

## 6.4: Robust Statistics

A *robust statistic* or *resistant statistic* is one that is less affected by outliers than a non-robust or non-resistant statistic. If you look at the numbers in Example 6.2 you can see that the value of the MD (and IQR) is completely unaffected by the value of the outlier data point 50. The mean and the standard deviation will, however, be greatly affected by the value of the outlier. So while some people may identify outliers as those being (say)  $3\sigma$  from the mean, we see that that is a non-robust way of identifying outliers. In summary:

Measures of central tendency and dispersion	
Robust	Non-robust
MD	$\bar{x}$
IQR	$s$

It would seem that inferential statistics based on robust statistics would be better than statistics based on non-robust values. Maybe. But, traditionally, statistical analysis like the  $t$ -tests, ANOVA and regression, are based on the non-robust statistics of means and standard deviations (or variance). People tend to use robust statistics in “Exploratory Data Analysis” (EDA). With EDA one is not concerned so much with testing hypothesis as in trying to get an understanding of general trends in the data. The techniques, and statistics, the fall under the two categories are:

Traditional	Exploratory Data Analysis (EDA)
Frequency Tables	Stem and Leaf Plot
Histogram	Box Plot
Mean, $\bar{x}$	Median, MD
Standard Deviation, $s$	Interquartile Range, IQR

You will find an EDA menu under Analyze → Descriptives in SPSS.

This page titled [6.4: Robust Statistics](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Gordon E. Sarty](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.