

04 Janvier 2023

PROJET BI

Analyse des fraudes par carte bancaire

Sommaire

- 1.Contexte.
- 2.Description du jeu de données.
- 3.Conception :
 - a.Questions analytiques.
 - b.DataMart en étoile.
- 4.Processus ETL
- 5.Visualisation avec Les tableaux de bord.

1. Contexte

Il y a un manque d'ensembles de données accessibles au public sur les services financiers et en particulier dans le domaine émergent des transactions d'argent mobile. Les ensembles de données financières sont importants pour de nombreux chercheurs et en particulier pour nous qui effectuons des recherches dans le domaine de la détection des fraudes. Une partie du problème réside dans la nature intrinsèquement privée des transactions financières, qui conduit à l'absence d'ensembles de données accessibles au public.

2. Description du dataset

Notre dataset tiré du site Kaggle:
<https://www.kaggle.com/datasets/realaxi/paysim1>
décrit des transactions d'argent
mobile généré par le simulateur
PaySim sur la base d'un échantillon de
transactions réelles.



2. Description du dataset

step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFraud
1	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155	0.0	0.0	0
1	TRANSFER	181.0	C1305486145	181.0	0.0	C553264065	0.0	0.0	1
1	DEBIT	5423.28	C2139168000	20281.0	14857.72	C367746789	11523.0	0.0	0
1	CASH_OUT	472385.25	C1193841543	26210.19	0.0	C392292416	1541224.63	3420103.09	0
1	CASH_IN	159334.62	C431164562	1746897.72	1906232.33	C242131142	15735.13	0.0	0

2. Description du dataset

Step : Mappe une unité de temps dans le monde réel. Dans ce cas, 1 étape correspond à 1 heure de temps.

type : ENCAISSEMENT, RETRAIT, DÉBIT, PAIEMENT et VIREMENT

Amount : Montant de la transaction en devise locale.

nameOrig : Client qui a commencé la transaction

oldbalanceOrg:Solde initial avant la transaction

newbalanceOrig :Solde du client après la transaction.

nameDest ;ID du destinataire de la transaction.

oldbalanceDest : Solde initial du destinataire avant la transaction.

newbalanceDest ; le solde du destinataire après la transaction.

isFraude : identifie une transaction frauduleuse (1) et non frauduleuse (0)

Fraude

isFraude	Description
1	Fraude
0	Non-Fraude

Temps

Heure	Jour	Annees	Libelle_jour	Mois
18:13:00	01-janv	2009	jeudi	Janvier
16:08:00	02-févr	2010	mardi	Février
00:16:00	03-mars	2011	jeudi	Mars
04:04:00	04-mai	2012	vendredi	Mai
22:59:00	05-juin	2013	mercredi	Juin
05:50:00	06-juil	2014	dimanche	Juillet
19:45:00	07-août	2015	vendredi	Août
08:32:00	08-sept	2016	samedi	Septembre

Localité

Ville	Pays
Casablanca	Morocco
Rabat	Morocco
Fes	Morocco
Paris	France
Toulouse	France
Lyon	France
Bamako	Mali
Koulikoro	Mali
Sikasso	Mali

Type transaction

Transaction	Description
PAYMENT	Paielement
TRANSFER	Transfert
CASH_OUT	Decaissement
DEBIT	Debit
CASH_IN	Encaissement

