu estimación de un intervalo de longitudes que contendría las longitudes del 25% más largo de los peces?

8. Estima la mediana de la longitud de las percas flavescens de la muestra. Explica cómo determinaste tu estimación.

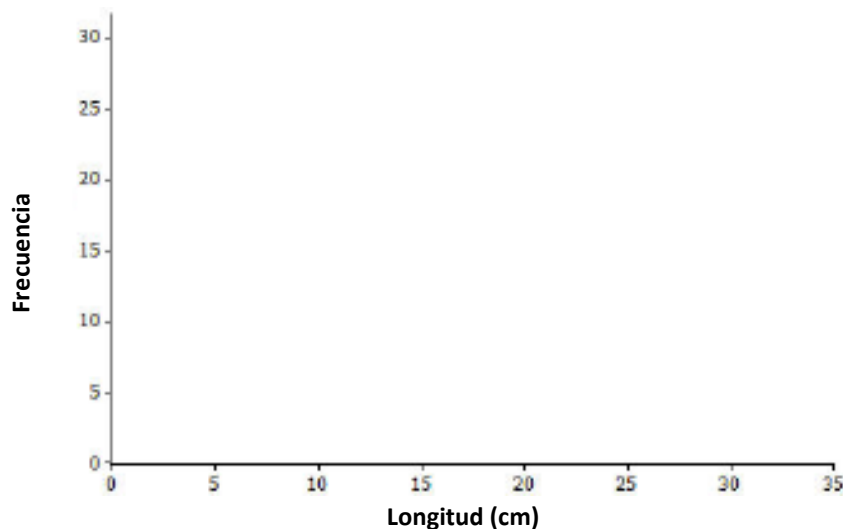
9. Según la forma de esta distribución de datos, ¿crees que la longitud media de una perca flavescens sería mayor, menor o igual que tu estimación de la mediana? Explica tu respuesta.

10. Recuerda que la longitud media es el punto de equilibrio de la distribución de longitudes. Estima la longitud media de esta muestra de percas flavescens.
11. La longitud de una perca flavescens se utiliza para estimar la edad de los peces. La perca flavescens suele crecer a lo largo de su vida. Las percas flavescens adultas tienen longitudes de entre 10 y 30 centímetros. ¿Cuántas percas flavescens de esta muestra se considerarían adultas? ¿Qué porcentaje de los peces de la muestra son adultos?

Ejemplo 2: ¿Cómo se vería una mejor distribución?

La perca flavescens es parte del alimento de peces más grandes y de otros tipos de animales salvajes de la región de los Grandes Lagos. ¿Por qué crees que los científicos se preocuparon cuando vieron el histograma anterior de longitudes de peces?

Haz un histograma que represente una muestra de 100 longitudes de percas flavescens que creas que indicaría que la perca no está en peligro de extinción.



Ejercicios 12 a 17: Estimación de la variabilidad de las longitudes de las percas flavescens

Estimaste la mediana de las longitudes de las percas flavescens de la primera muestra del ejercicio 8. También es útil describir la variabilidad de la longitud de la perca flavescens. ¿Por qué podría ser esto importante? Ten en cuenta las siguientes preguntas:

12. En varias lecciones anteriores, describiste una distribución de datos utilizando el resumen de cinco números. Utiliza el histograma y tus respuestas a las preguntas de ejercicios anteriores para proporcionar estimaciones de los valores del resumen de cinco números para esta muestra:

Valor mínimo (mín.) =

Valor del C1 =

Mediana =

Valor del C3 =

Valor máximo (máx.) =

13. Según el resumen de cinco números, ¿cuál es una estimación del valor del rango intercuartílico (R. I.C.) de esta distribución de datos?

14. Haz un diagrama de caja que represente las longitudes de las percas flavescens de esta muestra.



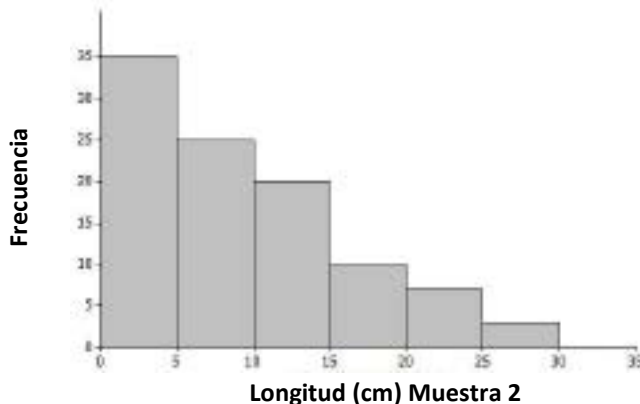
15. ¿Qué medida del centro (la mediana o la media) está más cerca del lugar donde tienden a agruparse las longitudes de las percas flavescens?
16. ¿Qué valor darías como longitud típica de las percas flavescens de esta muestra?
17. La desviación media absoluta (DMA) o el rango intercuartílico (R. I.C.) se utilizan para describir la variabilidad de una distribución de datos. ¿Qué medida de variabilidad utilizarías para esta muestra de percas? Explica tu respuesta.

