Vendor: Microsoft

Exam Code: DP-500

Exam Name: Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI

Version: DEMO

QUESTION 1
Case Study 1 - Litware, Inc

Overview
Litware, Inc. is a retail company that sells outdoor recreational goods and accessories. The company sells goods both online and at its stores located in six countries.

Existing environment
Azure Resources
Litware has the following Azure resources:
- An Azure Synapse Analytics workspace named synapseworkspace1
- An Azure Data Lake Storage Gen2 account named datalake1 that is associated with synapseworkspace1
- A Synapse Analytics dedicated SQL pool named SQLDW

Dedicated SQL Pool
SQLDW contains a dimensional model that contains the following table.

Name	Relevant column	Description
dbo.Customer	CustomerKey, CustomerID, CustomerEmail	The table currently contains 250,000 rows. Each row identifies a unique customer.
dbo.Product	ProductKey, ProductID, ProductName, ProductCategory, isActive	The table currently contains 2,500 rows. Each row identifies a unique product.
dbo.Date	Date, Month, Year	The table currently contains 3,653 rows. Each row identifies a unique date.
dbo.SalesTransactions	CustomerKey, ProductKey, SalesDate, SalesChannelKey, SalesAmount, QuantitySold	The table currently contains 75 million rows. Each row identifies the purchase of a single product in a sales transaction.
dbo.SalesChannel	SalesChannelKey, SalesChannel	The table currently contains two rows to identify whether a sale occurred online or in a store.

SQLDW contains the following additional tables.

Name	Relevant column	Description
MLModel	Model, Model_Name	The table contains a machine learning model named PredictPurchase that predicts the likelihood of customers purchasing a specific product based on their past purchases.
CustomersWithProductScore	CustomerID, CustomerEmail, ProductID, ProductName, Score	The Score column contains the results from calling the predictive model.

SQLDW contains a view named dbo.CustomerPurchases that creates a distinct list of values from dbo.Customer [customeriD], dbo.Customer[CustomerEmail], dbo.ProductfProductID] and dbo.Product[ProductName]. The sales data in SQLDW is updated every 30 minutes. Records in dbo.SalesTransactions are updated in SQLDW up to three days after being created. The records do NOT change after three days.

You need to configure the Sales Analytics workspace to meet the ad hoc reporting requirements.

What should you do?

A. Grant the sales managers the Build permission for the existing Power BI datasets.
B. Grant the sales managers admin access to the existing Power BI workspace.
C. Create a deployment pipeline and grant the sales managers access to the pipeline.
D. Create a PBIT file and distribute the file to the sales managers.

Answer: A
Explanation:
Template file will require data loading and will lead to separate Power BI dataset and reports. To minmize the effort, should give users build access to existing dataset.

QUESTION 2
Case Study 2 - Contoso, Ltd

Overview
Contoso, Ltd. is a company that sells enriched financial data to a variety of external customers. Contoso has a main office in Los Angeles and two branch offices in New York and Seattle.

Data Infrastructure
Contoso has a 50-TB data warehouse that uses an instance of SQL Server on Azure Virtual Machines. The data warehouse populates an Azure Synapse Analytics workspace that is accessed by the external customers. Currently, the customers can access alt the data. Contoso has one Power BI workspace named FinData that contains a single dataset. The dataset contains financial data from around the world. The workspace is used by 10 internal users and one external customer. The dataset has the following two data sources: the data warehouse and the Synapse Analytics serverless SQL pool.
Users frequently query the Synapse Analytics workspace by using Transact-SQL.

User Problems
Contoso identifies the following user issues:
- Some users indicate that the visuals in Power BI reports are slow to render when making filter selections.
- Users indicate that queries against the serverless SQL pool fail occasionally because the size of tempdb has been exceeded.
- Users indicate that the data in Power BI reports is stale. You discover that the refresh process of the Power BI model occasionally times out.

