

Parent Packet

HAUPPAUGE MATH

DEPARTMENT

CCLS

Grade 6

MODULE 6

<http://www.hauppauge.k12.ny.us/math>

Grade 6 Module 6

Statistics

In this module, students move from simply representing data into analysis of data. Students begin to think and reason statistically, first by recognizing a statistical question as one that can be answered by collecting data. Students learn that the data collected to answer a statistical question has a distribution that is often summarized in terms of center, variability, and shape. Throughout the module, students see and represent data distributions using dot plots and histograms. They study quantitative ways to summarize numerical data sets in relation to their context and to the shape of the distribution. As the module ends, students synthesize what they have learned as they connect the graphical, verbal, and numerical summaries to each other within situational contexts, culminating with a major project.

Grade 6 • Module 6

Statistics

OVERVIEW

In Grade 5, students used bar graphs and line plots to represent data and then solved problems using the information presented in the plots (5.MD.B.2). In this module, students move from simply representing data into analysis of data. In Topic A, students begin to think and reason statistically, first by recognizing a statistical question as one that can be answered by collecting data (6.SP.A.1). Students learn that the data collected to answer a statistical question has a distribution that is often summarized in terms of center, variability, and shape (6.SP.A.2). Beginning in Topic A, and throughout the module, students see and represent data distributions using dot plots and histograms (6.SP.B.4).

In Topics B and C, students study quantitative ways to summarize numerical data sets in relation to their context and to the shape of the distribution. The mean and mean absolute deviation (MAD) are used for data distributions that are approximately symmetric, and the median and interquartile range (IQR) are used for distributions that are skewed. Students apply their experience in writing, reading, and evaluating expressions in which letters stand for numbers (6.EE.A.2) as they learn to compute and interpret two pairs of statistical measures for center and spread (6.SP.A.5).

In Topic B, students study mean as a measure of center and mean absolute deviation as a measure of variability. Students learn that these measures are preferred when the shape of the distribution is roughly symmetric. Then, in Topic C, students study median as a measure of center and interquartile range as a measure of variability. Students learn that these measures are preferred when the shape of the distribution is skewed. Students develop in Topic B, and reinforce in Topic C, the idea that a measure of center provides a summary of all its values in a single number, while a measure of variation describes how values vary, also with a single number (6.SP.A.3). Measures of center and variability for distributions that are approximately symmetric (mean and MAD) are covered before measures (median and IQR) for skewed data distributions. This choice was made because it is easier for students to understand measuring center and variability in the context of symmetric distributions.

For students, box plots are the most difficult of the graphical displays covered in this module. This is because they differ from dot plots and histograms in that they are not really a display of the data but rather a graph of five summary measures (minimum, lower quartile, median, upper quartile, and maximum). This graph conveys information on center and variability but is more difficult for students to interpret because, unlike histograms, where large area corresponds to many observations, in a box plot, large area indicates spread and small area indicates a large number of observations in a small interval. Box plots also require the calculation of quartiles and are best covered after quartiles have been

introduced and used to calculate the IQR. For these reasons, box plots are introduced late in the module after the IQR and after students have already developed some fundamental understanding of data distributions, which is easier to do in the context of dot plots and histograms. In Topic D, students synthesize what they have learned as they connect the graphical, verbal, and numerical summaries to each other within situational contexts, culminating with a major project (6.SP.B.4, 6.SP.B.5).

Students implement the four-step investigative process with their projects by stating their statistical questions, explaining the plan they used to collect data, analyzing data numerically and with graphs, and interpreting their results as related to their questions. The Mid-Module Assessment follows Topic B. The End-of-Module Assessment follows Topic D.

