



Vendor: Microsoft

Exam Code: DP-700

Exam Name: Implementing Data Engineering Solutions
Using Microsoft Fabric

Version: DEMO

QUESTION 1

Case Study 1 - Contoso, Ltd

Overview. Company Overview

Contoso, Ltd. is an online retail company that wants to modernize its analytics platform by moving to Fabric. The company plans to begin using Fabric for marketing analytics.

Overview. IT Structure

The company's IT department has a team of data analysts and a team of data engineers that use analytics systems.

The data engineers perform the ingestion, transformation, and loading of data. They prefer to use Python or SQL to transform the data.

The data analysts query data and create semantic models and reports. They are qualified to write queries in Power Query and T-SQL.

Existing Environment. Fabric

Contoso has an F64 capacity named Cap1. All Fabric users are allowed to create items.

Contoso has two workspaces named WorkspaceA and WorkspaceB that currently use Pro license mode.

Existing Environment. Source Systems

Contoso has a point of sale (POS) system named POS1 that uses an instance of SQL Server on Azure Virtual Machines in the same Microsoft Entra tenant as Fabric. The host virtual machine is on a private virtual network that has public access blocked. POS1 contains all the sales transactions that were processed on the company's website.

The company has a software as a service (SaaS) online marketing app named MAR1. MAR1 has seven entities. The entities contain data that relates to email open rates and interaction rates, as well as website interactions. The data can be exported from MAR1 by calling REST APIs. Each entity has a different endpoint.

Contoso has been using MAR1 for one year. Data from prior years is stored in Parquet files in an Amazon Simple Storage Service (Amazon S3) bucket. There are 12 files that range in size from 300 MB to 900 MB and relate to email interactions.

You need to ensure that the data analysts can access the gold layer lakehouse.

What should you do?

- A. Add the DataAnalyst group to the Viewer role for WorkspaceA.
- B. Share the lakehouse with the DataAnalysts group and grant the Build reports on the default semantic model permission.
- C. Share the lakehouse with the DataAnalysts group and grant the Read all SQL Endpoint data permission.
- D. Share the lakehouse with the DataAnalysts group and grant the Read all Apache Spark permission.

Answer: C

Explanation:

Data Analysts' Access Requirements must only have read access to the Delta tables in the gold layer and not have access to the bronze and silver layers.

The gold layer data is typically queried via SQL Endpoints. Granting the Read all SQL Endpoint data permission allows data analysts to query the data using familiar SQL-based tools while restricting access to the underlying files.

QUESTION 1

The storage team has just deployed a SAN from a new vendor. The operations team has asked for it to be monitored, but a vendor specific management pack must be downloaded and installed first.

Where should this task be completed?

- A. In the administration interface of vRealize Log Insight
- B. Under the Infrastructure menu in Cloud Assembly
- C. Under Cloud Accounts in the Administration tab in vRealize Operations Manager
- D. In the Marketplace feature of vRealize Suite Lifecycle Manager

Answer: C

Explanation:

Monitoring and managing storage devices like a SAN, including the installation of vendor-specific management packs, is typically done within the administration interface of vRealize Operations Manager. The "Cloud Accounts" section is not directly related to managing storage devices and management packs; it's more about managing cloud-related accounts and configurations for vRealize Operations Manager.

QUESTION 2

Case Study 2 - Litware, Inc

Overview

Litware, Inc. is a publishing company that has an online bookstore and several retail bookstores worldwide. Litware also manages an online advertising business for the authors it represents.

Existing Environment. Fabric Environment

Litware has a Fabric workspace named Workspace1. High concurrency is enabled for Workspace1.

The company has a data engineering team that uses Python for data processing.

Existing Environment. Data Processing

The retail bookstores send sales data at the end of each business day, while the online bookstore constantly provides logs and sales data to a central enterprise resource planning (ERP) system.

Litware implements a medallion architecture by using the following three layers: bronze, silver, and gold. The sales data is ingested from the ERP system as Parquet files that land in the Files folder in a lakehouse. Notebooks are used to transform the files in a Delta table for the bronze and silver layers. The gold layer is in a warehouse that has V-Order disabled.

