

Quiz - Data Quality

Started on 12/04/2025 at 6:41 PM - Ended on 12/04/2025 at 6:50 PM

Progression: Managed by the host

Question 1

Question

Under a Closed World Assumption, a code not in the valid list is:

Answers (32)

- a. Incorrect by definition (88% - 29) - Correct answer
- b. Auto-corrected by the DBMS (6% - 2)
- c. Possibly correct, needs more evidence (0% - 0)
- d. Ignored at query time (3% - 1)

Question 2

Question

Under an Open World Assumption, when a value isn't in the known list:

Answers (30)

- a. It's "not possible to tell for sure" (85% - 28) - Correct answer
- b. It must be dropped (3% - 1)
- c. It's definitely wrong (0% - 0)
- d. It should be coerced to NULL (3% - 1)

Question 3

Question

Which best separates syntactic vs. semantic accuracy?

Answers (32)

- a. Both are the same in practice (0% - 0)
- b. Syntactic = user needs; Semantic = DB constraints (0% - 0)
- c. Syntactic = meaning; Semantic = format (9% - 3)
- d. Syntactic = format/rules; Semantic = meaning/reality correspondence (88% - 29) - Correct answer

Question 4

Question

Which statement about completeness is true?

Answers (30)

- a. Missing values can be acceptable, depending on the task/user needs (85% - 28) - Correct answer
- b. Completeness is independent of user tasks (0% - 0)
- c. Missing values always imply inconsistency (0% - 0)
- d. All missing values are completeness errors (6% - 2)

Question 5

Question

Referential integrity mainly prevents:

Answers (31)

- a. Orphans/mismatched references across tables (91% - 30) - Correct answer

- b. NULLs in numeric fields (0% - 0)
- c. Duplicate rows (3% - 1)
- d. Out-of-date values (0% - 0)

Question 6

Question

Comparing categories of different population sizes using counts creates bias. The fix is to compare:

Answers (31)

- a. Densities/normalized values (e.g., per 100k) (91% - 30) - Correct answer
- b. Cumulative sums (3% - 1)
- c. Pie charts only (0% - 0)
- d. Raw totals (0% - 0)

Question 7

Question

Mixing estimates/projections with final values without disclosure primarily harms:

Answers (30)

- a. Currency (0% - 0)
- b. Accessibility only (0% - 0)
- c. Precision only (3% - 1)
- d. Understandability (and can introduce consistency issues) (88% - 29) - Correct answer

Question 8

Question

Two polls with different party lists, different “Undecided/NA” handling, and different poll dates are compared as equal. The main flaw is:

Answers (30)

- a. Inconsistent entities/timeframe/method (91% - 30) - Correct answer
- b. Too many decimals (0% - 0)
- c. Over-use of colors (0% - 0)
- d. Low sample size by definition (0% - 0)

Question 9

Question

“Currency” is best defined as:

Answers (30)

- a. Data’s immunity to bias (0% - 0)
- b. Data’s exchange value (15% - 5)
- c. Data’s compliance with ISO 8601 (0% - 0)
- d. Data being up-to-date for reality and the task at hand (76% - 25) - Correct answer

Question 10

Question

Precision should provide:

Answers (30)

- a. Whatever the chart library defaults to (0% - 0)
- b. Enough detail to discriminate, but not overload (88% - 29) - Correct answer
- c. No decimals ever (0% - 0)
- d. Maximum decimals always (3% - 1)