Now that we have been introduced to the dataset, it is time to start plotting graphs that will help us understand what information is stored within the data. There are many different questions that can be answered. What is the GPA distribution like? Were the students included in the survey top students with very high GPAs? Do we have the same number of students from the business and engineering schools? Do engineering students have a higher GPA than business students? Did males engage in more academic misconduct than females? Do students with lower GPAs justify cheating more than students with a high GPA?

As you can see, we can think of a lot of questions, and finding the answer to these questions is actually interesting. All of the above questions can be answered using numbers. However, as humans, visualizing things allow us to better understand what is going on. This is why if you look at the slides used in most presentations, you will probably see more graphs than numbers. The right graph can show us the answer to more than one question. Although there are people who know how to produce graphs, many of them do not know when to use the right graph. Plotting graphs is not enough. There are many types of graphs and our choice depends on two factors. The first factor is what is the nature of the information that is stored in the variable? The variable *gpa* for example stores a number than can take on any value from 40 to 100 (in the three universities included in this survey, the lowest grade that a student can get on a course is 40). The variable *major* on the other hand can take on one of eight values. The same graph does not work for both variables.

The second question that affects what type of graph we should use is what is the question that we are seeking to answer. When we do research, we start with a question, and then we determine he best tools used to answer the question. It's just like when someone asks you what is the best computer. It depends on what you will use it for. Will you use it for games? For social media? Or just to take notes in class? The tool depends on the question.

In the next section, we will start by looking at just a single variable. Sometimes our question involves just one variable, just like when we wondered whether the students in our dataset had high or low GPAs. In data analytics, understanding the individual variables is the first step in understanding the dataset. If you want to have some information about a country you can read an article that summarizes the economy, the political situation, and the status of the education industry in that country. However, if you want to really to get to know the country, then you should go there and meet the individual people in order to get to know them. Let us now go and meet the individual variables in our dataset.