

# Data Warehousing in the Cloud

## Outline

### Introduction

#### Module 1: The Modern Data Warehouse

- Data challenges in the new world
- Evolution of the data warehouse
- Scale, complexity, and risk
- Operating cost in the cloud
- Data warehousing architecture in the cloud
- Provider landscape
- Pricing
- Storage types and tradeoffs
- Multi-cloud
- Trends

#### Module 2: Features of the Modern Data Warehouse Platform

- Parallelism
- Elasticity
- Partitioning, clustering, and indexing
- Query processing
- Joins
- Workload management

#### Module 3: Key Cloud Data Warehouse Platforms

- AWS
- Databricks
- Google Cloud Platform
- Microsoft Azure Synapse
- Oracle Autonomous Data Warehouse (ADW)
- Snowflake
- Other options
- Summary product matrix

#### Module 4: Analytics and Machine Learning

- Business analytics, edge analytics, and streaming analytics
- Machine learning
- Deep learning

#### Module 5: Cybersecurity and Privacy

- Are we secure?
- NIST SP.800-53 – 206 standards
- Breached! What next?

## **Module 6: Data Lakehouses and the Future**

- Future
- Advances in computing
- Pitfalls and risks
- Use cases
- Checklist