

# AZURE CLOUD - DATA ENGG &&DATA SCIENCE

## INTRODUCTION TO AZURE CLOUD

- Introduction to Azure Cloud
- Overview of Azure Account
- Azure Free Tier Account
- Creation of Azure Portal
- Overview of Billing and credits

## AZURE CORE COMPONENTS

- Resource Groups
- Subscriptions
- Resources Manager
- Tags

## AZURE CORE RESOURCES

- Identity and Access Management
- Networking
- Compute
- Databases
- Storage
- Analytics
- DevOps
- IoT
- Machine Learning

## AZURE MANAGEMENT TOOLS

- Azure Portal
- CLI- Command Line Interface
- Azure Console
- Azure PowerShell Scripts
- Azure Cloud Shell Scripting
- Azure Advisor
- Azure Resource Manager (ARM) templates
- Azure Monitor

## IDENTITIES – AZURE ACTIVE DIRECTORY(AD)

- Azure Active Directory
- Create/Manager users and Groups
- Users, Groups and Roles in Azure Active Directory
- Manage Azure AD Groups
- Manager Azure AD Users

- Create/Manage a Managed Identity for Azure Resources
- Create/Manage Administrative Units
- Role-based Access control (RBAC) authorization
- Configure Azure Policies
- Implement Conditional Access Policies
- Configure Azure Roles
- Create and assign roles

## **NETWORKING - AZURE VIRTUAL NETWORKS**

- Create and configure virtual Networks
- Azure Peering
- Private and Public IP addresses
- Implement Subnets
- Configure endpoints on subnets
- Configure private endpoints
- Configure Azure DNS
- DNS Servers resolution with VNET

## **NETWORKING - ADVANCED AZURE VIRTUAL NETWORKS**

- VNET endpoint resolution with Public IP
- Azure Network Security Groups
- Create Security Rules
- Creating Public and Private subnets
- Azure Load Balancers
- Configure Azure Application Gateways
- Configure an internal or public load balancer
- Azure Monitor for Networks
- Azure Network Watcher
- Azure VPN Gateway
- Azure ExpressRoute

## **COMPUTE - AZURE COMPUTE**

- Provision Virtual Machines
- RDP to Windows VM
- Deploy Linux VM
- SSH to Linux VM
- Configure Azure Disk Encryption
- Add data disks
- Configure high availability
- Virtual Machine Scale Sets
- Configure, validate and deploy ARM templates
- Configure a virtual hard disk(VHD) template

- Modify an Azure Resource Manager template
- Deploy from a template
- Deploy virtual machines extensions

## STORAGE - AZURE STORAGE

- Create and configure Storage accounts
- Configure Blob object replication
- Create an Azure File share
- Configure Blob Storage
- Configure Storage tiers
- Storage Account Endpoints
- Generating an Access Key
- Generate shared access signature(SAS) tokens
- Manage access Keys
- Configure access to Azure Files
- Export & Import from Azure Job
- Install and use Azure Storage Explorer
- Copy data using AZCopy
- Implement Azure Storage Replication

## SQL DATABASES - AZURE SQL DATABASES

- What is Azure Database Service?
- Benefits of Using Azure Databases
- Create an Azure Database for MySQL
- Configure the server firewall
- Use MySQL command-line tool to create a database
- Load sample data
- Query data
- Update/Restore data

## NoSQL DATABASES – AZURE Cosmos DB

- About NoSQL Databases
- Benefits of Using NoSQL Databases
- Creation of Azure Cosmos DB
- Connection to DB
- Writing Data to Cosmos DB
- Load sample data
- Query data
- Update data
- Restore data

## PAAS - AZURE APP SERVICES

- Create an App Service Plan
- Configure Scaling Settings in an App Service Plan
- Create an App Service
- Secure an App Service
- Configure Custom Domain Names
- Configure backup for an App Service
- Configure Networking Settings
- Configure Deployment Settings
- Deploy code to web app
- Configure Web App settings
- SSL/API/Connection Strings
- Scheduled Autoscaling
- Operational Metric Autoscaling

## PAAS - AZURE FUNCTIONS

- Create and Deploy Azure Functions
- Implement input and output bindings for a function
- Implement functions Triggers – Data Operation
- Implement functions Triggers – Timers
- Implement functions Triggers – Webhooks
- Implement Azure Durable functions

## MONITORING - AZURE MONITOR

- Configure and interpret Metrics
- Configure Azure Monitor Logs
- Query and analyze logs
- Setup up alerts and actions
- Configure application Insights
- Create and customize rules in Azure Sentinel
- Configure connectors in Azure Sentinel
- Evaluate alerts and incidents in Azure Sentinel

# DATA ENGG AND DATA SCIENCE ON AZURE

## INTRODUCTION TO DATA ENGINEERING

- Introduction to Data Engineering
- What we do in Data Engg?
- Roles we have in Data Engg
- Who are Data Engineers/Scientists?
- What is Data Engg Pipeline
- Different stages in Data Pipelines

## LEARNING ABOUT BIG DATA CONCEPTS

- Learning about Big Data Concepts
- Learning about Batch Data
- Learning about Streaming Data
- Storage for Big Data
- Processing Big Data
- Databases for Big Data

## LEARNING ABOUT DATA STORAGE

- Learning about Data Storage
- Learning about Data Ware Houses
- Learning about Data Lakes
- Storages for Structured Data
- Storages for Unstructured Data

## LEARNING ABOUT DATA PROCESSING

- Determine data processing solutions
- Learning about ETL technique's
- Learning about ELT technique's

## LEARNING ABOUT DATABASES

- What is ACID in Databases?
- What is BASE in Databases?
- What are SQL Databases?
- What are NoSQL Databases?
- SQL Databases vs NoSQL Databases

