Lesson 2: Plotting data

Goals

- Create a representation of single-variable data using a box plot, histogram, or dot plot, and compare and contrast (orally) these representations with a scatter plot.
- Create a scatter plot from a table of data, and describe (orally and in writing) the trend of the data.
- Create a table of collected data, and explain (orally) how to organise the data.

Learning Targets

- I can draw a scatter plot to show data that has two paired variables.

Lesson Narrative

In earlier years, students represented the distribution of a single statistical variable using dot plots, histograms, and box plots. In this lesson, students review those ways of displaying data and compare them with representing the relationship between *two* variables in a scatter plot. They aggregate data about hand span, arm span, and height for all students in the class, and use the data to create plots of a single variable as well as a scatter plot. (The data will be used again in later lessons, so should be kept in a spreadsheet or a permanent visual display in the classroom.) They notice that scatter plots can convey information about the relationship between two variables that representations of each variable separately do not reveal. They understand that every point in a scatter plot represents two measures of a single individual in the population.

In this lesson, students choose an appropriate way to display a data set and see that the structure in a scatter plot can reveal information about a data set that is not visible in other representations.

Building On

- Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

Addressing

- Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect
- Compare and Connect
- Discussion Supports

Required Materials

Measuring tapes
Rulers marked with centimetres

Required Preparation

For the Gathering Data activity, each pair of students will need to measure some lengths in centimetres. Rulers, metre sticks, or tape measures would work for this purpose, so long as they are marked in centimetres.

At the end of the Gathering Data activity, students are instructed to add their data to a class table. It would be fine for students to record their individual data on the board for all to see. However, this data will be used again in a later lesson, so you need a way of preserving it for later. A useful mechanism might be to set up a shared spreadsheet that students access through a browser.

Student Learning Goals

Let's collect and display some data about the class.

2.1 Representing Data

Warm Up: 5 minutes

In earlier years, students studied different ways of representing sets of data. The purpose of this warm-up is for students to recall ways that two different sets of data can be represented. Data visualisation is very useful for understanding patterns that are not visible in other ways. Different representations can highlight different aspects of the data and lead the viewer to see different patterns.

Students are given two scenarios and asked for appropriate representations. One scenario contains quantitative data (distance run) while the other has qualitative data (favourite colour). Each scenario can be represented in more than one way. As students work, monitor for different students using a variety of display choices.

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect

Launch

Give students 2 minutes of quiet work time followed by a whole-class discussion.

Ask students to brainstorm different graphical representations of data they have used or seen in the past. Record and display answers for all to see.

Student Task Statement

Lin surveyed 30 students about the longest time they had ever run. Andre asked them about their favourite colour. How could Lin and Andre represent their data sets? Would they represent them in the same way? Why or why not?

Student Response

Answers vary. Sample response: For the longest run, Lin could use a dot plot, a histogram, or a box plot. For the colours, Andre can use a bar graph or a pie chart. They represent the data in different ways because Lin's data is quantitative and Andre's is qualitative.

Activity Synthesis

The purpose of the discussion is for students to think about different representations of data and to emphasise the difference in qualitative and quantitative data types.

Select previously identified students to share their responses and record for all to see. To highlight the differences between the representations, after each student shares ask:

- "What patterns would Lin be able to see if she represented her data in that way?" (A dot plot would show all of the data, so would be good for showing details. A histogram would give a good overall picture of the data, but would hide details. A box plot would also be useful for showing an overview of the data and would highlight the median, but would not show the individual data values.)
- "What patterns would Andre be able to see if he represented his data in that way?" (A bar graph would be useful for quickly seeing the relative order of colour preferences. A pie chart would be useful for quickly seeing how preferred each colour is when compared to the whole.)

After the responses have been shared, ask, "Why couldn't Lin and Andre use the same graph type to represent their data?" (Lin has quantitative data and Andre has qualitative data.)

2.2 Gathering Data

10 minutes

Data collection is an important part of any statistics unit. Here, students collect data about a partner using the appropriate tools and submit the data for future use.

In a lesson later in the unit, students analyse the class data they collect here using the methods they learn in the next few lessons. A copy of the class's data should be kept. It may be convenient to set up a shared spreadsheet if students have access to Internet-enabled devices.

