In Lesson 26 you learned about random samples. In this lesson you will look at random samples to make predictions. Take a look at this problem.

Ms. Jennings held a “mystery bag” full of marbles up in front of her class. She said, “There are 100 marbles in this bag. Some are red, and all of the others are blue. I want you to estimate how many red marbles are in the bag without looking in the bag and counting all of them.” She then let Joe pick 10 marbles from the bag without looking.

Explore It

Use the math you already know to solve the problem.

- Suppose Joe had 4 red marbles in his sample. Write a ratio representing the number of red marbles in Joe’s sample.

- If Joe is not allowed to pull out any more marbles, what do you think his best estimate will be for the total number of red marbles in the mystery bag? Explain.

- Ms. Jennings had Joe put his 10 marbles back in the bag and shook it up. Then she let Angela pull 10 marbles from the bag without looking. Angela had 6 red marbles in her sample. Why do you think Angela’s sample was different from Joe’s?

- If Angela is not allowed to draw any more samples, and she does not know about Joe’s sample, what do you think her best estimate would be for the total number of red marbles in the mystery bag? Explain.

- Ms. Jennings had Angela put her marbles back in the bag, shook the bag up again, and let Isabella reach in and draw out 10 marbles without looking. Isabella got 9 reds in her sample. Why might Isabella’s sample be different from Joe’s and Angela’s?
Random samples can differ from one another due to **random variation**. The amount of variation may be small, as in the difference between Joe’s and Angela’s samples from the mystery bag. Occasionally, the variation is much larger, as in the difference between Joe’s and Isabella’s samples.

In order to get a better estimate of the number of red marbles in the mystery bag, one thing that can help is to draw more random samples. That would help you judge how typical the samples from Joe, Angela, and Isabella are. If you see a number of samples clustering around the same value, you can be more confident about using that value to estimate what you would like to know about the population.

Random sampling is helpful in situations beyond the mystery bag scenario. It can be useful, for example, when conducting surveys. Suppose, for example, you want to make a prediction about who will win a school election, but you don’t have the time or resources to survey everyone in school. You could draw a random sample of names from a list of students in school and administer the survey to them. The random sample could be as large as time and resources allow.

**Reflect**

1. Suppose Joe, Angela, and Isabella had drawn samples of 25 marbles instead of 10. How might this change their results?
Read the problem below. Then explore different ways to represent distributions of
statistics from random samples.

Ms. Jennings decided to let all 20 students in her class draw random samples from
the mystery bag. Each student was allowed to draw a sample of 10 marbles, count
the number of reds in the sample, and then put the 10 marbles back in the bag.
Angela thinks the extra samples will help the class make a better estimate of the
number of reds in the mystery bag. How should students organize their results to
get a good estimate?

Model It

Joe suggested organizing the results in the following table.

<table>
<thead>
<tr>
<th>Number of Reds</th>
<th>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of Reds</td>
<td>0.8 1.0 0.7 1.0 0.7 0.5 0.9 0.8 0.7 0.9 0.7 0.6 0.8 0.8 0.4 0.7 0.5 0.9 0.8 0.6</td>
</tr>
</tbody>
</table>

The table shows the proportion of reds obtained in each sample.

Model It

Angela thought it would be easiest to organize the data with a dot plot.

Model It

Isabella decided to construct a box plot to organize the class data.
Now you will represent your results of taking random samples.

2 Make your own mystery bag with 100 marbles (or similar objects). Decide how many objects will be red and how many will be blue. Record the number of reds in the bag and the number of blues in the bag here. ________________________________

3 Pretend to be each of the 20 students in Ms. Jennings’ class. Repeat the experiment. Take 10 marbles out of the bag without looking. Count the number of reds and record that number in a table. Replace the marbles and repeat until you have 20 samples.

4 Use your table to make a dot plot.

5 Use your dot plot to make a box plot.

6 Imagine someone who doesn’t know how many red marbles you put into the mystery bag. Which representation (table, dot plot, or box plot) would best help them estimate the number of reds in the mystery bag? Explain.

Try It

Use what you just learned about representing data to solve this problem.

7 Suppose one of your classmates used a different number of reds. Her dot plot is shown below. What is a good estimate for the number of reds in her bag? Explain.

Marbles Experiment 2

<table>
<thead>
<tr>
<th>Proportion of Red Marbles</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9</td>
</tr>
</tbody>
</table>

______________________________

______________________________

______________________________
Read the problem below. Then explore different ways to compare distributions of statistics from random samples when the samples are different sizes.

