

# Descriptive statistics (location)...

Indicator	Definition	Formula	In Excel	In Stata	In R
Location					
Mean	The mean is the sum of the observations divided by the total number of observations. It is the most common indicator of central tendency of a variable	$\bar{X} = \frac{\sum X_i}{n}$	=AVERAGE(range of cells) For example: =AVERAGE(J2:J31)	-tabstat var1, s(mean) or - sum var1	summary(x) mean(x) sapply(x, mean, na.rm=T)
Median	The median is another measure of central tendency. To get the median you have to order the data from lowest to highest. The median is the number in the middle. If the number of cases is odd the median is the single value, for an even number of cases the median is the average of the two numbers in the middle. It is not affected by outliers. Also known as the 50 <sup>th</sup> percentile.  <div style="text-align: center;"> <p>2 6 <u>7</u> 8 9</p> <p>2 6 <u>7 8</u> 9 10</p> </div>		=MEDIAN(range of cells)	- tabstat var1, s(median) or - sum var1, detail	summary(x) median(x) sapply(x, median, na.rm=T) #median
Mode	The mode refers to the most frequent, repeated or common number in the data		=MODE(range of cells)	mmodes var1	table(x) (frequency table)

**NOTE:** For `mmodes` you may have to install it by typing `ssc install mmodes`. You can estimate all statistics in Excel using “Descriptive Statistics” in “Analysis Toolpack”. In Stata by typing all statistics in the parenthesis `tabstat var1, s(mean median)`. In R see [http://www.ats.ucla.edu/stat/r/faq/basic\\_desc.htm](http://www.ats.ucla.edu/stat/r/faq/basic_desc.htm)