Table initiale

Heure	Ville	type	amount	nameOrig	oldbalanceOrig	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFraud
18:13:00	Casablanca	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155	0.0	0.0	0
16:08:00	Rabat	PAYMENT	1864.28	C1666544295	21249.0	19384.72	M2044282225	0.0	0.0	0
00:16:00	Fes	TRANSFER	181.87	C1305486145	181.0	0.0	C553264065	0.0	0.0	1
04:04:00	Paris	CASH_OUT	181.07	C840083671	181.0	0.0	C38997010	21182.0	0.0	1
22:59:00	Toulouse	PAYMENT	11668.14	C2048537720	41554.0	29885.86	M1230701703	0.0	0.0	0
05:50:00	Lyon	PAYMENT	7817.71	C90045638	53860.0	46042.29	M573487274	0.0	0.0	0
19:45:00	Bamako	PAYMENT	7107.77	C154988899	183195.0	176087.23	M408069119	0.0	0.0	0
08:21:00	Koulikoro	PAYMENT	7861.64	C1912850431	176087.23	168225.59	M633326333	0.0	0.0	0
19:23:00	Sikasso	PAYMENT	4024.36	C1265012928	2671.0	0.0	M1176932104	0.0	0.0	0
03:14:00	Ouagadougou	DEBIT	5337.77	C712410124	41720.0	36382.23	C195600860	41898.0	40348.79	0
14:03:00	Abidjan	DEBIT	9644.94	C1900366749	4465.0	0.0	C997608398	10845.0	157982.12	0
01:48:00	Abiko	PAYMENT	3099.97	C249177573	20771.0	17671.03	M2096539129	0.0	0.0	0

3.Conception

3.a. Questions Analytiques :

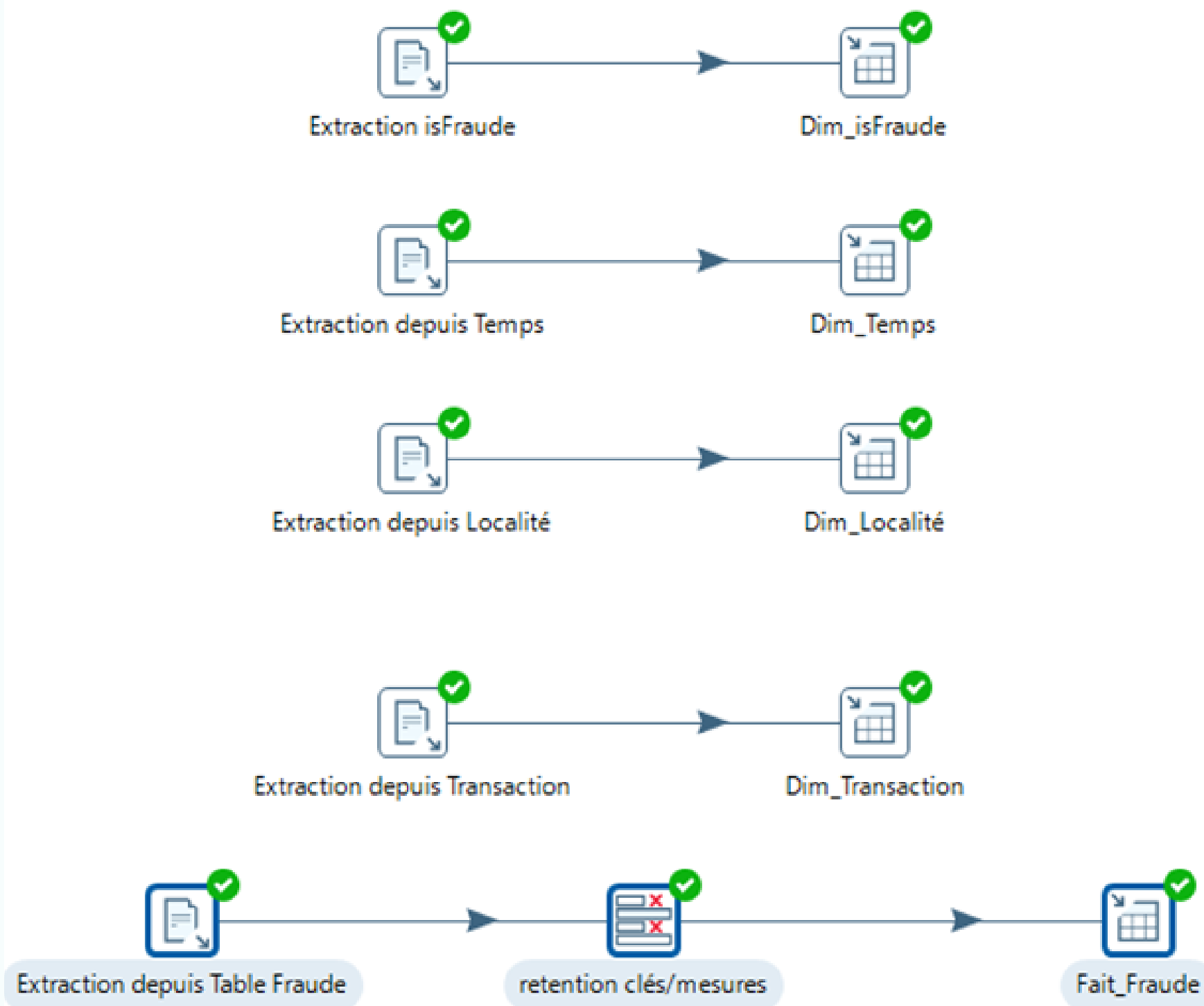
Quelles sont les periodes les plus susceptible d'avoir de fraude ?

