

BDSCP Module 4: Fundamental Big Data Analysis & Science

This course provides an in-depth overview of essential topic areas pertaining to data science and analysis techniques relevant and unique to Big Data with an emphasis on how analysis and analytics need to be carried out individually and collectively in support of the distinct characteristics, requirements and challenges associated with Big Data datasets.

The following primary topics are covered:

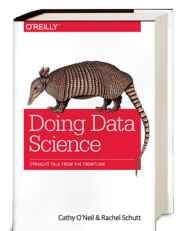
- Data Science, Data Mining & Data Modeling
- Big Data Dataset Categories
- Exploratory Data Analysis (EDA) (including numerical summaries, rules & data reduction)
- EDA analysis types (including univariate, bivariate & multivariate)
- Essential Statistics (including variable categories & relevant mathematics)
- Statistics Analysis (including descriptive, inferential, correlation, covariance & hypothesis testing)
- Data Munging & Machine Learning
- Variables & Basic Mathematical Notations
- Statistical Measures & Statistical Inference
- Distributions & Data Processing Techniques
- Data Discretization, Binning, Clustering
- Visualization Techniques & Numerical Summaries
- Correlation for Big Data
- Time Series Analysis for Big Data

Duration: 1 Day

For more information, visit www.bigdatascienceschool.com.

Text Book

This BDSCP course module covers a range of in-depth topics that are described in the course booklet(s) and further elaborated by detailed technical coverage and case study examples in the accompanying *Doing Data Science* text book.



Self-Study Kit

The materials for this course module can be purchased separately as part of the Module 4 Self-Study Kit, which includes additional materials and study aids. These materials are designed to prepare you for Exam B90.04 and are also suitable for general remote, self-paced study purposes.

For ordering information, visit www.bigdatascienceschool.com/store.

