Class Size Paradox

There are at least two ways to obtain the mean class size, and they can have very different results. At one college, if we take the numbers of students in 737 classes, we get a mean of 40 students. But if we were to compile a list of the class sizes for each student and use this list, we would get a mean class size of 147. This large discrepancy is due to the fact that there are many students in large classes, while there are few students in small classes. Without changing the number of classes or faculty, we could reduce the mean class size experienced by students by making all classes about the same size. This would also improve attendance, which is better in smaller classes.

Notation

Σ	denotes the *sum* of a set of data values.
x	is the *variable* usually used to represent the individual data values.
n	represents the *number of data values* in a *sample*.
N	represents the *number of data values* in a *population*.
$\bar{x} = \frac{\Sigma x}{n}$	is the mean of a set of *sample* values.
$\mu = \frac{\Sigma x}{N}$	is the mean of all values in a *population*.

Example 1 Mean

Table 3-1 includes counts of chocolate chips in different cookies. Find the mean of the first five counts for Chips Ahoy regular cookies: 22 chips, 22 chips, 26 chips, 24 chips, and 23 chips.

Solution

The mean is computed by using Formula 3-1. First add the data values, then divide by the number of data values:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{22 + 22 + 26 + 24 + 23}{5} = \frac{117}{5}$$
$$= 23.4 \text{ chips}$$

The mean of the first five chip counts is 23.4 chips.

Median

The median can be thought of loosely as a "middle value" in the sense that about half of the values in a data set are less than the median and half are greater than the median. The following definition is more precise.

DEFINITION The **median** of a data set is the measure of center that is the *middle value* when the original data values are arranged in order of increasing (or decreasing) magnitude.

Calculation and Notation of the Median The median is often denoted by $\tilde{x}$ (pronounced "*x*-tilde"). To find the median, first *sort* the values (arrange them in order), and then follow one of these two procedures:

1. If the number of data values is odd, the median is the number located in the exact middle of the sorted list.
2. If the number of data values is even, the median is found by computing the mean of the two middle numbers in the sorted list.

Important Properties of the Median

- The median does not change by large amounts when we include just a few extreme values (so the median is a *resistant* measure of center).
- The median does not use every data value.

Example 2 Median

Find the median of the five sample values used in Example 1: 22 chips, 22 chips, 26 chips, 24 chips, and 23 chips.