✓ **Congratulations! You passed!**

Grade received 87.50%   **To pass** 80% or higher

[Go to next item]

ℹ **You're viewing this assessment in its original language**
You can switch back to view this content in your preferred translation if you'd prefer. You won't lose any progress if you change languages. **Show Vietnamese version**   ✕

---

1. Which of the following are common options for processing data in Microsoft Azure?   **1 / 1 point**

Choose all options that apply.

☑ Azure Data Factory

✓ **Correct**
The most common options for processing data in Azure include Azure Databricks, Azure Data Factory, Azure Synapse Analytics, and Azure Data Lake.

☐ Azure Data Studio

☑ Azure Databricks

✓ **Correct**
The most common options for processing data in Azure include Azure Databricks, Azure Data Factory, Azure Synapse Analytics, and Azure Data Lake.

☑ Azure Synapse Analytics

✓ **Correct**
The most common options for processing data in Azure include Azure Databricks, Azure Data Factory, Azure Synapse Analytics, and Azure Data Lake.

---

2. Azure Synapse Analytics is a generalized analytics service. Which technologies does it support to process data?   **1 / 1 point**

Choose all options that apply.

☐ PL/SQL

☑ Transact-SQL

✓ **Correct**
This is the same dialect of SQL used by Azure SQL Database, with some extensions for reading data from external sources, such as databases, files, and Azure Data Lake storage.

☑ Spark

✓ **Correct**
This is the same open-source technology used to power Azure Databricks. You write your analytical code using notebooks in a programming language such as C#, Scala, Python, or SQL.

---

3. What do you use to write Databricks applications?   **1 / 1 point**

◉ A Notebook

○ A Yaml File

○ An XML File

✓ **Correct**
You write Databricks applications using a Notebook. A notebook contains a series of steps (cells), each of which contains a block of code.

---

4. Azure HDInsight is a managed analytics service in the cloud and is based on which of the following?   **1 / 1 point**

○ Apache Spark

◉ Apache Hadoop

○ Apache Hive

✓ **Correct**
Azure HDInsight is based on Apache Hadoop, a collection of open-source tools and utilities that enable you to run processing tasks over large amounts of data.

5. HDInsight can be used to analyze data using a collection of open-source tools. The ability to provide a scalable, fault-tolerant platform for running real-time data processing applications is provided by which service?

0 / 1 point

○ Apache Kafka

◉ Hadoop Map/Reduce

○ Apache Storm

○ Apache Hive

⊗ **Incorrect**
Hadoop Map/Reduce uses a simple framework to split a task over a large dataset into a series of smaller tasks over subsets of the data that can be run in parallel, and the results then combined.

6. True or False?

1 / 1 point

Azure Data Lake Store provides a hierarchical file system, granular security over data and supports both raw data (blobs) and structured data.

◉ True

○ False

⊘ **Correct**
Data Lake Store provides a file system that can store near-limitless quantities of data. It uses a hierarchical organization (like the Windows and Linux file systems), and can hold massive amounts of raw data (blobs) and structured data. It is optimized for analytics workloads.

7. Azure Data Lake Analytics is an on-demand analytics job service that you can use to process big data. You write jobs that contain queries to transform data and extract insights. What language do you use to define a job?

1 / 1 point

◉ U-SQL

○ PL-SQL

○ T-SQL

⊘ **Correct**
You define a job using a language called U-SQL. This is a hybrid language that takes features from both SQL and C#, and provides declarative and procedural capabilities that you can use to process data.

8. Azure Synapse Analytics is an integrated analytics service that allows organizations to gain insights quickly from all their data from both data warehouses and big data analytics systems.

1 / 1 point

Azure Synapse is composed of which of the following elements?

Choose all options that apply.

☐ Synapse NoSQL pool

☑ Synapse Pipelines

⊘ **Correct**
A Synapse pipeline is a logical grouping of activities that together perform a task. The activities in a pipeline define actions to perform on your data.

☑ Synapse SQL pool

⊘ **Correct**
Synapse SQL pool is a collection of servers running Transact-SQL. You write your data processing logic using Transact-SQL.

☑ Synapse Spark pool

⊘ **Correct**
Synapse Spark pool is a cluster of servers running Apache Spark to process data. You write your data processing logic using one of the four supported languages: Python, Scala, SQL, and C#.