Note: Some students may be self-conscious about their body measurements. On the other hand, it can be motivating for students to use their own measurements to analyse

statistically. Depending on your class, consider collecting the data and analysing a similar data set in a future lesson (measurements from the staff, a different class, or invented data similar to the data collected).

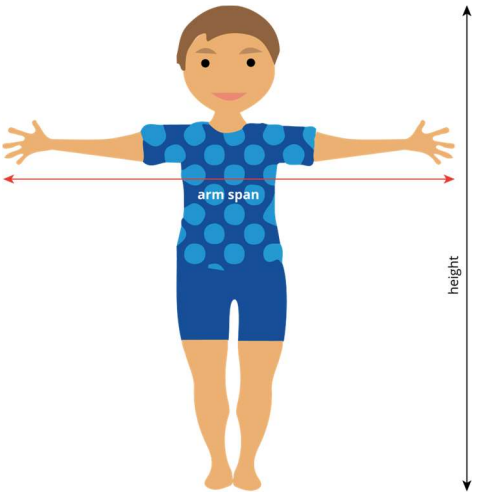
Instructional Routines

- Discussion Supports

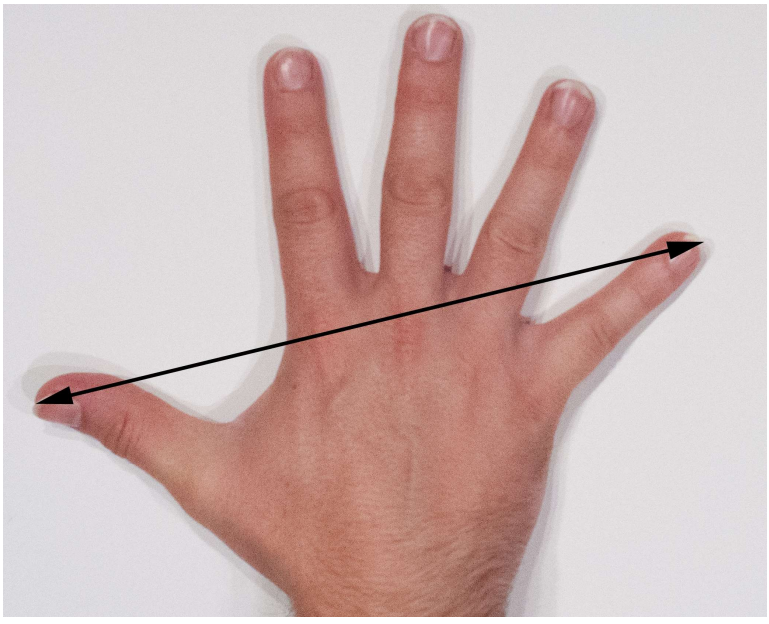
Launch

Arrange students in groups of 2. Provide access to rulers marked in centimetres.

It may be helpful to demonstrate how to measure arm span as defined in the task. An image is included here to aid understanding.



It may also be useful to demonstrate how to measure hand span as defined in the task. An image of this is also included.



Prepare a table with enough rows for the entire class to include their data for all to see. The data will be needed again in a future lesson, so it might be helpful to keep a copy digitally, on a large piece of chart paper, or something that could be brought out again later. It is important to collect this data so that the class will have enough data to look for patterns later, but the data does not need to be connected to individuals, so collecting names attached to the data is not important.

As groups complete their tables, select one student from each group to include their group's data in the table for the entire class. Each student should also have access to the class data for the next activity as well as for a future lesson.

Engagement: Develop Effort and Persistence. Provide prompts, reminders, guides, rubrics, or checklists that focus on increasing the length of on-task orientation in the face of distractions. For example, create a checklist for students to complete as they measure each person's height, arm span, and hand span and record the data in the table.
Supports accessibility for: Attention; Social-emotional skills

Student Task Statement

Are older students always taller? Do taller students tend to have bigger hands? To investigate these questions, the class will gather data.

- A person's *arm span* is the distance between the tips of their index fingers, when their arms are fully spread out.
- A person's *hand span* is the distance from the tip of their thumb to the tip of their little finger, when their fingers are fully spread out.

