

dbt with Snowflake - Training Syllabus

Course Overview

Duration: 4-5 Hours (Real-time Session)

Format: Interactive Live Session with Hands-on Labs

Target Audience: Data Engineers, Analytics Engineers, Data Analysts

Prerequisites: Basic SQL knowledge, familiarity with data warehousing concepts

Module 1: Introduction to Modern Data Stack (30 minutes)

1.1 Understanding the Modern Data Stack

- Evolution from ETL to ELT
- Role of data transformation in analytics
- Introduction to cloud-native data platforms
- Benefits of combining dbt with Snowflake

1.2 dbt Fundamentals

- What is dbt (data build tool)?
- Core dbt concepts: Models, Sources, Tests, Documentation
- dbt vs traditional ETL tools
- dbt Cloud vs dbt Core

1.3 Snowflake Overview

- Snowflake architecture: Compute, Storage, Services
- Key features: Multi-cluster, Auto-scaling, Data sharing
- Snowflake editions and pricing model
- Integration capabilities

Module 2: Environment Setup and Configuration (45 minutes)

2.1 Snowflake Environment Setup

- Creating Snowflake trial account
- Setting up warehouses, databases, and schemas
- User roles and permissions management
- Loading sample data (Jaffle Shop dataset)

2.2 dbt Cloud Setup

- Creating dbt Cloud account
- Connecting dbt Cloud to Snowflake (Partner Connect)
- Configuring development credentials
- Setting up Git repository integration

2.3 Project Initialization

- Creating new dbt project
- Understanding dbt project structure
- Configuration files: `dbt_project.yml`, `profiles.yml`
- Development vs Production environments

Hands-on Lab: Complete environment setup and connection testing

Module 3: dbt Models and Sources (60 minutes)

3.1 Building Your First Models

- Writing SQL models in dbt
- Materialization types: Table, View, Incremental, Ephemeral
- Model organization and naming conventions
- Understanding dbt compilation process

3.2 Working with Sources

- Declaring sources in `sources.yml`
- Using the `source()` function
- Source freshness testing
- Source documentation best practices

3.3 Staging Models

- Purpose of staging models
- Data cleaning and standardization
- Column renaming and type casting
- Creating reusable staging models

Hands-on Lab: Build staging models for customer and order data

Module 4: Advanced dbt Features (75 minutes)

4.1 Macros and Jinja Templating

- Introduction to Jinja in dbt
- Writing custom macros
- Built-in dbt macros and functions
- Dynamic SQL generation

4.2 Incremental Models

- When to use incremental materialization
- Configuring incremental strategies
- Handling late-arriving data
- Performance optimization techniques

4.3 Snapshots (Slowly Changing Dimensions)

- Understanding Type 2 SCDs
- Configuring dbt snapshots
- Timestamp vs check strategy
- Managing historical data changes

4.4 Seeds and Analyses

- Using CSV seeds for reference data
- Creating analyses for ad-hoc queries
- Best practices for seed management

Hands-on Lab: Create incremental models and implement snapshots

Module 5: Testing and Documentation (45 minutes)

5.1 Data Testing Framework

- Generic tests: unique, not_null, accepted_values, relationships
- Custom tests and assertions
- Configuring test severity levels
- Test results interpretation

5.2 Documentation Generation

- Auto-generated documentation
- Adding descriptions and metadata
- Column-level documentation
- Creating data lineage diagrams

5.3 Data Quality Best Practices

- Implementing data quality checks
- Monitoring data freshness
- Setting up alerting mechanisms
- Quality gates in CI/CD pipelines

Hands-on Lab: Add comprehensive tests and generate documentation

Module 6: Production Deployment and Orchestration (45 minutes)

6.1 Environment Management

- Development vs staging vs production
- Environment-specific configurations
- Variable management across environments
- Security considerations

6.2 Job Scheduling and Orchestration

- Setting up dbt Cloud jobs
- Scheduling strategies and best practices
- Job notifications and alerts
- Integration with external orchestrators (Airflow, Prefect)

6.3 CI/CD Implementation

- Git workflow for dbt projects
- Automated testing in pull requests
- Deployment strategies
- Rollback procedures

Hands-on Lab: Deploy models to production and schedule jobs

Module 7: Snowflake-Specific Features (30 minutes)

7.1 Leveraging Snowflake Features

- Using Snowflake-specific functions
- Working with semi-structured data (JSON, AVRO)
- Time travel and data recovery
- Clustering and optimization

7.2 Cost Optimization

- Warehouse sizing strategies
- Query optimization techniques
- Resource monitoring and alerting
- Best practices for cost management

7.3 Advanced Integrations

- Snowflake Data Marketplace integration
- External data sources
- Data sharing capabilities
- Security and compliance features

Module 8: Real-world Project Implementation (60 minutes)

8.1 End-to-End Project

- Business requirements analysis
- Data modeling design decisions
- Implementation of complete data pipeline
- Performance optimization

8.2 Troubleshooting and Debugging

- Common dbt errors and solutions
- Snowflake query optimization
- Performance monitoring
- Log analysis techniques

8.3 Maintenance and Scaling

- Model refactoring strategies
- Handling schema changes
- Team collaboration workflows
- Enterprise deployment considerations

Hands-on Lab: Complete end-to-end project implementation

Module 9: Best Practices and Advanced Topics (30 minutes)

9.1 dbt Style Guide

- SQL formatting standards
- Model organization patterns
- Naming conventions
- Code review processes

9.2 Performance Optimization

- Query performance tuning
- Materialization strategy selection
- Warehouse optimization
- Monitoring and alerting setup

9.3 Enterprise Considerations

- Multi-tenant architectures
- Security and access control
- Compliance requirements
- Disaster recovery planning

Module 10: Q&A and Wrap-up (30 minutes)

10.1 Common Challenges Discussion

- Real-world implementation challenges
- Industry-specific use cases
- Integration with BI tools
- Migration strategies

10.2 Career Development

- dbt certification paths
- Snowflake certification alignment
- Building analytics engineering skills
- Community resources and learning

10.3 Next Steps

- Advanced learning resources
- Community engagement opportunities
- Project ideas for practice
- Building a professional portfolio

Resources and Materials

Sample Datasets

- Jaffle Shop (e-commerce data)
- Knoema Economy Data Atlas
- Custom trading performance data

Tools and Platforms

- dbt Cloud account
- Snowflake trial account
- Git repository
- Sample dbt projects

Additional Learning Resources

- dbt Documentation Hub
- Snowflake University courses
- Community forums and Slack channels
- YouTube tutorials and webinars

Certification Preparation

- dbt Analytics Engineering Certification
- Snowflake SnowPro Core Certification
- Practice exams and study guides

Assessment and Certification

Practical Assessments

- Hands-on lab completion (70%)
- Final project implementation (20%)
- Participation and engagement (10%)

Certificate of Completion

- Issued upon successful completion of all modules
- Digital badge for LinkedIn profile
- Reference for professional development

Post-Session Support

Follow-up Resources

- Access to recorded sessions
- Sample code repositories
- Community Slack channel access
- Monthly office hours for questions

Continued Learning Path

- Advanced dbt techniques
- Snowflake optimization workshops
- Real-world project mentoring

- Industry-specific implementations

Register at: [Cloudhill.in](https://cloudhill.in)

Session Date: September 1st, 2025

Time: 8:30 AM IST

Transform your data career with industry-leading tools and expert guidance!