

Rapport d'Anomalies des Données - Projet BottleNeck

Date : 28 mai 2024

Auteur : Flavien BARDET

Contexte : Ce rapport synthétise les anomalies et incohérences de données identifiées lors de l'analyse croisée des fichiers ERP (erp.xlsx), Web (web.xlsx) et Liaison (liaison.xlsx) de BottleNeck. L'objectif est de fournir une vision claire des problèmes de qualité des données pour prioriser les corrections et améliorer la fiabilité des analyses commerciales.

F par Flavien Bardet

Synthèse Exécutive

L'analyse a révélé plusieurs types d'anomalies impactant la cohérence et la fiabilité des données entre les systèmes ERP et Web. Celles-ci incluent des incohérences internes à l'ERP (statuts/quantités de stock, données négatives), des problèmes de qualité et de format dans les données Web et de liaison (identifiants manquants ou non standards), des erreurs critiques de liaison entre les systèmes (produits 'online' sans lien, produits 'offline' avec lien), et des erreurs de données spécifiques ayant un impact direct sur les indicateurs clés (marge négative). Ces anomalies affectent la capacité à obtenir une vue unifiée et précise du catalogue, des stocks et des performances, et nécessitent des actions correctives ciblées. Environ **111 produits ERP (13%)** n'ont pu être inclus dans l'analyse finale croisée en raison de ces problèmes.

Détail des Anomalies Constatées (Partie 1)

#	Anomalie	Source(s)	Identification / Exemples	Cause(s) Probable(s)
1	Incohérence statut de Stock vs Quantité	ERP	2 produits (ID: 4039 stock=3/outofstock, 4885 stock=0/instock).	Erreur de saisie manuelle du statut. Désynchronisation lors de mise à jour (statut non recalculé après mouvement de stock).
2	Prix de vente négatifs ou nuls	ERP	3 produits (ID: 4233 prix=-20, 5017 prix=-8, 6594 prix=-9.1).	Erreur de saisie, d'import/export de données. Valeur par défaut incorrecte lors de la création produit.
3	Quantités de stock négatives	ERP	2 produits (ID: 4973 stock=-10, 5700 stock=-1).	Erreur de saisie lors d'un inventaire ou réception. Vente enregistrée sans stock physique (gestion des commandes/réservations). Problème de synchronisation (vente traitée avant réception).
4	Identifiants sku manquants (NaN) dans les Données Web	Web	85 lignes sans sku. 83 d'entre elles sont quasi-vides.	Artefacts d'export Web (lignes techniques, tests). Produits Web spécifiques non liés à l'ERP (ex: services, anciens produits). Erreur lors de la création du produit sur le site Web.

Détail des anomalies constatées (Partie 2)

#	Anomalie	Source(s)	Identification / Exemples	Cause(s) Probable(s)
5	Formats sku / id_web Non Standards	Web, Liaison	sku Web: '13127-1', 'bon-cadeau-25-euros'. id_web Liaison: '14680-1', 'bon-cadeau-25-euros'.	Produits spéciaux (bons cadeaux) utilisant un code spécifique. Erreur de saisie ou convention de nommage. Identifiant technique (ex: variation produit).
6	Produits ERP 'Online' sans lien Web (id_web NaN)	ERP vs Liaison	3 produits (ID: 4594, 5070, 5075). Note: Stock = 0 pour ces produits.	Oubli création/MàJ lien dans liaison lors activation web. Processus désynchronisé (ERP activé avant lien). Règle métier implicite ? Lien supprimé/non créé si stock = 0 ?
7	Produits ERP 'Offline' avec lien Web (id_web Existant)	ERP vs Liaison	21 produits (ex: 4200, 4289, 4568, incluant '14680-1').	Oubli de désactivation web (onsale_web=0) après suppression du lien. Désynchronisation inverse (produit retiré du web mais pas marqué offline dans ERP). Erreur statut ERP.

