

Statistics and Probability

Statistics is the study of how data are collected, summarized, and presented. **Probability** is the study of chance.

Suppose that a student rolls two dice, adds the outcomes, and records the result; rolls again, adds, and records; and continues until 100 sums have been recorded. It is likely that the sums will range from 2 to 12 and that more sums will fall toward the middle of this range than will fall toward either end. The collection of sums is a data *distribution*.

The focus in statistics and probability at this level is to teach students about statistical distributions. Students will learn that a distribution can be described in terms of its center and spread. In the dice example, the center of the distribution is 7, and the range, which is a measure of spread, is $12 - 2 = 10$. The distribution will be roughly symmetric about the center. If students plot the results of the dice example using a histogram or dot plot, they will see how the shape of the distribution provides a picture of these features.

The distribution of sums can be used to make a probability statement about the next roll. But such a statement cannot be made about all kinds of data. In the dice example, every roll is an identical experiment. But, a distribution of the number of student absences in a typical school year does not tell us the probability of any particular student being absent, let's say, 3 times this year. The conditions and behaviors of one student are not identical to those of the next student. But, the distribution could tell us, for example, that about 4 out of 10 students will have 3 absences this year.

The Grade 6 Common Core State Standards for Statistics and Probability specify that students should—

- Develop understanding of statistical variability.
- Summarize and describe distributions.

The study of statistics and probability gives students ample opportunity to reason about real-world questions, to construct arguments, and to critique the arguments of others. The following hands-on activities will enable teachers to engage students in these practices.

Statistics and Probability

Contents

Lesson 1 Distributions 120

Objective: Describe a distribution of data values.

Manipulative: XY Coordinate Pegboard

Lesson 2 Mean, Median, Mode, and Range 124

Objective: Find and compare the mean, median, mode, and range of a data set.

Manipulative: Centimeter Cubes

Lesson 3 Graphical Representations: Histograms and Circle Graphs 128

Objective: Present data in both a histogram and a circle graph.

Manipulative: Centimeter Cubes; Deluxe Rainbow Fraction® Circles Set

Objective

Describe a distribution of data values.

Common Core State Standards

- **6.SP.2** Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
- **6.SP.4** Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
- **6.SP.5a** Summarize numerical data sets in relation to their context, such as by reporting the number of observations.

Statistics and Probability

Distributions

Students learn early that data, such as that collected from a survey, can be displayed and analyzed in a table, chart, or graph. In this lesson, students explore how to analyze data to answer questions about how the data are distributed. A distribution can be described by its center (mean, median, and mode), spread (range), and overall shape (skewed or symmetric).

Try It! Perform the Try It! activity on the next page.

Talk About It

Discuss the Try It! activity.

- Elicit that the data are concentrated in the middle and “tail off” to both sides. **Say:** *Some distributions are concentrated to one side and tail off to the other side. These are called skewed distributions.* Explain that being skewed left means being concentrated to the left and tailing off to the right. Being skewed right means the opposite. Point out that data plots can have other shapes, too.
- Have the class think of examples of data that might be distributed in different ways. For example, a survey on the number of siblings students have might produce data that’s skewed left because the majority of students might have 0, 1, or 2 siblings. The number of siblings will tail off to the right.

Solve It

Reread the problem with students. Have students write a description of the shape of the data. Have them determine the number of games (observations) that are represented and identify the median and the range. Have students write an explanation for why the number of shots on goal might be distributed the way it is.

More Ideas

For other ways to teach about describing distributions—

- Have students do the activity by drawing the x-axis on a piece of paper and using Centimeter Cubes to plot the counts.
- Have students create different shapes of distributions using Centimeter Cubes. For each distribution, have students transfer the data to a table. Ask students to explain what the data represent and why the distribution is shaped the way it is.

Formative Assessment

Have students try the following problem.

What shape would the distribution of scores on an easy test probably have?

- A. symmetric B. skewed right C. skewed left D. uniform

Try It! 30 minutes | Groups of 4

Here is a problem about describing distributions.