You need to recommend a solution to add new fields to the financial data Power BI dataset with data from the Microsoft SQL Server data warehouse.
What should you include in the recommendation?

A. Azure Purview
B. Site-to-Site VPN
C. an XMLA endpoint
D. the on-premises data gateway

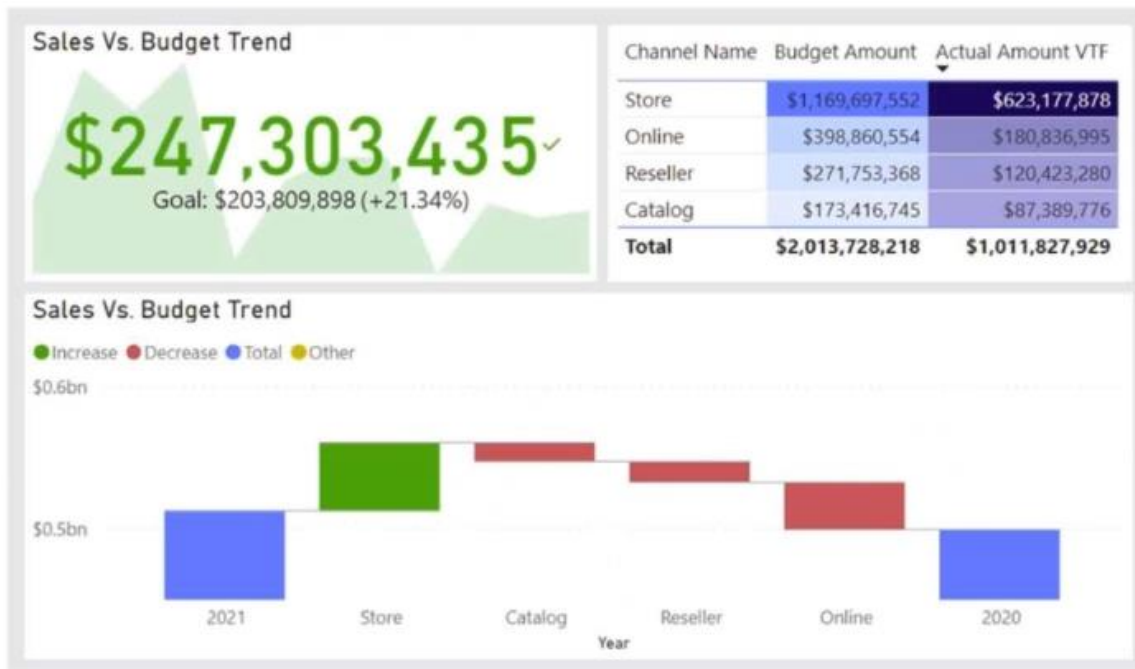**Answer:** D
Explanation:
Since the data is located in Microsoft SQL Server DWH, we need on-premises data gateway.

QUESTION 3
You are configuring a Power BI report for accessibility as shown in the following table.

You need to change the default colors of all three visuals to make the report more accessible to users who have color vision deficiency.
Which two settings should you configure in the Customize theme window? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Theme colors
B. Sentiment colors
C. Divergent colors
D. First-level elements colors

Answer: BC
Explanation:
Divergent colors - Colors used in conditional formatting to show where a data point falls in a range.
Sentiment colors - Colors used in key performance indicator (KPI) visuals and waterfall charts to indicate positive, negative, or neutral results.

QUESTION 4
You have a Power BI report that contains the visual shown in the following exhibit.

Product	Sales
Amarilla	17,747,116.06
Carretera	13,815,307.89
Montana	15,390,801.88
Paseo	33,011,143.95
Velo	18,250,059.47
VTT	20,511,921.02
Total	118,726,350.26

You need to make the visual more accessible to users who have color vision deficiency.
What should you do?

