

# Documentation Projet : Optimisation de la Gestion des Données – BottleNeck



## 1. Contexte du Projet

En tant que Data Analyst pour BottleNeck, un marchand de vins prestigieux, j'ai été missionné pour moderniser des outils d'analyse jusqu'ici artisanaux. La problématique centrale résidait dans le cloisonnement des données entre l'outil de gestion commerciale (ERP) et la plateforme de vente en ligne (CMS **Wordpress**).

L'objectif était double :

1. Réaliser l'unification des flux de données (Data Mashup).
2. Produire des indicateurs de performance (KPIs) pour le Comité de Direction (CODIR).

## 2. Environnement Technique

- **Langage** : Python 3
- **Interface** : Jupyter Notebook
- **Bibliothèques clés** :
  - Pandas : Pour la manipulation, le nettoyage et les jointures de données.
  - Matplotlib & Seaborn : Pour la détection visuelle des anomalies (Boxplots) et les corrélations.
  - Plotly : Pour la création de graphiques interactifs (Top ventes, Stocks).

## 3. Phase 1 : Audit et Qualité des

Une étape cruciale de "Data Cleaning" a été menée pour corriger plus de 8 types d'erreurs identifiées :

- **Erreurs de structure** : Identifiants clés différents (product\_id vs sku/id\_web) nécessitant une table de liaison.
- **Champs orphelins** : 91 références présentes dans une table mais absentes de l'autre.
- **Doublons** : Identification et suppression des doublons d'IDs web.
- **Valeurs manquantes (NaN)** : Gestion des cellules vides pour les produits non publiés en ligne.
- **Anomalies de Stock** : Détection de stocks négatifs.
- **Incohérences de prix** : Identification de produits dont le prix d'achat était supérieur au prix de vente.
- **Problèmes de formatage** : Conversion des types de données pour les calculs (numérique vs objet).
- **Erreurs de statut** : Incohérences entre l'affichage "En stock" et la quantité réelle à 0.

**Solutions proposées** : Automatisation des imports via une clé unique partagée dès la saisie et mise en place de masques de saisie pour éviter les prix ou stocks négatifs.

### 3. Phase 2 : Analyses Stratégiques et Performance

#### Analyse du Chiffre d'Affaires (CA)

- **Calcul par produit** : Multiplication des quantités vendues par le prix unitaire.
- **CA Total** : Agrégation globale pour une vision macro de la performance du site.
- **Top Références (Loi de Pareto)** : Analyse 20/80 pour identifier les produits générant l'essentiel du revenu.

#### Analyse des Prix et Valeurs Aberrantes (Outliers)

Utilisation de la méthode de l'**Écart Interquartile (IQR)** et du **Z-Score** pour isoler les prix atypiques :

- Visualisation via un **Boxplot** pour mettre en évidence les vins de luxe ou les erreurs de saisie potentielles.
- Conclusion : Confirmation que la majorité des prix élevés correspondent à des vins de prestige et non à des erreurs système.

## Gestion des Stocks et Marge

- **Rotation des stocks** : Calcul du nombre de mois de couverture pour identifier le "stock dormant".
- **Analyse des marges** : Étude du taux de marge par catégorie de produit.
- **Alerte Stock** : Création d'un Top 5 des produits avec le plus de stock dormant (en mois) pour orienter les actions de déstockage marketing.

## Analyse de Corrélation

Étude des relations entre les variables quantitatives :

- **Prix vs Ventes** : Existe-t-il une élasticité-prix marquée ?
- **Stock vs Ventes** : Optimisation du réapprovisionnement.
- **Résultat** : Mise en évidence des leviers de croissance (ex : corrélation entre prix élevé et faible rotation, mais forte marge unitaire).

## 4. Livrables inclus

1. **NoteBook.ipynb** : Code source complet documenté, incluant les jointures (merge), le nettoyage et les calculs.
2. **Présentation.pdf** : Support visuel synthétique destiné au CODIR avec graphiques d'aide à la décision.
3. **Jeux de données** : ERP, Web et Table de liaison.

## Recommandations Stratégiques

Sur la base de cette analyse, je préconise à **BottleNeck** d'investir dans l'harmonisation de ses systèmes d'information (API entre ERP et Web) et de lancer des campagnes ciblées sur les produits à forte marge mais à faible rotation de stock pour optimiser la trésorerie.