



SNOWFLAKE SNOWPARK DATAFRAME PROGRAMMING

ONE-DAY COURSE

23129

OVERVIEW

This one-day course covers key Snowpark concepts, features, and programming constructs intended for practitioners who will be building DataFrame data solutions in Snowflake. This course consists of lectures, demos, labs, and discussions.

ACQUIRED SKILLS

- Describe Snowpark's client-side and server-side capabilities.
- Connect to Snowflake using a Snowpark Session object.
- Query data sources as Snowpark DataFrame objects.
- Perform basic and advanced data transformations using a library of DataFrame functions.
- Action DataFrame objects to process results client-side or persist results in Snowflake.
- Create shareable and reusable code as User-Defined Functions (UDFs).
- Encapsulate a sequence of operations or conditional logic into a single, reusable object with Stored Procedures.

WHO SHOULD ATTEND

- Data Engineers
- Data Scientists
- Data Application Developers
- Database Architects
- Database Administrators
- Data Analysts with programming experience

PREREQUISITES

- Snowflake Hands-on Essentials: "Data Warehousing Workshop" or equivalent knowledge required.
- Previous data warehouse knowledge is assumed.
- Basic proficiency writing code in one of the following languages: Java, Scala, or Python.
- Familiarity with Snowflake objects and basic SQL.

DELIVERY FORMAT

Instructor-led Public or Private classes are available.

TOPICS COVERED

Snowpark Overview

- Snowpark Technical Overview
- Getting Started with the Snowpark API
- Setting Up Snowflake Connections and Exploring Multiple Authentication Methods
- Discovering What DataFrames are in Snowpark and How They Run on Snowflake's Elastic Compute Engine

Creating Snowpark DataFrames

- Exploring Multiple Methods to Create a DataFrame Object
- Key Concepts of Programming in Snowpark DataFrames Including Schemas, Data Types, and Lazy Evaluation
- Constructing Basic Create Statements

Transforming DataFrames: Basic and Advanced Operations

- Applying Column Operations for Filtering and Transforming Data
- Using Scalar Functions and Operators
- Sorting and Limiting Results
- Performing Aggregate and Set-based Operations on DataFrames
- Transforming Semi-structured Data in DataFrames

Actions on DataFrames: Evaluating and Persisting

- Identifying the Differences Between and How to Use DataFrame Actions and Transformations
- Evaluating DataFrame Transformations with Actions that Return Data to the Client-side
- Publishing Logical DataFrame Operations as Views
- Creating and Appending Snowflake Tables with DataFrame Results

Creating and Registering User-Defined Functions (UDFs)

- Writing a Basic UDF in Snowpark
- Registering and Granting Access to UDFs to Share Code with Others

Authoring a Stored Procedure

- Making Dependencies Available to Your Code
- Using a Python Worksheet to Create and Deploy a Stored Procedure