

Unit 8 Caregiver Support

Unit Overview + Narrative Connections

In Unit 8, students will learn about the many ways to display and analyze data. They will use important information about birds and other creatures to learn how to create and interpret a dot plot and a histogram. After reading about the declining bee population, they will further their knowledge about measures of center. They will spend the end of the unit engaged with stories about endangered species as they learn about variability, the five-number summary, and mean absolute deviation.



Prior Learning	Current Learning	Future Learning
<ul style="list-style-type: none"> • Bar graphs • Line plots • Frequency comparisons 	<ul style="list-style-type: none"> • Analyze models of data and describe shape and spread • Determine measures of center • Measures of variability 	<ul style="list-style-type: none"> • Statistical sampling • Probability

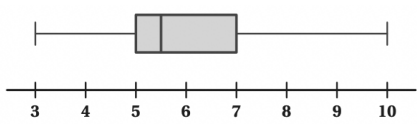
Key Ideas

- A question that can be answered by data that has *variability* is called a *statistical question*.
- Organizing data is helpful when trying to describe the data set or look for patterns.
- *Mean* and *median* are two important, but different, *measures of center* that can be used to draw conclusions about data.
- The variability of a data set can be viewed in multiple ways, for example, with numbers and plots.

Vocabulary

categorical data	Responses that are words or labels and can be sorted into categories	“Where do elephants live?” This IS a statistical question.
numerical data	Responses that are measurements or quantities that can be meaningfully compared	“What is your shoe size?” This is NOT a statistical question.

statistical question	A question that can be answered by collecting data that has variability	
distribution	The combination of all of the values in a data set and their frequencies	
dot plot	Representation of numerical data that uses a number line and stacked dots to represent the frequencies of each value in the data set	
frequency	A count of how often a value shows up in a data set	
mode	The “most common” or “most frequent” value in the data set	
center	A value in the middle of a data set that represents a typical value	
spread	A description of how alike or different the values in the data set are, often in relationship to the center	
histogram	A way to represent frequencies of data values that have been grouped into intervals, called <i>bins</i> , along a number line	
maximum	In a data set, the data value that is the greatest	
minimum	In a data set, the data value that is the least	
average	The sum of all the values in a set of numbers divided by how many numbers are in the set	
mean	A measure of center for a numerical data set that is equal to the average of all its values	
measure of center	Uses a single number to summarize a data set in terms of its typical values	
median	A measure of center for a distribution that represents the middle	

	value in a data set when the values are listed in order. Half of the values are less than or equal to the media, and half of the values are greater than or equal to the median.
variability	A description of how far away, or how spread out, data values generally are from the center
measure of variability	A way to summarize how the values in a data set vary with a single number
mean absolute deviation (MAD)	A measure of variability calculated by determining the average of the distances between each data value and the mean
five-number summary	Summarizes a distribution by five specific values: its minimum, first quartile, median, third quartile, and maximum
interquartile range (IQR)	A measure of variability that is calculated as the difference between the third quartile (Q3) and the first quartile (Q1)
quartile	One of three numbers (Q1, Q2, Q3) that divide a data set into 4 sections so that each contains the same number of data values
range	A measure of variability that is calculated as the difference between the maximum and minimum values in a data set
box plot	<p>A representation of the five-number summary for a numerical data set, positioned above a number line</p> 

Example Problems + Discussion Prompts

Sub-Unit 1

Problem	Sample Solution
<p style="text-align: center;">Lesson 2</p> <p>Which of the following is NOT a statistical question:</p> <ul style="list-style-type: none"> a) What is Mrs. Fleek’s favorite color? b) Which color - red, blue, yellow - is the most frequent color on national country flags? c) Which country has the most colors on its flag? 	<p>The answer is a: What is Mrs. Fleek’s favorite color?</p> <p>Answers b and c are statistical questions because they both produce variability in their answers.</p>

Discuss this question with your student:

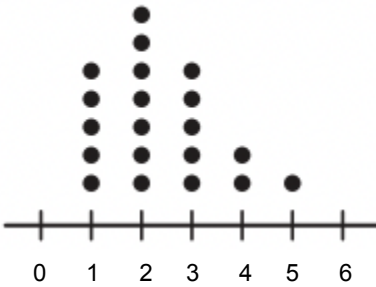
- Would your answer change if answer c was “Which country has the color red on its flag?”

