

**1. A cloud data analyst is working with a dataset that is the result of a data derivation process. What data problem should the analyst consider when using derived data?**

- The data can have NULLs values as products of the transformation
  - ☒ **The data is not always as accurate as the original data**  
*Correct: Derived data depends on algorithms and base data quality—errors in either can affect accuracy.*
  - The data can have personally identifiable information
  - The data is not always available to the cloud data analyst
- 

**2. A cloud data analyst is tasked with creating a report. The data they need for creating the report is not found directly in the raw data. As a next step, they decide to use data derivation. What does the cloud data analyst do in the data derivation process?**


- ☒ **The cloud data analyst creates an algorithm to transform the raw data**  
*Correct: Data derivation involves applying logic or calculations to base data to generate new insights.*
  - The cloud data analyst deletes the duplicate records in the data
  - The cloud data analyst performs statistical calculations with the raw data
  - The cloud data analyst standardizes the format of the raw data to be consistent
- 

**3. A cloud data analyst is joining two tables using an outer join. As a part of the process, the result table has some rows with NULLs values. What do the NULLs values represent?**

- The zeros or blank data
  - The unused fields
  - ☒ **The unmatched values**  
*Correct: NULLs in outer joins indicate missing matches from one of the joined tables.*
  - The missing data
-

**4. A cloud data analyst is joining two tables. For their project, the cloud data analyst needs only the information that matches both tables.**

**What type of join should the cloud data analyst use?**

- Full outer join
- Left outer join
- Outer join
-  **Inner join**

*Correct: Inner joins return only rows with matching values in both tables.*