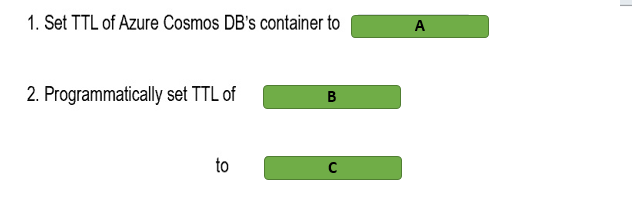
# Question153 Optimize Azure data solutions-

Your company’s quality control system uses a single container in Azure Cosmos DB to store information about quality checks (successful checks and defect detection) of finished products.  
  
Review of the data shows little value in storing successful check records long term. Management asks you to enable a data retention policy in Azure Cosmos DB, so that successful check records are automatically deleted after one hour, while defect detection records are retained indefinitely.  
  
You need to update Time-to-Live (TTL) settings in Azure Cosmos DB to enable the requested data retention policy.  
  
How should you configure TTL settings? To answer, select the appropriate options from the drop-down menus.

Choose the correct options



A)

1. 3600
2. -1

B)

1. Defect detection record
2. Successful check records

C)

1. -1
2. null

# Question154 Optimize Azure data solutions-

You are configuring a star-schema model in Azure SQL Data Warehouse for a new Operations Analytics solution.  
  
Telemetry data from factories will be loaded into the fact table, and grow by about 5 GB every month. Product details, which change infrequently, will be loaded into dimension tables, with the total size of about 50 MB.  
  
Daily shipment updates from an external website are to be fast loaded into the temporary staging table first, before being processed and then moved into the fact table.  
  
You need to choose the right data distribution method for the fact, dimension, and staging tables to optimize future data query operations.  
  
Which methods should you use? To answer, drag the appropriate data distribution method to each table type. A distribution method may be used once, more than once, or not at all.

Drag and drop the answers

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_63827/ltMS_DP-200_SelectPlace_3(6).jpeg

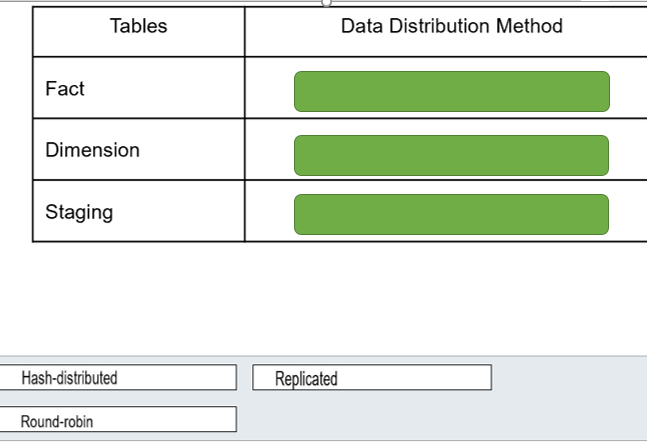
https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_63827/ltMS_DP-200_SelectPlace_3(7).jpeg

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_63827/ltMS_DP-200_SelectPlace_3(9).jpeg

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_63827/ltMS_DP-200_SelectPlace_3(6).jpeg

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_63827/ltMS_DP-200_SelectPlace_3(7).jpeg

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_63827/ltMS_DP-200_SelectPlace_3(9).jpeg



# Question155 Optimize Azure data solutions - D

Your company’s quality control system uses Azure Cosmos DB to store information about quality checks (successful checks and defect detection) of finished products.  
  
Management asks you to enable a data retention policy in Azure Cosmos DB, so that successful check entries are automatically deleted after one hour, while defect detection records are retained indefinitely.  
  
You need to update the Time-To-Live (TTL) settings in Azure Cosmos DB to enable this data retention policy.  
  
What should you do?

Choose the correct answer

Set the TTL settings of the Azure Cosmos DB container to -1, and programmatically set the TTL of defect detection items to null.

Set the TTL settings of the Azure Cosmos DB container to -1, and programmatically set the TTL of successful check items to null.

Set the TTL settings of the Azure Cosmos DB container to 3600, and programmatically set the TTL of defect detection items to -1.

Set the TTL settings of the Azure Cosmos DB container to 3600, and programmatically set the TTL of successful check items to -1.

# Question156 Optimize Azure data solutions - D

You manage an Azure SQL Database service that uses a single database. A particular query is taking a long time to run.  
  
You need to determine what about the query is causing it to take so long.  
  
What should you do?

Choose the correct answer

Run SQL Profiler.

Display the actual execution plan.

Drop all clustered indexes.

Display the estimated execution plan.

# Question157 Optimize Azure data solutions – Not complete question

You manage an Azure SQL Database. You run the query and display the execution plan shown in the exhibit. There are 10 rows in the Department table.  
  
You need to determine whether you can improve the performance of the query.  
  
What should you conclude?

Choose the correct answer

You need to remove the clustered index from the Department table.

Performance is optimal.

You need to add a non-clustered index on the Name column of the Department table.

You need to remove the clustered index from the Employee table.