Angela asked Ms. Jennings if they could re-do the mystery bag activity. Angela wanted to change just one thing about it: she asked if each student could draw 25 marbles instead of just 10 marbles. Ms. Jennings agreed, and the class re-did the activity in the way Angela asked. How do the results of this new activity compare to the results of the previous one?

---

Model It

You can stack dot plots to compare the two distributions of results.

Marbles Experiment 1

Marbles Experiment 3

Proportion of Red Marbles

You can stack box plots to compare the two distributions of results.

Marbles Experiment 1

Marbles Experiment 3

Proportion of Red Marbles
Lesson 27

Part 3: Guided Instruction

Connect It

Now you will use the dot plots and box plots to compare the two data distributions.

8 Look at the models on the previous page. Compare the shapes of the two distributions of data shown in the dot plots from the two experiments.

9 Compare the centers of the two distributions to one another.

10 Compare the spreads. Which distribution looks more spread out? Why?

11 If you could only take one sample from the mystery bag in order to make a prediction, would you rather draw a sample of 10 marbles or a sample of 25 marbles? Explain your choice by referring to the representations of the two distributions.

Try It

Use what you just learned about data distributions of data to solve this problem.

12 Perform the activity again, but this time draw 20 samples of only 5 marbles each. Record your results in a dot plot. How does the distribution compare to the dot plots you produced when you drew samples of 10?
Read the situation described below. Then solve problems 13–15.

Tammy was running for class president. She wanted to know if she had a good chance of winning the election, so she decided to have her friends help her with surveys.

Jonathan surveyed a random sample of 10 students from the school. 7 of them said they would vote for Tammy. If there are 230 students in the school, and all of them vote, what is Jonathan’s best estimate of the number of students who would vote for Tammy?

Look at how you could show your work.

\[
\frac{7}{10} \text{ of the students said they would vote for Tammy.} \]

\[
\frac{7}{10} = 0.7
\]

\[
0.7 \times 230 = 161
\]

Solution: 161

13 Kimberly had more time to conduct surveys and decided to survey a random sample of 30 students from the school. In Kimberly’s sample, 24 students said they would vote for Tammy. What is Kimberly’s best estimate of the number of students in the school who would vote for Tammy?

Solution: _____________________________

Pair/Share

What are some reasons someone would have to draw a small sample rather than a large one?
14 Who is likely to have a better estimate of the number of students who will vote for Tammy: Jonathan or Kimberly? Explain your choice by using what you learned in this lesson.

Solution: 

Whose sample size was larger?

15 Which size random sample is likely to provide the most trustworthy results?

A 5
B 10
C 30
D There is no difference.

Katie chose D as the correct answer. How did she get that answer?

How big do you think a “good” random sample has to be?

How does sample size affect results?

Pair/Share How would you help Katie understand her error?
Solve the problems. Mark your answers to problems 1–3 on the Answer Form to the right. Be sure to show your work.

1. Which distribution is likely to have the greatest amount of spread?
   A. 50 proportions computed from random samples of size 30
   B. 50 proportions computed from random samples of size 15
   C. 50 proportions computed from random samples of size 5
   D. The spread would likely be the same in each case.

2. Thomas wanted to estimate how many of the 534 students in his school owned cell phones. He took a random sample of 25 students. In his sample, 15 students said they owned cell phones. What is the best estimate for the number of students in the school who own cell phones?
   A. 320
   B. 509
   C. 890
   D. 15
3 Jamie wanted to estimate the mean word length in her science textbook. She did not have time to count every word and compute the average, so she took a random sample of 50 words. Which of these is the best thing to do to estimate the mean word length in the book?

A  Find the average word length for the words in her sample and then multiply by 50.
B  Find the average word length for the words in her sample and use it as the prediction for average word length for the entire book.
C  Use a smaller sample of words to reduce the amount of spread in the distribution of averages.
D  Draw a random sample of 50 words from a different science textbook and compare the new results to the original results.

4 Suppose you flipped a fair coin 15 times. Then, 11 of your friends did the same thing. Without actually flipping a coin, make a table to show a realistic set of data for this situation. Your table should show the percentage of heads obtained by each person doing the coin flips.

Organize the data in the table you constructed using a dot plot.

Suppose you and each of your friends decided to flip the coin 50 times each instead of 15. Explain how the dot plot above would change and sketch a predicted dot plot for the situation below.
Lesson 27 (Student Book pages 254–263)

Making Statistical Inferences

LESSON OBJECTIVES

• Use data from two samples to write ratios that can be easily used to make an estimate about a population.
• Compare estimates made from multiple samples of the same size to gauge the variation in the estimates.
• Predict the accuracy of the estimates made by various samples.

