

Présentation de notre projet **TALEND**



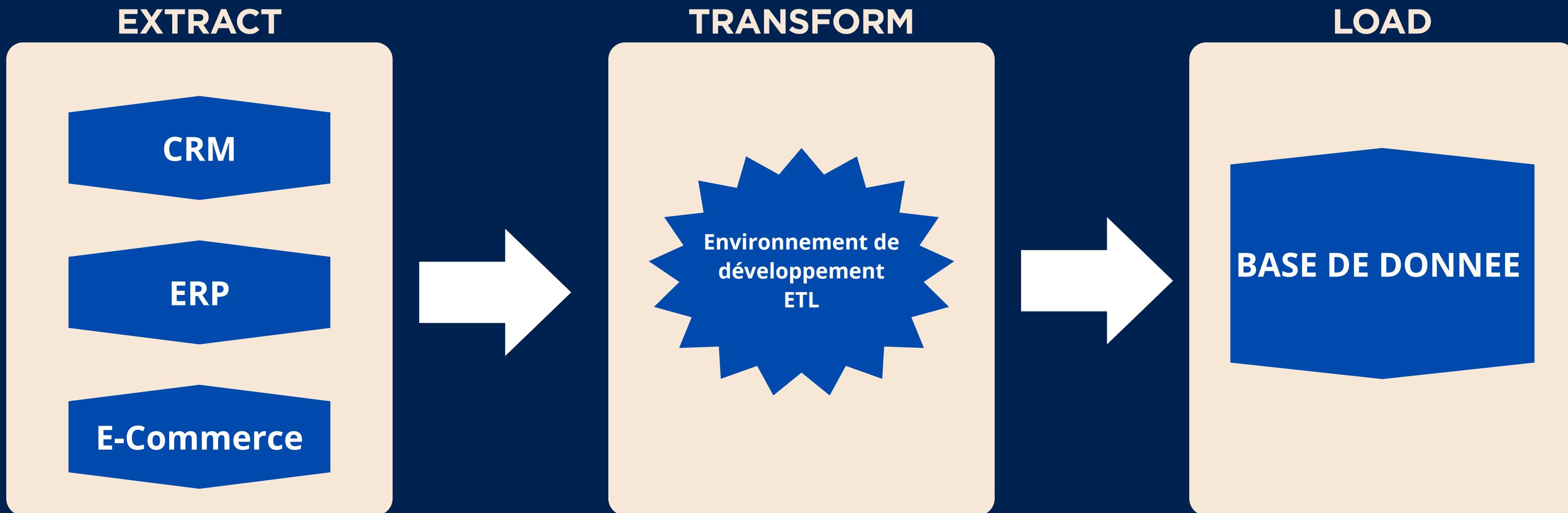
GROUPE 2

Qui sommes-nous ?	1
Contexte	2
Nos technologies	3
Import de données	4
Transformation des données	5
Jointures	6
Mise en base de données	7
Point d'améliorations	8

QUI SOMMES-NOUS ?

CONTEXTE

- Sources de données clients différentes
- Création d'un processus ETL
- Données consolidées et cohérentes pour les analyses