1. Each partner should:
   - Measure the other partner's height, arm span, and hand span for their right hand to the nearest centimetre.
   - Record the other partner's measurements and age (in months) in the table.

	height (cm)	arm span (cm)	hand span (cm)	age (months)
partner A				
partner B				

2. One partner records the data from your table in a table of data for the entire class.

Student Response

Answers vary. Sample response:

	height (cm)	arm span (cm)	hand span (cm)	age (months)
partner A	160	157	19	162, because 13 years is 156 months,

				plus 6 more months.
partner B	152	155	18	165

Activity Synthesis

The goal for this activity is for students to see the importance of careful data collection as well as an organisational strategy for keeping up with the data as it is collected.

Consider asking some of the following questions:

- "Do you notice any patterns in the data from the table?"
- "Is the class data organised in a way that is useful for noticing patterns?" (Not likely.)
- "How could we reorganise the table to make it more useful?" (Sort from shortest to tallest or youngest to oldest.)
- "What questions do you have about the data?" (Are some of the items related to one another? Do height and hand span have any relationship?)

Speaking: Discussion Supports. As students describe what they notice in the table, revoice student ideas to demonstrate mathematical language use. Press for details in students' explanations by requesting that students challenge an idea, elaborate on an idea, or give an example. This will help students to produce and make sense of the language needed to communicate their own ideas, and identify the importance of careful data collection.
Design Principle(s): Support sense-making; Optimise output (for explanation)

2.3 Scatter Plots

20 minutes (there is a digital version of this activity)

In this activity, students choose appropriate ways to display data using data that comes from the items students measured in the previous activity.

After drawing the distributions of the individual variables and then the scatter plot, students look for patterns in the data that might not have been visible in the tabular format.

Instructional Routines

- Compare and Connect

Launch

Keep students in groups of 2. For the sake of time, you may tell students that groups should discuss a plan for working on problems 1 and 2, then each student can create the display individually.

Brainstorm the ways that students can display the distribution of height data. Possibilities include dot plots, histograms, and box plots. Remind students to label axes and include units of measurement.

If using the digital materials: The digital tools allow students to input the data and quickly view the difference between many graphical representations. Have students discuss the benefits of different graphs for this specific data (height).

Action and Expression: Provide Access for Physical Action. Provide access to tools and assistive technologies such as a the digital materials or graphing software. Some students may benefit from a checklist or list of steps to be able to use the digital materials or software.
Supports accessibility for: Organisation; Conceptual processing; Attention

Anticipated Misconceptions

Some students may struggle to find a good way to number the axes so that the data is visible, but not misleading. As students have already seen in scatter plots for this unit, it is not essential to start from 0 on a scatter plot. To accurately show the data, ask students to find the minimum and maximum values in the data set and use those to help think about reasonable boundaries for the left, right, bottom, and top sides of the scatter plot. The increments should be chosen based on the minimum and maximum values of the boundaries.

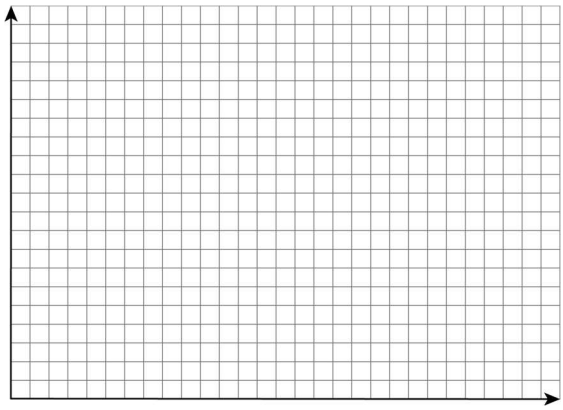
For example, if the minimum value is 8 532 and the maximum value is 13 456, then the left side of the graph might be 8 000 or 8 500 and the right side 13 500 or 14 000. The grid divisions should not represent steps of 1 (There would be thousands of grid lines!), but rather 300 or so. These values allow all of the data points to fit without them being too clustered together and indistinguishable.

In most cases, estimates of position can be used to plot the points in a scatter plot, but using technology makes it more precise.

Student Task Statement

1. What types of graphical representations could be used to show the class's height measurements? Make a graphical representation of the class's height measurements.

2. Choose a colour and use it to plot a point on the coordinate plane that represents your own height and hand span. Then, in the same colour, plot a second point that represents your partner's height and hand span.



3. In a different colour, plot the height and hand span of each student in your class, making a scatter plot of the heights and hand spans for the entire class.

