

Python for Data Sciences Training

Alta3 Research Authorized Training · Applied Technology Academy

In this course you will learn to use Python, the most popular programming language for data sciences, for data analysis and data visualization. Explore Python libraries to more easily sort and analyze data sets for emerging trends. Quickly produce Excel quality visualizations appropriate for displaying data in real time monitoring systems.

LEVEL	DURATION	EXPERIENCE	AVERAGE SALARY	LABS
Beginner	5 Days	1 year: Python	\$141,658	Yes

Course Overview

Intro to data science using Python libraries like pandas and numpy to identify trends within datasets. Create rich visualizations with matplotlib, folium and seaborn. Use open source toolset scipy for mathematics, science, and engineering applications. Introduction to scikit-learn, a machine learning tool for datasets.

Course Outline

- **Lesson 1: Introduction to Python Libraries for Data Sciences**

- Python with Jupyter Notebook overview Live code
- Equations
- Data cleaning
- Transformation
- Numerical simulation
- Statistical modeling
- Data visualization
- Machine Learning
- Pandas Filter DataFrames
- Dictionaries to DataFrames
- CSV to DataFrames
- Excel to DataFrames
- Numpy Work across arrays
- Requests Pull from RESTful APIs
- JSON

- **Lesson 2: Sort, Analyze, and Visualize Data with Python**

- Matplotlib Line Plots
- Area Plots
- Histograms

- Bar Charts
- Pie Charts
- Box Plots
- Scatter Plots
- Bubble Plots
- Waffle Charts
- Word Clouds
- Seaborn visualization techniques Relational
- Categorical
- Distributions
- Regressions
- Folium interactive leaflet maps
- rich vector/raster/HTML visual markers
- Saving visualizations output in various formats
- **Lesson 3: Python and Databases**
 - Creating a database engine in Python
 - sqlite3
 - Looking at tables in a database
 - Querying relational databases
 - MySQL and Python
 - SQL Queries Filtering with SQL WHERE
 - Ordering with SQL ORDER BY
 - Querying with pandas
 - Table relationships with INNER JOIN
 - MongoDB Understanding noSQL
 - Python and MongoDB
 - Pymongo Query
 - Find
 - Delete
 - Update
 - Limit
- **Lesson 4: Introduction to Machine Learning with Python**
 - scipy open ecosystem numerical integration
 - Optimization
 - linear algebra
 - statistics
 - Scikit-learn Applications of Machine Learning

- Training vs Testing sets
- Supervised vs Unsupervised Learning
- Python libraries suitable for Machine Learning
- Loading an example dataset
- Learning and predicting
- **Lesson 5: Introduction to Machine Learning with Python (continued)**
 - Scikit-learn Model persistence
 - Conventions
 - Refitting and updating parameters
 - Multiclass vs. multilabel fitting
 - Moving output to remote systems Streaming (push) to real-time dashboard APIs
 - Move data with SFTP
 - Labs
 - Lab 01 – Using vim
 - Lab 02 – Making and Syncing a Github account
 - Lab 03 – Using Jupyter Notebook
 - Lab 04 – Working with Local Files
 - Lab 04 – Pandas DataFrames
 - Lab 05 – CSV to DataFrames
 - Lab 06 – Excel to DataFrames
 - Lab 06 – Numpy Array
 - Lab 07 – Requests and APIs
 - Lab 08 – Getting JSON from RESTful APIs
 - Lab 09 – Matplotlib and Line Plots
 - Lab 10 – Matplotlib and Histograms
 - Lab 11 – Matplotlib and Pie Charts
 - Lab 12 – Matplotlib and Scatter Plots
 - Lab 13 – Matplotlib and Bubble Plots
 - Lab 14 – Matplotlib and Bar Charts
 - Lab 15 – Seaborn and Relational Visualizations
 - Lab 16 – Seaborn and Categorical Visualizations
 - Lab 17 – Seaborn and Distributions
 - Lab 18 – Seaborn and Regression Models
 - Lab 19 – Folium and Leaflet Maps
 - Lab 20 – Filtering SQL
 - Lab 21 – Ordering SQL
 - Lab 22 – Querying SQL with Pandas
 - Lab 23 – Querying MongoDB with pymongo

- Lab 24 – Pymongo Find, Delete, Update, Limit
- Lab 25 – Scipy and numerical integration
- Lab 26 – Scipy and linear algebra
- Lab 27 – Scipy and statistics
- Lab 28 – Scikit-learn and machine learning
- Lab 29 – Training vs Testing sets
- Lab 30 – Scikitlearn and supervised learning
- Lab 31 – Scikitlearn and unsupervised learning
- Lab 32 – Pushing data to real time dashboard APIs
- Lab 33 – Moving data with SFTP
- Lab 34 – Emailing with Attachments

Intended Audience

This course was written for professionals interested in Python and Data Sciences.

This includes:

- Engineers,
- Mathematicians,
- Actuaries,
- Network Specialists,
- System Admins, and developers.

Prerequisites

Keyboard proficiency, and some previous python coding experience is the only hard requirement. Students with some previous exposure to Python, or any another scripting experience, will take the most from the course. In lieu of previous experience, Alta3 research's Python Basics course is recommended.

Recommended Prerequisite: Python Basics (5 days)

Follow-on Courses

Recommended Follow up:

APIs and API Design with Python (5 days)

Recommended Follow up:

Machine Learning with Python (5 days)

Ready to enroll? Call 800.674.3550 or email Info@AppliedTechAc.com — private team cohorts, GSA MAS purchasing, military credentialing funding, VA GI Bill and VR&E and student financing available.