NOS TECHNOLOGIES



OUTIL ETL



BASE DE DONNÉES

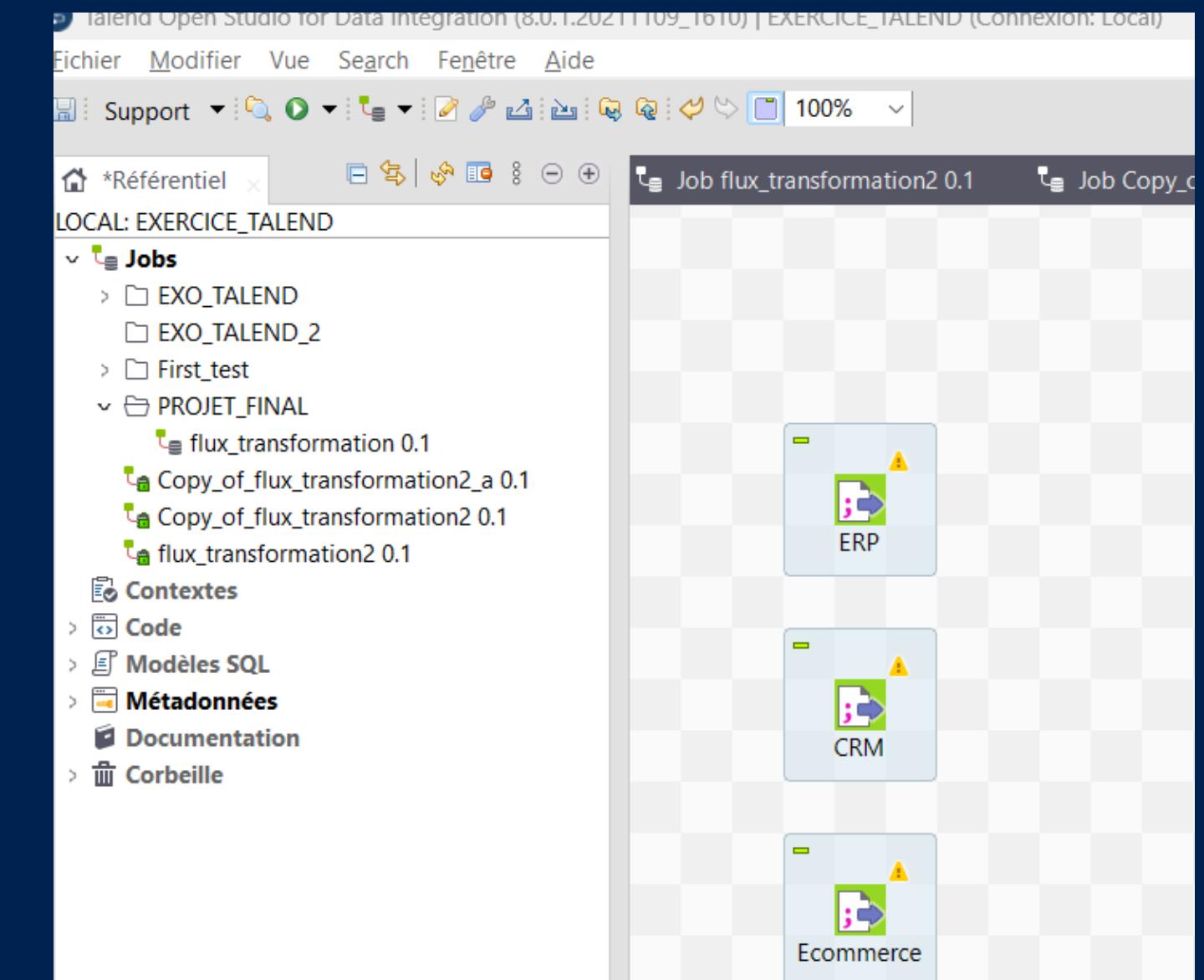
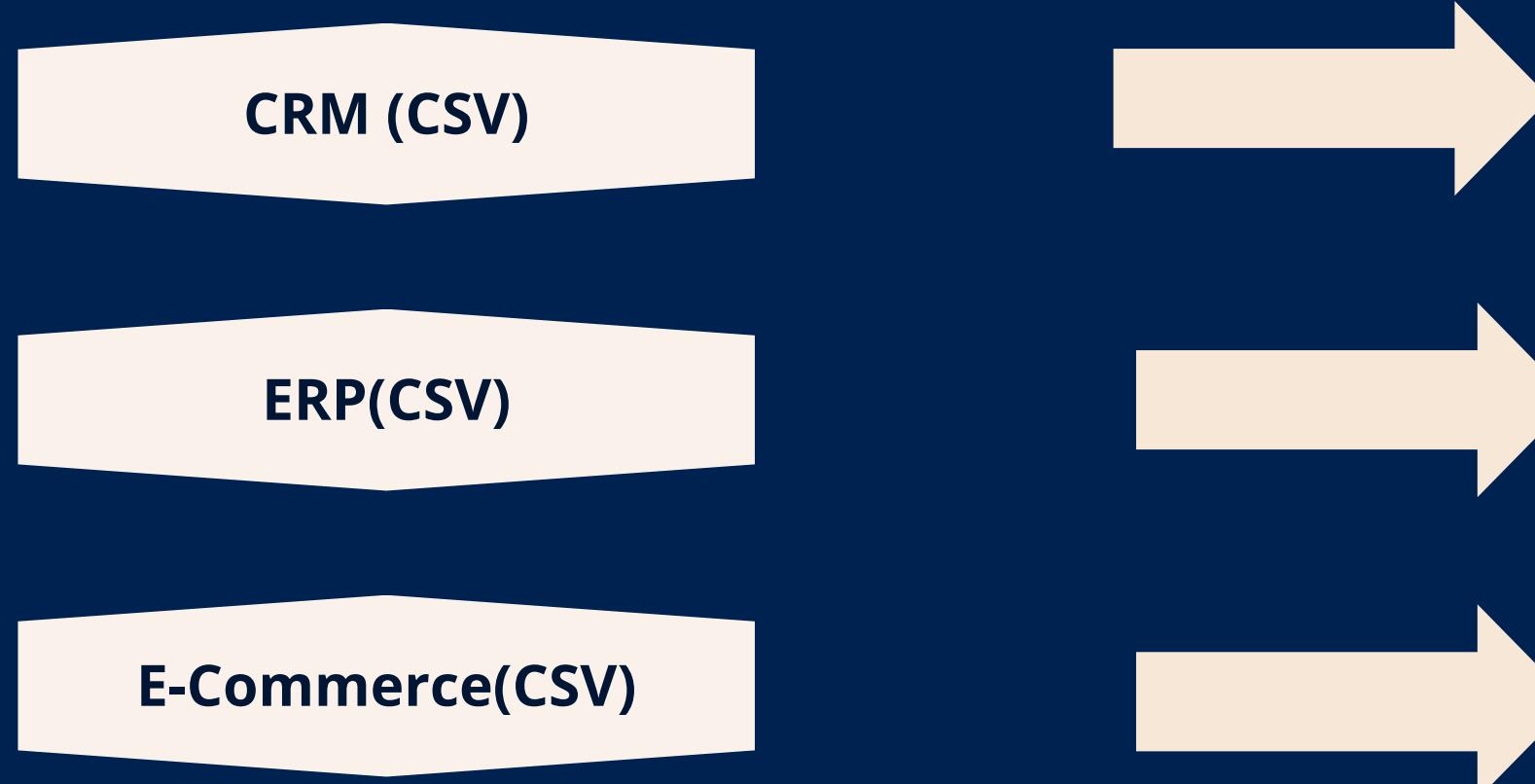


