

Unit 8 Caregiver Support

Unit Overview + Narrative Connections

In this unit, students will graph and analyze bivariate data on scatter plots. They begin by identifying patterns in data, such as whether there is an association between a krill's eye distance and its height. The analysis of data is descriptive at first, focusing on associations, and becomes increasingly quantitative as students use linear models to make predictions. Finally, students explore categorical data using two-way tables and bar charts. Statistics can help us study and preserve the balance of our ecosystems!

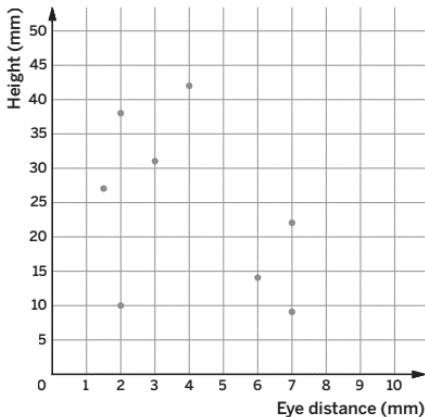
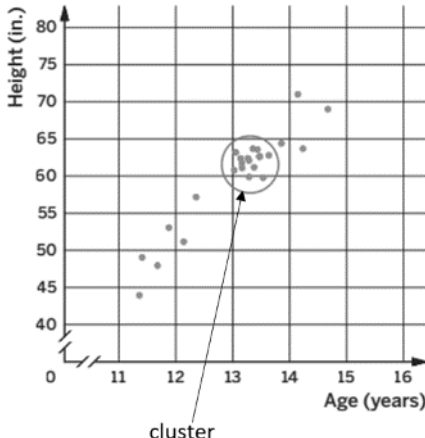


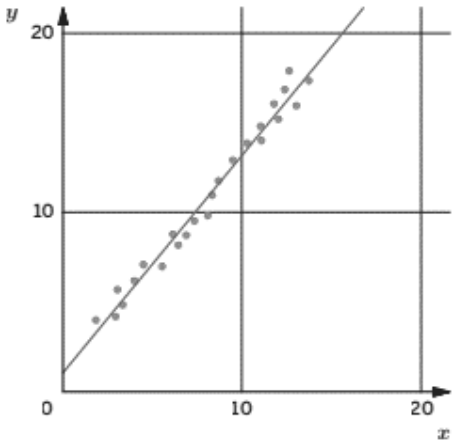
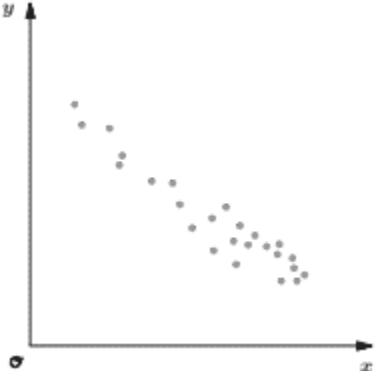
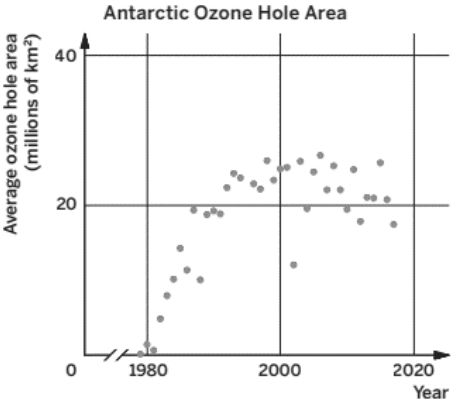
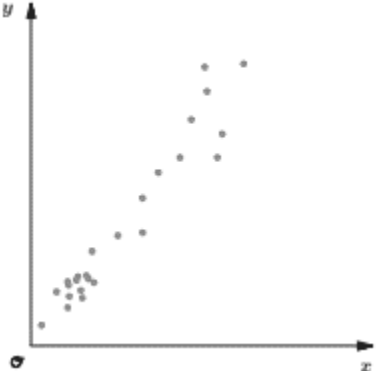
Prior Learning	Current Learning	Future Learning
<ul style="list-style-type: none"> Plotting points in the coordinate plane Graphing linear functions Ratios represented as percentages Looking at shape, center, spread, and outliers of data in one variable 	<ul style="list-style-type: none"> Describe associations between two variables Analyze numerical data in two variables using scatter plots and fitted lines Use two-way tables, bar graphs, and segmented bar graphs to analyze categorical data 	<ul style="list-style-type: none"> Model bivariate data with nonlinear functions Use correlation coefficients to describe associations Distinguish between correlation and causation