Détail des anomalies constatées (Partie 3)

#	Anomalie	Source(s)	Identification / Exemples	Cause(s) Probable(s)
8	Produit ERP 'Offline' présent dans l'analyse Finale	Fusion (df_final)	1 produit (ID: 4200, sku: 16295). Génère 81.20€ de CA.	Conséquence directe de l'anomalie #7 : Le produit 4200 (offline avec lien) a trouvé une correspondance sku valide dans les données web nettoyées.
9	Marge négative aberrante	Calcul Marge	1 produit (ID: 4355) avec taux de marge de -635%. Prix vente TTC = 12.65€, Prix Achat = 77.48€.	Erreur critique sur le Prix de Vente dans l'ERP. Moins probable : Erreur critique sur le Prix d'Achat.
10	Outliers de Prix élevés (Observation)	Analyse Prix	31 produits avec prix > 84.09€ (seuil IQR). Ex: ID 4352 (225€), 5001 (217.5€).	Produits de luxe/rares/grands formats (Champagne millésimé, Grand Cru, Cognac XO). Possible erreur de prix ponctuelle.

Actions préconisées pour les anomalies critiques



Prix de Vente négatifs ou nuls (criticité: Élevée)

1. Correction Données : Corriger immédiatement les prix erronés dans l'ERP.
2. Prévention ERP : Ajouter une règle de validation interdisant les prix ≤ 0 .
3. Audit Processus : Vérifier les interfaces d'import/export.



Quantités de stock négatives (criticité: Élevée)

1. Correction données : Réaliser un inventaire physique pour ces produits et corriger le stock dans l'ERP (probablement à 0 si rupture).
2. Prévention ERP : Bloquer les transactions (ventes) si stock $<$ quantité demandée (selon politique commerciale).
3. Analyse Processus : Revoir le flux logistique et sa traduction dans l'ERP.



Produits ERP 'Online' sans lien web (criticité: Élevée)

1. Investiguer les règles métiers : Confirmer si l'absence de lien pour stock=0 est normale ou une erreur.
2. Correction (si erreur) : Créer les liens manquants dans liaison.xlsx.
3. Processus : Assurer la création du lien en même temps que l'activation web (onsale_web=1).



Marge négative aberrante (criticité: Critique)

1. Correction Urgente Données : Vérifier et corriger immédiatement le prix de vente (et/ou achat) du produit 4355 dans l'ERP.
2. Prévention : Mettre en place une alerte si marge $<$ seuil minimum acceptable (ex: 0% ou 5%).

Actions préconisées pour les anomalies de priorité moyenne

Incohérence statut de Stock vs Quantité

1. Système ERP : Implémenter une mise à jour automatique du stock_status basée sur stock_quantity.
2. Processus : Vérifier périodiquement la cohérence statut/quantité.



Identifiants sku manquants (NaN)

1. Nettoyage données : Confirmer la suppression des 83 lignes vides.
2. Investigation : Analyser les 2 lignes avec données partielles pour identifier/corriger si possible.
3. Processus web : Rendre le champ SKU obligatoire pour les 'products' sur le site.

Produits ERP 'Offline' avec lien Web

1. Audit cohérence : Vérifier le statut réel de ces 21 produits sur le site web (via leur id_web).
2. Correction données : Aligner le statut onsale_web dans l'ERP avec la réalité (mettre à 1 si actif) OU supprimer le lien dans liaison si réellement inactif.

Conclusion et impact Général

Les anomalies identifiées, notamment les problèmes de liaison (anomalies 6, 7, 8) et les erreurs de données critiques (prix négatifs/nuls, stocks négatifs, marge aberrante), limitent la capacité à effectuer des analyses fiables et complètes sur l'ensemble du catalogue. Elles peuvent conduire à :



Une sous-estimation/surestimation du CA et des marges "online".



Une vision incorrecte de la disponibilité réelle des produits en ligne.



Des décisions potentiellement erronées sur la gestion des stocks et la politique de prix.

Il est recommandé de traiter en priorité les anomalies à criticité "Élevée" et "Critique" pour restaurer l'intégrité des données de base, puis d'adresser les problèmes de cohérence et de processus pour une amélioration durable.