Quel est le montant de fraudes(Journalier, Mensuel, Annuel, horaire par type de transaction ?

Quelles regions connaissent plus de fraudes ?

Quel est le nombre de fraudes et son montant par type de transaction ?

4. PROCESSUS ETL



On alimente nos dataMarts en utilisant les données provenant des divers fichiers csv.

L'etape *retention clés/mesures* est utilisée pour sélectionner les colonnes à inclure dans la table de faits.

Dim isFraude

sFraude	Description
0	Non-Fraude
1	Fraude

Ville	Pays
Abidjan	Ivory Coast
Abiko	Japan
Abilene	United States
Abohar	India

Dim Localité

FAIT FRAUDE

Heure	Ville	Type_Transaction	Amount	isFraude
00:03:00	Abidjan	CASH_IN	212604	0
00:03:00	Abidjan	CASH_OUT	91633	0
00:03:00	Abidjan	PAYMENT	10831	0
00:03:00	Abidjan	TRANSFER	618123	0
00:03:00	Abiko	CASH_IN	170210	0
00:03:00	Abiko	CASH_OUT	311415	0

DM alimenté

Dim type trans

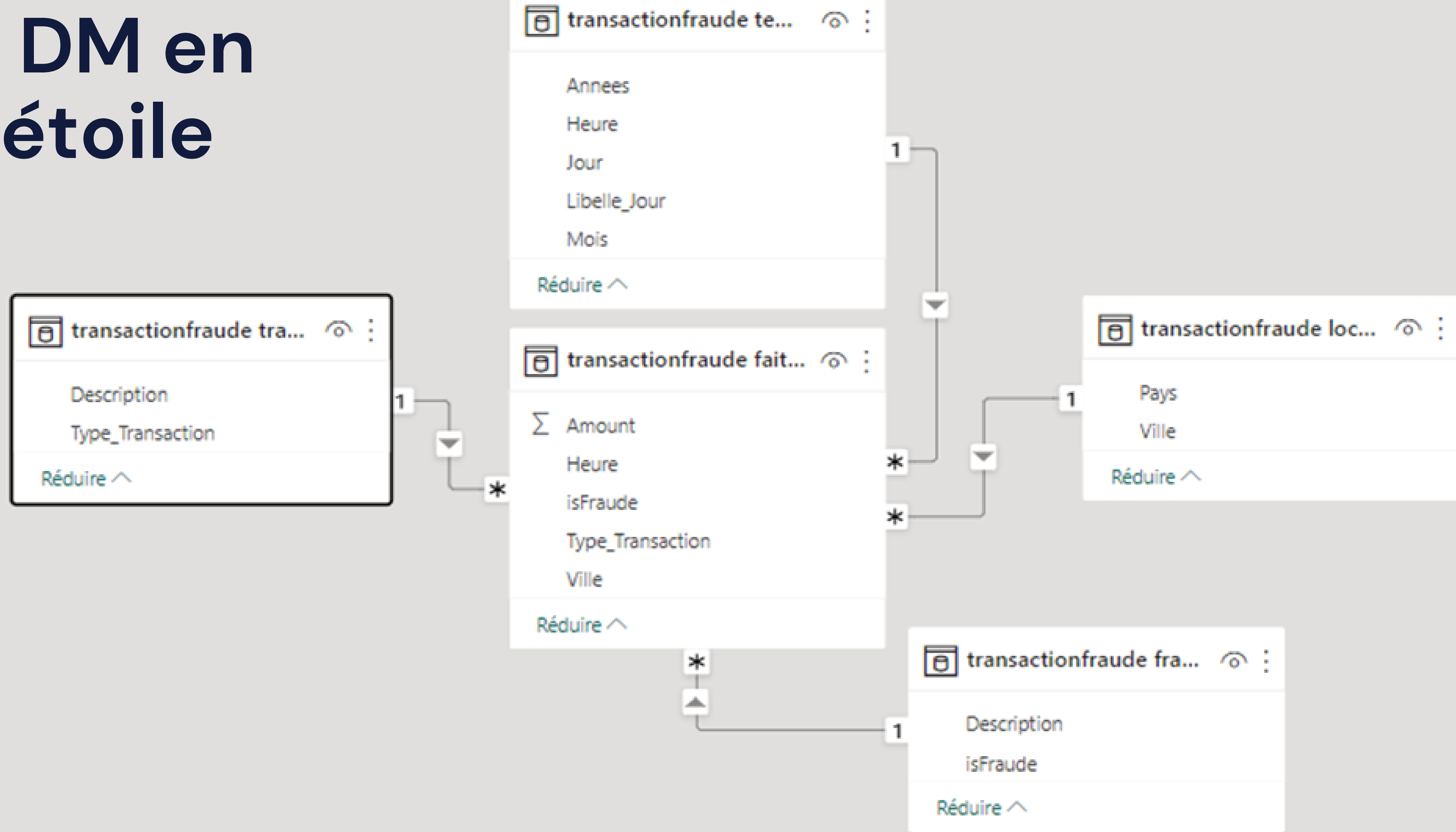
Type_Transaction	Description
CASH_IN	Encaissement
CASH_OUT	Decaissement
DEBIT	Debit

Dim Temps

Heure	Jour	Annees	Libelle_Jour	Mois
00:03:00	07-juin	2018	jeudi	Juin
00:09:00	26-août	2011	vendredi	Août
00:15:00	01-juil	2010	jeudi	Juin
00:16:00	03-mars	2011	jeudi	Mars
00:20:00	05-févr	2013	mardi	Février