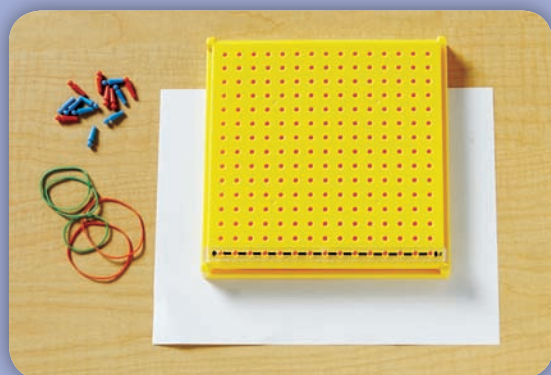
Coach Blackwell collected data on his soccer team for three seasons. The table shows his data for shots on goal. Make a display of the data. Describe the distribution of values. Is the distribution symmetric? Explain.

Shots on Goal	2	3	4	5	6	7	8	9	10	11	12
Number of Games	1	1	2	4	4	5	4	4	2	1	1

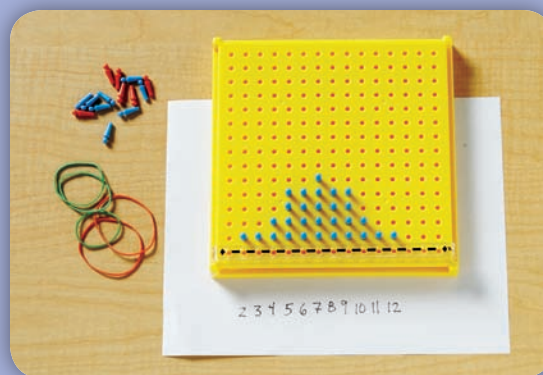
Introduce the problem. Then have students do the activity to solve the problem. Distribute pegboards to students.

Materials

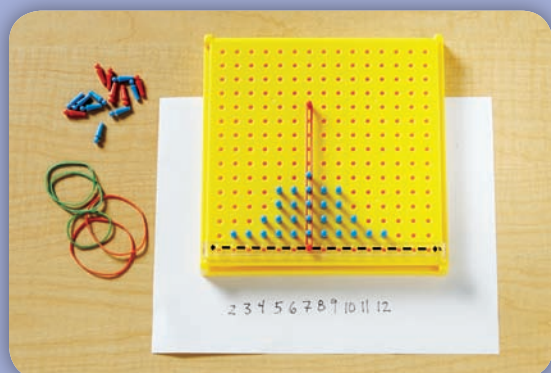
- XY Coordinate Pegboards



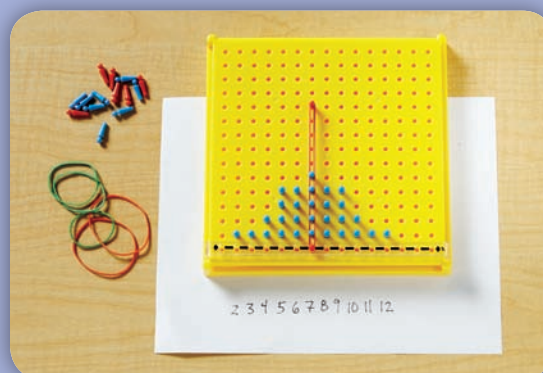
1. Say: Let's use the pegboard to display the data given in the problem. Remove the y-axis from the pegboard. Position the x-axis near the bottom. Help students remove the y-axis if necessary.



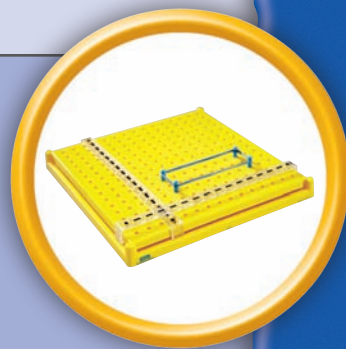
2. Say: Use the x-axis for the number of shots on goal. Plot the number of games for each value of shots on goal. Help students create a line plot using pegs for the counts. For example, have students insert 1 peg above $x = 2$ (to represent 1 game having 2 shots on goal), 1 peg above $x = 3$, 2 pegs above $x = 4$, and so forth, until every game has been accounted for.



