Anaconda

Welcome to this lesson on using Anaconda to manage packages and environments for use with Python. With Anaconda, it's simple to install the packages you'll often use in data science work. You'll also use it to create virtual environments that make working on multiple projects much less mind-twisting. Anaconda has simplified my workflow and solved a lot of issues I had dealing with packages and multiple Python versions.

Anaconda is actually a distribution of software that comes with `conda`, Python, and over 150 scientific packages and their dependencies. The application `conda` is a package and environment manager. Anaconda is a fairly large download (~500 MB) because it comes with the most common data science packages in Python. If you don't need all the packages or need to conserve bandwidth or storage space, there is also **Miniconda**, a smaller distribution that includes only `conda` and Python. Miniconda can do everything Anaconda does, but doesn't have the preinstalled packages. You can still install any of the available packages with `conda`, it just doesn't come with them, so either Anaconda or Miniconda are fine for this course.

`Conda` is a program you'll be using exclusively from the command line, so if you aren't comfortable using it, check out this command prompt tutorial for Windows or our Linux Command Line Basics course for OSX/Linux.

You probably already have Python installed and wonder why you need this at all. First, since Anaconda comes with a bunch of data science packages, you'll be all set to start working with data. Secondly, using `conda` to manage your packages and environments will reduce future issues dealing with the various libraries you'll be using.

Managing Packages