Resumen de la lección

Por lo general, las distribuciones de datos se describen en términos de forma, centro y dispersión. Para evaluar la forma, se utilizan visualizaciones gráficas tales como histogramas, diagramas de puntos y diagramas de caja. Según la forma de una distribución de datos, se utilizan diferentes medidas del centro y de la variabilidad para describir la distribución. La mediana se utiliza para describir un valor típico en una distribución asimétrica, mientras que la media se utiliza en distribuciones que son aproximadamente simétricas. El R. I.C. se utiliza para describir la variabilidad de una distribución de datos asimétrica, mientras que la DMA se utiliza para describir la variabilidad de distribuciones que son aproximadamente simétricas.

Conjunto de problemas

Se recogió otra muestra de percas *flavescens* de los Grandes Lagos de un lago diferente. A continuación, se muestra un histograma de las longitudes de los peces de esta muestra.



1. Si la longitud de una perca *flavescens* es un indicador de su edad, ¿en qué se diferencia esta segunda muestra de la que investigaste en los ejercicios? Explica tu respuesta.
2. ¿Este histograma representa una distribución de datos asimétrica o casi simétrica?
3. ¿Qué medida del centro utilizarías para describir una longitud típica de una perca *flavescens* de esta segunda muestra? Explica tu respuesta.
4. Supón que la perca más pequeña que se atrapó medía 2 centímetros de largo y que la perca más grande que se atrapó medía 29 centímetros de largo. Estima los valores del resumen de cinco números de esta muestra:
 - Valor mínimo (mín.) =
 - Valor del C1 =
 - Mediana =
 - Valor del C3 =
 - Valor máximo (máx.) =

5. Según la forma de esta distribución de datos, ¿crees que la longitud media de una perca flavescens de esta segunda muestra sería mayor, menor o igual que tu estimación de la mediana? Explica tu respuesta.
6. Estima el valor medio de esta distribución de datos.
7. ¿Cuál es tu estimación de una longitud típica de una perca flavescens de esta muestra? ¿Utilizaste la longitud media del problema 5 para esta estimación? Explica por qué sí o por qué no.
8. ¿Utilizarías la DMA o el R. I.C. para describir la variabilidad de la longitud de la perca flavescens de los Grandes Lagos de esta muestra? Estima el valor de la medida de variabilidad que seleccionaste.

Lección 21: Resumen de una distribución de datos mediante la descripción de centro, variabilidad y forma

Trabajo en clase

Cada lección de este módulo trata acerca de datos. ¿Qué son los datos? ¿Qué preguntas se pueden responder con datos? ¿Cómo representas la distribución de datos para que puedas comprender y describir su forma? ¿Qué nos indica la forma sobre la manera de resumir los datos? ¿Qué es un valor típico del conjunto de datos? Estas y muchas otras preguntas fueron parte de tu trabajo en los ejercicios e investigaciones. Todavía hay mucho por aprender acerca de lo que nos indican los datos. Continuarás trabajando con estadísticas y probabilidades en 7.º y en 8.º grado, y a lo largo de la escuela secundaria. No obstante, ya has comenzado a aprender cómo descubrir las historias que hay detrás de los datos.

Cuando comenzaste este módulo, se presentaron los cuatro pasos que se utilizaron para llevar a cabo un estudio estadístico.

Paso 1: Formula una pregunta que se pueda responder con datos.

Paso 2: Recaba los datos adecuados.

Paso 3: Resume los datos con gráficos y con resúmenes numéricos.

Paso 4: Responde la pregunta formulada en el paso 1 con resúmenes numéricos y con gráficos.

En esta lección, llevarás a cabo estos pasos utilizando un conjunto de datos proporcionado.

Desafío de exploración: Precipitaciones anuales en el estado de Nueva York

El Centro Nacional de Datos Climáticos recaba datos en todo Estados Unidos que se pueden utilizar para resumir el clima de una región. Puedes obtener datos climáticos de un estado, de una ciudad, de un condado o de una región. Si estuvieras interesado en investigar el clima de tu área, ¿qué datos recabarías? Explica por qué crees que estos datos serían importantes como estudio estadístico del clima de tu área.

Para esta lección, utilizarás los datos de precipitaciones anuales del estado de Nueva York que compiló el Centro Nacional de Datos Climáticos. Los siguientes datos representan la cantidad de pulgadas de agua de lluvia (como promedio de diferentes lugares del estado) entre 1983 y 2012 (30 años).

45	42	39	44	39	35	42	49	37	42	41	42	37	50	39
41	38	46	34	44	48	50	47	49	44	49	43	44	54	40

Utiliza los cuatro pasos para llevar a cabo un estudio estadístico utilizando estos datos.

Paso 1: Formula una pregunta que se pueda responder con datos.

¿Qué pregunta estadística crees que se puede responder con estos datos? Escribe tu pregunta en la plantilla proporcionada para esta lección.

Paso 2: Recaba los datos adecuados.

