✅ Quiz: Data Mapping & Ingestion Techniques

Question 1

A cloud data analyst is in the process of mapping their data. As a part of the process, they identify the fields that need to be mapped and standardize the fields names. What is the next step that the cloud data analyst should do?

- A. Combine the sources into a single dataset.
- ✅ **B. Create the data mapping rules.**
- C. Test the mapping in a subset of the data.
- D. Create a data map to relate the datasets.

📝 **Feedback:**
After identifying and standardizing fields, the next step is to **define mapping rules**—instructions for how fields from different sources will be matched.

Question 2

A cloud data analyst is designing a data pipeline. They are in the extract stage and are deciding which ingestion technique to use. What is a key factor in the decision?

- A. What the ingestion schedule is
- ✅ **B. If the data is time sensitive**
- C. How much storage is needed
- D. Where the data is available

📝 **Feedback:**
Time sensitivity is the key factor—**streaming ingestion** is used for real-time needs, while **batch ingestion** is used when timing is less critical.

Question 3

How does data mapping help improve data quality?

- A. By reducing the time needed for the analysis.
- ✅ **B. By ensuring that the data is standardized and consistent.**

- C. By reducing the amount of data that needs to be analyzed.
- D. By ensuring that the data fits the hypothesis being analyzed.

📝 **Feedback:**
Standardization and consistency are essential for accurate analysis, and data mapping ensures that fields align correctly across sources.

Question 4

A cloud data analyst is creating a data pipeline. As a part of the process, they are mapping the data. What does the cloud data analyst do when mapping data?

- ✅ **A. Match the fields between data sources.**
- B. Eliminate the duplicated data.
- C. Match the data between data sources.
- D. Eliminate the duplicated fields.

📝 **Feedback:**
Data mapping involves matching **fields** (not entire datasets) between sources to ensure they align for integration and analysis.

Question 5

A cloud data analyst is designing the first step in a data pipeline. They decide to use batch ingestion. Why does the cloud data analyst choose batch ingestion over streaming ingestion?
Select two answers.

- A. The data is on the cloud.
- ✅ **B. The data is not critical.**
- C. They want speedy results.
- D. The data is time sensitive.
- ✅ **E. They want lower costs.**

📝 **Feedback:**
Batch ingestion is ideal when data isn't time-sensitive and when **cost-efficiency** is a priority, as it uses fewer resources than streaming.