

LAPAGE

LIBRAIRIE

ANALYSE DES INDICATEURS ET
CHIFFRES CLÉS

CODIR 2025



SOMMAIRE

- 1. *Introduction*
- 2. *Les données initiales*
- 3. *Traitements des données*
- 4. *Jointures et fichier final*
- 5. *Analyses*
 - 5.1 Analyse du chiffre d'affaires
 - 5.2 Analyse des évolutions
 - 5.3 Analyse des clients
 - 5.4 Analyse des TOPS et FLOPS
- 6. *Corrélations*
- 7. *Conclusion*





1. INTRODUCTION

- **Contexte de l'entreprise**: Lapage, librairie physique avec plusieurs points de vente, a lancé un site e-commerce il y a 2 ans
- **Enjeu actuel**: Besoin de mieux comprendre les données pour piloter l'activité en ligne et physique
- **Objectif de la mission**: Analyser les indicateurs de vente et le comportement des clients afin de dégager des pistes stratégiques



2. LES DONNÉES INITIALES

client_id **sex** **birth**

| | client_id | sex | birth |
|---|-----------|-----|-------|
| 0 | c_4410 | f | 1967 |
| 1 | c_7839 | f | 1975 |
| 2 | c_1699 | f | 1984 |
| 3 | c_5961 | f | 1962 |
| 4 | c_5320 | m | 1943 |

Customers

id_prod **date** **session_id** **client_id**

| | id_prod | date | session_id | client_id |
|---|---------|----------------------------|------------|-----------|
| 0 | 0_1259 | 2021-03-01 00:01:07.843138 | s_1 | c_329 |
| 1 | 0_1390 | 2021-03-01 00:02:26.047414 | s_2 | c_664 |
| 2 | 0_1352 | 2021-03-01 00:02:38.311413 | s_3 | c_580 |
| 3 | 0_1458 | 2021-03-01 00:04:54.559692 | s_4 | c_7912 |
| 4 | 0_1358 | 2021-03-01 00:05:18.801198 | s_5 | c_2033 |

Transactions

id_prod **price** **categ**

| | id_prod | price | categ |
|---|---------|-------|-------|
| 0 | 0_1421 | 19.99 | 0 |
| 1 | 0_1368 | 5.13 | 0 |
| 2 | 0_731 | 17.99 | 0 |
| 3 | 1_587 | 4.99 | 1 |
| 4 | 0_1507 | 3.99 | 0 |

Products

3. TRAITEMENT DES DONNÉES

| | client_id | sex | birth |
|---|-----------|-----|-------|
| 0 | c_4410 | f | 1967 |
| 1 | c_7839 | f | 1975 |
| 2 | c_1699 | f | 1984 |
| 3 | c_5961 | f | 1962 |
| 4 | c_5320 | m | 1943 |

Customers

- 8621 lignes et 3 colonnes
- 0 doublons
- Pas de valeurs manquantes

| | id_prod | date | session_id | client_id |
|---|---------|----------------------------|------------|-----------|
| 0 | 0_1259 | 2021-03-01 00:01:07.843138 | s_1 | c_329 |
| 1 | 0_1390 | 2021-03-01 00:02:26.047414 | s_2 | c_664 |
| 2 | 0_1352 | 2021-03-01 00:02:38.311413 | s_3 | c_580 |
| 3 | 0_1458 | 2021-03-01 00:04:54.559692 | s_4 | c_7912 |
| 4 | 0_1358 | 2021-03-01 00:05:18.801198 | s_5 | c_2033 |

Transactions

- 687 534 lignes et 4 colonnes
- 0 doublons
- Pas de valeurs manquantes

| | id_prod | price | categ |
|---|---------|-------|-------|
| 0 | 0_1421 | 19.99 | 0 |
| 1 | 0_1368 | 5.13 | 0 |
| 2 | 0_731 | 17.99 | 0 |
| 3 | 1_587 | 4.99 | 1 |
| 4 | 0_1507 | 3.99 | 0 |

Products

- 3286 lignes et 3 colonnes
- 0 doublons
- Pas de valeurs manquantes

4. JOINTURES ET FICHIER FINAL

Etape 1

```
#Jointure des tables customers et transactions  
customers_transactions = pd.merge(customers, transactions, how='inner', on=['client_id'])
```

Etape 2

```
#Jointure de la table customers_transactions et products  
sales = pd.merge(customers_transactions, products, how='inner', on=['id_prod'])
```



Table obtenue

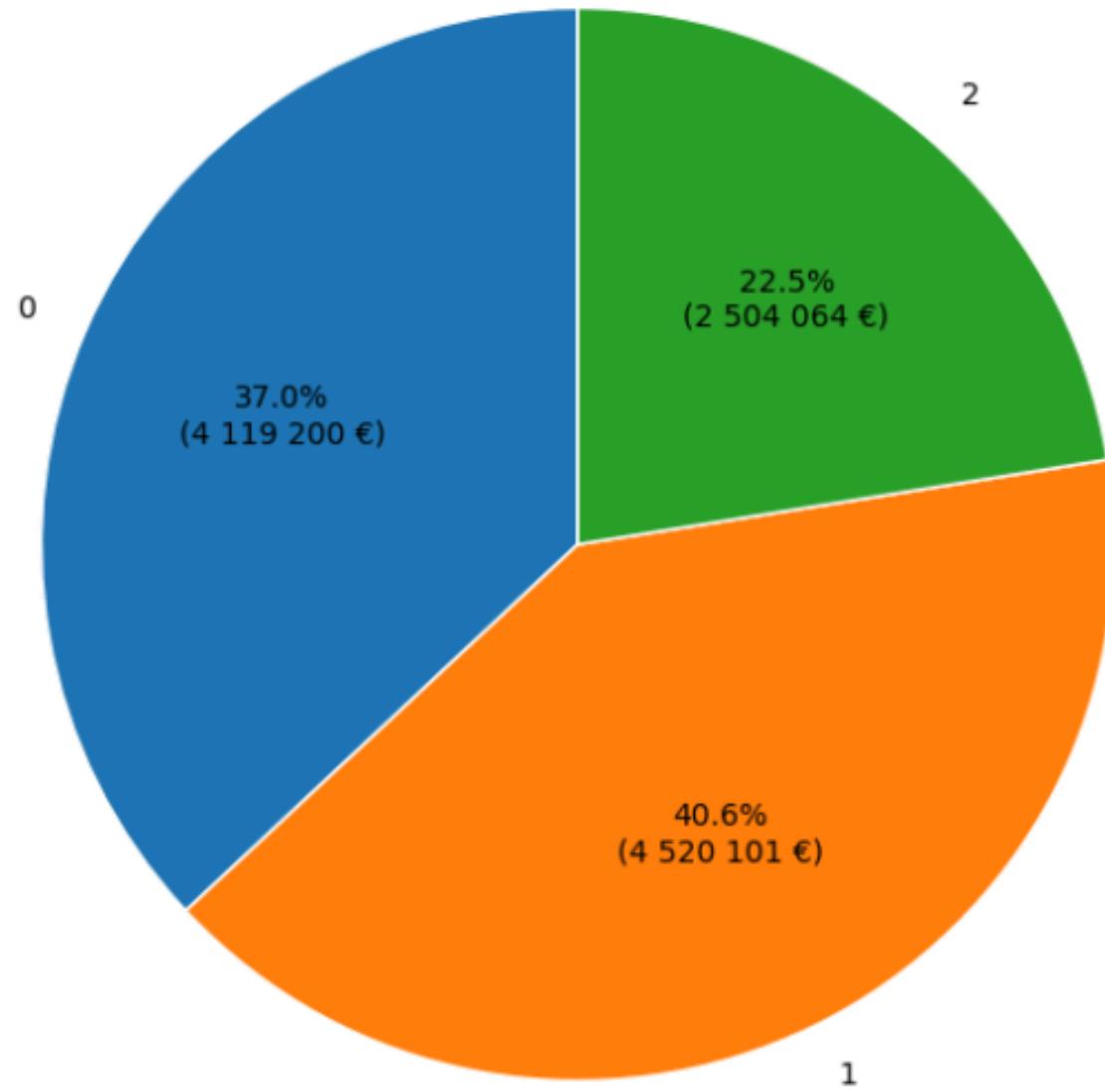
| | client_id | sex | birth | id_prod | | date | session_id | price | categ |
|---|-----------|-----|-------|---------|---------------------|---------|------------|-------|-------|
| 0 | c_4410 | f | 1967 | 1_483 | 2021-03-13 21:35:56 | s_5913 | 15.99 | 1 | |
| 1 | c_4410 | f | 1967 | 0_1111 | 2021-03-22 01:27:49 | s_9707 | 19.99 | 0 | |
| 2 | c_4410 | f | 1967 | 1_385 | 2021-03-22 01:40:23 | s_9707 | 25.99 | 1 | |
| 3 | c_4410 | f | 1967 | 0_1455 | 2021-03-22 14:29:25 | s_9942 | 8.99 | 0 | |
| 4 | c_4410 | f | 1967 | 0_1420 | 2021-03-22 22:31:26 | s_10092 | 11.53 | 0 | |

Table finale

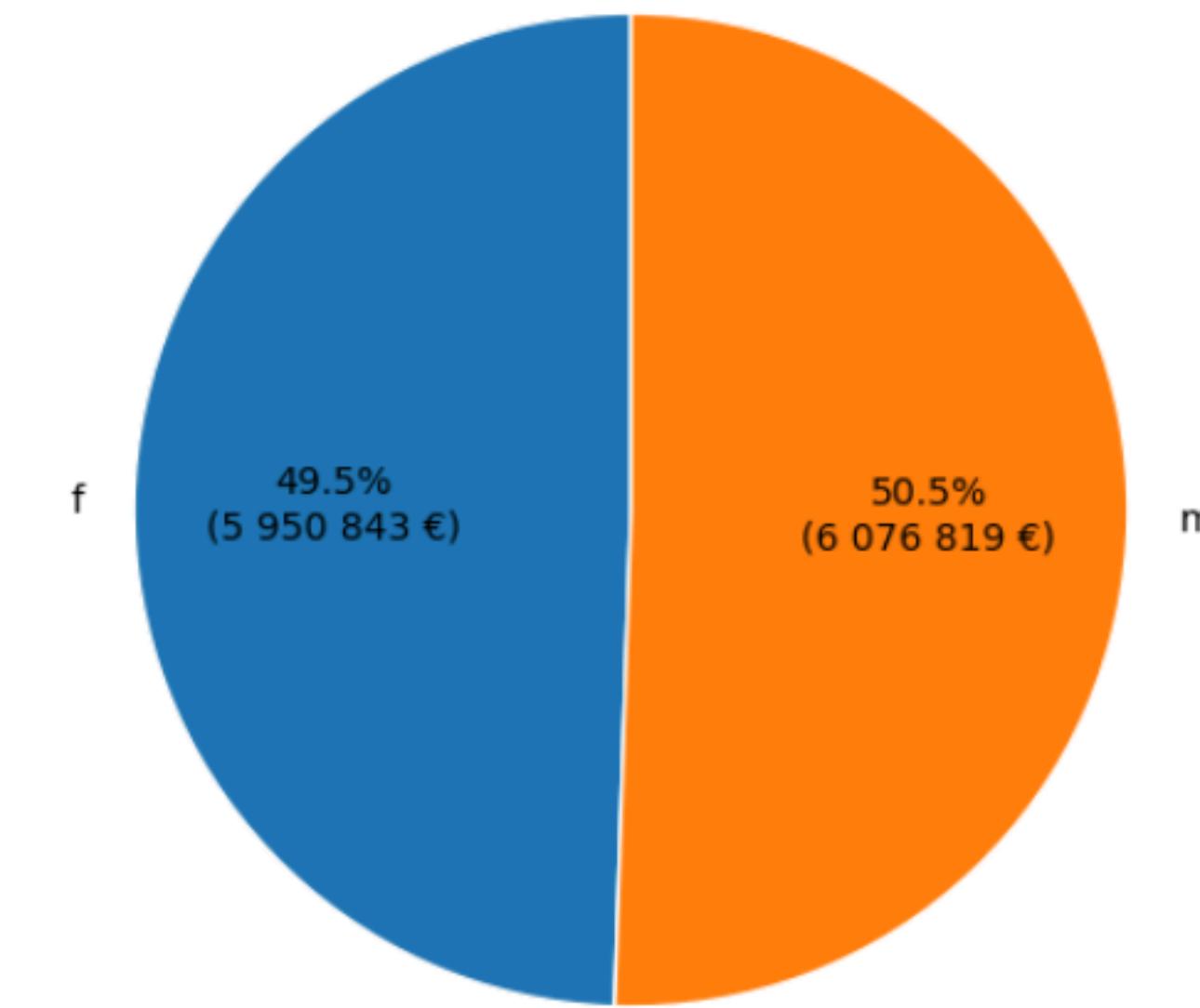
| | client_id | sex | birth | id_prod | | date | session_id | price | categ | year | month | age |
|---|-----------|-----|-------|---------|---------------------|---------|------------|-------|-------|------|-------|-----|
| 0 | c_4410 | f | 1967 | 1_483 | 2021-03-13 21:35:56 | s_5913 | 15.99 | 1 | 2021 | 3 | 54 | |
| 1 | c_4410 | f | 1967 | 0_1111 | 2021-03-22 01:27:49 | s_9707 | 19.99 | 0 | 2021 | 3 | 54 | |
| 2 | c_4410 | f | 1967 | 1_385 | 2021-03-22 01:40:23 | s_9707 | 25.99 | 1 | 2021 | 3 | 54 | |
| 3 | c_4410 | f | 1967 | 0_1455 | 2021-03-22 14:29:25 | s_9942 | 8.99 | 0 | 2021 | 3 | 54 | |
| 4 | c_4410 | f | 1967 | 0_1420 | 2021-03-22 22:31:26 | s_10092 | 11.53 | 0 | 2021 | 3 | 54 | |

