|  | **CAHIER DES CHARGES** |
| --- | --- |

| Objet : |  |
| --- | --- |

| RÉVISIONS | | | |
| --- | --- | --- | --- |
| Version | Date | Rédacteur | Nature |
| 1.0 | 01/10/2024 | A.KADMIRY | Initialisation du document |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Table des matières**

[**1 Introduction 3**](#_qsyposuchv3x)

[1.1 Contexte 3](#_3znysh7)

[**2 Les données 4**](#_7rkjly970dlm)

[2.1 Fichier Ticket 4](#_2jxsxqh)

[2.2 Fichier Article 4](#_z337ya)

[**3 Les livrables 5**](#_rtp9e57hvm5t)

[3.1 Application 5](#_u701wu5x4msx)

[3.2 Documentation 5](#_1t7igfdyqpq7)

[**4 Les attentes 6**](#_8aqgiwtvohua)

[4.1 Validation des données 6](#_m4b4d6x9clwi)

[4.2 Tableaux de bord et indicateurs 6](#_2bswstafg9i0)

[4.3 Préconisations 7](#_jc9uww5i2wo0)

# **Introduction**

## **Contexte**

Dans le cadre de l’expansion de ses activités à l’international, le groupe souhaite mener une analyse approfondie des ventes réalisées dans ses différents magasins à travers le monde. Pour atteindre cet objectif, il envisage de tirer parti des données contenues dans les tickets de caisse, une source précieuse d’informations sur les transactions réalisées.

Vous êtes donc chargé, à partir d’un premier échantillon de ces données, de concevoir un datawarehouse destiné à centraliser et structurer les informations de vente. Ce datawarehouse permettra de faciliter l’analyse des données selon divers axes (géographique, temporel, type de produit, etc.), offrant ainsi une vision détaillée et segmentée des performances commerciales.

Cependant, une partie des données étant saisie manuellement, la solution devra intégrer des mécanismes pour réduire le risque d’erreur et garantir la fiabilité des informations analysées.

# **Les données**

Pour mettre en place un processus d’intégration complet, il est nécessaire d’utiliser les 2 fichiers suivants :

## **Fichier Ticket**

Pour l’exploitation des données présentes dans ce fichier de tickets de caisse, voici les éléments clés de préparation :

* **Contenu** : Ce fichier contient le détail de chaque ticket de caisse, incluant notamment la date d’achat, le code article, la quantité, le prix, et d’autres informations transactionnelles.
* **Colonnes** : Le fichier comporte 33 colonnes, couvrant potentiellement des aspects variés de la transaction
* **En-tête** : Le fichier contient une ligne d’en-tête, qui fournit les noms des colonnes et facilite l’identification des champs.
* **Format** : Le fichier est au format CSV, mais avec un séparateur de champs spécifique, le caractère “|”, qui différencie chaque colonne.

## **Fichier Article**

Pour l’exploitation des données présentes dans ce fichier de référentiel article, voici les éléments clés de préparation :

* **Contenu** : Ce fichier contient le détail des articles vendu, incluant notamment le libellé du produit, la catégorie, le genre et d’autres informations.
* **Colonnes** : Le fichier comporte 13 colonnes
* **En-tête** : Le fichier contient une ligne d’en-tête, qui fournit les noms des colonnes et facilite l’identification des champs.
* **Format** : Le fichier est au format CSV, mais avec un séparateur de champs spécifique, le caractère “;”, qui différencie chaque colonne.

# **Les livrables**

Pour répondre aux exigences de ce projet, les livrables attendus sont organisés en deux catégories : l’application et la documentation. Voici le détail des livrables nécessaires :

## **Application**

1. **Datawarehouse (DWH)** : Un entrepôt de données structuré pour centraliser et organiser les informations de ventes, avec une architecture permettant des analyses multidimensionnelles (produit, géographie, temps, etc.).
2. **Process d’alimentation** : Un ensemble de procédures ETL (extraction, transformation, chargement) pour l’alimentation du DWH. Ce processus inclura des étapes de validation des données, garantissant leur qualité (complétude, exactitude, et format conforme) avant leur intégration.
3. **Tableau de bord (indicateurs)** : Conception et développement de tableaux de bord interactif qui affichera les indicateurs clés de performance

Ces tableaux de bord comprendront des visualisations adaptées (graphiques en barres, courbes, diagrammes de répartition, etc.) pour faciliter la compréhension et la prise de décision.

