

## Data Science and Big Data Analytics

Dieser Kurs vermittelt praxisorientiertes Wissen, um an Projekten mitzuwirken, bei denen große Datenvolumen (Big Data) verarbeitet werden und bei denen besondere Anforderungen an Analysefunktionen gestellt werden. Neben einer Einführung zu Big Data sowie zum Data Analytics Lifecycle vermittelt er grundlegende und fortgeschrittene Analysemethoden sowie eine Einführung in die Technologien und Werkzeuge, die zur Analyse von Big Data eingesetzt werden, wie MapReduce und Hadoop. Ausführliche Laborübungen sorgen für den Bezug zur Praxis. Der Kurs folgt einem Technologie-neutralen Ansatz und beinhaltet eine abschließende Übung im Labor, bei der die Teilnehmer die im Kurs erlernten Konzepte im Data Analytics Lifecycle-Umfeld anwenden. Er bereitet die Teilnehmer auf die Zertifizierung als Proven™ Professional Data Scientist Associate (EMCDSA) vor und vermittelt Grundlagen zu Data Science, welche durch weitere Schulungen und praktische Erfahrung vertieft werden können.

### Kursinhalt

- Introduction to Big Data Analytics
- Data Analytics Lifecycle
- Basic Data Analytic Methods Using R
- Advanced Analytics - Theory and Methods
- Advanced Analytics - Technologies and Tools
- The Endgame, or Putting it All Together

Jeder Teilnehmer erhält die englischsprachigen Original-Unterlagen von Dell Technologies.

### Zielgruppe

- Teamleiter, deren Aufgabengebiete Informationsdienstleistungen für Unternehmen und Datenanalyse umfassen, sowie alle Personen, die sich professionell mit Big Data auseinandersetzen.
- Analytiker, die sich mit Geschäftsprozessen und Daten beschäftigen und ihre Analysefähigkeiten für Big Data erweitern möchten.
- Daten- und Datenbankexperten, die ihre Analysefähigkeiten für Big Data erweitern möchten.
- Quereinsteiger, die sich in Data Science und den Bereich Big Data einarbeiten wollen.
- Personen, welche die Zertifizierung als EMC Proven Professional Data Scientist Associate (EMCDSA) erwerben wollen.

### Voraussetzungen

- Umfassendes Grundlagenwissen zu quantitativer Analyse und Statistik, wie es in einem Statistics 101 Level-Kurs vermittelt wird.
- Erfahrung mit einer Skriptsprache, z.B. Java, Perl, oder Python (oder R).
- Erfahrung mit SQL (in einigen im Kurse verwendeten Beispielen wird PSQL eingesetzt).

Für die Buchung eines Dell Technologie Kurses benötigen Sie vorab einen Account bei Dell. Diesen Account können Sie sich unter <https://education.dell.com> im Bereich Sign in (oben rechts) einrichten. Bitte beachten Sie, dass Sie sich mit Ihrer persönlichen E-Mailadresse Ihres Unternehmens anmelden müssen.

### Dieser Kurs im Web



Alle tagesaktuellen Informationen und Möglichkeiten zur Bestellung finden Sie unter dem folgenden Link: [www.experteach.de/go/EMDA](http://www.experteach.de/go/EMDA)

### Vormerkung

Sie können auf unserer Website einen Platz kostenlos und unverbindlich für 7 Tage reservieren. Dies geht auch telefonisch unter 06074 4868-0.

### Garantierte Kurstermine

Für Ihre Planungssicherheit bieten wir stets eine große Auswahl garantierter Kurstermine an.

### Ihr Kurs maßgeschneidert

Diesen Kurs können wir für Ihr Projekt exakt an Ihre Anforderungen anpassen.

Training		Preise zzgl. MwSt.	
<b>Termine in Deutschland</b>	<b>5 Tage</b>	<b>€ 2.995,-</b>	
<b>Online Training</b>	<b>5 Tage</b>	<b>€ 2.995,-</b>	
<b>Termin/Kursort</b>	<b>Kurssprache Englisch </b>		
12.05.-16.05.25  Online	11.08.-13.08.25  Online		
23.06.-27.06.25  Online	01.09.-05.09.25  Online		

Stand 07.05.2025



# Inhaltsverzeichnis

## Data Science and Big Data Analytics

### Introduction to Big Data Analytics

- . Big Data Overview
- . State of the Practice in Analytics
- . The Data Scientist
- . Big Data Analytics in Industry Verticals

### Data Analytics Lifecycle

- . Discovery
- . Data Preparation
- . Model Planning
- . Model Building
- . Communicating Results
- . Operationalizing

### Review of Basic Data Analytic Methods Using R

- . Using R to Look at Data iV Introduction to R
- . Analyzing and Exploring the Data
- . Statistics for Model Building and Evaluation

### Advanced Analytics iV Theory And Methods

- . K Means Clustering
- . Association Rules
- . Linear Regression
- . Logistic Regression
- . Naive Bayesian Classifier
- . Decision Trees
- . Time Series Analysis
- . Text Analysis

### Advanced Analytics - Technologies and Tools

- . Analytics for Unstructured Data - MapReduce and Hadoop
- . The Hadoop Ecosystem
- o In-database Analytics iV SQL Essentials
- o Advanced SQL and MADlib for In-database Analytics

### The Endgame, or Putting it All Together

- . Operationalizing an Analytics Project
- . Creating the Final Deliverables
- . Data Visualization Techniques
- . Final Lab Exercise on Big Data Analytics

