

Data Science and Engineering for Leaders

Duration: 5 days (40 hours)

Course Overview

This five-day leadership program is designed to equip managers responsible for Data Science, Data Engineering and Analytics functions with the leadership and technical knowledge required to effectively supervise and manage highly technical teams working across Microsoft Fabric and Azure. The program covers the evaluation of data architecture and platform choices, management of delivery and technical risks, data governance, cloud cost optimization, and alignment of data and technology initiatives with organizational priorities and business outcome

Prerequisites

- Learner's own subscription for Azure and Fabric Contributor access
- Experience in managing or leading technical teams, preferably in Data Engineering & Data Science.
- A basic understanding of cloud computing and enterprise technology concepts.
- Familiarity with data platforms, databases, data pipelines, analytics, or machine learning is recommended.
- Basic awareness of Microsoft Azure and/or Microsoft Fabric is helpful but not mandatory.
- No hands-on programming or advanced technical implementation experience is required.

Session Agenda

Module 1: Data Leadership Landscape

- Evolution of enterprise data platforms
- Modern data ecosystem and operating models
- Data Engineering, Data Science, Analytics, and Data Platform roles
- End-to-end data lifecycle
- Responsibilities of a data manager vs technical lead
- Assessing organizational data maturity
- Aligning data capabilities with business strategy

Module 2: Microsoft Fabric & Azure Data Ecosystem

- Microsoft Fabric capabilities from a manager's perspective
- Azure data services and their roles

- Fabric vs Azure platform considerations
- OneLake, Lakehouse, and Warehouse concepts
- Data integration and orchestration
- Enterprise data platform patterns
- Build vs buy vs reuse decisions
- Platform modernization considerations

Module 3: Data Strategy & Operating Model

- Developing a data strategy
- Current-state vs target-state assessment
- Data operating models
- Centralized, decentralized, and federated data teams
- Defining data ownership and accountability
- Prioritizing strategic data initiatives
- Building a data transformation roadmap
- Aligning technology investments with business priorities

Module 4: Data Architecture Decision-Making

- Evaluating data architecture proposals
- Scalability, performance, reliability, and maintainability
- Batch vs streaming architectures
- Lakehouse vs warehouse approaches
- Data pipelines and orchestration
- Integration and interoperability
- Architecture trade-offs
- Technical debt and modernization decisions
- Questions managers should ask during architecture reviews

Module 5: Managing Data Engineering Delivery

- Structuring Data Engineering teams
- Planning and prioritizing data initiatives
- Agile delivery for data projects
- Estimation and capacity planning
- Managing dependencies
- Identifying delivery bottlenecks
- Managing technical debt
- Balancing project delivery and platform operations
- Measuring engineering productivity

Module 6: Data Platform Operations & Reliability

- Data platform reliability and availability
- Pipeline monitoring and failure management
- Data incident management
- Root-cause analysis
- Observability and alerting
- Data lineage
- Operational SLAs and SLOs
- Disaster recovery and business continuity
- Balancing reliability, speed, and cost

Module 7: Data Quality & Data Management

- Data quality as a leadership responsibility
- Data quality dimensions
- Critical data elements
- Data profiling and validation
- Data quality monitoring
- Data ownership and stewardship
- Managing recurring quality issues
- Building trust in enterprise data
- Establishing data quality KPIs

Module 8: Data Governance, Security & Compliance

- Enterprise data governance principles
- Governance across Fabric and Azure
- Data ownership and stewardship models
- Identity and access management
- Role-based access control
- Sensitive and regulated data
- Data classification and protection
- Data cataloging and discoverability
- Lineage and auditability
- Balancing governance with engineering agility

Module 9: Analytics & Data Product Leadership

- Modern analytics ecosystem
- BI and analytics consumption patterns

- Data products and reusable data assets
- Self-service vs governed analytics
- Semantic models and business definitions
- Analytics adoption and usage
- Managing stakeholder expectations
- Measuring business value from data products

Module 10: Managing Data Science & Advanced Analytics Teams

- Data Science lifecycle
- Data Scientist vs Data Engineer vs Analyst roles
- Problem definition and experimentation
- Model and RAG pipeline
- Model evaluation and interpretation

-
-
-
- Moving from experimentation to production
- Reproducibility and operationalization
- Model monitoring and performance
- Managing Data Science delivery
- Setting realistic expectations for Data Science initiatives

Module 11: Building & Managing High-Performance Data Teams

- Designing effective Data Engineering and Data Science organizations
- Defining roles and responsibilities
- Hiring and capability assessment
- Identifying skill gaps
- Upskilling and reskilling teams
- Managing senior technical specialists
- Team capacity and workload management
- Setting meaningful team KPIs
- Creating a culture of engineering and analytical excellence
- Collaboration between technical and business teams

Module 12: Data Strategy, Cost & Executive Decision-Making

- Connecting data initiatives to business outcomes
- Identifying and prioritizing high-value data opportunities
- Data platform cost drivers
- Fabric and Azure capacity and resource management
- FinOps principles for data teams
- Cost vs performance vs reliability trade-offs
- Building business cases for data investments
- Measuring ROI and value realization
- Executive-level data KPIs
- Data leadership roadmap and capstone decision-making exercise

-
-
-