## **Documentation**

### 

1. **Spécifications fonctionnelles** : Un document décrivant en détail l’architecture du DWH, les technologies utilisées, les fonctionnalités attendues du système et les règles de gestion appliquées à l'alimentation du DWH et des tableaux de bord.
2. **Spécifications techniques** : Une documentation technique expliquant les choix techniques pour le processus d’alimentation (ETL), les mappings utilisés ainsi que les modalités de gestion des erreurs et contrôles de qualité.
3. **Code SQL de création du DWH** : Le code SQL détaillé pour créer et structurer le DWH, incluant la création des tables, des relations entre les tables, et toute autre configuration nécessaire au bon fonctionnement de l’entrepôt de données.
4. **Cahier de recette** : Un document de validation destiné aux tests d'alimentation du DWH et des tableaux de bord. Il inclura des scénarios de tests, des critères de validation pour chaque fonctionnalité, et des cas de tests pour le contrôle de la qualité des données.
5. **Planning de réalisation** : Un calendrier précis pour chaque étape du projet, incluant les délais pour la création du DWH, le développement des processus d’alimentation, la création des tableaux de bord, et la documentation.

Ces livrables fourniront une solution complète, répondant aux besoins d’analyse des ventes et assurant une documentation exhaustive pour une maintenance future efficace.

# **Les attentes**

Pour mettre en œuvre une solution d’analyse des ventes efficace et évolutive, voici les attentes essentielles du projet :

## **Validation des données**

### 

* **Vérification des formats de données** : Avant chaque chargement dans le *datawarehouse*, il sera essentiel de vérifier que les données respectent les formats attendus. Cela inclut le bon formatage des montants, des dates, des identifiants de produits et autres champs critiques.
* **Contrôle de complétude et exactitude** : Les données ne seront chargées que lorsqu’elles auront passé un contrôle de qualité, confirmant qu’elles sont complètes, cohérentes, et exemptes d’erreurs. Ce processus garantira des analyses fiables et exploitables en aval.

## **Tableaux de bord et indicateurs**

### 

Pour fournir une vision précise et exploitable des performances commerciales, plusieurs tableaux de bord seront développés, chacun accompagné de visualisations claires pour faciliter l’interprétation des données. Les indicateurs à inclure sont :

1. **Chiffre d'affaires en Euro** : Présentation du chiffre d'affaires à la fois *toutes taxes comprises* (TTC) et *hors taxes* (HT) afin d’analyser les ventes de manière exhaustive.
2. **Panier moyen** : Calcul et suivi du montant moyen par transaction, permettant de mesurer la valeur moyenne des ventes réalisées et de suivre les variations dans le temps.
3. **Top 10 des meilleures ventes** : Un classement des produits les plus vendus, permettant d'identifier les articles générant le plus de revenus et orienter la gestion des stocks et la stratégie commerciale.
4. **Top 10 des meilleurs magasins** : Un tableau de bord pour les magasins les plus performants permettra d’analyser les réussites géographiques et de focaliser les efforts d’optimisation des ventes.
5. **Autres indicateurs** : Des indicateurs additionnels pourront être proposés en fonction des besoins de l’analyse et des spécificités commerciales.

## **Préconisations**

À mesure que le volume des données augmentera, il sera nécessaire de garantir que la solution reste performante et évolutive. Les recommandations à venir porteront sur :

* **L'optimisation de l’architecture** pour supporter un flux de données en croissance constante.
* **L’automatisation des processus de vérification et de chargement** afin de réduire les délais et les risques d'erreurs.

En mettant en place ces indicateurs et en anticipant les besoins futurs, le groupe disposera d’un système d’analyse robuste et évolutif, à même de répondre aux enjeux stratégiques de son développement commercial.