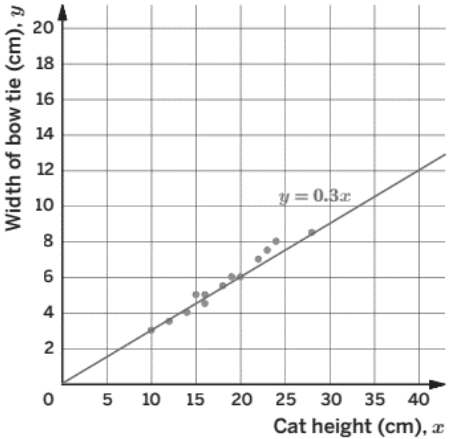
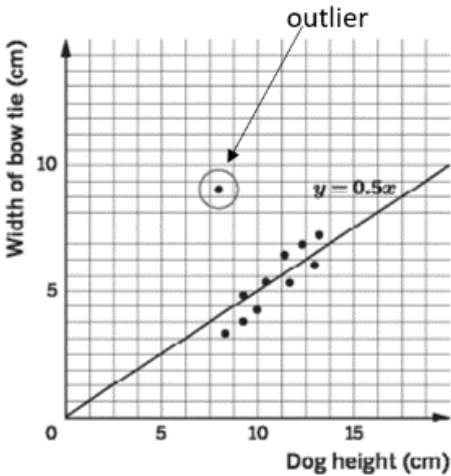
Key Ideas

- Scatter plots can show a *positive association*, a *negative association*, or no association, and they can also show a *linear association* or a *nonlinear association*.
- Data in a scatter plot that can be represented with a straight line shows a linear association.
- When appropriate, using a *linear model* to represent data in a scatter plot helps to interpret trends and make predictions.
- Different representations of *bivariate categorical data* — *two-way tables* and *segmented bar graphs* — can show *relative frequencies*, which can be used to determine associations in the data.

Vocabulary

<p>bivariate data</p>	<p>Data that has two variables</p>	<p>Example: The table shows the eye distances and heights (two variables) for five different krill.</p> <table border="1" data-bbox="1097 365 1471 688"> <thead> <tr> <th>Name</th> <th>Eye distance (mm)</th> <th>Height (mm)</th> </tr> </thead> <tbody> <tr> <td>Blue</td> <td>2</td> <td>30</td> </tr> <tr> <td>Orange</td> <td>4</td> <td>10</td> </tr> <tr> <td>Green</td> <td>8</td> <td>10</td> </tr> <tr> <td>Purple</td> <td>6</td> <td>20</td> </tr> <tr> <td>Red</td> <td>8</td> <td>20</td> </tr> </tbody> </table>	Name	Eye distance (mm)	Height (mm)	Blue	2	30	Orange	4	10	Green	8	10	Purple	6	20	Red	8	20
Name	Eye distance (mm)	Height (mm)																		
Blue	2	30																		
Orange	4	10																		
Green	8	10																		
Purple	6	20																		
Red	8	20																		
<p>scatter plot</p>	<p>A scatter plot is a graph that shows the values of two variables on a coordinate plane</p>																			
<p>cluster</p>	<p>A cluster represents data values that are grouped closely together</p>																			

<p>linear association</p>	<p>If a straight line can model the data, the data have a linear association</p>	 <p>A scatter plot on a coordinate plane with x and y axes ranging from 0 to 20. The data points show a clear upward trend, and a straight line of best fit is drawn through them, starting near the origin and extending towards the top right.</p>
<p>negative association</p>	<p>A negative association is a relationship between two quantities where one tends to decrease as the other increases</p>	 <p>A scatter plot on a coordinate plane with x and y axes. The data points show a downward trend, indicating that as the x-value increases, the y-value tends to decrease.</p>
<p>nonlinear Association</p>	<p>If a straight line cannot model the data, the data have a nonlinear association</p>	 <p>Antarctic Ozone Hole Area Average ozone hole area (millions of km²) Year</p> <p>A scatter plot showing the average ozone hole area in millions of km² from 1980 to 2020. The y-axis ranges from 0 to 40, and the x-axis ranges from 0 to 2020 with a break between 0 and 1980. The data points show a sharp increase from 1980 to around 2000, followed by a period of fluctuation between 20 and 30 million km².</p>
<p>positive association</p>	<p>A positive association is a relationship between two quantities where one tends to increase as the other increases</p>	 <p>A scatter plot on a coordinate plane with x and y axes. The data points show an upward trend, indicating that as the x-value increases, the y-value also tends to increase.</p>

