

[Rapport d'Analyse]

Date: [Date]

Table des matières

[Introduction 2](#_Toc164848905)

[● 1.1 Objectif du rapport 2](#_Toc164848906)

[● 1.2 Contexte de l'analyse 2](#_Toc164848907)

[État des lieux et évaluation de la pertinence 2](#_Toc164848908)

[● 2.1 Description de la situation actuelle 2](#_Toc164848909)

[● 2.2 Évaluation de la pertinence des données existantes 2](#_Toc164848910)

[● 2.3 Identification des principaux problèmes ou enjeux 2](#_Toc164848911)

[Besoin d'outils 3](#_Toc164848912)

[● 3.1 Identification des outils nécessaires pour collecter les données 3](#_Toc164848913)

[● 3.2 Identification des outils nécessaires pour traiter les données 3](#_Toc164848914)

[● 3.3 Identification des outils nécessaires pour analyser les données 3](#_Toc164848915)

[Solutions d'extraction, de traitement et de visualisation 3](#_Toc164848916)

[● 4.1 Solutions proposées pour l'extraction des données 3](#_Toc164848917)

[● 4.2 Solutions proposées pour le traitement des données 4](#_Toc164848918)

[● 4.3 Solutions proposées pour la visualisation des données 4](#_Toc164848919)

[Cohérence des solutions avec le besoin 4](#_Toc164848920)

[● 5.1 Explication détaillée de chaque solution proposée 4](#_Toc164848921)

[● 5.2 Alignement des solutions avec les besoins identifiés 5](#_Toc164848922)

[● 5.3 Avantages et limites de chaque approche 5](#_Toc164848923)

[Conclusion 6](#_Toc164848924)

# Introduction

## 1.1 Objectif du rapport

Ce présent rapport d’analyse a pour objectif de permettre la mise en place d’un système de data visualisation afin de mettre à disposition les données de l’entreprise.

## 1.2 Contexte de l'analyse

Une première mission de nettoyage de la base de données a permis de mettre à jour les informations afin de faciliter l’analyse.

Une fois les travaux de modification des bases de données réalisées, il est maintenant possible de réfléchir à un projet de mise à disposition d’un tableau de bord de data visualisation.

# État des lieux et évaluation de la pertinence

## 2.1 Description de la situation actuelle

Bottleneck présente néanmoins un certain retard quant à l’utilisation des données qu’elle détient. Cela met en danger sa compétitivité, son organisation interne ainsi que sa prévision de croissance commerciale.

## 2.2 Évaluation de la pertinence des données existantes

Bottleneck enregistre les informations relatives à

* Sa base de données financières : identification des produits et de ses prix
* Sa base de données des ventes : liste des ventes du mois et actualisation du stock
* Sa base de données de la boutique en ligne : liste de tous les produits en ligne et de ses caractéristiques
* Sa base de données de promotion : les promotions proposées avec les dates de début et de fin d’offre.

## 2.3 Identification des principaux problèmes ou enjeux

* Les données ne sont pas en temps réel. Il faut tout anticiper car les résultats se font en fin de mois. Il y a de ce fait un risque de rupture de stock, une mauvaise anticipation des résultats financiers. L’enjeux serait de rendre disponible les données à fréquence plus proches (quotidiennes, hebdomadaires, bimensuelles)
* Utilisation de deux outils différents pour gérer les bases de données. Wordpress et l’ERP développé n’ont pas les mêmes règles de disponibilités des données. L’enjeux serait de se détacher de Wordpress afin d’avoir une totale mise à disposition des bases de données.

# Besoin d'outils

## 3.1 Identification des outils nécessaires pour collecter les données

* Les données de la base de données Web sont issues du site sous Wordpress
  + La collecte des données se fait par le biais d’un export XML qui peut donner lieu à un fichier CSV.
* Les données relatives aux finances, aux ventes et gestion du stock ainsi que des promotions proposées sont gérées par le biais de l’ERP en place
  + La collecte des données s’effectue par le biais d’un export CSV. Il faudrait se rapprocher du service informatique afin de savoir si une connexion SQL est envisageable.

