

**Nom et prénoms : SOSSOU Didi Orlog**

**Parcours : IA et Big Data**

## **Analyse des Comportements Financiers et Détection d'Anomalies**

### **1. Introduction**

Ce rapport présente l'analyse d'un ensemble de 20 000 transactions financières. À l'aide de techniques de clustering, nous avons segmenté les transactions et identifié les comportements inhabituels.

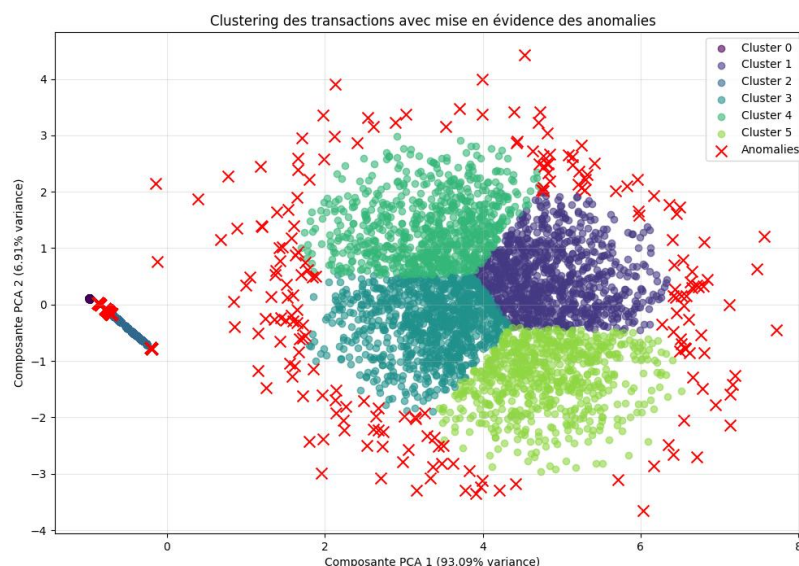
### **2. Méthodologie**

Notre analyse s'est déroulée en trois étapes principales :