5.1 ANALYSE DU CHIFFRE D'AFFAIRES

Répartition du chiffre d'affaires par catégorie



Répartition du chiffre d'affaires par genre



Le chiffre d'affaires total est de: 12 027 663 €

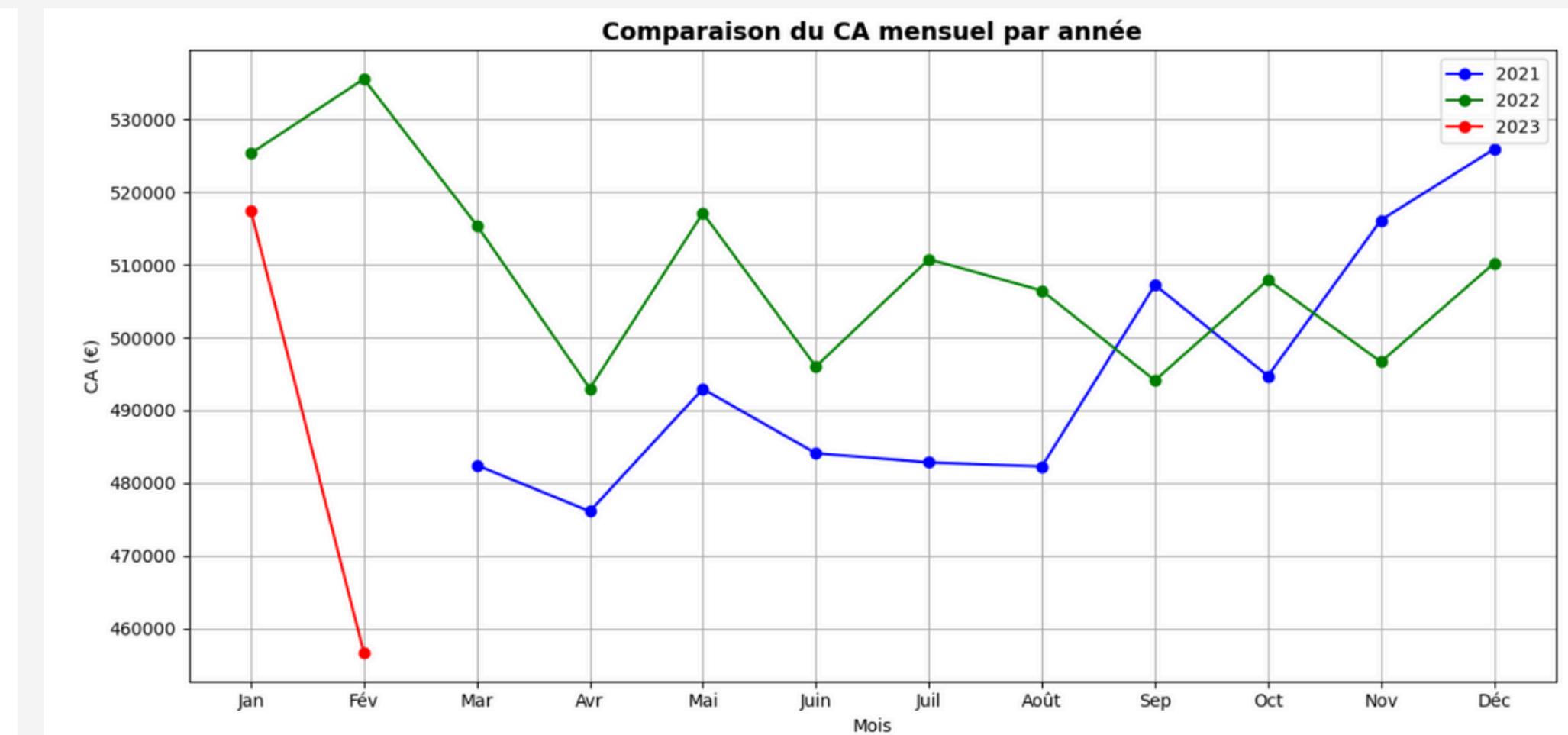
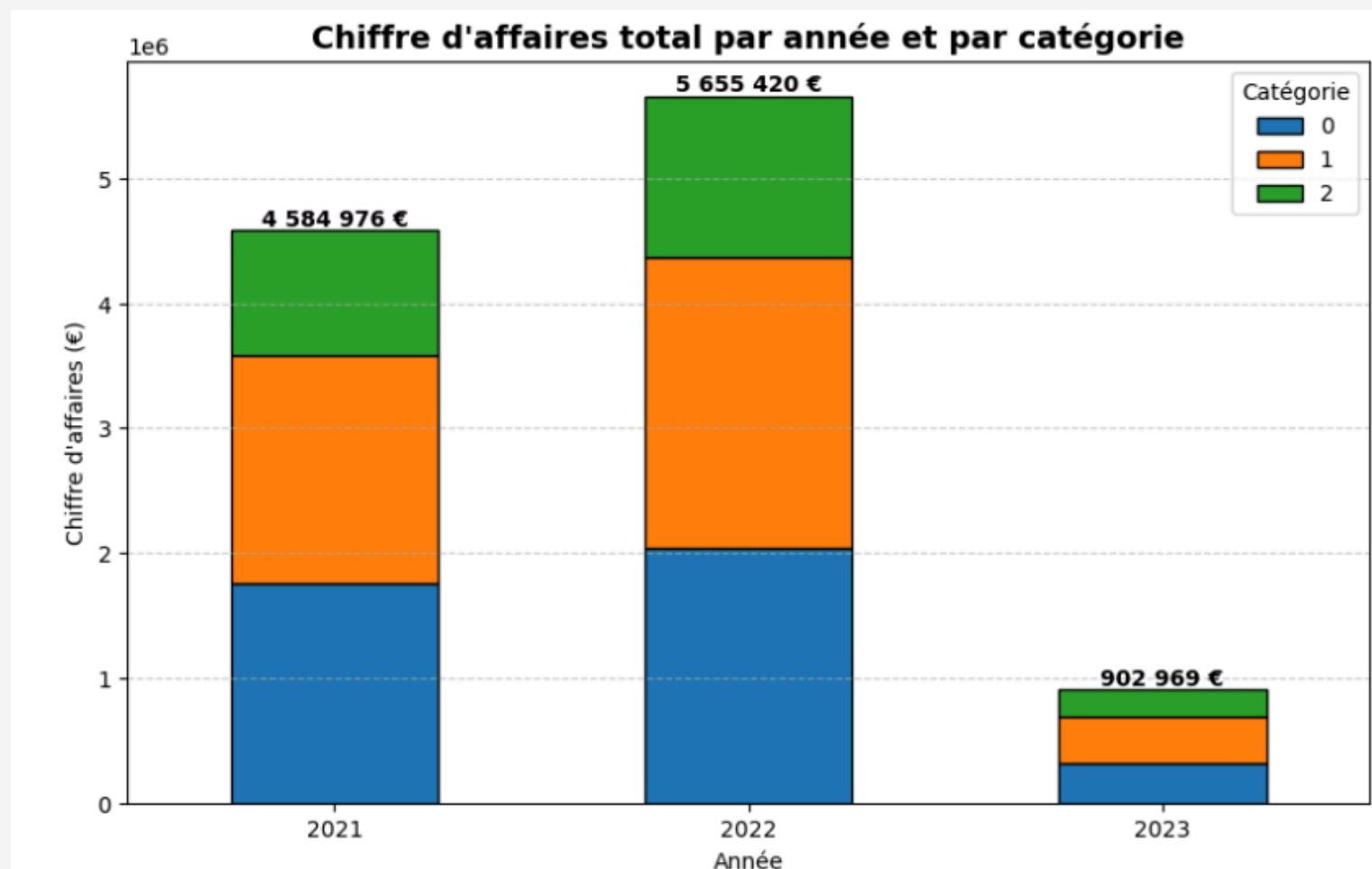
Nombre de commandes : 345505

