



## **Microsoft**

### **Exam Questions DP-500**

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

## About ExamBible

### *Your Partner of IT Exam*

## Found in 1998

ExamBible is a company specialized on providing high quality IT exam practice study materials, especially Cisco CCNA, CCDA, CCNP, CCIE, Checkpoint CCSE, CompTIA A+, Network+ certification practice exams and so on. We guarantee that the candidates will not only pass any IT exam at the first attempt but also get profound understanding about the certificates they have got. There are so many alike companies in this industry, however, ExamBible has its unique advantages that other companies could not achieve.

## Our Advances

### \* 99.9% Uptime

All examinations will be up to date.

### \* 24/7 Quality Support

We will provide service round the clock.

### \* 100% Pass Rate

Our guarantee that you will pass the exam.

### \* Unique Gurantee

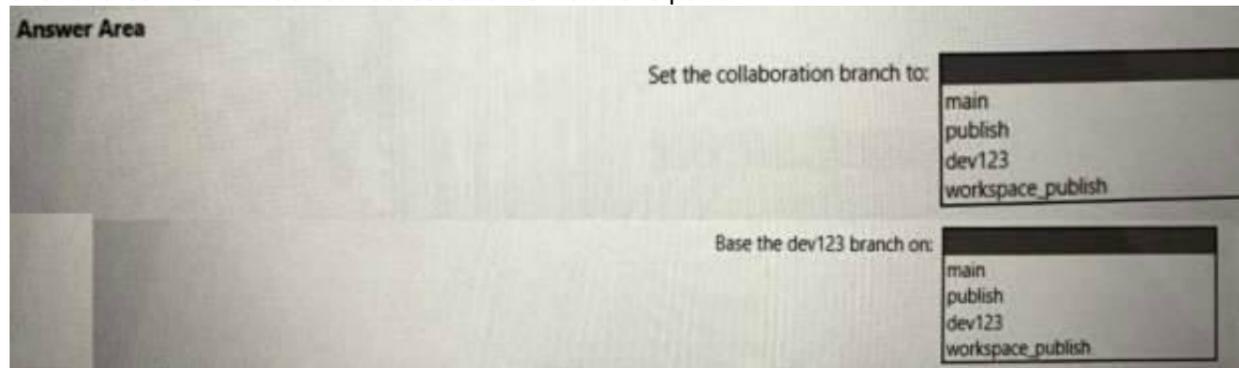
If you do not pass the exam at the first time, we will not only arrange FULL REFUND for you, but also provide you another exam of your claim, ABSOLUTELY FREE!

**NEW QUESTION 1**

- (Exam Topic 3)

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.



- A. Mastered
- B. Not Mastered

**Answer: A**

**Explanation:**

Box 1: main

Code must always be merged to the main branch before being published, and the main branch must be used for publishing resources.

Collaboration branch - Your Azure Repos collaboration branch that is used for publishing. By default, its master. Change this setting in case you want to publish resources from another branch. You can select existing branches or create new.

Each Git repository that's associated with a Synapse Studio has a collaboration branch. (main or master is the default collaboration branch).

Box 2: workspace\_publish

A branch named dev123 will be created to support the development of a new feature. The workspace templates must be stored in the publish branch.

Creating feature branches

Users can also create feature branches by clicking + New Branch in the branch dropdown.

By default, Synapse Studio generates the workspace templates and saves them into a branch called workspace\_publish. To configure a custom publish branch, add a publish\_config.json file to the root folder in the collaboration branch.

Reference: <https://docs.microsoft.com/en-us/azure/synapse-analytics/cicd/source-control>

**NEW QUESTION 2**

- (Exam Topic 3)

You are using a Python notebook in an Apache Spark pool in Azure Synapse Analytics. You need to present the data distribution statistics from a DataFrame in a tabular view. Which method should you invoke on the DataFrame?

