

Knowledge check

Total points 6

1. Question 1

Much of the functionality of Azure Data Factory appears in Azure Synapse Analytics as a feature called pipelines. You can use it to integrate data pipelines between which of the following?

Select all options that apply.

1 / 1 point

☒ SQL Serverless

Correct

Pipelines enable you to integrate data pipelines between SQL Pools, Spark Pools, and SQL Serverless.

☒ SQL Pools

Correct

Pipelines enable you to integrate data pipelines between SQL Pools, Spark Pools, and SQL Serverless.

☐ Apache Hive

☒ Spark Pools

Correct

Pipelines enable you to integrate data pipelines between SQL Pools, Spark Pools, and SQL Serverless.

2. Question 2

Which of the following provides a cloud-based data integration service that orchestrates the movement and transformation of data between various data stores and compute resources?

1 / 1 point

☒ Azure Data Factory

☐ Azure Databricks

☐ Azure HDInsight

☐ Azure SQL Database

Correct

Azure Data Factory is the platform that solves such data scenarios. It is the cloud-based ETL and data integration service that allows you to create data-driven workflows for orchestrating data movement and transforming data at scale.

3. Question 3

ADF has native functionality to ingest and transform data, sometimes it will instruct another service, such as Databricks, to perform the actual work required on its behalf. Which of the following terms best describes this process?

1 / 1 point



Orchestration



Integration



Transformation

Correct

ADF orchestrates the execution of a query, and then provides the pipelines to move the data onto the next step or destination.

4. Question 4

Which of the following terms describes analyzing past data patterns and trends by looking at historical data and customer insights?

1 / 1 point



Descriptive Analytics



Predictive Analytics



Prescriptive Analytics

Correct

Descriptive Analytics tells you what happened in the past.

5. Question 5

Microsoft Azure provides a variety of data platform services that enables you to perform different types of analytics. Predictive analytics can be implemented through which of the following features?

Select all options that apply

0.75 / 1 point



Machine Learning Services

Correct

Machine Learning Services enables you to perform different types of analytics including predictive analytics.



HDInsight



Azure Databricks

Correct

Azure Databricks enables you to perform different types of analytics including predictive analytics.

- ☐ Azure Data Lake Storage Gen2

You didn't select all the correct answers

6. Question 6

Data integration includes extraction, transformation, and loading of data. It is commonly referred to as Extract-Transform-Load or ETL.

At which stage in the ETL process is the splitting, combining, deriving, adding, and removing data carried out?

1 / 1 point

- ☒ Transform

Correct

Data transformation operations can include splitting, combining, deriving, adding, removing, or pivoting columns.

- ☐ Extract
- ☐ Load

Knowledge check

Total points 6

1. Question 1

You are creating a new Azure Data Factory instance. The instance name must be unique within which of the following?

1 / 1 point



Globally within Azure



The Resource Group



The Azure Subscription



The region

Correct

The name of the Azure data factory instance must be globally unique within Azure.

2. Question 2

How would you define an Azure factory dataset?

1 / 1 point



A Dataset does not exist in Azure Data factory



A Dataset is a table that holds a copy of the data for the destination.



A dataset is a named view that points to, or references, the data.



A dataset is a table that holds a copy of the data from the source.

Correct

A dataset is a named view of data that simply points to, or references, the data you want to use in your activities as inputs and outputs.

3. Question 3

How would you define an Azure Data Factory pipeline?

1 / 1 point



A pipeline contains the transformation logic, or the analysis commands, of the Azure Data Factory's work.



A Pipeline is a named view of data.



A pipeline is a logical grouping of activities that together perform a task.

Correct

The activities in a pipeline define actions to be performed on your data.

4. Question 4

When working with Azure Data factory, how many linked services are required to copy data from Blob storage to a SQL database?

1 / 1 point

☐ 0

☒ 2

☐ 3

☐ 1

Correct

To copy data from Blob storage to an Azure SQL Database, you must create two linked services; Azure Storage and Azure SQL Database.

5. Question 5

What are the three categories of activities within Azure Data Factory that define the actions to be performed on the data?

Select all options that apply.

1 / 1 point

☐ Data movement

☒ Control

Correct

You can use the control flow to orchestrate pipeline activities, including chaining activities in a sequence, branching, defining parameters at the pipeline level, and passing arguments while invoking the pipeline on-demand or from a trigger.

☒ Data transformation

Correct

Data transformation activities can be performed natively within the authoring tool of Azure Data Factory using the Mapping Data Flow. Alternatively, you can call a compute resource to change or enhance data through transformation or perform analysis of the data.

☐ Linked Service

6. Question 6

When graphically authoring ADF solutions, you can use the control flow within the design to orchestrate which of the following pipeline activities?

Select all options that apply.

1 / 1 point



Execute Pipeline Activity

Correct

Execute Pipeline activity allows an Azure Data Factory pipeline to invoke another pipeline.



Parameters Activity



WebActivity

Correct

Web Activity can be used to call a custom REST endpoint from an Azure Data Factory pipeline. You can pass datasets and linked services to be consumed and accessed by the activity.



ForEach Activity

Correct

ForEach Activity defines a repeating control flow in your pipeline. This activity is used to iterate over a collection and execute specified activities in a loop.

Knowledge check

Total points 5

1. Question 1

Which of the following methods provides the ability to build code-free data ingestion pipelines with zero transformation during the extraction of the data?

0 / 1 point

☐ Azure Machine Learning Studio

☒ Azure Databricks

☐ Azure SQL Data Warehouse

☐ Azure Copy Activity

Incorrect

Azure Data Factory can call on compute resources to process data by a data platform service that may be better suited to the job. Azure Databricks is more suitable for use with Notebook, Jar and Python files.

2. Question 2

Connectors are Azure Data Factory objects that enable your Linked Services and Datasets to connect to a wide variety of data sources and sinks. Which of the following are supported file formats?

Select all options that apply.

1 / 1 point

☒ ORC format

Correct

The ORC format is a supported file format.

☒ Parquet format

Correct

The Parquet format is a supported file format.

☐ YAML Format

☒ Binary format

Correct

The Binary format is a supported file format.

3. Question 3

You want to ingest data from a SQL Server database hosted on an on-premises Windows Server. What integration runtime is required for Azure Data Factory to ingest data from the on-premises server?

1 / 1 point



Self-Hosted Integration Runtime

Correct

A self-hosted integration runtime can run copy activities between a cloud data store and a data store in a private network. It also can dispatch transform activities against compute resources in an on-premises network or an Azure virtual network.



Azure-SSIS Integration Runtime



Azure Integration Runtime

4. Question 4

In Azure Data Factory when creating a new Copy Activity, which of the following steps do you carry out?

1 / 1 point



Click on Author, and create a new copy activity



Click on Monitor and add a new pipeline and copy activity



Click on Manage, and add a new copy activity



Open the authoring canvas, create a new Pipeline, or use an existing pipeline and then add a new copy activity

Correct

The correct process is to open the authoring canvas, create a new Pipeline and then add a new copy activity.

5. Question 5

Azure Data Factory provides the ability to lift and shift existing SQL Server Integration Services workloads and natively execute these packages. Which of the following would you select to do this?

1 / 1 point



An Azure-SSIS Integration Runtime



Integrate with Azure Databricks



Use Azure Machine Learning



An Azure Function

Correct

Azure Data Factory provides the ability to lift and shift existing SSIS workload by creating an Azure-SSIS Integration Runtime to natively execute SSIS packages.

Knowledge check

Total points 6

1. Question 1

Azure Data Factory provides a range of methods to perform transformations. Which of the following would you use to perform transformations without writing code?

1 / 1 point

- ☐ Azure Machine Learning Studio
- ☐ HDInsight Pig activity
- ☐ Stored procedure activity

☒ Mapping Data Flows

Correct

Mapping data flows are visually designed data transformations in Azure Data Factory. Data flows allow data engineers to develop graphical data transformation logic without writing code. The resulting data flows are executed as activities within Azure Data Factory pipelines that use scaled-out Spark clusters. Data flow activities can be operationalized via existing Data Factory scheduling, control, flow, and monitoring capabilities.