INTERFACE SGBD

IMPORT DE DONNEES

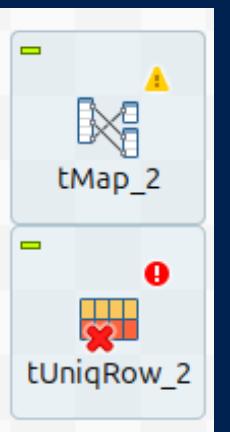
Préparation des flux d'entrée

Importation de données de chaque source sur Talend en format CSV.

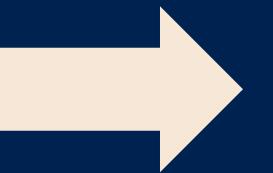
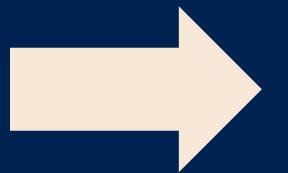


TRANSFORMATION DES DONNÉES

Composant principaux : tMap, tUniq



Données source



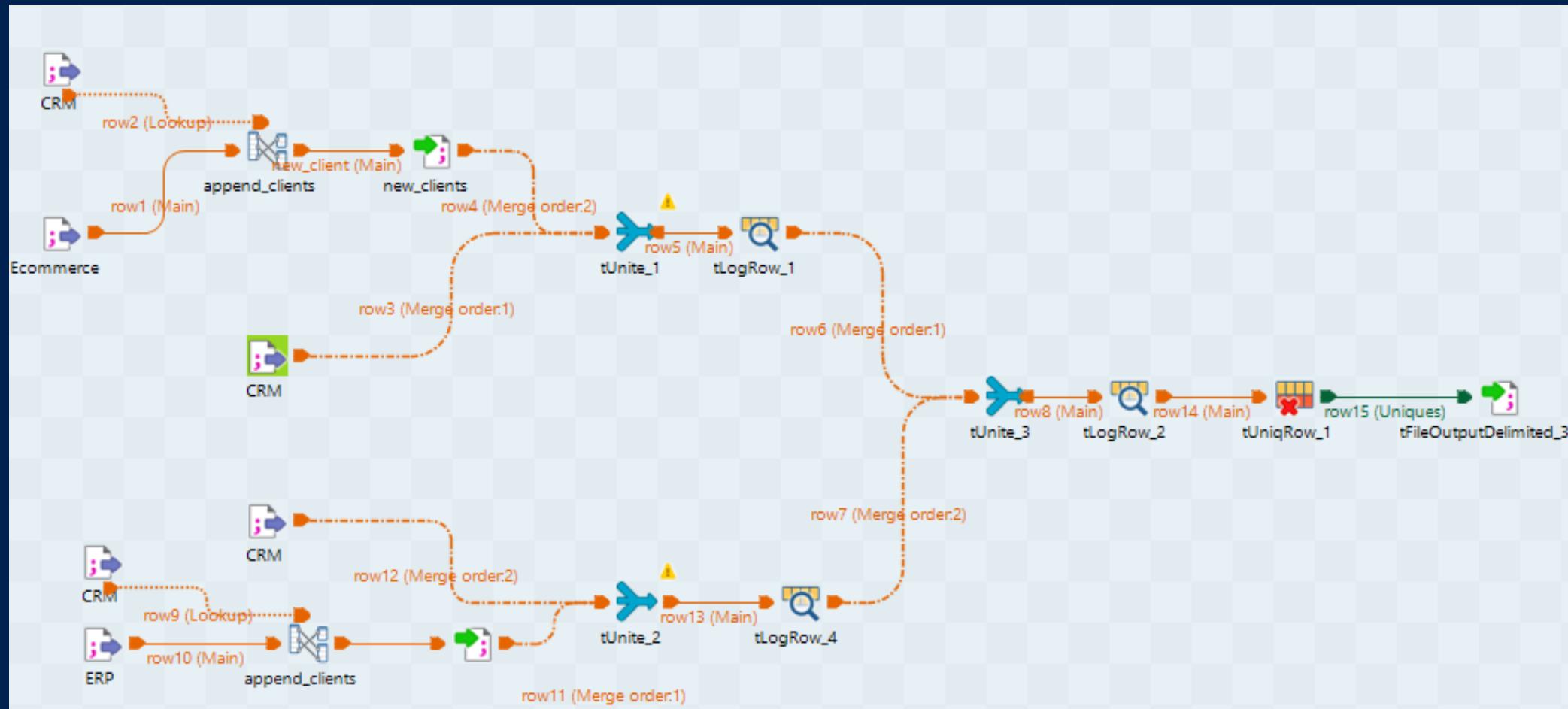
Données transformées
Bons types et bons formats

