



# SNOWFLAKE COURSE CONTENT

## 1. Introduction to Cloud

- What is Cloud?
- Different Cloud Vendors
- Advantages of Cloud over On-Premise
- Bulk data loading/unloading from/to
- Amazon S3
- Bulk data loading from Microsoft Azure
- Bulk data loading from Google cloud

## 2. DataWarehousing Concepts

- What is a Data Warehouse?
- Database Vs Data Warehouse
- Why do we need Data Warehouse?
- OLTP Vs OLAP
- What is ETL?

## 3. Introduction to Snowflake

- Supported Cloud Platforms
- Snowflake Architecture
  - Cloud Service Layer
  - Compute or Query processing Layer
  - Storage Layer
- Snowflake Editions
- Setting up virtual warehouses
- Scaling policy

## 4. Connecting to Snowflake

- User Interface
- SnowSQL (CLI Client)
- Using python
- Using 3rd party tools

## 5. Data Loading and Unloading in Snowflake

- File formats supported in Snowflake
- Creating file formats
- Types of data loading
  - Limited data
  - Bulk data
  - Continuous data
- Limited data loading/unloading using Web Interface
- Bulk data loading/unloading from/to local file

## 6. Working with Semi-structured data in Snowflake

- Creating stage & raw file
- Parsing JSON
- Handling nested data
- Dealing with hierarchy
- Insert final data
- Loading PARQUET data

## 7. Stages

- What is a stage?
- Types of stages
  - Table stages
  - User stages
  - Internal names stages
  - External stages
- Creating stages
- Data Loading/Unloading into/from stages
- Querying Data in Staged Files

## 8. Snowflake Continuous Data Ingestion Service (Snow pipe)

- Introduction to Snow pipe
- Load CSV data from S3 using Snow pipe
- Load JSON data from S3 using Snow pipe
- Load Parquet data from S3 using Snow pipe

## 9. Sharing Data in Snowflake

- Introduction to Data Sharing
- Data Providers
- Data Consumers
- Create first share





- Adding or removing consumer accounts from share
- Create database from share in consumer account
- Sharing data with non-Snowflake customers (Reader Account)
- Create and Configure Reader Account
- Managing Reader Accounts
- Dropping shares

#### 10. Tasks in Snowflake – Scheduling Service

- Introduction to Tasks
- Task tree
- Implement standalone task
- Check task history

#### 11. Snowflake Streams – Change Data Capture (CDC)

- Introduction to streams
- Insert operation in streams
- Delete operation in streams
- Update operation in streams

#### 12. Time travel

- Introduction to Time Travel
- Configure Retention time property
- Query historical results
- Clone historical objects
- Restore the dropped objects

#### 13. Fail Safe

- Introduction to Fail safe
- Check fail safe storages consumption
- Fail safe storage charges

#### 14. Zero Copy Cloning

- Introduction to Zero Copy Cloning
- Clone Database, Pipes, Streams, Tasks & Stages
- Clone Schema and Tables
- DML changes in cloned table
- Cloning using Time Travel
- Swap the Tables

#### 15. SnowSQL

- Downloading and installing SnowSQL
- Creating database, tables in SnowSQL
- Loading and unload data using SnowSQL

#### 16. Managing Your Snowflake Account

- System Usage & Billing
- Understanding Snowflake Credit and Storage Usage

#### 17. Understanding Snowflake Data Transfer Billing

- Monitoring Credit and Storage Usage
- Resource Monitors
- Parameter Management
- User Management

#### 18. Types of Tables in Snowflake

- Different table types
- Permanent tables & databases
- Transient tables & databases
- Temporary tables & databases

#### 19. Micro-Partitioning in Snowflake

- Partitioning in Traditional Data Warehouses
- Partitioning in Snowflake – Micro Partitioning
- Advantages of Micro-Partitioning
- Structure of Micro-Partition
- Query Processing in Snowflake

#### 20. Managing Security in Snowflake

- System defined User roles
  - Custom roles
  - Creating Users
  - Creating Custom Roles
  - Granting roles to users
  - Multi-Factor Authentication (MFA)
- Access Control in Snowflake (DAC & RBAC)
- Data Encryption in Snowflake
- Data Encryption at customer side

#### 21. Snowflake Pricing

- What are Snowflake Credits?
- Server less Features cost
- Storage Cost
- Check storage cost for different regions
- Virtual Warehouses – Compute Cost
- Cloud Services Cost
- Data Transfer Cost
- Check consumed CREDITS on your account





## 22. Performance Tuning

- Performance in Snowflake
- Create dedicated virtual warehouses
- Scaling up and Scaling down
- Caching
- Clustering

## 23. Caching in Snowflake

- Cache introduction
- Types of Cache in Snowflake

## 24. Views, Secure Views and Materialized views

- Views in Snowflake
- Secure views in Snowflake
- Views vs. Secure Views
- Understand and creating materialized views
- Refresh materialized views
- When to use materialized views

## 25. Dynamic Data Masking

- Understanding data masking
- Creating a masking policy
- Recreate and Unset Masking Policy
- Apply Masking policies on Views
- Alter Masking Policies

## 26. Resource monitors (Track Compute Consumption)

- Resource Monitors - Introduction
- Properties of Resource Monitor
- Actions or Triggers of Resource Monitors
- Suspension or Resumption
- Create a Resource Monitor using Web UI
- SQL queries to Create Resource Monitor