2. Question 2

Which of the following transformations are directly available in the Mapping Data Flows activity?

Select all options that apply.

1 / 1 point

☒ Exists

Correct

Exists checks whether your data exists in another source or stream.

☒ Conditional Split

Correct

Conditional Split can be used to route rows of data to different streams based on matching conditions.



Insert



Aggregate

Correct

Define different types of aggregations such as SUM, MIN, MAX, and COUNT grouped by existing or computed columns.

3. Question 3

Mapping Data Flow follows an extract, load, transform (ELT) approach and works with staging datasets that are in Azure. Which of the following datasets can be used in a source transformation?

Select all options that apply.

0.5 / 1 point



Azure Synapse Analytics



Azure CosmosDB

Correct

Mapping Data Flow supports Azure Cosmos DB



Azure Blob Storage (JSON, Avro, Text, Parquet)

Correct

Mapping Data Flow supports Azure Blob Storage (JSON, Avro, Text, Parquet)



Microsoft SQL Server

This should not be selected

Mapping Data Flow currently supports staging datasets that are all in Azure.

4. Question 4

Mapping Data Flows support Debug so that you can interactively watch how the data transformations are executing. When enabling Debug you are be prompted to select the Integration runtime to use. If you select AutoResolveIntegrationRuntime a cluster will be made available automatically. How many Cores will be made available in the cluster and what will the time to live be in minutes?

1 / 1 point



4 Cores and 60 Minutes



8 cores and 60 minutes



8 Cores and 30 Minutes

☐ 4 Cores and 30 Minutes

Correct

Selecting AutoResolveIntegrationRuntime, will create a cluster of 8 cores and a time to live of 60 Minutes

5. Question 5

When working with Azure Data Factory and wrangling data flows. You add a Source dataset for your wrangling data flow and select a sink dataset. If you use an Azure Data Lake Storage Gen2 Connector what type of authentication is supported in this instance?

Select all options that apply.

1 / 1 point

☒ Account Key

Correct

Azure Data Lake Storage Gen2 supports Data Formats of CSV and Parquet and Authentication of Account Key and Service Principal.

☐ SQL authentication

Azure Data Lake Storage Gen2 supports Data Formats of CSV and Parquet and Authentication of Account Key and Service Principal.

☒ Service Principal

Azure Data Lake Storage Gen2 supports Data Formats of CSV and Parquet and Authentication of Account Key and Service Principal.

Correct

☐ Anonymous Access

6. Question 6

In Azure Data Factory Transformations which transformation in the Mapping Data Flow is used to route data rows to different streams based on matching conditions?

1 / 1 point

☐ Lookup

☒ Conditional Split

☐ Select

☐ Union

Correct

A Conditional Split transformation routes data rows to different streams based on matching conditions. The conditional split transformation is similar to a CASE decision structure in a programming language.

Knowledge check

Total points 5

1. Question 1

Star schema design theory refers to common SCD types. In this design approach, the Dimension Table always reflects the latest values, and when changes in source data are detected, the dimension table data is overwritten. What SCD Type is being referred to in this scenario?

1 / 1 point

- ☐ Type 6 SCD
- ☐ Type 3 SCD
- ☒ Type 1 SCD
- ☐ Type 2 SCD

Correct

A Type 1 SCD always reflects the latest values, and when changes in source data are detected, the dimension table data is overwritten. This design approach is common for columns that store supplementary values, like the email address or phone number of a customer.

2. Question 2

Star schema design theory refers to common SCD types. In this design approach, the Dimension Table supports storing two versions of a dimension member as separate columns. What SCD Type is being referred to in this scenario?

1 / 1 point

- ☐ Type 2 SCD
- ☒ Type 3 SCD
- ☐ Type 1 SCD
- ☐ Type 6 SCD

Correct

A Type 3 SCD supports storing two versions of a dimension member as separate columns. The table includes a column for the current value of a member plus either the original or previous value of the member. So, Type 3 uses additional columns to track one key instance of history, rather than storing additional rows to track each change like in a Type 2 SCD.

3. Question 3

You want to load a dimension table in Azure Synapse from source data in your Azure Synapse database using mapping data flows. Which of the following options would you choose in Synapse Studio Hub to create a Type 1 SCD in a Mapping Data Flow pipeline activity?

1 / 1 point

- ☐ Data
- ☐ Integrate
- ☒ Develop
- ☐ Manage

Correct

The Develop hub contains your SQL scripts, notebooks, data flows, and Spark job definitions.

4. Question 4

Which SCD type would you use to keep history of changes in dimension members by adding a new row to the table for each change?

1 / 1 point

- ☐ Type 1 SCD
- ☐ Type 6 SCD
- ☐ Type 3 SCD
- ☒ Type 2 SCD

Correct

When a value changes, Type 2 SCD will add a new row for the entity with a start date, end date, and unique key which will join back to any transactions in the fact table within the effective data range.

5. Question 5

Which SCD type would you use to update the dimension members without keeping track of history?

1 / 1 point

- ☐ Type 2 SCD
- ☒ Type 1 SCD

- ☐ Type 3 SCD
- ☐ Type 6 SCD

Correct

When a value changes, Type 1 SCD will update the existing record without keeping history.

Knowledge check

Total points 6

1. Question 1

Control flow is the orchestration of pipeline activities. The If-condition activity which is similar to an if-statement provided in programming languages is an example of which type of control flow activity?

1 / 1 point

- ☐ Looping Containers
- ☐ Trigger based flows
- ☒ Branching Activity
- ☐ Chaining Activity

Correct

A branching activity evaluates a set of activities, and when the condition evaluates to true, a set of activities are executed. When it evaluates to false, then an alternative set of activities is executed.

2. Question 2

In Azure Data Factory activities in which a combination of ingesting and cleaning log data alongside a mapping data flow that analyzes the log data is referred to as which of the following?

1 / 1 point

- ☐ Branching activity
- ☐ Looping containers
- ☐ Custom State passing



Pipeline

Correct

A pipeline in Azure Data Factory represents a logical grouping of activities where the activities together perform a certain task. An example of a combination of activities in one pipeline can be, ingesting and cleaning log data in combination with a mapping data flow that analyzes the log data that has been cleaned.

3. Question 3

In Azure Data factory there are many activities that are possible in a pipeline. These activities are grouped into 3 categories. Which of the following are categories of activities?

Select all options that apply.

1 / 1 point



Control activities

Correct

Examples of control flow activities are 'get metadata', 'For Each', and 'Execute Pipeline'.



Data movement activities

Correct

The Copy Activity in Data Factory copies data from a source data store to a sink data store.



Branching Activities



Data transformation activities

Correct

Azure Data Factory supports transformation activities such as Data Flow, Azure Function, Spark, and others that can be added to pipelines either individually or chained with another activity.

4. Question 4

In Azure Data factory activity dependency defines how subsequent activities depend on previous activities. In order to complete a task, an activity that depends on one or more previous activities, can have different dependency conditions.

What are the four dependency conditions?

Select all options that apply.

1 / 1 point



Completed with Errors



Failed

Correct

Failed is a valid dependency.



Succeeded

Correct

Succeeded is a valid dependency.



Skipped

Correct

Skipped is a valid dependency.

☒ Completed

Correct

Completed is a valid dependency.

☐ Pending

5. Question 5

True or False

In Azure Data Factory, Pipelines or activities must be published prior to being debugged.

1 / 1 point

☐ True

☒ False

Correct

In Azure Data Factory, there is no need to publish changes in the pipeline or activities before you want to debug. This is helpful in a scenario where you want to test the changes and see if it works as expected before you actually save and publish them.

6. Question 6

True or False

In Azure Data Factory you enable Debug by selecting the Debug slider. What features will be enabled when selecting the Slider?

1 / 1 point

☒ Enables the Debug feature and run the pipeline

☐ Enables the Debug feature runs the Pipeline and publishes the pipeline

☐ Enables the Debug feature only

☐ Enables the Debug feature and publishes the pipeline

Correct

Enabling the Debug feature will run the pipeline.