PREREQUISITE SKILLS

• Recognize a statistical question as one that assumes variability.
• Recognize samples that represent a population and those that are biased.
• Understand how to obtain a sample of a population.
• Write and solve proportions.

VOCABULARY

random variation: a variable is subject to random variation if its value is not predictable

THE LEARNING PROGRESSION

In Grade 6, students have been using data, both categorical and measurement, to answer simple statistical questions, but have paid little attention to how the data were selected. Students identified questions that result in statistical variability and those that do not. They analyzed data using measures of central tendency and used data displays to present information.

Earlier in Grade 7, students developed an understanding of samples of a population. They learned about bias in a sample and learned that a more accurate prediction will result when you use a sample that reflects the population. In this lesson, students will use data from multiple samples to make estimations, gauge the variation of the estimates, and predict the accuracy of the estimates.

In Grade 8, students will construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities.

CCLS Focus

7.SP.2 Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions. For example, estimate the mean word length in a book by randomly sampling words from the book; predict the winner of a school election based on randomly sampled survey data. Gauge how far off the estimate or prediction might be.

STANDARDS FOR MATHEMATICAL PRACTICE: SMP 1–3, 5–7 (see page A9 for full text)
Part 1: Introduction

AT A GLANCE

Students read a word problem and explore the use of random samples to make predictions.

STEP BY STEP

• Tell students that this page models how to use random samples to make predictions.

• Have students read the problem at the top of the page. Have them restate it in their own words.

• Work through Explore It as a class.

• Discuss why it makes sense for the ratio of red marbles to total number of marbles in Joe’s sample to be equivalent to the ratio of red marbles to total number of marbles in his prediction.

SMP Tip: When students use samples to predict the total number, they must reason abstractly and quantitatively (SMP 2). Encourage them to apply math concepts they already have, such as the concept of equivalent ratios, in new contexts.

• As students consider why samples are often different, mention the idea of chance or luck in drawing a sample.

• Have students compare the three samples. Note that it is possible, but not likely, for samples to be as different as Joe’s and Isabella’s are.

• Ask student pairs or groups to explain their answers for the questions about sampling.

ELL Support

• Have students suppose you have several types of juice. Ask, What do I mean if I ask you to sample each one? [Take just a sip to see what it tastes like.]

• Discuss how a sample is different from the whole. Note that people sample a food or beverage to get a taste or an idea of what it is like.

• Connect the idea of sampling without eating the whole thing to sampling a population without surveying everyone.

Mathematical Discourse

• Suppose the jar of 100 marbles contains 60 red marbles and you close your eyes and pick 10. Name a likely and an unlikely number of red marbles that you would get. Explain.

  Students may say that 6 is likely because \( \frac{6}{10} = \frac{60}{100} \). They may also list numbers close to 6. Numbers not close to 6 are unlikely because the ratio they form is not especially close to \( \frac{60}{100} \).

• Suppose the jar of 100 marbles contains 5 blue marbles. You close your eyes and pick 10. Name a likely and an impossible number of blue marbles that you would get. Explain.

  Students may say that it is likely you will get 0 or 1 because \( \frac{5}{100} < \frac{1}{10} \). They should say that it is impossible to get more than 5 blue marbles because only 5 are in the jar.
Students learn about random variation and how to use samples to estimate total populations.

Random samples can differ from one another due to random variation. The amount of variation may be small, as in the difference between Joe's and Angela's samples from the mystery bag. Occasionally, the variation is much larger, as in the difference between Joe's and Isabella's samples.

In order to get a better estimate of the number of red marbles in the mystery bag, one thing that can help is to draw more random samples. That would help you judge how typical the samples from Joe, Angela, and Isabella are. If you see a number of samples clustering around the same value, you can be more confident about using that value to estimate what you would like to know about the population.

Random sampling is helpful in situations beyond the mystery bag scenario. It can be useful, for example, when conducting surveys. Suppose, for example, you want to make a prediction about who will win a school election, but you don't have the time or resources to survey everyone in school. You could draw a random sample of names from a list of students in school and administer the survey to them. The random sample could be as large as time and resources allow.

Suppose Joe, Angela, and Isabella had drawn samples of 25 marbles instead of 10. How might this change their results?

Possible answer: Their estimates for the number of reds in the whole bag would probably be closer together.

