**7.6F
7.12B**   **The student is expected to use data from a random sample to make inferences about a population.**

ℹ TELL ME MORE...

When information is being gathered about a group, the entire group of objects, individuals, or events is called the **population**.

A **sample** is part of the population that is chosen to represent the entire group. A sample in which every person, object, or event has an equal chance of being selected is called a **random sample**.

Sampling bias may skew the results of a survey. If the sample is created in such a way that some members of the intended population are less likely to be included than others, then the results may have sampling bias and may not reflect the characteristics of the entire population.

Data From a Random Sample
A vegetable garden has 36 tomato plants arranged in a 6-by-6 array. The gardener wanted to know the average number of tomatoes on the plants.

Random or Bias?

For a sample to be random, every member of the population must have an equal chance of being selected for the sample.

Suppose you want to know the most popular song among 7th grade students at your school. The **population** is all 7th grade students, so your **sample** must be randomly selected from this population.

Plan 1: Ask all students in the choir to tell you their favorite song.

This survey will have bias, because not all 7th grade students are in the choir. This survey will only tell you which song is the favorite of most choir members.

Plan 2: Ask every 20th student at the main entrance of the 7th grade wing of the school in the morning to tell you their favorite song.

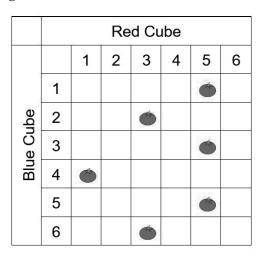*This sample is random since every 7th grader has an equal chance of being selected.*

Counting the number of tomatoes on *all* of the plants is too time-consuming, so the gardener decided to choose plants at random, count the number of tomatoes of those plants that are chosen, then calculate the mean number of tomatoes on each plant from the data he has collected. He rolled one red number cube and blue number cube. The gardener made a table where the column indicates the number shown on the red cube and the row represents the number shown on the blue cube. He repeated this process six times to generate a random sample. He counted the tomatoes on the plants where the two numbers rolled intersect on the table.

The tomato images at right show the results of rolling the number cubes 6 times. The gardener counted the number of tomatoes on the 6 plants, then used that data to find the mean number of tomatoes on each plant.

$$\frac{12 + 20 + 18 + 9 + 17 + 14}{6} = \frac{90}{6} = 15$$

Based on his random sample, the gardener can make a conclusion or **infer** that each tomato plant will produce about 15 tomatoes.

	Red Cube					
Blue Cube	1	2	3	4	5	6
1					🍅	
2			🍅			
3					🍅	
4	🍅					
5					🍅	
6			🍅			

✓ EXAMPLES

EXAMPLE 1: Ms. Attia teaches six classes of 7th grade math with a total of 150 students. She wants to find out how much time her students are spending on homework each week, so she randomly selected 10 students and asked them how many hours they spend each week on math homework. The results of her survey are shown.

Which of the following inferences are supported by the data?

Math Homework

			📖	
			📖	
		📖	📖	📖
📖	📖	📖	📖	📖
0	**1**	**2**	**3**	**4+**

Number of Hours per Week
📖 = 1 student

    A. Most of Ms. Attia's students spend 3 hours or less on math homework each week.
    B. At least 15 of Ms. Attia's students spend 4 or more hours per week on math homework.
    C. Most of Ms. Attia's students dislike doing math homework each week.

STEP 1 Evaluate whether or not the data support Inference A.

- According to the graph, 1 student spends 0 hours, 2 students spend 1 hour, 2 students spend 2 hours, and 4 students spend 3 hours.
- $1 + 2 + 2 + 4 = 9$, so 9 out of 10 students spend 3 hours or less on math homework.

Inference A is supported by the data.

STEP 2 Evaluate whether or not the data support Inference B.

- According to the graph, 1 student spends 4 hours or more on math homework.
- $\frac{1}{10} = \frac{1 \times 10}{10 \times 10} = \frac{10}{100} = 10\%$, so 10% of the students surveyed spend 4 hours or more on math homework each week.
- Determine 10% of the 150 students that Ms. Attia teaches.
  - $\frac{10}{100} = \frac{x}{150}$
  - $\frac{10 \times 1.5}{100 \times 1.5} = \frac{15}{150}$
  - 15 students is 10% of the total of 150 students Ms. Attia teaches.
- A random survey that is sufficiently large tends to be representative of the population.

Inference B is supported by the data.

STEP 3 Evaluate whether or not the data support Inference C.

- The graph shows how much time students spend on their math homework.
- We do not know how students feel about their math homework from this survey.

Inference C is not supported by the data.
Inferences A and B are supported by the data.

YOU TRY IT!

Jamonte wants to conduct a survey on how often people who belong to a gym exercise each week. He has two plans on how to gather this data.
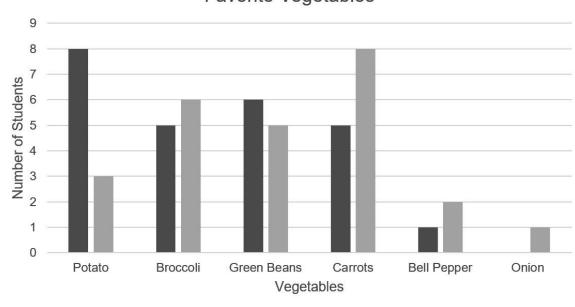
Plan 1: Go to a shopping mall and ask people to respond to the survey as they exit the mall.