Knowledge Check

Total points 6

1. Question 1

In SQL Server, which of the following is a unit of deployment for SSIS solutions?

1 / 1 point



SSIS Projects



SSIS Packages

Correct

A project is the unit of deployment for SSIS solutions. You can deploy projects to an SSIS catalog in a SQL Server instance, and configure project-level parameter values and connections as appropriate for execution environments.

2. Question 2

In Azure data factory an Integration Runtime (IR) can be defined as which of the following?

1 / 1 point



The Action to be performed



A target data store on a compute service



The bridge between the activity and the linked service

Correct

An Integration Runtime (IR) provides the bridge between the Activity and Linked Services.

3. Question 3

When creating a new SSIS Integration Runtime which of the following pieces of information are required on the General setting page?

Select all options that apply.

0.8 / 1 point



Name

Correct

Name is a required entry.



Location

Correct

Location is a required entry.



Node Number



Node Size

Correct

Node Size is a required entry.



Description

You didn't select all the correct answers

4. Question 4

When you migrate your database workloads from SQL Server on premises to Azure SQL database services, you may have to migrate SSIS packages as well. The first step required is to perform an assessment of you current SSIS packages to make sure that they are compatible in Azure. Which of the following tools would be most appropriate to check for compatibility issues?

1 / 1 point



Data Migration Assistant (DMA)



SQL Server Management Studio (SSMS)



SQL Server Data Tools (SSDT)

Correct

The first step required is to perform an assessment of your current SSIS packages to make sure that they are compatible in Azure. You can use the Data Migration Assistant (DMA) to perform an assessment of the SSIS packages that exist and identify any compatibility issues with them.

5. Question 5

An SSIS Project is the unit of deployment for SSIS solutions. In which version of SQL Server was SSIS Projects introduced?

1 / 1 point



SQL Server 2016



SQL Server 2012

Correct

SSIS Projects were introduced in SQL Server 2012 and is the unit of deployment for SSIS solutions.



SQL Server 2008

6. Question 6

Which tool is used to create and deploy SQL Server Integration Packages on an Azure-SSIS integration runtime, or for on-premises SQL Server?

0.6666666666666666 / 1 point



SQL Server Management Studio



SQL Server Data Tools

Correct

SQL Server Data Tools is typically used to create and deploy SQL Server Integration Services (SSIS) packages.



dtexec

This should not be selected

The dtexec command prompt utility is used to configure and execute SQL Server Integration Services packages.

Knowledge check

Total points 6

1. Question 1

In Azure Data Factory you can programmatically interact with multiple languages and SDKs. Which of the following languages and SDKs can you use?

Select all options that apply.

1 / 1 point

☐

Yaml

☒

Python

Correct

In Azure Data Factory you can programmatically interact with languages and SDKs such as Python.

☒

PowerShell

Correct

In Azure Data Factory you can programmatically interact with languages and SDKs such as PowerShell.

☒

.NET

Correct

In Azure Data Factory you can programmatically interact with languages and SDKs such as .NET.

☒

REST APIs

Correct

In Azure Data Factory you can programmatically interact with languages and SDKs such as REST API's.

2. Question 2

Azure Data Factory allows you to configure a Git repository with either Azure Repos or GitHub.

Which of the following are advantages of using GIT Integration?

Select all options that apply.

1 / 1 point

☐

PowerShell or an SDK changes are published directly to GIT

☒

Better CI/CD

Correct

With GIT integration makes certain actions easier such as Configuring release pipeline to trigger automatically as soon as there are any changes made to your 'dev' factory.

☒

Source control

Correct

Using GOT Source Control provides benefits such as the ability to track/audit changes and revert changes that introduced bugs.

☒

Partial saves

Correct

Configuring a git repository allows you to save changes, letting you only publish when you have tested your changes to your satisfaction.

3. Question 3

Which of the following would you select when you need to merge changes in your feature Branch to your collaboration branch?

0 / 1 point



Create Push Request



Create Merge Request



Create pull request

Incorrect

There is no Create Push request when you need to merge changes in your feature Branch to your collaboration branch.

4. Question 4

What effect does removing a Git configuration from a Data Factory have on the repository?

1 / 1 point



It does not delete anything from the repository and the factory will still contain all published resources and you can continue to edit the factory directly against the service.



It deletes everything from the repository. The factory will also lose access to all published resources, and you cannot continue to edit the factory directly against the service.



It deletes everything from the repository. The factory will still contain all published resources and you can continue to edit the factory directly against the service.

Correct

Removing Git configuration from a data factory doesn't delete anything from the repository. The factory will contain all published resources. You can continue to edit the factory directly against the service.

5. Question 5

Mapping data flows allow you to build code-free data transformation logic that runs at scale. When building your logic, you can turn on a debug session to interactively work with your data. Which of the following is created when turning on Debug?

1 / 1 point



Data Lake Storage Account



Spark Cluster



Blob Storage



MS SQL Cluster

Correct

When building your logic, you can turn on a debug session to interactively work with your data using a live Spark cluster.

6. Question 6

In Azure Data factory what is a supported connector for built-in parameterization?

1 / 1 point



Azure Synapse Analytics



Azure Data Lake Storage Gen2

☐ Azure Key Vault

Correct

Azure Synapse Analytics is a supported connector for built-in parameterization for Linked Services in Azure Data Factory.

Knowledge check

Total points 6

1. Question 1

True or False

Azure Data Factory provides support for both code-free Extract Transform Load (ETL) and Extract Load Transform (ELT) processes.

1 / 1 point

☐ False

☒ True

Correct

Azure Data Factory gives you the opportunity to perform code-free Extract Transform Load (ETL) and Extract Load Transform (ELT).

2. Question 2

As a data provider, you can share datasets with third parties or between departments within your organization. You would like to give your data consumers the ability to automatically get refreshed data, in this scenario, which feature would you enable?

1 / 1 point

☒ Snapshots

☐ Data Ingestion

Correct

You can configure a snapshot schedule, so that your data consumers have an option to automatically get refreshed data.

3. Question 3

Azure Data Share provides open and flexible data sharing, including the ability to share from and to different data stores. Which of the following Data stores support incremental snapshots?

Select all that apply

1 / 1 point



Azure Data Lake Storage Gen2

Correct

Azure Data Share supports incremental snapshots for Azure Data Lake Storage Gen2.



Azure Blob Storage

Correct

Azure Data Share supports incremental snapshots for Blob Storage.



Azure Data Lake Storage Gen1

Correct

Azure Data Share supports incremental snapshots for Azure Data Lake Storage Gen1.



Azure SQL Database

4. Question 4

When creating an Azure Data Share from within the portal you must specify which of the following?

Select all options that apply.

1 / 1 point



Location

Correct

You must select a Location for the Data Share.



Name

Correct

You must specify a Name for the Data Share.



Tag



Version



Resource Group Name

Correct

Resource Group name must be specified by either selecting an existing Resource Group or creating a new Resource Group.

5. Question 5

You are defining the connection information for Azure Data Factory to an external resource that you want to connect with, such as an Azure SQL Database. In this scenario, which of the following do you need to create to connect to an external resource?

0 / 1 point



Linked Service



Pipeline



Dataset

Incorrect

a dataset is a named view of data that simply points or references the data you want to use in your activities as inputs and outputs.

6. Question 6

When creating Azure Data Shares which of the following should be selected to enable access to data at source?

1 / 1 point



In-Place sharing



Snapshot sharing

Correct

Choose in-place sharing to provide access to data at its source. Supported data stores include Azure Data Explorer.

Knowledge check

Total points 6

1. Question 1

A Mapping Data Flow is a visually designed transformation service in Azure Data Factory. With the ability to visually design things such as different types of joins, aggregate functions etc. Which of the following Join types are supported in Mapping Data Flow?

Select all options that apply.

1 / 1 point



Inner Join

Correct

Mapping data flows supports Inner join which only outputs rows that have matching values in both tables.



