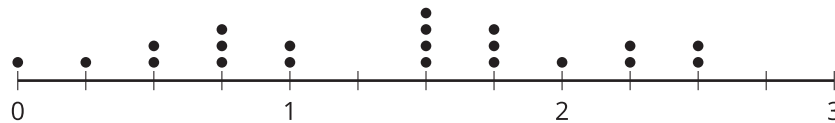


Representing Data

Let's explore different kinds of data.

1.1 Dots of Data

Here is a **dot plot** for a data set.



- Determine if each of these would be an appropriate label to represent the data in the dot plot? Be prepared to explain your reasoning.
 - Number of children per class.
 - Distance between home and school, in miles.
 - Favorite subject in school.
 - Weight of elephants, in pounds.
 - Points received on a homework assignment.
- Think of another label that can be used with the dot plot.
 - Write it below the scale of the dot plot. Be sure to include the unit of measurement.
 - In your scenario, what does one dot represent?
 - In your scenario, what would a data point of 0 mean? What would a data point of $3\frac{1}{4}$ mean?

1.2 What's in the Data?

Ten sixth-grade students at a school were each asked five survey questions. Their answers to each question are shown here.

Data Set A	0	1	1	3	0	0	0	2	1	1
Data Set B	12	12	12	12	12	12	12	12	12	12
Data Set C	6	5	7	6	4	5	3	4	6	8
Data Set D	6	6	6	6	6	6	6	6	6	6
Data Set E	3	7	9	11	6	4	2	16	6	10

1. Here are the five survey questions. Match each question to a data set that could represent the students' answers. Explain your reasoning.

- Question 1: Flip a coin 10 times. How many heads did you get?
- Question 2: How many books did you read in the last year?
- Question 3: What grade are you in?
- Question 4: How many dogs and cats do you have?
- Question 5: How many inches are in 1 foot?

2. How are survey Questions 3 and 5 different from the other questions?



1.3

Been There, Done That!

Priya wants to know if basketball players on 2 teams have had prior experience in international competitions. She gathers data on the number of times the players were on a team before 2016.

Team 1

3 0 0 0 0 1 0 0 0 0 0 0

Team 2

2 3 3 1 0 2 0 1 1 0 3 1

1. Did Priya collect categorical or numerical data?
2. Organize the information on the two basketball teams into these tables.

Team 1

number of prior competitions	frequency (number)
0	
1	
2	
3	
4	

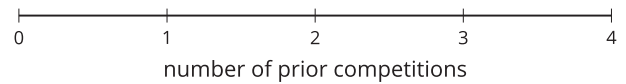
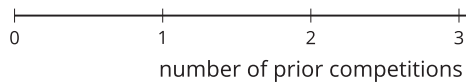
Team 2

number of prior competitions	frequency (number)
0	
1	
2	
3	
4	

3. Make a dot plot for each table.

Team 1

Team 2



4. Study your dot plots. What do they tell you about the competition participation of:

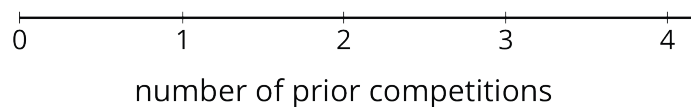
a. The players on team 1?

b. The players on team 2?

5. Explain why a dot plot is an appropriate representation for Priya's data.

Are you ready for more?

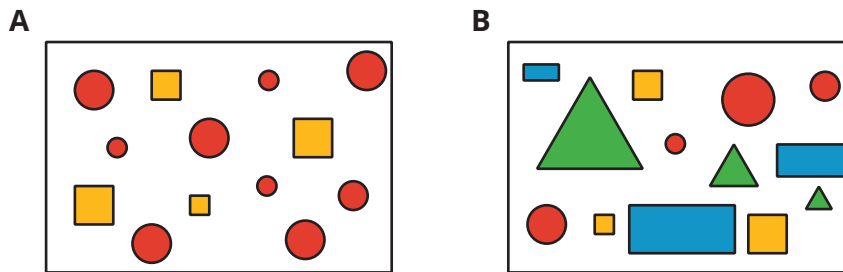
Combine the data for the players on both teams and represent it as a single dot plot. What can you say about the repeat participation of the basketball players?



Lesson 1 Summary

We often collect data to answer questions about something. The data we collect may show **variability**, which means the data values are not all the same.

Some data sets have more variability than others. Here are two sets of figures.



Set A has more figures with the same shape, color, or size. Set B shows more figures with different shapes, colors, or sizes, so Set B has greater variability than Set A.

When a question can be answered only by using data, and we expect that data to have variability, we call it a **statistical question**. Here are some examples.

- Who is the most popular musical artist at your school?
- When do students in your class typically eat dinner?
- Which classroom in your school has the most books?
- How much do the dogs at a pet adoption center weigh?

The table contains data about 10 dogs.

- The weights of the dogs are an example of **numerical data**, which are data that are numbers, quantities, or measurements. The weights of the dogs are measurements in kilograms.
- The dog breeds are an example of **categorical data**, which are data containing values that can be sorted into categories. In this case, there are three categories for dog breeds: pug, beagle, and German shepherd.

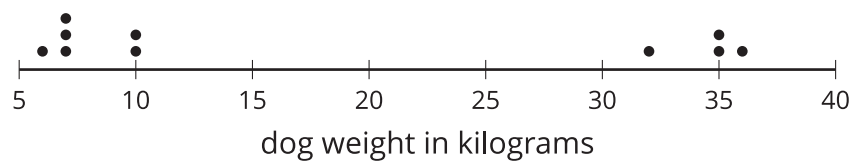
dog name	weight (kg)	breed
Duke	36	German shepherd
Coco	6	pug
Pierre	7	pug
Ginger	35	German shepherd
Lucky	10	beagle
Daisy	10	beagle
Buster	35	German shepherd
Pepper	7	pug
Rocky	7	beagle
Lady	32	German shepherd

When we analyze data, we are often interested in the **distribution**, which is information that shows all the data values and how often they occur. We can see the distribution of the dog weights in a table such as this one.

weight in kilograms	frequency
6	1
7	3
10	2
32	1
35	2
36	1

The term **frequency** refers to the number of times a data value occurs. In this case, we see that there are 3 dogs that weigh 7 kilograms, so 3 is the frequency for the value “7 kilograms.”

Like a frequency table, a **dot plot** also shows the distribution of a data set. This dot plot shows the distribution of dog weights.



A dot plot uses a horizontal number line. We show the frequency of a value by the number of dots drawn above that value. Here, the two dots above the number 35 tell us that there are two dogs that weigh 35 kilograms.