4. Based on your scatter plot, answer these questions:

    a. Do taller students in your class tend to have bigger hands? Explain how you know.

    b. Is hand span a linear function of height? Explain how you know.

Student Response

1. Answers vary. Sample response. Histogram, a dot plot, or a box graph.

2. Graphs vary based on class data. The graph should be labelled with height, in cm, along one axis and hand spans, in cm, along the other axis. If the heights are along the horizontal axis, then each student should be represented by a single point with coordinates in the form of (height, hand span).

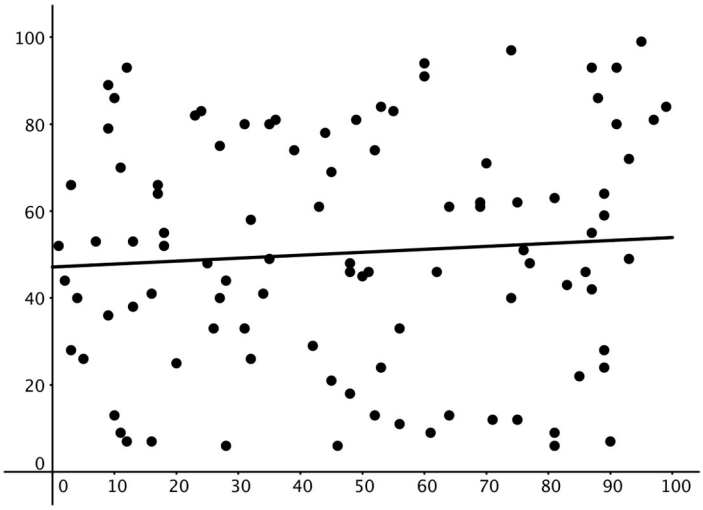
3. Graphs vary based on class data.

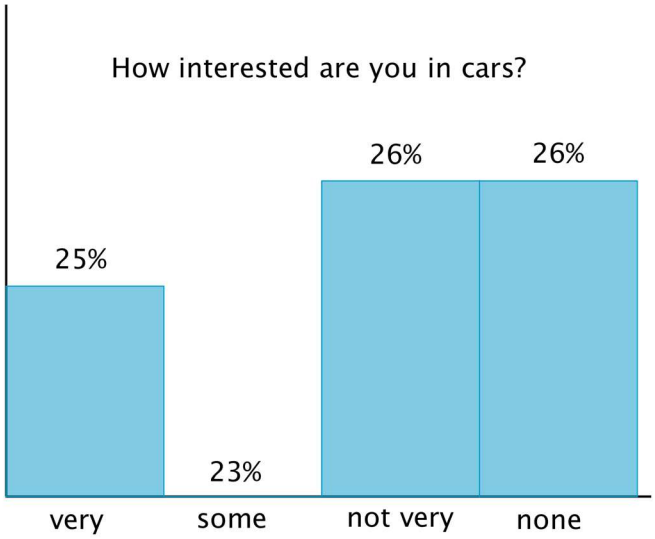
    a. Yes, taller students tend to have greater hand spans. On the graph, as the heights increase, the hand spans tend to increase as well.

    b. No, the hand span is not a linear function of height, because I cannot draw a straight line through all of the data points. Sometimes there are different hand span measurements for students with the same height.

Are You Ready for More?

Although the data may be accurate, displaying the data incorrectly can tell the wrong story. What is wrong with each of these graphic representations of the data?





Student Response

Answers vary. Sample response: In the scatter plot, the axes are not labelled, so it is not clear what the data represents. The line also doesn't seem to make sense with the rest of the points. In the bar graph, the heights of the bars do not match the values written above them. All four responses are pretty close to one another, but the bars make it look like there is a major difference.

Activity Synthesis

The goal of the discussion is for students to explore appropriate data representations including one of the main focuses of the unit: scatter plots.

Select students to share their graphs for height measurements and their scatter plots for heights and hand spans in the class. In particular, look for students who chose to represent the data in a different way.

For example, to graph height data, students could group different heights into bins to graph with a histogram or graph each student's height individually on a dot plot.

Additionally, for the scatter plot, height could be along the x-axis and hand span along the y-axis or this could be reversed. If no students have tried these methods, mention them as a possibility. Use *Compare and Connect* to ask what is similar and what is different about the graph when height and hand span are on different axes.

