

Byte Academy Data Science Syllabus

Prerequisites: All levels, no prior data science experience required.



1. Intro To Data Science

- Data Science Fundamentals
 - What is it? Applicability/Importance?
- Data Science Workflow
 - Python Data Science Stack
 - Tools: IPython Notebook, Anaconda
 - Modules: Panda, numpy, scikit-learn, scipy, nltk, etc.
 - Collaboration: GitHub

2. Data Acquisition, Exploration, and Wrangling

- Web Scraping, Regular Expressions, Reshaping
 - BeautifulSoup, re, pandas
- Exploratory Data Analysis
- Data Visualization
 - i. Seaborn, matplotlib, ggplot, bokeh
- GeoSpatial Data Analysis
 - Data Types
 - Point, Polygon, MultiPolygon, etc.
 - Analysis & Plotting
 - Modules: Geojsonio, shapely, geopandas
- Small Project with Data Analysis

3. Data Storage and Management

- Intro to Databases
- SQL
 - Relational Database Design/Modeling
 - SQL specifics
 - SELECT, UPDATE, INSERT, DELETE, WHERE, GROUP BY, JOINS, etc.
 - Foreign keys, indices,
- NoSQL
 - Documents, Key-Value stores, Auto-sharding, etc.
 - MongoDB
- Database Small Project

4. Big Data Technologies

- Fundamental Concepts
 - Map-Reduce, Data Lakes
- Cluster and Cloud Computing
 - Amazon AWS, Google
- Hadoop
- Apache Spark
- Small Project with Map-Reduce and Spark on AWS

5. Prediction and Machine Learning

- Statistical Modeling & Inference
 - Probability, Distributions
 - Classical/Frequentist Statistics
 - Regression and Bias
 - Scipy.stats & statsmodel packages
 - Small Project
- Machine Learning
 - Standard Machine Learning Algorithms
 - Regression Models, Classification, SVM, Decision Trees, Random Forests
 - Scikit-learn
 - Intro to Deep Learning
 - RNNs & CNNs
 - Intro to TensorFlow
 - Other Tools: Theano, Lasagne
 - NLP & Text Mining
 - Nltk Corpora
 - Brown
 - Stop Words
 - Nltk Package
 - Lexical analysis
 - N-gram models
 - Parts-of-speech tagger
 - Small Project

6. Data Science Project