**CM4 : Qualité des données**

**6 dimensions : Politique de conformité :**

| Exactitude | Reflet de la réalité ? |  | Certification des données |
| --- | --- | --- | --- |
| Exhaustivité | Informations clés manquantes ? |  | Gestion de la confidentialité |
| Cohérence | Formatage cohérent ? |  | Lignée des données: Source et transformations subies |
| Disponibilité | Au moment nécessaire et attendu ? |  | Limitation de l’utilisation |
| Validité | Format attendu ? |  | Source de référence unifiée |
| Unicité | Doublons ? |  |  |

**Procédures relatives à la qualité des données** (exemples) :

| Modèles normalisés d'inspection des données |
| --- |
| Qualité opérationnelle des données |
| Suivi des problèmes et remédiation |
| Intervention manuelle si nécessaire |
| Intégrité de l'échange de données |
| Plan d'urgence |
| Validation des données |

**Impact sur les entreprises** (si mauvaise qualité des données) :

| Augmentation des coûts | Données redondantes ou inutiles: espace stockage++, durée processus++, retards++, charge++ |
| --- | --- |
| Baisse des revenus | Mauvaise adresse: retards++. Mauvaise facturation/sous-facturation. |
| Baisse de confiance | Conflits, non utilisation de l’application |
| Risque accru | Législation, santé, confidentialité, fraudes… |

Qualité des données est significative uniquement lorsque la "qualité des données" est liée à

l'utilisation prévue des données. Certaines données sont bonnes ("haute qualité") pour certains usages, mais

mauvaises ("basse qualité") pour d'autres.

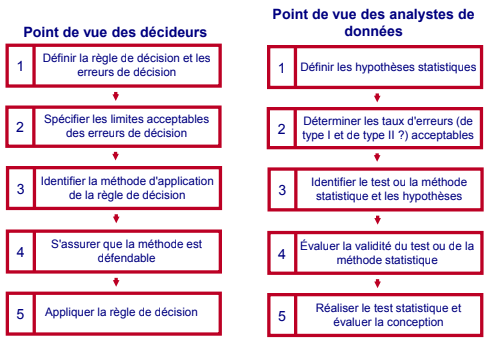
**Evaluation de la Qualité des Données (EQD)**

| Vérification des données - processus d'évaluation de l'exhaustivité, de l'exactitude et de la  conformité d'un ensemble de données spécifique par rapport à la méthode, aux procédures ou aux  exigences contractuelles. |
| --- |
| Validation des données - processus spécifique à l'analyste et à l'échantillon qui étend l'évaluation  des données au-delà de la méthode, de la procédure ou de la conformité contractuelle (c'est-à-dire  la vérification des données) afin de déterminer la qualité analytique d'un ensemble de données  spécifique. |
| Évaluation de la qualité des données - processus visant à déterminer si les données sont adaptées  à une utilisation spécifique. |

**Les cinq étapes de l'évaluation de la qualité des données :**

| Examiner les objectifs et le plan d'échantillonnage | Réviser l’étendu du problème si besoin | Traduire les objectifs de l'utilisateur, Définir les limites tolérables d’erreurs et leurs effets, Examiner le plan d'échantillonnage, Les critères de décision et de performance |
| --- | --- | --- |
| Procéder à un examen préliminaire des données | Transformer ou modifier les données si besoin | Examiner les rapports d'assurance qualité, Calculer des quantités statistiques standard, Représenter les données graphiquement |
| Choisir la méthode statistique | Choisir un nouveau test statistique si besoin | Méthode statistique appropriée pour les données et les objectifs du projet, Lister des hypothèses pour la méthode statistique |
| Vérifier les hypothèses de la méthode statistique | Revenir en arrière si besoin | Déterminer l’approche et le niveau de signification approprié, Effectuer des tests |
| Tirer des conclusions à partir des données |  | Évaluer les résultats, Tenir compte du contexte |

**Deux points de vues :**

****