- A. freqItems
- B. explain
- C. rollup
- D. describe

**Answer: D**

**Explanation:**

The aggregating statistic can be calculated for multiple columns at the same time with the describe function. Example:

titanic[["Age", "Fare"]].describe() Out[6]:

```
Age Fare
count 714.000000 891.000000
mean 29.699118 32.204208
std 14.526497 49.693429
min 0.420000 0.000000
25% 20.125000 7.910400
50% 28.000000 14.454200
75% 38.000000 31.000000
max 80.000000 512.329200
```

Reference: [https://pandas.pydata.org/docs/getting\\_started/intro\\_tutorials/06\\_calculate\\_statistics.html](https://pandas.pydata.org/docs/getting_started/intro_tutorials/06_calculate_statistics.html)

**NEW QUESTION 3**

- (Exam Topic 3)

You use Vertipaq Analyzer to analyze a model.

The Relationships tab contains the results shown in the following exhibit.

Relationship	Type	Max From Cardinality	Max to Cardinality	1:M Ratio	% Missing Keys	Invalid Rows	Relationships Size	Bid. Filters	MMR
'Date' [Date] ==<-1 'LocalDateTable_39c22ddb-27f3-4e6c-8a44-a3380850fcb4' [Date]	M:1	84	2,557	3044.05%	0	0	4,056		
<b>Fact</b>	<b>M:1</b>	<b>90</b>	<b>327</b>	<b>0.69%</b>	<b>22</b>		<b>184</b>		
'Fact' [BU Key] ==<-1 'BU' [BU Key]	M:1	26	164	0.34%	0	0	32		
'Fact' [Customer Key] ==<-1 'Customer' [Customer]	M:1	90	327	0.69%	21	1,804	112		
'Fact' [Product Key] ==<-1 'Product' [Product Key]	M:1	7	6	0.01%	1	6,577	8		
'Fact' [Scenario Key] ==<-1 'Scenario' [Scenario Key]	M:1	2	2	0.00%	0	0	8		
'Fact' [YearPeriod] ==<-1 'Date' [YearPeriod]	M:1	16	84	0.18%	0	0	24		
<b>Grand Total</b>	<b>M:1</b>	<b>90</b>	<b>2,557</b>	<b>3044.05%</b>	<b>27</b>		<b>4,320</b>		

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.  
NOTE: Each correct selection is worth one point.

## Answer Area

The [answer choice] table is missing records needed by the Fact table.

	▼
BU Key	
Customer	
Date	
Scenario	

There are [answer choice] blank values created by missing dimensional relationships.

	▼
22	
1,804	
6,577	
8,381	

- A. Mastered
- B. Not Mastered

**Answer:** A

### Explanation:

Box 1: Customer

There are 1804 invalid rows (records) in the Customer table. Box 2: 22

There are 22 missing keys.

Note: VertiPaq Analyzer in DAX Studio is useful in identifying referential integrity violations which slow down your DAX codes. It helps you determine which table or column needs to be optimized and improved. Reference: <https://blog.enterprisedna.co/vertipaq-analyzer-tutorial-relationships-referential-integrity/>

### NEW QUESTION 4

- (Exam Topic 3)

You have a file named File1.txt that has the following characteristics:

- A header row
- Tab delimited values
- UNIX-style line endings

You need to read File1.txt by using an Azure Synapse Analytics serverless SQL pool. Which query should you execute?

- A. SELECT\*  
 FROM OPENROWSET (  
 BULK 'file1.txt',  
 DATA\_SOURCE = 'Sql1',  
 FORMAT = 'CSV', PARSER\_VERSION = '2.0',  
 FIELDTERMINATOR = '\t',  
 ROWTERMINATOR = '0x0a',  
 FIRSTROW= 2  
 )
- B. SELECT\*  
 FROM OPENROWSET (  
 BULK 'file1.txt',  
 DATA\_SOURCE = 'Sql1',  
 FORMAT = 'CSV', PARSER\_VERSION = '2.0',  
 FIELDTERMINATOR = ',',  
 ROWTERMINATOR = '\n',  
 FIRSTROW= 2  
 )
- C. SELECT\*  
 FROM OPENROWSET (  
 BULK 'file1.txt',  
 DATA\_SOURCE = 'Sql1',  
 FORMAT = 'CSV', PARSER\_VERSION = '2.0',  
 FIELDTERMINATOR = ',',  
 ROWTERMINATOR = '0x0a',  
 FIRSTROW= 2  
 )
- D. SELECT\*  
 FROM OPENROWSET (  
 BULK 'file1.txt',  
 DATA\_SOURCE = 'Sql1',  
 FORMAT = 'CSV', PARSER\_VERSION = '2.0',  
 FIELDTERMINATOR = '\t',  
 ROWTERMINATOR = '0x0a',  
 FIRSTROW= 1  
 )

- A. Option A  
 B. Option B  
 C. Option C  
 D. Option D

**Answer:** A

**Explanation:**

Use FIELDTERMINATOR = '\t' for tab.

