



[Rapport d'Analyse]

Date: mai 2025

Introduction

- **Objectif du rapport**

Ce rapport a pour objectif de proposer une solution de data visualisation adaptée aux besoins de l'entreprise Bottleneck. Cette solution doit permettre une meilleure exploitation des données existantes afin d'optimiser la prise de décision, notamment en matière de gestion des stocks, de suivi des ventes, et d'analyse du comportement client.

- **Contexte de l'analyse**

Après une première mission de nettoyage des données liée à la gestion des stocks, l'entreprise Bottleneck souhaite désormais aller plus loin dans sa transformation numérique en mettant en place un outil de data visualisation. L'objectif est de rendre les données accessibles et compréhensibles pour les équipes, afin de faciliter l'analyse et le pilotage de l'activité.

État des lieux et évaluation de la pertinence

- **2.1 Description de la situation actuelle**

L'entreprise dispose d'une base de données contenant 4 tables : Web, Finance, Promo et Sales. La table Web contient des données sur tous les articles du magasin (référence, moyenne des notes, url des produits, etc...). La table Finance contient les prix de ventes et d'achats notamment. La table Promo recense les promotions pour les produits en question (date de début et de fin de promotion, etc). Enfin, la table Sales contient des informations comme le stock au dernier jour du mois. Les données sont accessibles via une base SQLite, mais elles ne sont pas encore utilisées dans un outil de visualisation.

Il sera donc nécessaire de faire des extractions des CSV pour les utiliser dans l'outil de visualisation de données.

- **2.2 Évaluation de la pertinence des données existantes**

Les données existantes sont globalement exploitables, mais il y a quelques points de vigilance. Certaines tables présentent des doublons ou des valeurs manquantes. Dans la table *sales* on constate des valeurs manquantes dans la colonne [Orders], quand il n'y a pas de commandes la valeur est bien à 0 mais certaines sont totalement vides).

La structuration des données est correcte mais pourrait être optimisée pour une utilisation analytique (normalisation, typage clair, etc.). On pourrait remplacer des virgules par des points pour les valeurs décimales. Certaines colonnes n'ont pas le bon format, par exemple: web[Average_rating], [Post_author], [Post_date], [Post_date_gmt]. Dans la table *finance*, la colonne [Onsale_web] doit contenir uniquement des 0 ou des 1 (selon le dictionnaire de donnée fournie) tandis qu'il y a une ligne contenant un 4.

- **2.3 Identification des principaux problèmes ou enjeux**

Absence de centralisation pour l'analyse (données dispersées, parfois extraites manuellement).

Difficulté d'accès aux données pour les équipes non techniques.

Manque de visualisations synthétiques pour la direction et les responsables de service.

Risque d'erreurs ou d'incohérences en cas d'extractions manuelles non automatisées.

Besoin d'outils

- **3.1 Identification des outils nécessaires pour collecter les données**

La collecte de données pourrait se faire directement via le site via des API par exemple puis elles seraient structurées via des tables SQL par exemple. Une fois ces tables correctement structurées nous pourrions les disposer dans un google drive par exemple pour que tout le monde y ait accès.

- **3.2 Identification des outils nécessaires pour traiter les données**

Le traitement des données pourrait se faire avec le langage SQL et/ou python. Leur faciliter d'usage permettrait un nettoyage préalable des tables si nécessaire, le typage des colonnes sera également plus simple. Une fois ce nettoyage réalisé nous pourrions exporter ces tables sous format csv. Ce fichier pourrait éventuellement être converti en un tableau excel pour des modifications manuelles si besoin.

- **3.3 Identification des outils nécessaires pour analyser les données**

Deux outils sortent du lot pour analyser les données : Power Bi et Tableau. D'un point de vue coût Power Bi s'en sort mieux car la licence est moins coûteuse que celle de Tableau.

L'intégration Microsoft (Excel, Google Drive) se fait de manière plus simple sur Power Bi et donc serait plus à même de répondre à notre besoin. Le connecteur avec Excel et Google Drive est aussi possible sur Tableau mais moins fluide.

