

# DATA SCIENCE AND BIG DATA ANALYTICS

## An 'open' course to unleash the power of Big Data



“We live in a data-driven world. Increasingly, the efficient operation of organizations across sectors relies on the effective use of vast amounts of data. Making sense of big data is a combination of organizations having the tools, skills and more importantly, the mindset to see data as the new “oil” fueling a company. Unfortunately, the technology has evolved faster than the workforce skills to make sense of it and organizations across sectors must adapt to this new reality or perish.”

- [Andreas Weigend](#), Ph.D Stanford, Head of the Social Data Lab at Stanford, former Chief Scientist, Amazon.com



### ORDER YOUR STARTERKIT TODAY!

Become a certified Data Science Associate (EMCDSA)

- Immerse yourself in a near-classroom experience without the need for traveling
- Learn from top EMC subject-matter experts at your own pace
- Access complete student materials including slides and hand outs
- Leverage modular content tailored to your unique learning needs
- View recorded walk-through lab demonstrations
- Prepare for the E20-007 Data Science and Big Data Analytics certification exam
- Ask questions or discuss topics with experts and peers via online community

Visit: <http://education.EMC.com/DataScience>



## DATA SCIENCE AND BIG DATA ANALYTICS COURSE

An 'open' course and certification focused on concepts and principles applicable to any technology environment and industry.

This course is intended for:

- Business and data analysts looking to add big data analytics skills
- Managers of business intelligence, analytics, or big data groups
- Database professionals looking to enrich their analytic skills
- College graduates considering data science as a career field

The course provides a hands-on\* practitioner's approach to the techniques and tools required for analyzing Big Data.

The course is designed to enable students to:

- Become an immediate contributor on a data science team
- Assist reframing a business challenge as an analytics challenge
- Deploy a structured lifecycle approach to data analytics problems
- Apply appropriate analytic techniques and tools to analyze big data
- Tell a compelling story with the data to drive business action
- Use open source tools such as R, Hadoop, and Postgres
- Prepare for EMC Proven™ Professional Data Scientist certification

Visit <http://education.EMC.com/DataScience> to

- Register for the Data Science and Big Data Analytics StarterKit
- Access additional information on the course and the certification

\* The Video ILT includes recorded walk-through lab demonstrations.

# DATA SCIENCE AND BIG DATA ANALYTICS COURSE OUTLINE

Applying a hands-on practitioner's approach to the techniques and tools required for Big Data Analytics.



Big Data Overview

State of the practice in analytics

The role of the Data Scientist

Big Data Analytics in industry verticals

Introduction to Big Data Analytics



Key roles for a successful analytics project

Main phases of the lifecycle

Developing core deliverables for stakeholders

End-to-end data analytics lifecycle



Introduction to R

Analyzing and exploring data with R

Statistics for model building and evaluation

Using R to execute basic analytics methods



K-Means Clustering

Association Rules

Linear and Logistic Regression

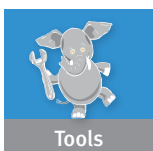
Naïve Bayesian Classifier

Decision Trees

Time Series Analysis

Text Analysis

Advanced analytics and statistical modeling for Big Data – Theory and Methods



Using MapReduce/Hadoop for analyzing unstructured data

Hadoop ecosystem of tools

In-database Analytics

MADlib and Advanced SQL Techniques

Advanced analytics and statistical modeling for Big Data – Technology and Tools



How to operationalize an analytics project

Creating the Final Deliverables

Data Visualization Techniques

Hands-on Application of Analytics Lifecycle to a Big Data Analytics Problem

Endgame, or Putting it all together

## CONTACT US

Engage your local Education Services Consultants for local pricing information.

Online: <http://education.EMC.com/Contact>

Phone: 1-888-EMC-TRNG (888-362-8764)



Follow us on Twitter!  
[twitter.com/EMCProven](http://twitter.com/EMCProven)



Connect with Our Community!  
<http://education.EMC.com/ProvenCommunity>



Like us on Facebook!  
[facebook.com/ProvenProfessional](http://facebook.com/ProvenProfessional)



Learn more and purchase Data Science and Big Data Analytics Video-ILT by scanning this QR code.

EMC<sup>2</sup>, EMC, EMC Proven, and the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners. © Copyright 2014 EMC Corporation. All rights reserved. Published in the USA. 06/14

EMC Corporation  
Hopkinton, Massachusetts 01748-9103  
1-508-435-1000 In North America 1-866-464-7381  
[www.EMC.com](http://www.EMC.com)