3. Say: Let's talk about the shape of the data plot. **Ask:** Does it have symmetry? Elicit that it does have symmetry. Have students show the line of symmetry using a rubber band. Explain that when data are distributed equally about a center value (or nearly so), the distribution is said to be *symmetric*.



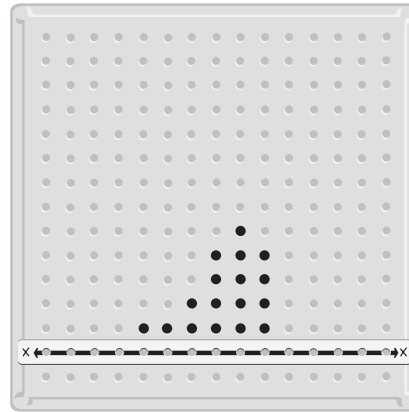
4. Ask: How many games are represented by the data? Elicit that 29 games are represented. **Say:** Find the median number of shots on goal. Find the range. Provide guidance if necessary. The median is 7 and the range is 10.



Use an XY Coordinate Pegboard to plot the data.
Use the plot to answer Questions 1–4.

(Check students' work.)

Score (x)	Number of Students (y)
5	1
6	1
7	2
8	4
9	5
10	4



- How many observations are there?
- What is the median score?
- What is the range of the scores?
- Describe the shape of the distribution.

17

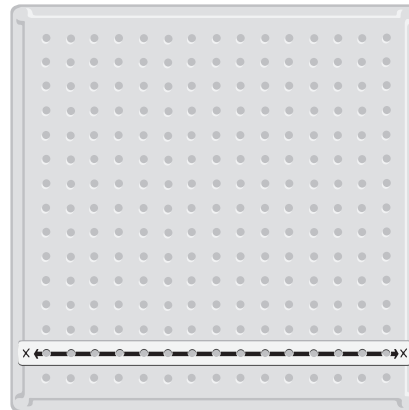
9

5

skewed right

Plot the data on the XY Coordinate Pegboard. Use the plot to answer Questions 5–8.

Score (x)	Number of Students (y)
4	1
5	1
6	2
7	4
8	3
9	1
10	1



- How many observations are there?
- What is the median score?
- What is the range of the scores?
- Describe the shape of the distribution.

13

7

6

symmetric

Answer Key

Challenge! Refer to the two distributions you used to answer Questions 1–8. Find the mean score in each case. Round your answers to the nearest hundredth. Show your work. Why is the mean closer to the median in the second distribution?

Challenge: 8.35; 7.08; It is more symmetric.

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. There are no margins, text, or other markings on the paper.

LESSON 2

Objective

Find and compare the mean, median, mode, and range of a data set.

Common Core State Standards

- **6.SP.3** Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.
- **6.SP.5c** Summarize numerical data sets in relation to their context, such as by giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.

Statistics and Probability

Mean, Median, Mode, and Range

Students might be familiar with the meaning of *average*. Here they learn that the average, or mean, is just one of the measures of center that can be ascribed to a set of data. The median and mode are also measures of center. The range is a measure of spread.

Try It! Perform the *Try It!* activity on the next page.

Talk About It

Discuss the Try It! activity.

- Talk about how a set of data can have more than one mode or no mode.
- **Ask:** *Why is it important to order the data first?* **Say:** *Sometimes there are two middle values, so you have to find the mean, or average, of the two middle values to find the median.* **Ask:** *How can you tell that a data set has two middle values?*
- **Ask:** *Which measures are most likely to change if you remove the least and greatest values from the data set? What are some situations in which you would want to know the mean, median, mode, or range?*

Solve It

Reread the problem with students. Ask them to describe each measure and how to find it. Then have them compare measures and explain their answers.

More Ideas

For other ways to teach mean, median, mode, and range—

- Provide data sets with six values such as 9, 6, 1, 10, 9, and 7. Demonstrate how to find the median when there are two middle values. Have students work in pairs using Color Tiles and/or Two-Color Counters to model the data and find the mean (7), median (8), mode (9), and range (9). Have students discuss how each measure represents the data.
- Give students a partial data set of values such as 4, 9, 7. Tell students the data set has five values total. The mean and median are 6, the mode is 4, and the range is 5. Have students use Color Tiles to find that the data set is 4, 4, 6, 7, and 9.

