DATA ENGINEERING | BEGINNER | 2 DAYS

Snowflake Fundamentals

Build a strong foundation in Snowflake for modern analytics and ELT workflows, from core architecture to performance and cost-aware operations. Learn how to load and transform structured and semi-structured data, apply governance basics, and deliver analytics-ready datasets using safe and scalable patterns.

Why Snowflake?

Faster analytics scaling

Separate compute and storage to scale workloads efficiently.

Better performance control

Right-size warehouses for predictable query speed and cost.

Lower operational overhead

Reduce maintenance compared to traditional data warehouse setups.

Modern data foundation

Enable BI, data science, and AI use cases on a unified platform.

Secure data collaboration

Share data across teams and partners with controlled access.

What you'll be able to do

- Understand Snowflake architecture and how it differs from traditional databases.
- Learn core Snowflake concepts including databases, schemas, tables, virtual warehouses, roles, and access control.
- Load data into Snowflake using practical ingestion methods and COPY INTO workflows.
- Write efficient SQL queries for analytics and transformations.
- Understand micro-partitions and Snowflake performance fundamentals.
- Use Time Travel and zero-copy cloning for safe recovery and experimentation.
- Understand data sharing concepts and governance basics (intro).
- Apply cost-awareness practices for compute usage using warehouse scaling and auto-suspend.
- Build an end-to-end workflow from staging to transformed analytics tables.
- Use Snowflake UI (Snowsight) features effectively for development workflows.

Who this is for

- Data Engineers
- Analytics Engineers
- Data Analysts moving to Snowflake
- BI Engineers
- Developers working with cloud data warehouses
- DevOps/Platform teams supporting data platforms

Curriculum

01 Snowflake Architecture (How It Differs from Traditional Databases)

- Snowflake cloud-native architecture overview (storage, compute, services)
- Separation of storage and compute and why it matters
- Virtual warehouses and independent scaling concepts
- How Snowflake supports multi-tenant, secure analytics workloads
- **Hands-on:** Activity: Compare a traditional database design vs Snowflake architecture for an analytics use case

02 Core Snowflake Concepts (DBs, Schemas, Warehouses, Roles)

- Databases, schemas, tables, and naming conventions
- Virtual warehouses (sizes, scaling, concurrency basics)
- Roles, grants, and access control model fundamentals
- Best practices for environment separation (dev/test/prod concept)
- **Hands-on:** Lab: Create a database, schema, warehouse, roles, and assign permissions

03 Using Snowsight (UI Workflow and Productivity)

- Navigating Snowsight (worksheets, databases, query history)
- Using UI for role switching and context management
- Query profiling basics in Snowsight
- Managing warehouses and monitoring usage from UI
- **Hands-on:** Lab: Run queries, inspect query history, and use the UI to profile performance

04 Loading Data into Snowflake (COPY INTO and Practical Ingestion)

- Staging concepts (internal stage vs external stage overview)
- File formats and ingestion readiness (CSV, JSON basics)
- COPY INTO workflows and common options
- Error handling basics (bad records, on_error concepts)
- **Hands-on:** Lab: Create a stage, define file format, and load raw datasets using COPY INTO

05 Analytics SQL and Transformations (Efficient Query Patterns)

- Writing efficient analytics queries (filters, joins, aggregations)
- CTEs and reusable transformation patterns
- Window functions for analytics logic (rank, dedup concepts)
- Building transformed tables (CTAS patterns, views overview)
- **Hands-on:** Lab: Transform raw tables into curated analytics tables using SQL best practices

06 Performance Fundamentals (Micro-Partitions and Query Behavior)

- Micro-partitions concept and how Snowflake stores data
- Partition pruning and why filtering matters
- Clustering concept and when it becomes relevant
- Using query profile to find bottlenecks and optimize
- **Hands-on:** Lab: Compare query performance with and without pruning-friendly filters and review query profile

07 Time Travel and Zero-Copy Cloning (Safe Recovery and Experimentation)

- Time Travel use cases (recover deleted/updated data)
- Querying historical data using Time Travel syntax
- Zero-copy cloning for dev/test experiments
- Rollback and recovery workflows for analytics pipelines
- **Hands-on:** Lab: Clone a table, run changes safely, and recover data using Time Travel

08 Governance and Data Sharing (Intro)

- Data sharing concepts and common enterprise scenarios
- Secure access controls with roles and least privilege
- Governance basics (ownership, auditing, usage tracking concepts)
- Sharing readiness checklist (what to expose, what to restrict)
- **Hands-on:** Activity: Design a role model and sharing policy for analysts vs external consumers

09 Cost Awareness and Warehouse Management

- Snowflake cost model basics (compute vs storage)
- Warehouse scaling strategy (right size, concurrency handling)
- Auto-suspend and auto-resume best practices
- Monitoring cost drivers with query history and warehouse usage
- **Hands-on:** Lab: Configure auto-suspend and validate cost impact with different workload patterns

10 End-to-End Workflow (Staging → Transform → Analytics Tables)

- Designing pipeline layers (raw/staging, cleansed, curated concepts)
- Load raw data into staging tables using COPY INTO
- Transform into analytics-ready tables (facts/dimensions concept)
- Validation checks (row counts, null checks, duplicates concepts)
- **Hands-on:** Capstone Lab: Build an end-to-end Snowflake workflow and produce final KPI queries