



[Rapport d'Analyse]

Date: [Date]

Introduction

- Objectif du rapport

Ce rapport vise à analyser et recommander la meilleure solution pour connecter les bases de données de Bottleneck (finance, promo, sales, web) à un outil de data visualisation, afin de créer un tableau de bord efficace pour le PDG Florian et les chefs de produits.

- Contexte de l'analyse

Suite à la mise à jour récente des données dans les différentes bases de données, Bottleneck souhaite mettre en place une solution de data visualisation. Cette initiative s'inscrit dans un projet plus large de mise à disposition des données de l'entreprise, lancé par le PDG Florian.

État des lieux et évaluation de la pertinence

- 2.1 Description de la situation actuelle

Base de données SQLite avec 4 tables relationnelles (finance, promo, sales, web) liées par des clés primaires. Ces dernières ont des noms différents, mais leur format est identique. Je modifierais dans le logiciel choisi.

Données historiques couvrant **octobre 2022 à septembre 2023**

- 2.2 Évaluation de la pertinence des données existantes

Atouts : Les données sont structurées de manière relationnelle, ce qui facilite leur analyse. Le schéma fourni et le dictionnaire de données permettent une bonne compréhension de la structure.

Limites : SQLite moins adapté aux requêtes concurrentielles lourdes.

Pas de mécanisme natif de mise à jour automatique pour la visualisation.

- 2.3 Identification des principaux problèmes ou enjeux

Besoin d'accès en temps réel aux données pour la prise de décision.

Nécessité d'une interface user-friendly pour les utilisateurs non techniques.

Sécurisation de l'accès aux données sensibles.

Capacité à gérer efficacement les mises à jour régulières des données.

Besoin d'outils

- 3.1 Identification des outils nécessaires pour collecter les données

Connecteurs de base de données, Scripts d'extraction : SQLite, PostgreSQL

Outils ETL (Extract, Transform, Load) : Knime, Power Query, Talend Open Studio (TOS)

- 3.2 Identification des outils nécessaires pour traiter les données

Outils de nettoyage et transformation de données : Python avec Pandas, SQL, R, Knime

Outils d'agrégation pour le calcul des KPI : Power Query, Knime, Tableau

- 3.3 Identification des outils nécessaires pour analyser les données

Outils de Business Intelligence : Power BI, Tableau, Looker Studio, Plotly via Python
(bibliothèque de visualisation)

Solutions d'extraction, de traitement et de visualisation

- 4.1 Solutions proposées pour l'extraction des données

Connexion directe à la base de données

Critère	SQLite Studio	PostgreSQL	KNIME	Power Query	Talend Open Studio (TOS)
Prix	Gratuit	Gratuit	Gratuit (version de base)	Gratuit avec Excel/Power BI	Gratuit (open source)
Accessibilité	Très facile à utiliser, pas de configuration serveur nécessaire	Nécessite une configuration serveur	Interface graphique intuitive	Intégré à Excel/Power BI, facile à utiliser	Interface graphique, mais courbe d'apprentissage plus raide
Modélisation des données	Basique, adaptée aux petites applications	Avancée	Modélisation visuelle puissante	Modélisation limitée, principalement pour la préparation des données	Avancée
Visualisations	Limitées (exploration des données)	Limitées, nécessite des outils tiers	Nombreuses options de visualisation intégrées	Limitées dans Power Query (excellentes dans Power BI)	Limitée (pour le monitoring des jobs)
Mise à jour	Facile pour les petites bases de données	Robuste, supporte les mises à jour concurrentes	Mises à jour régulières	Mises à jour fréquentes via Microsoft	Mises à jour régulières
Automatisation	Limitée, principalement via scripts	Excellente, supporte les procédures stockées et les triggers	Workflows automatisés puissants	Automatisation via Power Automate	Automatisation puissante des jobs ETL

- 4.2 Solutions proposées pour le traitement des données

Nettoyage et transformation de données, agrégation pour le calcul des KPI :