Nombre de produits vendus : 687534

Nombre moyen de produits vendus par commande : 1.99

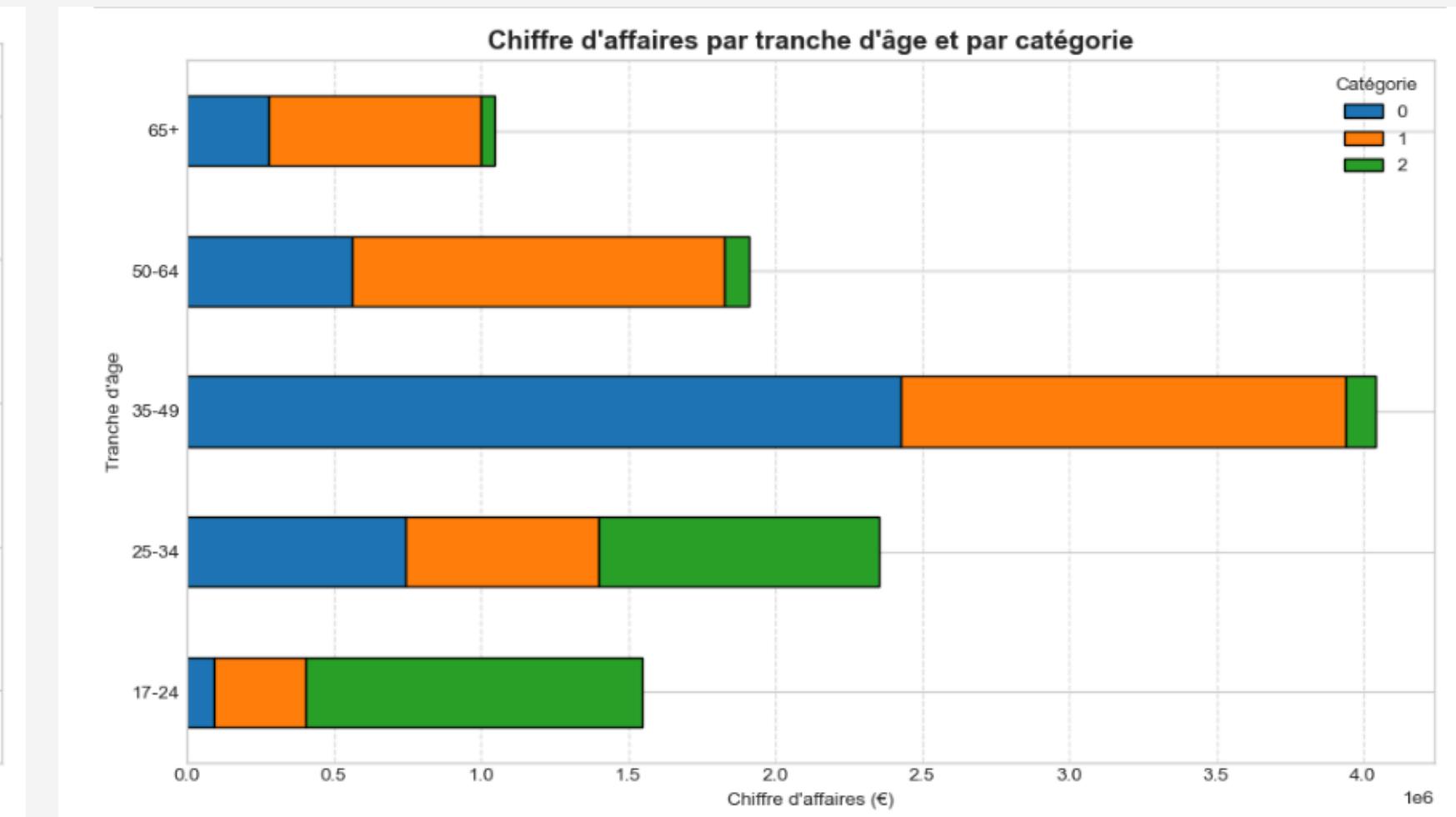
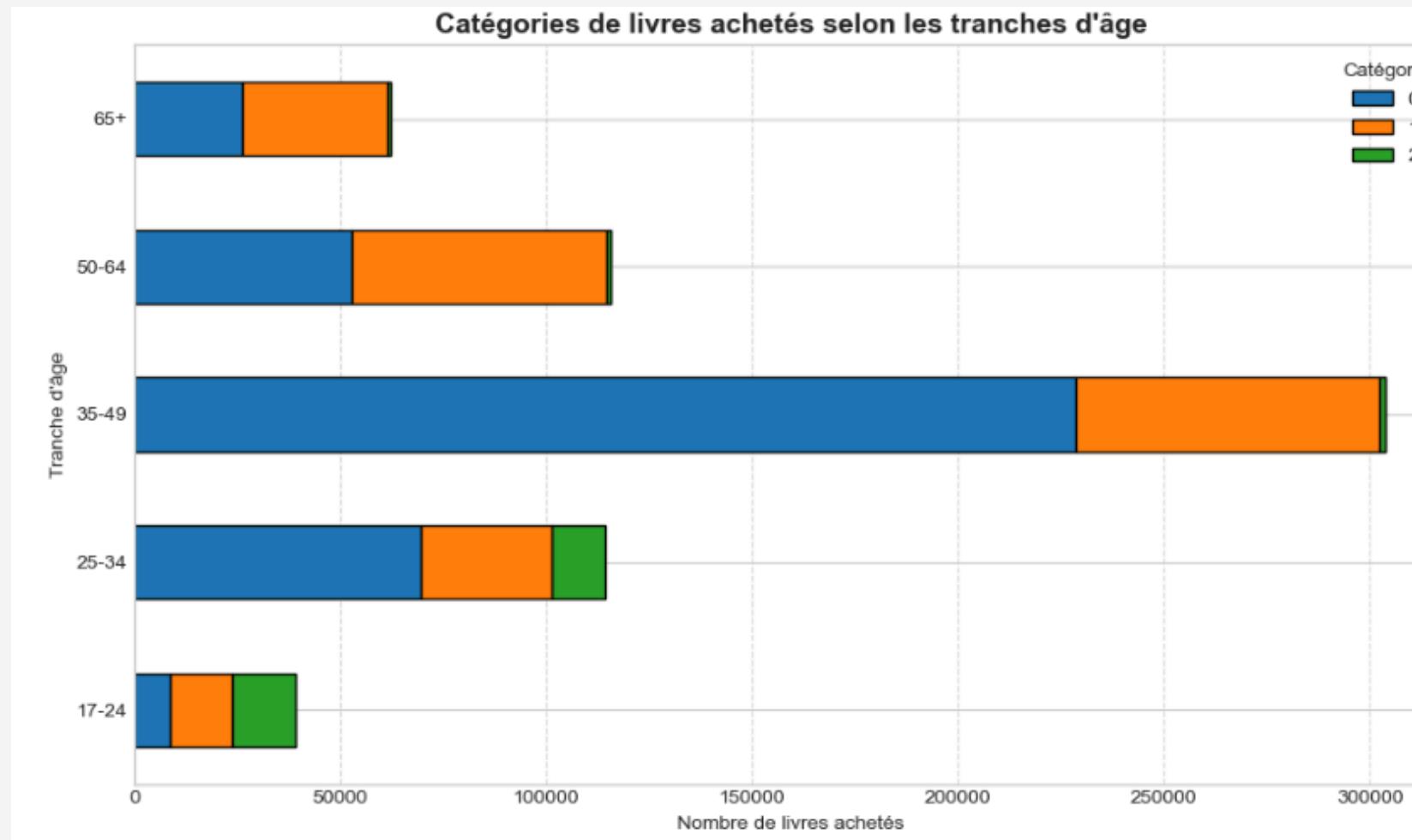
Panier moyen : 34.81

5.1 ANALYSE DU CHIFFRE D'AFFAIRES



5.2 ANALYSE DES CLIENTS

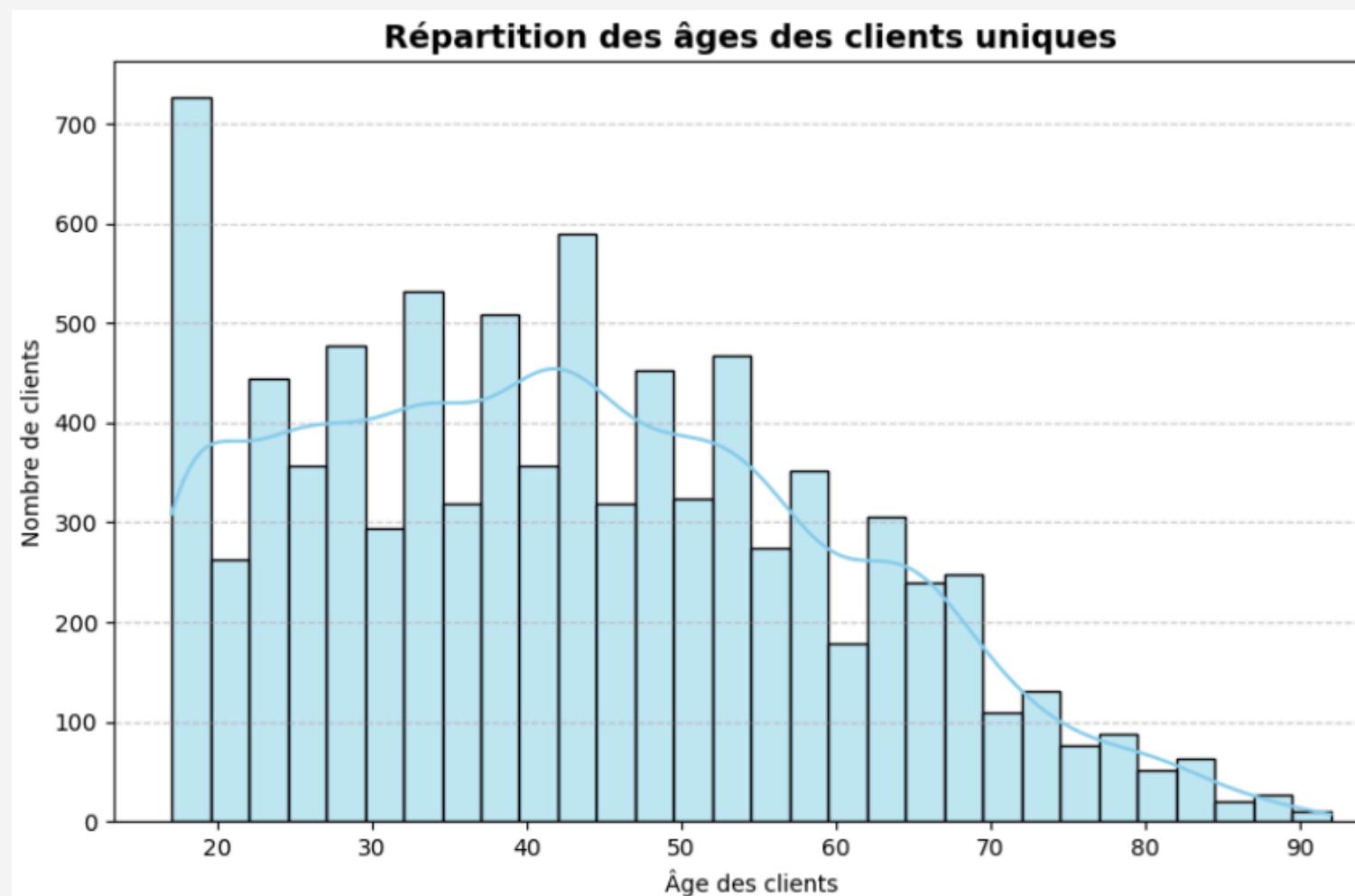
L'AGE DES CLIENTS



- **Catégorie d'âge: 35-49:** représente le plus gros CA avec un nombre de commande dépassant les 300 000
- **Catégorie d'âge: 25-34 :** 2ème tranche d'âge avec le plus haut CA
- **Catégorie d'âge: 17-24** achètent moins de livres mais le CA est **non négligeable**

