

Etape 1

Définitions des contrôles pour chaque colonne d'une table.

Email

- % de valeurs manquantes
- % de valeurs dupliquées
- ...

Etape 2

Le % d'anomalies de la colonne est la moyenne des anomalies mesurées par ses contrôles.

- 13% de valeurs manquantes
- 5% de valeurs dupliquées
- ...



9% d'anomalies en moyenne

Etape 3

Le % d'anomalies de la table est la moyenne des anomalies mesurées par les contrôles de ses colonnes.

- 3% de valeurs manquantes sur la colonne 1
- 5% de valeurs dupliquées sur la colonne 2
- ...



4% d'anomalies en moyenne

Etape 4

Le % d'anomalies du périmètre est la moyenne des anomalies mesurées par les contrôles des colonnes de ses tables.

- 8% de valeurs manquantes sur la colonne 1 de la table 1
- 6% de valeurs dupliquées sur la colonne 1 de la table 2
- ...



7% d'anomalies en moyenne

La variation du % d'anomalies permet d'observer l'évolution entre le % d'anomalies du dernier contrôle et le % d'anomalies de l'avant dernier contrôle $[(\% \text{ anomalies } J - \% \text{ anomalie } J-1) / \% \text{ anomalie } J-1]$.

Une variation positive signifie plus d'anomalies et donc une dégradation de la qualité des données.