Critère	Python (Pandas)	R	KNIME	Power Query	Tableau
Prix	Gratuit	Gratuit	Gratuit (version de base)	Gratuit avec Excel/Power BI	Payant
Accessibilité	Nécessite des compétences en programmation	Nécessite des compétences en programmation	Interface graphique intuitive	Intégré à Excel/Power BI, facile à utiliser	Interface intuitive
Modélisation des données	Très flexible	Puissante pour l'analyse statistique	Modélisation visuelle puissante	Modélisation limitée (pour la préparation des données)	Modélisation avancée
Visualisations	Nombreuses options via bibliothèques	Nombreuses options via packages	Nombreuses options intégrées	Limitées dans Power Query, mais excellentes dans Power BI	Excellentes capacités de visualisation
Mise à jour	Fréquente	Fréquente	Mises à jour régulières	Mises à jour fréquentes via Microsoft	Mises à jour régulières
Automatisation	Très puissante via scripts	Puissante via scripts	Workflows automatisés puissants	Automatisation via Power Automate	Automatisation limitée

- 4.3 Solutions proposées pour la visualisation des données

Critère	Power BI	Tableau	Looker Studio	Plotly (Python)
Prix	Abordable, plusieurs options tarifaires	Plus coûteux, licence par utilisateur	Gratuit (version de base)	Gratuit (open-source)
Accessibilité	Interface intuitive, facile à prendre en main	Interface conviviale, légère courbe d'apprentissage	Facile à utiliser, orienté utilisateur final	Nécessite des compétences en programmation Python
Modélisation des données	Puissante avec DAX et Power Query	Robuste, avec langage VizQL	Utilise LookML pour la modélisation	Flexible, dépend des compétences en Python
Visualisations	Large gamme de visualisations	Excellentes capacités de visualisation	Bonnes options de visualisation	Hautement personnalisable
Mise à jour	Fréquentes mises à jour par Microsoft	Mises à jour régulières	Mises à jour régulières	Mises à jour fréquentes
Automatisation	Power Automate, API REST	Tableau Prep, API	Intégration avec Google Cloud	Hautement automatisable

Cohérence des solutions avec le besoin

- 5.1 Explication détaillée de chaque solution proposée

1. Collecte des données : SQLite Studio

- Les données ont été fournies au format SQL, ce qui rend SQLite Studio parfaitement adapté.
- C'est un outil gratuit et facile à utiliser, ne nécessitant pas de configuration serveur complexe.
- Il permet une exploration rapide et efficace des données SQL.
- Pour un projet de petite à moyenne taille, il offre une solution légère et efficace.

2. Traitement des données : Power Query

- Il est intégré à Excel et Power BI, des outils largement utilisés en entreprise.
- Son interface utilisateur intuitive facilite son adoption par des utilisateurs non techniques.
- Il offre des capacités puissantes de transformation et de nettoyage des données.
- L'intégration transparente avec l'écosystème Microsoft permet une transition fluide entre la préparation des données et leur analyse.

3. Analyse des données : Power BI

- Il s'intègre parfaitement avec Power Query, assurant un flux de travail cohérent du traitement à l'analyse.
- Il offre une large gamme de visualisations interactives et personnalisables.
- Son interface intuitive permet aux utilisateurs de créer facilement des rapports et des tableaux de bord.
- Il bénéficie de mises à jour fréquentes et d'un support solide de Microsoft.
- Ses capacités de modélisation des données avec DAX permettent des analyses complexes.

Cette combinaison d'outils (SQLite Studio, Power Query, Power BI) offre une solution cohérente et intégrée pour répondre aux besoins de collecte, de traitement et d'analyse des données. Elle permet un flux de travail fluide, de l'extraction initiale des données à la création de visualisations et de rapports interactifs, tout en restant accessible aux utilisateurs ayant différents niveaux d'expertise technique.

5.2 Alignement des solutions avec les besoins identifiés

Intégration et cohérence :

La combinaison SQLite Studio, Power Query et Power BI offre une solution intégrée et cohérente.

Cela facilite le flux de travail de la collecte à l'analyse des données.

Flexibilité et évolutivité :

Ces outils peuvent s'adapter à l'évolution des besoins de l'entreprise.

Power BI, en particulier, offre des options d'évolutivité pour gérer des volumes de données croissants.

Sécurité des données :

Power BI propose des fonctionnalités de sécurité robustes, essentielles pour la protection des données sensibles.

Collaboration :

Power BI facilite le partage et la collaboration autour des rapports et tableaux de bord

Coût-efficacité :

La combinaison d'outils gratuits (SQLite Studio) et abordables (Power BI) optimise le rapport coût-bénéfice.