<p style="text-align: center;">Lesson 6</p> <p>This is the data for the number of screen time hours of a group of 6th graders in one weekend day. Draw a histogram to display the data.</p> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tbody> <tr><td>0</td><td>2</td><td>2</td><td>3</td><td>3</td><td>3</td><td>3</td></tr> <tr><td>3</td><td>5</td><td>5</td><td>7</td><td>7</td><td>7</td><td>7</td></tr> <tr><td>7</td><td>7</td><td>8</td><td>8</td><td>9</td><td>9</td><td>9</td></tr> <tr><td>10</td><td>10</td><td>11</td><td>12</td><td>14</td><td>14</td><td>14</td></tr> </tbody> </table>	0	2	2	3	3	3	3	3	5	5	7	7	7	7	7	7	8	8	9	9	9	10	10	11	12	14	14	14	<p>To create the histogram we need to count how many numbers are in each <i>bin</i>. For this example we did <i>bins</i> of 0-2, 3-5, 6-8, 9-11, and 12-14. I counted the number of values in each bin to be:</p> <ul style="list-style-type: none"> 3 in 0-2 7 in 3-5 8 in 6-8 6 in 9-11 4 in 12-14
0	2	2	3	3	3	3																							
3	5	5	7	7	7	7																							
7	7	8	8	9	9	9																							
10	10	11	12	14	14	14																							

Discuss this question with your student:

- What is the difference between a histogram and a dot plot?

Sub-Unit 2

Problem	Sample Solution
<p style="text-align: center;">Lesson 9</p> <p>What is the mean value of the data shown on the dot plot?</p> 	<p>To find the mean we need to first calculate the sum of all the data values. To do this we will add the value for each dot.</p> $1+1+1+1+1+2+2+2+2+2+2+2+3+3+3+3+3+4+4+5=47$ <p>After we have the sum, we need to divide it by the number of values in the dot plot. To do this we will count the number of dots. Number of dots: 20</p> <p>To get the mean, we take the sum (47) and divide it by the number of values (20).</p> $\frac{47}{20} = 2.35$
<p>Discuss these questions with your student:</p> <ul style="list-style-type: none"> • How would the mean change if there was another dot above the 5? • How would the mean change if there was a dot above the zero? 	

Sub-Unit 3

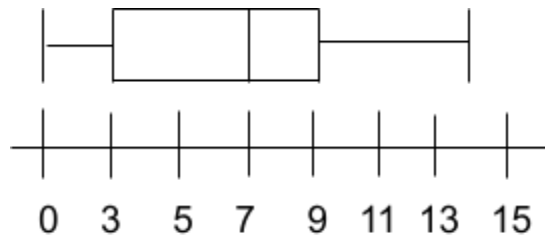
Problem	Sample Solution																												
<p style="text-align: center;">Lesson 14</p> <p>Find the five-number summary for the data below:</p> <table border="1" data-bbox="110 1360 792 1642"> <tbody> <tr> <td>0</td><td>2</td><td>2</td><td>3</td><td>3</td><td>3</td><td>3</td> </tr> <tr> <td>3</td><td>5</td><td>5</td><td>7</td><td>7</td><td>7</td><td>7</td> </tr> <tr> <td>7</td><td>7</td><td>8</td><td>8</td><td>9</td><td>9</td><td>9</td> </tr> <tr> <td>10</td><td>10</td><td>11</td><td>12</td><td>14</td><td>14</td><td>14</td> </tr> </tbody> </table>	0	2	2	3	3	3	3	3	5	5	7	7	7	7	7	7	8	8	9	9	9	10	10	11	12	14	14	14	<p>To find the five-number summary we need to find the:</p> <p>Minimum (smallest number): 0 Quartile 1 (median of the first half): 3 Median (middle of the data): 7 Quartile 3 median of the second half): 9.5* Maximum (largest number): 14</p> <p>*The third quartile for this data falls between values 9 and 10. So, we need to find the average of those two numbers to get our third quartile value.</p> $9 + 10 = 19, \frac{19}{2} = 9.5$
0	2	2	3	3	3	3																							
3	5	5	7	7	7	7																							
7	7	8	8	9	9	9																							
10	10	11	12	14	14	14																							