Union



Right Outer

Correct

Mapping data flows supports Right outer join which returns all rows from the right stream and matched records from the left stream.

☒ Left Outer

Correct

Mapping data flows supports Left outer join which returns all rows from the left stream and matched records from the right stream.

2. Question 2

In Azure Data factory after you have finished joining and transforming a dataset through a mapping data flow, it is important to write the newly created dataset into a destination store. Which of the following options would you use to do this?

1 / 1 point

☒ Sink

☐ Copy

☐ Union

☐ Merge

Correct

In a mapping data flow after you finish transforming your data, you write it into a destination store by using the sink transformation.

3. Question 3

In a Mapping Data Flow after you finish transforming your data, you write it into a destination store by using the sink transformation. When using sink transformations which of the following statements is True?

1 / 1 point

☒ Every data flow requires at least one sink transformation, and you can write to multiple sinks as necessary.

☐ Every data flow requires at least one sink transformation, but cannot write to multiple sinks.

Correct

Every data flow requires at least one sink transformation, but you can write to as many sinks as necessary to complete your transformation flow. To write to additional sinks, create new streams via new branches and conditional splits.

4.

Question 4

In Azure Data Factory, what is a logical grouping of activities referred to as?

1 / 1 point

☐ Data Ingestion

☐ A Linked Service

☐ A Sink

☒ A Pipeline

Correct

A pipeline in Azure Data Factory is a logical grouping of activities such as copy to perform a task. The activity defines the operation that you're performing on the data (therefore, a copy means copying the same data to another data store).

5. Question 5

To add or remove datasets created with Azure Data Share which of the following is True?

1 / 1 point

- ☐ It is only possible to remove or add datasets before it's sent within Azure Data Share.
- ☐ It is not possible to add or remove datasets if created with Azure Data Share.

☒ It is possible to add or remove datasets within Azure Data Share after it has been created.

Correct

It is possible to add or remove datasets after it has been created in Azure Data Share.

6. Question 6

What must be done when a connector in Azure Data Factory is not supported in a mapping data flow task to transform data from one of these sources?

1 / 1 point

☒ Ingest the data into a supported source using the copy activity.

- ☐ Use an aggregate transformation in the Dataflow activity.
- ☐ Use a group by activity in the Dataflow activity.

Correct

If a connector in Data factory is not supported, create a copy activity of the source data into a supported data source in mapping dataflow and continue the transformations from there.

Test prep

Latest Submission Grade 90%

1. Question 1

Which of the following processes will allow data to be extracted and loaded in its native format?

Select all options that apply.

1 / 1 point

☐ ETLL

☒ ELTL

Correct

In an ELT process, data is extracted and loaded in its native format. This change reduces the time required to load the data into a destination system. The difference with ELTL is that it also has a final load into a destination system.

☐ ETL

☒ ELT

Correct

In an ELT process, data is extracted and loaded in its native format. This change reduces the time required to load the data into a destination system.

2. Question 2

Select all options that apply.

Pipelines in Azure Data Factory typically perform the four distinct steps. Identify these steps?

1 / 1 point

☒ Monitor

Correct

Azure Data Factory has built-in support for pipeline monitoring via Azure Monitor, API, PowerShell, Azure Monitor logs, and health panels on the Azure portal, to monitor the scheduled activities and pipelines for success and failure rates.

☒ Connect and Collect

Correct

The first step is to define and connect all the required sources of data together, such as databases, file shares, and FTP web services. The next step is to ingest the data as needed to a centralized location for subsequent processing.

☒ Publish

Correct

After the raw data has been refined into a business-ready consumable form from the transform and enrich phase, you can load the data into the Azure Data Warehouse.

☒ Transform and Enrich

Correct

Compute services such as Databricks and Machine Learning can be used to prepare or produce transformed data on a maintainable and controlled schedule to feed production environments with cleansed and transformed data.

☐ Data Analysis

3. Question 3

To create and manage child resources including publishing in the Azure portal, which Data Factory role must you belong to at the resource group level or above?

1 / 1 point

☐ Data Factory User

☐ Data Factory Reader

☒ Data Factory Contributor

☐ Data Factory Writer

Correct

Users must have the Data Factory contributor role on the Resource Group that contains the Data Factory to be able to publish.

4. Question 4

When setting up Azure Data factory through the Azure Portal. Which of the following pieces of Information are required?

Select all options that apply.

1 / 1 point

☒ Resource Group name

Correct

When setting up Azure Data Factory you are required to supply the Resource Group Name.

☒ Subscription

Correct

When setting up Azure Data Factory you are required to supply the subscription in which the ADF instance is created.

☒ Region

Correct

When setting up Azure Data Factory you are required to supply the Region Information.

☒ Instance Name

Correct

When setting up Azure Data Factory you are required to supply the name of the Azure Data Factory instance.

☐ Management Group Name

5. Question 5

Azure Data Factory integration runtime enables bridging between activities and linked Services objects. Which of the following are types of Integration Runtime in Azure data factory?

Select all options that apply.

1 / 1 point

☐ Azure-SQL

☒ Self Hosted

Correct

There are three types of Integration Runtime available in Azure Data Factory, including Self-hosted.

☒ Azure-SSIS

Correct

There are three types of Integration Runtime available in Azure Data Factory, including Azure-SSIS.

☒ Azure

Correct

There are three types of Integration Runtime available in Azure Data Factory, including Azure.

☐ Azure-Spark

6. Question 6

An integration runtime provides which of the following?

1 / 1 point

☐ A target data store or a compute service.

☒ The infrastructure for the activity and linked services.

☐ The action to be performed.

Correct

An integration runtime provides the infrastructure for the activity and linked services.

7. Question 7

Data Factory offers three types of Integration Runtime.

- Azure
- Self-hosted
- Azure-SSIS

Which of these provides support for both Public Networks and Private Networks?

1 / 1 point

☐ Azure-SSIS only

☒ Self Hosted and Azure-SSIS

☐ Azure only

☐ Azure and Self Hosted

☐ Azure and Azure-SSIS

☐ Self Hosted only

Correct

Self Hosted and Azure-SSIS support both Public and Private networks

8. Question 8

You are moving data from an Azure Data Lake Gen2 store to Azure Synapse Analytics. Which Azure Data Factory integration runtime would be used in a data copy activity?

1 / 1 point

- ☐ Azure-SSIS
- ☐ Self-hosted
- ☒ Azure

Correct

When moving data between Azure data platform technologies, the Azure Integration runtime is used when copying data between two Azure data platforms.

9. Question 9

Which Azure Data Factory component orchestrates a data movement command?

0 / 1 point

- ☐ Linked Services
- ☒ Activities
- ☐ Datasets

Incorrect

Try going back and reviewing the Set up and manage Azure Data Factory lesson.

10. Question 10

Which Azure Data Factory process involves using compute services to produce data to feed production environments with cleansed data?

1 / 1 point

- ☒ Transform and enrich
- ☐ Publish
- ☐ Connect and collect
- ☐ Monitor

Correct

Transform and enrich is used to invoke compute services to produce data to feed production environments with cleansed data

Test prep

Latest Submission Grade 82.5%

1. Question 1

Azure Data Factory can call on compute resources to process data by a specific data platform service that may be better suited to the job. Which compute resource can carry out the following activities, Spark, MapReduce and Hadoop Streaming activities?

1 / 1 point

☐ Azure Machine Learning

☒ On-demand HDInsight cluster

☐ Azure Batch

☐ Azure Databricks

Correct

On-demand HDInsight cluster or your own HDInsight cluster can carry out activities such as Hive, Pig, Spark, MapReduce, Hadoop Streaming.

2. Question 2

What are the benefits of Azure-SSIS Integration Runtime?

Select all options that apply.

1 / 1 point

☐ Create On-demand HDInsight clusters

☐ Process Azure Databricks Activities

☒ Natively execute existing SSIS packages

Correct

Azure Data Factory provides the ability to lift and shift existing SSIS workload, by creating an Azure-SSIS Integration Runtime to natively execute SSIS packages, and will enable you to deploy and manage your existing SSIS packages with little to no change.

