Snowflake SnowPro Advanced: Data Engineer

Course Number: INNO-SP-DATA-ENG

Duration: 3 days

Prerequisite Courses

Snowflake - Snowflake Fundamentals

View Details

Course Description

This course prepares attendees to take the SnowPro Advanced Data Engineer exam. The advanced-level certification validates the skills and knowledge required to design and implement complex Snowflake data solutions. This class covers advanced data modeling, advanced SQL, data integration, advanced data transformation, data quality, data visualization and reporting, performance tuning and optimization, and Snowflake deployment and management.

Skills Gained

The focus is on leveraging Snowflake with common data engineering competencies. This includes:

- Source data from Data Lakes, APIs, and on-premises
- Transform, replicate, and share data across cloud platforms
- Design end-to-end near real-time streams
- Design scalable compute solutions for DE workloads
- Evaluate performance metrics.

Who Can Benefit

The audience for this class is Data Engineers.

Course Details

Exam Overview

- Quick Introductions
- Discuss exam prerequisites, format, scoring, and passing requirements
- Review exam objectives and identify key areas to focus on
- Ensure participants have the necessary course materials

Data Ingestion

Load Data into Snowflake

- Discuss considerations for data loading
- Data loading features and potential impacts
- Ingest data of various formats (CSVs, Parquet Files, etc.)
- Outline stages
- Troubleshooting data ingestion

Data Pipelines and Snowpark

Design, build, and troubleshoot continuous data pipelines

- Design a data pipeline that forces uniqueness but is not unique
- Discuss and implement Stages, Tasks, Streams, Snowpipe, Auto-Ingest vs. Rest API
- Understand Snowpark architecture
- Deploy a UDF and a stored procedure using Snowpark
- Design and use the Snowflake SQL API

Managing Data in Snowflake

- Using connectors to connect to Snowflake
- Design and build data sharing solution - Implement a Data Share, create a secure View, implement row-level filtering
- External Tables - When to use and how they work, partitioning
- Materialized Views
- Partitioned data unloading

Performance Optimization

Troubleshooting Underperforming Queries

- How to identify underperforming queries
- Understanding metrics around query operations
- Increasing efficiency and cost considerations
- Identifying the root cause of underperforming queries

Understanding performance improvement levers available in Snowflake

- Scale-Out vs. Scale In
- Clustering vs. increasing warehouse size
- Search optimization in Snowflake
- Performance of Materialized Views
- Discuss query complexity
- Micro-Partitions and the impact of clustering
- Discuss caching in Snowflake
- How to monitor continuous data pipelines - Snowpipe, Stages, Tasks, and Streams

Storage and Data Protection

- Implement data recovery features in Snowflake (Time Travel + Fail Safe)
- Impact of Streams on Time Travel
- Using system functions to analyze micro-Partitions (Cluster Depth + Cluster Keys)
- Using Time Travel in DevOps
- Creating a backup database
- Test changes before deployment
- How to rollback using Time Travel

Security

- Authentication Methods: Single-Sign-On, Key Authentication, Username/password, Multi-Factor Authentication
- Dive into Role-Based Access Control in Snowflake with examples
- Combining column-level security, and data masking with RBAC to secure sensitive data
- Discuss System Defined Roles - usage and best practices (SYSADMIN, SECURITYADMIN, USERADMIN, ACCOUNTADMIN, ORGADMIN)
- Data Governance
- Column Level Security - Dynamic Data Masking and External Tokenization
- Row Level Security - snowflake row access policies

- DDL for Dynamic Data Masking + row access policies
- Object Tagging (method and best practices)
- Masking Policies (method and best practices)

User Defined Functions

- Create a secure UDF
- SQL and Javascript UDFs
- Returning table vs. scalar values in UDFs
- Define and create external functions (Secure External Functions)

Stored Procedures and Data Transformation

- Stored Procedures - design, build, and leverage
- Transaction management in Snowflake
- Handle and transform semi-structured data
- Convert semi-structured data to structured data
- Convert structured data to semi-structured data

Snowpark for Data Transformation

- Query and filter data using the Snowpark library
- Perform data transformation using Snowpark (aggregations, etc.)
- Joining Snowpark data frames

Exam Preparation

- Mock Questions from the Exam
- Group Discussion on Questions

Related Certifications

Authorized training from Ascendient is created and maintained by the vendor who also creates the certification exams. While it may not be a requirement of certification to attend a vendor-authorized training class, doing so will put you in the best position to successfully complete the related exams. Start training and begin working towards one of the following certifications today.

SnowPro Certifications
SnowPro Advanced Data Engineer
View Certification arrow_forward
View All Certifications arrow_forward