Formative Assessment

Have students try the following problem.

Ryan sent the following numbers of e-mails each day last week: 4, 12, 10, 6, 3, 12, and 9. Find the mean, median, mode, and range of the data set. Which measure is the greatest?

- A. mean B. median C. mode D. range

Try It! 20 minutes | Pairs

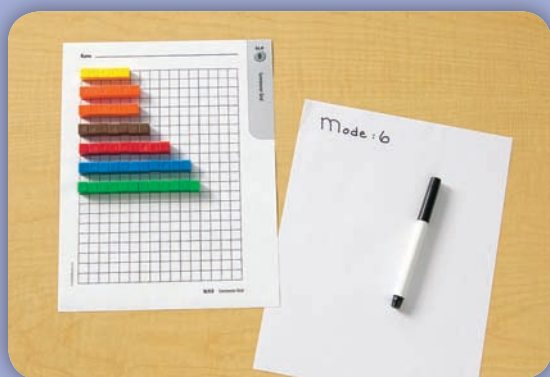
Here is a problem about the mean, median, mode, and range of a data set.

A pet store has the following numbers of fish in seven different aquariums: 6, 12, 7, 5, 11, 6, and 9. Compare the mean, median, mode, and range of the fish in the aquariums. The pet store owner doesn't want any aquarium to have too many more fish than another. Which data measure would he be interested in?

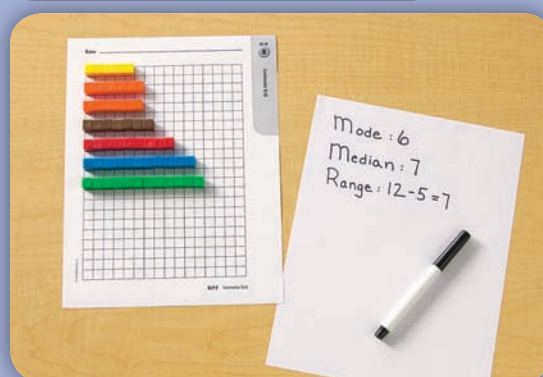
Introduce the problem. Then have students do the activity to solve the problem. Distribute Centimeter Cubes, grid paper, paper, and pencils. Define mean, median, mode, and range.

Materials

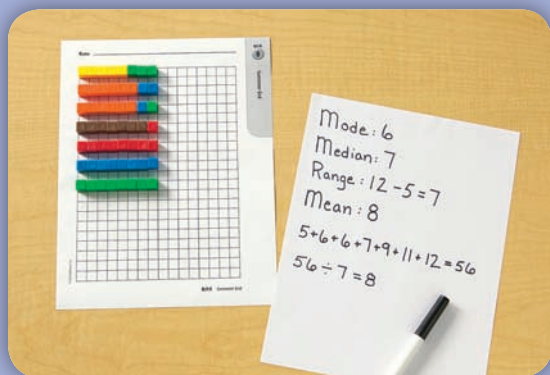
- Centimeter Cubes (60 per pair)
- Centimeter Grid Paper (BLM 10; 1 per pair)
- paper (1 sheet per pair)
- pencils (1 per pair)



1. Have students use Centimeter Cubes to show the number of fish in each aquarium and arrange the cubes on the grid in rows using one color of cube for each value. Then have them order the values from least to greatest.
Say: The mode is the value that occurs most often. **Ask:** What is the mode of this data set?



2. Say: The median is the middle value.
Ask: What is the median of this data set?
Say: The range is the difference between the greatest value and the least value. **Ask:** What is the range of this data set?



3. Have students arrange the cubes so that there is an equal number in each row. Explain how this number is the mean, or average.
Say: You can also compute the mean. Add all of the values in the data set and divide the sum by the total number of data values. Help students compute the mean.

Look Out!

Watch for students who confuse median and mean. Emphasize that when the data items are listed in order according to size, the median is the middle number. The mean is the average. Have students arrange two sets of cubes for the data set 3, 5, 5, 8, and 9, with one set of cubes arranged to determine the median and the other arranged to demonstrate the mean. Then have them compare and contrast the mean and median of the data set.



