



DP-203^{Q&As}

Data Engineering on Microsoft Azure

Pass Microsoft DP-203 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.geekcert.com/dp-203.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers





QUESTION 1

You are designing database for an Azure Synapse Analytics dedicated SQL pool to support workloads for detecting ecommerce transaction fraud.

Data will be combined from multiple ecommerce sites and can include sensitive financial information such as credit card numbers.

You need to recommend a solution that meets the following requirements:

Users must be able to identify potentially fraudulent transactions. Users must be able to use credit cards as a potential feature in models. Users must NOT be able to access the actual credit card numbers.

What should you include in the recommendation?

- A. Transparent Data Encryption (TDE)
- B. row-level security (RLS)
- C. column-level encryption
- D. Azure Active Directory (Azure AD) pass-through authentication

Correct Answer: C

Use Always Encrypted to secure the required columns. You can configure Always Encrypted for individual database columns containing your sensitive data. Always Encrypted is a feature designed to protect sensitive data, such as credit card numbers or national identification numbers (for example, U.S. social security numbers), stored in Azure SQL Database or SQL Server databases.

Reference: <https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/always-encrypted-database-engine>

QUESTION 2

You have an Azure data factory named ADF1.

You currently publish all pipeline authoring changes directly to ADF1.

You need to implement version control for the changes made to pipeline artifacts. The solution must ensure that you can apply version control to the resources currently defined in the Azure Data Factory Studio for ADF1.

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

NOTE: Each correct selection is worth one point.

- A. From the Azure Data Factory Studio, run Publish All.
- B. Create an Azure Data Factory trigger.
- C. Create a Git repository.
- D. Create a GitHub action.



E. From the Azure Data Factory Studio, select Set up code repository.

F. From the Azure Data Factory Studio, select Publish.

Correct Answer: CE

Source control in Azure Data Factory

C: By default, the Azure Data Factory user interface experience (UX) authors directly against the data factory service. This experience has the following limitations:

The Data Factory service doesn't include a repository for storing the JSON entities for your changes. The only way to save changes is via the Publish All button and all changes are published directly to the data factory service.

The Data Factory service isn't optimized for collaboration and version control.

The Azure Resource Manager template required to deploy Data Factory itself is not included.

E: To provide a better authoring experience, Azure Data Factory allows you to configure a Git repository with either Azure Repos or GitHub. Git is a version control system that allows for easier change tracking and collaboration. Connect to a Git repository There are four different ways to connect a Git repository to your data factory for both Azure Repos and GitHub. After you connect to a Git repository, you can view and manage your configuration in the management hub under Git configuration

in the Source control section Configuration method 1: Home page

1.

In the Azure Data Factory home page, select Set up code repository at the top.

2.

Etc.

Reference: <https://docs.microsoft.com/en-us/azure/data-factory/source-control>

QUESTION 3

You have an Azure data factory named ADF1 that contains a pipeline named Pipeline1.

Pipeline1 must execute every 30 minutes with a 15-minute offset.

You need to create a trigger for Pipeline1. The trigger must meet the following requirements:

1.

Backfill data from the beginning of the day to the current time.

2.

If Pipeline1 fails, ensure that the pipeline can re-execute within the same 30-minute period.

3.

Ensure that only one concurrent pipeline execution can occur.



4.

Minimize development and configuration effort. Which type of trigger should you create?

- A. schedule
- B. event-based
- C. manual
- D. tumbling window

Correct Answer: D

Tumbling window triggers are a type of trigger that fires at a periodic time interval from a specified start time, while retaining state. Tumbling windows are a series of fixed-sized, non-overlapping, and contiguous time intervals. A tumbling window trigger has a one-to-one relationship with a pipeline and can only reference a singular pipeline. Tumbling window trigger is a more heavy weight alternative for schedule trigger offering a suite of features for complex scenarios.

The following table provides a comparison of the tumbling window trigger and schedule trigger:

*

Backfill scenarios Tumbling: Supported Scheduled: Not supported

*

Retry capability

Tumbling: Supported Failed pipeline runs have a default retry policy of 0, or a policy that's specified by the user in the trigger definition.

Scheduled: Not supported

* Etc.

Reference: <https://learn.microsoft.com/en-us/azure/data-factory/how-to-create-tumbling-window-trigger>
<https://learn.microsoft.com/en-us/azure/data-factory/concepts-pipeline-execution-triggers>

QUESTION 4

DRAG DROP

You have an Azure subscription that contains an Azure Synapse Analytics workspace named workspace1. Workspace1 connects to an Azure DevOps repository named repo1. Repo1 contains a collaboration branch named main and a development branch named branch1. Branch1 contains an Azure Synapse pipeline named pipeline1.

In workspace1, you complete testing of pipeline1.

You need to schedule pipeline1 to run daily at 6 AM.

Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.

Select and Place:

Actions

Answer Area

- Create a new branch in Repo1.
- Merge the changes from branch1 into main.
- Associate the schedule trigger with pipeline1.
- Switch to Synapse live mode.
- Create a schedule trigger.
- Publish the contents of main.

Correct Answer:

Actions

Answer Area

- | | |
|-------------------------------|--|
| Create a new branch in Repo1. | Create a schedule trigger. |
| | Associate the schedule trigger with pipeline1. |
| | Merge the changes from branch1 into main. |
| Switch to Synapse live mode. | Publish the contents of main. |
| | |
| | |

QUESTION 5



You are designing a streaming data solution that will ingest variable volumes of data. You need to ensure that you can change the partition count after creation.

Which service should you use to ingest the data?

- A. Azure Event Hubs Dedicated
- B. Azure Stream Analytics
- C. Azure Data Factory
- D. Azure Synapse Analytics

Correct Answer: A

You can't change the partition count for an event hub after its creation except for the event hub in a dedicated cluster.

Reference: <https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-features>

[DP-203 VCE Dumps](#)

[DP-203 Practice Test](#)

[DP-203 Study Guide](#)