Use ROWTERMINATOR = '0x0A ' for UNIX-style line endings Use FIRSTROW= 2 for a header row

Note: Using Row Terminators

The row terminator can be the same character as the terminator for the last field. Generally, however, a distinct row terminator is useful. For example, to produce tabular output, terminate the last field in each row with the newline character (\n) and all other fields with the tab character (\t).

If you want to output a line feed character only (LF) as the row terminator - as is typical on Unix and Linux computers - use hexadecimal notation to specify the LF row terminator. For example:

bcp -r '0x0A' FIRSTROW

FIRSTROW = first\_row Specifies the number of the first row to load. The default is 1. This indicates the first row in the specified data file. The row numbers are determined by counting the row terminators. FIRSTROW is 1-based.

Reference:

<https://docs.microsoft.com/en-us/sql/relational-databases/import-export/specify-field-and-row-terminators-sql-se>

<https://docs.microsoft.com/en-us/sql/t-sql/functions/openrowset-transact-sql>

**NEW QUESTION 5**

- (Exam Topic 3)

You have an Azure Synapse Analytics serverless SQL pool.

You need to catalog the serverless SQL pool by using Azure Purview.

Which three actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Create a managed identity in Azure Active Directory (Azure AD).  
 B. Assign the Storage Blob Data Reader role to the Azure Purview managed service identity (MSI) for the storage account associated to the Synapse Analytics workspace.  
 C. Assign the Owner role to the Azure Purview managed service identity (MSI) for the Azure Purview resource group.  
 D. Register a data source.  
 E. Assign the Reader role to the Azure Purview managed service identity (MSI) for the Synapse Analytics workspace.

**Answer:** ABE

**Explanation:**

Authentication for enumerating serverless SQL database resources

There are three places you'll need to set authentication to allow Microsoft Purview to enumerate your serverless SQL database resources:

The Azure Synapse workspace The associated storage

The Azure Synapse serverless databases

The steps below will set permissions for all three. Azure Synapse workspace

In the Azure portal, go to the Azure Synapse workspace resource. On the left pane, select Access Control (IAM).

Select the Add button.

Set the Reader role and enter your Microsoft Purview account name, which represents its managed service identity (MSI).

Select Save to finish assigning the role

Azure Synapse Analytics serverless SQL pool catalog Purview Azure Purview managed service identity Storage account

In the Azure portal, go to the Resource group or Subscription that the storage account associated with the Azure Synapse workspace is in.

On the left pane, select Access Control (IAM). Select the Add button.

Set the Storage blob data reader role and enter your Microsoft Purview account name (which represents its MSI) in the Select box.

Select Save to finish assigning the role. Azure Synapse serverless database

Go to your Azure Synapse workspace and open the Synapse Studio. Select the Data tab on the left menu.

Select the ellipsis (...) next to one of your databases, and then start a new SQL script.

Add the Microsoft Purview account MSI (represented by the account name) on the serverless SQL databases. You do so by running the following command in your SQL script:

SQL

```
CREATE LOGIN [PurviewAccountName] FROM EXTERNAL PROVIDER;
```

Apply permissions to scan the contents of the workspace

You can set up authentication for an Azure Synapse source in either of two ways. Select your scenario below for steps to apply permissions.

Use a managed identity Use a service principal

Reference: <https://docs.microsoft.com/en-us/azure/purview/register-scan-synapse-workspace?tabs=MI>

**NEW QUESTION 6**

- (Exam Topic 3)

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:**

C: Power BI Desktop also comes with Power Query Editor. Use Power Query Editor to connect to one or many data sources, shape and transform the data to meet your needs, then load that model into Power BI Desktop.

D: Common uses for parameters

Here are some of the most common ways to use parameters. Control paginated report data

\* Filter paginated report data at the data source by writing dataset queries that contain variables.

\* Etc.

