

Catalogue de Données

Imad Elmanser

Avril 2025

Introduction

Ce document présente le catalogue de données extrait de deux sources principales :

- **orders.csv** : Données relatives aux commandes des clients.
- **logs.json** : Logs des événements utilisateurs sur la plateforme.

Chaque table est décrite en détail avec les métadonnées suivantes : nom du champ, type, description, et présence éventuelle de valeurs manquantes.

1. Table : orders.csv

Description : Cette table contient l'historique des commandes passées par les clients.

Nombre de lignes : 100

Nombre de colonnes : 7

Clé primaire potentielle : order_id

Clé étrangère potentielle : customer_id

Nom du champ	Type	Description	Valeurs man-quant
order_id	UUID	Identifiant unique de la commande	0
customer_id	UUID	Identifiant du client ayant passé la commande	0
order_date	Date	Date de la commande	0
status	Texte	Statut de la commande (DELIVERED, CANCELLED, etc.)	0
product_id	UUID	Identifiant du produit commandé	0
quantity	Entier	Quantité commandée	0
price	Décimal	Prix total de la commande	0

2. Table : logs.json

Description : Cette table contient les enregistrements des interactions des utilisateurs sur le site.

Nombre de lignes : 200

Nombre de colonnes : 4

Clé primaire : event_id
Clé étrangère potentielle : user_id

Nom du champ	Type	Description	Valeurs manquantes
event_id	UUID	Identifiant unique de l'événement	0
user_id	UUID	Identifiant de l'utilisateur	0
event_type	Texte	Type d'événement (click, view, purchase)	0
timestamp	Date/Heure	Date et heure de l'événement	0

3. Relations entre les tables

- Le champ customer_id de la table des commandes peut être relié au champ user_id des logs.
- Il est donc possible de faire des analyses croisant les achats et les comportements utilisateurs.

4. Métadonnées globales

- **Responsable des données :** Équipe projet PFE 2025
- **Date de collecte :** Avril 2025
- **Origine :** Interne (simulation projet GED)
- **Fréquence de mise à jour :** Hebdomadaire