## LEARNING ABOUT DATA ANALYTICS

- Analytics techniques
- Descriptive analytics
- Diagnostic analytics
- Predictive analytics
- Prescriptive analytics
- Data Visualization
- Determine data processing solutions

## DESIGN NON-RELATIONAL CLOUD DATA STORES

- Designing a Data Partition in Azure
- Designing for Global Distribution in Azure
- Designing for Consistency in Cosmos DB

- Choosing the Appropriate Cosmos DB API
- Designing for Disaster Recovery and High Availability
- Refresh on Data Lake Gen 2
- Designing a Solution That Utilizes Cosmos DB, Data Lake Gen 2, or Blob Storage
- Setting Partition Keys in Azure Cosmos DB

## IMPLEMENT NON-RELATIONAL DATA STORES

- Azure Data Lake Introduction
- Cosmos DB Essentials
- Implementing Consistency in Cosmos DB
- Partitioning and Horizontal Scaling in Cosmos DB
- Selecting and Implementing API in Cosmos DB
- Implementing Security in Cosmos DB
- Intro to Using Azure Blob Storage
- Provisioning a Cosmos DB Instance in Azure
- Provisioning a Gen 2 Azure Data Lake

## DESIGN RELATIONAL CLOUD DATA STORES

- Partitioning Tables
- Designing for SQL DB and SQL DW
- Designing Data Distribution in Azure
- Designing for Ingestion with PolyBase
- Designing for Scale in SQL DB
- Designing for Scale in SQL DW
- Designing for Disaster Recovery and High Availability
- SQL Database Automated Backup Strategies
- Design a Disaster Recovery Strategy
- Design for DR and HA in Azure SQL Database
- Managing an Azure SQL Database

## IMPLEMENT RELATIONAL DATA STORES

- Introduction to Azure Synapse Analytics
- High Availability
- Disaster Recovery
- Global Distribution in Azure Synapse Analytics
- Data Distribution and Partitioning in Azure Synapse
- Introduction to PolyBase
- Ingesting Data in PolyBase

## **BIG DATA HADOOP & SPARK CLUSTERS - AZURE HD INSIGHTS**

- Deploy HDInsight clusters
- Deploy and secure multi-user HDInsight clusters
- Ingest data for batch and interactive processing
- Configure HDInsight clusters
- Manage and debug HDInsight jobs
- Implement batch solutions with Hive and Apache Pig
- Design batch ETL solutions: big data with Spark
- Operationalize Hadoop and Spark

## **BIG DATA INTEGRATION SOLUTIONS - AZURE SYNAPSE ANALYTICS**

- Understand Azure Synapse Analytics
- Deploy an Azure Synapse Analytics Data Warehouse Instance
- Load and Analyze Data with Azure Synapse Analytics Data Warehouse
- Deploy an Azure Synapse Analytics Workspaces Instance
- Load and Analyze Data with Azure Synapse Studio
- Manage Compute for Azure Synapse Analytics
- Secure Data by Implementing Data Masking
- Encrypt Data at Rest and in Motion
- Develop Batch Processing Solutions
- Batch Processing Overview

## **BIG DATA PROCESSING SOLUTIONS [SPARK]– AZURE DATABRICKS**

- Databricks Introduction
- Implementing Azure Databricks Clusters Part 1
- Implementing Azure Databricks Clusters Part 2
- Working With Notebooks in Databricks

## **DEVELOP STREAMING SOLUTIONS – AZURE STREAM ANALYTICS**

- Introduction to Azure Stream Analytics
- Azure Stream Windowing Functions
- Creating a Job in Stream Analytics
- Using Reference Data for Lookups in Stream Analytics

## **DEVELOP STREAMING SOLUTIONS– AZURE EVENT HUB**

- Introduction
- Implement a Data Streaming Solution with Azure Streaming Analytics
- Architect message brokering and serverless applications in Azure
- Create an Event Hub using the Azure CLI
- Use the Azure CLI to Create an Event Hub
- Configure applications to send or receive messages through an Event Hub

- Evaluate the performance of the deployed Event Hub using the Azure portal

## DEVELOP INTEGRATION SOLUTIONS - AZURE DATAFACTORY

- Introduction to Azure Data Factory
- Create and Schedule Triggers
- Developing Pipelines in Azure Data Factory
- Implementing Integration Runtime for Data Factory
- Developing a Pipeline in Azure Data Factory

## MONITOR DATA STORAGE

- Monitoring a Storage Account in the Azure Portal
- An Introduction to Azure Monitor
- Azure Storage Metrics in Azure Monitor
- Implement Cosmos DB Monitoring
- Implement Monitoring in Blob Storage
- Implement Monitoring in Azure Synapse Analytics
- Working with Azure Log Analytics
- Implementing Azure Monitor

## MONITOR DATA PROCESSING

- Monitoring Stream Analytics
- Monitoring Data Factory
- Monitoring Azure Databricks
- 

## DESIGN SECURITY FOR SOURCE DATA ACCESS

- Planning for Secure Endpoints
- Authenticate with Azure Active Directory
- Authenticate with Access Keys
- Authenticate with Shared Access Signatures

## MIGRATION TO AZURE DATABASES

- Environment setup
- Migration process overview
- Assess a database by using Data Migration Assistant
- Migration with downtime
- Migrate a database schema by using Data Migration Assistant



- Migrate a database with downtime by using Azure Database Migration Service

## MONITORING AND LOG ANALYTICS

- Introduction
- Design for Azure Monitor data sources
- Design for Log Analytics
- Design for Azure Workbooks and Azure Insights
- Design for Azure Data Explorer
- Monitor resources for performance efficiency