Reference: <https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-query-overview> <https://docs.microsoft.com/en-us/learn/modules/dax-power-bi-add-calculated-tables/1-introduction>

**NEW QUESTION 7**

- (Exam Topic 3)

You manage a dataset that contains the two data sources as shown in the following table.

Data source	Type of data	Privacy level
Azure SQL database	Sensitive company data	Private
Microsoft SharePoint folder	Non-sensitive company data	Private

When you attempt to refresh the dataset in powerbi.com, you receive the following error message: “[Unable to combine data] Add Columns is accessing data sources that have privacy levels which cannot be used together. Please rebuild this data combination.”

You discover that the dataset contains queries that fold data from the SharePoint folder to the Azure SQL database.

You need to resolve the error. The solution must provide the highest privacy possible.

Which privacy level should you select for each data source? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Azure SQL database:  ▼

SharePoint folder:  ▼

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Box 1: Private

This Formula.Firewall error is the result of Power Query's Data Privacy Firewall (aka the Firewall)

Note: Folding is a term that refers to converting expressions in M (such as filters, renames, joins, and so on) into operations against a raw data source (such as SQL, OData, and so on).

Box 2: Organizational

Organizational Limits the visibility of a data source to a trusted group of people. It is isolated from all Public data sources, but is visible to other Organizational data sources. A common example is a Microsoft Word document on an intranet SharePoint site with permissions enabled for a trusted group.

Reference:

<https://support.microsoft.com/en-us/office/set-privacy-levels-power-query-cc3ede4d-359e-4b28-bc72-9bee7900>

**NEW QUESTION 8**

- (Exam Topic 3)

You have a Power BI workspace named Workspace1 that contains five dataflows.

You need to configure Workspace1 to store the dataflows in an Azure Data Lake Storage Gen2 account. What should you do first?

- A. Delete the dataflow queries.
- B. From the Power BI Admin portal, enable tenant-level storage.
- C. Disable load for all dataflow queries.
- D. Change the Data source settings in the dataflow queries.

**Answer:** B

**Explanation:**

Configuring Azure connections is an optional setting with additional properties that can optionally be set:

\* Tenant Level storage, which lets you set a default, and/or

\* Workspace-level storage, which lets you specify the connection per workspace

You can optionally configure tenant-level storage if you want to use a centralized data lake only, or want this to be the default option.

Reference:

<https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-azure-data-lake-storage-integra>

**NEW QUESTION 9**

- (Exam Topic 3)

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 section, 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 Power BI Desktop, you group the measures in a display folder.

Does this meet the goal?

- A. Yes
- B. No

**Answer:** B

**Explanation:**

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

A grouping is a set of discrete values that are used to group measure fields. Reference: <https://docs.microsoft.com/en-us/power-bi/developer/visuals/capabilities>

**NEW QUESTION 10**

- (Exam Topic 3)

You are configuring Azure Synapse Analytics pools to support the Azure Active Directory groups shown in the following table.

Name	Requirement
Group1	Analyze data to create and train machine learning models in Synapse Analytics.
Group2	Execute complex queries with multiple joins against relational data. Results will be exported by using PolyBase.
Group3	Query and load data from Apache Parquet files stored in Azure Data Lake Storage Gen2. Costs must be based on the amount of data processed.

Which type of pool should each group use? To answer, drag the appropriate pool types to the groups. Each pool type may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.



- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Box 1: Apache Spark pool

An Apache Spark pool provides open-source big data compute capabilities. After you've created an Apache Spark pool in your Synapse workspace, data can be loaded, modeled, processed, and distributed for faster analytic insight.