Use Centimeter Cubes to model the set of data shown. Write the numbers in the set. Determine the mode, median, and range. Rearrange your cubes to find the mean.

(Check students' work.)

1.



Data: 3, 1, 7, 5, 4

Mode: no mode

Median: 4

Range: 6

Mean: 4

Using Centimeter Cubes, model the set of data. Sketch the set of data. Determine the mode, median, and range. Rearrange the cubes to find the mean. Sketch the set of data.

2. 9, 8, 3, 2, 3

Mode: 3 Median: 3

Range: 7 Mean: 5

Find the measure of central tendency named for each set of data.

3. 4, 7, 3; range

4

4. 6, 4, 2, 8, 5; mean

5

5. 15, 9, 3; median

9

6. 3, 2, 7, 9, 6, 3; mode

3

7. 14, 3, 8, 4, 1; mean

6

8. 16, 12, 13, 5; median

12.5

Answer Key

Challenge! Which measure of central tendency can you find by simply looking at the set of data? Explain.

Challenge: (Sample) The mode is the number that occurs most often in a set. So, you can tell by looking at the set which number is the mode.

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

LESSON 3

Objective

Present data in both a histogram and a circle graph.

Common Core State Standards

- **6.SP.4** Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
- **6.SP.5a** Summarize numerical data sets in relation to their context, such as by reporting the number of observations.

Statistics and Probability

Graphical Representation: Histograms and Circle Graphs

Students have already worked with numerous bar graphs, histograms, and circle graphs in their previous study of mathematics, science, and social studies. It is important for them to understand how circle graphs are created. In this activity, the creation of a circle graph is directly related to a histogram that presents the same data.

Try It! Perform the Try It! activity on the next page.

Talk About It

Discuss the Try It! activity.

- **Ask:** What does each cube in the histogram represent?
- **Ask:** What can you say about the heights of the people at the picnic?
- **Ask:** Which graph do you find easier to read? Explain.

Solve It

Reread the problem with students. Have them create a histogram from the data given. When they have finished, ask them to create a circle graph using the same data.

More Ideas

For another way to teach about histograms and circle graphs—

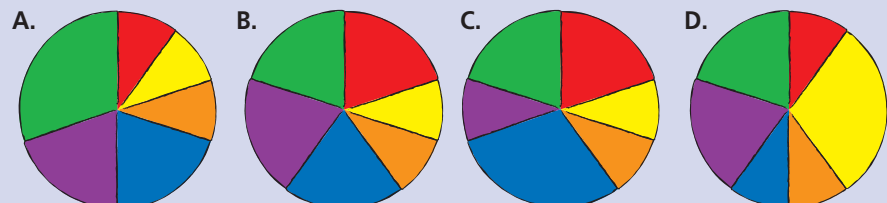
- Have students repeat the activity by collecting data regarding each student's shoe size. Ask students to create a histogram, using a range of two whole shoe sizes. Have them convert their histograms to circle graphs.

Formative Assessment

Have students try the following problem.

Favorite Color	Red	Green	Blue	Yellow	Orange	Purple
Number of Votes	10	30	20	10	10	20

Which graph represents the data above?



Try It! 30 minutes | Groups of 4

Here is a problem about representing data in histograms and circle graphs.

Merlene collected data about the heights of the people at a father-daughter picnic. Her results are shown below:

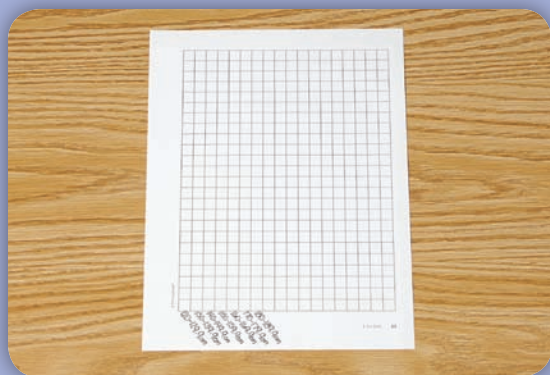
Heights (cm)	120–129.9	130–139.9	140–149.9	150–159.9	160–169.9	170–179.9	180–189.9
Number of Picnickers	6	12	9	12	6	18	9

Merlene would like to display her findings in two formats. Help Merlene create a histogram of her data. Then use the histogram to create a circle graph.

