Azure Data Engineer

Cloud Fundamentals

- Data Engineer and Role
- on-premises Vs cloud computing
- Clouddeployment models
- CloudKey characteristics
- Typesofservices in cloud computing
- Azureportal walk through
- AccessManagement

Storageaccount – ADLS

- Container
- Life cycle management
- Formats of data

Logic apps

- Workflow
- Actions

Azure Data Factory

- Author
  - Pipeline
    - Re-usable pipeline
    - Parameters, variables
  - Activities
    - Set variable
    - Append variable
    - If condition
    - switch

- ForEach
- Until Filter
- Validation
- Delete
- Get metadata
- Lookup Stored
- procedure Script
- Copydata web

o Dataset
o dataflow
- Transformations
  - Source
  - Sink
  - Select
  - Derived column
  - Conditional split
  - Filter
  - Sort
  - Aggregate
  - Join
  - Union
  - Lookup
  - Exists
  - Surrogate key
  - Window

- Monitor
  - o Pipeline runs
  - o Debug runs
  - o Trigger runs
- Manage
  - o Integration runtime
    - Azure managed
    - SHIR SSIS

- Linked service
- Trigger
- Schedule
- Event based
- Tumbling window
- Globalparameters

ADF Integrationwith GitHub

ADF use cases

ADF Ingestionframework - controltable driven

Azure databricks

- Databricks GUI
  - o Workspace
  - o Compute
    - Cluster management
      - Driver/worker nodes
  - o Catalog
  - o Workflows
- Spark architecture
- Lazy evaluation
  - o Transformations
    - Narrow transformations
    - Wider transformations
  - o Actions
- Dbutils commands
- Mounting
- Service principal
- Dataframe
  - o Column operations
  - o Row operations
  - o Transformations
  - o Joins
  - o Aggregations
  - o Window functions
  - o Set operators
  - o File read / write
  - o Spark SQL

- Schema
- User definedfunction (UDF)
- Delta lake
  - ACID
  - Versioning
- Timetravel
- Schema evaluation
- Merge

Data bricksintegration withGitHub

Genie - AIfromdatabricks

SQL Editor-AIplatform in databricks

Databricks use cases

PerformanceTuningconcepts

SCD type – 1,2and3

Artificial Intelligence

- What is AI
- Machine learning, deep learning
- Foundation model, large language model
- Generative AI
- Natural language processing
- Prompt engineering
  - Zero shot / single shot / few shot prompting
- Context engineering
- Train & Tune the model
- Azure AI Foundry
  - AI Agents
    - Knowledge base
    - Actions Tools

- Multi (connected) AIagentscreation

AI agents creation_____

SQL

- Data, Types of Data
- Database, schema, table, columns
- Constraints
- Azure SQL DB, Azure studio
- Operators
- Aggregate functions
- CTE
- Window functions
- Sub-query, co-related sub query
- Joins
- Set operators
- Functions
- Indexes
- Views

_____

Python

- Collection Data types
  - List, tuple, set, dictionary
- Operators
- If condition, loops
- Functions, Lambdafunctions

_____

Schedule forthe next training session

- Number of hours : 55To60
- Start date : 2026-01-22
- End date : 2026-03-12(Few sessions goes in parallel)
- Timing : 9 To 10 PMIST (15 minutes buffer time)
- Working days : Mondaythrough Saturday
  - Saturdays - will spend more time

Technical knowledgesharing

- Cloudfundamentals -03hours
- AzureData Factory -12hours
- ADFIngestion Framework -05hours
- AzureData Bricks -15hours
- Azuredata bricks use cases -04hours
- Logicapps, storage account -01hours
- Artificial Intelligence -04hours
- SQL (parallel sessions) -05hours
- Python(parallel sessions) -03hours

Provides

Every Day session videos ADF ingestion framework ADB 40-50 complex scenarios AI Agents creation process Data Engineering Interview Questions Data Engineering Notes