## 3.2 Identification des outils nécessaires pour traiter les données

Le traitement des données peut être réalisé par :

* Requêtes SQL (PostGre SQL, Sqlite, SQL Server)
* Notebook Python (Jupyter par exemple)
* Excel (lecture des fichiers CSV et XML)

## 3.3 Identification des outils nécessaires pour analyser les données

Les données traitées peuvent ensuite être analysées par plusieurs moyens :

* Notebook Python : création de tables et calculs via la bibliothèque Pandas, création de graphiques via les bibliothèques Seaborn et Matplotlib
* Microsoft Excel : calculs, création de tableaux croisés dynamiques, créations de graphiques
* Microsoft PowerBi : création de tables, tableaux croisés dynamiques et visuels

# Solutions d'extraction, de traitement et de visualisation

## 4.1 Solutions proposées pour l'extraction des données

Il existe 3 solutions d’extraction :

* Extraction par requête SQL (Création en sortie d’un fichier CSV) – Uniquement pour le logiciel ERP
* Excel / PowerBi : Connection à un fichier CSV
* Extraction automatique par le biais de PowerBi (Importer ou Direct Query)
  + DirectQuery est une connexion ne conservant aucune donnée au sein du logiciel mais ne nécessitant pas d’actualisation
  + Importer permet de stocker les données mais l’actualisation (automatique ou manuelle) est nécessaire.

## 4.2 Solutions proposées pour le traitement des données

Les données extraites peuvent être traitées de la manière suivante :

* Les données extraites par requêtes SQL devront par la suite être traitées par un autre logiciel qui peut être Excel, Notebook Python ou PowerBi.
* Les fichiers CSV peut être traité par le biais de tableurs de type Excel, via un Notebook Python ou utilisé sur PowerBi.
* La connexion PowerBi à la base de données permet de pouvoir les traiter immédiatement.

## 4.3 Solutions proposées pour la visualisation des données

La visualisation des données peut être réalisée par différents moyens :

* Par Excel, en présentant ensuite les graphiques sous un fichier PowerPoint
* Par un Notebook Python, en présentant ensuite les graphiques sous un fichier PowerPoint
* Par PowerBi, embarquant un système de datavisualisation

# Cohérence des solutions avec le besoin

## 5.1 Explication détaillée de chaque solution proposée

Solution 1 : Analyse des fichiers CSV via Python ou Excel

* Tous les débuts de mois il faudra effectuer les différents exports (ERP et Wordpress) afin d’avoir des fichiers CSV et/ou Excel exploitables sur Excel ou Notebook Python. L’analyse se fera sur ces outils et une restitution devra être réalisée sur un fichier PowerPoint qui pourra ensuite être diffusé.

Solution 2 : Connexion des fichiers CSV sur PowerBi

* *Tous les débuts de mois il faudra effectuer les différents exports (ERP et Wordpress) afin de les stocker dans un répertoire défini. Le nom des fichiers devra être identique et les données précédentes seront écrasées. PowerBi s’actualisera sur la connexion à ces fichiers et les graphiques s’actualiseront en autonomie. Il faudra ensuite charger les données sur la plateforme en ligne PowerBi afin que chaque utilisateur autorisé ait accès aux données à jour.*

Solution 3 : Connexion aux bases de données via PowerBi

* Deux options :
  + Sans passerelle : il faudra ouvrir le fichier PowerBi tous les débuts de mois et actualiser les données. Les tables et graphiques s’actualiseront en autonomie. Il faudra ensuite charger les données sur la plateforme en ligne PowerBi afin que chaque utilisateur ait accès aux données à jour.
  + Avec passerelle : la passerelle sera paramétrée de telle manière que tous les débuts de mois les données PowerBi seront mises à jour.

## 5.2 Alignement des solutions avec les besoins identifiés

Solution 1 : Analyse des fichiers CSV via Python ou Excel

* Les résultats sont présentés sous forme d’un fichier PowerPoint manque d’ergonomie. Il n’est pas possible de personnaliser les données, les filtrer.

Solution 2 : Connexion des fichiers CSV sur PowerBi

* Les résultats sont présentés via un dashboard PowerBi consultable librement avec possibilité de filtrer

Solution 3 : Connexion aux bases de données via PowerBi

* Les résultats sont présentés via un dashboard PowerBi consultable librement avec possibilité de filtrer

## 5.3 Avantages et limites de chaque approche

Solution 1 : Analyse des fichiers CSV via Python ou Excel

* **Avantages** :
  + Toute personne à l’aise avec les outils Microsoft (Excel/PowerPoint) peut réaliser l’analyse. Python nécessite néanmoins un peu plus de connaissance.
* **Inconvénients** :
  + Nécessite l’autorisation d’extraire les données ou demander à fréquence régulière au service informatique les fichiers CSV à jour
  + Chronophage, car l’analyse devra être refaite mensuellement, ainsi que le fichier PowerPoint de restitution
  + Peu de personnalisation des résultats, sauf demande particulière.