Encourage students to think of everyday situations in which people must predict how many of an item they will need.

Examples: Restaurant managers predict how much food customers will order, store managers predict how many of each color and size of an item to order, politicians predict what percentage of various groups will vote for them.
Students use tables and graphs to represent the results of multiple samples.

STEP BY STEP

• Read the problem at the top of the page as a class.
• Look at the first Model It. Have a student explain how he or she could use the number of red marbles to find the proportion of red marbles. Discuss why the table makes it difficult to picture the data as a whole.
• Have students look at the dot plot in the second Model It. Ask a student to explain what is meant by the word distribution. Have students describe the distribution shown by the dot plot. Model the word cluster to help them describe it efficiently.
• Have students compare and contrast the dot plot with the box plot in the third Model It. Discuss the different types of information given by each.

Hands-On Activity

Understand dot plots to show results of multiple trials.

Materials: 2 number cubes and 5 sticky dots for each pair of students

• Draw a number line on the board. Number it from 0 through 10. Title it “How Many Nines?”
• Tell students that in a game, the goal is to roll a sum of 9 the most times on 2 number cubes in 10 rolls. Have them play the game 5 times and record the number of 9s they roll each time.
• Have students go to the board and use sticky dots to record how many times they rolled 0 nines, 1 nine, 2 nines, and so on.
• Help students interpret the box plot. Stress that the number line shows the number of nines rolled in 10 rolls. The dots show how many times that number occurred.

Mathematical Discourse

• How do you make each type of table or graph?
  Students should describe how to find the proportions when making the table. They should explain how to place dots on a number line to record the frequencies of red marbles. They should go through the steps of making a box plot, including finding the mean.
• What type of questions could you answer using the information from each graph or table?
  Students may say that the table shows the proportions but are not in any special order. The dot plot would answer questions about how often certain frequencies occurred and which are most common as well as questions about spread. The box plot could be used to name the median, maximum, minimum, outliers, and spread.
Students revisit the problem on page 256 and conduct their own experiments using random samples.

### STEP BY STEP

- **Read Connect It as a class.**
- **Read the first two items.** Have students explain in their own words what they are supposed to do.
- **Once students have taken and recorded 20 samples,** instruct them to display their results using a table, a dot plot, and a box plot. Review how to make each if necessary.

**SMP Tip:** When students use dynamic graphing software or graphing calculators to create box plots, they are learning to use appropriate tools strategically (SMP 5). The tools allow them to focus on the distribution of the data rather than the creation of box plots.

- **As students evaluate the models to see which is most useful for estimating the number of red marbles in the bag,** make sure they explain the reason for their choice.

### TRY IT SOLUTIONS

**7 Solution:** A number between 50 and 60; Students may note the clusters of dots at 0.5 and 0.6, which correspond to 50 and 60 red marbles.

**ERROR ALERT:** Students who wrote 0.5 or 0.6 found the proportion of marbles that are red rather than the number of marbles which should be a whole number.
Students use models to compare samples of different sizes.

**STEP BY STEP**

- Read the problem at the top of the page as a class.
- Have students explain how the experiment is different from the one on the previous page.
- Study the dot plots in the first Model It. As students compare the two plots, encourage them to look at the modes and the ranges.
- Ask students to contrast the two box plots in the second Model It. Talk about the width of the interquartile ranges and the presence of outliers.

**Visual Model**

- On the board, draw approximately 30 red circles, 20 green circles, 10 yellow circles, and 5 blue circles.
- Tell students to pretend that these are marbles in a bag and you want to get an idea of the colors of the marbles without looking at all of them.
- Ask, If I close my eyes and choose 1 marble, will I get a good idea of the contents of the bag? Why or why not? [No, you would only see one color.]
- Say, There are 4 colors, so if I choose 4 marbles, will I get a good idea of the contents of the bag? Why or why not? [No, you are more likely to get more than 1 red than a single blue.]
- Continue with the line of questioning until students see that as you take more marbles out of the bag, you get a better idea of its contents.

**Mathematical Discourse**

- The number lines for both plots are the same. Why does that work if one plot displays the results of samples of 10 and the other displays the results of samples of 25? Students should communicate that it is the number lines show the ratio of reds to total in the sample as a decimal. The plots are comparing ratios, not absolute numbers.
- How do you write the ratios as decimals? You can find an equivalent ratio with a denominator of 100 and then write it as a decimal. You can also divide the numerator by the denominator.
Part 3: Guided Instruction

### AT A GLANCE

Students revisit the problem on page 258 and compare the distribution of data.