Box 2: Dedicated SQL Pool

Dedicated SQL Pool - Data is stored in relational tables

Box 3: Serverless SQL pool

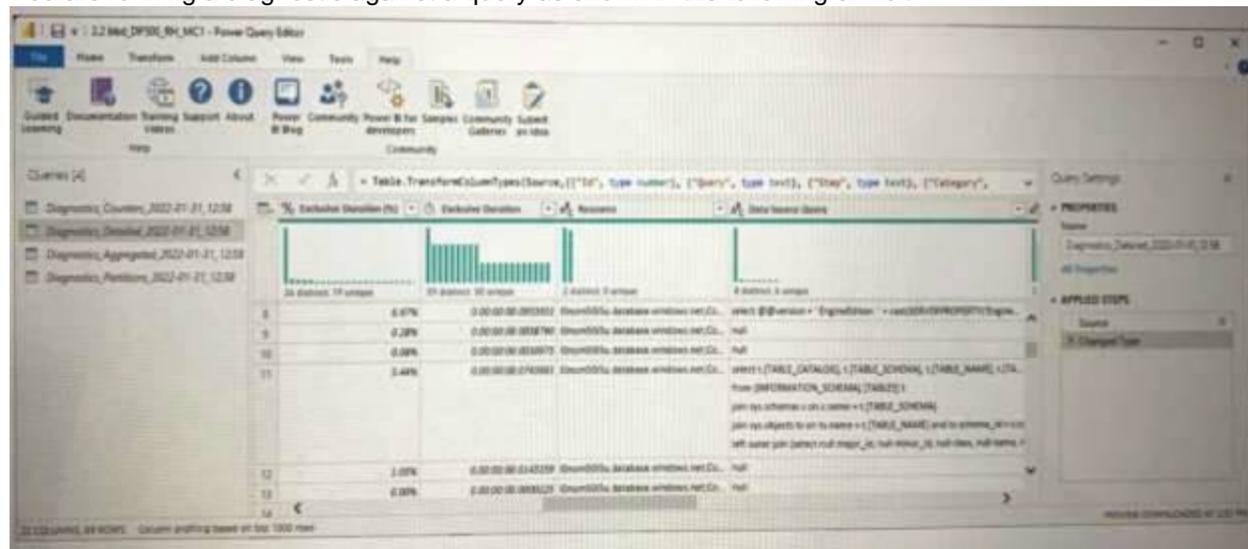
Serverless SQL pool - Cost is incurred for the data processed per query

Reference:  
<https://docs.microsoft.com/en-us/azure/synapse-analytics/quickstart-create-apache-spark-pool-portal>  
<https://www.royalcyber.com/blog/data-services/dedicated-sql-pool-vs-serverless-sql/>

**NEW QUESTION 10**

- (Exam Topic 3)

You are running a diagnostic against a query as shown in the following exhibit.



What can you identify from the diagnostics query?

- A. All the query steps are folding.
- B. Elevated permissions are being used to query records.
- C. The query is timing out.
- D. Some query steps are folding.

**Answer:** A

**Explanation:**

Understanding folding with Query Diagnostics

One of the most common reasons to use Query Diagnostics is to have a better understanding of what operations were 'pushed down' by Power Query to be performed by the back-end data source, which is also known as 'folding'. If we want to see what folded, we can look at what is the 'most specific' query, or queries, that get sent to the back-end data source. We can look at this for both ODATA and SQL.

Reference: <https://docs.microsoft.com/en-us/power-query/querydiagnosticsfolding>

**NEW QUESTION 15**

- (Exam Topic 3)

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 section, 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 Tabular Editor, you create a calculation group.

Does this meet the goal?

- A. Yes
- B. No

**Answer: B**

**Explanation:**

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

A grouping is a set of discrete values that are used to group measure fields. Reference: <https://docs.microsoft.com/en-us/power-bi/developer/visuals/capabilities>

**NEW QUESTION 16**

- (Exam Topic 3)

You have an Azure Synapse Analytics serverless SQL pool.

You need to return a list of files and the number of rows in each file.

How should you complete the Transact-SQL statement? To answer, drag the appropriate values to the targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Values	Answer Area
APPROX_COUNT_DISTINCT	<pre> SELECT   asa.filename() AS [filename] , [ ] (*) AS [rows] FROM   [ ]   BULK 'parquet/production/year=2017/month=9/*.parquet',   DATA_SOURCE = 'DataLake1',   FORMAT= 'PARQUET' ) asa GROUP BY [filename]           </pre>
COUNT_BIG	
OPENDATASOURCE	
OPENJSON	
OPENQUERY	
OPENROWSET	

- A. Mastered
- B. Not Mastered

**Answer: A**

**Explanation:**

Box 1: APPROX\_COUNT\_DISTINCT

The APPROX\_COUNT\_DISTINCT function returns the approximate number of unique non-null values in a group.