Litware has image files of book covers in Azure Blob Storage. The files are loaded into the Files folder.

Existing Environment. Sales Data

Month-end sales data is processed on the first calendar day of each month. Data that is older than one month never changes.

In the source system, the sales data refreshes every six hours starting at midnight each day.

The sales data is captured in a Dataflow Gen1 dataflow. When the dataflow runs, new and historical data is captured. The dataflow captures the following fields of the source:

- Sales Date
- Author
- Price

- Units
- SKU

A table named AuthorSales stores the sales data that relates to each author. The table contains a column named AuthorEmail. Authors authenticate to a guest Fabric tenant by using their email address.

You need to implement the solution for the book reviews.

Which should you do?

- A. Create a Dataflow Gen2 dataflow.
- B. Create a shortcut.
- C. Enable external data sharing.
- D. Create a data pipeline.

Answer: B

Explanation:

The requirement specifies that Litware plans to make the book reviews available in the lakehouse without making a copy of the data. In this case, creating a shortcut in Fabric is the most appropriate solution. A shortcut is a reference to the external data, and it allows Litware to access the book reviews stored in Amazon S3 without duplicating the data into the lakehouse.

QUESTION 3

You have a Fabric workspace that contains a lakehouse named Lakehouse1. Data is ingested into Lakehouse1 as one flat table. The table contains the following columns.

Name	Description
TransactionID	Contains a unique ID for each transaction
Date	Contains the date of a transaction
ProductID	Contains a unique ID for each product
ProductColor	Contains a descriptive attribute that describes the color of each product
ProductName	Contains a unique name for each product
SalesAmount	Contains the sales amount of a transaction

You plan to load the data into a dimensional model and implement a star schema. From the original flat table, you create two tables named FactSales and DimProduct. You will track changes in DimProduct.

You need to prepare the data.

Which three columns should you include in the DimProduct table? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Date
- B. ProductName
- C. ProductColor
- D. TransactionID

- E. SalesAmount
- F. ProductID

Answer: BCF

Explanation:

In a star schema, the DimProduct table serves as a dimension table that contains descriptive attributes about products. It will provide context for the FactSales table, which contains transactional data. The following columns should be included in the DimProduct table:

ProductName: The ProductName is an important descriptive attribute of the product, which is needed for analysis and reporting in a dimensional model.

ProductColor: ProductColor is another descriptive attribute of the product. In a star schema, it makes sense to include attributes like color in the dimension table to help categorize products in the analysis.

ProductID: ProductID is the primary key for the DimProduct table, which will be used to join the FactSales table to the product dimension. It's essential for uniquely identifying each product in the model.

QUESTION 4

You have a Fabric workspace named Workspace1 that contains a notebook named Notebook1. In Workspace1, you create a new notebook named Notebook2. You need to ensure that you can attach Notebook2 to the same Apache Spark session as Notebook1. What should you do?

- A. Enable high concurrency for notebooks.
- B. Enable dynamic allocation for the Spark pool.
- C. Change the runtime version.
- D. Increase the number of executors.

Answer: A

Explanation:

To ensure that Notebook2 can attach to the same Apache Spark session as Notebook1, you need to enable high concurrency for notebooks. High concurrency allows multiple notebooks to share a Spark session, enabling them to run within the same Spark context and thus share resources like cached data, session state, and compute capabilities. This is particularly useful when you need notebooks to run in sequence or together while leveraging shared resources.

QUESTION 5

You have a Fabric workspace named Workspace1 that contains a lakehouse named Lakehouse1. Lakehouse1 contains the following tables:

- Orders
- Customer
- Employee

The Employee table contains Personally Identifiable Information (PII). A data engineer is building a workflow that requires writing data to the Customer table, however, the user does NOT have the elevated permissions required to view the contents of the Employee table. You need to ensure that the data engineer can write data to the Customer table without reading data from the Employee table.

