✅ Quiz: Data Pipelines and ELT/ETL Concepts

Question 1

In which stage of a data pipeline does a cloud data analyst check the data for duplicate entries or obvious errors?

- A. Extract stage
- B. Load stage
- ✅ **C. Transform stage**
- D. Process stage

📝 **Feedback:**
The **Transform** stage is where data is cleaned and standardized, including removing duplicates and correcting errors.

Question 2

A cloud data analyst is tasked with creating a pipeline to solve a data problem. Based on the characteristics of the problem, they choose an ELT pipeline. What are two reasons why the cloud data analyst chose the ELT pipeline over the ETL pipeline?
Select two answers.

- A. They are working with historical data.
- ✅ **B. The data needs to be tracked as soon as it is available.**
- C. The data is in more than four data sources.
- ✅ **D. They need to have flexibility to transform the data as needed.**
- E. They need to use the cloud as data storage.

📝 **Feedback:**
ELT is ideal for **real-time tracking** and **flexible transformation**, especially in cloud environments where scalability matters.

Question 3

A cloud data analyst is solving a data problem using a pipeline. As part of the process, they move the data into a temporary staging area. In what stage of the data pipeline is the cloud data analyst working?

- A. Transform
- B. Load
- ✅ **C. Extract**
- D. Process

📝 **Feedback:**
During the **Extract** stage, raw data is gathered and moved to a **staging area** before transformation begins.

Question 4

A cloud data analyst is solving a data problem. After careful consideration, they decide to create an ELT pipeline. What process illustrates the correct sequence for an ELT pipeline?

- A. Extract the data from the company databases, transform the data from the smart sensors around the company, clean and format the data, and then load the data into the destination system.
- B. Extract the data from the company databases, clean the data, load the data from the smart sensors around the company, transform the data, and then load the data into the destination system.
- C. Extract the data from the company databases and the smart sensors around the company, clean and format the data, and then load the data into the destination system.
- ✅ **D. Extract the data from the company databases and the smart sensors around the company, load the data into the destination system, and then clean and format the data.**

📝 **Feedback:**
In **ELT**, data is **loaded first**, then **transformed** later—ideal for cloud-based, scalable environments.

Question 5

What are two key benefits of a data pipeline?
Select two answers.

- A. Helps organizations reduce the use of coding languages
- ✅ **B. Helps organizations save time and resources**
- C. Helps organizations identify data problems
- ✅ **D. Automates the processes of extract, load, and transform**
- E. Automates the process of storage on the cloud

📝 **Feedback:**

Data pipelines **automate ETL/ELT processes** and help organizations **save time and resources** by streamlining data movement and transformation.