Discuss this question with your student:

- What is the interquartile range of the data?

Lesson 15

Using the five-number summary from the previous question, create a box plot.



To complete the box plot we create a box that uses the Quartile 1, Median and Quartile 3 values. The ends of the box plot are marked by the minimum and maximum values.

Discuss these questions with your student:

- Between what two numbers would you find the first 75% of the data?
- Between which two numbers would you find the greatest 50% of the data?

Sample Answers to Discussion Questions

Answers may vary.

- Would your answer change if answer c was: Which country has the color red on its flag?
 - *The answer would change to include that question because it is NOT statistical. There is only one single answer possible.*
- What is the difference between a histogram and a dot plot?
 - *A histogram uses bars to represent the number of data pieces in each bin. The data is grouped into the bins. A dot plot shows each value as a dot, giving all of the exact data pieces.*
- How would the mean change if there was another dot above the 5?
 - *If there was another dot above the 5, the sum of all values would be 52 and the total number of dots would be 21. This would make the mean: $\frac{52}{21} = 2.48$*

- How would the mean change if there was a dot above the zero?
 - *If there was a dot above the zero, the sum of all values would still be 47, but the total number of dots would be 21. This would make the mean: $\frac{47}{21} = 2.24$*

- What is the interquartile range of the data?
 - *The IQR is found by subtracting the Q1 from the Q3 value: $9.5 - 3 = 6.5$*

- Between what two numbers would you find the first 75% of the data?
 - *Between 0 (minimum) and 9.5 (Q3)*

- Between which two numbers would you find the greatest 25% of the data?
 - *Between 9.5 (Q3) and 14 (maximum)*

Apoyo para cuidadores/as, Unidad 8

Vista general de la unidad + Conexiones narrativas

En la Unidad 8, los/as estudiantes conocerán las múltiples formas de mostrar y analizar datos. Utilizarán información importante sobre las aves y otras criaturas para aprender a crear e interpretar un gráfico de puntos y un histograma. Después de leer sobre la disminución de la población de abejas, ampliarán sus conocimientos sobre las medidas de centro. Pasarán el final de la unidad ocupándose de historias sobre especies en peligro de extinción mientras aprenden sobre la variabilidad, el resumen de cinco números y la desviación media absoluta.



Aprendizaje previo	Aprendizaje actual	Aprendizaje futuro
<ul style="list-style-type: none">• Gráficas de barras• Diagrama de líneas• Comparaciones de frecuencia	<ul style="list-style-type: none">• Analizar modelos de datos y describir la forma y la propagación• Determinar las medidas del centro• Medidas de variabilidad	<ul style="list-style-type: none">• Muestreo estadístico• Probabilidad

Ideas clave

- Una pregunta que puede responderse con datos que tienen *variabilidad* se llama *pregunta estadística*.
- La organización de los datos es útil cuando se intenta describir el conjunto de datos o buscar patrones.
- *La media y la mediana* son dos importantes, aunque diferentes, *medidas de centro* que pueden utilizarse para sacar conclusiones sobre los datos.
- La variabilidad de un conjunto de datos puede verse de múltiples maneras, por ejemplo, con números y gráficas.

