Quiz: Data Manipulation, Profiling, and Cleaning

Question 1

A cloud data analyst is working with a dataset. They notice that the fields containing dates do not follow the same format, and they find format inconsistencies in a few other fields. What manipulation technique can the cloud data analyst use to fix these format issues?

- A. Data conversion
- B. Data standardization
- C. Data enrichment
- D. Data profiling

Feedback:

Data standardization ensures that all fields follow a consistent format, making the dataset easier to process and analyze.

Question 2

What is the term for the manipulation technique that changes the format of data to improve compatibility?

- A. Data enrichment
- **V** B. Data conversion
- C. Data standardization
- D. Data processing

Feedback:

Data conversion changes the format of data (e.g., from CSV to Parquet) to improve compatibility, readability, or efficiency.

Ouestion 3

A cloud data analyst is analyzing a dataset. As part of the process, they find the names and types of the 8 dataset columns. They also identify that the data has duplicate records and some of the data in one of the columns does not follow the same format. Which data process is the cloud data analyst using?

- A. Data processing
- **V** B. Data profiling
- C. Data duplication
- D. Data cleaning

Feedback:

Data profiling involves exploring the structure, format, and quality of data to identify issues before cleaning or transformation.

Question 4

A cloud data analyst is working with a data project. They decide they need to enrich the dataset to obtain more valuable information. What can the cloud data analyst do to enrich the dataset?

Select two answers.

- A. Join the data with other data sources
- B. Store the dataset in the cloud
- C. Standardize the data to a common format
- D. Apply mathematical conversions
- Z E. Add new fields to the data

Feedback:

Data enrichment involves adding new fields or joining external data sources to enhance the dataset's value and depth.

Question 5

A cloud data analyst is analyzing a dataset. As part of the process, they identify duplicate records, inconsistencies in the data format, and missing values. What should the cloud data analyst do next?

- A. Collect more data
- B. Clean the data
- C. Process the data
- D. Profile the data

Feedback:

After profiling reveals issues, the next step is **data cleaning**—fixing or removing errors to ensure accuracy and completeness.