6:7 Song Length Distribution

Teacher Notes

Central math concepts

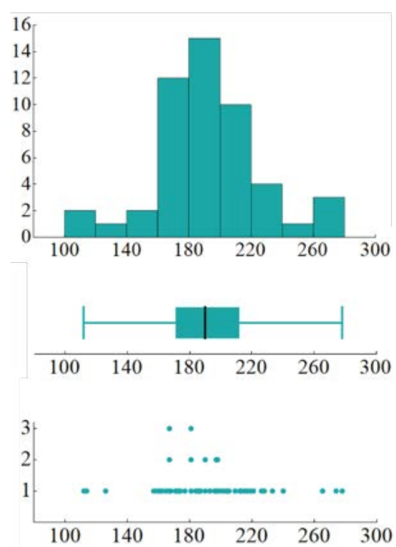
As explained in the relevant *Progression document*, statistical reasoning is a four-step investigative process:

- Formulate questions that can be answered with data
- Design and use a plan to collect relevant data
- Analyze the data with appropriate methods
- Interpret results and draw valid conclusions from the data that relate to the questions posed.†

In contrast to mathematical questions about mathematical objects, *statistical questions* anticipate variability in the data related to the question and account for it in the answers. CCSS 6.SP.A.1 gives the following example:

> *"How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.*

A set of data collected to answer a statistical question has a *distribution* that can be described by its center, spread, and overall shape. Ways of picturing distributions include creating histograms, box plots and dot plots. (See the figure for an example of each kind of display using the song length data.)

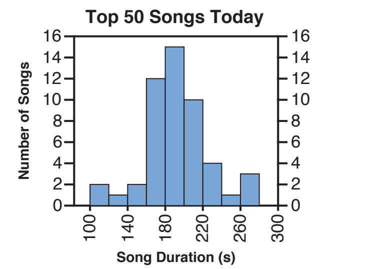


All three types of plot have a measurement scale along the horizontal axis. In a dot plot or histogram, the vertical scale is a count scale. In a box plot, the vertical scale is not numerically meaningful.

Students used dot plots in the elementary grades. In a dot plot, which is useful primarily for smaller data sets, every data point is shown. In a histogram, data values are grouped into bins and counted. The bins are not given and must be imposed on the data, usually by technology based on an algorithm. In a box plot, the data values are not shown; but by indicating the extreme values, the quartiles, and the median, the box plot visually indicates the width and clustering of the distribution of data values.

Box plots and histograms are sophisticated representations of data. In either case, reading a single fact from the display involves thoroughly comprehending the data, the context, and the conventions of the representation. For example, the tallest rectangle in the histogram represents the fact that 15 songs in the data set had durations of at least 180 seconds but less than 200 seconds.

6:7 **(1)** Look up the 50 top songs on a music streaming service. Type each song's duration into a spreadsheet. **(2)** Write a sentence about the data giving a measure of center and a measure of variability. **(3)** Make a histogram of the data.* **(4)** Write a sentence describing the overall pattern of the distribution and any striking deviations from the overall pattern. **(5)** Imagine that one year from now, you go back online and repeat (1)–(4). In what ways would you expect the data distribution to look similar? What differences would you expect to see?



*Use this histogram for (4) and (5) if you don't do (3).

Answer

(1) Data sets may vary; see example. **(2)** Answers may vary. Examples: "The mean song duration was 192 seconds, with a mean absolute deviation of 25 seconds"; "The median song duration was 190 seconds, with an interquartile range of 40 seconds." **(3)** Answers may vary; for an example, see the histogram provided in the task. **(4)** Answers may vary but should include the observation that song durations cluster near the center. **(5)** Answers may vary but could include hypotheses about whether the mean/median song duration would be similar or different to today's, and whether the variability in top song durations would be similar or different to today's. An opinion could be expressed about whether extreme song length (very short or very long) is a disadvantage for popularity. Answers should not suggest that the data will look precisely the same one year from now.

Click here for a student-facing version of the task.

A measure of center for a data set summarizes all of its values with a single number, and a measure of variability describes how its values vary with a single number (CCSS 6.SP.A.3). Inherent in the notion of a *summary* is loss of information; for example, if we say that "The mean song duration was 192 seconds, with a mean absolute deviation of 25 seconds," then we have left out an enormous number of details about the data set, but we have also said something useful about the data set. Measures of center and measures of variability thus allow us to refer to complicated realities in simplified terms. In that sense, these measures are important mathematical models.

In practical terms, measures of center and measures of variability are most useful when the data set in question has a large number of values. If there are only a few data values, then no summary is needed. Authentic statistics education should therefore include some work with relatively large data sets, including using technology to record, organize, analyze, and display those data sets.

Relevant prior knowledge

The following mathematics knowledge may be activated, extended, and deepened while students work on the task: thinking about data in context; calculating measures of center and measures of variation; describing patterns in distributions of univariate measurement data; and using technology.

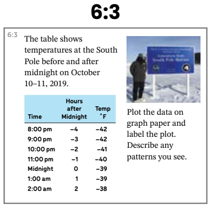
Extending the task

How might students drive the conversation further?

- Students could discuss the relative effectiveness of a dot plot, box plot, and histogram for displaying the distribution in ways that promote understanding of the patterns.
- Students could compare the distributions from two different genres to see if the distributions are appreciably different (CCSS 7.SP.B.3).