Which three actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Share Lakehouse1 with the data engineer.
- B. Assign the data engineer the Contributor role for Workspace2.
- C. Assign the data engineer the Viewer role for Workspace2.
- D. Assign the data engineer the Contributor role for Workspace1.
- E. Migrate the Employee table from Lakehouse1 to Lakehouse2.
- F. Create a new workspace named Workspace2 that contains a new lakehouse named Lakehouse2.
- G. Assign the data engineer the Viewer role for Workspace1.

Answer: ADE

Explanation:

To meet the requirements of ensuring that the data engineer can write data to the Customer table without reading data from the Employee table (which contains Personally Identifiable Information, or PII), you can implement the following steps:

Share Lakehouse1 with the data engineer.

By sharing Lakehouse1 with the data engineer, you provide the necessary access to the data within the lakehouse. However, this access should be controlled through roles and permissions, which will allow writing to the Customer table but prevent reading from the Employee table. Assign the data engineer the Contributor role for Workspace1. Assigning the Contributor role for Workspace1 grants the data engineer the ability to perform actions such as writing to tables (e.g., the Customer table) within the workspace. This role typically allows users to modify and manage data without necessarily granting them access to view all data (e.g., PII data in the Employee table).

Migrate the Employee table from Lakehouse1 to Lakehouse2. To prevent the data engineer from accessing the Employee table (which contains PII), you can migrate the Employee table to a separate lakehouse (Lakehouse2) or workspace (Workspace2). This separation of sensitive data ensures that the data engineer's access is restricted to the Customer table in Lakehouse1, while the Employee table can be managed separately and protected under different access controls.

QUESTION 6

You have a Fabric warehouse named DW1. DW1 contains a table that stores sales data and is used by multiple sales representatives.

You plan to implement row-level security (RLS).

You need to ensure that the sales representatives can see only their respective data. Which warehouse object do you require to implement RLS?

- A. STORED PROCEDURE
- B. CONSTRAINT
- C. SCHEMA
- D. FUNCTION

Answer: D

Explanation:

To implement Row-Level Security (RLS) in a Fabric warehouse, you need to use a function that defines the security logic for filtering the rows of data based on the user's identity or role. This function can be used in conjunction with a security policy to control access to specific rows in a table.

In the case of sales representatives, the function would define the filtering criteria (e.g., based on a column such as SalesRepID or SalesRepName), ensuring that each representative can only see their respective data.

QUESTION 7

You have a Fabric deployment pipeline that uses three workspaces named Dev, Test, and Prod.

You need to deploy an eventhouse as part of the deployment process.

What should you use to add the eventhouse to the deployment process?

- A. GitHub Actions
- B. a deployment pipeline
- C. an Azure DevOps pipeline

Answer: B

Explanation:

A deployment pipeline in Fabric is designed to automate the process of deploying assets (such as reports, datasets, eventhouses, and other objects) between environments like Dev, Test, and Prod. Since you need to deploy an eventhouse as part of the deployment process, a deployment pipeline is the appropriate tool to move this asset through the different stages of your environment.

QUESTION 8

You have a Fabric workspace named Workspace1 that contains a warehouse named Warehouse1. You plan to deploy Warehouse1 to a new workspace named Workspace2. As part of the deployment process, you need to verify whether Warehouse1 contains invalid references. The solution must minimize development effort. What should you use?

- A. a database project
- B. a deployment pipeline
- C. a Python script
- D. a T-SQL script

Answer: C

Explanation:

A deployment pipeline in Fabric allows you to deploy assets like warehouses, datasets, and reports between different workspaces (such as from Workspace1 to Workspace2). One of the key features of a deployment pipeline is the ability to check for invalid references before deployment. This can help identify issues with assets, such as broken links or dependencies, ensuring the deployment is successful without introducing errors. This is the most efficient way to verify references and manage the deployment with minimal development effort.

QUESTION 9

You have a Fabric workspace that contains a Real-Time Intelligence solution and an eventhouse. Users report that from OneLake file explorer, they cannot see the data from the eventhouse. You enable OneLake availability for the eventhouse. What will be copied to OneLake?