Support et formation :

L'écosystème Microsoft offre de nombreuses ressources de formation et un support solide.

- 5.3 Avantages et limites de chaque approche

La solution ETL SQLite/Power Query et PowerBi offre le meilleur équilibre entre flexibilité, sécurité et facilité d'utilisation.

Avantages :

Intégration transparente : Power Query s'intègre parfaitement avec Power BI, facilitant le flux de travail de l'ETL à la visualisation.

Automatisation : Power Query permet d'automatiser les tâches de préparation des données, libérant du temps pour l'analyse.

Flexibilité des sources de données : Power Query peut se connecter à une grande variété de sources de données, au-delà de SQLite.

Limites :

Performances : Pour de très grands ensembles de données, les performances de Power Query peuvent être limitées comparées à des solutions ETL dédiées.

Gouvernance des données : La séparation entre le modèle de données et les rapports peut être difficile à maintenir, ce qui peut poser des problèmes de gouvernance à long terme.

Évolutivité : Pour des projets plus complexes ou à plus grande échelle, cette approche pourrait nécessiter une transition vers des solutions ETL plus robustes.

Il est important de noter que bien que cette solution offre un bon équilibre pour de nombreux cas d'utilisation, elle peut nécessiter une réévaluation à mesure que les besoins de l'entreprise évoluent et que la complexité des données augmente.

Conclusion

- Synthèse des principales conclusions de l'analyse

L'analyse a révélé que la meilleure approche pour Bottleneck est d'utiliser une solution ETL combinée à un outil de visualisation, spécifiquement SQLite Studio, Power Query et Power BI.

Cette combinaison offre un excellent équilibre entre flexibilité, sécurité et facilité d'utilisation pour gérer les données complexes de l'entreprise tout en fournissant une interface conviviale pour les utilisateurs finaux.

Plus précisément :

1. **SQLite Studio** servira de base de données légère et efficace, idéale pour stocker et gérer les données localement.

2. **Power Query** s'intègre parfaitement avec Power BI, facilitant le flux de travail de l'ETL à la visualisation. Il permet d'automatiser les tâches de préparation des données, libérant du temps pour l'analyse.
3. **Power BI** offre des capacités puissantes de visualisation et d'analyse des données, avec une interface intuitive pour créer des rapports et des tableaux de bord interactifs.

L'approche choisie permet une gestion flexible des données, offrant la possibilité de réaliser des tâches complexes sans dépendre excessivement de l'équipe IT.

Cette solution permettra à Bottleneck de bénéficier d'une approche agile, évolutive et conviviale pour répondre à ses besoins en matière de gestion et d'analyse de données, tout en optimisant le temps et les ressources de l'équipe.

- Recommandations pour la mise en œuvre des solutions proposées

Pour la mise en œuvre des solutions proposées (SQLite Studio, Power Query et Power BI), voici quelques recommandations clés :

Préparation et configuration

Installez les outils nécessaires : SQLite Studio, Power BI Desktop

Utilisez l'Éditeur Power Query pour nettoyer et transformer efficacement les données avant de les charger dans Power BI

Modélisation et visualisation

Créez des mesures et des colonnes calculées dans Power BI pour faciliter l'analyse et la visualisation des KPI.

Maintenance et évolutivité

Documentez soigneusement les processus ETL et les transformations effectuées dans Power Query.

Mettez en place un processus de mise à jour régulière des données, en utilisant les fonctionnalités d'actualisation automatique de Power BI

Formation et support

Formez les utilisateurs finaux à l'utilisation de Power BI pour l'exploration et l'analyse des données.

Mettez en place un support technique pour résoudre rapidement les problèmes éventuels liés à la connexion SQLite Studio ou aux transformations Power Query.

[Annexes]

- Liste des sources de données utilisées

Base de données SQLite (bottleneck_2022_2023.db)

Schéma relationnel de la BDD

Dictionnaire de données

- Description détaillée des outils recommandés

Power Query : Outil ETL open-source avec interface graphique

Power BI : Outil de BI gratuit de Microsoft

- Exemples de visualisations de données proposées

Tableau de bord des ventes par région et par produit

Analyse de l'efficacité des promotions

Suivi des KPIs financiers (marge brute, ROI)

Analyse du parcours client sur le site web