Raw data consists of long lists of numbers and labels that don't seem to be very informative. Raw data lacks context. Exploratory data analysis is what you use to make sense of the data. You do this by converting data from its raw form, to a form that makes sense, that has context, that tells the story you want to tell. Basically, exploratory data analysis consists of organizing and summarizing raw data, looking for important features and patterns in the data, looking for any striking deviations from those patterns, and interpreting your findings in the context of the problem or research question. We'll begin exploratory data analysis by looking at one variable at a time, also called univariate analysis. In order to convert raw data into useful information, we need to summarize and then examine the distribution of any variables of interest. By distribution of a variable, we mean what values the variable takes, and how often the variable takes those values. >> If we were studying a small number of observations, we could do this with a pencil and paper, a calculator, or even in our heads. The data sets you're working with, often have thousands of observations. Working with such large samples is only achievable if we use statistical software. These software programs require the use of syntax or formal code to retrieve, analyze, and manipulate data. Learning to write code, learning the proper use of syntax can really expand your capacity for engaging in statistical applications. And is a central skill you will learn in this course. This skill will also greatly expand your capacity for engaging in deeper levels of quantitative reasoning about data. >> For this course, you will be using Python. Python is a widely used general purpose language which is designed to be more readable. That is, the code is easier to read and write than in other general purpose languages, such as C++ or Java. While Python was not developed specifically for data analysis pandas and other libraries provide data analysis tools for use with the Python language. Looking at all the windows, options,, menus and features though, can be rather daunting. So it's important for you to realize, this course will introduce yo to the basics. You will learn what you need to know to get started asking and answering interesting questions about data. >> In the beginning, you may feel like you're learning another language. Basically, you are. As you work through your project, you should begin to feel more comfortable implementing the various decisions you'll be making about the data. When you need help, seek it from your

course moderator, professor, your peers, or from course discussions.
>> But do I write r? >> No, no, no. >> Now you need to install Python.
Follow the available instructions
in the document called Getting Started with Python.