Vocabulario

datos categóricos	Respuestas que son palabras o etiquetas y que se pueden clasificar en categorías	"¿Dónde viven los elefantes?" Esta ES una pregunta estadística.
datos numéricos	Respuestas que son medidas o cantidades que pueden ser comparadas de manera significativa	"¿Cuál es su talla de zapatos?" Esta NO es una pregunta estadística.
pregunta estadística	Una pregunta que puede responderse recogiendo datos que tengan variabilidad	
distribución	La combinación de todos los valores de un conjunto de datos y sus frecuencias	
diagrama de puntos	Representación de datos numéricos que utiliza una línea numérica y puntos apilados para representar las frecuencias de cada valor en el conjunto de datos	
frecuencia	Un recuento de la frecuencia con la que aparece un valor en un conjunto de datos	
modo	El valor "más común" o "más frecuente" del conjunto de datos	
centro	Un valor en el centro de un conjunto de datos que representa un valor típico	
extender	Una descripción de la similitud o diferencia de los valores del conjunto de datos, a menudo en relación con el centro	
histograma	Una forma de representar las frecuencias de los valores de los datos que se han agrupado en intervalos, llamados <i>intervalos de clase</i> , a lo largo de una línea numérica	
máximo	En un conjunto de datos, el valor de los datos que es el mayor	

mínimo	En un conjunto de datos, el valor de los datos que es el menor	
promedio	La suma de todos los valores de un conjunto de números dividida por cuántos números hay en el conjunto	
media	Una medida de centro para un conjunto de datos numéricos que es igual a la media de todos sus valores	
medida del centro	Utiliza un único número para resumir un conjunto de datos en términos de sus valores típicos	
mediana	Medida del centro de una distribución que representa el valor medio de un conjunto de datos cuando los valores están ordenados. La mitad de los valores son menores o iguales a la media, y la mitad de los valores son mayores o iguales a la mediana.	
variabilidad	Una descripción de la distancia, o la dispersión, de los valores de los datos en general con respecto al centro	
medida de variabilidad	Una forma de resumir cómo varían los valores de un conjunto de datos con un solo número	
desviación absoluta media (DAM)	Una medida de variabilidad calculada mediante la determinación de la media de las distancias entre cada valor de los datos y la media	
resumen de cinco números	Resume una distribución por cinco valores específicos: su mínimo, primer cuartil, mediana, tercer cuartil y máximo	
rango intercuartil (RIC)	Una medida de variabilidad que se calcula como la diferencia entre el tercer cuartil (C3) y el primer cuartil (C1)	
cuartil	Uno de los tres números (C1, C2, C3) que dividen un conjunto de datos en 4 secciones para que cada una contenga el mismo número de valores de datos	
rango	Una medida de variabilidad que se calcula como la diferencia entre los valores máximos y mínimos de un conjunto de datos	
diagrama de cajas	Una representación del resumen de cinco números para un conjunto de datos numéricos, colocado sobre una línea numérica	

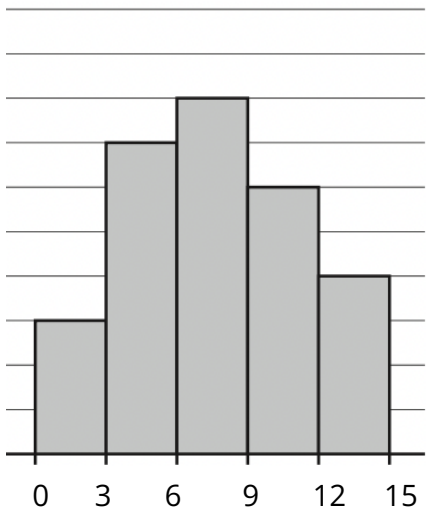
Problemas de ejemplo + Temas de discusión

