BDSCP MODULE 04 | Fundamental Big Data Analysis & Science

Arcitura® CERTIFIED *Big Data Scientist*

OVERVIEW

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
- High-Volume, High-Velocity, High-Variety, High-Veracity, High-Value Datasets
- Exploratory Data Analysis (EDA)
- EDA Numerical Summaries, Rules and Data Reduction
- EDA analysis types, including Univariate, Bivariate and Multivariate
- Essential Statistics, including Variable Categories and Relevant Mathematics
- Statistics Analysis, including Descriptive, Inferential, Covariance, Hypothesis Testing, etc.
- Measures of Variation or Dispersion, Interquartile Range & Outliers, Z-Score, etc.
- Probability, Frequency, Statistical Estimators, Confidence Interval, etc.
- Data Munging and Machine Learning
- Variables and Basic Mathematical Notations
- Statistical Measures and Statistical Inference
- Confirmatory Data Analysis (CDA)
- CDA Hypothesis Testing, Null Hypothesis, Alternative Hypothesis, Statistical Significance, etc.
- Distributions and Data Processing Techniques
- Data Discretization, Binning and Clustering
- Visualization Techniques, including Bar Graph, Line Graph, Histogram, Frequency Polygons, etc.
- Prediction Linear Regression, Mean Squared Error and Coefficient of Determination R2, etc.
- Clustering k-means, Cluster Distortion, Missing Feature Values, etc.
- Numerical Summaries

Duration: 1 Day

TRAINING + CERTIFICATION

This course is part of a training and accreditation program through which official certification can be achieved and for which official training and certification badges can be issued.

www.arcitura.com/bdscp/scientist
www.arcitura.com/bdscp/consultant

eLEARNING + PRINTED STUDY KITS

This course is available as part of a study kit that includes course materials and supplemental self-study materials. The study kit is available in full-color printed format and eLearning format. Both types of study kits include the contents listed below.

- Workbook
- Self-Study Guide
- Supplement
- Mind Map Poster
- Flashcards
- Audio Tutor Recordings

Arcitura®

Pearson | VUE | OnVUE | Arcitura On-Site EXAM PROCTORING | acclaim | Credly

BDSCP MODULE 04 | Fundamental Big Data Analysis & Science

	Certified Big Data Professional*	Certified Big Data Science Professional	Certified Big Data Scientist	Certified Big Data Consultant	Certified Big Data Engineer	Certified Big Data Architect	Certified Big Data Governance Specialist
MODULE 01 Fundamental Big Data	●	●	●	●	●	●	●
MODULE 02 Big Data Analysis & Technology Concepts	●	●	●	●	●	●	●
MODULE 03 Big Data Analysis & Technology Lab		●		●			
MODULE 04 Fundamental Big Data Analysis & Science			●	●			
MODULE 05 Advanced Big Data Analysis & Science			●				
MODULE 06 Big Data Analysis & Science Lab			●				
MODULE 07 Fundamental Big Data Engineering				●	●		
MODULE 08 Advanced Big Data Engineering					●		
MODULE 09 Big Data Engineering Lab					●		
MODULE 10 Fundamental Big Data Architecture						●	
MODULE 11 Advanced Big Data Architecture						●	
MODULE 12 Big Data Architecture Lab						●	
MODULE 13 Fundamental Big Data Governance							●
MODULE 14 Advanced Big Data Governance							●
MODULE 15 Big Data Governance Lab							●

** The Certified Big Data Professional designation is automatically issued when achieving any other BDSCP certification. It can also be achieved by receiving passing grades on Exams B90.01 + B90.02.*

Arcitura Big Data School (www.arcitura.com/big-data-school)