A. Change the font color of values in the Sales column to white.
B. Change the red background color to orange.
C. Add icons to represent the sales status of each product.
D. Add additional measures to the table values.

Answer: C
Explanation:
For color-blind people, the green and red combo is bad, even if you change red to orange, it won't matter much.
https://learn.microsoft.com/en-us/power-bi/create-reports/desktop-accessibility-creating-reports

QUESTION 5
You use Azure Synapse Analytics and Apache Spark notebooks to explore native visuals.
You need to use PySpark to gain access to the visual libraries.
Which Python libraries should you use?

A. Seaborn only
B. Matplotlib and Seaborn
C. Matplotlib only
D. Matplotlib and TensorFlow
E. TensorFlow only
F. Seaborn and TensorFlow

Answer: B
Explanation:
Seaborn is a data visualization library based on Matplotlib, mainly used for Statistical graphs.
Matplotlib is used for line plots, scatter plots, histograms, bar charts, pie charts and many other different plots.
https://learn.microsoft.com/en-us/azure/synapse-analytics/spark/apache-spark-data-visualization#additional-libraries

QUESTION 6
You have a Power BI workspace that contains one dataset and four reports that connect to the dataset.
The dataset uses Import storage mode and contains the following data sources:

```
- A CSV file in an Azure Storage account
- An Azure Database for PostgreSQL database
```

You plan to use deployment pipelines to promote the content from development to test to production. There will be different data source locations for each stage.
What should you include in the deployment pipeline to ensure that the appropriate data source locations are used during each stage?

A. parameter rules
B. selective deployment
C. auto-binding across pipelines
D. data source rules

Answer: A
Explanation:
In the best practices for deployment, there is written that you can use both data source rules and parameters as well, however, they recommend parameters.
https://learn.microsoft.com/en-us/power-bi/create-reports/deployment-pipelines-best-practices#use-parameters-in-your-model
Also, data source rules is supported for limited data sources where I can not find Azure storage Account so I think data source won't be applicable to our case.
https://learn.microsoft.com/en-us/power-bi/create-reports/deployment-pipelines-get-started#supported-data-sources-for-dataflow-and-dataset-rules

QUESTION 7
You have a 2-GB Power BI dataset.
You need to ensure that you can redeploy the dataset by using Tabular Editor. The solution must minimize how long it will take to apply changes to the dataset from powerbi.com.
Which two actions should you perform in powerbi.com? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Enable service principal authentication for read-only admin APIs.
B. Turn on Large dataset storage format.
C. Connect the target workspace to an Azure Data Lake Storage Gen2 account.
D. Enable XMLA read-write.

Answer: BD
Explanation:
You need XMLA read-write for Tabular Editor for metadata operations:
https://learn.microsoft.com/en-us/power-bi/enterprise/service-premium-connect-tools
And as pointed out below, you need to enable the Large dataset storage format when using XMLA endpoint write operations

QUESTION 8
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this question, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are using an Azure Synapse Analytics serverless SQL pool to query a collection of Apache Parquet files by using automatic schema inference. The files contain more than 40 million rows of UTF-8-encoded business names, survey names, and participant counts. The database is configured to use the default collation.

The queries use OPENROWSET and infer the schema shown in the following table.

name	system_type_name	max_length
businessName	varchar(8000)	8000
surveyName	varchar(8000)	8000
participants	int	4

You need to recommend changes to the queries to reduce I/O reads and tempdb usage.

Solution: You recommend using OPENROWSET WITH to explicitly define the collation for businessName and surveyName as Latin1_General_100_BIN2_UTF8.

Does this meet the goal?

A. Yes
B. No

Answer: A
Explanation:
Make sure that you are explicilty specifying some UTF-8 collation (for example Latin1_General_100_BIN2_UTF8) for all string columns in WITH clause or set some UTF-8 collation at database level.
https://learn.microsoft.com/en-us/azure/synapse-analytics/sql/query-parquet-files

QUESTION 9
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this question, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Power BI dataset named Dataset1.
In Dataset1, you currently have 50 measures that use the same time intelligence logic.