<p>linear model</p>	<p>A linear equation that models a relationship between two quantities</p>													
<p>outlier</p>	<p>Outliers are points that are far away from their predicted values</p>													
<p>two-way table</p>	<p>A two-way shows two categorical variables. It shows one of the variables across the top and the other down one side. Each entry in the table is the frequency or relative frequency of the category shown by the column and row headings</p>	<table border="1" data-bbox="1062 1108 1510 1241"> <thead> <tr> <th></th> <th>Rides a bicycle</th> <th>Does not ride a bicycle</th> <th>Total</th> </tr> </thead> <tbody> <tr> <th>Uses a reusable water bottle</th> <td>12</td> <td>3</td> <td>15</td> </tr> <tr> <th>Does not use a reusable water bottle</th> <td>4</td> <td>6</td> <td>10</td> </tr> </tbody> </table>		Rides a bicycle	Does not ride a bicycle	Total	Uses a reusable water bottle	12	3	15	Does not use a reusable water bottle	4	6	10
	Rides a bicycle	Does not ride a bicycle	Total											
Uses a reusable water bottle	12	3	15											
Does not use a reusable water bottle	4	6	10											
<p>relative frequency</p>	<p>The relative frequency is the ratio of the number of times an outcome occurs in a set of data</p>	<table border="1" data-bbox="1062 1455 1510 1587"> <thead> <tr> <th></th> <th>Rides a bicycle</th> <th>Does not ride a bicycle</th> <th>Total</th> </tr> </thead> <tbody> <tr> <th>Uses a reusable water bottle</th> <td>80%</td> <td>20%</td> <td>100%</td> </tr> <tr> <th>Does not use a reusable water bottle</th> <td>40%</td> <td>60%</td> <td>100%</td> </tr> </tbody> </table>		Rides a bicycle	Does not ride a bicycle	Total	Uses a reusable water bottle	80%	20%	100%	Does not use a reusable water bottle	40%	60%	100%
	Rides a bicycle	Does not ride a bicycle	Total											
Uses a reusable water bottle	80%	20%	100%											
Does not use a reusable water bottle	40%	60%	100%											

<p>segmented bar graph</p>	<p>A segmented bar graph shows two categories of data. The whole bar represents all the data within one category. Then, each bar is separated into parts (segments) that show the percentage of each part in the second category</p>	
-----------------------------------	--	--

Example Problems + Discussion Prompts

Sub-Unit 1

Problem	Sample Solution
<p>Lesson 3</p> <p>Select all the following terms that describe the association in the scatter plot.</p> <p>A. linear association B. nonlinear association C. positive association D. negative association E. no association</p>	<p>The graph shows that as the age increases, the weight tends to increase. A positive association is shown between two quantities where one tends to increase as the other increases. A straight line can be used to model the data, so the data has a linear association.</p> <p>A. linear association C. positive association</p>

Discuss this question with your student:

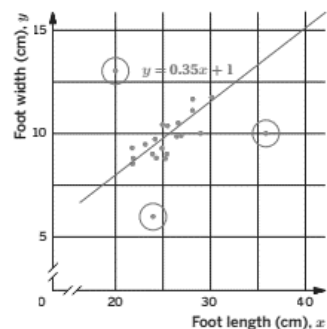
- How do you know if the quantities on a scatter plot show no association?

<p>Lesson 5</p> <p>The scatter plot shows several foot lengths and widths, along with a linear model represented by the equation $y = 0.35x + 1$.</p> <p>a. Use the linear model's equation to predict the width of a foot that is 50 cm long.</p>	<p>a. The equation $y = 0.35x + 1$ represents a linear model for the scatterplot, where x is the foot length in centimeters and y is the foot width in centimeters. Substitute the value for the foot length, $x = 50$, into the equation and solve for y, the foot width.</p> $y = 0.35(50) + 1$ $y = 17.5 + 1$ $y = 18.5$ <p>The predicted foot width is 18.5 cm.</p> <p>b. Outliers are points that are far away from their predicted values, so they would be far from the</p>
--	--

b. Does the scatter plot appear to have any outliers? If so, circle them and describe what they represent about foot length and foot width.

line that is drawn through the data.

The points that are outliers are located at about (20, 13), (24, 6), and (36, 10). The point (20, 13) shows a foot with a width much greater than predicted. The points (24, 6) and (36, 10) show feet with widths much smaller than predicted.



Discuss these questions with your student:

- Suppose you had a foot width of 12 cm. How would you find the predicted foot length?
- How do the data represented by the outliers compare to the data represented by the other points?