- A. only data added to new databases that are added to the eventhouse
- B. only the existing data in the eventhouse
- C. no data
- D. both new data and existing data in the eventhouse
- E. only new data added to the eventhouse

Answer: D

Explanation:

When you enable OneLake availability for an eventhouse, both new and existing data in the eventhouse will be copied to OneLake. This feature ensures that data, whether newly ingested or already present, becomes available for access through OneLake, making it easier for users to interact with and explore the data directly from OneLake file explorer.

QUESTION 10

You have two Fabric workspaces named Workspace1 and Workspace2.

You have a Fabric deployment pipeline named deployPipeline1 that deploys items from Workspace1 to Workspace2. DeployPipeline1 contains all the items in Workspace1.

You recently modified the items in Workspaces1.

The workspaces currently contain the items shown in the following table.

Workspace	Items
Workspace1	Model1 Notebook1 Report1 Lakehouse1 Pipeline1
Workspace2	Model1 Notebook2 Report1 Lakehouse2

Items in Workspace1 that have the same name as items in Workspace2 are currently paired.

You need to ensure that the items in Workspace1 overwrite the corresponding items in Workspace2.

The solution must minimize effort.

What should you do?

- A. Delete all the items in Workspace2, and then run deployPipeline1.
- B. Rename each item in Workspace2 to have the same name as the items in Workspace1.
- C. Back up the items in Workspace2, and then run deployPipeline1.
- D. Run deployPipeline1 without modifying the items in Workspace2.

Answer: D

Explanation:

When running a deployment pipeline in Fabric, if the items in Workspace1 are paired with the corresponding items in Workspace2 (based on the same name), the deployment pipeline will automatically overwrite the existing items in Workspace2 with the modified items from Workspace1. There's no need to delete, rename, or back up items manually unless you need to keep versions. By simply running deployPipeline1, the pipeline will handle overwriting the existing items in Workspace2 based on the pairing, ensuring the latest version of the items is deployed with minimal effort.

QUESTION 11

Hotspot Question

You have a Fabric workspace named Workspace1_DEV that contains the following items:

- 10 reports
- Four notebooks
- Three lakehouses
- Two data pipelines
- Two Dataflow Gen1 dataflows
- Three Dataflow Gen2 dataflows
- Five semantic models that each has a scheduled refresh policy

You create a deployment pipeline named Pipeline1 to move items from Workspace1_DEV to a new workspace named Workspace1_TEST.

You deploy all the items from Workspace1_DEV to Workspace1_TEST.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Answer Area

Statements	Yes	No
Data from the semantic models will be deployed to the target stage.	<input type="radio"/>	<input type="radio"/>
The Dataflow Gen1 dataflows will be deployed to the target stage.	<input type="radio"/>	<input type="radio"/>
The scheduled refresh policies will be deployed to the target stage.	<input type="radio"/>	<input type="radio"/>

Answer:

Answer Area

Statements	Yes	No
Data from the semantic models will be deployed to the target stage.	<input type="radio"/>	<input checked="" type="radio"/>
The Dataflow Gen1 dataflows will be deployed to the target stage.	<input checked="" type="radio"/>	<input type="radio"/>
The scheduled refresh policies will be deployed to the target stage.	<input type="radio"/>	<input checked="" type="radio"/>

Thank You for Trying Our Product

Lead2pass Certification Exam Features:

- ★ More than **99,900** Satisfied Customers Worldwide.
- ★ Average **99.9%** Success Rate.
- ★ **Free Update** to match latest and real exam scenarios.
- ★ **Instant Download** Access! No Setup required.
- ★ Questions & Answers are downloadable in **PDF** format and **VCE** test engine format.
- ★ Multi-Platform capabilities - **Windows, Laptop, Mac, Android, iPhone, iPod, iPad**.
- ★ **100%** Guaranteed Success or **100%** Money Back Guarantee.
- ★ **Fast**, helpful support **24x7**.



View list of all certification exams: <http://www.lead2pass.com/all-products.html>



10% Discount Coupon Code: ASTR14