



## Advanced-Data Engineering for Data Management Professionals

30 Sep - 04 Oct 2024  
Madrid (Spain)



# Advanced-Data Engineering for Data Management Professionals

**Ref.:** 15681\_320431 **Date:** 30 Sep - 04 Oct 2024 **Location:** Madrid (Spain) **Fees:** 4900 Euro

## Introduction:

This course is designed for professionals working as Data Managers or in related roles who seek to deepen their understanding and skills in Data Engineering. It will cover advanced topics, including data pipelines, big data architectures, data governance, and performance optimization. By the end of the program, participants will be equipped with the knowledge and tools to effectively manage and engineer data systems at an enterprise level.

## Targeted Groups:

- Data Managers are seeking to expand their technical expertise in Data Engineering.
- Senior Data Engineers aiming to refine and enhance their skills.
- IT professionals are transitioning into data management roles.
- Data Architects and Database Administrators looking to broaden their scope of knowledge.
- Technical leaders responsible for data strategy and infrastructure.

## Course Objectives:

At the end of this course, the participants will be able to:

- Design and implement robust data pipelines for complex data environments.
- Optimize data architectures for scalability and performance.
- Apply best practices in data governance and compliance.
- Utilize advanced tools and techniques for data processing and storage.
- Lead data engineering projects within their organizations with confidence.

## Targeted Competencies:

- Advanced data pipeline design and implementation.
- Mastery of big data technologies and frameworks.
- Expertise in data governance and security.
- Skills in performance tuning and system optimization.
- Leadership in data engineering and project management.

## Course Content:

### Unit 1: Advanced-Data Pipeline Design:

- Understanding Data Pipelines Explore the architecture and components of data pipelines, including sources, transformations, and destinations.
- Batch vs. Real-Time Processing: Delve into the differences, benefits, and challenges of batch and real-time data processing.
- ETL vs. ELT Processes Compare and contrast the Extract, Transform, Load ETL and Extract, Load, Transform ELT approaches and their use cases.
- Pipeline Orchestration Tools Study tools like Apache Airflow, Luigi, and Prefect for managing and orchestrating complex data workflows.
- Error Handling and Data Quality Implement strategies for error handling, retries, and ensuring data quality within pipelines.

### Unit 2: Big Data Architectures:

- Introduction to Big Data Understand big data's characteristics, challenges, and opportunities in modern enterprises.
- Distributed Data Processing Learn about frameworks like Apache Hadoop, Apache Spark, and Apache Flink for processing large datasets.
- Data Storage Solutions Explore storage options like HDFS, NoSQL databases e.g., Cassandra, MongoDB, and cloud storage solutions.
- Scalability and performance: Analyze strategies for scaling big data solutions and optimizing performance.
- Stream Processing Dive into tools like Apache Kafka, Apache Pulsar, and Kinesis to process streaming data.

### Unit 3: Data Governance and Compliance:

- Fundamentals of Data Governance Define data governance and its importance in ensuring data integrity, security, and compliance.
- Regulatory Compliance Study key regulations such as GDPR, CCPA, and HIPAA and how to implement compliant data practices.
- Data Cataloging and Lineage: Learn how to maintain a comprehensive data catalog and track data lineage for transparency and accountability.
- Access Control and Security Implement robust access control mechanisms to safeguard sensitive data.
- Data Stewardship Understand the role of data stewards in maintaining data quality and governance.

## **Unit 4: Performance Optimization and Tuning:**

- Database Optimization Techniques Explore advanced techniques for optimizing relational and non-relational databases.
- Query Performance Tuning Learn how to analyze and optimize SQL queries for improved performance.
- Indexing Strategies Study various indexing techniques to enhance data retrieval speeds.
- Resource Management Manage computational resources effectively to ensure efficient processing in big data environments.
- Monitoring and Troubleshooting Implement monitoring tools and techniques to proactively identify and resolve performance issues.

## **Unit 5: Leadership in Data Engineering:**

- Project Management for Data Engineering: Apply project management principles to lead data engineering initiatives successfully.
- Team Collaboration and Communication Foster effective communication and collaboration within data engineering teams.
- Data Strategy Development Develop and execute a comprehensive data strategy aligned with organizational goals.
- Change Management Manage change within data projects, including stakeholder engagement and risk management.
- Emerging Trends in Data Engineering: Stay updated on the latest trends, tools, and technologies shaping the future of data engineering.



**Registration form on the :  
Advanced-Data Engineering for Data Management Professionals**

**code:** 15681 **From:** 30 Sep - 04 Oct 2024 **Venue:** Madrid (Spain) **Fees:** 4900 **Euro**

Complete & Mail or fax to Mercury Training Center at the address given below

**Delegate Information**

Full Name (Mr / Ms / Dr / Eng):

Position:

Telephone / Mobile:

Personal E-Mail:

Official E-Mail:

**Company Information**

Company Name:

Address:

City / Country:

**Person Responsible for Training and Development**

Full Name (Mr / Ms / Dr / Eng):

Position:

Telephone / Mobile:

Personal E-Mail:

Official E-Mail:

**Payment Method**

Please invoice me

Please invoice my company