Module 2
Big Data Analysis & Technology Concepts

Big Data Analysis Lifecycle Stages

- Business Case Evaluation
- Data Identification — of: external datasets, internal datasets
- Data Acquisition & Filtering
- Data Extraction
- Data Validation & Cleansing
- Data Aggregation & Representation
- Data Analysis
- Data Visualization
- Utilization of Analysis Results — via: alerts, applications, business process optimization

Big Data Technology Components & Concepts

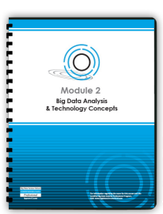
- Clusters
- File Systems & Distributed File Systems
- NoSQL
- Distributed Data Processing
- Parallel Data Processing
- Processing Workloads — types: batch, transactional
- Cloud Computing

Big Data Analysis Techniques

- Statistical Analysis: A/B Testing, Correlation, Regression
- Semantic Analysis: Natural Language Processing, Sentiment Analysis, Text Analytics
- Machine Learning: Classification, Clustering, Outlier Detection, Filtering
- Visual Analysis: Heat Maps, Network Analysis, Spatial Data Analysis, Time Series Analysis

Big Data Mechanisms

- Analytics Engine
- Coordination Engine
- Data Transfer Engine — types: Ingress, Egress
- Processing Engine
- Query Engine
- Resource Manager
- Storage Device
- Workflow Engine



Big Data Science Certified Professional (BDSCP) Program
Module 2: Big Data Analysis & Technology Concepts
Official Mind Map Supplement
www.bigdatascienceschool.com

Big Data Science School
CERTIFIED BIG DATA
Professional

Big Data Science School™
Big Data Science Certified Professional (BDSCP) Program
www.arcitura.com • www.bigdatascienceschool.com
Copyright © Arcitura Education Inc.