Topic A

Understanding Distributions

In Topic A, students begin a study of statistics by learning to recognize a statistical question. They develop an understanding of what data could be collected to answer a statistical question and anticipate variability in the data collected to answer the question. In Lesson 1, statistical questions are introduced in the context of a four-step process for posing and answering questions based on data. As students begin to explore data, they see the need to organize and summarize data. In Lesson 2, students are introduced to the idea that a data distribution can be represented graphically and that there are several different types of graphs, including dot plots and histograms, commonly used to represent a distribution of numerical data. This lesson then builds on students' previous work with line plots, introducing them to dot plots. In Lesson 3, students construct dot plots and begin to describe data distributions. In Lesson 4, students are introduced to histograms as another way of representing a data distribution graphically and the advantages and disadvantages of histograms relative to dot plots are discussed. Additionally, students begin to consider the shape of a data distribution (symmetric versus skewed) in this lesson and are introduced to the idea that different numerical summary measures of center and variability are used to describe data distributions that are approximately symmetric than the measures used to describe data distributions that are skewed. This is an important distinction and is the basis for the content introduced in Topics B and C. Lesson 5 gives students additional practice in constructing and describing histograms and introduces relative frequency histograms.

Topic B

Summarizing a Distribution that is Approximately Symmetric Using the Mean and Mean Absolute Deviation

In Topic B, students begin to summarize data distributions numerically. In Topic A, students have represented data distributions graphically and have described distributions informally in terms of shape, center, and variability. In this topic, students are introduced to a measure of center (the mean) and a measure of variability (the mean absolute deviation (MAD)) that are appropriate for describing data distributions that are approximately symmetric. In Lesson 6, students learn to calculate the mean and to understand the “fair share” interpretation of the mean. In Lesson 7, students develop an understanding of the mean as a balance point of a data distribution—the point where the sum of distances of points to the right of the mean and the sum

of distances of points to the left of the mean are equal. This understanding provides a foundation for considering distances from the mean, which are used in calculating the MAD, a measure of variability around the mean. Lessons 8 and 9 introduce the MAD as a measure of variability, and students calculate and interpret the value of the MAD. Lessons 10 and 11 give students the opportunity to use both graphical and numerical summaries to describe data distributions, to compare distributions, and to answer questions in context using information provided by a data distribution.

Topic C

Summarizing a Distribution that is Skewed Using the Median and the Interquartile Range

In Topic C, students are introduced to a measure of center (the median) and a measure of variability (the interquartile range (IQR)) that are appropriate for describing data distributions that are skewed. Box plots are also introduced in this topic. In Lesson 12, students learn to calculate and interpret the median. Quartiles are introduced in Lesson 13, and the quartiles are then used to calculate the IQR. Students also learn to interpret the IQR as a measure of variability in a data distribution. Lessons 14–16 introduce box plots. Boxplots are often difficult for students to interpret, as they are not a graph of a data distribution (as are dot plots and histograms), but rather are a graph of five key summary statistics of a data set (the minimum, lower quartile, median, upper quartile, and the maximum). Lesson 16 has students use box plots to compare groups, setting the stage for future work on comparing groups in Grade 7.

Topic D

Summarizing and Describing Distributions

In Topic D, students integrate what they have learned about graphical and numerical data summaries in the previous topics. They match dot plots and histograms to numerical measures of center and variability. Students estimate means and medians from graphical representations of data distributions. They also estimate mean absolute deviation (MAD) and interquartile range (IQR) from graphical representations based on an understanding of data distributions in terms of shape, center, and variability.

Terminology

New or Recently Introduced Terms

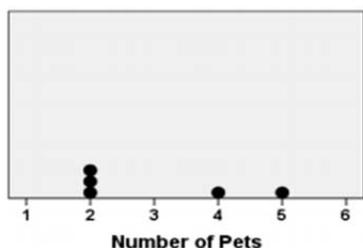
- **Statistical Question** (A question that anticipates variability in the data that would be collected in order to answer the question.)
- **Median** (A measure of center appropriate for skewed data distributions. It is the middle value when the data are ordered from smallest to largest if there are an odd number of observations and half way between the middle two observations if the number of observations is even.)
- **Mean** (A measure of center appropriate for data distributions that are approximately symmetric. It is the average of the values in the data set. Two common interpretations of the mean are as a “fair share” and as the balance point of the data distribution.)
- **Dot Plot** (A plot of numerical data along a number line.)
- **Histogram** (A graphical representation of a numerical data set that has been grouped into intervals. Each interval is represented by a bar drawn above that interval that has a height corresponding to the number of observations in that interval.)
- **Box Plot** (A graph of five numerical summary measures: the minimum, lower quartile, median, upper quartile, and the maximum. It conveys information about center and variability in a data set.)
- **Variability** (Variability in a data set occurs when the observations in the data set are not all the same.)
- **Deviations from the Mean** (The differences calculated by subtracting the mean from the observations in a data set.)
- **Mean Absolute Deviation (MAD)** (A measure of variability appropriate for data distributions that are approximately symmetric. It is the average of the absolute value of the deviations from the mean.)
- **Interquartile Range (IQR)** (A measure of variability appropriate for data distributions that are skewed. It is the difference between the upper quartile and the lower quartile of a data set and describes how spread out the middle 50% of the data are.)

