

QUIZ 2 — LAB 3.2

Data Quality Dashboard

1 Pourquoi mesurer le taux de rejet est plus important que le nombre brut de lignes rejetées ?

2 Si le taux de rejet passe de 3 % à 18 % en une journée, quelles hypothèses peux-tu formuler ?

3 Quelle transformation Hop permet de calculer des agrégations comme AVG, COUNT ou MAX ?

4 Quelle différence entre :

- COUNT(*)
 - COUNT(colonne)
-

5 Pourquoi un dashboard de qualité doit-il être historisé ?

6 Quel est le risque d'utiliser uniquement la moyenne pour analyser `total_amount` ?

7 À partir de quel seuil une donnée devient-elle un problème ?

Qui décide ce seuil ?

8 Quelle différence entre :

- métrique descriptive
- métrique d'alerte

9 Pourquoi exporter les métriques dans une table SQL peut être préférable à un CSV ?

10 Si les données sont propres mais incohérentes d'un point de vue métier, le dashboard le détectera-t-il ? Pourquoi ?