Describing Distributions

Alignments to Content Standards:  6.SP.A.2  6.SP.B.4

Task

Data Set 1 consists of data on the time to complete an assignment (in minutes) for 25 sixth graders. Data Set 2 consists of data on the time to complete an assignment for 25 seventh graders. Dot plots of the two data sets are shown below.

1. Describe the data distribution of times for seventh graders (Data Set 2). Be sure to comment on center, spread and overall shape.

2. Are Data Set 1 and Data Set 2 centered in about the same place? If not, which one has the greater center?

3. Which of Data Set 1 and Data Set 2 has greater spread?

4. Were sixth graders (Data Set 1) or seventh graders (Data Set 2) more consistent in
their times to complete the task?

Data Set 3 consists of data on the number of text messages sent in one month for 100 teenage girls who have a cell phone. Data Set 4 consists of data on the number of text messages sent in one month for 100 teenage boys who have a cell phone. Histograms of the two data sets are shown below.

5. Describe the data distribution of number of text messages for the girls (Data Set 3). Be sure to comment on center, spread and overall shape.
6. Are Data Set 3 and Data Set 4 centered in about the same place? If not, which one has the greater center?

7. Which of Data Set 3 and Data Set 4 has greater spread?

8. On average, did the girls (Data Set 3) or the boys (Data Set 4) send more text messages?

Data Set 5 consists of data on the scores on a video game for 100 teenage girls. Data Set 6 consists of the scores on a video game for 100 teenage boys. Box plots of the two data sets are shown below.

9. Describe the data distribution of Data Set 5. Be sure to comment on center, spread and overall shape.

10. Are Data Set 5 and Data Set 6 centered in about the same place? If not, which one has the greater center?

11. Which of Data Set 5 and Data Set 6 has greater spread?

12. On average, did the girls (Data Set 5) or the boys (Data Set 6) tend to have higher scores?
IM Commentary

This task is designed to assess student thinking related to content standard 6.SP.A.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 shape) and 6.SP.B.4 (Display numerical data in plots on a number line, including dot plots, histograms, and box plots). Although students are not asked to construct graphical displays in this task, they must be familiar with the different graphical displays mentioned in standard 6.SP.B.4 and understand how each of these displays summarizes a data distribution.

In this task, students are asked to describe data distributions in terms of center, spread and overall shape and to also compare data distributions in terms of center and spread by selecting which of two distributions has a greater center and which has a greater spread. Students who can do this show an understanding of how center and variability in a data distribution is reflected in each of the three types of graphical displays (dot plots, histograms, and box plots).

A productive way to use this task is as the basis for work in small groups. This encourages students to “talk statistics”. Once the groups have had a chance to work together on the questions posed in this task, they could share their answers as part of a whole class discussion.

Solution

1. The distribution of times is approximately symmetric. A typical time for seventh graders is around 54 or 55 minutes. The times for seventh graders range from 50 to 60 minutes, and are fairly evenly spread out over that interval.

2. No, Data Set 1 and Data Set 2 are not centered in about the same place. Data Set 1 is centered at around 50 minutes, while Data Set 2 is centered at about 54 or 55 minutes.
3. Data Set 1 is more spread out.

4. The times for seventh graders were closer together than the times for sixth graders, so seventh graders were more consistent in their times.

5. The distribution of number of text messages sent for the girls is approximately symmetric. A typical number of text messages sent for girls is about 100. The values in the data set range from about 75 to 125 text messages, with more near 100 than farther away from 100.

6. No, Data Set 3 and Data Set 4 are not centered in about the same place. Data Set 3 is centered at around 100 text messages, while Data Set 4 is centered at about 80 text messages.

7. Data Set 4 is more spread out with the number of text messages sent ranging from around 30 to around 150. The range for Data Set 3 is much smaller and the values in the data set tend to be closer together.

8. On average, the girls sent more text messages.

9. The distribution of video game scores for the girls is approximately symmetric. A typical score for the girls is around 105. The values in the data set range from about 0 to 225.

10. No, Data Set 5 and Data Set 6 are not centered in about the same place. Data Set 5 is centered at around 105, while Data Set 6 is centered at about 75.

11. Data Set 5 is more spread out. Both the interquartile range and the range are smaller for the girls’ scores.

12. On average, the girls tended to have higher scores. The boxplot for girls is centered at around 105, while the boxplot for the boys is centered at about 75.