Lesson 8

A farmer brings his produce to the farmer's market and records whether people buy lettuce, apples, both, or something else. Complete the two-way table with the missing values. Then create a double bar graph to represent the data.

	Bought apples	Did not buy apples	Total
Bought lettuce	14		72
Did not buy lettuce		29	
Total	22		109



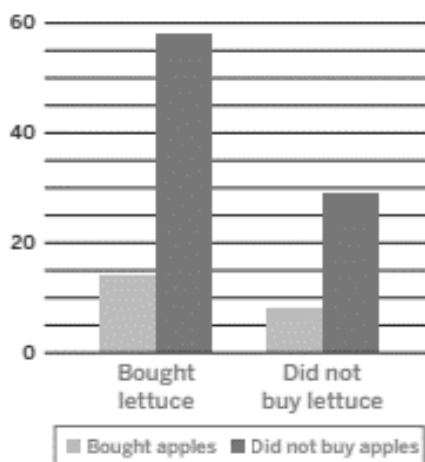
The total of the values and one other value in the *Bought lettuce* row are given. These values can be used to find the second value in that row: $72 - 14 = 58$. That means 58 is the first value in the *Did not buy apples* column. So the total in the *Did not buy apples* column is $58 + 29 = 87$.

The total and one value in the *Bought apples* column are given. So the value in the second row of that column is: $22 - 14 = 8$. That means 8 is the first value in the *Did not buy lettuce* row. So the total in the *Did not buy lettuce* row is $8 + 29 = 37$.

	Bought apples	Did not buy apples	Total
Bought lettuce	14	58	72
Did not buy lettuce	8	29	37
Total	22	87	109

A double bar graph can be created showing those who bought lettuce and apples with a height of 14, those who bought lettuce and did not buy apples with a height of 58, those who did not buy lettuce

and bought apples with a height of 8, and those who did not buy lettuce and did not buy apples with a height of 29.



Discuss this question with your student:

- How can you check the answers in the two-way table?

Sample Answers to Discussion Questions

Answers may vary.

- How do you know if the quantities on a scatter plot show no association?
 - *If the points on a scatter plot do not form any pattern; that is, the points do not show any increasing or decreasing trend, there is no association between the quantities.*
- Suppose you had a foot width of 11 cm. How would you find the predicted foot length?
 - *Since the y-value in the equation $y = 0.35x + 1$ represents the foot width and x represents the foot length, I would substitute the known foot width for y and solve for x .*
- How do the data represented by the outliers compare to the data represented by the other points?
 - *The data for the outliers do not follow the same trends as the other data points. The outliers represent a combination of foot length and width that are significantly different from the others.*
- How can you check the answers in the two-way table?
 - *I can check the answers in the two-way table by adding the data in each of the two rows to make sure I get the same totals listed in each row. I can add the data in each of the three columns to make sure I get the same totals listed in each column.*

Apoyo para cuidadores/as, Unidad 8

Vista general de la unidad + Conexiones narrativas

En esta unidad, los estudiantes van a graficar y analizar datos bivariados en diagramas de dispersión. Comienzan identificando patrones en los datos, como si existe una asociación entre la distancia del ojo de un krill y su altura. El análisis de los datos es descriptivo al principio, se centra en las asociaciones y se vuelve cada vez más cuantitativo a medida que los estudiantes usan modelos lineales para hacer predicciones. Finalmente, los estudiantes exploran datos categóricos utilizando tablas de doble entrada y gráficas de barras. ¡Las estadísticas pueden ayudarnos a estudiar y mantener el equilibrio de nuestros ecosistemas!



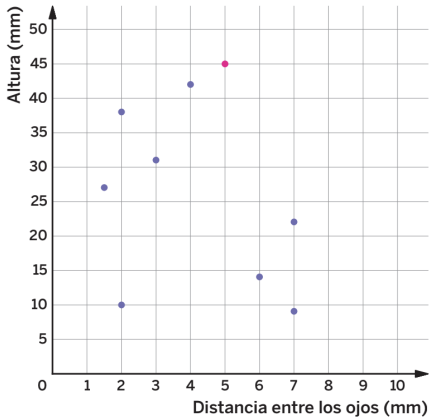
Aprendizaje previo	Aprendizaje actual	Aprendizaje futuro
<ul style="list-style-type: none"> Trazar puntos en el plano de coordenadas Graficar funciones lineales Proporciones representadas como porcentajes Mirar la forma, el centro, la dispersión y los valores atípicos de los datos en una variable 	<ul style="list-style-type: none"> Describir asociaciones entre dos variables Analizar datos numéricos en dos variables usando diagramas de dispersión y líneas ajustadas Usar tablas de doble entrada, gráficas de barras y gráficas de barras segmentadas para analizar datos categóricos 	<ul style="list-style-type: none"> Modelar datos bivariados con funciones no lineales Usar coeficientes de correlación para describir asociaciones Distinguir entre correlación y causalidad.

