Microsoft Azure Data Factory

Module 1: Cloud Computing Concepts

- What is the "Cloud" ?
- Why cloud services
- Types of cloud models
  - Deployment Models
  - private Cloud deployment model
  - public Cloud deployment model
  - hybrid cloud deployment model
- Types of cloud services
- Infrastructure as a Service,
- Platform as a Service,
- Software as a Service
- Comparing Cloud Platforms
  - Microsoft Azure,
  - Amazon Web Services,
  - Google Cloud Platform
- characteristics of cloud computing
  - On-demand self-service
  - Broad network access
  - Multi-tenancy and resource pooling
  - Rapid elasticity and scalability
  - Measured service
- Cloud Data Warehouse Architecture
- Shared Memory architecture
- Shared Disk architecture
- Shared Nothing architecture

Module 2: Core Azure services

- Core Azure Architectural components
- Core Azure Services and Products
- Azure solutions
- Azure management tools

Module 3: Security, Privacy, Compliance

- Securing network connectivity
- Core Azure identity services
- Security tools and features

Microsoft Azure Data Factory

- Azure Governance methodologies
- Monitoring and reportings
- Privacy, compliance, and data protection standards

Module 4: Azure Pricing and Support

- Azure subscriptions
- Planning and managing costs
- Azure support options
- Azure Service Level Agreements (SLAs)
- Service Lifecycle in Azure

Module 5: Azure SQL Database

- Introduction Azure SQL Database.
- Comparing Single Database
- Managed Instance
- Creating and Using SQL Server
- Creating SQL Database Services.
- Azure SQL Database Tools.
- Migrating on premise database to SQL Azure.
- Purchasing Models
- DTU service tiers
- vCore based Model
- Serverless compute tier
- Service Tiers
  - General purpose / Standard
  - Business Critical / Premium
  - Hyperscale
- Deployment of an Azure SQL Database
- Elastic Pools.
- What is SQL elastic pools
  - Choosing the correct pool size
- Creating a New Pool
- Manage Pools
- Monitoring and Tuning Azure SQL Database
- Configure SQL Database Auditing
- Export and Import of Database
- Automated Backup
- Point in Time Restore

Microsoft Azure Data Factory

- Restore deleted databases
- Long-term backup retention
- Active Geo Replication
- Auto Failover Group

Module 6: Azure Storage Service

- Storage Service and Account
- Creating a Storage Account
- Standard and Premium Performance
- Understanding Replication
- Hot, Cold and Archive Access Tiers
- Working with Containers and Blobs
- Types of Blobs
- Block Blobs,
- Append Blobs
- Page Blobs
- Blob Metadata
- Soft Delete
- Azure Storage Explorer
- Access blobs securely
- Access Key
- Account Shared Access Token
- Service Shared Access Token
- Shared Access Policy
- Storage Service Encryption
- Azure Key Vault

Module 7: Azure Data Lake

- Introduction to Azure Data Lake
- What is Data Lake ?
- What is Azure Data Lake ?
- Data Lake Architecture
- Working with Azure Data Lake
- Provisioning Azure Data Lake.
- Explore Data Lake Analytics
- Explore Data Lake Store
- Uploading Sample File
- Using Azure Portal

Microsoft Azure Data Factory

- Using Storage Explorer
- Using Azure CLI

Module 8: Azure Data Factory

- What is Data Factory ?
- Data Factory Key Components
- Pipeline and Activity
- Linked Service o Data Set
- Integration Runtime Provision Required Azure Resources
- Create Resource Group
- Create Storage Account
- Provision SQL Server and Create Database
- Provision Data Factory

Module 9: Working with Copy Activity

- Understanding Data Factory UI
- Copy Data from Blob Storage to SQL Database
- Copy data from storage account to storage account
- Create Linked service o Create Dataset
- Create Pipeline Integration Service
- Copy Data from on-premise SQL Server to Blob Storage Working with Activities
- Understanding Lookup Activity
- Understanding for Each Activity
- Filter Activity
- Get Metadata Activity Azure
- Lift and Shift
- Provisioning Azure - SSIS Integration Runtime
- Execute SSIS Packages from Azure
- Execute SSIS Packages from SSISDB Triggers,
- Monitoring Pipeline
- Debug Pipeline
- Trigger pipeline manually
- Monitor pipeline
- Trigger pipeline on schedule