Box 2: OPENROWSET

OPENROWSET function in Synapse SQL reads the content of the file(s) from a data source. The data source is an Azure storage account and it can be explicitly referenced in the OPENROWSET function or can be dynamically inferred from URL of the files that you want to read. The OPENROWSET function can optionally contain a DATA\_SOURCE parameter to specify the data source that contains files.

The OPENROWSET function can be referenced in the FROM clause of a query as if it were a table name OPENROWSET. It supports bulk operations through a built-in BULK provider that enables data from a file to be read and returned as a rowset.

Reference: <https://docs.microsoft.com/en-us/sql/t-sql/functions/approx-count-distinct-transact-sql> <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-openrowset>

**NEW QUESTION 17**

- (Exam Topic 3)

You are implementing a reporting solution that has the following requirements:

- Reports for external customers must support 500 concurrent requests. The data for these reports is approximately 7 GB and is stored in Azure Synapse Analytics.
  - Reports for the security team use data that must have local security rules applied at the database level to restrict access. The data being reviewed is 2 GB.
- Which storage mode provides the best response time for each group of users?

- A. DirectQuery for the external customers and import for the security team.
- B. DirectQuery for the external customers and DirectQuery for the security team.
- C. Import for the external customers and DirectQuery for the security team.
- D. Import for the external customers and import for the security team.

**Answer: A**

**Explanation:**

With DirectQuery, queries are sent back to your Azure Synapse Analytics in real time as you explore the data. Real-time queries, combined with the scale of Synapse Analytics enables users to create dynamic reports in minutes against terabytes of data.

Need import for the security team for local security rules. Reference:

<https://docs.microsoft.com/en-us/power-bi/connect-data/service-azure-sql-data-warehouse-with-direct-connect>

### NEW QUESTION 20

- (Exam Topic 3)

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:

Using the compute engine to improve performance

Take the following steps to enable workloads trigger the compute engine, and always improve performance: For computed and linked entities in the same workspace:

Ensure you perform the operations that fold, such as merges, joins, conversion, and others.

For ingestion focus on getting the data into the storage as fast as possible, using filters only if they reduce the overall dataset size. It's best practice to keep your transformation logic separate from this step, and allow the engine to focus on the initial gathering of ingredients. Next, separate your transformation and business logic into a separate dataflow in the same workspace, using linked or computed entities; doing so allows for the engine to activate and accelerate your computations. In our analogy, it's like food preparation in the kitchen: food preparation is typically a separate and distinct step from gathering your raw ingredients, and a pre-requisite for putting the food in the oven. Similarly, your logic needs to be prepared separately before it can take advantage of the compute engine.

Reference:

<https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-premium-workload-configurati>

### NEW QUESTION 25

- (Exam Topic 3)

You have a sales report as shown in the following exhibit.



The sales report has the following characteristics: The measures are optimized.

The dataset uses import storage mode.

Data points, hierarchies, and fields cannot be removed or filtered from the report page. From powerbi.com, users experience slow load times when viewing the report.

You need to reduce how long it takes for the report to load without affecting the data displayed in the report. Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Change the report theme to monochromatic.
- B. Replace the single-value cards with a multi-row card.
- C. Replace the product category charts with a bar chart for sales and a hierarchy of Category and Sub Category on the axis.
- D. Replace all the filters on the Filters pane with visual slicers on the report page.

**Answer: BC**

### NEW QUESTION 26

- (Exam Topic 3)

You are planning a Power BI solution for a customer.

The customer will have 200 Power BI users. The customer identifies the following requirements:

- Ensure that all the users can create paginated reports.
- Ensure that the users can create reports containing AI visuals.
- Provide autoscaling of the CPU resources during heavy usage spikes.

You need to recommend a Power BI solution for the customer. The solution must minimize costs. What should you recommend?

- A. Power BI Premium per user
- B. a Power BI Premium per capacity
- C. Power BI Pro per user
- D. Power BI Report Server

**Answer:** A

#### Explanation:

Announcing Power BI Premium Per User general availability and autoscale preview for Gen2. Power BI Premium per user features and capabilities

\* Pixel perfect paginated reports are available for operational reporting capabilities based on SSRS technology. Users can create highly formatted reports in various formats such as PDF and PPT, which are embeddable in applications and are designed to be printed or shared.