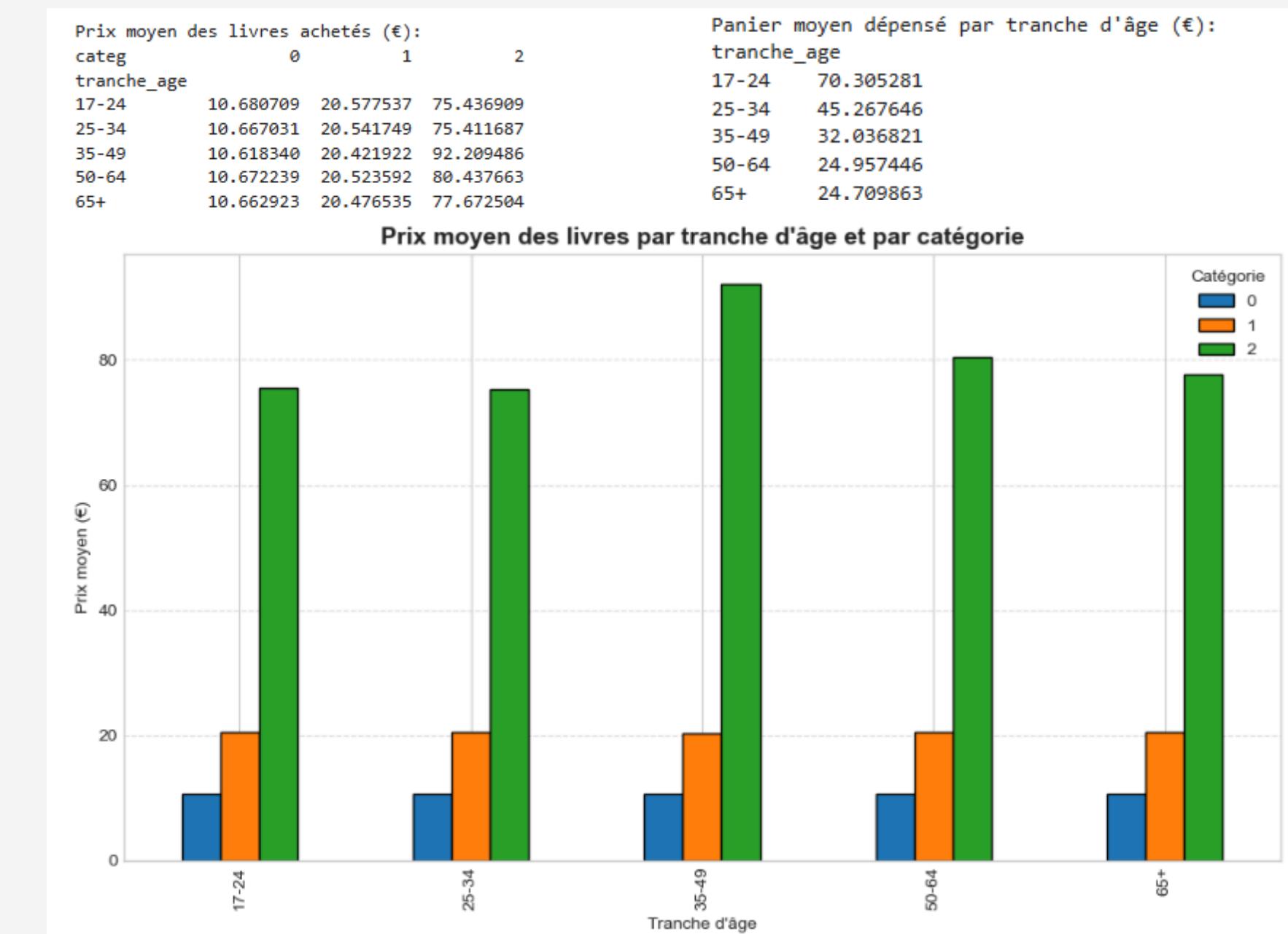
5.2 ANALYSE DES CLIENTS

L'ÂGE ET DEPENSES DES CLIENTS



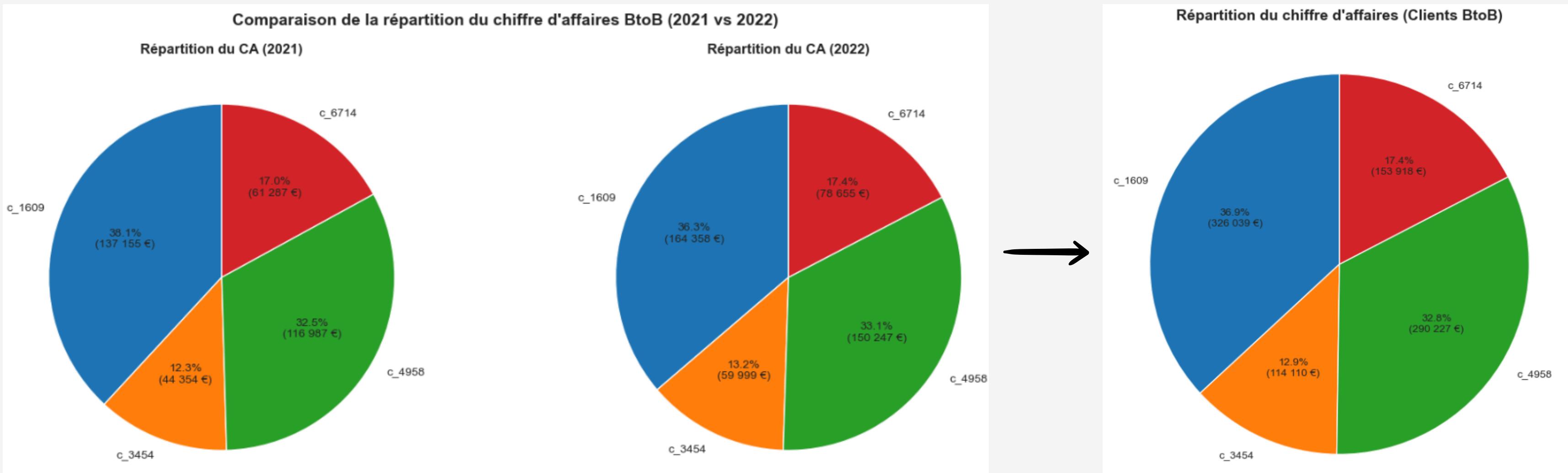
Nombre de clients de moins de 20 ans : 726

Nombre total de clients uniques : 8600



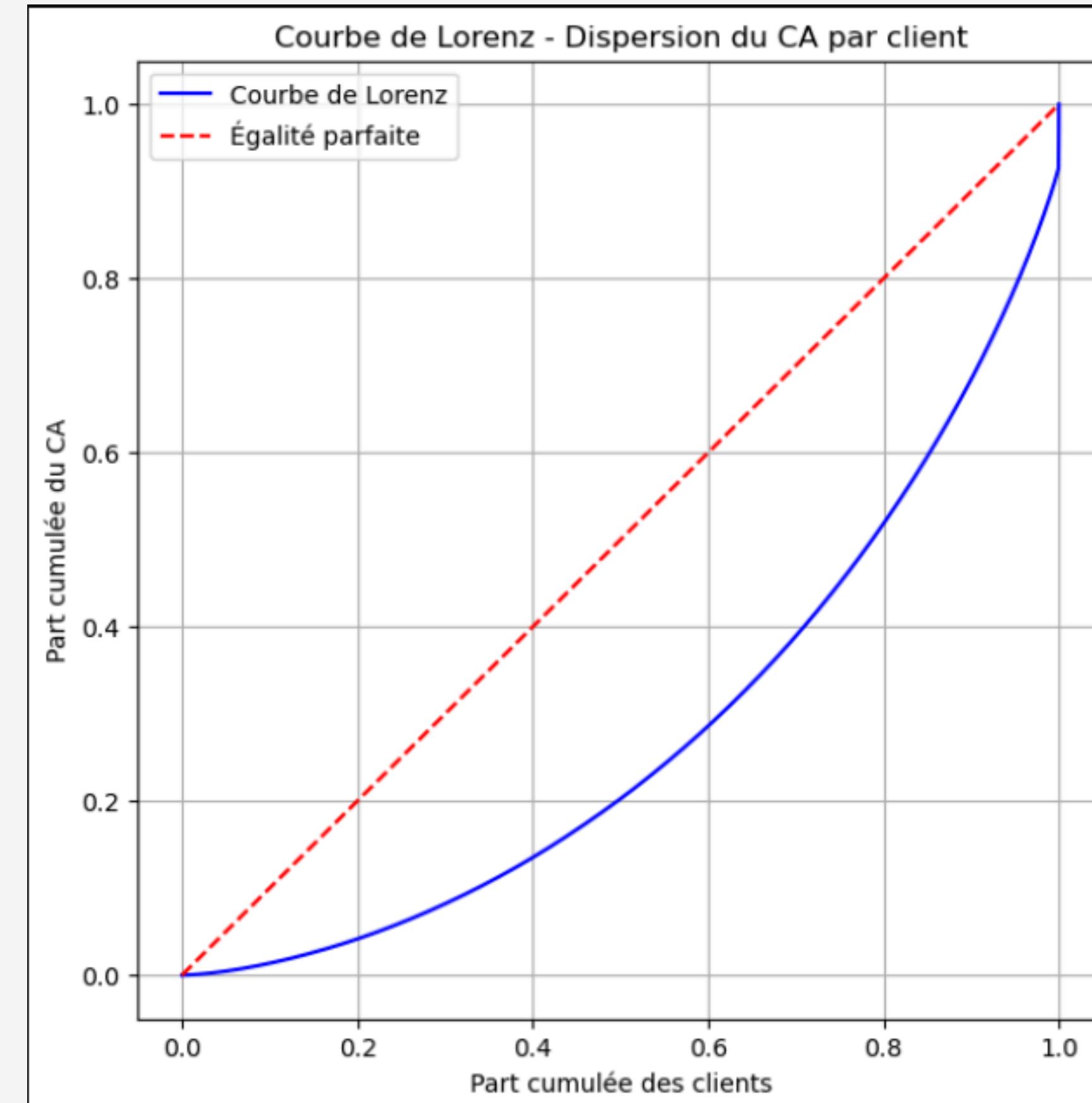
5.2 ANALYSE DES CLIENTS

LES CLIENTS BTOB

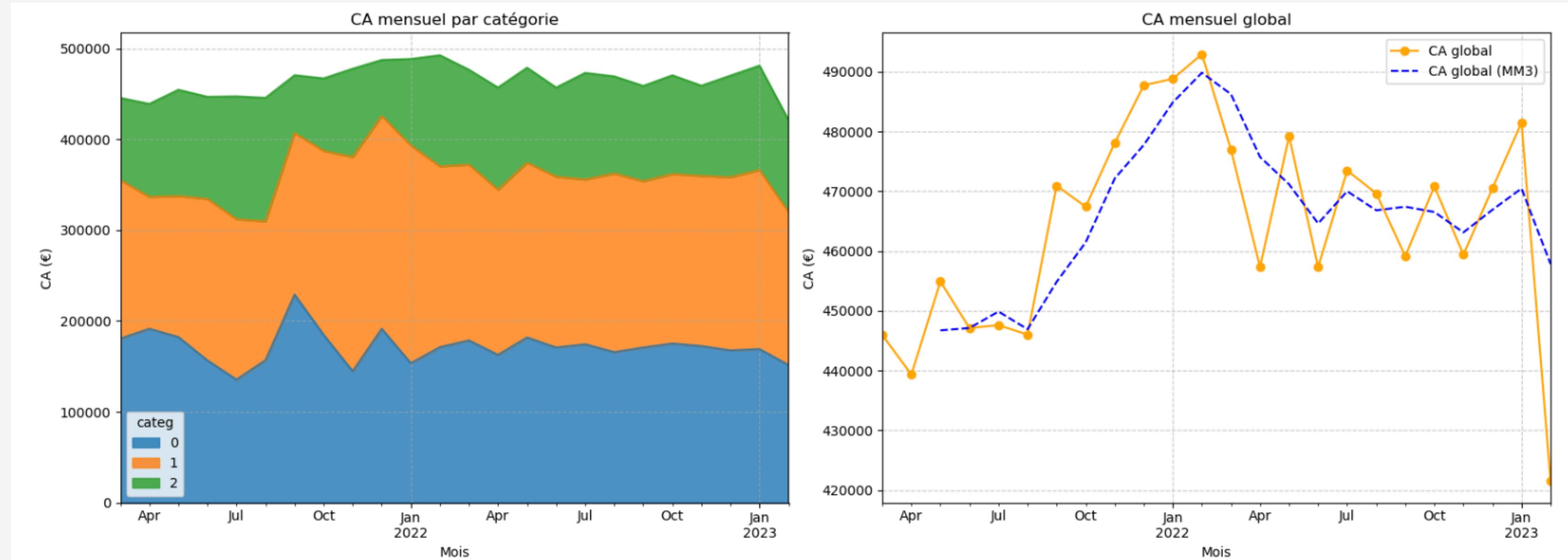


Globalement, la répartition du CA des clients BtoB
est sensiblement la même en 2021 et 2022

