

Descriptive Analysis Definitions
to accompany
Descriptive Analysis Worksheet

- **Class Limits** are numbers that tell you what data values are to be tabulated in what intervals. You'll find several different definitions of class limits in different elementary statistics books. We will use the endpoints of a half-open interval $[a,b)$ as class limits. This is the approach used by many software packages and your graphing calculator. For example, your first three class limits might be $[0,3)$, $[3,6)$ and $[6,9)$. In that case, the data values 2, 2.9, 2.99 and 2.999 should be placed in Class 1. The values 3, 3.5, 5.9, and 5.9999 should go in class 2, etc.
- The **Class Mark** of a class is a representative number within the class. The mid-point of each class is usually chosen to be its class mark.
- The **Tally** column is used to count the number of data values that fall into a class.
- The **Frequency** of a class is the final tally count of the number of data values in that class.
- The **Relative Frequency** of a class is the fraction (or percent) of data values that fall into that class. In other words, if there are n total data values in the data set and the frequency of a class is f , then the relative frequency of the class is f/n written as a fraction or as a percent.
- The **Cumulative Frequency** of a class is the total number of data values in that class or a previous class. Example: If the Frequency of Class 1 is 5, that of Class 2 is 6 and that of Class 3 is also 6, then the Cumulative Frequency of Class 3 is $5+6+6=17$.
- The numerical statistics (mean, median, etc.) are defined in your text. We will use those definitions.
- A (Frequency) **Histogram** is a specialized bar graph with the Class Limits on the horizontal axis, bar height equal to the frequency of the corresponding class, and no spaces between bars. The total area of the bars in a histogram is n , the total number of data values. We may also at times construct a Relative Frequency Histogram which has Relative Frequency instead of Frequency on the vertical axis. The total area of the bars in a Relative Frequency Histogram is 1.
- An **Ogive** is a line graph of the Cumulative Frequency versus the UPPER Class Limit. That is, upper class limits are displayed on the horizontal axis; cumulative frequency on the vertical axis.
- A **Box Plot** or Box-and-Whiskers Plot has vertical marks drawn on a number line at the first and third quartiles, forming left and right ends of a box, another vertical mark at the median, vertical marks at the minimum and maximum data values and line segments called "whiskers" drawn from the box to the max and min marks. Box plots, a.k.a. "five number summary plots", are often used to compare two or more data sets.