# Question158 Optimize Azure data solutions-

You manage an Azure SQL Data Warehouse. You have three tables:  
  
DimProduct: Small table that changes infrequently  
DimCustomer: Small table that changes infrequently  
FactSales: Large table that changes frequently  
  
All three tables use round-robin distribution. Queries join all three tables.   
  
You need to optimize these tables for performance.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You should use a replicated table for DimProduct. |  |  |
| You should use a replicated table for DimCustomer. |  |  |
| You should use a replicated table for FactSales. |  |  |

# Question159 Optimize Azure data solutions-

You manage an Azure SQL Data Warehouse. You have three tables:  
  
DimExam: 2000 rows, 200 megabytes (MB), changes infrequently  
DimCandidate: 400 rows, 40 MB, changes infrequently  
FactScores: 100,000,000 rows, 500 gigabytes (GB), changes frequently  
  
All three tables use round-robin distribution. Queries join all three tables.   
  
You need to optimize these tables for performance.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You should use hash distribution for DimExam. |  |  |
| You should use hash distribution for DimCandidate. |  |  |
| You should use hash distribution for FactScores. |  |  |

# Question160 Optimize Azure data solutions-

You manage an Azure SQL Data Warehouse. You have four tables:  
  
DimProduct: 2000 rows, 200 megabytes (MB), changes infrequently  
DimCustomer: 400 rows, 40 megabytes (MB), changes infrequently  
FactSales: 100,000,000 rows, 500 gigabytes (GB), changes frequently  
FactOrders: 100,000,000 rows, 500 gigabytes (GB), changes frequently  
  
All three tables use hash distribution. Queries join the DimProduct, DimCustomer, and FactSales tables. The FactOrders table contains all the data it needs.   
  
You need to optimize these tables for performance.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You should use a replicated table for DimProduct. |  |  |
| You should use a replicated table for DimCustomer. |  |  |
| You should use a replicated table for FactSales. |  |  |
| You should use round-robin distribution for FactOrders. |  |  |

# Question161 Optimize Azure data solutions-

You manage an Azure Streaming Analytics job.  
  
You need to administer the job so that it uses optimal performance.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You should start with two Streaming Units (SUs) for queries that do not use PARTITION BY. |  |  |
| You should keep the Streaming Unit (SU) metric below 80 percent. |  |  |
| You should allocate more SUs than you need. |  |  |

# Question162 Optimize Azure data solutions - D

You manage an Azure Data Lake Storage Gen2 account. Your source data is stored on file servers and SQL Server on-premises. You anticipate that it will take a long time to copy the data from your company to Data Lake over the public internet.  
  
You need to ensure optimal performance when copying the source data to Azure.  
  
What should you do?

Choose the correct answer

Create an Azure Data Lake Analytics account.

Create an Azure Data Factory account.

Install Active Directory (AD) Connect on-premises.

Use ExpressRoute.

# Question163 Optimize Azure data solutions-

You are a data engineer for your company. You have several blob storage accounts that you need to manage.  
  
You need to match the access tier with the blob usage.  
  
Which access tier should you use for each scenario? To answer, drag the appropriate access tier to each blob usage. An access tier may be used once, more than once, or not at all.

Drag and drop the answers

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_64227/Archive.png

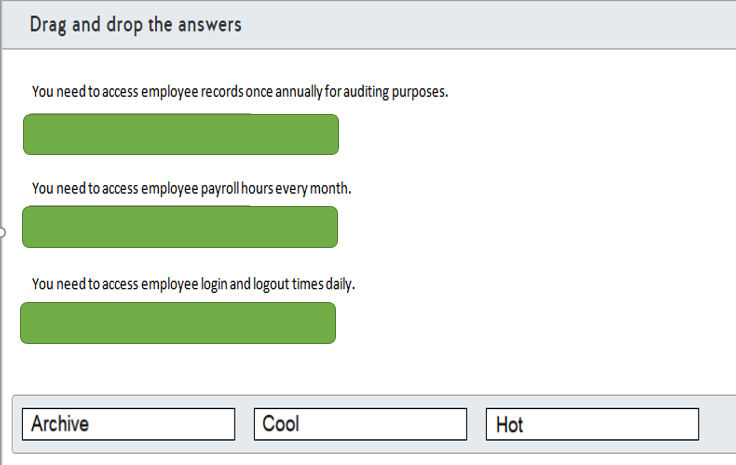
https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_64227/Cool.png

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_64227/Hot.png

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_64227/Archive.png

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_64227/Cool.png

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_64227/Hot.png



# Question164 Optimize Azure data solutions-

Your company has an application for technical writers to share, review and publish exam questions. The application allows live reviews for questions and provides reports about question performance after the questions are published. All questions and reviews and stored in Azure SQL Database.   
  
Writers are able to report question metrics in an exam results table with millions of rows. Reports include the percentage of correct answers, average question rating and answer count per question.  
  
After a question is published, live reviews are disabled and previous reviews are archived in a separate database.  
  
Writers report the following performance issues with the application:

* Some delays are occurring during the live reviews.
* Reports are taking too long to be generated.

You need to resolve the performance issues.  
  
