Simplify pipeline for batch and stream pipelines using Delta Live Tables.

Workshop Details:

|  |  |
| --- | --- |
| Duration: | 4 Hours |
| Description: | This webinar will address the use cases of data pipeline orchestration to improve the efficiency and performance of Data Estate Management to streamline live real-time and batch processing for data cleaning, transformation, aggregation in one or different data pipelines. You define the transformations to perform on your data, and Delta Live Tables manages task orchestration, cluster management, monitoring, data quality, and error handling. Define Batch and Stream processing pipelines and integrate them together using Delta Tables. |
| Objectives: | Understanding the need of Delta Tables and its implementation as well as orchestration using Delta Live Tables in Databricks. |
| Participants’ Entry Profile: | Participants attending this course must have development experience on:   * PySpark, Batch and Stream Processing using PySpark. * Azure Databricks platform |
| Training Methodology: | The workshop will follow Synergetics methodology of   * Concept Visualization * Active Experimentation * Active Demonstration |

Setup Requirements:

|  |  |
| --- | --- |
| Hardware and Software Requirements: | Participant’s as well as Trainer’s Machine are required to have:  Hardware   * Intel Core i5 or better * 8 GB RAM * 100 GB HDD space * LAN connectivity * Good Internet connectivity and bandwidth   Software [Installed]   * Windows 10 or later * Azure Subscription * Visual Studio Code |
| Training Lab Requirements: | Whiteboard 6 feet by 4 feet (minimum)  Whiteboard markers – Red, Blue, Green, Black  Video Projector (1024 X 768 resolutions) |
| Virtual Lab Requirements:  [Optional] | Virtual labs can be provided for participants, that provides completely configured platform to work with. |

Course Contents:

Understanding Lake House and Delta Tables

* Brief understanding of Lake House
* The Delta Lake Architecture
* Brief understanding of Delta Tables in Azure Databricks
* Implementation of Delta Tables in Azure Databricks

Understanding Delta Live Tables

* What is Delta Live Table?
* Overview of concepts: Orchestration, Pipeline, Queries, Expectations
* Implementing Delta Live Tables using PySpark
* Implementing Delta Live Tables using SQL