DP-700 Exam Prep

Overview

As a candidate for DP-700 exam, you should have subject matter expertise with data loading patterns, data architectures, and orchestration processes. Your responsibilities for this role include:

- Ingesting and transforming data.
- Securing and managing an analytics solution.
- Monitoring and optimizing an analytics solution.

You work closely with analytics engineers, architects, analysts, and administrators to design and deploy data engineering solutions for analytics.

You should be skilled at manipulating and transforming data by using Structured Query Language (SQL), PySpark, and Kusto Query Language (KQL).

This course will help you review the content and go through the resources that will help you prepare well for the exam.

Duration : 8 hours

Skills measured

- Implement and manage an analytics solution
- Ingest and transform data
- Monitor and optimize an analytics solution

Content Breakup :

Implement and manage an analytics solution (30–35%)

- Configure Microsoft Fabric workspace settings
 - Configure Spark workspace settings
 - Configure domain workspace settings
 - Configure OneLake workspace settings
 - Configure data workflow workspace settings

• Implement lifecycle management in Fabric

- Configure version control
- Implement database projects
- Create and configure deployment pipelines
- Configure security and governance
 - Implement workspace-level access controls
 - Implement item-level access controls
 - Implement row-level, column-level, object-level, and filelevel access controls
 - Implement dynamic data masking
 - Apply sensitivity labels to items
 - Endorse items
- Orchestrate processes
 - Choose between a pipeline and a notebook
 - Design and implement schedules and event-based triggers

 Implement orchestration patterns with notebooks and pipelines, including parameters and dynamic expressions

Ingest and transform data (30–35%)

- Design and implement loading patterns
 - Design and implement full and incremental data loads
 - Prepare data for loading into a dimensional model
 - Design and implement a loading pattern for streaming data

• Ingest and transform batch data

- Choose an appropriate data store
- Choose between dataflows, notebooks, and T-SQL for data transformation
- Create and manage shortcuts to data
- Implement mirroring
- Ingest data by using pipelines
- Transform data by using PySpark, SQL, and KQL
- Denormalize data
- Group and aggregate data
- Handle duplicate, missing, and late-arriving data
- Ingest and transform streaming data
 - Choose an appropriate streaming engine
 - Process data by using eventstreams
 - Process data by using Spark structured streaming
 - Process data by using KQL

• Create windowing functions

Monitor and optimize an analytics solution (30–35%)

- Monitor Fabric items
 - Monitor data ingestion
 - Monitor data transformation
 - Monitor semantic model refresh
 - Configure alerts

• Identify and resolve errors

- Identify and resolve pipeline errors
- Identify and resolve dataflow errors
- Identify and resolve notebook errors
- Identify and resolve eventhouse errors
- Identify and resolve eventstream errors
- Identify and resolve T-SQL errors

• Optimize performance

- Optimize a lakehouse table
- Optimize a pipeline
- Optimize a data warehouse
- Optimize eventstreams and eventhouses
- Optimize Spark performance
- Optimize query performance