

**ONLINE COURSE**



**ForumDE**  
Code your career

# MASTER AZURE DATA ENGINEERING

**WITH JOB  
ASSISTANCE**

## WHAT WE OFFER?

- ✓ Azure Databricks
- ✓ Azure Data Lake
- ✓ Azure Data Factory
- ✓ Azure SQL Database
- ✓ PySpark
- ✓ Azure Key Vault
- ✓ Azure Logic Apps
- ✓ Azure DevOps



**MR. ARUN KUMAR**

Ex-TCS, Ex-Tredence  
Ex-Harman, Ex-WNS

**NEW**



**Batch Starting on**

**DEMO**



**24th August 2024**

**CLASS**



**04:00 PM - 05:30 PM**

**CLASS**



**25th August 2024**

**BEGINS**



**01:00 PM - 02:30 PM**

**ONLY AT**  
**₹ 19,999/-**

**CONTACT US:-**



**+91-78709 70617**



**<https://forumde.in/>**

# AZURE DATA ENGINEERING SYLLABUS



## ETL CONCEPT

- Lake house Concept
- Data Warehouse Concept
- Difference between Database, Data Warehouse and Lake house?
- OLAP vs OLTP systems Concept
- Fact Table and Dimension Table
- Data Modeling-Star Schema vs Snowflake Schema Concept

---

## SPARK

- Architecture of Spark
- Job, Stages and Task in Spark
- Pyspark concepts
- Spark SQL concepts
- Transformation and Action
- Lineage

**ENROLLMENT:** Contact: 9608703348, 7870970617 | Website: <https://forumde.in/>

- Lazy Evaluation
- rdd vs dataframe vs dataset
- Immutability Concept
- operations on RDD
- operations on dataframe
- Difference file formats like parquet, avro, orc and their differences
- Understanding Shuffling in Spark
- Narrow Transformation
- Wide Transformation
- Accumulator
- Broadcast variable
- Partition By
- repartition and coalesce
- partition pruning
- cache vs persist
- User Defined Functions(UDF)

---

## AZURE

- Furnishing basic information about services of Azure with perspective of Data Engineering and how it is used in industry and real time scenario.

## AZURE DATABRICKS

- Clusters
- Widgets
- Databricks CLI overview
- Service Principal
- Mount Point
- JDBC connection with Azure SQL Server
- Read and write data in Azure Databricks
- Data processing in Azure Databricks
- Introductory notebooks
- Working with Dataframes in Azure Databricks
- Working with different types of file formats like CSV, Parquet, Delta, Avro
- dbutils commands
- Databricks Runtime
- Concept of Delta Table
- Get started with Delta Lake
- Time Travel in Delta Lake
- Schema Enforcement and Schema Evolution
- MERGE statement
- Ingest data into Data Lake
- Table batch reads and writes
- Table delete, update, and merge
- Handling Bad Records
- Scheduling jobs in Databricks Notebook
- Medallion architecture

**ENROLLMENT:** Contact: 9608703348, 7870970617 | Website: <https://forumde.in/>

# AZURE DATA LAKE GEN2 (ADLSGEN2)

- How to provision ADLS Gen2
- Theoretical concept of Storage Azure Data Lake Gen2
- How do we store data in ADLS Gen2
- How to read/write data in ADLS Gen2 from databricks
- Hot, Cold, Cool and Archive Access Tiers
- Difference between ADLS Gen2 vs ADLS Gen1 vs Blob Storage
- ZRS,LRS,GRS,GZRS

# AZURE DATA FACTORY (ADF)

- Pipelines
- Activities
- Datasets
- Linked Services
- Integration Runtimes
- Auto Resolve Infrastructure Runtime
- Self Hosted Infrastructure Runtime(SHIR)
- SSIS Infrastructure Runtime
- Pipeline Run
- Scheduling and Trigger of pipeline
- Parameters
- Variables
- Orchestration of Pipeline in ADF
- Triggers
- Schedule Trigger and its System Variables
- Event Based Trigger and its System Variables
- Tumbling Widow Trigger and its System Variables
- Data Flow Pipeline
- Pipeline Level Parameters
- Global Parameters
- Creating Dynamic Pipeline creation
- How to monitor ADF data pipelines?
- ARM Template
- Copy one table from My SQL (on-premise) to Azure SQL Server.
- Copy filtered data of one table from My SQL (on-premise) to Azure SQL Server.
- Copy All tables from My SQL Server(on-premise) to Azure SQL Server
- Copy selected tables from My SQL Server(on-premise) to Azure SQL Server
- Pipeline for Incremental Load
- Copy all tables from Azure SQL Database to ADLS Gen2
- Copy selected tables from Azure SQL Database to ADLS Gen2
- Error Logging of ADF pipelines in Azure SQL Server.

# DEPLOYMENT

- Deploy Azure Databricks and Azure Data Factory pipelines using Azure DevOps.

## AZURE KEY-VAULT SERVICE

- How to securely store Keys and Password in Azure Key-Vault
- How to access Keys and Password for Azure Key-Vault
- How to connect AKV from ADF(Azure Data Factory) and read the secrets of AKV.
- How to connect AKV from ADB(Azure Databricks) and read the secrets of AKV.

## AZURE LOGIC APP

- How to send mail alert for pipeline failure and success using logic app.

## AZURE SQL SERVER

- Connect Azure SQL server from databricks
- Connect Azure SQL server from Azure Data Factory

## SOFTWARE INSTALLATIONS

- Installation of My SQL
- Installation of SSMS and connection with Azure SQL Server
- Installation of Azure Data Studio and connection with Azure SQL Server
- Installation of Azure Storage Explorer and connection with ADLS Gen2
- Installation of Notepad++

**APART FROM ABOVE COURSE DETAILS, WE WILL ALSO DELIVER BELOW: -**

- In addition to this we will do one end to end data pipeline on Azure using all the major components of Azure using Python as programming language.
- We will guide you what kind of questions can be asked in interview.
- We will help you in building your resume.
- We will take mock interview to prepare you for real time interview.

**Sharing of videos to another person is strictly prohibited.**

**Contact us For Admission or Queries: -**



**+91 78709 70617**



**+91 98018 30173**