Expressions permettant de normaliser
au format voulu la donnée d'entrés

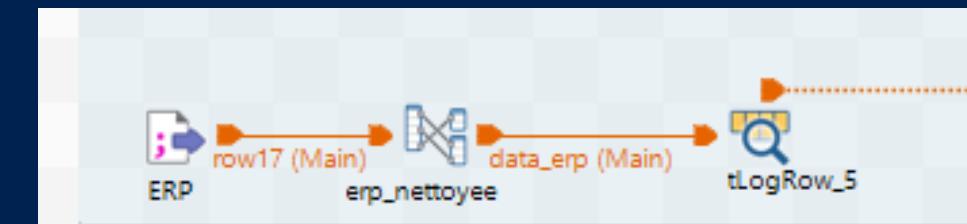


TRANSFORMATION DES DONNÉES

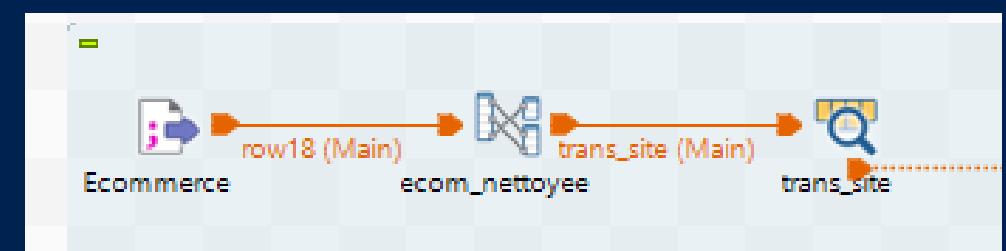
Flux de transformation d'un nouveau fichier clients