Select a single student's information from the data and ask students to identify where that person's data is represented in each of the displays.

Make sure that students understand the distinction between representations of the distribution of a *single* variable (height data alone) and this new representation that

contains information about two variables at once (height and hand span). Tell students that we will be using scatter plots to look at the relationship between different variables that cannot be seen as easily when we represent the distributions separately.

Speaking, Listening: Compare and Connect. As students share their visual displays, ask students to identify what is the same and what is different about each representation. Draw students' attention to the ways individual data points were represented in each of the displays. In this discussion, emphasise the mathematical language used to make sense of the different graphical representations to display the data. These exchanges strengthen students' mathematical language use and reasoning used to represent data.
Design Principle(s): Maximise meta-awareness

Lesson Synthesis

Compare and contrast the different kinds of representations of data that students worked with in this lesson. Some represent a single qualitative variable. Some a single quantitative variable. Scatter plots show the relationship between *two* variables. Note that we will be using scatter plots extensively in this unit to investigate the relationship between different variables in many different contexts.

To highlight the different representations students created and saw today, ask:

- "When do we use histograms or box plots to represent a data set? When do we use a scatter plot?" (Histograms or box plots are useful when displaying numerical data from a single variable. Scatter plots are more useful when looking for a relationship between two variables that produce numerical data.)
- "What does a point in a scatter plot represent?" (Two measures about a single person or object.)

2.4 Right Side Measurements

Cool Down: 5 minutes

Students create a scatter plot from data given in a given table. Then they circle the data point in the scatter plot that corresponds to a particular entry in the table, reinforcing the connection between these representations.

Student Task Statement

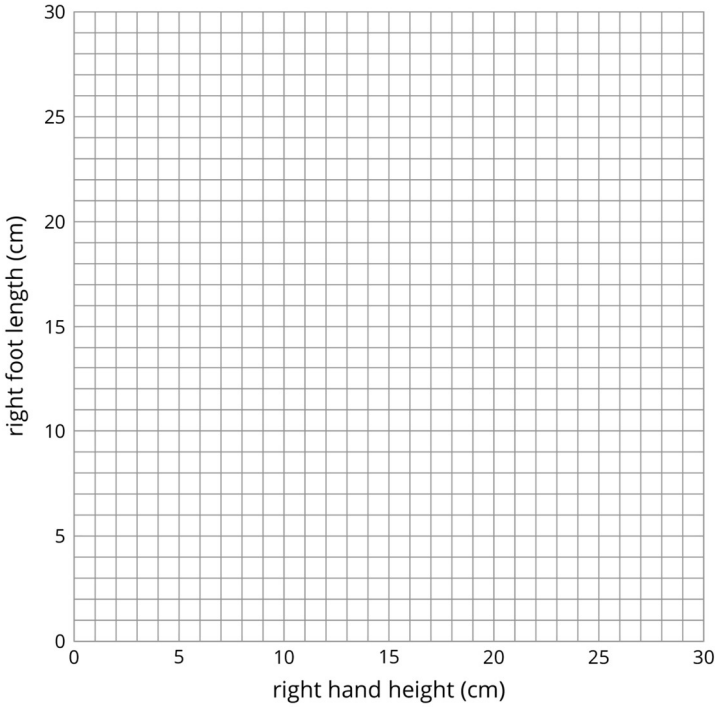
The table shows measurements of right hand length and right foot length for five people.

	right hand length (cm)	right foot length (cm)
person A	19	27
person B	21	30
person C	17	23
person D	18	24