Introduce the problem. Then have students do the activity to solve the problem. Distribute Centimeter Cubes, Deluxe Rainbow Fraction Circles, and grid paper to students.

Materials

- Centimeter Cubes
- Deluxe Rainbow Fraction® Circles
- Centimeter Grid Paper (BLM 10; 1 per group)



1. Have students label columns on their grid paper with the ranges of heights shown in Merlene's data table.



2. Ask students to build a histogram representing this data using Centimeter Cubes.

Heights (cm)	120–129.9	130–139.9	140–149.9	150–159.9	160–169.9	170–179.9	180–189.9
Number of Picnickers	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{8}$

3. Have students calculate the fraction of people in each of the ranges. **Ask:** How many people were at the picnic? What fraction of them falls into each range?

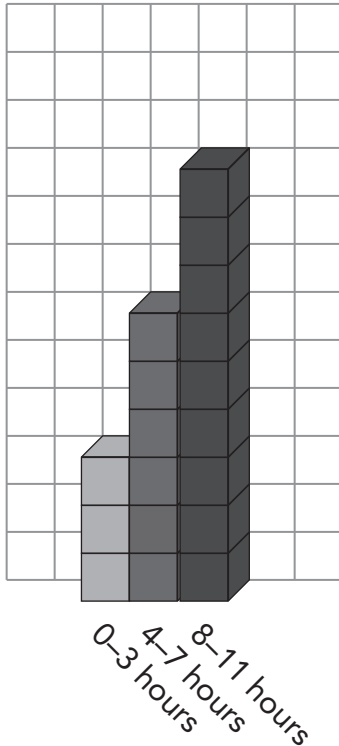


4. Have students use the Fraction Circles and the fractions they found for each range to create a circle graph.



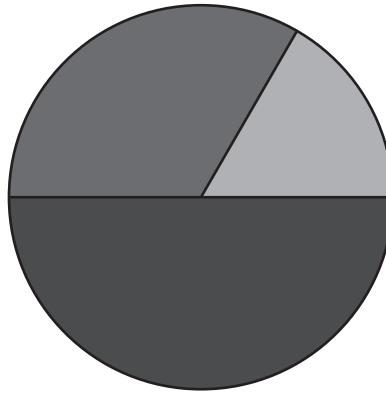
Use Centimeter Cubes to make a histogram. Use Fraction Circles to make a circle graph of the same data. Complete the table.

1.



(Check students' work.)

Number of Hours Slept Each Night by People Surveyed



How many people were surveyed?

18

Number of Hours	Fraction of People
0-3	$\frac{1}{6}$
4-7	$\frac{1}{3}$
8-11	$\frac{1}{2}$

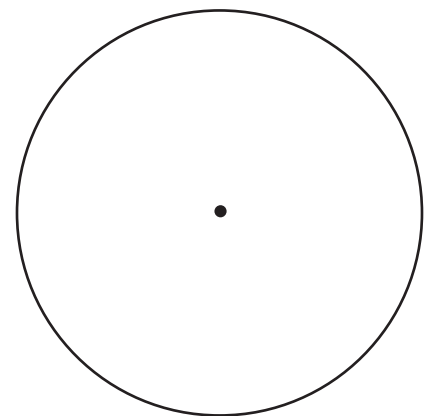
Using Centimeter Cubes and Fraction Circles, make a histogram and circle graph for the data in the table. Sketch the graphs. Complete the table.

2.

Number of Pets	Number of People	Fraction of People
0-1	10	$\frac{5}{12}$
2-3	9	$\frac{3}{8}$
4-5	3	$\frac{1}{8}$
6 or more	2	$\frac{1}{12}$



Check students' graphs.



Answer Key

Challenge! Explain how the tallest bar in a histogram is related to the sections of the corresponding circle graph.

Challenge: (Sample) The tallest bar represents the greater number, so the range with the tallest bar will also be in the largest section of the circle graph.

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.