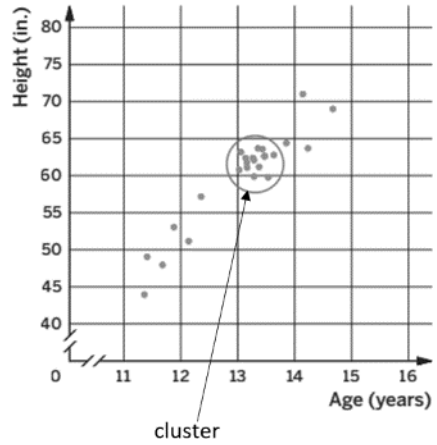
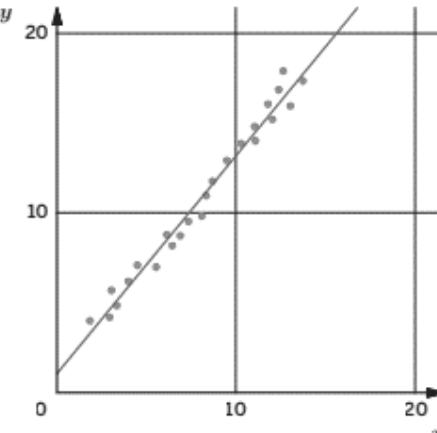
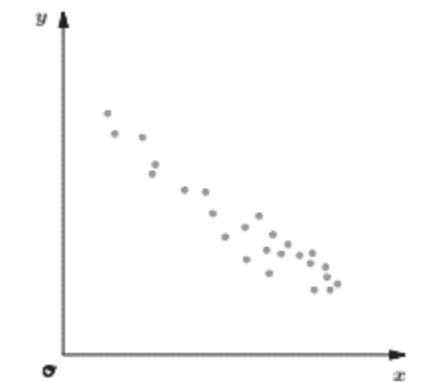
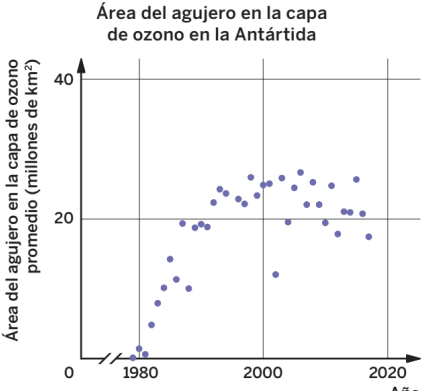
Ideas clave

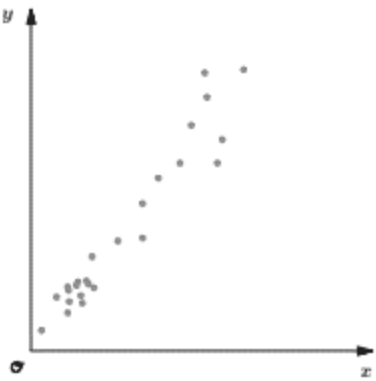
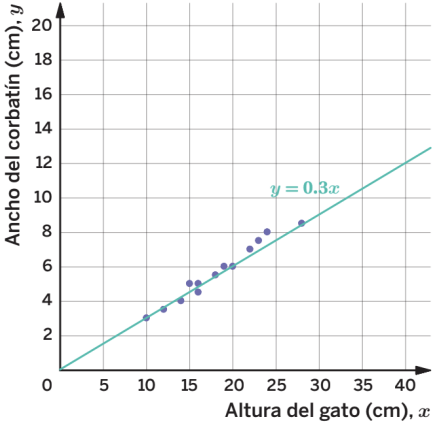
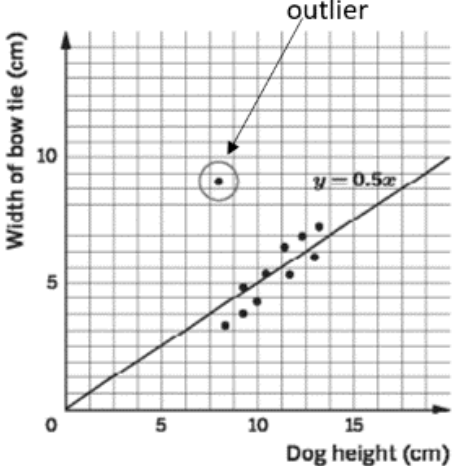
- Los *diagramas de dispersión* pueden mostrar una *asociación positiva*, una *asociación negativa* o *asociación ni positiva ni negativa*, y también pueden mostrar una *asociación lineal* o *no lineal*.
- Los datos en un gráfico de dispersión que se pueden representar con una línea recta muestran una asociación lineal.

