6. Databricks - 3 Days

- Databricks overview, setup, and environment
    - Databricks architecture
    - Databricks platforms overview – Full vs Community edition
    - Databricks Control plane – Notebooks, Jobs & queries, Cluster management
    - Databricks Data plane – Databricks File System (DBFS) overview and commands, Spark cluster overview
    - Databricks workspaces – Interactive data science
    - Databricks clusters – Spark Cluster monitoring, Spark UI, Job metrics etc
    - Databricks notebooks – Working with Notebooks (Scala, Python, SQL)
    - Explore and work with Databricks sample datasets and notebooks
    - Data Visualization in Notebooks
    - Write Spark/Scala programs and Run notebooks

- Databricks notebooks with Spark/Scala/SQL
    - Work with Spark RDDs and DataFrames on Databricks notebook
    - Work with Apache Spark SQL on Databricks notebook
    - Work with Logs, CSV, JSON files in Databricks notebook (read, transform, query, visualize and write)
    - Create databases, tables and SQL Deep dive on Databricks
    - Databricks Job scheduling

- Databricks advanced features on AWS/Azure
    - Setting up Databricks Cluster on AWS/Azure
    - Advanced features – Databricks security, permissions etc
    - Databricks workflows – Jobs scheduling & workflow automation
    - Structured Streaming with Databricks overview
    - Streaming data sources and sinks on AWS/Azure

- Structured streaming in Production
- Streaming data visualizations in Real-time
- Delta Lake Overview

8. Cloud (Azure) - 3 Days

- Introduction to Azure
  - Managing Azure with Windows PowerShell
  - Overview of Azure Resource Manager
  - Azure management services
  - Lab: Managing Microsoft Azure
  - Using Azure Resource Manager features via the Azure portal.
    - Using Azure PowerShell.
  - Using Azure CLI

- Azure Storage
  - Core Storage Services
    - Virtual Machines
    - Disks
    - Snapshots
  - Types of Storage Accounts
  - Securing the Data
  - Implementing Azure Backup
  - Planning for and implementing Azure Site Recovery
  - Lab: Planning and implementing storage
  - Creating and manage Storage Accounts
  - Create and manage containers
  - Create and manage Blobs, Queues, Files and Tables

- Copying data with AzCopy
- Secure your Azure Storage Account

- Azure Networking Service
  - Virtual Network
  - Lab: Exploring Azure Networking
  - Creating Virtual Network and creating VMs inside it

- Planning and implementing Azure Relational Databases
  - Planning and deploying Azure Database for Postgres SQL
  - Implementing and managing Azure Database for Postgres SQL
  - Lab: Planning and implementing Azure SQL Database
  - Creating, securing Azure Database for PostgresSQL
  - Connecting to DB Resources
  - Best Practices for Database Services

- Planning and implementing Azure Non-Relational Databases
  - Planning and deploying Cosmos DB
  - Implementing and managing Cosmos DB
  - Lab: Planning and implementing Azure SQL Database
  - Creating, securing Azure ComsosDB
  - Creating CosmosDB Resource
  - Connecting to DB Resources
  - Best Practices for Database Services
  - Deploying Sample Application on VM connecting to a Non-Relational Database on Azure

- Azure Data Engineering Services
  - What is Azure Data Lake Storage?
  - What is Azure Data Factory?

- Building Blocks of Data Factory
- What is Azure Synapse Analytics?
- Understanding Synapse SQL Pool Architecture
- Lab: Exploring Azure Data Engineering Services
  - ADF - Copy data tool
- Synapse Analytics - Load Data using Azure Data Factory

- Azure Key Vault
  - Introduction to Azure Key Vault
  - Key Vault Terminology
  - Key Vault Security
  - Authentication, Requests and Responses
  - Lab: Planning and implementing Azure Key Vault
  - Creating Key Vault
  - Use Azure Key Vault with a virtual machine in Python