\* Automated machine learning (AutoML) in Power BI enables business users to build ML models to predict outcomes without having to write any code.

\* Etc. Note:

Power BI empowers every business user and business analyst to get amazing insights with AI infused experiences. With Power BI Premium, we enable business analysts to not only analyze and visualize their data, but to also build an end-to-end data platform through drag and drop experiences. Everything from ingesting and transforming data at scale, to building automated machine learning models, and analyzing massive volumes of data is now possible for our millions of business analysts.

Reference:

<https://powerbi.microsoft.com/nl-be/blog/announcing-power-bi-premium-per-user-general-availability-and-auto>

### NEW QUESTION 29

- (Exam Topic 3)

You are creating a Python visual in Power BI Desktop.

You need to retrieve the value of a column named Unit Price from a DataFrame. How should you reference the Unit Price column in the Python code?

- A. pandas.DataFrame('Unit Price')
- B. dataset['Unit Price']
- C. data = [Unit Price]
- D. ('Unit Price')

**Answer:** A

#### Explanation:

You can retrieve a column in a pandas DataFrame object by using the DataFrame object name, followed by the label of the column name in brackets.

So if the DataFrame object name is dataframe1 and the column we are trying to retrieve the 'X' column, then we retrieve the column using the statement, dataframe1['X'].

Here's a simple Python script that imports pandas and uses a data frame: import pandas as pd

```
data = [['Alex',10],['Bob',12],['Clarke',13]]
```

```
df = pd.DataFrame(data,columns=['Name','Age'],dtype=float) print (df)
```

When run, this script returns: Name Age

```
0 Alex 10.0
```

```
1 Bob 12.0
```

```
2 Clarke 13.0
```

Reference:

<http://www.learningaboutelectronics.com/Articles/How-to-retrieve-a-column-from-a-pandas-dataframe-object-i>

### NEW QUESTION 32

- (Exam Topic 3)

You are using an Azure Synapse Analytics serverless SQL pool to query network traffic logs in the Apache Parquet format. A sample of the data is shown in the following table.

source		destination	
name	ip	name	ip
Network01	192.168.0.1	Internet	0.0.0.0

You need to create a Transact-SQL query that will return the source IP address.

Which function should you use in the select statement to retrieve the source IP address?

- A. JSON\_VALUE
- B. FOR.JSON
- C. CONVERT
- D. FIRST VALUE

**Answer:** A

### NEW QUESTION 35

- (Exam Topic 3)

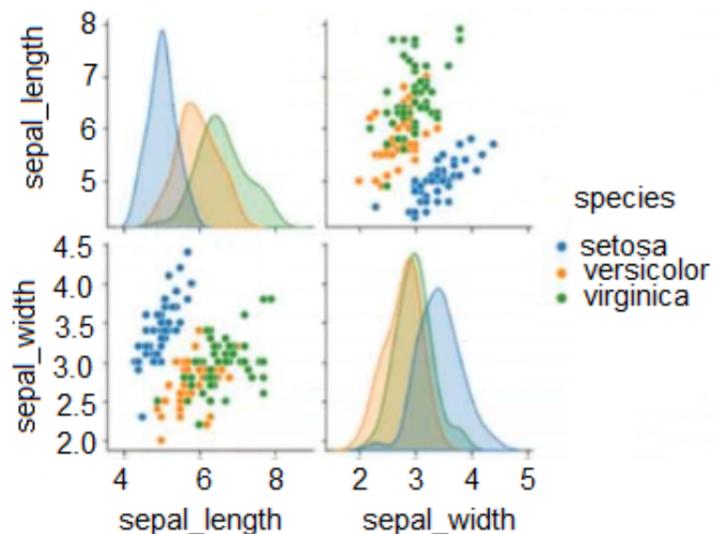
You are using an Azure Synapse notebook to create a Python visual. You run the following code cell to import a dataset named Iris.

```
iris = sns.load_dataset("iris")
iris.head()
```

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

index	sepal_length	sepal_width	species
0	5.1	3.5	setosa
2	4.9	3	setosa
145	6.7	3	virginica
156	6.3	2.5	virginica

You need to create the visual shown in the exhibit. (Click the Exhibit tab.)



