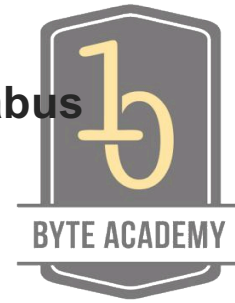


Byte Academy Data Science in Python Syllabus



Common topics throughout course

- POSIX Compatible Operations
- Git workflow
- Unix/Linux usage
- Debugging
- Pandas
- PEP 8 Compliance
- Group weekend projects

Phase 0: Pre Course Work

- Environment setup (Setup Ubuntu, GitHub & Python Development Environment)
- Fundamentals of Python & Javascript programming language
- Fundamentals of Writing Software and System Operations (Iteration, Control Flow, Editing and Executing Files)
- Foundations of Mathematics and Statistics for Data Science

Phase 1-1: Python Basics


- Learning Bash commands and Bash Development Environment
- Programming Fundamentals Review: Iteration, Control flow, Python Data Structures
- Python Modular Programming: Python Standard Library (standard libraries typically include definitions for commonly used algorithms, data structures, and mechanisms for input and output.)
- Python Functional Programming: Anonymous Functions, Decorator Functions, Iterators, Generators, Functional Objects
- Python Object Oriented Programming: Classes, Objects, Inheritance
- Best Practices: Keeping it simple, DRY code, naming conventions, comments and documentation
- Python mini project: Well- documented Python module; (those are one to two-hour mini project which could be finished in class so that the students can discuss with peers and get feedback from the instructor.)

Phase 1-2: Computer Science, Beyond the basics

- Introduction to Computer Science
- Big O Notation, Data Structures, Sorts and Searches
- MVC - Model Views Controller
- Introduction to Database Management System and SQL Introduction
- (Weekend) - Building a terminal application utilizing the MVC Design pattern and a SQL Database for persistent data



Phase 1-3: Databases

- SQL Relationships, Joins
- CRUD operations and Introduction to APIs
- DOM (Document Object Model)
- Data Formats – XML, JSON, CSV
- Object Relational Mapping, ETL Concepts
-  SQL (MongoDB)
- (Weekend) - Building a terminal application with the MVC Design pattern, persisting data in SQL, and utilizing APIs to grab data in JSON format

Phase 1-4: Web Scraping and Phase 1 Assessment

- Introduction to Web Scraping using Beautiful Soup, Requests and Selenium
- Phase 1 Assessment

Phase 2: Core Data Science Concepts

Phase 2-1: Hypothesis Testing and Linear Regression(s)

- Probability Distributions
- Hypothesis testing (test of statistical significance)
- Ordinary Least Squares
- BLUE Assumptions
- Evaluating performance
- Python Libraries – statsmodels, linalg

Phase 2-2: More Regression(s), Regularizations, Optimization

- Logistic Regression
- Minimizing Error
- Ridge and Lasso Regressions
- Gradient Descent
- Python Libraries – Scikit-learn

Phase 2-3: Classification methods

- Naive Bayes
- KNN
- LDA
- Support Vector Machines
- Cross-validation techniques for tuning



Phase 2-4: Ensemble methods

- Decision Trees
- Bagging and boosting
- Perceptron
- Supervised Neural Networks

Phase 2-5: Unsupervised and Deep Learning

- K-means Clustering
- Dimensionality Reduction
- Auto encoding, Convolution Neural Networks
- Recurrent Neural Networks
- LSTM

Phase: 2-6: Natural Language Processing

- Preparing textual data using Regular Expressions
- Entity Extraction, Lemmatization
- Textual Classifiers (Naive Bayes, SVM)
- Sentiment Analysis

Phase 3: Big Data and Final Projects

- Foundations
 - Monolithic & Micro Service Architecture
 - Cluster & Cloud Computing principles
- Big Data
 - Map Reduce
 - Hadoop
 - Spark
- Cloud Computing Platforms:
 - Digital Ocean, AWS
 - Google Cloud
 - Azure
- Final Project Demo