- Cuando corresponda, el uso de un *modelo lineal* para representar datos en un diagrama de dispersión ayuda a interpretar tendencias y hacer predicciones.
- Diferentes representaciones de *datos categóricos bivariados* (*tablas de doble entrada* y *gráficas de barras segmentadas*) pueden mostrar *frecuencias relativas*, que se pueden usar para determinar asociaciones en los datos.

Vocabulario

<p>datos bivariados</p>	<p>Datos que tienen dos variables</p>	<p>Ejemplo: La tabla muestra las distancias y alturas de los ojos (dos variables) para cinco krill diferentes.</p> <table border="1" data-bbox="1094 636 1474 963"> <thead> <tr> <th>Nombre</th> <th>Distancia entre los ojos (mm)</th> <th>Altura (mm)</th> </tr> </thead> <tbody> <tr> <td>Azul</td> <td>2</td> <td>30</td> </tr> <tr> <td>Naranja</td> <td>4</td> <td>10</td> </tr> <tr> <td>Verde</td> <td>8</td> <td>10</td> </tr> <tr> <td>Morado</td> <td>6</td> <td>20</td> </tr> <tr> <td>Roja</td> <td>8</td> <td>20</td> </tr> </tbody> </table>	Nombre	Distancia entre los ojos (mm)	Altura (mm)	Azul	2	30	Naranja	4	10	Verde	8	10	Morado	6	20	Roja	8	20
Nombre	Distancia entre los ojos (mm)	Altura (mm)																		
Azul	2	30																		
Naranja	4	10																		
Verde	8	10																		
Morado	6	20																		
Roja	8	20																		
<p>diagrama de dispersión</p>	<p>Un diagrama de dispersión es una gráfica que muestra los valores de dos variables en un plano de coordenadas</p>																			

<p>agrupación</p>	<p>Una agrupación representa valores de datos que se agrupan muy de cerca</p>	 <p>A scatter plot with 'Age (years)' on the x-axis (0 to 16) and 'Height (in.)' on the y-axis (40 to 80). A cluster of points is circled and labeled 'cluster' with an arrow pointing to it. The cluster is centered around age 13 and height 60-65.</p>
<p>asociación lineal</p>	<p>Si una línea recta puede modelar los datos, los datos tienen una asociación lineal</p>	 <p>A scatter plot with x and y axes ranging from 0 to 20. The data points show a clear positive linear trend, and a straight line of best fit is drawn through them.</p>
<p>asociación negativa</p>	<p>Una asociación negativa es una relación entre dos cantidades donde una tiende a disminuir a medida que la otra aumenta.</p>	 <p>A scatter plot with x and y axes. The data points show a clear negative linear trend, sloping downwards from left to right.</p>
<p>asociación no lineal</p>	<p>Si una línea recta no puede modelar los datos, los datos tienen una asociación no lineal</p>	<p>Área del agujero en la capa de ozono promedio (millones de km²)</p>  <p>A scatter plot with 'Año' on the x-axis (1980 to 2020) and 'Área del agujero en la capa de ozono promedio (millones de km²)' on the y-axis (0 to 40). The data points show a non-linear trend, peaking around 2000 and then slightly declining.</p>

<p>asociación positiva</p>	<p>Una asociación positiva es una relación entre dos cantidades donde una tiende a aumentar a medida que la otra aumenta</p>													
<p>modelo lineal</p>	<p>Una ecuación lineal que modela una relación entre dos cantidades</p>													
<p>valor atípico</p>	<p>Los valores atípicos son puntos que están alejados de sus valores predichos</p>													
<p>tabla de doble entrada</p>	<p>Una tabla de doble entrada muestra dos variables categóricas. Muestra una de las variables en la parte superior y la otra en un lado. Cada entrada en la tabla es la frecuencia o frecuencia relativa de la categoría mostrada por los encabezados de columna y fila</p>	<table border="1" data-bbox="1057 1528 1533 1665"> <thead> <tr> <th></th> <th>Anda en bicicleta</th> <th>No anda en bicicleta</th> <th>Total</th> </tr> </thead> <tbody> <tr> <th>Usa una botella de agua reutilizable</th> <td>12</td> <td>3</td> <td>15</td> </tr> <tr> <th>No usa una botella de agua reutilizable</th> <td>4</td> <td>6</td> <td>10</td> </tr> </tbody> </table>		Anda en bicicleta	No anda en bicicleta	Total	Usa una botella de agua reutilizable	12	3	15	No usa una botella de agua reutilizable	4	6	10
	Anda en bicicleta	No anda en bicicleta	Total											
Usa una botella de agua reutilizable	12	3	15											
No usa una botella de agua reutilizable	4	6	10											

