INDATAI.COM
DATA TO INSIGHTS

SNOWFLAKE DATA ENGINEERING

Description

This training program is specifically designed for Data Engineers seeking to master Snowflake. It covers Snowflake's architecture, core features, and advanced functionalities, along with its integration with AI to enhance and streamline modern data engineering workflows. Participants will gain in-depth knowledge of data ingestion, efficient data loading, transformation processes, schema evolution, and AI-driven pipeline automation. The program is intended for professionals who will be building and managing scalable data pipelines, as well as designing and maintaining data infrastructure to efficiently load data into the Snowflake platform from various sources, supporting analytics, reporting, and other critical business tasks.
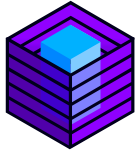
Format

Mostly workshops

Duration

2 days

Prerequisites

- Basic SQL knowledge
- Understanding of key data engineering concepts and architecture
- Familiarity with cloud computing
- Basics of software development
- Interest in data loading and data transformation

Target audience

This training is ideal for:

- **Data Engineers** looking to enhance their skills in Snowflake and cloud-based data platforms.
- **ETL/ELT Developers** focused on optimizing data ingestion, transformation, and loading processes.
- **Data Architects** aiming to design robust data pipelines and infrastructure for Snowflake.
- **Cloud Engineers** working with data solutions in cloud environments.
- **Database Administrators** interested in managing scalable data infrastructures.
- **Data Scientists** looking to streamline data access and integration for AI/ML workflows.

Workshop program

1. Snowflake and DWH fundamentals
   - Overview
   - Snowflake architecture and layers
   - Snowsight
   - Virtual warehouses
   - Snowflake objects, databases, schemas
   - Datatypes
   - Structured and semi-structured data
   - Snowflake roles and access control

INDATAI Inc.
2021 Midwest Road, Suite 200
Oak Brook, IL 60523
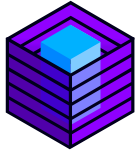
2. Connection options
   - Authentication types
   - Connectors
   - Snowflake Connector for Python
   - SnowSQL
   - Snowflake Native Connectors
   - Snowflake REST SQL API
   - Snowflake Connector for Kafka

3. Snowflake features for Data Loading
   - Storage
   - Internal stages
   - External stages
   - Streams
   - Tasks
   - Serverless Tasks

4. Data ingestion
   - ETL vs ELT
   - Pull vs Push
   - Bulk vs. Continuous Data Load
   - Change Data Capture
   - Snowpipe
   - Snowpipe Streaming
   - Snowflake Connector for Kafka
   - Snowflake Connector for Kafka with Streaming
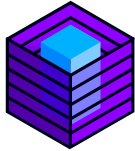
5. Data Transformation
   - Moving data from stage to table
   - Transformation during load
   - COPY command
   - Snowpipe transformation
   - Dynamic tables
   - UDF & Stored Procedures

6. Schema detection and evolution
   - Schema detection - INFER_SCHEMA
   - Object creation based on metadata
   - Automatic schema evolution

7. AI in Data Engineering
   - Snowflake Cortex
   - Copilot
   - AI enhanced Data Pipelines
   - AI generated Data Pipelines
   - Anomaly detection
   - Data Cleaning
   - Data Anonymization
   - Conversion of unstructured data into structured or semi-structured data
   - Sentiment analysis
   - RAG (Retrieval augmented generation) overview