Lesson 1: Posing Statistical Questions

Statistics is about using data to answer questions. In this module, the following four steps will summarize your work with data:

Step 1: Pose a question that can be answered by data.

Step 2: Determine a plan to collect the data.

Step 3: Summarize the data with graphs and numerical summaries.

Step 4: Answer the question posed in Step 1 using the data and summaries.

You will be guided through this process as you study these lessons. This first lesson is about the first step – what is a statistical question, and what does it mean that a question can be answered by data?

Classwork

Example 1: What is a Statistical Question?

Jerome, a 6th grader at Roosevelt Middle School, is a huge baseball fan. He loves to collect baseball cards. He has cards of current players and of players from past baseball seasons. With his teacher's permission, Jerome brought his baseball card collection to school. Each card has a picture of a current or past major league baseball player, along with information about the player. When he placed his cards out for the other students to see, they asked Jerome all sorts of questions about his cards. Some asked:

- How many cards does Jerome have altogether?
- What is the typical cost of a card in Jerome's collection?
- Where did Jerome get the cards?

Exercises 1–5

1. For each of the following, determine whether or not the question is a statistical question. Give a reason for your answer.

   a.   Who is my favorite movie star?

   b.   What are the favorite colors of 6th graders in my school?

c.   How many years have students in my school's band or orchestra played an instrument?

d.   What is the favorite subject of 6th graders at my school?

e.   How many brothers and sisters does my best friend have?

2.   Explain why each of the following questions is not a statistical question.

a.   How old am I?

b.   What's my favorite color?

c.   How old is the principal at our school?

3.   Ronnie, a 6th grader, wanted to find out if he lived the farthest from school. Write a statistical question that would help Ronnie find the answer.

4.   Write a statistical question that can be answered by collecting data from students in your class.

5.   Change the following question to make it a statistical question: "How old is my math teacher?"

Example 2: Types of Data

We use two types of data to answer statistical questions: numerical data and categorical data. If we recorded the age of 25 baseball cards, we would have numerical data. Each value in a numerical data set is a number. If we recorded the team of the featured player for 25 baseball cards, you would have categorical data. Although you still have 25 data values, the data values are not numbers. They would be team names, which you can think of as categories.

Exercises 6–7

6. Identify each of the following data sets as categorical (C) or numerical (N).

    a. Heights of 20 6th graders _____

    b. Favorite flavor of ice cream for each of 10 6th graders _____

    c. Hours of sleep on a school night for 30 6th graders _____

    d. Type of beverage drank at lunch for each of 15 6th graders _____

    e. Eye color for each of 30 6th graders _____

    f. Number of pencils in each desk of 15 6th graders _____

7. For each of the following statistical questions, students asked Jerome to identify whether the data are numerical or categorical. Explain your answer, and list four possible data values.

    a. How old are the cards in the collection?

    b. How much did the cards in the collection cost?

    c. Where did you get the cards?

Lesson Summary

A **statistical question** is one that can be answered by collecting data that vary (i.e., not all of the data values are the same).

There are two types of data: numerical and categorical. In a **numerical data set**, every value in the set is a number. **Categorical data sets** can take on non-numerical values, such as names of colors, labels, etc. (e.g., "large," "medium," or "small").

Statistics is about using data to answer questions. In this module, the following 4 steps will summarize your work with data:

Step 1: Pose a question that can be answered by data.

Step 2: Determine a plan to collect the data.

Step 3: Summarize the data with graphs and numerical summaries.

Step 4: Answer the question posed in Step 1 using the data and the summaries.

Problem Set

1. For each of the following, determine whether the question is a statistical question. Give a reason for your answer.
   a. How many letters are in my last name?
   b. How many letters are in the last names of the students in my 6th grade class?
   c. What are the colors of the shoes worn by the students in my school?
   d. What is the maximum number of feet that roller coasters drop during a ride?
   e. What are the heart rates of the students in a 6th grade class?
   f. How many hours of sleep per night do 6th graders usually get when they have school the next day?
   g. How many miles per gallon do compact cars get?

2. Identify each of the following data sets as categorical (C) or numerical (N). Explain your answer.
   a. Arm spans of 12 6th graders
   b. Number of languages spoken by each of 20 adults
   c. Favorite sport of each person in a group of 20 adults
   d. Number of pets for each of 40 3rd graders
   e. Number of hours a week spent reading a book for a group of middle school students

3. Rewrite each of the following questions as a statistical question.
   a. How many pets does your teacher have?
   b. How many points did the high school soccer team score in its last game?
   c. How many pages are in our math book?
   d. Can I do a handstand?

4.   Write a statistical question that would be answered by collecting data from the 6th graders in your classroom.

5.   Are the data you would collect to answer that question categorical or numerical? Explain your answer.