Los datos ya se han recabado para esta lección. ¿Cómo crees que se recabaron estos datos? Recuerda que los datos representan la cantidad de pulgadas de agua de lluvia (como promedio de diferentes lugares del estado) entre 1983 y 2012 (30 años). En la plantilla de esta lección, escribe un resumen acerca de la manera en la que crees que se recabaron los datos.

Paso 3: Resume los datos con gráficos y con resúmenes numéricos.

Un buen primer paso podría ser resumir los datos con un diagrama de puntos. ¿Qué otro gráfico podrías hacer? Haz un diagrama de puntos u otro gráfico adecuado en la plantilla de esta lección. ¿Qué resúmenes numéricos calcularás? ¿Qué medida del centro utilizarás para describir un valor típico de estos datos? ¿Qué medida de la variabilidad calcularás y utilizarás para resumir la dispersión de datos? Calcula los resúmenes numéricos y escríbelos en la plantilla de esta lección.

Paso 4: Responde tu pregunta estadística utilizando los resúmenes numéricos y los gráficos.

En la plantilla de esta lección, escribe un resumen que responda la pregunta que formulaste.

Plantilla para la Lección 21

Paso 1: ¿Cuál es tu pregunta estadística?

Paso 2: ¿Cómo crees que se recabaron los datos?

Paso 3: Haz gráficos y calcula los resúmenes numéricos de los datos.
Haz, al menos, un gráfico de la distribución de datos. Calcula los resúmenes numéricos adecuados de los datos. Además, indica por qué seleccionaste estos resúmenes.

Paso 4: Responde tu pregunta estadística utilizando tus gráficos y resúmenes numéricos.

Resumen de la lección

La Estadística trata acerca del uso de datos para responder preguntas. Los cuatro pasos utilizados para llevar a cabo un estudio estadístico incluyen formular una pregunta que se pueda responder con datos, recabar los datos adecuados, resumir los datos con gráficos y resúmenes numéricos, y utilizar los datos, los gráficos y los resúmenes para responder la pregunta estadística.

Conjunto de problemas

En la Lección 17, formulaste una pregunta estadística y un plan para recabar datos para responder tu pregunta. También hiciste gráficos y calculaste resúmenes numéricos de tus datos. Revisa los datos recabados y tus resúmenes.

De acuerdo con las instrucciones de tu maestro, haz un cartel o un bosquejo para una presentación utilizando tus propios datos. En tu cartel, indica tu pregunta estadística. Además, indica un breve resumen acerca de la manera en la que recabaste tus datos de acuerdo con el plan que propusiste en la Lección 17. Incluye un gráfico que muestre la forma de tu distribución de datos, junto con las medidas de resumen del centro y de la variabilidad. Finalmente, responde tu pregunta estadística de acuerdo con los gráficos y los resúmenes numéricos.

Comparte el cartel que presentarás en la Lección 22 con tu maestro. Si se te indica preparar un bosquejo de la presentación, compártelo con tu maestro.

Lección 22: Presentación de un resumen de un proyecto estadístico

Trabajo en clase

Un estudio estadístico implica el siguiente proceso de investigación de cuatro pasos:

Paso 1: Formula una pregunta que se pueda responder con datos.

Paso 2: Recaba los datos adecuados.

Paso 3: Resume los datos con gráficos y con resúmenes numéricos.

Paso 4: Responde la pregunta formulada en el paso 1 con resúmenes numéricos y con gráficos.

Ahora te toca a ti ser un investigador y presentar tu propio estudio estadístico. En la Lección 17, formulaste una pregunta estadística, propusiste un plan para recabar datos para responder la pregunta y recabaste los datos. En la Lección 21, hiciste un cartel o un bosquejo de una presentación que incluía lo siguiente: la pregunta estadística, el plan que utilizaste para recabar los datos, gráficos y resúmenes numéricos de los datos, y una respuesta a la pregunta estadística basada en tus datos. Utiliza la siguiente tabla para organizar tu presentación.

Puntos para tener en cuenta:	Notas para incluir en tu presentación:
(1) Describir tu pregunta estadística.	
(2) Explicar a tu audiencia por qué estabas interesado en esta pregunta.	
(3) Explicar el plan que utilizaste para recabar los datos.	
(4) Explicar cómo organizaste los datos que recabaste.	

(5) Explicar los gráficos que preparaste para tu presentación y por qué hiciste estos gráficos.	
(6) Explicar qué medida del centro y qué medida de la variabilidad seleccionaste para resumir tu estudio. Explicar por qué seleccionaste estos valores.	
(7) Describir qué aprendiste de los datos. (Asegúrate de incluir una respuesta a la pregunta del paso (1) anterior).	

Resumen de la lección

La Estadística trata acerca del uso de datos para responder preguntas. Los cuatro pasos utilizados para llevar a cabo un estudio estadístico incluyen formular una pregunta que se pueda responder con datos, recabar los datos adecuados, resumir los datos con gráficos y resúmenes numéricos, y utilizar los datos, los gráficos y los resúmenes para responder la pregunta estadística.