

Dear Family,

The next unit in your child's mathematics class this year is ***Data Distributions: Describing Variability and Comparing Groups***. Students will learn to choose among a variety of representations to display distributions and will analyze, describe, and compare sets of data.

UNIT GOALS

Exploring statistics as a process of data investigation involves a set of four interrelated components (Graham, 1987).

- Posing the question: formulating the key question(s) to explore and deciding what data to collect to address the question(s);
- Collecting the data: deciding how to collect the data as well as actually collecting it;
- Analyzing the data: organizing, representing, summarizing, and describing the data and looking for patterns in the data; and
- Interpreting the results: predicting, comparing, and identifying relationships and using the results from the analyses to make decisions about the original question(s).

This dynamic process often involves moving back and forth among the four components.

HELPING WITH HOMEWORK

You can help with homework and encourage sound mathematical habits as your child studies this unit by asking questions such as:

- Is there anything that surprises you about the data and their distribution?
- Where do the data cluster in the distribution?
- How can I use the mean or median and range to help me understand and describe a data distribution?
- What strategies can I use to compare two different data sets?

In your child's notebook, you can find worked-out examples from problems done in class, notes on the mathematics of the unit, and descriptions of the vocabulary words.

HAVING CONVERSATIONS ABOUT THE MATHEMATICS IN DATA DISTRIBUTIONS

You can help your child with his or her work for this unit in several ways:

- Look with your child for uses of data in magazines, newspapers, and on TV.
- Point out examples of graphical displays and ask your child questions about the information shown.
- Ask your child about the data studied in class. What were the typical (mode, median, or mean) values for these data?
- Look over your child's homework and make sure all questions are answered and that explanations are clear.

A few important mathematical ideas that your child will learn in *Data Distributions* are given on the back. As always, if you have any questions or concerns about this unit or your child's progress in class, please feel free to call.

Sincerely,

Important Concepts and Examples

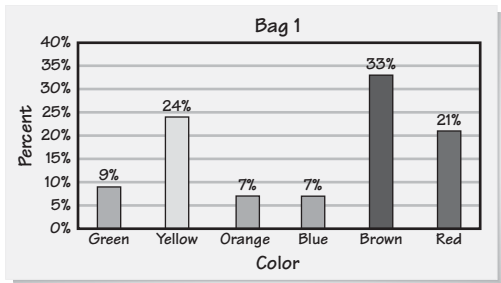
Representing Data Distributions

Statisticians use representations or summary statistics during the analysis part of the process of statistical investigation to describe the data distribution.

READING STANDARD DATA REPRESENTATIONS

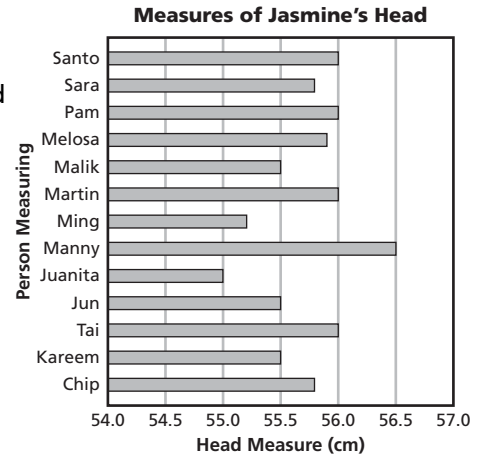
- Reading the data involves “lifting” information from a graph to answer explicit questions.
- Reading between the data includes the interpretation and integration of information presented in a graph.
- Reading beyond the data involves extending, predicting, or inferring from data to answer implicit questions.

FREQUENCY BAR GRAPH A bar’s height is the number (frequency) of cases that all have that value.



VALUE BAR GRAPH

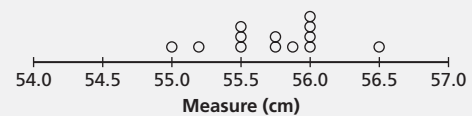
Each case is represented by a separate bar whose relative length corresponds to the magnitude or value of that case.



DOT PLOT (OR LINE PLOT)

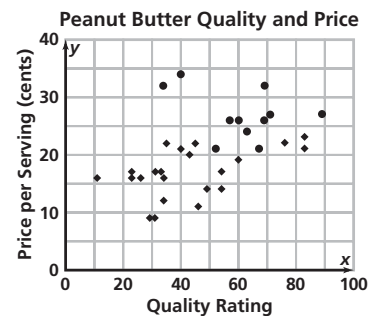
Each case is represented as a dot (or an “x”) positioned over a labeled number line.

Line Plot with Dots: Measures of Jasmine's Head



SCATTERPLOT

The relationship between two different attributes is explored by plotting values of two numeric attributes on a Cartesian coordinate system.



Measures of Central Tendency or Location (Mode, Median, Mean)

MODE The data value or category that occurs with greatest frequency. It is not usually used for summarizing numerical data.

Number of siblings: 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 3, 5, 6.

The mode is 2.

MEDIAN The numerical value that marks the middle of an ordered distribution. It is not influenced by extreme data values. Graphically, the median marks the location that divides a distribution into two equal parts.

The median for the data set 3, 4, 4, 7, 8, 9 is $5\frac{1}{2}$, the number halfway between 4 and 7.

For 4, 5, and 7, the median is 5.

MEAN The numerical value that marks the balance point of a distribution; it is influenced by all values of the distribution including extremes and outliers. It is a good measure to use when working with distributions that are roughly symmetric.

Number of people in household: 2, 3, 3, 4, 6, 6.

The mean (average) number of people in these households is 4. There are 24 people “shared” among 6 households.

Using Measures of Variability

Measures of variability are used to describe how widely spread or closely clustered the individual data values are.

Range depends on two values, the greatest and the smallest. Range is the difference between the greatest value and the least value in the data.

On the **CMP Parent Web Site**, you can learn more about the mathematical goals of each unit, see an illustrated vocabulary list, and examine solutions of selected ACE problems. <http://PHSchool.com/cmp2parents>