# BDSCP Module 7: Fundamental Big Data Engineering

This course explores introductory topics pertaining to the field of developing data processing solutions–data engineering–in the context of Big Data environments. Specifically it covers concepts, techniques and technologies related to the processing and storage of Big Data datasets including MapReduce and NoSQL. It highlights the unique challenges faced when processing and storing Big Data datasets. The MapReduce data processing engine, which is the de facto framework for batch processing of large amounts of data, is also explained in detail.

The following primary topics are covered:

- Big Data Engineering – Big Data Engineering Challenges
- Big Data Storage Terminologies (including sharding, replication, CAP theorem, ACID, BASE)
- Big Data Storage Requirements
- On-Disk Storage (including distributed file system – databases)
- Introduction to NoSQL – NewSQL
- NoSQL Rationale – Characteristics
- NoSQL Database Types (including key-value, document, column-family and graph databases)
- Big Data Processing Requirements
- Big Data Processing (including batch mode and realtime mode)
- Introduction to MapReduce for Big Data Processing (batch mode)
- MapReduce Explained (including map, combine, partition, shuffle and sort, and reduce)
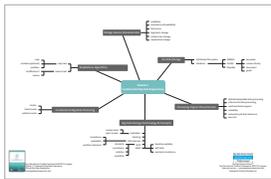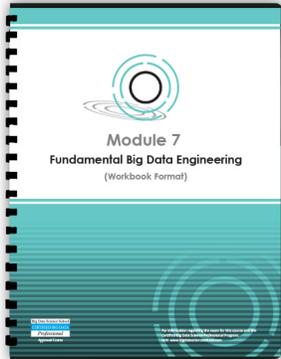
Duration: 1 Day

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 *NoSQL: A Brief Guide to the Emergine World of Polygot Persistence Distilled Text Book* text book.
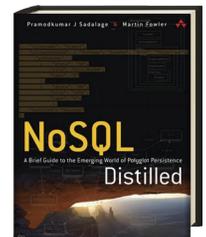
Self-Study Kit

The materials for this course module can be purchased separately as part of the Module 7 Self-Study Kit, which includes additional materials and study aids. These materials are designed to prepare you for Exam B90.07 and are also suitable for general remote, self-paced study purposes.

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

Arcitura Big Data Science School - Big Data Science Certified Professional (BDSCP) Official Curriculum
For more information, visit www.bigdatascienceschool.com
Copyright © Arcitura Education Inc

PEARSON VUE

Big Data Science School
CERTIFIED BIG DATA
Professional