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

**Answer Area**

```
sns.  (iris, hue= '', height=2.5)

plt.show()
```

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Box 1: pairplot

A pairs plot allows us to see both distribution of single variables and relationships between two variables. Pair plots are a great method to identify trends for follow-up analysis and, fortunately, are easily implemented in Python!

Example, let's plot data using pairplot:

From the picture below, we can observe the variations in each plot. The plots are in matrix format where the row name represents x axis and column name represents the y axis. The main-diagonal subplots are the univariate histograms (distributions) for each attribute.

A picture containing diagram Description automatically generated

```
In [4]: sns.pairplot(df)
```

```
Out[4]: <seaborn.axisgrid.PairGrid at 0x1a4b1b64ba8>
```



Box 2: sepal\_width

sepal\_width is displayed with a height of 2.5 (between 2.0 and 4.5).

Reference: <https://medium.com/analytics-vidhya/pairplot-visualization-16325cd725e6>

**NEW QUESTION 39**

- (Exam Topic 3)

You have a dataset that contains a table named UserPermissions. UserPermissions contains the following data.

User	Region
CONTOSO\User1	1
CONTOSO\User2	2
CONTOSO\User3	1
CONTOSO\User4	3
CONTOSO\User4	5

You plan to create a security role named User Security for the dataset. You need to filter the dataset based on the current users. What should you include in the DAX expression?

- A. [UserPermissions] - USERNAME()
- B. [UserPermissions] - USERPRINCIPALNAME()
- C. [User] = USERPRINCIPALNAME()
- D. [User] = USERNAME()
- E. [User] = USEROBJECTID()

**Answer: D**

**Explanation:**

USERNAME() returns the domain name and username from the credentials given to the system at connection time.

It should be compared to column name of User, which in DAX is expressed through [User]. Reference: <https://docs.microsoft.com/en-us/dax/username-function-dax>

**NEW QUESTION 42**

- (Exam Topic 1)

How should you configure the Power BI dataset refresh for the dbo.SalesTransactions table?

- A. an incremental refresh of Product where the ModifiedDate value is during the last three days.
- B. an incremental refresh of dbo.SalesTransactions where the SalesDate value is during the last three days.
- C. a full refresh of all the tables
- D. an incremental refresh of dbo.SalesTransactions where the SalesDate value is during the last hour.

**Answer: B**

**Explanation:**

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.

**NEW QUESTION 44**

- (Exam Topic 1)

You need to populate the CustomersWithProductScore table.

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

**Answer Area**

```

DECLARE @model
SELECT model
FROM MLModel
WHERE model_name = 'PredictPurchase'
);
INSERT INTO CustomersWithProductScore (
CustomerID
, CustomerEmail
, ProductID
, ProductName
, Score
)
SELECT d.CustomerID
, d.CustomerEmail
, d.ProductID
, d.ProductName
, p.score
FROM PREDICT(MODEL = @model, DATA =
WITH (score FLOAT) AS p;
    
```

Box 1 (next to @model): BIT, FLOAT, NVARCHAR(1000), VARBINARY(max)

Box 2 (next to DATA =): dbo.Customer, dbo.CustomerPurchases, dbo.CustomersWithProductScore, dbo.Product

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

Box 1: FLOAT

Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.

FLOT is used in the last statement of the code: WITH (score FLOAT) as p; From syntax: MODEL

The MODEL parameter is used to specify the model used for scoring or prediction. The model is specified as a variable or a literal or a scalar expression.

Box 2: dbo.CustomerWithProductScore

Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.

Only table CustomerWithProductScore has the required filed score.

From the syntax: DATA

The DATA parameter is used to specify the data used for scoring or prediction. Data is specified in the form of a table source in the query. Table source can be a table, table alias, CTE alias, view, or table-valued function.

Reference: <https://docs.microsoft.com/en-us/sql/t-sql/queries/predict-transact-sql>

**NEW QUESTION 48**

.....

## Relate Links

**100% Pass Your DP-500 Exam with ExamBible Prep Materials**

<https://www.exambible.com/DP-500-exam/>

## Contact us

We are proud of our high-quality customer service, which serves you around the clock 24/7.

Viste - <https://www.exambible.com/>