

# **Procédures de documentation**

## **Objectif**

Montrer ta capacité à formaliser les processus pour qu'un projet data soit reproductible, sécurisé et maintenable.

## **A. Organisation des procédures**

1. **Collecte des données**
  - Identifier les sources.
  - Configurer les connecteurs ETL.
  - Vérifier la fréquence et la complétude.
2. **Nettoyage et préparation**
  - Standardisation des formats.
  - Gestion des valeurs manquantes.
  - Documentation des choix (glossaire technique).
3. **Stockage**
  - Structuration des répertoires dans le Data Lake.
  - Application des règles d'accès.
  - Plan de sauvegarde automatisé.
4. **Analyse et modélisation**
  - Workflow d'expérimentation (train/test/validation).
  - Versionnage des scripts.
  - Documentation des hyperparamètres.
5. **Restitution**
  - Création et validation des dashboards.
  - Production de rapports exécutifs et techniques.
  - Archivage des livrables.
6. **Sécurité**
  - Chiffrement des données sensibles.
  - Gestion des droits d'accès.
  - Journalisation des accès et modifications.

## **B. Format des documents**

- **Guide utilisateur** : étapes simples avec captures d'écran.
- **Documentation technique** : architecture, scripts commentés, schémas de flux.
- **Dossier projet** : résumé exécutif + livrables.