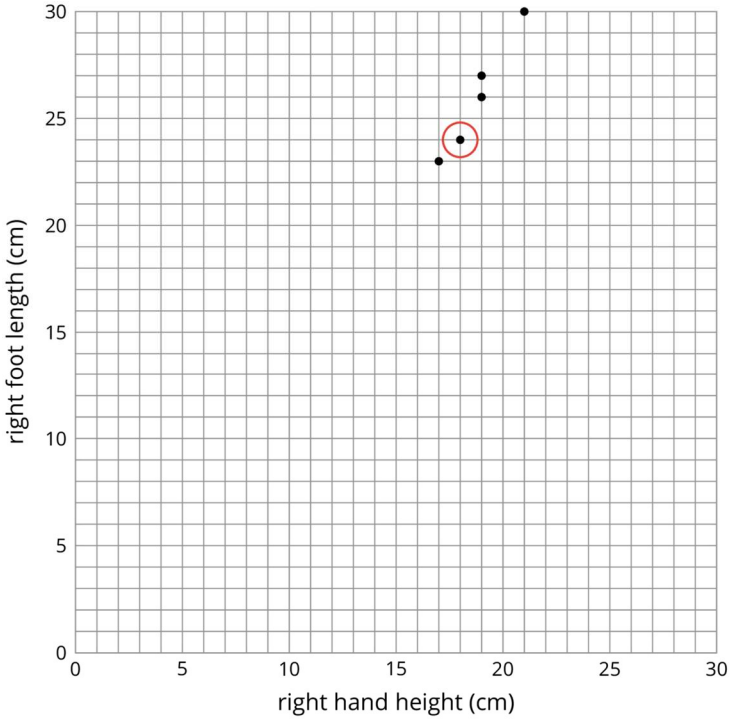
person E	19	26

1. Draw a scatter plot for the data.

2. Circle the point in the scatter plot that represents Person D's measurements.



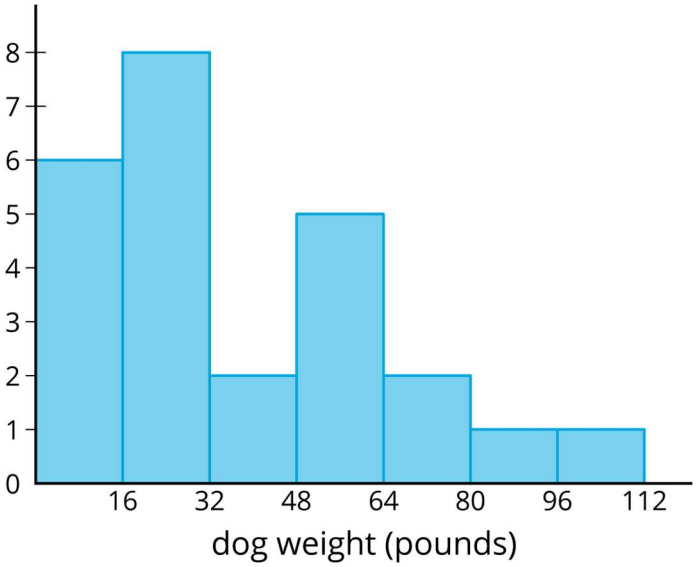
Student Response



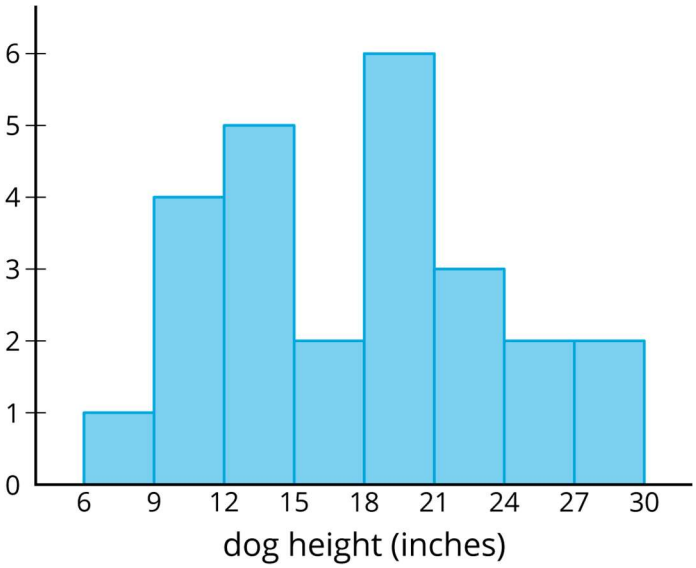
Student Lesson Summary

Histograms show us how measurements of a single attribute are distributed. For example, a veterinarian saw 25 dogs in her clinic one week. She measured the height and weight of each dog.

This histogram shows how the weights of the dogs are distributed.