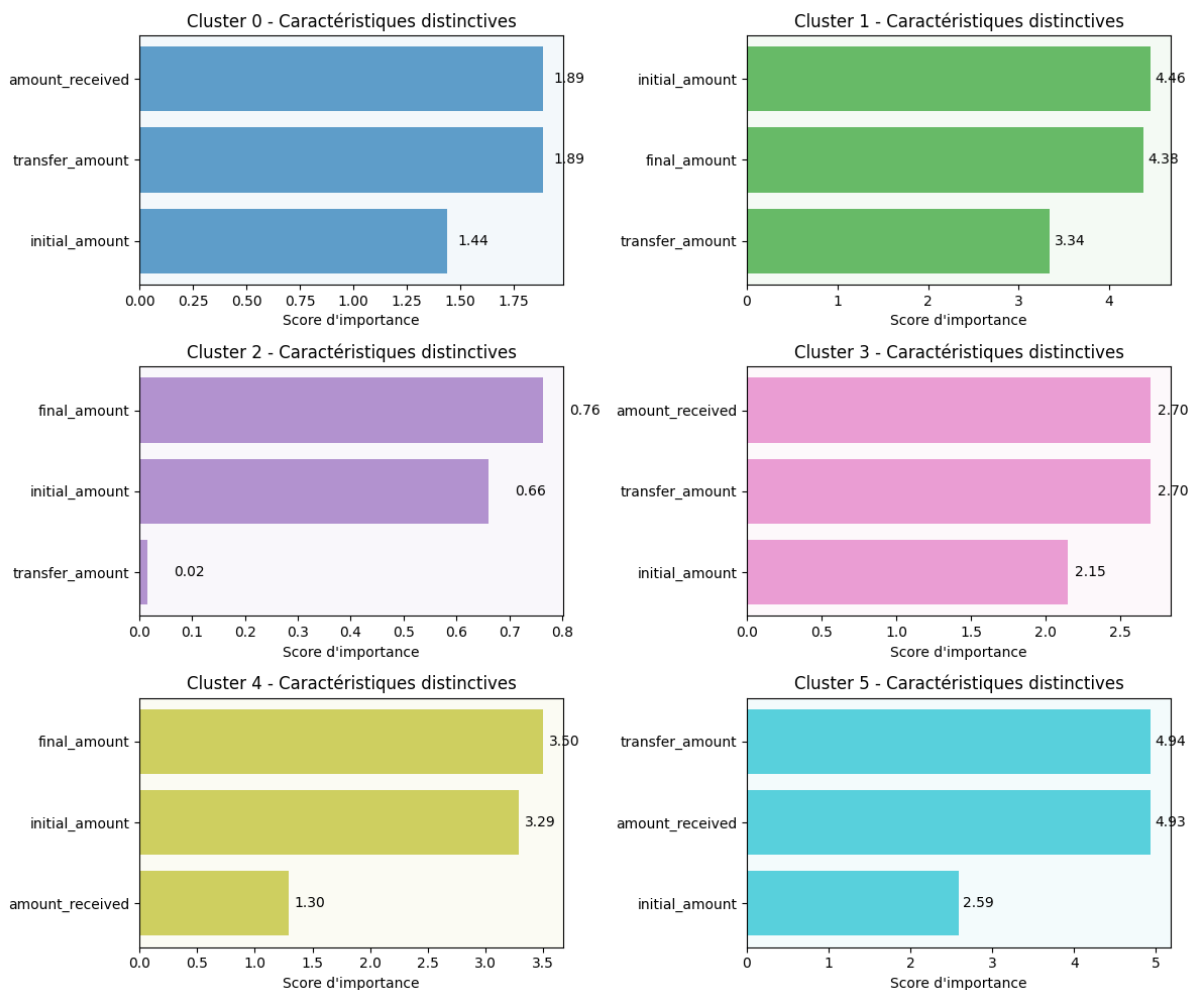
- Sélection et normalisation des caractéristiques financières pertinentes
- Segmentation des transactions par clustering K-means
- Détection d'anomalies via la détermination de seuils dynamiques par cluster en utilisant la méthode de Tukey (IQR)

### **3. Analyse des Clusters**

L'algorithme K-means a identifié 6 clusters distincts, chacun représentant un comportement financier spécifique :



Contribution de chaque variable à la création des clusters :



#### Cluster 0 : Clients très modestes (67.14% des transactions)

- **Profil financier** : Solde initial modeste (~390 Unités), transferts de montants similaires (~369 Unités), solde final très faible (~51 Unités). Utilisateurs dépensant presque l'intégralité de leur solde disponible pour des achats ou paiements courants. Ces transactions épuisent presque entièrement le solde disponible (87% utilisé en moyenne).
- **Comportement** : Normal

#### ➤ Cluster 1 : Clients riches dépensiers (4.83% des transactions)

- **Profil financier** : Soldes extrêmement élevés (~247 083), transferts importants mais proportionnellement modérés (~33 090). Clients fortunés effectuant des

virements substantiels tout en maintenant une forte position de liquidités. Investisseurs ou grandes entreprises gérant des réserves importantes avec des transferts pour investissements ou opérations majeures. Ce type de client maintient un ratio prudent de transfert/solde (~13%), conservant une forte réserve de liquidités.

- **Comportement** : High-risk (montants très élevés)

➤ **Cluster 2 : Clients frôlant le découvert (15.38% des transactions)**

- **Profil financier** : Solde initial modéré (~2 256 Unités), transferts largement excédentaires (-4 175 Unités). Utilisateurs dépensant systématiquement au-delà de leurs moyens, générant des découverts significatifs. Ils sont potentiellement en difficulté financière. Le montant moyen transféré représente 285% du solde initial, entraînant un découvert équivalent à 185% du solde de départ.
- **Comportement** : Overdraft