DM alimenté

DM en étoile



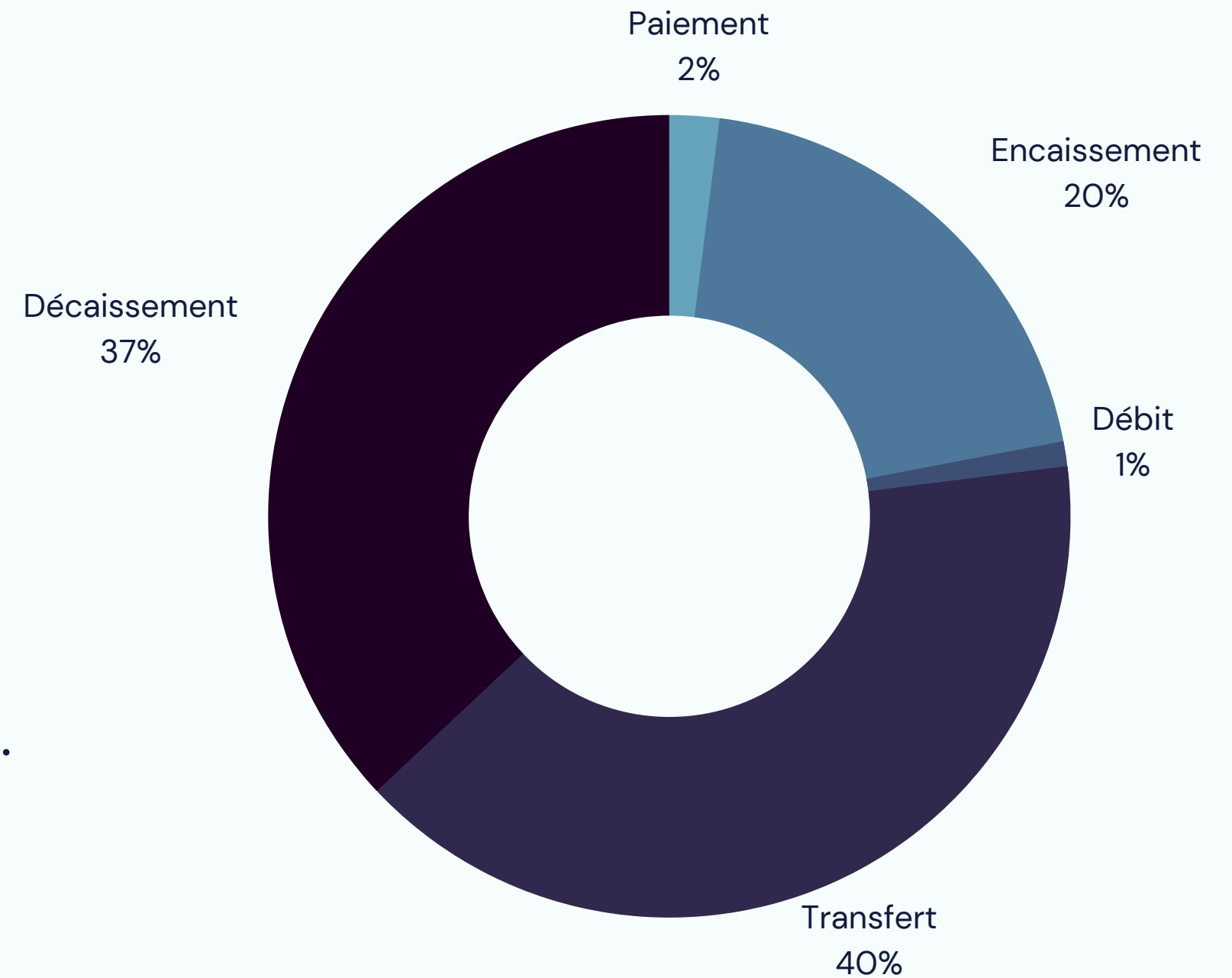
Power BI

Visuels utilisés

Dans cette partie on expose les différents visuels utilisés et dans quels buts.

