
[← Pre-Course Work](#)

6 / 11

Python and Data Science

Even if we narrow our focus to data science, it's difficult to encapsulate the breadth of Python's applications. Python is used to collect data from sources like databases and APIs, clean and organize that data, then build, tune, and evaluate machine learning models.

Python has many libraries to assist with these types of functions, including:

- [Pandas](#): Facilitates the cleaning, organization, and structuring of data. Think of this as a souped-up Excel for use within Python.
- [Scikit-learn](#): Includes a wide variety of statistical and machine learning models, allowing data scientists to try multiple models on their data sets without the time commitment of building them independently.
- [TensorFlow](#): Is growing in popularity as neural networks become more widespread and broadly used. There are increasingly few areas of coding in which Python would not be a valid choice as the chief programming language.
- [NumPy](#) and [StatsModels](#): Allow for advanced mathematical calculations and statistical analysis.

With so many functions available in these easy-to-use libraries, it's no wonder that nearly every industry uses Python code in some capacity!