You need to reduce the number of measures, while maintaining the current functionality.

Solution: From DAX Studio, you write a query that uses grouping sets.

Does this meet the goal?

A. Yes
B. No

Answer: B
Explanation:

Using grouping sets in DAX Studio allows you to create groups of measures, but it does not reduce the number of measures in the dataset. Grouping sets are used to aggregate data across multiple dimensions in a query result, but they do not help in reducing the number of individual measures.

QUESTION 10
You need to save Power BI dataflows in an Azure Storage account.
Which two prerequisites are required to support the configuration? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. The storage account must be protected by using an Azure Firewall.
B. The connection must be created by a user that is assigned the Storage Blob Data Owner role.
C. The storage account must have hierarchical namespace enabled.
D. Dataflows must exist already for any directly connected Power BI workspaces.
E. The storage account must be created in a separate Azure region from the Power BI tenant and workspaces.

Answer: BC
Explanation:
To bring your own ADLS Gen 2 account, you must have Owner permission at the storage account layer.
The storage account must be created with the Hierarchical Namespace (HNS) enabled.
https://learn.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-azure-data-lake-storage-integration

QUESTION 11
You are optimizing a dataflow in a Power BI Premium capacity. The dataflow performs multiple joins.
You need to reduce the load time of the dataflow.
Which two actions should you perform? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Reduce the memory assigned to the dataflows.
B. Execute non-foldable operations before foldable operations.
C. Execute foldable operations before non-foldable operations.
D. Place the ingestion operations and transformation operations in a single dataflow.
E. Place the ingestion operations and transformation operations in separate dataflows.

Answer: CE
Explanation:
You should split the ETL stage into two separate dataflows.
https://learn.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-premium-features?tabs=gen2#using-the-enhanced-compute-engine

QUESTION 12
Hotspot Question

You need to configure a source control solution for Azure Synapse Analytics. The solution must meet the following requirements:

```
- Code must always be merged to the main branch before being published,
and the main branch must be used for publishing resource
- The workspace templates must be stored in the publish branch.
- A branch named dev123 will be created to support the development of a
new feature.
```

What should you do? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Answer Area

Set the collaboration branch to:

▼
main
publish
dev123
workspace_publish

Base the dev123 branch on:

▼
main
publish
dev123
workspace_publish

Answer:

Answer Area

Set the collaboration branch to:

▼
main
publish
dev123
workspace_publish

Base the dev123 branch on:

▼
main
publish
dev123
workspace_publish

Explanation:

By default, Synapse Studio generates the workspace templates and saves them into a branch called workspace_publish.
https://learn.microsoft.com/en-us/azure/synapse-analytics/cicd/source-control#configure-publishing-settings

QUESTION 13
Hotspot Question

You use Advanced Editor in Power Query Editor to edit a query that references two tables named Sales and Commission.
A sample of the data in the Sales table is shown in the following table.

OrderID	SalesPerson	Amount
101	Tom	199.99
103	Eileen	279.99
108	Enrique	333.42

A sample of the data in the Commission table is shown in the following table.

Person	Commission
Tom	0.04
Eileen	0.05

You need to merge the tables by using Power Query Editor without losing any rows in the Sales table.

How should you complete the query? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

Combined = Table. [____▼____] (Sales, "SalesPerson", Commission, [____▼____]),
"Name", JoinKind.

| Combine |
| InsertRows |
| Join |
| TransformRow |

| LeftAnti |
| LeftOuter |
| RightAnti |
| RightOuter |

Answer:

Answer Area

Combined = Table. [____▼____] (Sales, "SalesPerson", Commission, [____▼____]),
"Name", JoinKind.

| Combine |
| InsertRows |
| **Join** |
| TransformRow |

| LeftAnti |
| **LeftOuter** |
| RightAnti |
| RightOuter |

Explanation:
https://learn.microsoft.com/en-us/power-query/merge-queries-left-anti
https://learn.microsoft.com/en-us/power-query/merge-queries-left-outer

QUESTION 14
You have a deployment pipeline for a Power BI workspace. The workspace contains two datasets that use import storage mode.