Solution 2 : Connexion des fichiers CSV sur PowerBi

* **Avantages** :
  + Personnalisation des données communiquées via filtres et/ou droits d’accès.
  + Dashboard ergonomique
  + Export possible en pdf pour garder une trace des résultats.
* **Inconvénients**:
  + Nécessite l’autorisation d’extraire les données ou demander à fréquence régulière au service informatique les fichiers CSV à jour
  + Nécessite d’actualiser manuellement le fichier PowerBi et penser à charger les données sur la plateforme en ligne.
  + Licence payante obligatoire pour le partage des résultats

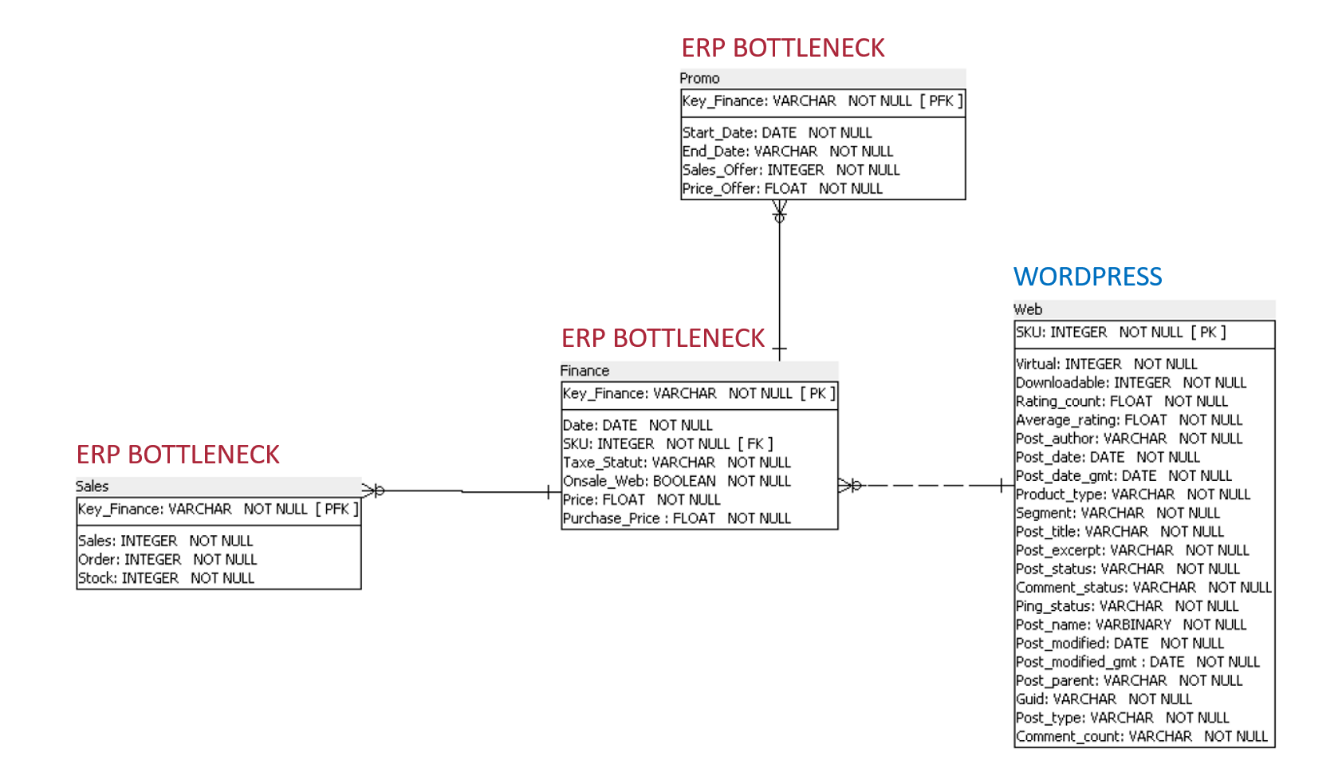
Solution 3 : Connexion aux bases de données via PowerBi

* **Avantages** :
  + Si utilisation de passerelle : totale autonomie, aucune action manuelle
  + Personnalisation des données communiquées via filtres et/ou droits d’accès.
  + Dashboard ergonomique
  + Export possible en pdf pour garder une trace des résultats.
* **Inconvénients** :
  + Si absence de passerelle : nécessaire d’actualiser les données et les charger sur la plateforme PowerBi en ligne
  + Licence payante obligatoire pour le partage des résultats ainsi que l’utilisation d’une passerelle
  + S’agissant des données de Wordpress, il ne sera pas possible de s’y connecter automatiquement, uniquement les données issues de l’ERP.