<p>frecuencia relativa</p>	<p>La frecuencia relativa es la razón del número de veces que ocurre un resultado en un conjunto de datos</p>	<table border="1"> <thead> <tr> <th></th> <th>Anda en bicicleta</th> <th>No anda en bicicleta</th> <th>Total</th> </tr> </thead> <tbody> <tr> <th>Usa una botella de agua reutilizable</th> <td>80%</td> <td>20%</td> <td>100%</td> </tr> <tr> <th>No usa una botella de agua reutilizable</th> <td>40%</td> <td>60%</td> <td>100%</td> </tr> </tbody> </table>		Anda en bicicleta	No anda en bicicleta	Total	Usa una botella de agua reutilizable	80%	20%	100%	No usa una botella de agua reutilizable	40%	60%	100%
	Anda en bicicleta	No anda en bicicleta	Total											
Usa una botella de agua reutilizable	80%	20%	100%											
No usa una botella de agua reutilizable	40%	60%	100%											
<p>gráfica de barras segmentada</p>	<p>Una gráfica de barras segmentada muestra dos categorías de datos. La barra completa representa todos los datos dentro de una categoría. Luego, cada barra se separa en partes (segmentos) que muestran el porcentaje de cada parte en la segunda categoría</p>													

Problemas de ejemplo + Temas de discusión

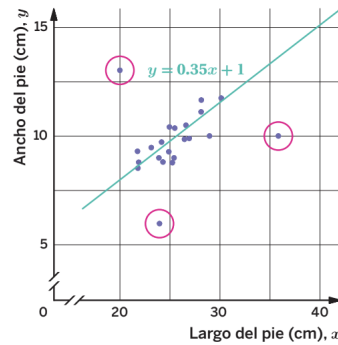
Subunidad 1

<p>Problema</p>	<p>Solución de ejemplo</p>
<p>Lección 3</p> <p>Selecciona todos los siguientes términos que describan la asociación en el diagrama de dispersión.</p> <p>A. asociación lineal B. asociación no lineal C. asociación positiva D. asociación negativa E. asociación ni positiva ni negativa</p>	<p>La gráfica muestra que a medida que aumenta la edad, el peso tiende a aumentar. Se muestra una asociación positiva entre dos cantidades donde una tiende a aumentar a medida que la otra aumenta. Se puede usar una línea recta para modelar los datos, por lo que los datos tienen una asociación lineal.</p> <p>A. Asociación lineal B. Asociación positiva</p>
<p>Comente esta pregunta con su estudiante:</p>	

- ¿Cómo sabes si las cantidades en un diagrama de dispersión muestran asociación ni positiva ni negativa?

Lección 5

El diagrama de dispersión muestra varias longitudes y anchos de pie, junto con un modelo lineal representado por la ecuación $y = 0.35x + 1$.



a. Usa la ecuación del modelo lineal para predecir el ancho de un pie que mide 50 cm de largo.

b. ¿Pareciera que el diagrama de dispersión tiene valores atípicos? Si es así, encierra en un círculo y describe lo que representan sobre la longitud y el ancho del pie.

a. La ecuación $y = 0.35x + 1$ representa un modelo lineal para el diagrama de dispersión, donde x es la longitud del pie en centímetros, e y es el ancho del pie en centímetros. Sustituye el valor de la longitud del pie, $x = 50$, en la ecuación y resuelve y , el ancho del pie.

$$y = 0.35(50) + 1$$

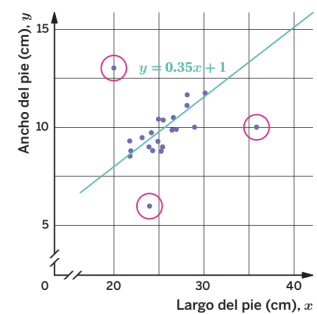
$$y = 17.5 + 1$$

$$y = 18.5$$

El ancho del pie predicho es **18.5 cm**.

b. Los valores atípicos son puntos que están alejados de sus valores predichos, por lo que estarían alejados de la línea que se dibuja a través de los datos.