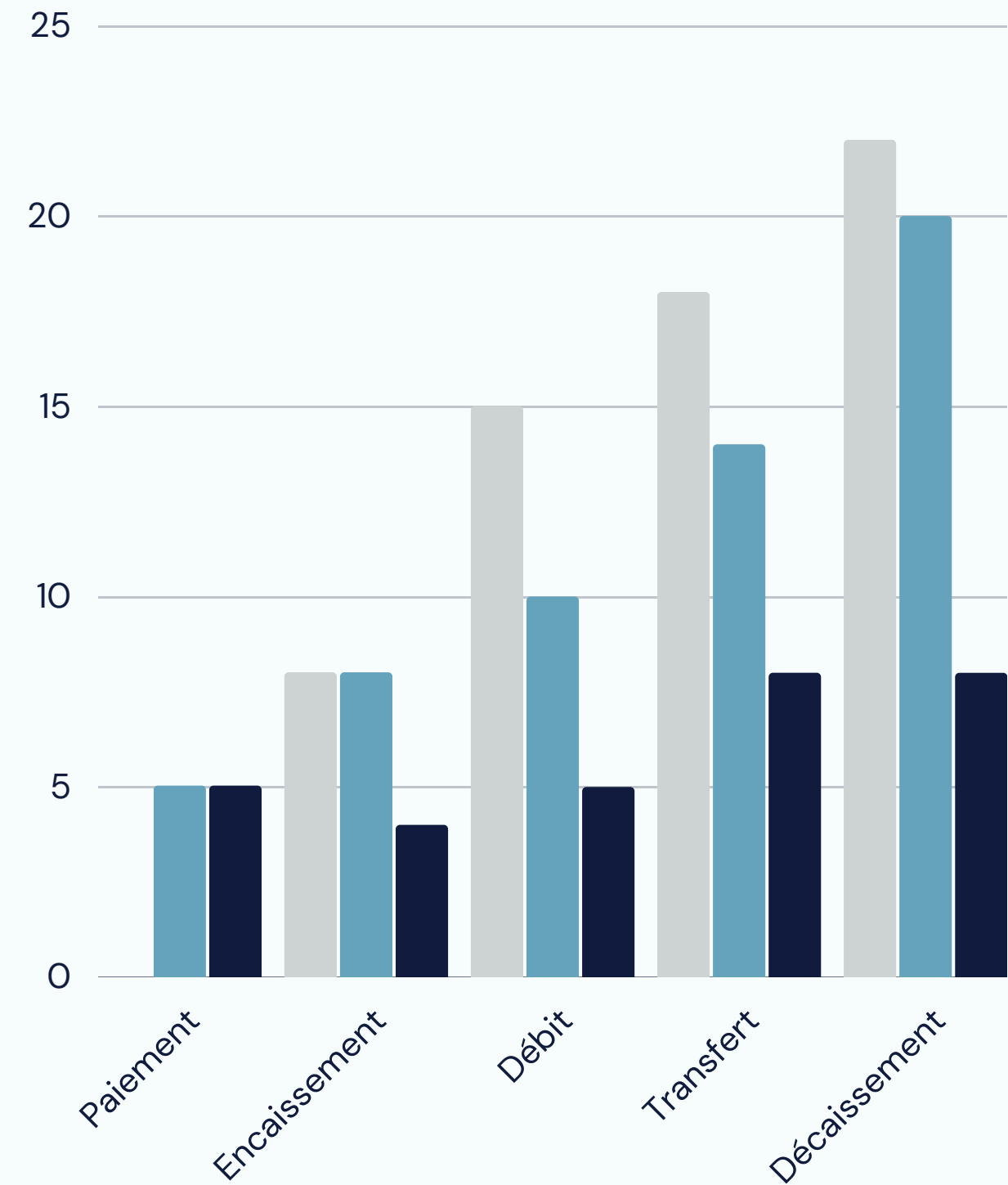
Graphique en anneau

- La somme des montants par transaction.
- Le pourcentage des transactions frauduleuses/non frauduleuses.



Histogramme groupé/empilé

- Le nombre de fraude par pays.
- Le nombre de fraude par type de transaction.



Graphique en courbes

- L'évolution du nombre de transactions frauduleuses au cours des mois.