Subunidad 1

Problema	Solución de ejemplo
<p style="text-align: center;">Lección 2</p> <p>¿Cuál de las siguientes opciones NO es una pregunta estadística?</p> <p>a) ¿Cuál es el color favorito de la Sra. Fleek?</p> <p>b) ¿Qué color - rojo, azul, amarillo - es el más frecuente en las banderas de los países?</p> <p>c) ¿Qué país tiene más colores en su bandera?</p>	<p>La respuesta es a: ¿Cuál es el color favorito de la Sra. Fleek?</p> <p>Las respuestas b y c son preguntas estadísticas porque ambas producen variabilidad en sus respuestas.</p>

Comente esta pregunta con su estudiante:

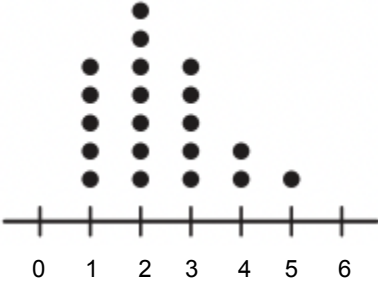
- ¿Cambiarías tu respuesta si la respuesta c fuera "¿Qué país tiene el color rojo en su bandera"?

<p style="text-align: center;">Lección 6</p> <p>Estos son los datos del número de horas de pantalla de un grupo de alumnos de 6° grado en un día de fin de semana. Dibuja un histograma para mostrar los datos.</p> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tbody> <tr> <td>0</td><td>2</td><td>2</td><td>3</td><td>3</td><td>3</td><td>3</td> </tr> <tr> <td>3</td><td>5</td><td>5</td><td>7</td><td>7</td><td>7</td><td>7</td> </tr> <tr> <td>7</td><td>7</td><td>8</td><td>8</td><td>9</td><td>9</td><td>9</td> </tr> <tr> <td>10</td><td>10</td><td>11</td><td>12</td><td>14</td><td>14</td><td>14</td> </tr> </tbody> </table>	0	2	2	3	3	3	3	3	5	5	7	7	7	7	7	7	8	8	9	9	9	10	10	11	12	14	14	14	<p>Para crear el histograma tenemos que contar cuántos números hay en cada <i>intervalo de clase</i>. En este ejemplo, hemos hecho <i>intervalos de clase</i> de 0-2, 3-5, 6-8, 9-11 y 12-14. Conté el número de valores en cada intervalo de clase para ser:</p> <p>3 in 0-2 7 in 3-5 8 in 6-8 6 in 9-11 4 in 12-14</p> 
0	2	2	3	3	3	3																							
3	5	5	7	7	7	7																							
7	7	8	8	9	9	9																							
10	10	11	12	14	14	14																							

Comente esta pregunta con su estudiante:

- ¿Cuál es la diferencia entre un histograma y un diagrama de puntos?

Subunidad 2

Problema	Solución de ejemplo
<p style="text-align: center;">Lección 9</p> <p>¿Cuál es el valor medio de los datos mostrados en el diagrama de puntos?</p> 	<p>Para hallar la media tenemos que calcular primero la suma de todos los valores de los datos. Para ello sumaremos el valor de cada punto.</p> $1+1+1+1+1+2+2+2+2+2+2+2+3+3+3+3+3+4+4+5=47$ <p>Una vez obtenida la suma, tenemos que dividirla por el número de valores en el diagrama de puntos. Para ello, contaremos el número de puntos.</p> <p>Número de puntos: 20</p> <p>Para obtener la media, tomamos la suma (47) y la dividimos por el número de valores (20).</p> $\frac{47}{20} = 2.35$

Comente estas preguntas con su estudiante:

- ¿Cómo cambiaría la media si hubiera otro punto por encima del 5?
- ¿Cómo cambiaría la media si hubiera un punto por encima del cero?