Module 10: Practical Scenarios and Use Cases

Microsoft Azure Data Factory

- ADF Introduction
- Important Concepts in ADF
- Create Azure Free Account for ADF
- Integration Runtime and Types
- Integration runtime in ADF-Azure IR
- Create Your First ADF
- Create Your First Pipeline in ADF
- Azure Storage Account Integration with ADF
- Copy multiple files from blob to blob
- Filter activity__Dynamic Copy Activity
- Get File Names from Folder Dynamically
- Deep dive into Copy Activity in ADF
- Copy Activity Behavior in ADF
- Copy Activity Performance Tuning in ADF
- Validation in ADF
- Get Count of files from folder in ADF
- Validate copied data between source and sink in ADF
- Azure SQL Database integration with ADF
- Azure SQL Databases - Introduction Relational databases
- Creating Your First Azure SQL Database
  - o 1) Deployment Models
  - o 2) Purchasing Modes
- Overwrite and Append Modes in Copy Activity
- Full Load in ADF
- Copy Data from Azure SQL Database to BLOB in ADF
- Copy multiple tables in Bulk with Lookup & ForEach in Data Factory
- Logging and Notification Azure Logic Apps
- Log Pipeline Executions to SQL Table using ADF
- Custom Email Notifications Send Error notification with logic app
- Use Foreach loop activity to copy multiple Tables- Step by Step Explanation
- Incremental Load in ADF
- Incremental Load or Delta load from SQL to Blob Storage in ADF
- Multi Table Incremental Load or Delta load from SQL to Blob Storage
- Incrementally copy new and changed files based on Last Modified Date
- Azure Key Vault integration with ADF
- Azure Key Vault, Secure secrets, keys & certificates in Azure Data
- ADF Triggers:
- Event Based Trigger in ADF
- Tumbling window trigger dependency & parameters

Microsoft Azure Data Factory

- Schedule Trigger
- Self Hosted Integration Runtime
- Copying On Premise data using Azure Self Hosted integration Runtime
- Data Migration from On premise SQL Server to cloud using ADF
- Load data from on premise sql server to Azure SQL DB
- Data Migration with polybase and Bulk insert
- Copy Data from sql server to Azure SQL DW with polybase & Bulk Insert
- Data Migration from On premise File System to cloud using ADF
- Copy Data from on-premise File System to ADLS Gen2
- ToCopying data from REST API using ADF
- Loop through REST API copy data TO ADLS Gen2-Linked Service Parameters
- AWS S3 integration with ADF
- Migrate Data from AWS S3 Buckets to ADLS Gen2
- Activities in ADF
- Switch Activity-Move and delete data
- Until Activity-Parameters & Variables
- Copy Recent Files From Blob input to Blob Output folder without LPV
- Snowflake integration with ADF
- Copy data from Snowflake to ADLS Gen2
- Copy data from ADLS Gen2 to Snowflake
- Azure CosmosDB integration with ADF
- Copy data from Azure SQLDB to CosmosDB
- Copy data from blob to cosmosDB
- Advanced Concepts in ADF
- Nested ForEach -pass parameters from Master to child pipeline
- High Availability of Self Hosted IR &Sharing IR with other ADF
- Data Flows Introduction
- Azure Data Flows Introduction
- Setup Integration Runtime for Data Flows
- Basics of SQL Joins for Azure Data Flows
- Joins in Data Flows
- Aggregations and Derive Column Transformations
- Joins in Azure DataFlows
- Advanced Join Transformations with filter and Conditional Split
- Data Flows - Data processing use case1
- Restart data processing from failure
- Remove Duplicate Rows &Store Summary Credit Stats
- Difference Between Join vs.Lookup Transformation& Merge Functionality

Microsoft Azure Data Factory

- Dimensions in Data Flows
- Slowly Changing Dimension Type1 (SCD1) with HashKey Function
- Flatten Transformation
- Rank, Dense_Rank Transformatios
- Data Flows Performance Metrics and Data Flow Parameters
- How to use pivot and unpivot Transformations
- Data Quality Checks and Logging using Data Flows
- Batch Account Integration with ADF
- Custom Activity in ADF
- Azure Functions Integration with ADF
- Azure HDInsight Integration with ADF
- Azure HDInsight with Spark Cluster
- Azure Databricks Integration with ADF
- ADF Integration with Azure Databricks
- Azure Data Lake Analytics integration with ADF

Module 11: Spark Basics

- Spark Architecture
- Spark RDD
- Spark SQL
- Spark SQL Functions
- Spark SQL Advanced

Module 12: Azure Integration Services

- Azure Functions
- Azure service bus
- Azure logic apps
- Azure Event Grid
- Azure Event Hub
- Azure API Management