# Frequency Distributions and Relative Frequency Distributions
### (a.k.a. Constructing a Frequency Distribution Table and a Relative Frequency Distribution Table)

Many errors occur when making frequency distributions and relative frequency distributions. Frequency is the number of occurrences or the count while relative frequency is the proportion which can be represented as a fraction, a decimal value between zero and one, inclusive, or a percentage. Since the world relates to proportions represented as percentages, you will be expected to represent relative frequencies as percentages on examinations and quizzes; carefully read the directions for MyMathLab questions so that you will provide the appropriate format (fraction, decimal value between zero and one, inclusive, or percentage) when relative frequencies are required/requested.

Read and work through the examples provided below, taking careful note of the finished frequency distribution table and relative frequency distribution table that I have provided for you. In addition, be careful to use appropriate terminology when discussing percentages: a percentage is a quantity while percent is a unit of measure.

## <u>Frequency Distribution</u>

When you construct a frequency distribution, you need to count the number of times each category or value of the variable is present in the data set. For this, you will find it helpful to keep a tally. While the tally will provide you with the frequency (the number of times that the value of the variable is present in the data set), this tally (and its associated hash marks) is NOT part of the frequency distribution. The tally is a way of keeping track of the count. So, a completed frequency distribution table will NOT have a column for the tally. These tally columns that you will find online and in the text are provided to help you to construct the frequency distribution – they are scratch work.

A completed frequency distribution has two columns, one for the variable (the left-hand column) and the other for the associated frequencies (the right-hand column) for the values of the variable, each column has a meaningful column heading (the name of the variable for the left-hand column and "Frequency" for the right-hand column), and an appropriate, meaningful title. Please be careful to make sure that you do not represent the frequency distribution by creating what looks like unrelated columns sitting next to each other on a piece of paper. To make a table for your frequency distribution, simply make a horizontal line and write the name of the variable above the line on the left-hand side and write "Frequency" above the line on the right-hand side. Line up the values of the variable below the name of the variable to create the column of the table containing the values of the variable and, then, write the frequencies *corresponding* to these values of the variable below the word "Frequency" (and on the line corresponding to the associated value of the variable) to create the frequency column. Please see the example below.

### Frequency Distribution for the Class Year for the Students in MATH 117

| Class Year | Frequency |
|------------|-----------|
| Freshman   | 5         |
| Sophomore  | 10        |
| Junior     | 5         |
| Senior     | 1         |

Please notice that there is no total for the frequency column and no tally column. If the variable had units then these units must be recorded with the name of the variable. For example, if the variable in the first column were the height then our column heading would be "Height, in inches" if the heights of the individuals had been measured in inches.

The purpose for the title for *any* table or graph is to tell the reader about the information provided in the table or graph – to put the table or graph *in context*. A template for an easy title is *what you are making* (here, Frequency Distribution), *for what variable* (here, the class year), *and in what context or for whom* (here, for the students in MATH 117). Since data provided without context is meaningless, it is vital that you always put the information provided in tables and graphs in context.

**CAUTION**: A properly constructed, complete frequency distribution is a table with two (2) columns, one for the variable and one for the corresponding frequency for each value of the variable. Each column is labeled

appropriately, and the table has a meaningful title.  A properly constructed frequency distribution does not include a column for a tally (again, a tally is scratch work and does not belong in this table).  Please see Table 2 on page 63, Tables 14 on page 82, and Table 15 on page 82 of your text book for what *NOT* to give as final-form frequency distribution; these include scratch work and do not have a title.


**Relative Frequency Distribution**

When you construct a relative frequency distribution, you need the frequency distribution information as a base:  you need the values of the variable and their associated frequencies within the data set.  However, a relative frequency distribution table does NOT have a column for the frequency or a column for the tally or *any* scratch work.

Relative frequency is the ratio of the frequency for a value of the variable to the total frequency for the variable – the fraction created by taking the number of times that the value of interest of the variable appears in the data set and dividing it by the total frequency (the sum of the frequencies for all the values of the variable).  In our example, since there are five (5) freshmen, ten (10) sophomores, five (5) juniors, and one (1) senior, the total frequency is $5 + 10 + 5 + 1 = 21$.

To determine the relative frequency for the freshmen, we take the quotient of the frequency for the freshmen, 5, and the total frequency, 21, to obtain

- $\dfrac{5}{21}$ , if we are expressing the relative frequency as a fraction,

- $\dfrac{5}{21} \approx 0.238095238095238095238095523809524$ , if we express the relative frequency as a decimal

  (of course, we must determine the number of decimal places to use), and

- $\dfrac{5}{21} \approx 23.8\%$ , if we express the relative frequency as a percentage using one decimal place.


So, the base representation for relative frequency is the fraction, here, $\dfrac{5}{21}$ , which can then be expressed

as a decimal value between zero and once, inclusive, by dividing 5 by 21, 0.238 if expressed to three (3) decimal places, and, then, multiplying 0.238 by 100%, we obtain the relative frequency expressed as a percentage, 23.8%, the first relative frequency recorded in the table below.  Therefore, according to this table, 23.8 percent of the students in MATH 117 are freshmen.

Relative Frequency Distribution for the
Class Year for the Students in MATH 117

| Class Year | Relative Frequency, in percent |
|---|---|
| Freshman | 23.8 |
| Sophomore | 47.6 |
| Junior | 23.8 |
| Senior | 4.8 |

We would proceed in a similar manner in order to obtain the relative frequencies for the sophomores, juniors, and seniors.  Once the relative frequencies have been determined for each value of the variable, we can construct the relative frequency distribution table.

A relative frequency distribution table is constructed in a similar manner to that of the frequency distribution table.  Like the frequency distribution table, the relative frequency distribution table has two columns, one for the variable (the left-hand column) and the other for the associated relative frequencies (the right-hand column) for the values of the variable, each column has a meaningful column heading (the name of the variable for the left-hand column and "Relative Frequency, in percent" for the right-hand column), and an appropriate, meaningful title.

As previously discussed, we will record the relative frequency as a percentage and to indicate this, we use a column heading of "Relative Frequency, in percent". Since we represent the relative frequency as a percentage, writing "in percent" in the heading informs the reader that the relative frequency has been provided as a percentage and provides the units for each value in the column. (Remember, the units for a percentage are percent.)

As with the frequency distribution table, if the variable in the first column were height, for example, then the column heading would be recorded with its associated units as well. The total (sum) of the relative frequencies is NOT recorded at the bottom of the relative frequency column and, since the units for the relative frequency are recorded at the top of the column as part of the column heading, we do not write a percent sign (%) after the values of the relative frequency. Since the units for the relative frequency are provided in the column heading, writing a percent sign (%) after the entries in the relative frequency column would make no sense; if we included a percent sign after 23.8, for example, the entry in the table below would actually be 23.8%% (please note the two percent signs, one would be the from the column heading and the other would be from writing a percent sign on the entry).

**CAUTION**: A properly constructed, complete relative frequency distribution is a table with two (2) columns, one for the variable and one for the corresponding relative frequency for each value of the variable, each column is labeled with a meaningful column heading, and the table has a representative, meaningful title. A properly constructed relative frequency distribution does not include a column for corresponding frequency for each value of the variable or for a tally (the frequency and tally are scratch work for the creation of a relative frequency distribution and do not belong in this table). Please see Table 3 on page 64, Table 7 on page 69, Table 9 on page 79, and Table 13 on page 81 of your text book for what NOT to give as final-form relative frequency distribution; ; these include scratch work and do not have a title.