

# MAPPING COMPLET

SmartMarket\_raw.xlsx → Data Warehouse

FEUILLE 1 : TRANSACTIONS (100 lignes × 11 colonnes)

N°	Colonne Excel	Type	Table DW	Champ DW	Transformation Requise
1	Transaction_ID	object	FAIT_VENTES	ID_Vente	Nettoyer doublons (95 uniques/100)
2	Date	object	DIM_TEMPS	Date_Complete	<b>CRITIQUE</b> : 99 formats différents → Standardiser ISO
3	Client_ID	object	DIM_CLIENT	ID_Client	42 clients uniques
4	Produit_ID	object	DIM_PRODUIT	ID_Produit	30 produits uniques
5	Quantite	int64	FAIT_VENTES	Quantite	Vérifier > 0 (min=1, max=9)
6	Remise_%	float64	FAIT_VENTES	Taux_Remise	<b>CRITIQUE</b> : Corriger -10% et 150% → [0-100]
7	Canal_Vente	object	DIM_CANAL	Nom_Canal	<b>CRITIQUE</b> : 'En ligne' vs 'en ligne' → Standardiser
8	Mode_Paiement	object	(optionnel)	-	4 valeurs : Virement, CB, cheque, Espèce
9	Statut_Livraison	object	FAIT_VENTES	Indicateur_Retard	Mapper : Livré→0, En attente→1, Annulé→NULL
10	Date_Livraison	object	FAIT_VENTES	Delai_Livraison_Jours	Calculer : Date_Livraison - Date
11	Revenu_Total	float64	FAIT_VENTES	Revenu_Reel	<b>15 valeurs NULL</b> → Recalculer

TABLE 1: Mapping Transactions

FEUILLE 2 : SATISFACTION (60 lignes × 4 colonnes)

N°	Colonne Excel	Type	Table DW	Champ DW	Transformation Requise
1	Client_ID	object	DIM_CLIENT	ID_Client	Jointure avec Transactions (39 clients)
2	Date_Enquete	object	DIM_TEMPS	Date_Complete	Standardiser format date
3	Score	object	FAIT_VENTES	Note_Satisfaction	<b>CRITIQUE</b> : 'Bon'→4, 'Très satisfait'→5, Numérique→OK
4	Commentaire	object	-	-	Analyse qualitative (optionnel)

TABLE 2: Mapping Satisfaction

FEUILLE 3 : FOURNISSEURS (10 lignes × 6 colonnes)

N°	Colonne Excel	Type	Table DW	Champ DW	Transformation
1	Fournisseur_ID	object	DIM.FOURNISSEUR	ID.Fournisseur	OK (10 fournisseurs)
2	Nom_Fournisseur	object	DIM.FOURNISSEUR	Nom_Fournisseur	OK (4 noms uniques)
3	Pays	object	DIM.FOURNISSEUR	Pays_Origine	'cn'→'CHN', 'France'→'FRA'
4	Evaluation	float64	DIM.FOURNISSEUR	Note_Fiabilite	Normaliser [3-10] → [1-5]
5	Délai_Moyen_Livraison	object	DIM.FOURNISSEUR	Delai_Moyen_Livraison	'non défini' → NULL
6	Fiabilité_%	float64	DIM.FOURNISSEUR	Taux_Retard_Pct	<b>CRITIQUE</b> : 120% → Corriger à 100% max

TABLE 3: Mapping Fournisseurs

## CHAMPS CALCULÉS

Champ Calculé	Formule	Source
Prix_Unitaire	$\text{Revenu\_Total} / (\text{Quantite} \times (1 - \text{Remise\_}\% / 100))$	Transactions (rétro-calculé)
Montant_Brut	$\text{Prix\_Unitaire} \times \text{Quantite}$	Transactions
Montant_Remise	$\text{Montant\_Brut} \times \text{Remise\_}\% / 100$	Transactions
Delai_Livraison_Jours	$\text{Date\_Livraison} - \text{Date}$ (en jours)	Transactions
Indicateur_Retard	IF (Statut_Livraison = 'En attente', 1, 0)	Transactions
Note_Fiabilite (normalisée)	$((\text{Evaluation} - 3) / 7) \times 4 + 1$	Fournisseurs

TABLE 4: Champs calculés lors de l'ETL

## COLONNES MANQUANTES (à obtenir)

Dimension	Colonnes Manquantes	Action Requise
DIM.CLIENT	Nom, Prénom, Email, Téléphone, Ville, Code_Postal, Date_Inscription	Créer feuille "Clients" ou obtenir source externe
DIM.PRODUIT	Nom_Produit, Categorie, Sous_Categorie, Prix_Catalogue, Marque	Créer feuille "Produits" ou obtenir source externe
DIM.REGION	Département, Région_Administrative	Enrichir depuis Code_Postal

TABLE 5: Données manquantes

## PROBLÈMES QUALITÉ DÉTECTÉS

Priorité	Colonne	Problème	Solution
<b>CRITIQUE</b>	Remise_%	Valeurs -10% et 150%	Clipper à [0, 100]
<b>CRITIQUE</b>	Canal_Vente	Incohérence casse	Normaliser à Title Case
<b>CRITIQUE</b>	Score	Valeurs texte + numérique	Mapper texte → numérique
<b>CRITIQUE</b>	Fiabilité_%	Valeur 120%	Clipper à max 100%
ÉLEVÉE	Date	99 formats différents	Parser avec date flexible
ÉLEVÉE	Revenu_Total	15 valeurs NULL	Recalculer si données complètes
MOYENNE	Pays	Codes incohérents	Dictionnaire de mapping
MOYENNE	Evaluation	Échelle 3-10	Normaliser sur [1-5]

TABLE 6: Problèmes qualité par priorité