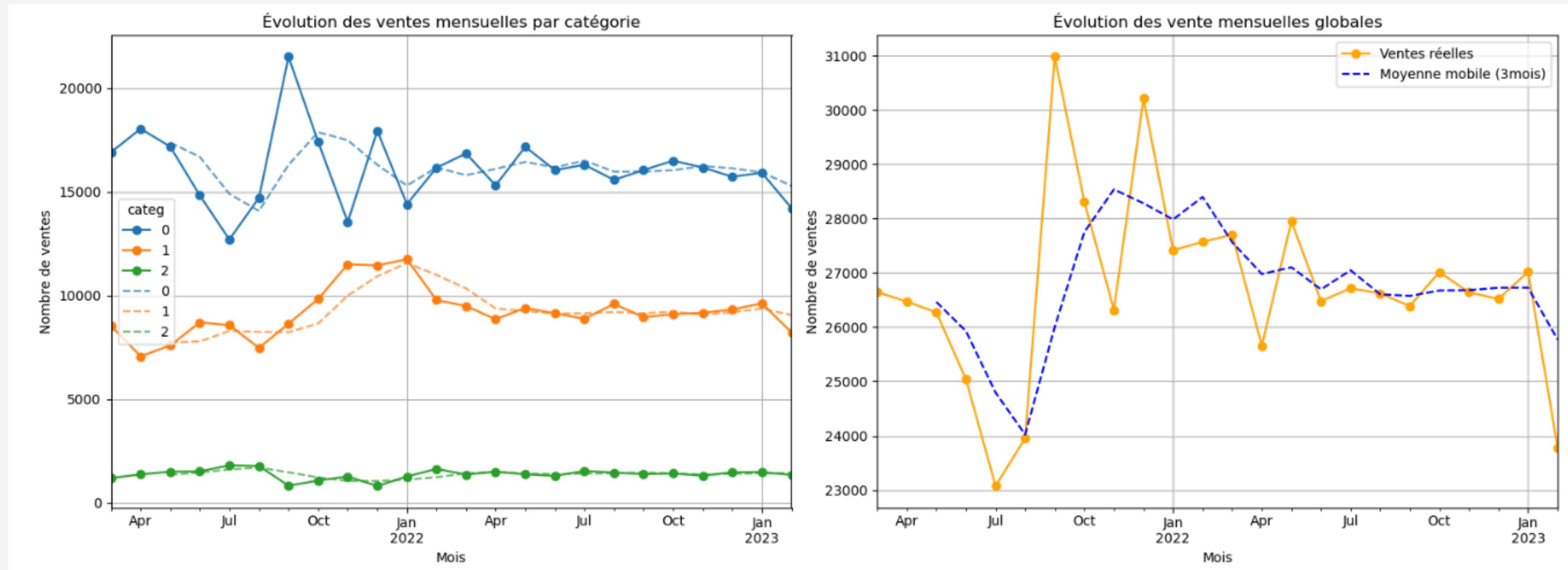
5.2 ANALYSE DES CLIENTS



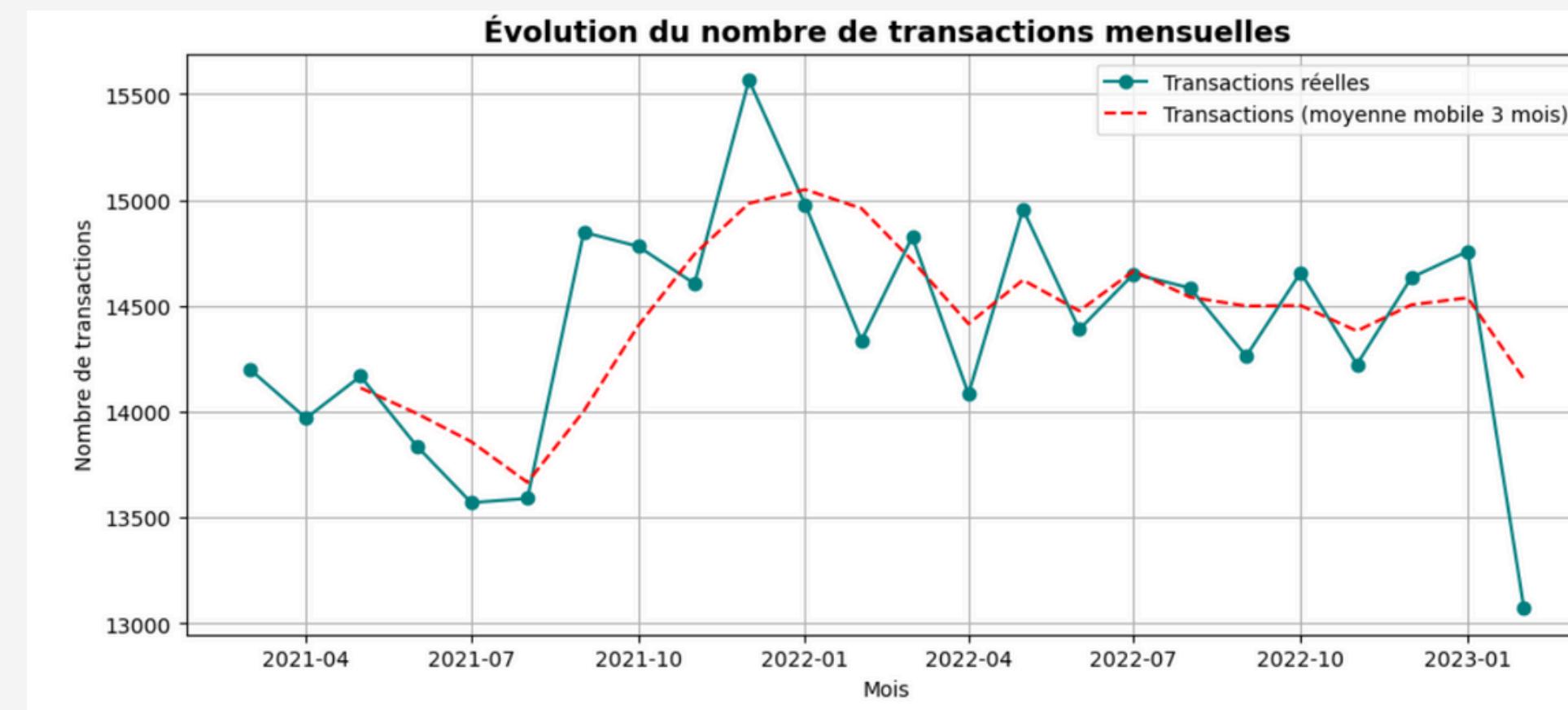
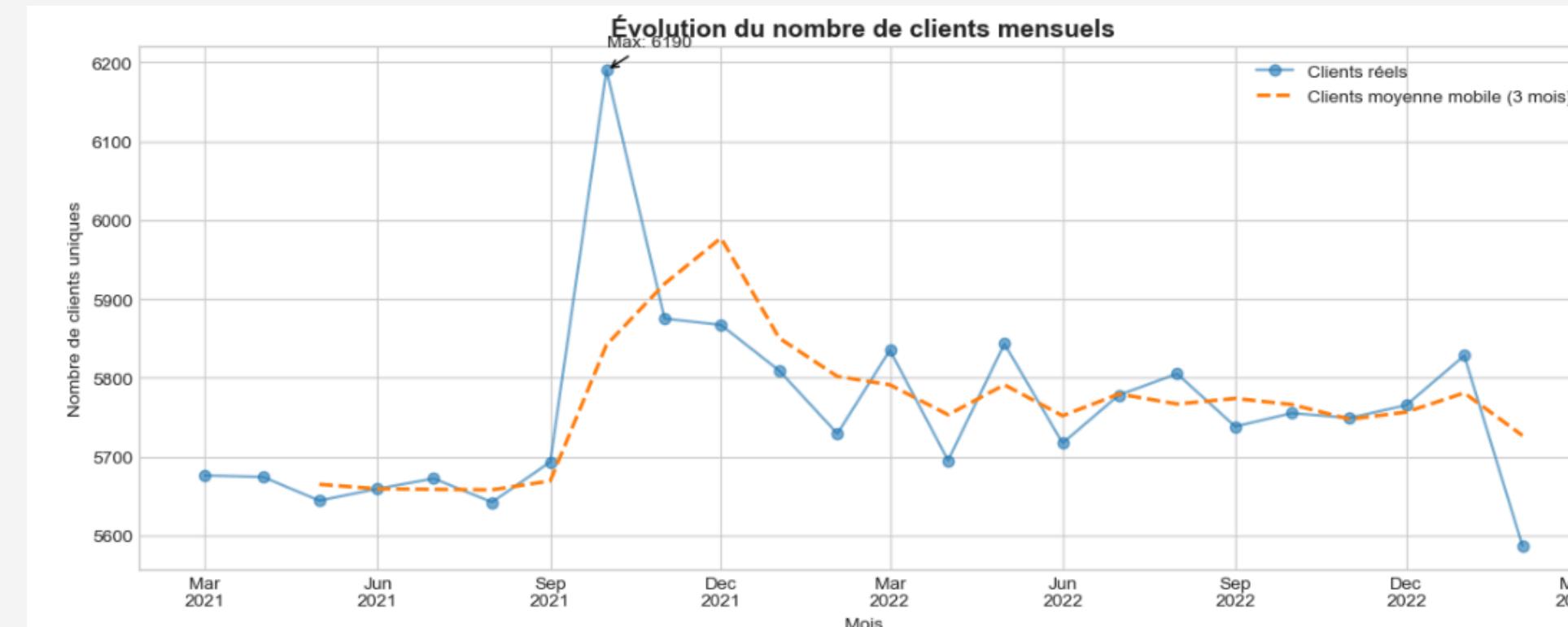
5.3 Analyse des évolutions



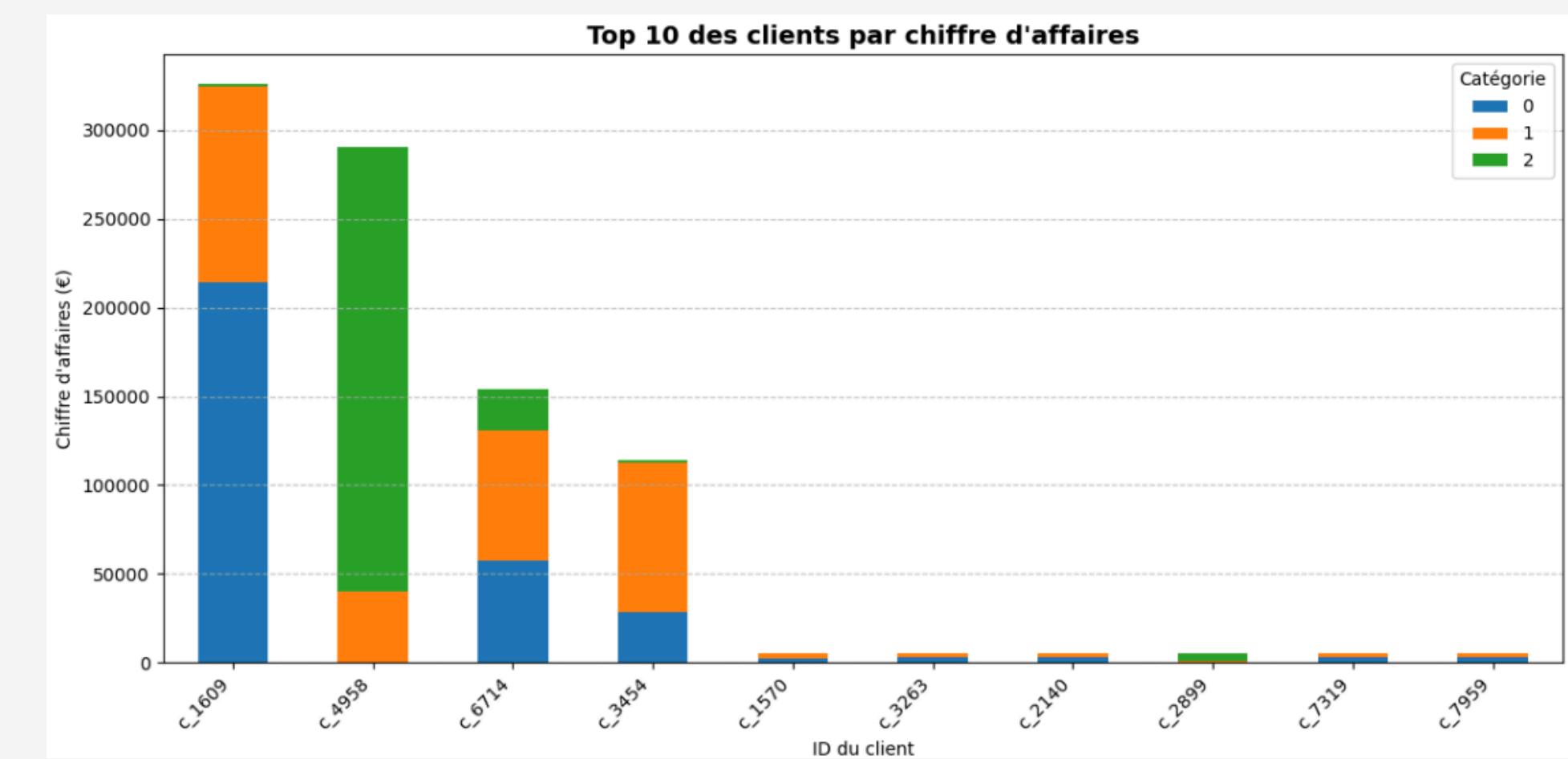
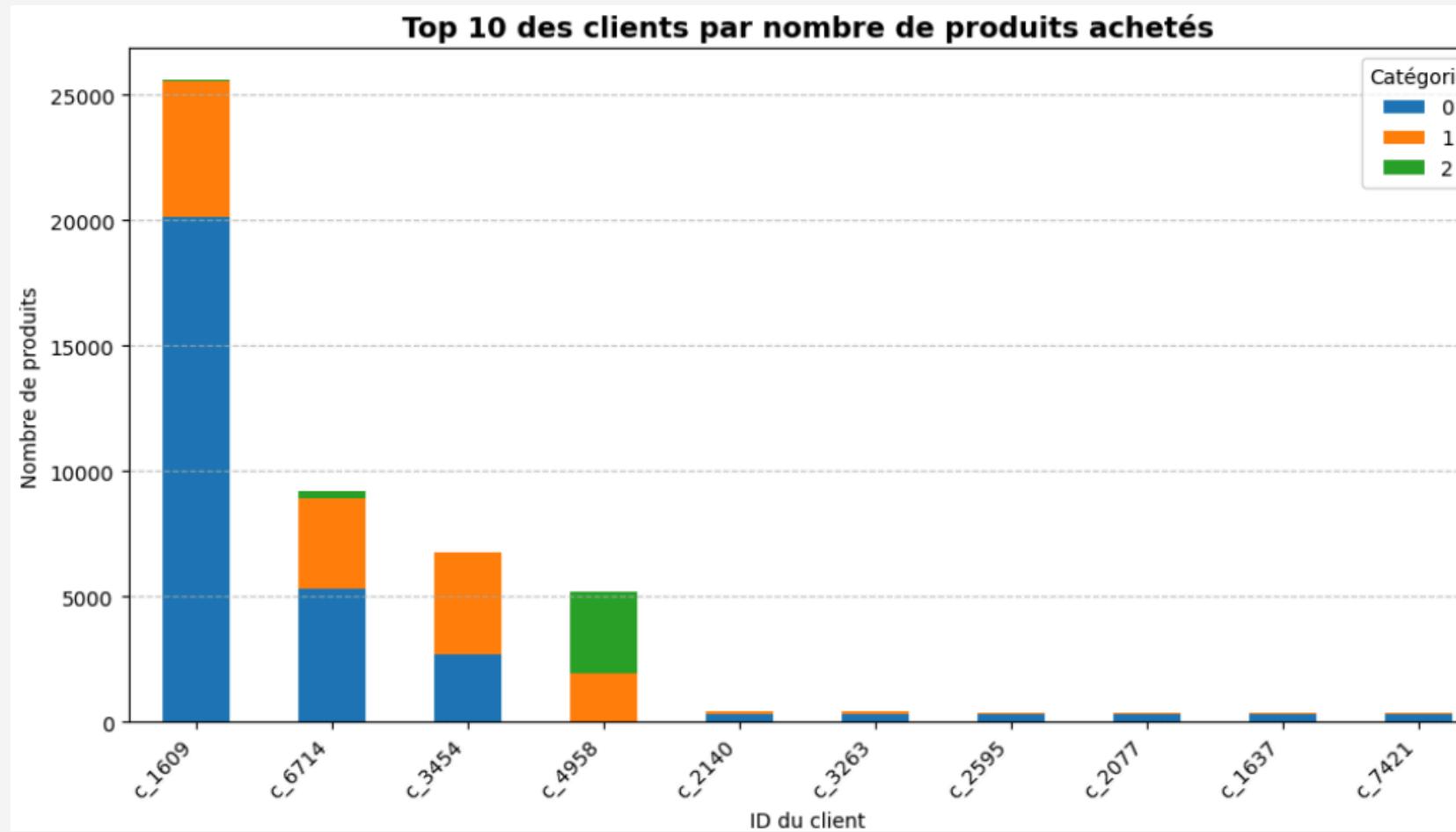
5.3 Analyse des évolutions



