Print to PDF

Python Fundamentals

Contents

To kickstart the training program, well start off with one of the most important tools in data science: python.

Python is used for diverse tasks in the data science pipeline. From data cleaning, data exploration, statistical analysis, visualization, model building, presentation, etc.

In this section of the training you will learn the basics of python programming.

Python is easy

Python is designed to be close to natural language. It was created to promote readability, maintainability and communication. Instead of dedicating the time to learn difficult syntax, you can easily jump into writing python scripts.

Python is well supported

Python is not only well-loved in data science, it is well-loved across many fields. Because of this python enjoys the support of a strong and passionate community that loves to contribute new libraries and tools.

Python tools for Data Science

Throughout this training course you will be learning python and the tools associated to python for data science.

pandas

pandas is used for all data cleaning, data exploration, data analysis. With pandas you're able to structure data in an intuitive way, you're able to manipulate data easily, and seamlessly

numpy

numpy is the heart of almost every involving vectorized operations in python. You will use numpy when you're interacting with numerical data,

matplotlib

matplotlib is used for creating rich customizable visualizatio. You will matplotlib to explore data, evaluate models, and present results

scikit-learn

scikit-learn offers a wide library of machine learning models. It allows you to tune, cross validate, and evaluate models

Table of Contents