### STEP BY STEP

- **Read Connect It as a class.** Be sure to point out that the questions refer to the problem on page 258.
- **Ask students to explain the difference between the two data sets used to make each pair of graphs.** Have them take this difference into account as they compare the centers and spread shown on the graphs.

**SMP Tip:** When students compare graphs to examine the effect of sample size on accuracy, they are modeling with mathematics (SMP 4). As you discuss the center and spread on each graph, be sure to relate these features to the context of the problem to reinforce the connection between the model and the actual situation.

- **Review the concept that bigger samples are more likely to result in more accurate representations of the actual data than smaller samples.**

### TRY IT SOLUTIONS

12 **Solution:** Answers will vary but may reflect that the shape and center of the data are the same, but the smaller samples result in a greater spread; Students may conduct 20 trials and display the results using a dot plot.

**ERROR ALERT:** Students who made a box plot did not read the instructions carefully.
Students use data from random samples to make predictions about a population.

**STEP BY STEP**

- Ask students to solve the problems individually and to show their work.
- When students have completed each problem, have them Pair/Share to discuss their solutions with a partner or in a group.

### SOLUTIONS

**Ex** The total number of students is multiplied by the fraction of students who supported Tammy in the survey.

13 **Solution:** 184; Students could solve the problem by multiplying the total number of students by the fraction of Tammy supporters in the survey.

14 **Solution:** Kimberly; Students could solve the problem by noting that larger random samples tend to give more accurate results.

15 **Solution:** C; Larger random samples are less spread out and tend to give more accurate estimates. Explain to students why the other two answer choices are not correct:

- A is not correct because it is less than 30.
- B is not correct because it is less than 30.
Students use what they know about random samples to answer questions and solve word problems that might appear on a mathematics test.

**STEP BY STEP**

- First, tell students that they will use what they know about random samples to answer questions and solve word problems. Then have students read the directions and answer the questions independently. Remind students to fill in the correct answer choices on the Answer Form.

- After students have completed the Common Core Practice problems, review and discuss correct answers. Have students record the number of correct answers in the box provided.

**SOLUTIONS**

1. **Solution:** A; Find the fraction of cell phone owners, \(\frac{15}{25}\) or 0.6, and multiply it by the total students, \(0.6 \times 534\)

2. **Solution:** For example, the new dot plot should show less spread than the previous one.
Assessment and Remediation

- Ask students to predict the number of purple jellybeans in a bag of 200 jellybeans if a random sample of 25 resulted in 4 purple jellybeans. [32]
- For students who are struggling, use the chart below to guide remediation.
- After providing remediation, check students’ understanding. Ask students to predict the number of pink jellybeans in a bag of 300 jellybeans if a random sample of 20 resulted in 3 pink jellybeans. [45]
- If a student is still having difficulty, use Ready Instruction, Level 7, Lesson 26.

<table>
<thead>
<tr>
<th>If the error is . . .</th>
<th>Students may . . .</th>
<th>To remediate . . .</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>have not connected the sample with the total population.</td>
<td>Have students reread the problem and explain what it means in their own words.</td>
</tr>
<tr>
<td>8</td>
<td>have multiplied by $\frac{1}{25}$, not $\frac{4}{25}$.</td>
<td>Have students find the fraction of jellybeans in the sample as both a fraction and a decimal.</td>
</tr>
<tr>
<td>16</td>
<td>have multiplied the fraction by 100, not 200.</td>
<td>Point out that there are 200 jellybeans in the bag.</td>
</tr>
</tbody>
</table>

Hands-On Activity

Make predictions using random samples.

Fill a bag with 100 colored marbles. At least 30 of them should be blue. Count how many blue marbles you use as you fill the bag, but do not tell the students. Tell students that they are to guess how many of the 100 marbles are blue.

Divide students into 3 groups. Students in the first group will each take a random sample of 10 marbles. Students in the second group will each take a random sample of 30 marbles. Students in the third group will each take a three random sample of 10 marbles each.

After students make their guesses based on their samples, reveal how many blue marbles are actually in the bag. Talk about which samples resulted in the most accurate guesses.

Challenge Activity

Use samples to make predictions.

Have each student think of a survey question about students’ preferences such as favorite type of music, books, or lunch. Have them survey between 20 and 50 students. Remind them about the importance of using a representative random sample. Tell them the total number of students enrolled in the school. Have them use the results to predict how many students in the entire school would give each response.