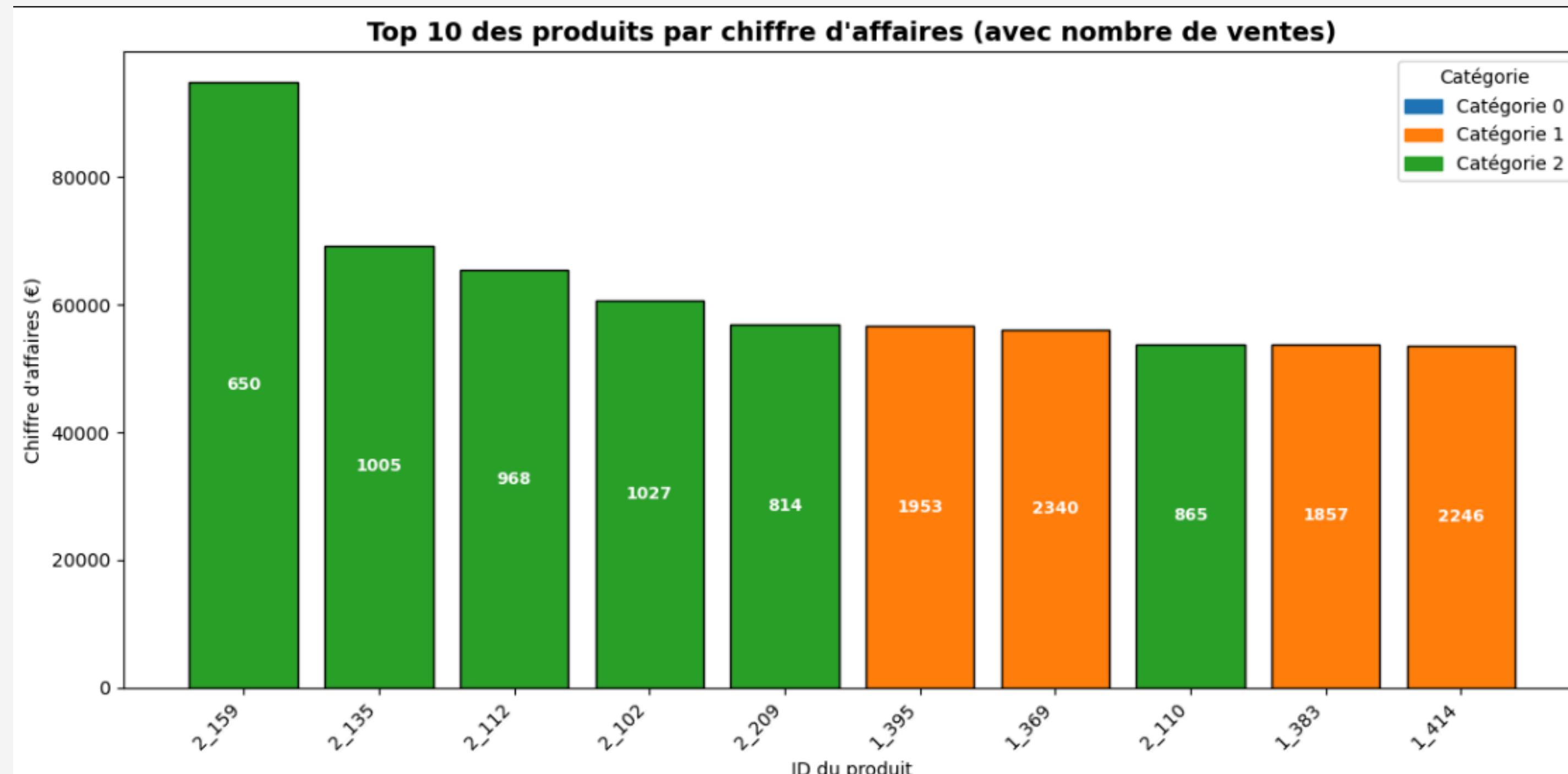
5.3 Analyse des évolutions



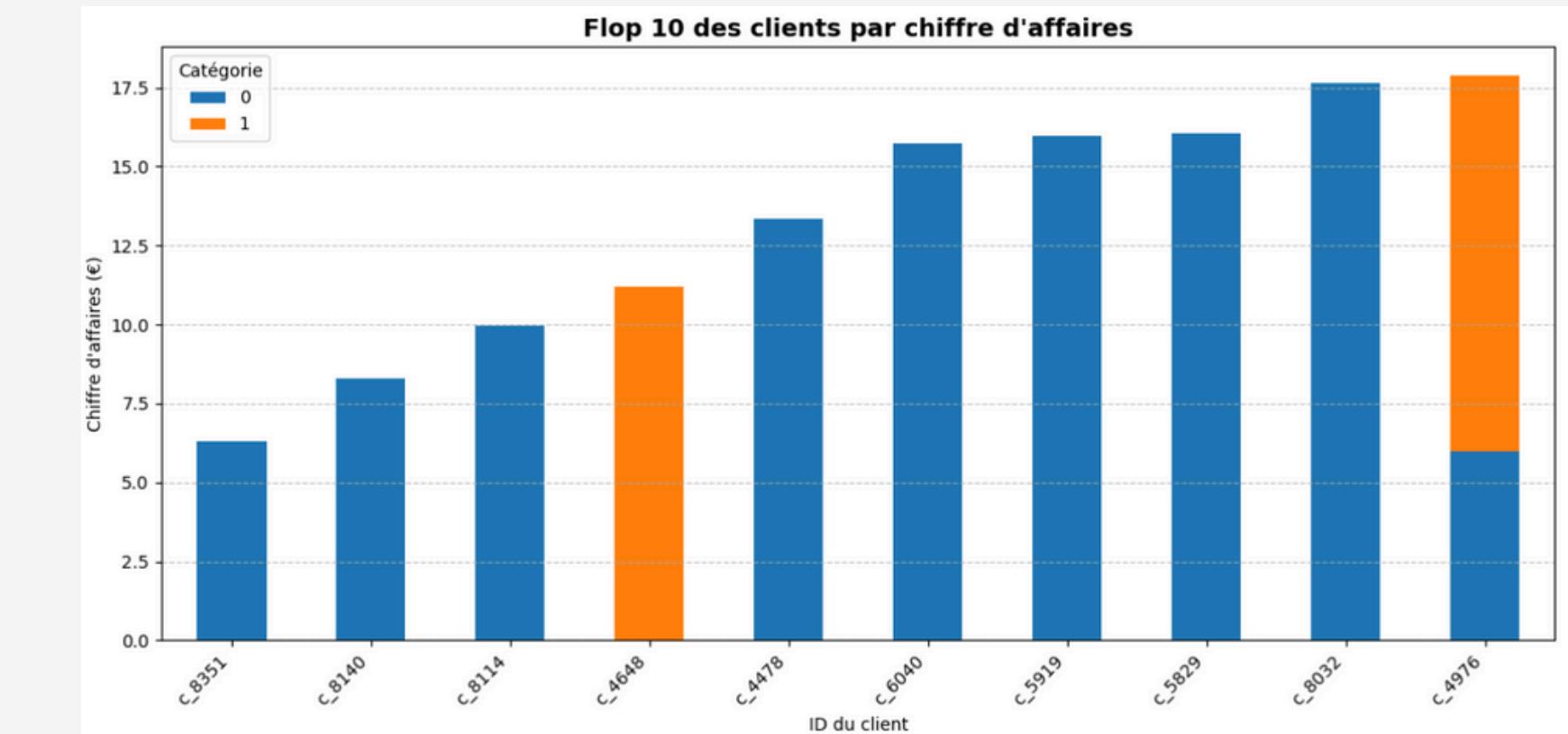
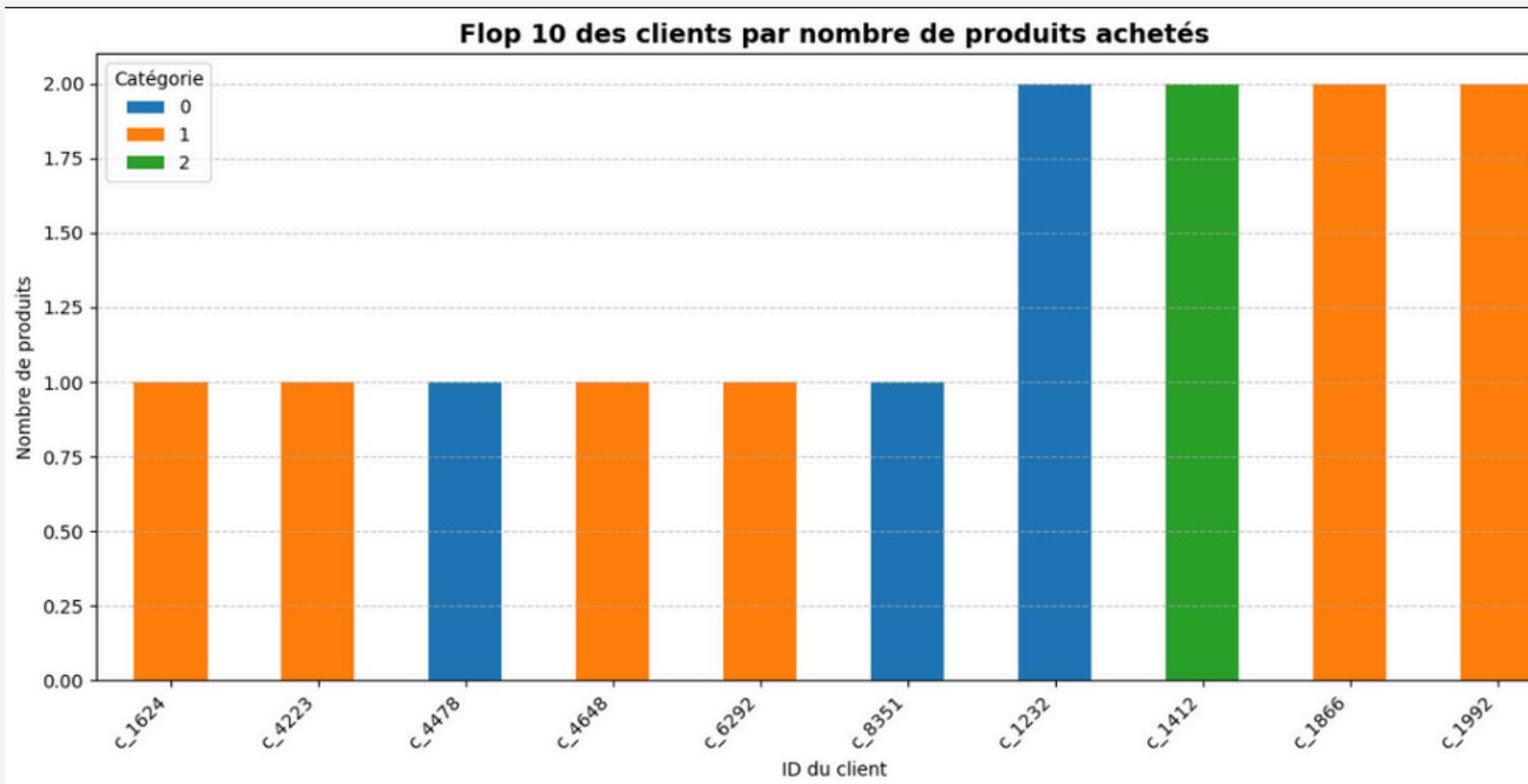
5.4 ANALYSE DES TOPS ET FLOPS



