

Alex Olteanu  
Assignment #5

1. Outliers are observations in our data samples that do not necessarily have many commonalities with most other observations in the data set. For example, while most values may be within one standard deviation from the mean, an outlier would be one of the few exceptions.
2. I would determine an outlier by how much it differs from the rest of the population regarding its mean, median and standard deviation placement.
3. You should delete an outlier when real world factors that are hard to quantify severely influenced this allegedly skewed result. Either this or if the outlier is value truly is statistically different enough from the rest of the data population.
4. Deleting outliers in this process would affect the mean depending on how many outliers there are. For example, one high and one low outlier being deleted would probably cancel out. On the other hand, if we have 1 low outlier and 2 high outliers, deleting one from each would still result in a mean that becomes skewed.