Splash! into Statistical Variability – Middle School (Grades 6-8)

Introduction	In these activities, students will explore the data that they gathered with Splash! and apply it to various mathematical tasks.
Time	1 class period
Grade	6-8
Lesson Preparation	Students will have visited the Tsongas Industrial History Center to participate in the Power to Production program. Students gathered data from the waterwheel test on the Splash! app. For this activity, teachers can download the class's data at www.tihcsplash.org. For each student or team of students, you will also need a copy of the Splash into Statistical Variability worksheet.
Vocabulary	Mean Center Median Spread Mode Variability Range
Anticipated Student Preconceptions/ Misconceptions	Students will have studied mean, median, mode and range. They may not have thought about the relationship these mathematical terms have to the commonly used term "average."
Frameworks	Massachusetts Math Standards 6.SP Develop understanding of statistical variability. 1. Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers. *For 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.* 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. 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 Summarize and describe distributions. 4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots. This activity also addresses the following Mathematical Practice Standards MPS.3 Construct viable arguments and critique the reasoning of others. MPS.4 Model with mathematics
Guiding Question	Why is variation in data important in statistics? How can we display and describe data in ways that makes sense to other people?

Objectives	Students will be able to: • Compute measures of central tendency for a given set of data. • Compute range for a given set of data. • Describe what might cause variation within a set of data and between sets of data. • Create a box plot to show data.

Activity	1. Group the students in pairs or groups of three.
	2. Be sure to display the data that students gather as part of their field trip. Inform students that the data sets come from the first 20 tests of wheel B. Distribute the worksheets to the students.
	3. Ask students to scan data set 1 and then ask what they notice about the data. Students may respond with comments about range, mode, etc.
	4. Tell students that they will calculate the mean, median, mode, and range for each data set and that they will create a box plot for each data set.
	5. When students have completed this, they can answer the following questions in their small groups, and then the teacher can lead a whole-class discussion of the questions.
	    • How would you describe the center of each set of data?
	    • Which measure, mean, median or mode, do you think best describes the "average" of this data?
	    • How would you describe the variability of each set of data?
	    • What explanations might there be for the variability of the data?
	    • How do the mean, median, mode and range of data set 1 compare to the mean, median, mode and range of data set 2? What might explain differences or similarities between the two?

Assessment	Student work produced during this lesson can serve as assessment. Students will have found mean, median, mode and range for two data sets and will have created boxplots for the two data sets.
Differentiated Suggestions	Review vocabulary and procedure for determining mean, median, mode and range. Review procedure for creating a box plot, or have anchor chart showing how a box plot is created. If students are working quickly, ask them to create box plots for the strength data (sets 3 and 4). In addition to answering the activity questions above, students can consider how the scale of a box plot changes the way in which it is interpreted. Ask students to consider how the data they gathered on the field trip would change the statistics they have calculated. Was all of their data within the range? Does including their data change the mean, median or mode?
Adapting the Activity for Other Grades	In upper middle school grades, teachers may want to present questions such as the following: 1. One measurement is added to this data set and the mean increases to 10.5. What would this measurement be? 2. Sofia knows the following about her set of data – it has 5 measurements, the mean is 3.5, the median is 3. What numbers could make up this set of data?

Splash! into Statistical Variability Worksheet

Data Set 1: Bucket Wheel-B Base-Speed

9.5	8	11	9	9	10	8	9	6	9
8	9	10	9	10	7	10	5	9	5

Mean: Median:

Mode: Range:

Draw a box plot to represent the data in the space below

Data Set 2: Paddle Wheel-B Base-Speed

8	7	10	7	7	11	8	7	15	8
5	8	5	5	5	5	8	9	7	7

Mean: Median:

Mode: Range:

Draw a box plot to represent the data in the space below

1. How would you describe the center of each set of data?
2. Which measure, mean, median or mode, do you think best describes the "average" of this data?
3. How would you describe the variability of each set of data?
4. What explanations might there be for the variability of the data?
5. How does the mean, median, mode and range of data set 1 compare to the mean, median, mode and range of data set 2? What might explain differences or similarities between the two?

Data Set 3: Paddle Wheel-B Base-Strength

5	5	2	3	5	3	3	3	6	6
6	7	5	5	5	5	4	7	5	2

Mean: Median:

Mode: Range:

Draw a box plot to represent the data in the space below

Data Set 4: Bucket Wheel-B Base-Strength

7	5	4	3	6	3	5	4	2	4
4	4	4	3	3	1	5	2	4	2

Mean: Median:

Mode: Range:

Draw a box plot to represent the data in the space below

1. How would you describe the center of each set of data?
2. Which measure, mean, median or mode, do you think best describes the "average" of this data?
3. How would you describe the variability of each set of data?
4. What explanations might there be for the variability of the data?
5. How does the mean, median, mode and range of data set 3 compare to the mean, median, mode and range of data set 4? What might explain differences or similarities between the two?