5.4 ANALYSE DES TOPS ET FLOPS

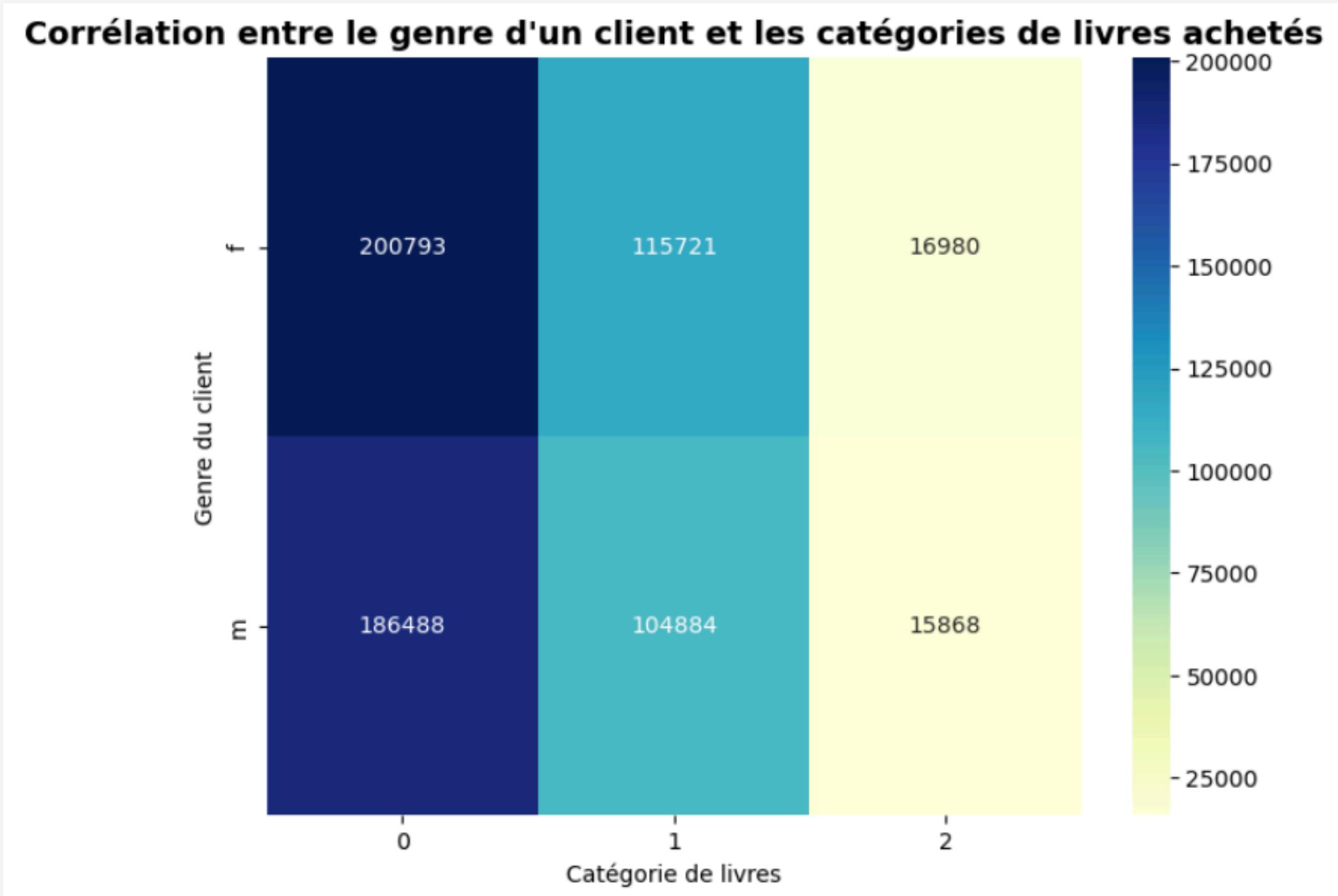


5.4 ANALYSE DES TOPS ET FLOPS



6. CORRÉLATIONS

1) CORRÉLATION ENTRE LE GENRE D'UN CLIENT ET LES CATÉGORIES DE LIVRES ACHETÉS



| categ | 0 | 1 | 2 |
|-------|--------|--------|-------|
| sex | | | |
| f | 200793 | 115721 | 16980 |
| m | 186488 | 104884 | 15868 |

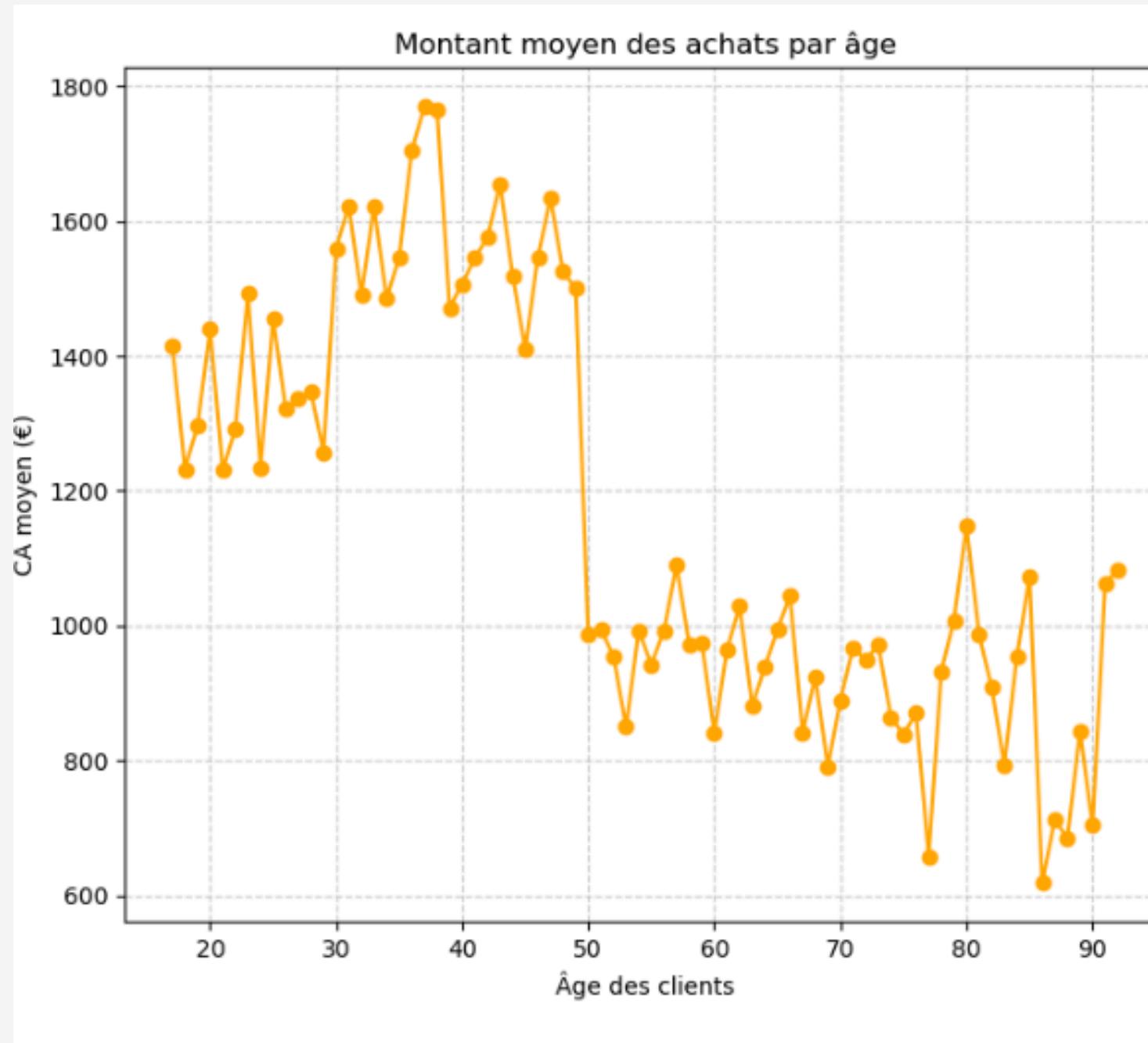
Statistique de Chi² : 22.669
Degrés de liberté : 2
p-value : 0.000012

Le test du Chi² montre qu'il existe une **différence significative** entre hommes et femmes dans la répartition des catégories