☒ Deploy and manage existing SSIS packages

Correct

Azure Data Factory provides the ability to lift and shift existing SSIS workload, by creating an Azure-SSIS Integration Runtime to natively execute SSIS packages, and will enable you to deploy and manage your existing SSIS packages with little to no change.

3. Question 3

In Azure Data factory integration runtime defines which of the following?

1 / 1 point

- ☐ The Target Data Store or Compute service
- ☒ The bridge between the activity and linked services
- ☐ The action to be performed

Correct

An integration runtime provides the bridge between the activity and linked services.

4. Question 4

Azure Data factory Copy Activity supports which of the following?

Select all options that apply.

1 / 1 point

- ☒ Code-free data ingestion pipelines.

Correct

The Copy Activity method can be used to build code-free data ingestion pipelines that do not require any transformation during the extraction of the data.

- ☒ Support for over 100 native connectors.

Correct

The Copy Activity has support for over 100 native connectors. This method can suit green-field projects that have a simple method of extraction to an intermediary data store.

- ☐ Complex business logic.
- ☐ Complex transformation during the extraction of data.

5. Question 5

In the Azure Data Factory Authoring Tool, where would you find the Copy data activity?

1 / 1 point

- ☐ Batch Service
- ☐ Databricks
- ☒ Move & Transform

Correct

The Move & Transform section contains activities that are specific to Azure Data Factory copying data and defining data flows.

6. Question 6

You can use which of the following tools to deploy and manage existing SSIS packages using Azure Data Factory?

Select all options that apply.

1 / 1 point

- ☐ Azure Copy Activity
- ☐ Azure Data Factory Pipelines

☒ SQL Server Data Tools (SSDT)

Correct

To deploy and manage your existing SSIS packages with little to no change using familiar tools such as SQL Server Data Tools (SSDT) and SQL Server Management Studio (SSMS), just like using SSIS on-premises.

☒ SQL Server Management Studio (SSMS)

Correct

To deploy and manage your existing SSIS packages with little to no change using familiar tools such as SQL Server Data Tools (SSDT) and SQL Server Management Studio (SSMS), just like using SSIS on-premises.

7. Question 7

To create Data Factory instances, the user account that you use to sign into Azure must be a member of which of the following roles?

0.5 / 1 point

☐ The Owner role

☒ Managed Application Operator Role

This should not be selected

Try going back and reviewing the Ingest Source Data lesson.

☒ The Contributor Role

Correct

The Contributor Role grants full access to manage all resources, but does not allow you to assign roles in Azure RBAC, manage assignments in Azure Blueprints.

☐ Monitoring Contributor

8. Question 8

When using Azure-SSIS Integration runtime, you are given the option to join a virtual network. This option enables Azure Data Factory to create network resources, an example of which is a Network Security Group. What port is configured by default to allow all traffic through the NSG?

0 / 1 point

- ☐ Port 80
- ☒ Port 21
- ☐ Port 443
- ☐ Port 3389

Incorrect

Try going back and reviewing the Ingest Source Data lesson.

9. Question 9

When configuring Azure Data Factory integration runtimes you can deny communication with known IP addresses. Which of the following must you enable to do so?

0.75 / 1 point

☐ Azure Identity and Access management (IAM)

☒ Distributed Denial of Service (DDoS)

Correct

You can deny communication with known IP addresses by enabling distributed denial of service (DDoS) protection

☐ Azure Security Center Integrated Threat Intelligence

☒ Azure Firewall with Threat Intelligence

Correct

Azure Firewall with Threat Intelligence can be used to control network access.
You didn't select all the correct answers

10. Question 10

True or False

Azure Data Factory can ingest data from an on-premises server, or a virtual machine that is hosted within Azure.

1 / 1 point

☒ True

☐ False

Correct

Azure Data Factory can ingest data from an on-premises server, or a virtual machine that is hosted within Azure.

Test prep

Latest Submission Grade 88.89%

1. Question 1

Azure Data Factory can also call on compute resources to transform data. Which Compute environment can be used with Notebook, Jar and Python Activities?

1 / 1 point

- ☐ Azure Machine Learning Studio Machine
- ☐ Azure SQL, Azure SQL Data Warehouse, SQL Server
- ☐ Azure Data Lake Analytics



Azure Databricks

Correct

Azure Databricks supports Notebook, Jar and Python activities.

2. Question 2

Star schema design theory refers to common SCD types. In this design approach, the Dimension Table always reflects the latest values, and when changes in source data are detected, the dimension table data is overwritten. What SCD Type is being referred to in this scenario?

1 / 1 point

- ☐ Type 2 SCD
- ☐ Type 6 SCD
- ☒ Type 1 SCD
- ☐ Type 3 SCD



Type 1 SCD

Correct

A Type 1 SCD always reflects the latest values, and when changes in source data are detected, the dimension table data is overwritten. This design approach is common for columns that store supplementary values, like the email address or phone number of a customer.

3. Question 3

When building data flows, you can enable debug mode allowing for interactive preview of the output of your transformation logic. Enabling Debug Mode turns on which of the following?

1 / 1 point



Spark Cluster

- ☐ SQL Cluster

Correct

When building data flows, you can enable debug mode, which turns on a small interactive Spark cluster.

4. Question 4

You are currently working on a Mapping Data Flow transformation for a Movie Database. You are interested in how a movie ranks within its year for its specific genre. Which of the following transformations should you use to generate the required transformation?

1 / 1 point

☒ Window Transformation

- ☐ Aggregate Transformation
- ☐ Alter Row Transformation
- ☐ A Derive Transformation

Correct

The Window transformation is where you will define window-based aggregations of columns in your data streams. In the Expression Builder, you can define different types of aggregations that are based on data or time windows.

5. Question 5

You want to load a dimension table in Azure Synapse from source data in your Azure Synapse database using mapping data flows. Which of the following options would you choose in Synapse Studio Hub to create a Type 1 SCD in a Mapping Data Flow pipeline activity?

1 / 1 point

- ☐ Integrate
- ☐ Data
- ☐ Manage
- ☒ Develop

Correct

The Develop hub contains your SQL scripts, notebooks, data flows, and Spark job definitions.

6. Question 6

When creating transformations in Azure Data factory which of the following uses a grid type interface for basic data preparation similar to Excel, known as an Online Mashup Editor, while also allowing complex data preparation using formulas?

1 / 1 point

☒ Wrangling Data Flow

- ☐ Mapping Data Flow
- ☐ Azure-SSIS Integration Runtime

Correct

Wrangling Data Flow uses a grid type interface for basic data preparation similar to Excel, known as an Online Mashup Editor, while also allowing complex data preparation using formulas.

7. Question 7

Which transformation is used to load data into a data store or compute resource?

1 / 1 point



Sink



Window



Source



Surrogate key

Correct

A Sink transformation allows you to choose a dataset definition for the destination output data. You can have as many sink transformations as your data flow requires.

8. Question 8

Which SCD type would you use to keep history of changes in dimension members by adding a new row to the table for each change?

0 / 1 point



Type 3 SCD



Type 1 SCD



Type 2 SCD



Type 6 SCD

Incorrect

Try going back and reviewing the Populate slowly changing dimensions in Azure Synapse Analytics pipelines lesson.

9. Question 9

Which of the following transformations are directly available in the Mapping Data Flows activity?

Select all options that apply.

1 / 1 point



Conditional Split

Correct

Try going back and reviewing the Transform and cleanse data with Data Factory lesson.



Aggregate

Correct

Define different types of aggregations such as SUM, MIN, MAX, and COUNT grouped by existing or computed columns.



Exists

Correct

Exists checks whether your data exists in another source or stream.



Insert

Test prep

Latest Submission Grade 87.5%

1. Question 1

In Azure Data Factory to get a view of the history of debug runs, or see all the active debug runs, you can navigate to the monitor tab. How many days does Azure Data Factory service keep Debug run history for?