Los puntos que son valores atípicos están ubicados aproximadamente en (20, 13), (24, 6) y (36, 10). El punto (20, 13) muestra un pie con un ancho mucho mayor al predicho. Los puntos (24, 6) y (36, 10) muestran pies con anchos mucho más pequeños que lo predicho.



Comente estas preguntas con su estudiante:

- Supongamos que tienes un ancho de pie de 11 cm. ¿Cómo encontrarías la longitud del pie predicha?
- ¿Cómo se comparan los datos representados por los valores atípicos con los datos representados por los otros puntos?

Lección 8

Un/a agricultor/a lleva su cosecha a un mercado de productores y registra si las personas compran lechuga, manzanas, ambas u otra cosa. Completa la tabla de doble entrada con los valores que faltan. Luego crea una gráfica de barras dobles para representar los datos.

	Bought apples	Did not buy apples	Total
Bought lettuce	14		72
Did not buy lettuce		29	
Total	22		109

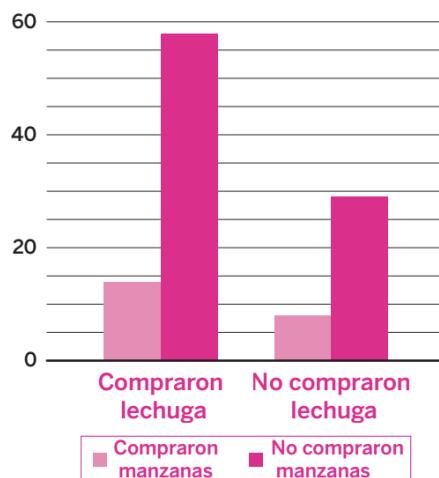


El total de los valores y otro valor más en la fila *Compraron lechuga* se dan. Estos valores pueden ser usados para encontrar el segundo valor de la fila: $72 - 14 = 58$. Eso significa que 58 es el primer valor en la columna *No compraron manzanas*. Por lo tanto, el total en la columna *No compraron manzanas* es $58 + 29 = 87$.

Se dan el valor total y un valor en la columna *Compraron manzanas*. Entonces, el valor en la segunda fila de esa columna es: $22 - 14 = 8$. Eso significa que 8 es el primer valor en la fila *No compraron lechuga*. Entonces, el total en la fila *No compraron lechuga* es $8 + 29 = 37$.

	Compraron manzanas	No compraron manzanas	Total
Compraron lechuga	14	58	72
No compraron lechuga	8	29	37
Total	22	87	109

Se puede crear una gráfica de barras dobles que muestre a quienes compraron lechuga y a quienes no compraron manzanas con una altura de 58, quienes no compraron lechuga y quienes no compraron manzanas con una altura de 8, y quienes no compraron lechuga y no compraron manzanas con una altura de 29.



Comente esta pregunta con su estudiante:

- ¿Cómo puedes verificar las respuestas en la tabla de doble entrada?

Respuestas de ejemplo a las preguntas de discusión

Puede haber varias respuestas.

- ¿Cómo sabes si las cantidades en un diagrama de dispersión muestran asociación ni positiva ni negativa?
 - *Si los puntos de un diagrama de dispersión no forman ningún patrón; es decir, los puntos no muestran ninguna tendencia creciente o decreciente, no hay asociación entre las cantidades.*
- Supongamos que tienes un ancho de pie de 11 cm. ¿Cómo encontrarías la longitud del pie predicha?
 - *Dado que el valor y en la ecuación $y = 0.35x + 1$ representa el ancho del pie y x representa la longitud del pie, yo sustituiría el ancho del pie conocido por y y resolvería para x .*
- ¿Cómo se comparan los datos representados por los valores atípicos con los datos representados por los otros puntos?
 - *Los datos de los valores atípicos no siguen las mismas tendencias que los otros puntos de datos. Los valores atípicos representan una combinación de la longitud y el ancho del pie que son significativamente diferentes de los demás.*
- ¿Cómo puedes verificar las respuestas en la tabla de doble entrada?
 - *Puedo verificar las respuestas en la tabla de doble entrada sumando los datos en cada una de las dos filas para asegurarme de obtener los mismos totales enumerados en cada fila. Puedo agregar los datos en cada una de las tres columnas para asegurarme de obtener los mismos totales enumerados en cada columna.*