



Rapport d'Analyse des solutions de Data Visualisation

Date : Juillet 2024

Auteur : Manon DARGET

## Table des matières

1. Introduction .....	2
1.1. Objectif du rapport .....	2
1.2. Contexte de l'analyse .....	2
2. État des lieux et évaluation de la pertinence .....	2
2.1. Description de la situation actuelle .....	2
2.2. Évaluation de la pertinence des données existantes .....	2
2.3. Identification des principaux problèmes ou enjeux .....	2
3. Besoin d'outils .....	3
3.1. Identification des outils nécessaires pour collecter les données .....	3
3.2. Identification des outils nécessaires pour traiter les données .....	3
3.3. Identification des outils nécessaires pour analyser les données .....	3
4. Solutions d'extraction, de traitement et de visualisation .....	4
4.1. Solutions proposées pour l'extraction des données .....	4
4.2. Solutions proposées pour le traitement des données .....	4
4.3. Solutions proposées pour la visualisation des données .....	4
5. Cohérence des solutions avec le besoin .....	5
5.1. Explication détaillée de chaque solution proposée .....	5
5.2. Alignement des solutions avec les besoins identifiés .....	5
5.3. Avantages et limites de chaque approche .....	5
6. Conclusion .....	6
6.1. Synthèse des principales conclusions de l'analyse .....	6
6.2. Recommandations pour la mise en œuvre des solutions proposées .....	6
7. Annexes .....	6

# 1. Introduction

## 1.1. Objectif du rapport

L'objectif de ce rapport est d'analyser les différentes solutions disponibles pour la mise en place d'un système de data visualisation pour notre entreprise. Ce rapport présente une évaluation des options de gestion des données, ainsi que des recommandations basées sur les besoins identifiés.

## 1.2. Contexte de l'analyse

Suite à la présentation sur le nettoyage des données, il a été décidé de lancer un projet de mise à disposition des données de l'entreprise, celles-ci étant aujourd'hui de meilleure qualité. Pour ce faire, nous devons choisir une solution de data visualisation adaptée à nos besoins.

# 2. État des lieux et évaluation de la pertinence

## 2.1. Description de la situation actuelle

Actuellement, les données de l'entreprise sont mises à jour dans les différentes bases de données (BDD) et nous avons un export de la BDD au format SQLite couvrant la période du 1er Octobre 2022 au 30 Septembre 2023.

## 2.2. Évaluation de la pertinence des données existantes

Les données existantes couvrent une année entière et semblent suffisamment complètes pour permettre une analyse approfondie. Cependant, une vérification plus détaillée des données est nécessaire pour s'assurer de leur qualité et de leur pertinence.

## 2.3. Identification des principaux problèmes ou enjeux

- ❖ Intégrité et cohérence des données entre les différentes sources.
- ❖ Choix d'un outil de visualisation qui s'intègre bien avec nos systèmes existants.
- ❖ Facilité d'utilisation et de maintenance de la solution choisie.

### 3. Besoin d'outils

#### 3.1. Identification des outils nécessaires pour collecter les données

- 🍷 Connexion directe à la base de données.
- 🍷 Extraction des données en CSV.
- 🍷 Utilisation d'un ETL (PowerQuery, Knime, etc.).

#### 3.2. Identification des outils nécessaires pour traiter les données

- 🍷 ETL pour transformer et nettoyer les données.
- 🍷 Outils de gestion de base de données pour la manipulation des données brutes.

#### 3.3. Identification des outils nécessaires pour analyser les données

- 🍷 Logiciels de data visualisation (Tableau, Power BI, etc.).



- 🍷 Langages de programmation (Python, R) pour des analyses plus personnalisées.



## 4. Solutions d'extraction, de traitement et de visualisation

### 4.1. Solutions proposées pour l'extraction des données

- 🍷 **Connexion directe** : Permet un accès en temps réel aux données mais nécessite une infrastructure réseau robuste.
- 🍷 **Extraction en CSV** : Simple à mettre en place mais peut devenir rapidement obsolète si les données sont fréquemment mises à jour.
- 🍷 **Utilisation d'un ETL** : Offre une flexibilité maximale pour transformer et nettoyer les données avant leur visualisation.

### 4.2. Solutions proposées pour le traitement des données

- 🍷 **PowerQuery** : Intégré dans Excel et Power BI, adapté pour des utilisateurs non techniques.



- 🍷 **Knime** : Offre des capacités avancées de traitement et de transformation des données.



### 4.3. Solutions proposées pour la visualisation des données

- 🍷 **Tableau** : Réputé pour sa capacité à créer des visualisations interactives et intuitives.



- 🍷 **Power BI** : Bien intégré avec les produits Microsoft et offre une excellente connectivité aux sources de données variées.



🍷 **D3.js** : Pour des visualisations web personnalisées mais nécessite des compétences en programmation.



## 5. Cohérence des solutions avec le besoin

### 5.1. Explication détaillée de chaque solution proposée

🍷 **Connexion directe** : Idéale pour des besoins en temps réel.

🍷 **Extraction en CSV** : Simple à mettre en œuvre, utile pour des analyses ponctuelles.

🍷 **ETL** : Essentiel pour des flux de données complexes et transformations nécessaires.

### 5.2. Alignement des solutions avec les besoins identifiés

🍷 Tableau et Power BI : Alignés avec les besoins de visualisation avancée et interactive.

🍷 Knime et PowerQuery : Alignés avec les besoins de traitement et de transformation de données.

### 5.3. Avantages et limites de chaque approche

🍷 **Connexion directe** :

- Avantages - Accès en temps réel.
- Limites - Dépendance à l'infrastructure réseau.

### **Extraction en CSV :**

- Avantages - Facilité d'utilisation.
- Limites - Maintenance et actualisation des données.

### **ETL :**



- Avantages - Flexibilité et puissance.
- Limites - Complexité de mise en œuvre.

## **6. Conclusion**

### **6.1. Synthèse des principales conclusions de l'analyse**

Pour une solution robuste et flexible, **l'utilisation** d'un **ETL** pour le traitement des données, **couplé avec Power BI** pour la visualisation, semble être la meilleure approche. Cela permet une intégration fluide avec nos systèmes existants et offre une grande flexibilité pour les analyses futures.

### **6.2. Recommandations pour la mise en œuvre des solutions proposées**

-  Mettre en place une phase pilote avec Power BI et un ETL comme Power Query pour évaluer la solution.
-  Former l'équipe aux outils choisis pour assurer une adoption fluide et efficace.

## **7. Annexes**

- Liste des sources de données utilisées
  - Base de données SQLite exportée.
  - Schéma relationnel de la BDD.
  - Dictionnaire des données.