

Data Engineering Mastery with Snowflake and dbt

Course Description

This program equips professionals with end-to-end skills in Snowflake and dbt, covering ingestion, modeling, governance, CI/CD, and advanced AI integrations. Learners will gain practical experience through hands-on labs and real-world scenarios, enabling them to design, optimize, and operationalize modern data pipelines.

Duration

5 Days

Pre-requisites

- Basic SQL knowledge
- Familiarity with cloud concepts (Azure/AWS/GCP)
- Understanding of data warehousing fundamentals
- Prior exposure to GitHub and CI/CD concepts (recommended)

Learning Objectives

By the end of this course, participants will:

- Understand Snowflake architecture, editions, and core features
- Perform ingestion and data modeling using Snowflake
- Implement dbt workflows for Silver and Gold layer transformations
- Apply governance, RBAC, and observability practices
- Configure CI/CD pipelines for dbt projects
- Explore advanced Snowflake Cortex and MCP integrations with AI agents

Day-wise Module Coverage

Day 1

Module 1: Snowflake Basics

- Snowflake architecture
- Editions & pricing
- UI navigation (roles, users, resource monitors, cost management)
- Table types
- Zero-copy cloning
- Time travel
- Basic SQL querying
- Best practices for schema design
- Performance tuning basics

Module 2: Ingestion

- Snowflake stages
- Data loading via Snowflake UI (COPY command)
- Data loading from Azure
- Automating ingestion pipelines
- Error handling and retries

Module 3: EDW & Data Modelling

- Enterprise data warehouse concepts
- Dimensional modelling (facts & dimensions)
- Medallion architecture (Bronze / Silver / Gold)
- Low-level design (LLD) walkthrough
- Best practices for scalable models
- Handling slowly changing dimensions

Day 2

Module 4: dbt Introduction & Setup

- dbt overview — models, materialization, seeds, sources
- Snapshots (SCD Type 2) & history snapshots
- Tests
- dbt project structure & configuration
- dbt documentation generation

Module 5: Transformations — Silver Layer

- Generating Silver LLD
- Generating dbt model scripts based on the LLD
- Validating transformations with tests
- Managing dependencies between models

Day 3

Module 6: Transformations — Gold Layer

- Generating Gold LLD & scripts from Silver
- Generating dbt model scripts based on the LLD
- Building aggregated fact tables
- Creating business-ready reporting views

Module 7: Cost, Observability & RBAC

- RBAC
- Cost management
- dbt audit logging
- Row access policies
- Dynamic data masking
- Query performance monitoring

- Resource monitors & alerts

Day 4

Module 8: CI/CD

- dbt environments creation
- Setting up environment variables
- Setting up CI and merge jobs
- Showcasing CI check restrictions on PRs in GitHub
- Automated testing in CI/CD pipelines
- Deployment strategies for dbt projects

Day 5

Module 9: Snowflake Cortex

- Cortex Search
- Cortex Analyst
- Cortex Agents
- Use cases for AI-driven analytics
- Integrating Cortex with BI tools

Module 10: MCP to Claude (Live Showcase)

- MCP server creation
- Adding required tools to MCP
- Establishing connection between Snowflake MCP and Claude
- Demonstrating AI-assisted query generation
- Real-time collaboration scenarios