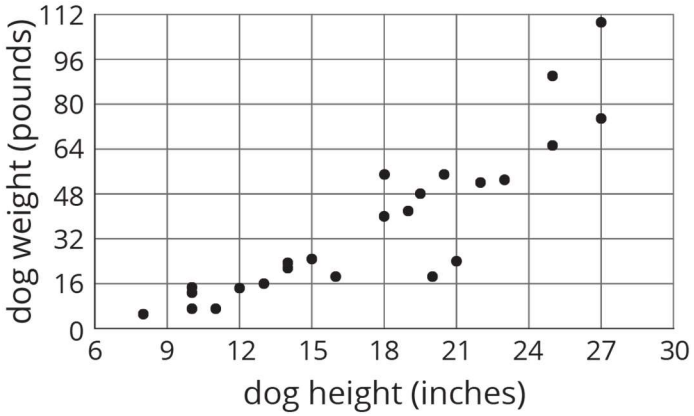
This histogram shows how the heights of the dogs are distributed.



These histograms tell us how the weights of the dogs and how the heights of dogs were distributed. But, they do not give any evidence of a connection between a dog's height and its weight.

Scatter plots allow us to investigate possible connections between two attributes. In this example, each plotted point corresponds to one of the 25 dogs, and its coordinates tell us

the height and weight of that dog. Examination of the scatter plot allows us to see a connection between height and weight for the dogs.



Lesson 2 Practice Problems

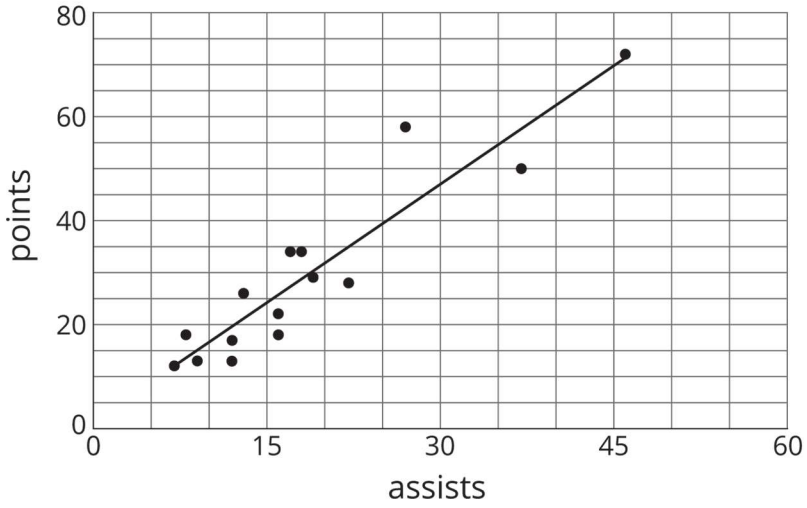
Problem 1 Statement

In hockey, a player gets credited with a "point" in their statistics when they get an assist or goal. The table shows the number of assists and number of points for 15 hockey players after a season.

assists	points
22	28
16	18
46	72
19	29
13	26
9	13
16	22
8	18
12	13
12	17
37	50
7	12
17	34
27	58
18	34

Make a scatter plot of this data. Make sure to scale and label the axes.

Solution



Problem 2 Statement

Select **all** the representations that are appropriate for comparing bite strength to weight for different carnivores.

a. Histogram

b. Scatter plot

c. Dot plot

d. Table

e. Box plot

Solution ["B", "D"]

Problem 3 Statement

When is it better to use a table? When is it better to use a scatter plot?

Solution

Answers vary. Sample response: Scatter plots are best when looking for an overall pattern (or lack of one). Tables are best when looking for the precise details of the data.

Problem 4 Statement

There are many cylinders with radius 6 metres. Let h represent the height in metres and V represent the volume in cubic metres.

a. Write an equation that represents the volume V as a function of the height h.

b. Sketch the graph of the function, using 3.14 as an approximation for π.

c. If you double the height of a cylinder, what happens to the volume? Explain this using the equation.

d. If you multiply the height of a cylinder by $\frac{1}{3}$, what happens to the volume? Explain this using the graph.

Solution

a. $V = 36\pi h$

b. The graph is a line starting from $(0,0)$ then through about $(1,113)$ and $(2,226)$.

c. If you double the height, the volume doubles. Replacing h with $2h$ in the equation gives $V = 36\pi \times 2h = 2(36\pi h)$, double the original volume.

d. If you multiply the height by $\frac{1}{3}$, the volume is also multiplied by $\frac{1}{3}$. On the graph this can be seen using similar triangles, or by noting the relationship is proportional.