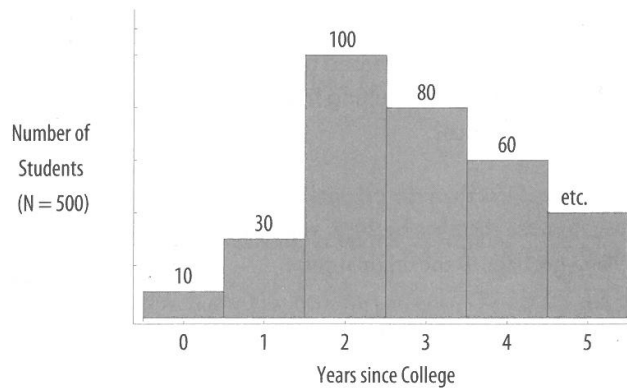


Descriptive Statistics

Say there are 500 people in your business school class and you want to think about the *number of years* each of you spent working between college graduation and business school.

To make things simple, you'll probably round to the nearest whole number (instead of having data like 5.25 years, 7.8 years, etc.). Whole numbers are *discrete* (or able to be separated and counted), so with this information, you can make a *histogram* to display the count in each category.



These three terms are very commonly tested on both IR and Quant:

1. Average (arithmetic mean)
2. Median
3. Mode

The arithmetic mean is the most important. You may already know the formula from the GMAT Quant section:

$$\text{Mean} = \frac{\text{Sum of all terms}}{\text{Number of terms}}$$

Add up everyone's "years since college" and divide that total by 500, the number of people in your class:

$$\begin{aligned} \text{Mean Years} &= \frac{\text{Total years}}{\text{Number of students}} \\ 3.36 &= \frac{1,680}{500} \end{aligned}$$

The **median** is the middle number, or the 50th percentile: half of the people have more years since college (or the same number), and half have fewer years (or the same number). If you have an odd number of terms, say {1, 2, 3, 4, 5}, then the median is the middle number (in this case, 3). If you have an even number of terms, say {1, 2, 3, 4}, then the median is the average of the two middle terms (in this case, 2.5). The median of the Years since College distribution is 4 years because the 250th and 251st students are both 4 years out of college.

The **mode** is the observation that shows up the most often, corresponding to the highest frequency on the histogram. If none of the years to the right of 4 have more than 100 people, then the histogram peaks at this entry and the mode is 2 years.