## 27. Data Sampling

- Why data sampling?
- Methods of data sampling
- Sampling data: Hands-on

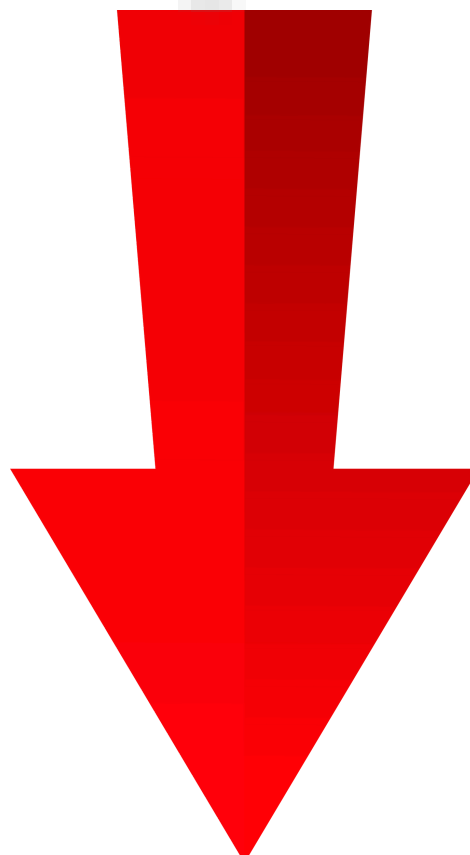
## 28. Snowflake Partner Connect

- Introduction to Partner Connect
- Know about available Tools
- Start Free Trial for Matillion ETL

## 29. Visualization – Power BI & Tableau

- Data Visualization
- Download & install Power BI
- Connect Power BI & Snowflake
- Working in Power BI
- Download & install Tableau
- Connect Tableau & Snowflake

**DBT**





# INTRODUCTION TO **DBT**

## Day 1 – Introduction to dbt

- What is dbt? Why dbt in modern data stack?
- ELT vs ETL (dbt's role in ELT)
- dbt core vs dbt Cloud
- Installing dbt (local setup with Snowflake profile)
- Understanding dbt project structure (models, seeds, macros, tests, analyses)

## Day 2 – Connecting dbt with Snowflake

- Configuring Snowflake profile (profiles.yml)
- Authentication methods (username/password, key-pair, SSO)
- Creating a dbt project (dbt init)
- Running dbt commands (dbt run, dbt debug, dbt test)
- First dbt model (SQL file) – build & materialize in Snowflake

## Day 3 – Models & Materializations

- What are dbt models?
- Materializations in dbt: view, table, incremental, ephemeral
- Model dependencies & DAG (ref function)
- dbt run order & dependency graph visualization
- Hands-on: Build staging models in Snowflake using dbt





## Day 4 – Jinja & Macros

- Introduction to Jinja templating
- Variables in dbt (var(), env\_var())
- Macros in dbt ({{ macro\_name() }})
- Using reusable SQL logic with macros
- Hands-on: Create custom macro for Snowflake SQL functions

## Day 5 – dbt Testing & Documentation

- Why testing in dbt?
- Schema tests: uniqueness, not null, accepted values, relationships
- Writing custom tests
- Auto-documentation with dbt (dbt docs generate, dbt docs serve)
- Hands-on: Build documentation site for your Snowflake models

## Day 6 – Sources & Seeds

- Defining sources (sources.yml) and referencing in models
- Source freshness checks
- What are seeds? Loading CSVs into Snowflake via dbt
- Hands-on: Define raw tables as sources, load seed data, and transform





## Day 7 – Advanced Modeling

- Staging layer, Intermediate layer, Mart layer (dbt best practices)
- Incremental models for large Snowflake datasets
- Snapshots (SCD Type 2 in dbt)
- Hands-on: Build an incremental model with partitioned Snowflake table

## Day 8 – dbt Deployment & Scheduling

- dbt Cloud overview vs CI/CD with dbt Core + GitHub Actions
- Job scheduling in dbt Cloud
- Version control integration (Git)
- Hands-on: Deploy dbt project on dbt Cloud connected to Snowflake

## Day 9 – dbt + Snowflake Performance Optimization

- dbt best practices for Snowflake
- Using Snowflake-specific configurations (cluster by, tags, secure views)
- Query optimization with dbt + Snowflake micro-partitioning
- Resource monitor alignment for dbt runs
- Hands-on: Optimize a slow dbt model in Snowflake





## Day 10 – Mini Project

- End-to-End Data Pipeline with dbt + Snowflake
- Raw data ingestion → Sources in dbt
- Staging layer → Transform raw JSON/CSV into clean tables
- Intermediate transformations → Joins, aggregations
- Mart layer → Final tables for reporting
- Build documentation & tests
- Deploy pipeline in dbt Cloud & schedule runs

**\*\*\* Python Integration and many more Integrations with Snowflake \*\*\***

**\*\*\* Mini Project - Setup Row Level & Column Level Security \*\*\***

➤ We will build mini - project to understand project structure in Snowflake.

**Project 1:** Real time major project – Complete End to End data pipe line (Real time data moving from **AWS** to Snowflake)

**Project 2:** Real time major project – Complete End to End data pipe line (Real time data moving from **AZURE** to Snowflake)

**Project 3:** Real time major project – Complete End to End data pipe line (Real time data moving from **GCP** to Snowflake)

❖ We will build Major project from scratch where we used most of features available in Snowflake and so that you will get complete real time analysis about snowflake projects.

*Thank You*

