

According to Mark Twain there are three kinds of lies: “lies, damned lies and statistics” while Winston Churchill reportedly said that “the only statistics you can trust are the ones you have falsified yourself”.



You might wonder why statistics have such a bad reputation. I guess one basic contributing factor is that a majority of people do not feel comfortable or confident dealing with numbers. Many suspect that arguments and reasons that include or rely on numbers are deliberately made more complicated than is necessary, that we are being manipulated. Take the term “average”, a commonly used word that is not as simple as we may be led to believe. In statistical terms, averages are measures of central tendency and there are three common measures – the mean, median and mode. However, within the mean there are three variations – simple, trimmed and weighted. Simple Average A simple mean is the sum of all the values in the population or sample being measured and is sensitive to extreme values, e.g. I have 100 people in a group: 10 have \$10,000 25 have \$1,000 15 have \$100 40 have \$10 10 have \$1 The total amount of money they have together is \$126,910 and the simple average per person is \$1,269.10 Trimmed Average A trimmed mean is calculated by excluding extreme values from the average. The drawback here is that the decision on what should be included or excluded is necessarily arbitrary and may create a misleading impression. If we exclude the top 10 values of \$10,000 each and bottom values of \$1 each: The total amount of money the remaining 80 people have is \$26,900 and the trimmed average is \$336.25 Weighted Average A weighted mean is computed by attributing a weight to each value. If the example above were weighted according to the proportion of people at each value (i.e. 10% have \$10,000; 25% have \$1,000; 10% have \$100; 40% have \$10, and 10% have \$1): The weighted sum is \$16,636, the weighted number is 26.5 and the weighted average is \$627.77 Median The median value is the middle value in a population or sample. If the population or sample contains an even number of items, the median is the simple average of the two middle values. In the example give above the median is the simple average of the 50th value which is \$100 and the 51st value which is \$10 = $\$100 + \$10 = \$110 / 2 = \55 . Mode The mode is the value that occurs with the greatest frequency – in the

example given, the mode is \$10 because there are 40 people with \$10. So now we have five distinct values for the average, depending on how we choose to calculate it: Simple average (mean) \$1,269.10 Trimmed average (mean) \$ 336.25 Weighted average (mean) \$ 627.77 Median \$ 55.00 Mode \$ 10.00 If you intend to use and/or rely on averages in planning your business, it is important to know what those averages actually mean. If I create my marketing plan assuming that the average person has \$1,269.10 and am not aware that 50% of this group have \$10 or less I am inviting disaster. Key questions to consider when looking at averages are:



1. What kind of average is this? 2. Is it based on a population or on a sample? 3. How is the population or sample defined? 4. How reliable is the data? I will talk about populations and samples in a later post, highlighting some of the main aspects of each as well as the relative advantages and disadvantages.