Subunidad 3

Problema	Solución de ejemplo																												
<p style="text-align: center;">Lección 14</p> <p>Encuentra el resumen de cinco números para los datos de abajo:</p> <table border="1" data-bbox="110 1524 792 1801"> <tbody> <tr> <td>0</td><td>2</td><td>2</td><td>3</td><td>3</td><td>3</td><td>3</td> </tr> <tr> <td>3</td><td>5</td><td>5</td><td>7</td><td>7</td><td>7</td><td>7</td> </tr> <tr> <td>7</td><td>7</td><td>8</td><td>8</td><td>9</td><td>9</td><td>9</td> </tr> <tr> <td>10</td><td>10</td><td>11</td><td>12</td><td>14</td><td>14</td><td>14</td> </tr> </tbody> </table>	0	2	2	3	3	3	3	3	5	5	7	7	7	7	7	7	8	8	9	9	9	10	10	11	12	14	14	14	<p>Para encontrar el resumen de cinco números necesitamos encontrar el:</p> <p>Mínimo (número más pequeño): 0 Cuartil 1 (mediana de la primera mitad): 3 Mediana (mitad de los datos): 7 Cuartil 3 (mediana de la segunda mitad): 9.5* Máximo (número más grande): 14</p> <p>*El tercer cuartil para estos datos cae entre los valores 9 y 10. Por lo tanto, tenemos que encontrar la media de esos dos números para obtener el valor del tercer cuartil.</p>
0	2	2	3	3	3	3																							
3	5	5	7	7	7	7																							
7	7	8	8	9	9	9																							
10	10	11	12	14	14	14																							

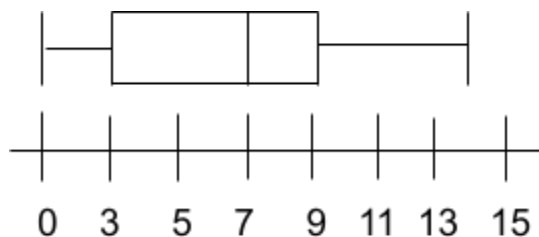
$$9 + 10 = 19, \frac{19}{2} = 9.5$$

Comente esta pregunta con su estudiante:

- ¿Cuál es el rango intercuartil de los datos?

Lección 15

Utilizando el resumen de cinco números de la pregunta anterior, crea un diagrama de caja.



Para completar el diagrama de caja, creamos una caja que utiliza los valores del cuartil 1, la mediana y el cuartil 3. Los extremos del diagrama de caja están marcados por los valores mínimo y máximo.

Comente estas preguntas con su estudiante:

- ¿Entre qué dos números se encuentra el primer 75% de los datos?
- ¿Entre qué dos números se encuentra el mayor 50% de los datos?

Respuestas de ejemplo a las preguntas de discusión

Puede haber varias respuestas.

- ¿Cambiarías tu respuesta si la respuesta c fuera "¿Qué país tiene el color rojo en su bandera"?
 - *La respuesta cambiaría para incluir esa pregunta porque NO es estadística. Sólo hay una única respuesta posible.*
- ¿Cuál es la diferencia entre un histograma y un diagrama de puntos?
 - *Un histograma utiliza barras para representar el número de piezas de datos en cada intervalo de clase. Los datos se agrupan en los intervalos de clase. Un diagrama de puntos muestra cada valor como un punto, dando todos los datos exactos.*
- ¿Cómo cambiaría la media si hubiera otro punto por encima del 5?
 - *Si hubiera otro punto por encima del 5, la suma de todos los valores sería 52 y el número total de puntos sería 21. Esto haría la media: $\frac{52}{21} = 2.48$*

- ¿Cómo cambiaría la media si hubiera un punto por encima del cero?
 - *Si hubiera un punto sobre el cero, la suma de todos los valores seguiría siendo 47, pero el número total de puntos sería 21. Esto haría la media: $\frac{47}{21} = 2.24$*

- ¿Cuál es el rango intercuartil de los datos?
 - *El RIC se obtiene restando el valor C1 del valor C3: $9.5 - 3 = 6.5$*

- ¿Entre qué dos números se encuentra el primer 75% de los datos?
 - *Entre 0 (mínimo) y 9.5 (C3)*

- ¿Entre qué dos números se encuentra el mayor 25% de los datos?
 - *Entre 9.5 (C3) y 14 (máximo)*