Plan 2: Go to *Workout World* gym and ask people to respond to the survey as they exit the gym.

Which sample is more random? Why?

EXAMPLE 2: A school cafeteria randomly asked 25 boys and 25 girls to identify their favorite vegetable from a list of potato, broccoli, green beans, carrots, bell pepper, and onion. Each day, approximately 1,500 students eat lunch in the cafeteria. The results of that survey are shown in the graph below.

Favorite Vegetables

Which of the following inferences are supported by the data?

    A.   Potato is the favorite vegetable for about 16% of all the boys at the school.

    B.   There are no students in the school whose favorite vegetable is corn.

    C.   Broccoli is the favorite vegetable for the same number of girls as potato, bell pepper, and onion combined.

STEP 1   Evaluate whether or not the data support Inference A.

- According to the graph, 8 boys out of the 25 boys surveyed chose potatoes as their favorite vegetable.

- $\frac{8}{25} = \frac{32}{100} = 32\%$, so 32% of the boys surveyed said that potato was their favorite vegetable.

- Since the sample is random, you can conclude that about 32% of all boys in the school would prefer potato.

Inference A is not supported by the data.

STEP 2   Evaluate whether or not the data support Inference B.

- Students were not given the option to select "corn" in this survey.

- Therefore, we do not have any information that would tell us how many students have corn as a favorite vegetable.

Inference B is not supported by the data.

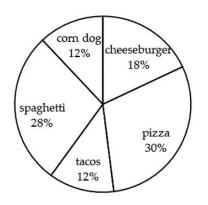**STEP 3** Evaluate whether or not the data support Inference C.

- According to the graph, 6 girls chose broccoli as their favorite vegetable. 3 girls chose potato, 2 girls chose bell pepper, and 1 girl chose onion.
- Add the numbers for potato, bell pepper, and onion: 3 + 2 + 1 = 6.
- The number of girls who have broccoli as a favorite vegetable is equal to the total number in this group.

Inference C is supported by the data.

PRACTICE

Use the circle graph for questions 1-3.

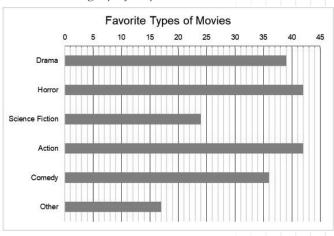**Favorite Lunch**

Use the bar graph for questions 4-6.

1. The cafeteria manager took a poll of 150 randomly selected students to find the most popular lunch item offered in the school cafeteria.
Which of the following statements is true about the data represented in the graph?

   A Most of the students polled were male.

   B About half of the students refused to eat pizza or spaghetti.

   C The same number of students prefer corn dogs as tacos.

   D Twice as many students prefer pizza as cheeseburgers.

2. Write your own valid inference from the data represented in the circle graph.

3. Validate your inference with data support from the information in the problem.

4. Paul surveyed 200 randomly selected people in the mall and asked them what genre of movie they prefer. Their results are shown in the bar graph.
Which of the following statements is **not** true about the results of the survey?

   F *Horror* and *Action* are the two most popular movie genres.

   G The number of people who selected *Other* represents 34% of the people surveyed.

   H 3 more people preferred *Drama* than *Comedy*.

   J *Science Fiction* and *Comedy* combined received 30% of the votes.

5. Write your own valid inference from the data represented in the bar graph.

6. Validate your inference with data support from the information in the bar graph.

7. Rhonda randomly selected 15 light bulbs out of a set of 100 that were being sold. She tested them to see how many days they would remain lit. She recorded the results in the table shown.

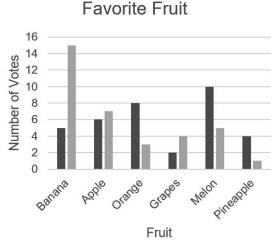| Number of Day	Number of Light Bulbs
10	1
11	2
12	2
13	5
14	3
15	2

Rhonda made the following conclusions about the data she collected.

I. A light bulb may remain lit as few as 10 days or as many as 15 days.

II. 20% of the light bulbs Rhonda tested remained lit for exactly 14 days.

III. The majority of the light bulbs tested remained lit for fewer than 14 days.

IV. The mean number of days a light bulb remained lit was between 12 and 13.

Which of the observations are true?

F I and III only

G I, II, and IV only

H II and IV only

J All are true.

8. Write your own valid inference from the data represented in the table.

9. Validate your inference with data support from the information in the problem.

10. A survey of 35 adults and 35 children was taken at a local farmers' market. One of the farmers asked each of them to select their favorite fruit from the fruit he was selling. The results are shown above. Explain why this survey could be considered bias.

Favorite Fruit