“

5. VISUALISATION AVEC LES TABLEAUX DE BORD.

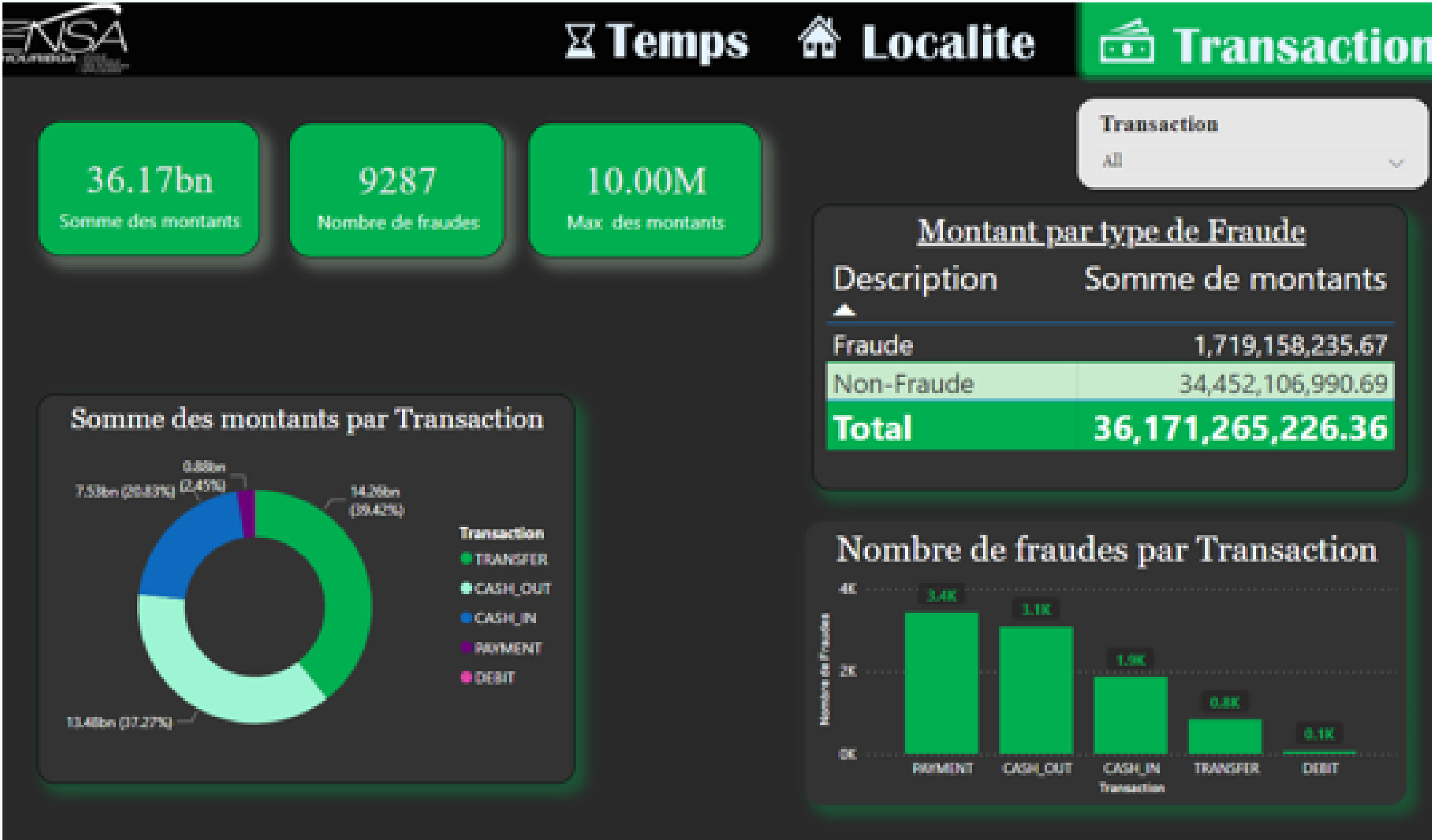
Notre Dash-Bord contient 3 pages

-Temps -Localité -Transaction

PAGE TRANSACTION

Cette page permet de répondre à la question suivante :

Quel est le nombre de fraudes et son montant par type de transaction ?



PAGE LOCALITE

cette page permet de répondre à la question suivante :