Which two features should you use? Each correct answer presents a complete solution.

Choose the correct answers

Heap table

Memory-optimized table

Nonclustered index

Columnstore index

Partitioned view

# Question165 Optimize Azure data solutions-

Your company manages an ecommerce platform for a large retailer. The platform is composed of thousands of web servers running on Azure virtual machines (VMs). The platform generates approximately three gigabytes (GB) of log data in a day.  
  
The log files for each server are stored in Azure Data Lake Storage Gen2 inside a single folder. This data is processed and analyzed on an HDInsight cluster.  
  
You need to provide some performance improvements while minimizing cost.  
  
Which two changes should you implement? Each correct answer presents a complete solution.

Choose the correct answers

Combine the daily log files into one file.

Use a cool tier for Azure Data Lake Storage Gen2.

Increase the number of worker nodes.

Separate the log files into a daily generated folder.

# Question166 Optimize Azure data solutions-

Your team manages an image recognition process for a security firm. The solution runs on an Azure virtual machine (VM) at a daily frequency.  
  
Images are uploaded to an Azure Blob Storage container. The VM processes the results using the data uploaded from the previous day and stores the results in the same blob storage.  
  
You must meet the following data retention requirements:

* Daily results must be accessible for one week.
* Data for the current month must be available but is rarely used.
* Current year data must be stored for auditing purposes.
* Audit data must be requested at least one day before being accessed.

You need to use a lifecycle policy that minimizes cost.  
  
How should you configure the lifecycle policy? To answer, drag a JSON segment to the most appropriate location. A JSON segment may be used once, more than once, or not at all.

Drag and drop the answers

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036a.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036b.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036d.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036f.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036g.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036a.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036b.gif

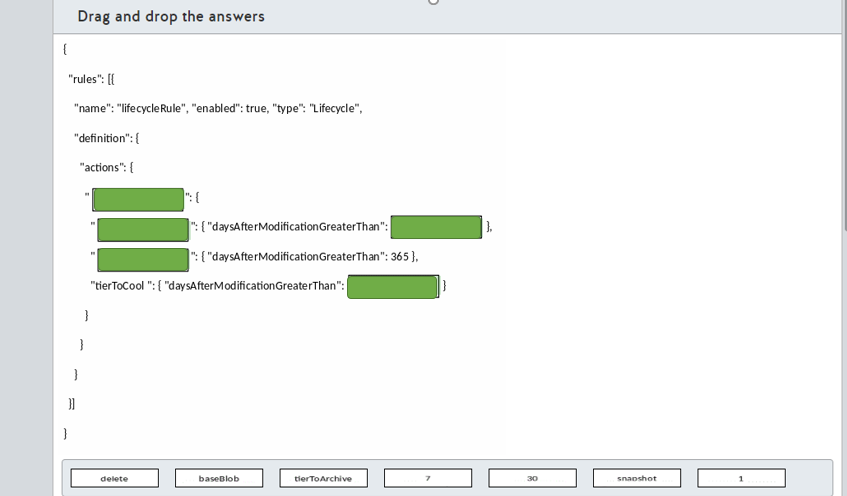
https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036d.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036f.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036g.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036c.gif

https://pts.measureup.com/web/instances/MUP/assets/images/DP-200/DP-200_65259/gsDP-200_036e.gif



# Question167 Optimize Azure data solutions-

You manage a real-time data analysis solution hosted in Azure.   
  
The solution uses Azure Event Hubs to ingest data and analyze the data in real-time with an Azure Stream Analytics job.  
  
The Stream Analytics job uses 18 Streaming Units (SU). Stream Analytics job metrics and SU percentage utilization metrics have been 90% in average over the last month.  
  
You need to optimize the Azure Stream Analytics job performance.  
  
Which two actions should you perform? Each correct answer presents part of the solution.

Choose the correct answers

Configure event ordering.

Decrease the SU count for the job.

Partition data for query parallelization.

Increase the SU count for the job.

Use Stream Analytics JavaScript user-defined functions.

# Question168 Optimize Azure data solutions-

You need to load data into an Azure Synapse Analytics SQL pool. The data is stored in a general purpose v2 Azure storage account as text files.  
  
You need to use the fastest data loading option.  
  
Which two options can you use to meet your goal? Each correct answer presents a complete solution.

Choose the correct answers

Use SQL Server Integration Services (SSIS).

Use a Copy Activity in Azure Data Factory.

Use the Bulk Copy Program utility (BCP).

Write and run PolyBase T-SQL commands.

Use a Copy Activity in Azure Data Factory with the bulk insert option.

Use the SQL BulkCopy API.

# Question169 Optimize Azure data solutions – Not complete question

You have an Azure Synapse Analytics SQL pool.  
  
The output of the EXPLAIN command is shown in the exhibit.  
  
You need to optimize the performance of a query while minimizing cost.  
  
What should you do?

Choose the correct answer

Update the statistics on the FactInternetSales table.

Upgrade the performance level.

Re-create the FactInternetSales table using hash-distribution on the ProductKey column .

Re-create the FactInternetSales table using round-robin distribution.

Re-create the FactInternetSales table using page compression.