1 / 1 point

- ☐ 365 Days
- ☒ 15 days
- ☐ 5 Days
- ☐ 30 Days

Correct

Azure Data Factory service keeps Debug run history for 15 Days.

2. Question 2

In Azure Data factory you need to connect to several different databases that are on the same SQL server. Which of the following will require the least administrative effort?

1 / 1 point

- ☐ Create a linked service for each database
- ☒ Use Parameters

Correct

The benefit of parameterizing the database name in the linked service definition is that you don't have to create a single linked service for each database that is on the same SQL Server. It is also possible to parameterize other properties of the linked service like a Username.

3. Question 3

In Azure Data Factory activities in which repetition occurs in a pipeline enabling you to iterate over a collection is referred to as which of the following?

0 / 1 point

- ☐ Custom State passing
- ☐ Pipeline
- ☒ Branching activity
- ☐ Looping containers

Incorrect

Try going back and reviewing the Orchestrate data movement and transformations lesson.

4. Question 4

One of the features of Azure Pipelines is the ability to create Continuous Integration and Continuous Delivery (CI/CD) to persistently build and test the code and finally ship a high-performing, high-quality product.

Which of the following definitions are correct?

Select all options that apply.

1 / 1 point



Continuous Integration produces deployable code, infrastructure, and other artifacts to be deployed.

Correct

Continuous Integration produces deployable code, infrastructure, and other artifacts to be deployed.



Continuous Delivery helps in deploying the integrated and built code into different delivery stages as new versions of the product.

Correct

Continuous Delivery helps in deploying the integrated and built code into different delivery stages as new versions of the product.



Continuous Delivery produces deployable code, infrastructure, and other artifacts to be deployed



Continuous Integration helps in deploying the integrated and built code into different delivery stages as new versions of the product.

5. Question 5

In Azure Data Factory which of the following is a supported connector for built-in parameterization for linked services?

1 / 1 point



Azure Synapse Analytics



Azure Data Lake Storage Gen2



Azure Key Vault

Correct

Azure Synapse Analytics is a supported connector for built-in parameterization for Linked Services in Azure Data Factory.

6. Question 6

In Azure Data factory which of the following is an example of a branching activity used in control flows?

1 / 1 point



Until-condition



The If-condition



Lookup- condition

Correct

An example of a branching activity is The If-condition activity which is similar to an if-statement provided in programming languages.

7. Question 7

In Azure Data Factory you enable Debug by selecting the Debug slider. Selecting the slider will enable which of the following features?

1 / 1 point



Enables the Debug feature and run the pipeline.



Enables the Debug feature runs the Pipeline and publishes the pipeline.



Enables the Debug feature and publishes the pipeline.



Enables the Debug feature only.

Correct

Enabling the Debug feature will run the pipeline.

8. Question 8

In Azure Data Factory a logical grouping of activities where the activities together perform a certain task is referred to as which of the following?

1 / 1 point



Custom State passing



Pipeline



Branching activity



Looping containers

Correct

A pipeline in Azure Data Factory represents a logical grouping of activities where the activities together perform a certain task. An example of a combination of activities in one pipeline can be, ingesting and cleaning log data in combination with a mapping data flow that analyzes the log data that has been cleaned.

Test prep

Latest Submission Grade 97.5%

1. Question 1

SQL Server Integration Services (SSIS) is a platform for building which of the following?

1 / 1 point



Extract Transform and Load (ETL) solutions



Extract Load Transform (ELT) Solutions

Correct

SQL Server Integration Services (SSIS) is a platform for building complex Extract Transform and Load (ETL) solutions.

2. Question 2

SSIS packages defining a workflow of tasks to be executed. The workflow of tasks in a package is referred to as which of the following?

1 / 1 point



SSIS Project



Control flow



Data Flow

Correct

SSIS packages, define a workflow of tasks to be executed. The workflow of tasks in a package is referred to as its control flow.

3. Question 3

To lift and shift an existing SSIS workload, you can create an Azure-SSIS IR to natively execute SSIS packages. The location for your Azure-SSIS IR is important to achieve high performance in your extract-transform-load (ETL) workflows. Select which of the following statements are true.

1 / 1 point



The location of your Azure-SSIS IR must be the same as the location of your data factory.



The location of your Azure-SSIS IR does not need to be the same as the location of your data factory.

Correct

To lift and shift an existing SSIS workload, The location of your Azure-SSIS IR does not need to be the same as the location of your data factory.

4. Question 4

When creating a new SSIS Integration Runtime which of the following pieces of information are required on the Deployment setting page?

Select all options that apply

1 / 1 point



Create Catalog (SSISDB)



Admin Username

Correct

Admin Username is a required entry.

☒ Subscription

Correct

Subscription must be selected.

☒ Catalog database Server Endpoint

Correct

Catalog database Server Endpoint must be completed.

5. Question 5

When migrating SSIS packages, you need to consider the location of the SSIS packages that you are migrating, as this can impact how you migrate the packages. Which of the following storage types are available when migrating SSIS packages?

Select all options that apply.

0.8 / 1 point

☒ SSIS Package Store

Correct

There are four types of storage including SSIS Package store.

☐ Azure Key Vault

☒ File system

☒ SSIS Catalog (SSISDB)

Correct

There are four types of storage including SSIS Catalog.

☒ MSDB Database in SQL Server

Correct

There are four types of storage including MSDB Database in SQL Server.

You didn't select all the correct answers

6. Question 6

Which tool is used to perform an assessment of migrating SSIS packages to Azure SQL Database services?

1 / 1 point

☒ Data Migration Assistant

Correct

The Data Migration Assistant is used to perform an assessment of migrating SSIS packages to Azure SQL Database services.

☐ Data Migration Service

☐ Data Migration Assessment

7. Question 7

You want to ingest data from a SQL Server database hosted on an on-premises Windows Server. What integration runtime is required for Azure Data Factory to ingest data from the on-premises server?

1 / 1 point

☐ Azure Integration Runtime

☐ Azure-SSIS Integration Runtime

☒ Self-Hosted Integration Runtime

Correct

A self-hosted integration runtime can run copy activities between a cloud data store and a data store in a private network. It also can dispatch transform activities against compute resources in an on-premises network or an Azure virtual network.

8. Question 8

When you migrate your database workloads from SQL Server on premises to Azure SQL database services, you may have to migrate SSIS packages as well. The first step required is to perform an assessment of your current SSIS packages to make sure that they are compatible in Azure. Which of the following tools would be most appropriate to check for compatibility issues?

1 / 1 point

☐ SQL Server Management Studio (SSMS)

☐ SQL Server Data Tools (SSDT)

☒ Data Migration Assistant (DMA)

Correct

The first step required is to perform an assessment of your current SSIS packages to make sure that they are compatible in Azure. You can use the Data Migration Assistant (DMA) to perform an assessment of the SSIS packages that exist and identify any compatibility issues with them.

Test prep

Latest Submission Grade 84%

1. Question 1

By default, the Azure Data Factory user interface experience (UX) authors directly against the data factory service. Which of the following are true in respect of the Azure Data Factory user interface experience?

Select all options that apply.

1 / 1 point



Does not include a repository for storing the JSON entities for changes.

Correct

The Data Factory service does not 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 is not optimized for collaboration and version control.

Correct

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



Includes a repository for storing the JSON entities for changes.



The Data Factory service is optimized for collaboration and version control.

2. Question 2

Using the Azure Data factory user interface, which of the following can be used to connect a Git repository to your data factory for both Azure Repos and GitHub?

Select all options that apply.

0.4 / 1 point



ADF Management hub



ADF Monitor Hub

This should not be selected

Try going back and reviewing the Use Azure Data Factory to build your data integration solutions lesson.



During factory creation

Correct

When creating a new data factory in the Azure portal, you can configure Git repository information in the Git configuration tab.



ADF Authoring Hub

Correct

In the Azure Data Factory UX authoring canvas, select the Data Factory drop-down menu, and then select Set up code repository.



