

Diploma in Snowflake (6 Months)

Duration: 6 Months (~240 Hours)

Mode: Live Online / Classroom

Tools & Technologies: Snowflake Console, SQL, Snowpipe, Streams, Tasks, Snowpark, External Functions, Spark Integration

Certifications Prepared: SnowPro Core, SnowPro Advanced: Administrator, SnowPro Advanced: Data Engineer

Course 1: Snowflake Administration (3 Months)

Syllabus

Week 1: Introduction & Architecture

- Data Warehousing concepts and modern cloud evolution
- Snowflake overview: multi-cluster, shared-data architecture
- Layers: Storage, Compute, Cloud Services
- Snowflake Editions & cloud platforms (AWS, Azure, GCP)
- Hands-on: Set up trial account, explore Snowsight UI and SnowSQL CLI
- Assignment

Week 2: User & Role Management (RBAC)

- Snowflake access model overview
- Creating users, roles, and role hierarchy
- Privileges and grants (database, schema, table level)
- Hands-on: Create admin, developer, and analyst roles with least privilege model
- Assignment

Week 3: Virtual Warehouses & Resource Monitors

- Virtual warehouse concepts: sizing, scaling, multi-cluster
- Auto-suspend & auto-resume configuration
- Monitoring compute credits with Resource Monitors
- Hands-on: Configure multiple warehouses with scaling policies
- Assignment

Week 4: Security & Governance

- Authentication methods (MFA, Key Pair, OAuth, SSO)
- Data security: Row-Level Security (RLS), Dynamic Data Masking, Column Security
- Network policies, IP whitelisting, and firewall integration
- Hands-on: Apply masking policy on sensitive column (e.g., PII data)
- Assignment
- Mock Interview 1

Week 5: Data Loading & Unloading

- Internal vs External Stages (S3, Blob, GCS)
- File formats: CSV, JSON, Avro, Parquet
- COPY INTO for bulk load
- Unloading data from Snowflake back to cloud storage
- Hands-on: Load JSON from AWS S3 into Snowflake tables
- Assignment

Week 6: Monitoring & Query Profiling

- Query History and Access History views
- Using Query Profile to troubleshoot query performance
- Result, Metadata, and Warehouse caching layers
- Hands-on: Profile a complex SQL query and identify bottlenecks
- Assignment

Week 7: Backup, Recovery & Cloning

- Time Travel for recovery
- Fail-safe retention
- Zero-Copy Cloning (databases, schemas, tables)
- Hands-on: Recover dropped table and clone environment for sandbox testing
- Assignment

Week 8: Capstone Project & Interview Prep

- Capstone: Configure secure, multi-warehouse Snowflake environment with RBAC, resource monitors, and data loading
- Assignment
- Mock Interview 2

Course 2: Snowflake Development – SQL & Core Features (3 Months)

Syllabus

Week 1: SQL Foundations

- Databases, Schemas, Tables; Snowflake-specific data types
- DDL & DML operations (CREATE, INSERT, UPDATE, DELETE)
- Hands-on: Create schema and load initial dataset
- Assignment

Week 2: Advanced SQL Queries

- Complex Joins & Subqueries; CTEs
- Set Operators: UNION, MINUS, INTERSECT
- Window Functions (RANK, ROW_NUMBER, LEAD, LAG)

- Hands-on: Build analytics queries with ranking functions
- Assignment

Week 3: Semi-Structured Data Handling

- VARIANT data type for JSON/XML
- Flattening nested JSON with FLATTEN function
- Querying arrays & objects in Snowflake SQL
- Hands-on: Parse nested JSON payloads
- Assignment

Week 4: Views & Materialized Views

- Standard vs Secure Views; Row Access Policies
- Materialized Views for performance optimization
- Hands-on: Create secure views for sensitive fields
- Assignment
- Mock Interview 1

Week 5: Stored Procedures & UDFs

- JavaScript Stored Procedures; SQL UDFs
- Scalar & Table Functions; Table UDFs
- Hands-on: Build procedure for ETL process and UDF for string manipulation
- Assignment

Week 6: Streams & Tasks

- Change Data Capture (CDC) using Streams
- Automating ETL workflows with Tasks
- Hands-on: Incremental load using Streams & Tasks
- Assignment

Week 7: Data Sharing & Marketplace

- Secure Data Sharing; Reader accounts
- Exploring Snowflake Data Marketplace
- Hands-on: Share dataset securely with another account
- Assignment

Week 8: Capstone Project & Interview Prep

- Capstone: Build CDC-enabled ETL pipeline with Streams, Tasks, and Views
- Assignment
- Mock Interview 2

Course 3: Snowflake Development – Data Engineering & Advanced Features (3 Months)

Syllabus

Week 1: Data Ingestion with Snowpipe

- Event-driven vs manual ingestion; file notifications
- Integration with S3 / Azure Blob / GCS
- Configure Stage + Pipe; validate ingestion
- Hands-on: Auto-ingest JSON from S3 via SNS
- Assignment

Week 2: Data Modeling in Snowflake

- Star & Snowflake schema design; SCD types
- Design for semi-structured data; best practices
- Hands-on: Build schema for retail analytics
- Assignment

Week 3: Performance Optimization

- Clustering keys & Micro-partitioning strategies
- Search optimization service; pruning and statistics
- Result caching strategies and warehouse tuning
- Hands-on: Optimize queries on large tables
- Assignment

Week 4: Snowpark for Developers

- Snowpark APIs (Python, Java, Scala)
- DataFrame operations and UDFs in Snowpark
- Packaging & deployment to Snowflake
- Hands-on: Build ETL pipeline using Snowpark Python
- Assignment
- Mock Interview 1

Week 5: External Functions & API Integrations

- Creating external functions; Secrets & API integration
- Secure integration with AWS Lambda / Azure Functions
- Hands-on: External function to call REST API for FX rates
- Assignment

Week 6: Streams & Advanced Tasks

- Chained tasks, scheduling, and error handling
- Notifications & retries; orchestration patterns
- Hands-on: Automated CDC pipeline using chained tasks
- Assignment

Week 7: Integration with Big Data & ML

- Connecting Snowflake with Spark (JDBC/Connector)

- Data prep for ML; model serving options
- Hands-on: Train ML model using Snowflake + Python (sklearn)
- Assignment

Week 8: Capstone Project & Interview Prep

- Capstone: Real-time ETL pipeline with Snowpipe, Snowpark, and Tasks
- Assignment
- Mock Interview 2

Learning Outcomes

- Administer Snowflake (RBAC, warehouses, security & governance)
- Develop SQL, UDFs, Stored Procedures; implement CDC pipelines
- Ingest and process structured/semi-structured data with Snowpipe & Streams
- Automate workflows using Tasks and integrate with APIs via External Functions
- Optimize performance using clustering, micro-partitioning, and caching
- Use Snowpark for advanced data engineering and ML pipelines
- Prepare for SnowPro Core & Advanced certifications (Admin, Data Engineer)