Donc **le genre et la variable catégorie ne sont pas indépendants**

6. CORRÉLATIONS

2) CORRÉLATION ENTRE L'ÂGE DES CLIENTS ET LE MONTANT TOTAL DES ACHATS



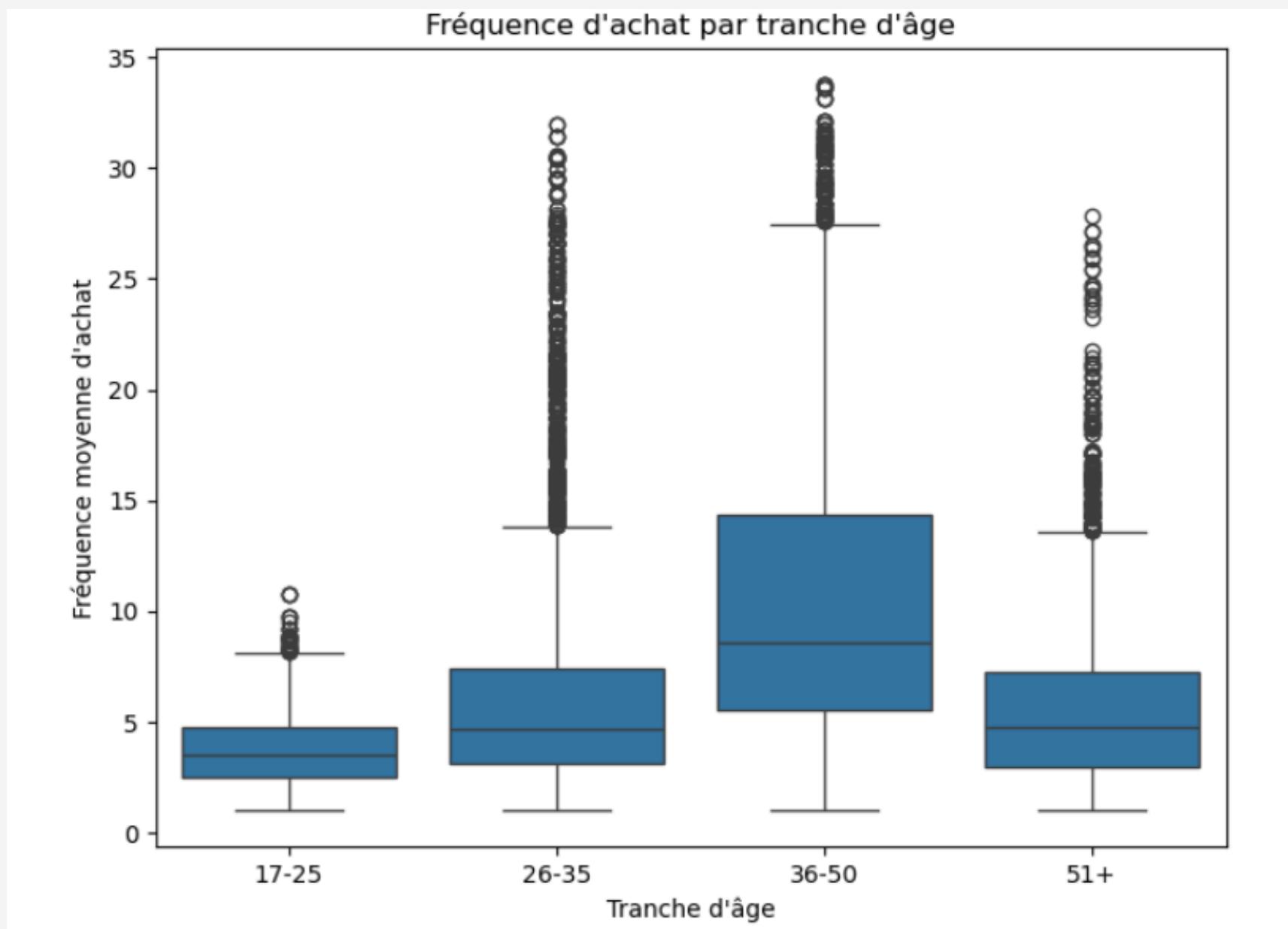
Corrélation de Spearman : $r = -0.185$, $p\text{-value} = 0.00$

Le coefficient de Spearman $r = -0.185$ indique une corrélation monotone négative faible : plus l'âge augmente, moins le montant total des achats est élevé.

La $p\text{-value} = 0.00$ (inférieure à 0.05) signifie que cette relation est statistiquement significative. Elle n'est pas due au hasard => Il y a une corrélation faible ce n'est pas une relation forte.

6. CORRÉLATIONS

3) CORRÉLATION ENTRE L'ÂGE DES CLIENTS ET LA FRÉQUENCE D'ACHAT



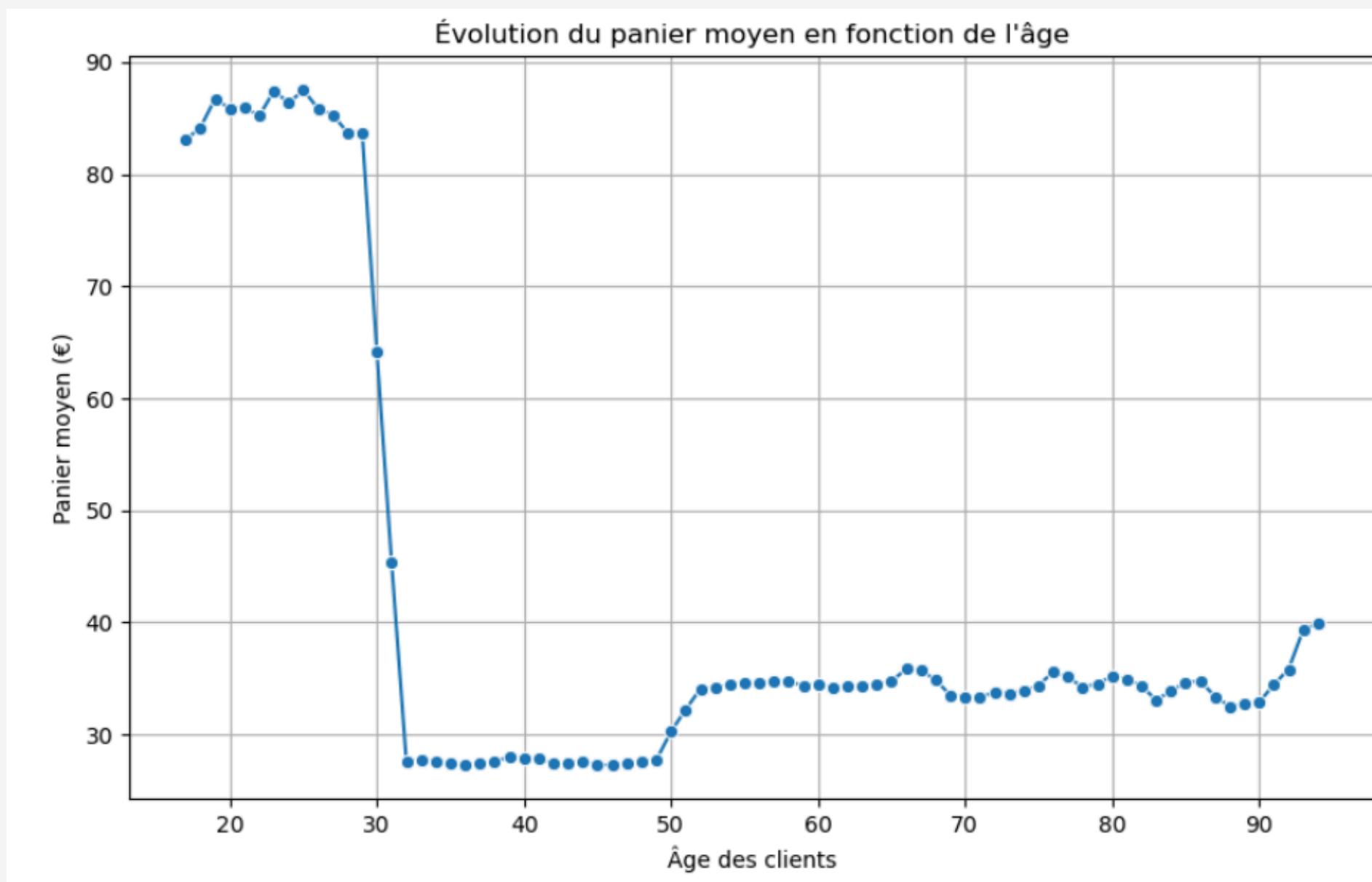
Corrélation de Spearman : 0.0899
p-value : 0.000000

Statistique ANOVA : 2198.2394
p-value : 0.000000

- L'ANOVA compare les moyennes de fréquence d'achat entre les tranches d'âge.
 - La statistique obtenue est très élevée (2198.24), ce qui montre une grande différence entre groupes.
 - La p-value est quasiment nulle (< 0.000001).
- => Cela signifie que l'âge a un impact significatif sur la fréquence d'achat.
- Les moyennes de fréquence d'achat ne sont pas identiques selon les tranches d'âge

6. CORRÉLATIONS

4) CORRÉLATION ENTRE L'ÂGE DES CLIENTS ET LA TAILLE DU PANIER MOYEN



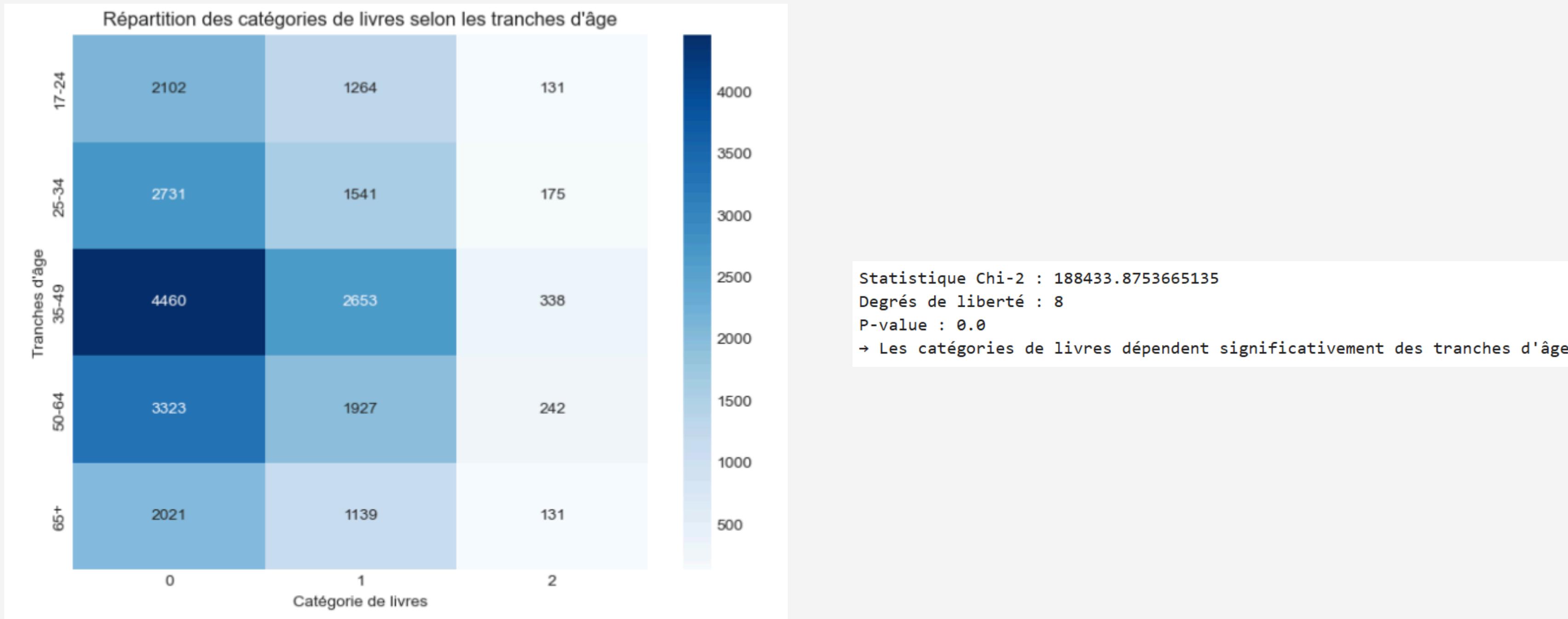
Statistique ANOVA : 4358.1033

p-value : 0.000000

- On rejette très fortement l'hypothèse d'égalité des moyennes
- Le panier moyen diffère significativement selon les tranches d'âge
- En d'autres termes, l'âge influence clairement le panier

6. CORRÉLATIONS

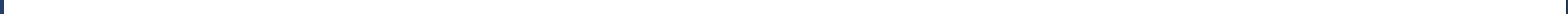
5) CORRÉLATION ENTRE L'ÂGE ET LA CATÉGORIE DE LIVRES ACHETÉS



7. CONCLUSION

- **Le chiffre d'affaires** est principalement **généré par les 25-50 ans**, avec une majorité situé dans la tranche 35-49 ans
- **Les catégories 0 et 1** sont les plus populaires auprès des clients
- **La catégorie 2** enregistre un **volume de ventes plus faible**, cependant c'est cette catégorie la plus achetée par **les 17 - 34 ans**
- **Les 17-24 ans achètent moins en volume** mais présentent un panier moyen élevé de **70 €**, suivi par les **25-34 avec un panier de 45€**

La tranche d'âge 17-34 représente donc une opportunité stratégique à explorer



**MERCI POUR VOTRE
ATTENTION**

