

## Data Engineering - Introduction

### Storage Services - Overview

- Storage Accounts - Azure Data Lake Gen2
- Azure SQL Databases
- File Formats
- Concepts in Storage Accounts
- Overview of Transact-SQL

### ***LAB - Storage Account through Portal***

### ***LAB - Azure SQL and Databases implementation through Portal***

### ***LAB - Storage account and database creation through Azure CLI***

## Azure Datalake Storage

- What is a data lake and Azure Data Lake Storage Gen2?
- Use Azure Data Lake Storage Gen2 in data analytics workloads
- Compare Azure Data Lake Store to Azure Blob storage

### ***LAB - Create Azure Datalake Storage through Portal***

## Azure Synapse Analytic Services

### Synapse Analytics Introduction

- Why need a data warehouse
- Synapse Analytics Basics and Architecture
- Azure Synapse Compute Options
- Serverless SQL - Overview
- Dedicated SQL - Overview
- Spark Pool - Overview

### ***LAB - Create Synapse Analytics Workspace***

### ***LAB - Create Dedicated SQL Pool***

### ***LAB - Create Spark Pool and data processing with Notebook***

### External Table

- External Tables Basics

### ***LAB - External Table - CSV data***

### ***LAB - External Table - Parquet data***

### Data Loading

- Data Loading into Dedicated SQL Pool and options
- Data Loading - COPY command, Polybase
- Data Loading best practices

### ***LAB - Data loading with Polybase Option***

### ***LAB - Data Loading with COPY command***

### Designing a Database Warehouse

- Data Modelling - Dimensional Model
- Fact Table and Dimension Table
- Star Schema
- Snowflake Schema

Surrogate Key - Overview  
Slowly Changing Dimensions - Overview

**LAB - Create Fact Table**

**LAB - Create Dimension Table**

**LAB - Implement Surrogate Key**

**LAB - Implement SCD Type-1**

## **Azure Synapse SQL Architecture**

Architecture - Serverless and Dedicated SQL

Understanding Table Types

Distributions

Round-Robin Tables

Hash-Distributed Tables

Replicated Tables

Design Tables - Selection of Tables and Points for Table designing

Windowing Functions - Overview

Heap Table - Overview

Table Partitions - Overview

Indexes

**LAB - Create Round Robin Table**

**LAB - Create Hash Distribution Table**

**LAB - Understand the difference between Round Robin and Hash distribution tables**

**LAB - Implement PowerBI and connect Synapse SQL data**

## **Azure Synapse Link**

Hybrid Transactional and Analytical Processing (HTAP) - Overview

Describe Azure Synapse Link

Implement Azure Synapse Link with Azure Cosmos DB

Implement Azure Synapse Link for SQL

**LAB - Implement Azure Synapse Link for SQL**

## **Azure Event Hubs**

Azure Event Hubs - Overview

Key architecture components

Event Hubs for Apache Kafka

Support for real-time and batch processing

## **Azure Stream Analytics**

Understanding Stream Analytics

Azure Stream Analytics jobs and clusters - Overview

Understand window functions - Overview

Data Streaming Solution with Azure Stream Analytics - Overview

Use a Power BI output in Azure Stream Analytics - Overview

**LAB - Explore Azure Stream Analytics in a sample solution that aggregates streaming data events**

**LAB - Data Streaming Solution with Azure Stream Analytics**

**LAB - Process data from your event hub using Azure Stream Analytics**

## Azure Data Factory

### Azure Data Factory Introduction

- ETL - Extract Transform and Load
- Azure Data Factory Architecture Overview
- Azure Data Factory - Components Overview
- Azure Data Factory - Activities Overview
- Azure Data Factory - Triggers Overview

### Data Flow

- Mapping Data Flow Overview
- Wrangling Data Flow Overview
- Data Cleansing Overview
- Self Hosted Runtime Overview
- Overview - Conditional Split, Schema Drift and Metadata Activity

### Azure Data Factory Parametrization

- How to parameterize Linked Service in Azure Data Factory
- How to parameterize Dataset in Azure Data Factory
- How to parameterize Pipeline in Azure Data Factory

### Azure Data Factory - Azure DevOps

- ADF - Git Configuration - Overview
- ADF - Release Configuration - Overview

### Data Security and Monitoring

- Azure Data Encryption Overview
- Azure Data Factory Monitoring Overview
- Azure Data Factory Monitoring - Alerts and Metrics Overview
- Azure Data Factory Monitoring - Custom Email Notification

### Data Factory - Realtime Scenarios Labs

- LAB - Convert CSV to Parquet*
- LAB - Loading JSON to SQL*
- LAB - Setting up Self Hosted Runtime*
- LAB - Self hosted runtime - Copy Activity from Fileserver*
- LAB - Changing the time format*
- LAB - Data cleansing using conditional split*
- LAB - Control Flow Activities - ForEach*
- LAB - Bulk SQL tables copy activity*
- LAB - Implement Incremental data transfer*
- LAB - Azure Data Factory - Custom Email Notification*

## Azure Data Bricks

### Basics

- Scala - Basics

Python - Bascis

## **Azure Data Bricks**

What is Azure Databricks and Architecture

Spark Architecture Overview

Clusters in Azure Databricks

Interactive Cluster and Job Cluster

Types of Interactive Cluster - Standard Cluster and High Concurrency Cluster

Autoscaling of a cluster

Notebooks overview

DataFrame Overview

DataBricks Filesystem - Overview

Visualization in NoteBook

***LAB - Create Databricks Workspace***

***LAB - Create Databricks Cluster***

***LAB - Basic Dataloading and and Visualization***

## **ADB - Delta Lake and Catalog Table**

Managing Delta Table

Time Travel

Data Streaming into Table

***LAB - Explore Delta Table and Catalog Tables***

***LAB - Query data using Spark SQL***

## **ADB - Azure Data Lake & ADF**

Getting Data from Azure Data Lake

Handling data wiith Azure Data Lake

***LAB - ADF - Running a notebook as pipeline***

## **ADB - SQL Warehouse**

SQL Warehouse - Overview

***LAB - Implement SQL Warehouse and Data loading***

## **ADB - Data Processing - Realtime Examples**

***LAB - Handling Date Functions***

***LAB - Filtering Null Values***

***LAB - Processing CSV data***

***LAB - Processing Parquet files***

***LAB - Processing JSON files***

***LAB - Removing Duplicate Rows***

***LAB - Interacting with SQL databases***

## **Data Security and Monitoring**

### **Data Security Introduction**

Azure Key Vault Services Overview

Azure Data Encryption Overview