L'interface de Tableau est plus visuelle mais moins faciles pour les débutants tandis que celle de Power Bi est plus intuitive.

D'un point de vue visualisation de données, interface graphique Tableau s'en sort mieux car plus de finesse, les graphiques sont plus malléables.

D'un point de vue volumétrie Tableau sort du lot.

Le langage de calcul DAX (power bi) est très puissant mais demande un certain apprentissage avant de le maîtriser et Tableau Calc (tableau) semble plus intuitif et donc idéal pour les débutants.

La mise en production des rapports est très simple avec Power Bi service tandis que Tableau demande d'avoir au préalable Tableau Server ou Tableau Cloud.

Au niveau de la documentation, les deux logiciels sont équivalents car très bien fournis.

Solutions d'extraction, de traitement et de visualisation

- **4.1 Solutions proposées pour l'extraction des données**

Il serait judicieux d'avoir une connexion directe à la base de données pour une meilleure automatisation.

- **4.2 Solutions proposées pour le traitement des données**

Nettoyage avec PowerQuery pour les données peu volumineuses.

Création de tables de faits et de dimensions si besoin d'un modèle en étoile (modèle Power Bi).

- **4.3 Solutions proposées pour la visualisation des données**

Power BI est plus adapté à Bottleneck car :

- Il est moins coûteux, ce qui est essentiel pour une PME.
- Il s'intègre très bien à Excel, facilitant la prise en main.
- Il propose toutes les fonctionnalités nécessaires (KPI, tableaux de bord dynamiques, filtres, etc.).
- Il permet une mise en œuvre rapide avec un bon niveau d'automatisation via Power Query.

Il nous permettra notamment de mettre en place un dashboard par service pour assurer la gestion des stocks, les ventes, la fidélisation des clients.
La visualisation est possible par produit, période, etc.

Cohérence des solutions avec le besoin

- **5.1 Explication détaillée de chaque solution proposée**

Connexion directe : permet un accès en temps réel aux données ; réduit les erreurs.
Power BI : outil adapté aux PME, permettant des rapports interactifs et partagés.
PowerQuery : intégré à Power BI, permet de nettoyer et transformer les données avant l'analyse.

- **5.2 Alignement des solutions avec les besoins identifiés**

Automatisation du flux de données pour gagner du temps et éviter de potentielles erreurs manuelles.

Des visualisations claires pour permettre une prise de décision adéquat et rapide.
Le rapport sera accessible aux équipes non techniques ce qui permettra un échange de connaissances et une démocratisation de l'usage des données.

- **5.3 Avantages et limites de chaque approche**

Approche	Avantages	Limites
Connexion directe	Temps réel, fiable	Nécessite configuration technique
Extraction CSV	Simple, rapide	Peu automatisé, risques d'erreur
ETL (PowerQuery)	Flexible, permet des traitements	Peut demander un temps d'apprentissage
Power BI	Intuitif, puissant	Licence nécessaire (pour partage avancé)
Excel	Connu de tous	Limité pour des gros volumes de données

Conclusion

- **Synthèse des principales conclusions de l'analyse**

La mise en place d'une solution de visualisation est non seulement faisable mais aussi stratégique pour Bottleneck. Les données existantes sont globalement exploitables. Une solution combinant Power BI, PowerQuery et une connexion directe à la base de données apparaît comme la plus cohérente.

- **Recommandations pour la mise en œuvre des solutions proposées**

- Mettre en place une connexion directe à la base Excel/OneDrive via Power BI.
- Créer un modèle de données propre (via PowerQuery).
- Construire des tableaux de bord par thématique (ventes, stocks, clients).
- Former les équipes à la lecture des dashboards.
- Prévoir un suivi mensuel pour améliorer le modèle selon les besoins évolutifs.

[Annexes]

- **Liste des sources de données utilisées**

Web, finance, sales, promo

- **Exemples de visualisations de données proposées**

- Évolution des ventes mensuelles par région
- Niveaux de stock par produit
- Taux de fidélité client par période
- Top 10 des produits les plus vendus