➤ **Cluster 3 : Clients riches modérés (4.92% des transactions)**

- **Profil financier** : Solde initial élevé (~153 769 Unités), transferts substantiels (~28 473 Unités), solde final encore confortable (~125 296 Unités). Clients aisés réalisant des transactions importantes mais mesurées par rapport à leur capacité financière. Probablement des professionnels, moyennes entreprises ou investisseurs effectuant des opérations régulières d'envergure. Ces clients utilisent environ 18.5% de leur solde pour chaque transaction, préservant une bonne marge de sécurité.
- **Comportement** : High-risk (montants très élevés)

➤ **Cluster 4 : Clients riches très conservateurs (4.09% des transactions)**

- **Profil financier** : Très haut solde initial (~213 191 Unités), transferts relativement modérés (~18 004 Unités), solde final presque intact (~195 187 Unités). Grands épargnants réalisant des opérations ponctuelles de montant limité par rapport à leurs avoirs. Rentiers, grands patrimoines ou entreprises avec forte trésorerie effectuant des dépenses courantes. Ces clients maintiennent un ratio transfert/solde très faible (~8.4%), indiquant une gestion très prudente et conservatrice des liquidités.

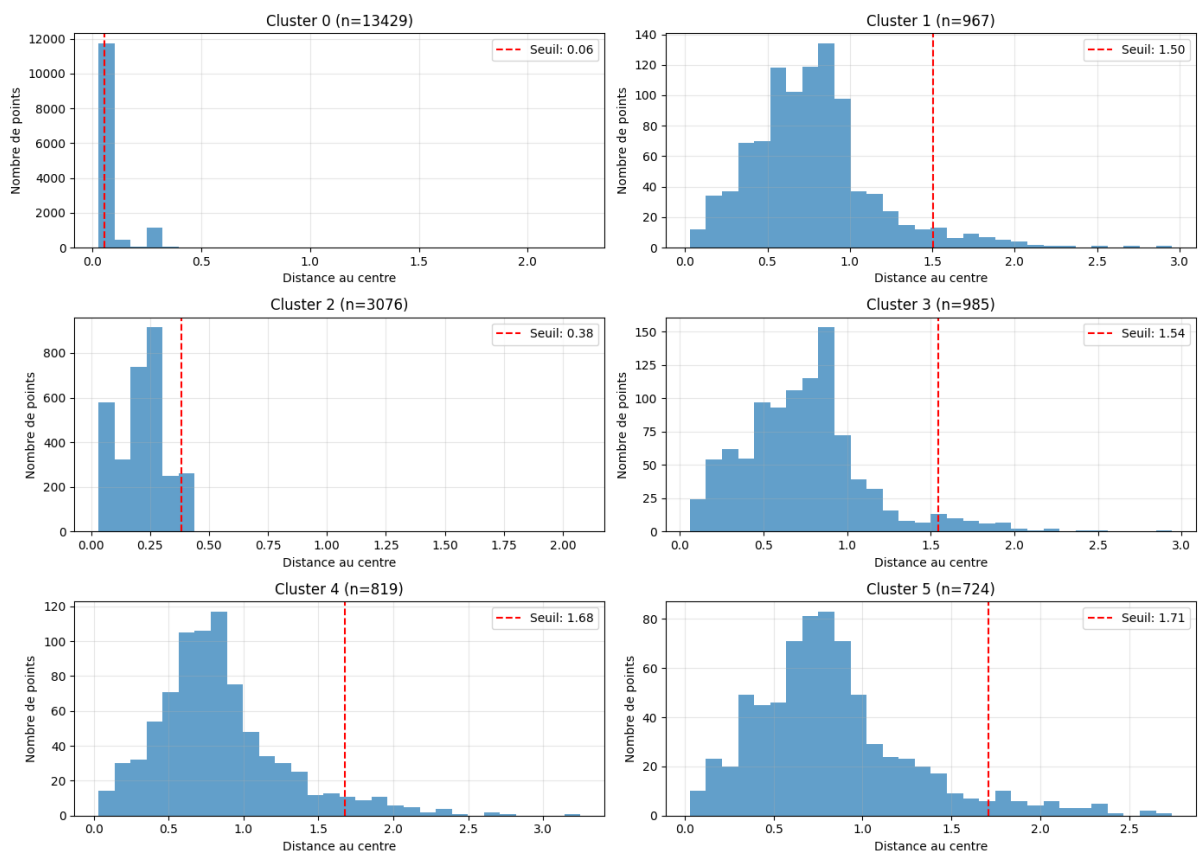
- **Comportement** : High-risk (montants très élevés)

