SNOWFLAKE SNOWPARK DATAFRAME PROGRAMMING

Prerequisites: Basic knowledge of java, scala or python is required, foundational Knowledge of snowflake

Duration: 1 Day (8 Hrs./Day)

Course Objective: 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

Lab Requirement: Koenig DC/Linux.

Module 1 - 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

Module 2 - 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

Module 3 - 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

Module 4 - 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

Module 5 - Creating and Registering User-Defined Functions (UDFs)

Writing a Basic UDF in Snowpark

Registering and Granting Access to UDFs to Share Code with Others

Module 6 - Authoring a Stored Procedure

Making Dependencies Available to Your Code

Using a Python Worksheet to Create and Deploy a Stored Procedure