Flux de nettoyage du fichier ERP



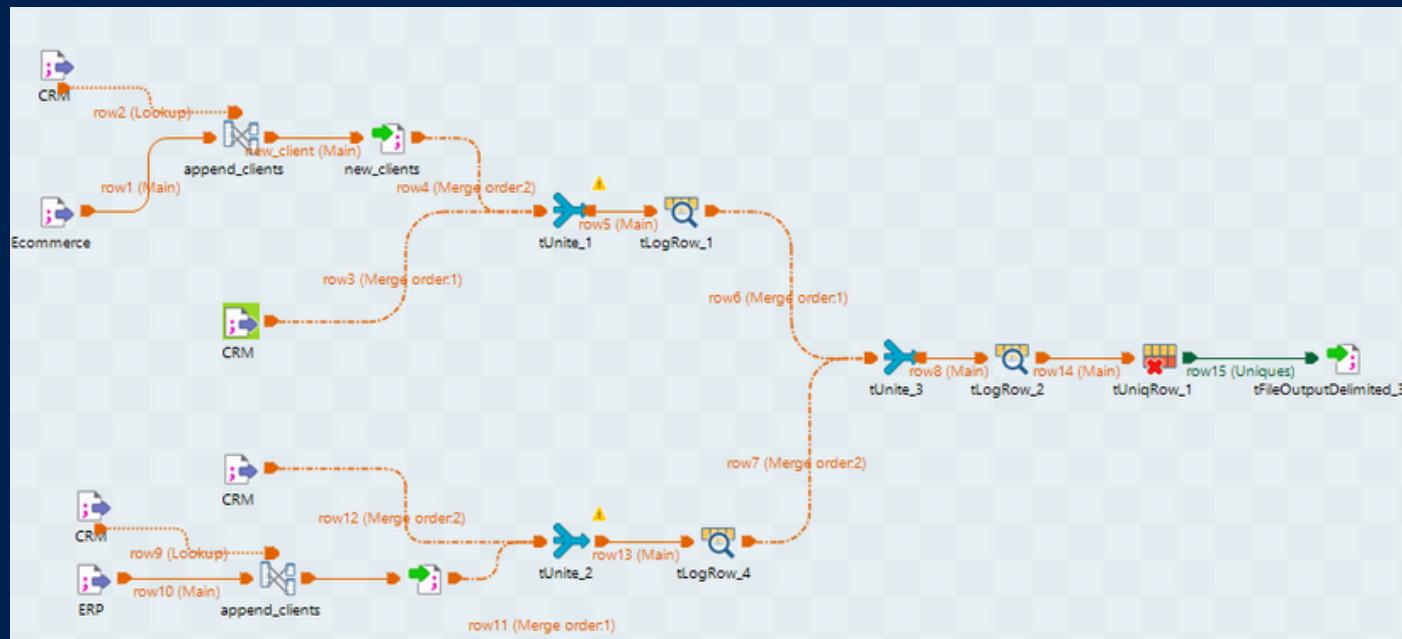
Flux de nettoyage du fichier Ecommerce



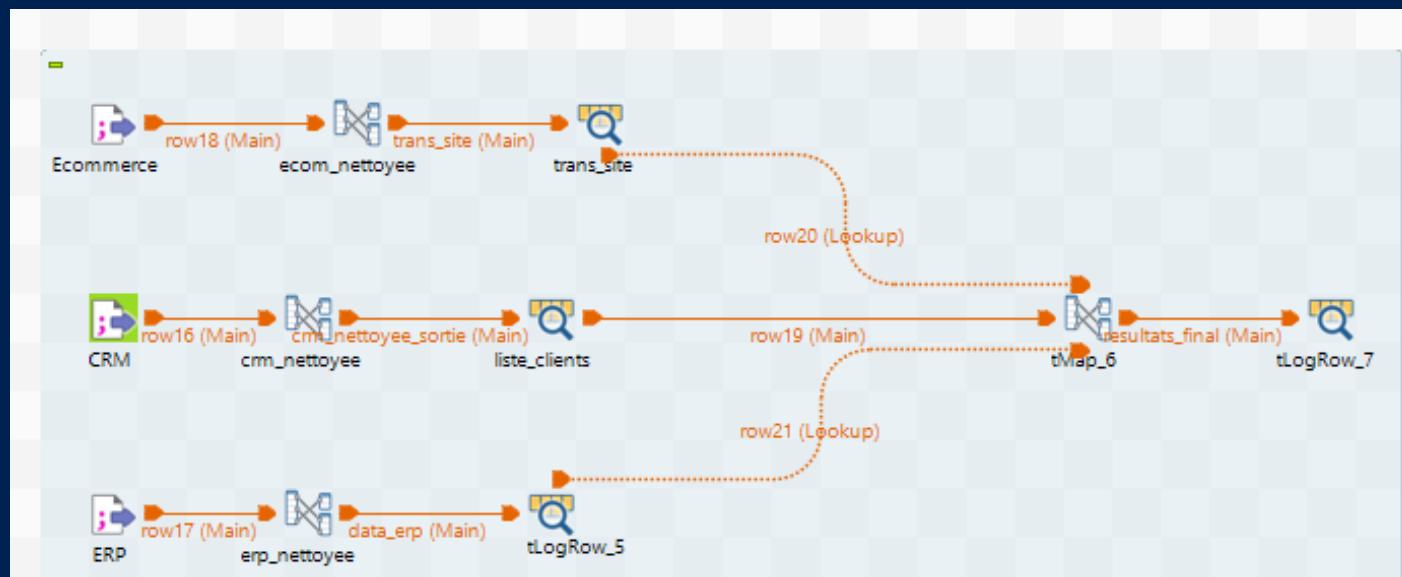
JOINTURES

Regroupement des données

Flux de création d'une nouvelle liste de clients en utilisant tMap et tUnité



Flux de jointure des tables avec tMap



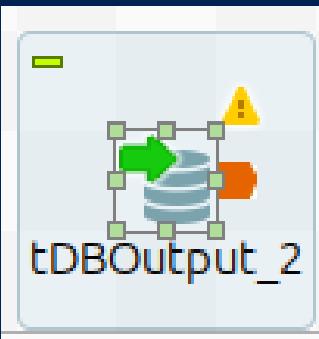
Un seul tableau final, complet et homogène, réunissant les données clients, ventes et facture