ADF Home Page

3. Question 3

Each Azure Repos Git repository that's associated with a data factory has a collaboration branch with a default name of main. Users can also create feature branches. Which of the following statements are true?

1 / 1 point

- ☐ You can publish to the Data Factory service from your Feature branch only.
- ☒ You can publish to the Data Factory service from your collaboration branch only.
- ☐ You can publish to the Data Factory service from both Collaboration and Feature Branches.

Correct

You are only allowed to publish to the Data Factory service from your collaboration branch.

4. Question 4

When implementing Git integration with your Azure Data Factory you typically don't want every team member to have permissions to update the Data Factory. To publish to the Data Factory a user must have which of the following permissions?

1 / 1 point

- ☐ User Access Administrator
- ☒ Data Factory contributor
- ☐ Managed Application Operator Role

Correct

Data Factory Contributor allows a user to create and manage data factories, as well as child resources within them.

5. Question 5

If your ADF development factory has an associated git repository, you can override the default Resource Manager template parameters of the Resource Manager template generated by publishing or exporting the template. However, there is a limit on the number of allowed parameters in a template. Which limit applies to Resource Manager template parameters?

1 / 1 point

- ☐ 512
- ☒ 256
- ☐ 1024
- ☐ 125

Correct

256 is the limit on the number of parameters allowed in an ARM Resource Manager template.

6. Question 6

Azure Data Factory can integrate with which version control software?

1 / 1 point

- ☒ Git repositories
- ☐ Team Foundation Server
- ☐ Source Safe

Correct

Azure Data Factory allows you to configure a Git repository with either Azure Repos or GitHub and is a version control system that allows for easier change tracking and collaboration.

7. Question 7

Which feature commits the changes of Azure Data Factory work in a custom branch created with the main branch in a Git repository?

1 / 1 point

- ☐ Repo
- ☐ Commit
- ☒ Pull request

Correct

After a developer is satisfied with their changes, they create a pull request from their feature branch to the master or collaboration branch to get their changes reviewed by peers.

8. Question 8

Which feature in alerts can be used to determine how an alert is fired?

0 / 1 point

- ☒ Add severity
- ☐ Add criteria
- ☐ Add rule

Incorrect

Try going back and reviewing the Use Azure Data Factory to build your data integration solutions lesson.

9. Question 9

Azure Data Factory allows you to configure a Git repository with either Azure Repos or GitHub. Which of the following are advantages of using GIT Integration?

Select all options that apply.

1 / 1 point

- ☒ Source control

Correct

Using GOT Source Control provides benefits such as the ability to track/audit changes and revert changes that introduced bugs.

- ☒ Partial saves

Correct

Configuring a git repository allows you to save changes, letting you only publish when you have tested your changes to your satisfaction.

- ☒ Better CI/CD

Correct

GIT integration makes certain actions easier such as Configuring release pipeline to trigger automatically as soon as there are any changes made to your 'dev' factory.

- ☐ PowerShell or an SDK changes are published directly to GIT

10. Question 10

Mapping data flows allow you to build code-free data transformation logic that runs at scale. When building your logic, you can turn on a debug session to interactively work with your data. Turning on Debug creates which of the following?

1 / 1 point

- ☐ MS SQL Cluster
- ☒ Spark Cluster
- ☐ Data Lake Storage Account
- ☐ Blob Storage

Correct

When building your logic, you can turn on a debug session to interactively work with your data using a live Spark cluster

Test prep

Latest Submission Grade 83.33%

1. Question 1

You want to add a data flow activity to a pipeline, which of the following options would you choose?

1 / 1 point



Move and Transform



General



Batch Service



Azure Data Explorer

Correct

The Move and Transform option in pipelines allows you to create a data flow activity.

2. Question 2

In Azure Data factory when you create multiple pipelines it is important to keep track of the pipelines. To do so, there is a Monitor tab in Azure Data Factory, which by default will store all the run information. How many days is the run information stored?

0 / 1 point



10 Days



45 Days



30 Days



365 Days

Incorrect

Try going back and reviewing the Join and transform shared and owned data lesson.

3. Question 3

Which transformation can be used to load data into a data store or compute resource?

1 / 1 point



Window



Source



Sink

Correct

A Sink transformation allows you to choose a dataset definition for the destination output data. You can have as many sink transformations as your data flow requires.

4. Question 4

When a connector in Azure Data Factory is not supported in a mapping data flow task to transform data from one of these sources what must be done?

1 / 1 point



Ingest the data into a supported source using the copy activity.



Use a group by activity in the Dataflow activity.

- ☐ Use an aggregate transformation in the Dataflow activity.

Correct

If a connector in Data factory is not supported, create a copy activity of the source data into a supported data source in mapping dataflow and continue the transformations from there.

5. Question 5

In a Mapping Data Flow after you finish transforming your data, you write it into a destination store by using the sink transformation when using sink transformations which of the following statements is True?

1 / 1 point

- ☐ Every data flow requires at least one sink transformation, but cannot write to multiple sinks.
- ☒ Every data flow requires at least one sink transformation, and you can write to multiple sinks as necessary.

Correct

Every data flow requires at least one sink transformation, but you can write to as many sinks as necessary to complete your transformation flow. To write to additional sinks, create new streams via new branches and conditional splits.

6. Question 6

As a data provider, you can share datasets with third parties or between departments within your organization. To enable your data consumers the ability to automatically get refreshed data. What feature should you enable?

1 / 1 point

- ☒ Snapshots
- ☐ Data Ingestion

Correct

You can configure a snapshot schedule, so that your data consumers have an option to automatically get refreshed data.

Course practice exam

Latest Submission Grade 83.33%

1. Question 1

Much of the functionality of Azure Data Factory appears in Azure Synapse Analytics as a feature called pipelines. You can use it to integrate data pipelines between which of the following?

1 / 1 point

☒ SQL Pools

Correct

Much of the functionality of Azure Data Factory appears in Azure Synapse Analytics as a feature referred to as Pipelines, which enables you to integrate data pipelines between SQL Pools, Spark Pools and SQL Serverless.

☐ Apache Hive

☒ SQL Serverless

Correct

Much of the functionality of Azure Data Factory appears in Azure Synapse Analytics as a feature referred to as Pipelines, which enables you to integrate data pipelines between SQL Pools, Spark Pools and SQL Serverless.

☒ Spark Pools

Correct

Much of the functionality of Azure Data Factory appears in Azure Synapse Analytics as a feature referred to as Pipelines, which enables you to integrate data pipelines between SQL Pools, Spark Pools and SQL Serverless.

2. Question 2

Microsoft Azure provides a variety of data platform services that enables you to perform different types of analytics. Predictive analytics can be implemented through which of the following features?

Select all options that apply.

0.75 / 1 point

☒ Machine Learning Services

Correct

Microsoft Azure provides a variety of data platform services that enables you to perform different types of analytics including predictive analytics within HDInsight, Azure Databricks or Machine Learning Services.

☒ Azure Databricks

Correct

Microsoft Azure provides a variety of data platform services that enables you to perform different types of analytics including predictive analytics within HDInsight, Azure Databricks or Machine Learning Services.

☐ Azure Data Lake Storage Gen2

☒ HDInsight

You didn't select all the correct answers

3. Question 3

When setting up Azure Data factory through the Azure Portal. Which of the following pieces of Information are required?

Select all options that apply.

1 / 1 point

☒ Region

Correct

When setting up Azure Data Factory you are required to supply the Region Information.

☒ Subscription

Correct

When setting up Azure Data Factory you are required to supply the subscription in which the ADF instance is created.

☒ Resource Group name

Correct

When setting up Azure Data Factory you are required to supply the Resource Group Name.

☐ Management Group Name

☒ Instance Name

Correct

When setting up Azure Data Factory you are required to supply the name of the Azure Data Factory instance.

4. Question 4

Which of the following is true in respect of Azure data factory datasets?

0 / 1 point

☐ A dataset is a named view that points or references the data.

☒ A dataset is a table that holds a copy of the data for the destination.