# Conclusion

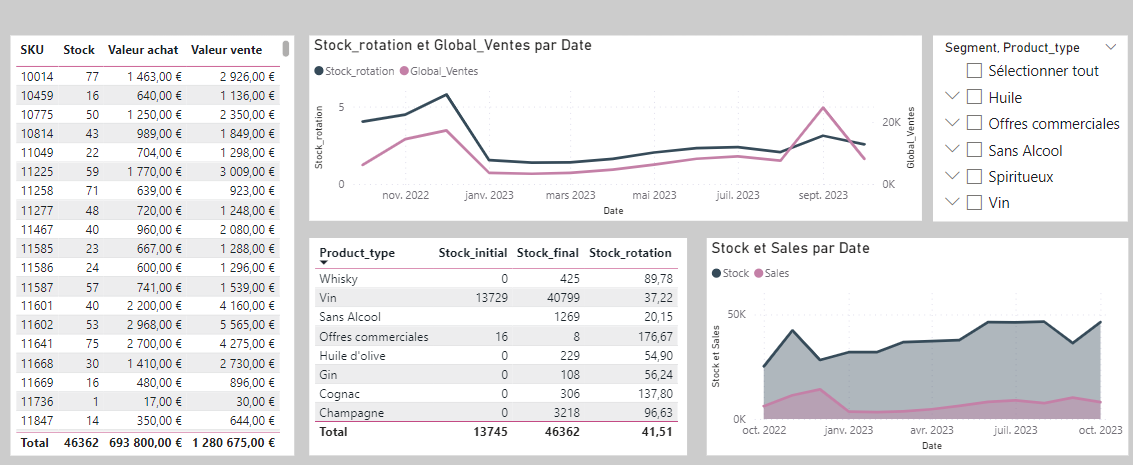
* **Synthèse des principales conclusions de l'analyse**
  + Dans le cadre d’une mise à disposition d’un dashboard pour que les différents acteurs puissent avoir une vue sur plusieurs KPI, la **solution 3** (connexion aux bases de données PowerBi) répond davantage aux exigences :
    - Droit d’accès
    - Ergonomie
    - Actualisation automatique possible
* **Recommandations pour la mise en œuvre des solutions proposées**
  + Il faudra penser à une solution de connexion automatique aux bases de données issues de Wordpress
  + Une licence « Pro » devra être fournie à tous les utilisateurs de l’outil
  + Une petite formation en interne peut être nécessaire pour comprendre le fonctionnement et l’accès au dashboard PowerBi
  + Une passerelle devra être mise en place afin de permettre une actualisation automatique

# [Annexes]

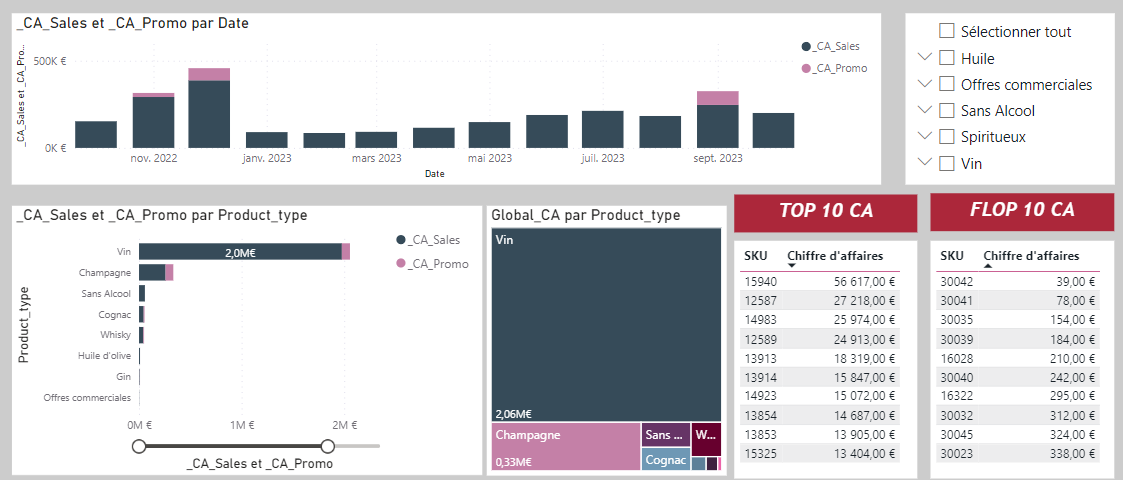
* Liste des sources de données utilisées
  + Nécessité d’avoir accès :
    - A la base de données du Wordpress
    - A la base de données de l’ERP
* Description détaillée des outils recommandés
  + Necessité d’avoir une licence PowerBi Pro minimum afin de profiter d’un maximum d’option de l’outil, la possiblité de communiquer en interne les résultats et d’avoir une passerelle permettant l’actualisation automatique.



* Exemples de visualisations de données proposées
  + Tables, courbes, aires



* + Histogrammes, treemap



* + Secteurs, cartes de résultats