Related Math Milestones tasks

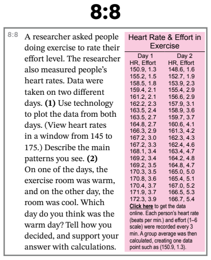
6:3


6:3 The table shows temperatures at the South Pole before and after midnight on October 10–11, 2019.

Time	Hours after Midnight	Temp °F
8:00 pm	−4	−42
9:00 pm	−3	−42
10:00 pm	−2	−41
11:00 pm	−1	−40
Midnight	0	−39
1:00 am	1	−39
2:00 am	2	−38

Plot the data on graph paper and label the plot. Describe any patterns you see.

8:8

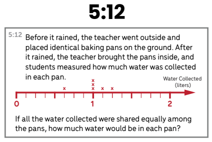

8:8 A researcher asked people doing exercise to rate their effort level. The researcher also measured people's heart rates. Data were taken on two different days. **(1)** Use technology to plot the data from both days. (View heart rates in a window from 145 to 175.) Describe the main patterns you see. **(2)** On one of the days, the exercise room was warm, and on the other day, the room was cool. Which day do you think was the warm day? Tell how you decided, and support your answer with calculations.

Heart Rate & Effort in Exercise

Day 1 HR, Effort	Day 2 HR, Effort
150.9, 1.3	148.6, 1.6
155.2, 1.5	152.7, 1.9
158.5, 1.8	153.9, 2.3
159.4, 2.1	155.4, 2.9
161.2, 2.1	156.6, 2.9
162.2, 2.3	157.9, 3.1
163.5, 2.4	158.9, 3.6
163.5, 2.7	159.7, 3.7
164.8, 2.7	160.6, 4.1
166.3, 2.9	161.3, 4.2
167.2, 3.0	162.3, 4.3
167.2, 3.3	162.4, 4.6
168.1, 3.4	163.4, 4.7
169.2, 3.4	164.2, 4.8
169.2, 3.5	164.8, 4.7
170.3, 3.5	165.0, 5.0
170.8, 3.6	165.4, 5.1
170.4, 3.7	167.0, 5.2
171.9, 3.7	166.5, 5.3
172.3, 3.9	166.7, 5.4

Click here to get the data online. Each person's heart rate (beats per min.) and effort (1–6 scale) were recorded every 3 min. A group average was then calculated, creating one data point such as (150.9, 1.3).

5:12


5:12 Before it rained, the teacher went outside and placed identical baking pans on the ground. After it rained, the teacher brought the pans inside, and students measured how much water was collected in each pan.

If all the water collected were shared equally among the pans, how much water would be in each pan?

Task **6:3 South Pole Temperatures** involves a distribution of bivariate measurement data.

In later grades, task **8:8 Heart Rate and Effort in Exercise** involves two distributions of bivariate data.

In earlier grades, task **5:12 Rain Measurements** involves a set of univariate measurement data with a calculation that prefigures the mean as a measure of center.

Refer to the Standards

6.SP; MP.3, MP.4, MP.5. Standards codes refer to www.corestandards.org. One purpose of the codes is that they may allow a task to shed light on the Standards cited for that task. Conversely, reading the cited Standards may suggest opportunities to extend a task or draw out its implications. Finally, Standards codes may also assist with locating relevant sections in curriculum materials, including materials aligned to comparable standards.

Aspect(s) of rigor:

Concepts, Procedural skill and fluency, Application

Additional notes on the design of the task

- Task 6:7 uses a selection criterion of the nationwide top 50 songs, but an alternative could be for students to choose a more specific music genre/chart they know about or enjoy.
- The data source for the histogram in task 6:7 was a music streaming service (accessed in 2019).

Curriculum connection

1. In which unit of your curriculum would you expect to find tasks like 6:7? Locate 2–3 similar tasks in that unit. How are the tasks you found similar to each other, and to 6:7? In what specific ways do they differ from 6:7?
2. Thinking about the curriculum unit you identified, at what point in the unit might a task like 6:7 help students converge toward grade-level thinking about the important mathematics in the task? What factors would you consider in choosing when to use such a task in the unit?*

† Common Core State Standards Writing Team (2011), *Progressions for the Common Core State Standards in Mathematics (draft) 6–8 Statistics and Probability.*

* Math Milestones™ tasks are not designed for summative assessment. Used formatively, the tasks can reveal and promote student thinking.

6:7 Song Length Distribution

Teacher Notes

Anticipating and responding to student thinking about the task

Imagine how students might think about the task, and what you might see and hear while they work.

On this page, you can write your thoughts on the following questions.

Solution Paths

- What solution paths might you expect to see?
- What representations might you see? What correspondences between those representations might be noticed by students (or be worth pointing out to students) and discussed by them?
- What misconceptions or partial understandings might be revealed as students work on the task? How could you respond to these positively and productively?

Language

- What might you expect to hear from students engaged with the task? What does that language reveal about their mathematical thinking, and how might you respond to different ways of thinking?
- If students are using early English or using multiple languages in an integrated communication system, how might you help their classmates see those mathematical ideas as valuable?
- Even when using nascent language, students are thinking and communicating their thinking. What might it look like to respond positively and productively to the mathematics in their thinking before giving feedback on the language used?

Identity, Agency, and Belonging

- How can you engage students' interests, experiences, or funds of knowledge?
- How can you build students' self-confidence as learners, thinkers, and doers of mathematics?
- What choices are there for a student to make in the task? How can you build students' agency to the point where they notice and make these choices to solve problems?
- How might one use feedback to build student agency? Where might there be opportunities to build students' self-confidence?