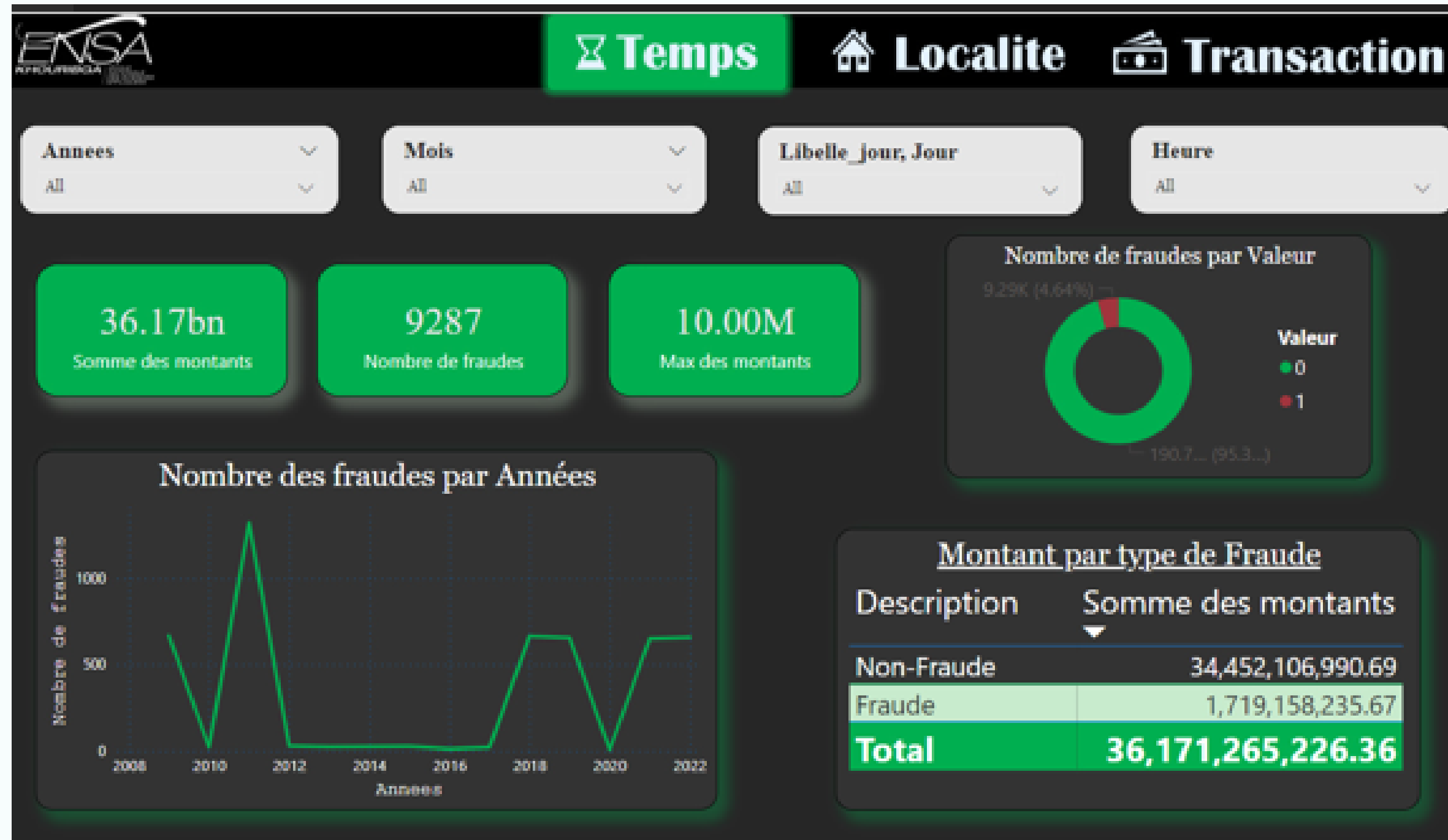
Quelles regions connaissent plus de fraudes ? et quel est son montant?



PAGE TEMPS

cette page permet de répondre à la question suivante :

Quelles périodes connaissent plus de transactions frauduleuses ?
Quel en est le montant correspondant ?



Conclusion

Grace à ce projet, nous avons pu nous familiariser avec l'environnement de Power BI, d'appliquer les notions vu au cours sur un data set assez volumineux et d'en tirer des connaissances. Nous adressons notre chaleureuse gratitude à notre professeur du module –Entrepôt de données– Mme. Nassima SOUSSI, pour sa disponibilité et pour la qualité de son enseignement durant ce module.

Merci !

de votre attention