➤ **Cluster 5 : Clients riches aux grosses transactions (3.62% des transactions)**

- **Profil financier** : Solde conséquent (~182 312 Unités), transferts de grande ampleur (~43 558 Unités), solde final diminué mais substantiel (~138 754 Unités). Entreprises effectuant des investissements stratégiques, acquisitions ou règlements de contrats importants. Les transferts représentent environ 23.9% du solde initial.
- **Comportement** : High-risk (montants très élevés)

#### 4. Analyse des Anomalies

En utilisant la méthode de Tukey (IQR) pour fixer les seuils, nous obtenons les seuils suivants pour chaque cluster :



Avec ces seuils dynamiques nous avons identifié 2 241 transactions anormales (11.2% du total), réparties comme suit :

Distribution des anomalies par cluster			
Cluster	Nb anomalies	% des anomalies	% du cluster
Cluster 0	1783	79.56%	13.28%
Cluster 1	49	2.19%	5.07%
Cluster 2	273	12.18%	8.88%
Cluster 3	45	2.01%	4.57%
Cluster 4	49	2.19%	5.98%
Cluster 5	42	1.87%	5.80%

### ➤ Types d'anomalies identifiées :

Parmi les anomalies détectées, on peut distinguer plusieurs typologies :

- **Ghost transactions (principalement Cluster 0):** Transactions où le montant transféré dépasse inexplicablement le solde disponible, créant des découverts non anticipés (ex: ID 13663, ID 13845).
- **Overdraft extrême (Cluster 2):** Découverts anormalement importants par rapport au solde initial, dépassant significativement le ratio moyen du cluster déjà caractérisé par des découverts.
- **Comportement inversé (Cluster 2):** Anomalies présentant un profil contraire à celui du cluster, comme des soldes largement positifs dans un cluster généralement à découvert (ex: ID 17106).
- **Transactions disproportionnées (tous clusters):** Transferts dont la proportion par rapport au solde initial s'écarte significativement de la norme du cluster d'appartenance.

## 5. Interprétation et Recommandations

### ➤ Patterns financiers identifiés :

- La majorité des transactions (67%) suit un profil standard, mais contient paradoxalement la plus grande proportion d'anomalies (79.56%)
- Les transactions à découvert représentent une part significative (15%) de l'activité, suggérant une pratique répandue
- Les profils de clients aisés (Clusters 1, 3, 4, 5) présentent des taux d'anomalies plus faibles (~5-6%), indiquant des comportements plus prévisibles mais néanmoins avec des écarts significatifs
- Certaines anomalies révèlent des comportements diamétralement opposés au profil type de leur cluster (notamment dans le Cluster 2)

➤ **Recommandations :**

- Il serait pertinent de mettre en place un système de contrôle pour les transactions du Cluster 0 qui dépassent significativement le solde initial moyen, ce qui pourrait signaler des comportements à risque.
- Pour le Cluster 2, nous recommandons d'établir des seuils de découvert adaptés ou un mécanisme d'autorisation préalable. Une analyse des habitudes financières des clients concernés permettrait de leur proposer des alternatives mieux adaptées.
- Nous proposons un système de vérification à deux étapes pour les transactions anormales détectées dans les Clusters 1, 4 et 5. La mise en place d'un protocole de confirmation pour les transferts excédant 30% des montants habituels pourrait réduire les risques de transactions frauduleuses.