Chapter 3 Measuring Central Tendency of Ungrouped Data

I. Introduction

A. Central tendency describes the middle of data. It represents a typical value.
B. Measures of central tendency are called averages.
C. The **arithmetic mean** is the most common average. It is used to measure grades, success in sports, business success, and many other interesting subjects.
D. Population parameters are represented by Greek capital letters.
E. Sample statistics are represented by Arabic lowercase letters.

II. The mean

A. The sample mean ($\overline{X}$)

1. Linda is interested in how many self-help videotapes she rented last year. If substantial, she will stock a larger variety of tapes. To make an estimate, she counted last week's self-help tape rentals and recorded the following sample data. The data is a sample because she only included part of last year's data.
2. Daily self-help tape rentals were: 3, 7, 7, 4, 1, 8, 5.

$\overline{X} = \frac{\sum x}{n}$ where $\overline{X}$, read x bar, is the sample mean. x is the variable being measured.

Σ is the Greek capital letter sigma. It is the symbol for addition. n is the sample size.

$$\overline{X} = \frac{\sum x}{n} = \frac{3+7+7+4+1+8+5}{7} = \frac{35}{7} = 5$$

B. The population mean (μ)

1. Had Linda used all of last year's data, this population mean formula would have been used.
2. μ is the Greek capital letter for M and it is read Mu.
3. N is the population size.

$$\mu = \frac{\sum x}{N}$$

C. A weighted mean ($\overline{X}_w$)

1. When a data set has a number of duplicate values, a weighted mean is often calculated.
2. Each variable occurring more than once is assigned a variable name consisting of capital x with a subscript and a weight (W) with a similar subscript.
3. Linda's Video Showcase rents tapes for \$2, \$3, and \$4. The weighted mean of receipts per tape for a day of 36, 18, and 6 respective rentals is calculated as follows:

$$\overline{X}_w = \frac{W_1X_1 + W_2X_2 + W_3X_3 + \bullet \bullet \bullet + W_nX_n}{W_1 + W_2 + W_3 \bullet \bullet \bullet + W_n} = \frac{\sum (W_x X_x)}{\sum w_x}$$

$$\overline{X}_w = \frac{(36)(\$2) + (18)(\$3) + (6)(\$4)}{36 + 18 + 6} = \frac{\$72 + \$54 + \$24}{60} = \frac{\$150}{60} = \$2.50$$

Note: W_1 refers to how often X_1 happens.

D. The sum of the deviations around a mean equals zero.

1. $\Sigma(x-\mu) = 0$
2. The mean of 1, 3, and 8 is 4.
3. The sum of the deviations around the mean would be calculated as follows:

$$\Sigma(x-\mu) = (1-4) + (3-4) + (8-4)$$
$$= (-3) + (-1) + (4) = 0$$

E. The primary disadvantage of using the mean as a measure of central tendency concerns it being severely affected by a few values at either extreme. Using the data at the top of this page as an example, the mean is small because a big snowstorm resulted in a day with only 1 rental.