nom	email	date_inscription	revenu_annuel	purchase_date
Client_1	client1@gmail.com	06/03/2022	0.0	null
Client_2	client2@gmail.com	15/03/2022	35000.0	null
Client_3	client3@gmail.com	08/01/2023	35000.0	null
Client_4	client4@gmail.com	28/04/2022	28000.0	null
Client_5	client5@gmail.com	25/09/2022	42000.0	null
Client_6	client6@gmail.com	26/03/2023	52000.0	05-07-2023
Client_7	client7@gmail.com	29/10/2022	52000.0	null
Client_8	client8@gmail.com	22/08/2022	52000.0	24-02-2023
Client_9	client9@gmail.com	30/08/2022	28000.0	07-01-2023
Client_10	client10@gmail.com	21/04/2023	28000.0	null
Client_11	client11@gmail.com	17/08/2023	42000.0	08-10-2023
Client_12	client12@gmail.com	20/05/2022	35000.0	05-06-2023
Client_13	client13@gmail.com	16/03/2023	0.0	16-07-2023
Client_14	client14@gmail.com	02/02/2023	35000.0	02-08-2023
Client_15	client15@gmail.com	08/11/2023	52000.0	null
Client_18			0.0	12-01-2023
Client_17			0.0	22-03-2023
Client_16			0.0	16-12-2023
Client_24			0.0	11-11-2023
Client_21			0.0	11-12-2023
Client_23			0.0	10-12-2023

MISE EN BASE DE DONNÉES

INFOS À RENSEIGNER

COMPOSANT UTILISÉ



Database MySQL Appliquer

Type de propriété Built-in

Version de la base de données MySQL 8

Utiliser une connexion existante

Hôte "" Port "3306"

Base de données ""

Utilisateur "" Mot de passe *****

Table ""

Action sur la table Par défaut Action sur les données Insert

Schéma Built-in Modifier le schéma

RÉSULTAT FINAL EN BASE

client_id	nom	email	date_inscription	revenu_annuel	purchase_date	total_spent	newsletter_optin	ville	chiffre_affaires	date_creation
9	Client_9	client9@gmail.com	30/08/2022	28000.0	07-01-2023	394.4	No	Toulouse	1620.8	2021-10-26
8	Client_8	client8@gmail.com	22/08/2022	52000.0	24-02-2023	238.3	No	Lyon	9382.1	2024-01-08
7	Client_7	client7@gmail.com	29/10/2022	52000.0				Toulouse	5926.0	2022-04-07
6	Client_6	client6@gmail.com	26/03/2023	52000.0	05-07-2023	230.1	No	Paris	2166.7	2021-07-07
5	Client_5	client5@gmail.com	25/09/2022	42000.0				Toulouse	3714.6	2022-12-27
4	Client_4	client4@gmail.com	28/04/2022	28000.0				Paris	9547.2	2023-10-13
3	Client_3	client3@gmail.com	08/01/2023	35000.0				Marseille	3022.8	2021-10-24
2	Client_2	client2@gmail.com	15/03/2022	35000.0						
1	Client_1	client1@gmail.com	06/03/2022	0.0						
15	Client_15	client15@gmail.com	08/11/2023	52000.0				Marseille	2568.6	2023-08-05
14	Client_14	client14@gmail.com	02/02/2023	35000.0	02-08-2023	329.8	No	Marseille	6062.3	2023-09-22
13	Client_13	client13@gmail.com	16/03/2023	0.0	16-07-2023	265.9	No			
12	Client_12	client12@gmail.com	20/05/2022	35000.0	05-06-2023	364.9	No		9509.2	2022-11-30
11	Client_11	client11@gmail.com	17/08/2023	42000.0	08-10-2023	436.8	Yes	Paris	2484.1	2022-03-20
10	Client_10	client10@gmail.com	21/04/2023	28000.0				Toulouse	8242.0	2022-02-22
	Client_18		0.0	12-01-2023	92.62	No	Paris	7221.6		2022-06-26
9	Client_17		0.0	22-03-2023	184.1	Yes	Toulouse	2333.8		2023-05-13
12	Client_16		0.0	16-12-2023	169.0	No				
14	Client_24		0.0	11-11-2023	497.6	No				
6	Client_21		0.0	11-12-2023	287.1	Yes				
	Client_23		0.0	10-12-2023	136.6	No				
5	Client_19		0.0	20-12-2023	64.91	No				
8	Client_22		0.0	01-11-2023	176.2	No				
	Client_20		0.0						5426.3	2021-08-16