Familiar Terms and Symbols²

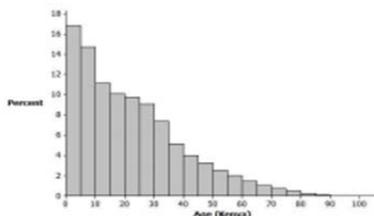
- Line Plot or Dot Plot

Suggested Tools and Representations

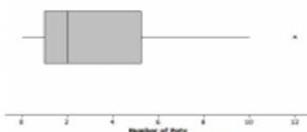
- Dot Plots
- Histograms
- Box Plots



Dot Plot



Histogram



Box Plot

Lesson 13

Objective: Describing Variability Using the Interquartile Range (IQR)

Example 1: Finding the IQR

Read through the following steps. If something does not make sense to you, make a note and raise it during class discussion. Consider the data: 1, 1, 3, 4, 6, 6, 7, 8, 10, 11, 11, 12, 15, 15, 17, 17, 17

Creating an IQR:

- I. Order the data: The data is already ordered.

1, 1, 3, 4, 6, 6, 7, 8, 10, 11, 11, 12, 15, 15, 17, 17, 17

- II. Find the minimum and maximum: The minimum data point is 1, and the maximum is 17.

(1), 1, 3, 4, 6, 6, 7, 8, 10, 11, 11, 12, 15, 15, 17, 17, (17)

- III. Find the median: There are 17 data points so the 9th one from the smallest or from the largest will be the median.

1, 1, 3, 4, 6, 6, 7, 8, (10), 11, 11, 12, 15, 15, 17, 17, 17

↑
median

- IV. Find the lower quartile and upper quartile: The lower quartile (Q1) will be half way between (the mean) the 4th and 5th data points (4 and 6), or 5 and the upper quartile (Q3) will be half way between the 13th and the 14th data points (15 and 15), or 15.

1, 1, 3, (4, 6), 6, 7, 8, 10, 11, 11, 12, (15, 15), 17, 17, 17

Q1 is 5 Q3 is 15

- V. Find the difference between Q3 and Q1: The $IQR = 15 - 5 = 10$.

Lesson 14

Objective: Summarizing a Distribution Using a Box Plot

Example 2: Making a Box Plot

Mr. S suggested that to be sure everyone had the same picture, statisticians developed a standard procedure for making the cut marks for the sections.

Mr. S. wrote the following on the board:

To make a box plot

- Find the median of all of the data
- Find Q1, the median of the bottom half of the data, and Q3, the median of the top half of the data.
- Draw a box that goes from Q1 to Q3, the two middle sections.
- Draw a line segment connecting the minimum value to the box and one that connects the maximum value to the box.

Now use the given number line to make a box plot of the data below.

20, 21, 25, 31, 35, 38, 40, 42, 44



The 5-number summary
is as follows:

Min = 20
Q1 = 23
Median = 35
Q3 = 41
Max = 44

Technology Resources

www.k-5mathteachingresources.com -This site provides an extensive collection of free resources, math games, and hands-on math activities aligned with the Common Core State Standards for Mathematics.

www.parccgames.com – fun games to help kids master the common core standards.

<http://www.mathplayground.com> –common core educational math games and videos.

www.learnzillion.com – math video tutorials.

www.ixl.com – practice common core interactive math skills practice.

www.mathnook.com –common core interactive math skill practice/ games, worksheets and tutorials.

www.adaptedmind.com – common core interactive practice, video lessons and worksheets

www.brainpop.com – animated tutorials of curriculum content that engages students. Can use a limited free version or buy a subscription.