A database administrator reports a drastic increase in the number of queries sent from the Power BI service to an Azure SQL database since the creation of the deployment pipeline.

An investigation into the issue identifies the following:

```
- One of the datasets is larger than 1 GB and has a fact table that
contains more than 500 million rows.
- When publishing dataset changes to development, test, or production
pipelines, a refresh is triggered against the entire dataset.
```

You need to recommend a solution to reduce the size of the queries sent to the database when the dataset changes are published to development, test, or production.

What should you recommend?

A. Turn off auto refresh when publishing the dataset changes to the Power BI service.
B. In the dataset, change the fact table from an import table to a hybrid table.
C. Enable the large dataset storage format for workspace.
D. Create a dataset parameter to reduce the fact table row count in the development and test pipelines.

Answer: B
Explanation:
By changing the fact table from an import table to a hybrid table, you can leverage the benefits of DirectQuery storage mode for the fact table. With DirectQuery, the data remains in the Azure SQL database, and Power BI sends queries directly to the database when users interact with the report. This approach can significantly reduce the amount of data transferred between Power BI and the Azure SQL database.

QUESTION 15
Your company is migrating its current, custom-built reporting solution to Power BI.

The Power BI tenant must support the following scenarios:

```
- 40 reports that will be embedded in external websites. The websites
control their own security.
- The reports will be consumed by 50 users monthly.
- Forty-five users that require access to the workspaces and apps in
the Power BI Admin portal. Ten of the users must publish and consume
datasets that are larger than 1 GB.
- Ten developers that require Text Analytics transformations and
paginated reports for datasets.
- An additional 15 users will consume the reports.
```

You need to recommend a licensing solution for the company. The solution must minimize costs.

Which two Power BI license options should you include in the recommendation? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. 70 Premium per user
B. one Premium
C. 70 Pro
D. one Embedded
E. 35 Pro
F. 35 Premium per user

Answer: BE
Explanation:
Since we need 45 users to have access to the workspaces and apps, 35 Premium Per user would not work because for the Premium Per User workspace content to be accessed we need that all users have Premium Per User. So even if we go for Premium Per User, then we need to purchase it for 70 users.
Also, since we want to publish reports to web and not embed to the app, I would not consider Embedded capacity.
Another important thing is to minimize the cost. We can use Premium capacity for publishing the reports to web, supporting Text Analytics transformations and paginated reports and having dataset more than 1GB.
For the users who will be actively busy with development we can purchase 35 pro licenses (IMO 20 pro account is enough).

QUESTION 16
You have two Power BI reports named Report1 and Report2.
Report1 connects to a shared dataset named Dataset1.
Report2 connects to a local dataset that has the same structure as Dataset1.
Report2 contains several calculated tables and parameters.
You need to prepare Report2 to use Dataset1.
Which two actions should you perform? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Remove the data source permissions.
B. Delete all the Power Query Editor objects.
C. Modify the source of each query.
D. Update all the parameter values.
E. Delete all the calculated tables.

Answer: CD
Explanation:
Modify the source of each query: You need to modify the source of each query in Report2 to use Dataset1 as the data source, instead of the local dataset. This will ensure that Report2 uses the data from Dataset1, rather than the data from the local dataset.
Calculated columns absolutely can work if the dataset has the same structure meaning same column names and data types. The calculation will work on the dataset 1 data just as it worked on dataset2 if al elements of the formula for the calculated column stay the same.

Thank You for Trying Our Product

Passleader 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.passleader.com/all-products.html

10% Discount Coupon Code: ASTR14