POINT D'AMELIORATIONS

- Remplacement des valeurs nulles
- Avoir une clé primaire dont un client_id

client_id	nom	email	date_inscription	revenu_annuel	purchase_date	total_spent	newsletter_optin	ville	chiffre_affaires	date_creation
9	Client_9	client9@gmail.com	30/08/2022	28000.0	07-01-2023	394.4	No	Toulouse	1620.8	2021-10-26
8	Client_8	client8@gmail.com	22/08/2022	52000.0	24-02-2023	238.3	No	Lyon	9382.1	2024-01-08
7	Client_7	client7@gmail.com	29/10/2022	52000.0				Toulouse	5926.0	2022-04-07
6	Client_6	client6@gmail.com	26/03/2023	52000.0	05-07-2023	230.1	No	Paris	2166.7	2021-07-07
5	Client_5	client5@gmail.com	25/09/2022	42000.0				Toulouse	3714.6	2022-12-27
4	Client_4	client4@gmail.com	28/04/2022	28000.0				Paris	9547.2	2023-10-13
3	Client_3	client3@gmail.com	08/01/2023	35000.0				Marseille	3022.8	2021-10-24
2	Client_2	client2@gmail.com	15/03/2022	35000.0						
1	Client_1	client1@gmail.com	06/03/2022	0.0						
15	Client_15	client15@gmail.com	08/11/2023	52000.0				Marseille	2568.6	2023-08-05
14	Client_14	client14@gmail.com	02/02/2023	35000.0	02-08-2023	329.8	No	Marseille	6062.3	2023-09-22
13	Client_13	client13@gmail.com	16/03/2023	0.0	16-07-2023	265.9	No			
12	Client_12	client12@gmail.com	20/05/2022	35000.0	05-06-2023	364.9	No		9509.2	2022-11-30
11	Client_11	client11@gmail.com	17/08/2023	42000.0	08-10-2023	436.8	Yes	Paris	2484.1	2022-03-20
10	Client_10	client10@gmail.com	21/04/2023	28000.0				Toulouse	8242.0	2022-02-22
	Client_18			0.0	12-01-2023	92.62	No	Paris	7221.6	2022-06-26
9	Client_17			0.0	22-03-2023	184.1	Yes	Toulouse	2333.8	2023-05-13
12	Client_16			0.0	16-12-2023	169.0	No			
14	Client_24			0.0	11-11-2023	497.6	No			
6	Client_21			0.0	11-12-2023	287.1	Yes			
	Client_23			0.0	10-12-2023	136.6	No			
5	Client_19			0.0	20-12-2023	64.91	No			
8	Client_22			0.0	01-11-2023	176.2	No			
	Client_20			0.0					5426.3	2021-08-16

.....FIN.....
MERCI!