- ☐ A dataset does not exist in Azure Data factory
- ☐ A dataset is a table that holds a copy of the data from the source.

Incorrect

Try going back and reviewing **Set up and manage Azure Data Factory**.

5. Question 5

Data Factory offers three types of Integration Runtime.

- Azure
- Self-hosted
- Azure-SSIS

Which of these provides support for both Public Networks and Private Networks?

Select all options that apply.

1 / 1 point

- ☐ Azure only
- ☐ Azure and Self Hosted
- ☒ Self Hosted and Azure-SSIS
- ☐ Azure and Azure-SSIS
- ☐ Azure-SSIS only
- ☐ Self Hosted only

Correct

Self Hosted and Azure-SSIS support both Public and Private networks.

6. Question 6

Which Azure Data Factory component contains the transformation logic or the analysis commands of the Azure Data Factory's work?

1 / 1 point

- ☐ Linked Services
- ☒ Activities
- ☐ Pipelines
- ☐ Datasets

Correct

Activities contains the transformation logic or the analysis commands of the Azure Data Factory's work.

7. Question 7

Azure Data Factory can call on compute resources to process data by a specific data platform service that may be better suited to the job. Which compute resource can carry out the following activities, Spark, MapReduce and Hadoop Streaming activities?

1 / 1 point

- ☐ Azure Databricks
- ☐ Azure Machine Learning
- ☐ Azure Batch

☒ On-demand HDInsight cluster

Correct

On-demand HDInsight cluster or your own HDInsight cluster can carry out activities such as Hive, Pig, Spark, MapReduce, Hadoop Streaming.

8. Question 8

You want to ingest data from a SQL Server database hosted on an on-premises Windows Server. What integration runtime is required for Azure Data Factory to ingest data from the on-premises server?

0.6666666666666666 / 1 point

☐ Azure Integration Runtime

☒ Azure-SSIS Integration Runtime

This should not be selected

Try going back and reviewing **Petabyte-scale ingestion with Azure Data Factory or Azure Synapse Pipeline**.

☒ Self-Hosted Integration Runtime

Correct

A self-hosted integration runtime can run copy activities between a cloud data store and a data store in a private network. It also can dispatch transform activities against compute resources in an on-premises network or an Azure virtual network.

9. Question 9

To create Data Factory instances, the user account that you use to sign into Azure must be a member of which of the following roles?

Select all options that apply.

0.75 / 1 point

☒ The Owner role

Correct

The Owner role grants full access to manage all resources, including the ability to assign roles in Azure RBA.

☐ Managed Application Operator Role

☒ The Contributor Role

Correct

The Contributor Role grants full access to manage all resources, but does not allow you to assign roles in Azure RBAC, manage assignments in Azure Blueprints.

☒ Monitoring Contributor

This should not be selected

Try going back and reviewing **Petabyte-scale ingestion with Azure Data Factory or Azure Synapse Pipeline**.

10. Question 10

Azure Data Factory provides a range of methods to perform transformations. You need to perform transformations without writing code. Which transformations would you use in this scenario?

0 / 1 point

- ☐ Stored procedure activity
- ☐ HDInsight Pig activity
- ☐ Mapping Data Flows

☒ Azure Machine Learning Studio

Incorrect

Feedback: In this scenario, you would not use Azure Machine Learning Studio. Try going back and reviewing **Perform code-free transformation at scale with Azure Data Factory or Azure Synapse Pipeline**.

11. Question 11

Which transformation is used to load data into a data store or compute resource?

1 / 1 point

☐ Source

☒ Sink.

☐ Window

☐ Surrogate key

Correct

A Sink transformation allows you to choose a dataset definition for the destination output data. You can have as many sink transformations as your data flow requires.

12. Question 12

You want to load a dimension table in Azure Synapse from source data in your Azure Synapse database using mapping data flows. Which of the following options would you choose in Synapse Studio Hub to create a Type 1 SCD in a Mapping Data Flow pipeline activity?

1 / 1 point

☒ Develop

☐ Integrate

☐ Manage

☐ Data

Correct

The Develop hub contains your SQL scripts, notebooks, data flows, and Spark job definitions.

13. Question 13

In Azure Data factory there are many activities that are possible in a pipeline. These activities are grouped into 3 categories. Which of the following are categories of activities?

Select all options that apply.

0.5 / 1 point

☒ Data transformation activities

☒ Control activities

Correct

Examples of control flow activities are 'get metadata', 'ForEach', and 'Execute Pipeline'.

☒ Branching Activities

This should not be selected

Try going back and reviewing **Orchestrate data movement and transformation in Azure Data Factory or Azure Synapse Pipeline.**

☒ Data movement activities

Correct

The Copy Activity in Data Factory copies data from a source data store to a sink data store.

14. Question 14

In Azure Data Factory to get a view of the history of debug runs, or see all the active debug runs, you can navigate to the monitor tab. Azure Data Factory service keeps Debug run history for how many days?

1 / 1 point



15 days



30 Days



365 Days



5 Days

Correct

Azure Data Factory service keeps Debug run history for 15 Days.

15. Question 15

To lift and shift an existing SSIS workload, you can create an Azure-SSIS IR to natively execute SSIS packages. The location for your Azure-SSIS IR is important to achieve high performance in your extract-transform-load (ETL) workflows. Select which of the following statements are true.

1 / 1 point



The location of your Azure-SSIS IR must be the same as the location of your data factory.



The location of your Azure-SSIS IR does not need to be the same as the location of your data factory.

Correct

To lift and shift an existing SSIS workload, The location of your Azure-SSIS IR does not need to be the same as the location of your data factory.

16. Question 16

When you migrate your database workloads from SQL Server on premises to Azure SQL database services, you may have to migrate SSIS packages as well. The first step required is to perform an assessment of your current SSIS packages to make sure that they are compatible in Azure. Which of the following tools would be most appropriate to check for compatibility issues?

1 / 1 point



SQL Server Data Tools (SSDT)



Data Migration Assistant (DMA)



SQL Server Management Studio (SSMS)

Correct

The first step required is to perform an assessment of your current SSIS packages to make sure that they are compatible in Azure. You can use the Data Migration Assistant (DMA) to perform an assessment of the SSIS packages that exist and identify any compatibility issues with them.

17. Question 17

An SSIS Project is the unit of deployment for SSIS solutions. In which version of SQL Server was SSIS Projects introduced?

1 / 1 point

- ☐ SQL Server 2008.
- ☐ SQL Server 2016

☒ SQL Server 2012

Correct

SSIS Projects were introduced in SQL Server 2012 and is the unit of deployment for SSIS solutions.

18. Question 18

Azure Data Factory allows you to configure a Git repository with either Azure Repos or GitHub. Which of the following are advantages of using GIT Integration?

Select all options that apply.

1 / 1 point

☒ Better CI/CD

Correct

GIT integration makes certain actions easier such as Configuring release pipeline to trigger automatically as soon as there are any changes made to your 'dev' factory.

☐ PowerShell or an SDK changes are published directly to GIT.

☒ Source control

Correct

Using GIT Source Control provides benefits such as the ability to track/audit changes and revert changes that introduced bugs.

☒ Partial saves

Correct

Configuring a git repository allows you to save changes, letting you only publish when you have tested your changes to your satisfaction.

19. Question 19

Mapping data flows allow you to build code-free data transformation logic that runs at scale. When building your logic, you can turn on a debug session to interactively work with your data. Turning on Debug creates which of the following?

1 / 1 point

- ☐ MS SQL Cluster
- ☐ Data Lake Storage Account
- ☒ Spark Cluster
- ☐ Blob Storage

Correct

When building your logic, you can turn on a debug session to interactively work with your data using a live Spark cluster.

20. Question 20

In Azure Data factory what is a supported connector for built-in parameterization?

1 / 1 point

- ☐ Azure Data Lake Storage Gen2.

☒ Azure Synapse Analytics

- ☐ Azure Key Vault

Correct

Azure Synapse Analytics is a supported connector for built-in parameterization for Linked Services in Azure Data Factory.