

**OPENCLASSROOMS**

**VENTES EN LIGNE**  
LIBRAIRIE LAPAGE



# INTRODUCTION

❑ METHODOLOGIE

❑ INDICATEURS DE VENTE

❑ ANALYSES CLIENTS



# LES DONNÉES

## - MÉTHODOLOGIE -

### ➤ Base de données

#### ➤ Anomalies détectées

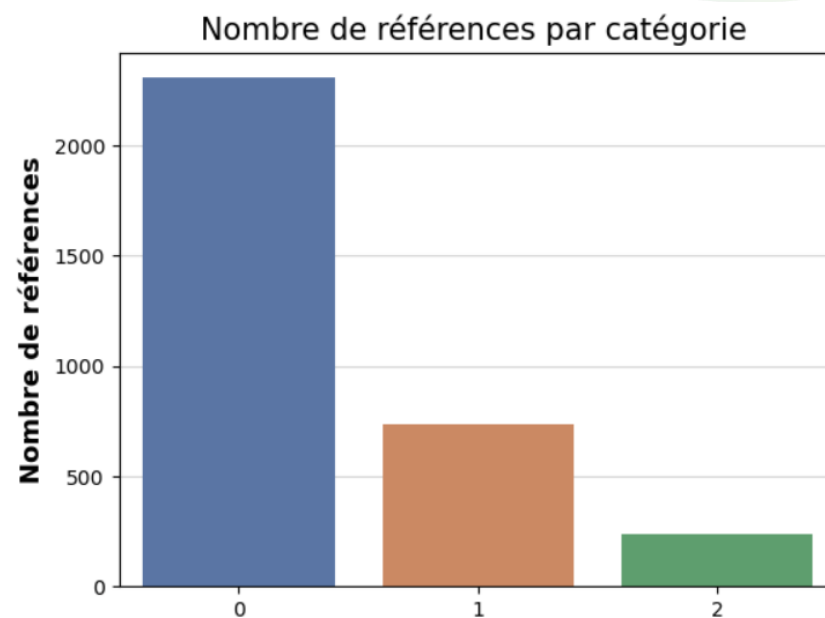
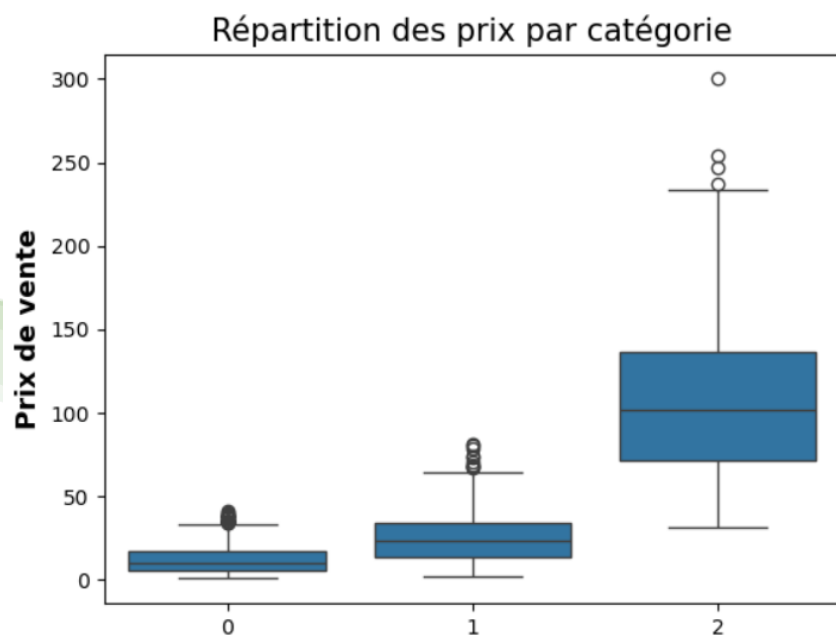
- 1 produit vendu n'apparaît pas au catalogue  
= pas de prix = pas comptabilisé
- ref client, ref produit, prix
- oct.2021 = catégorie 1



# LES DONNÉES

## - INDICATEURS DE VENTE -

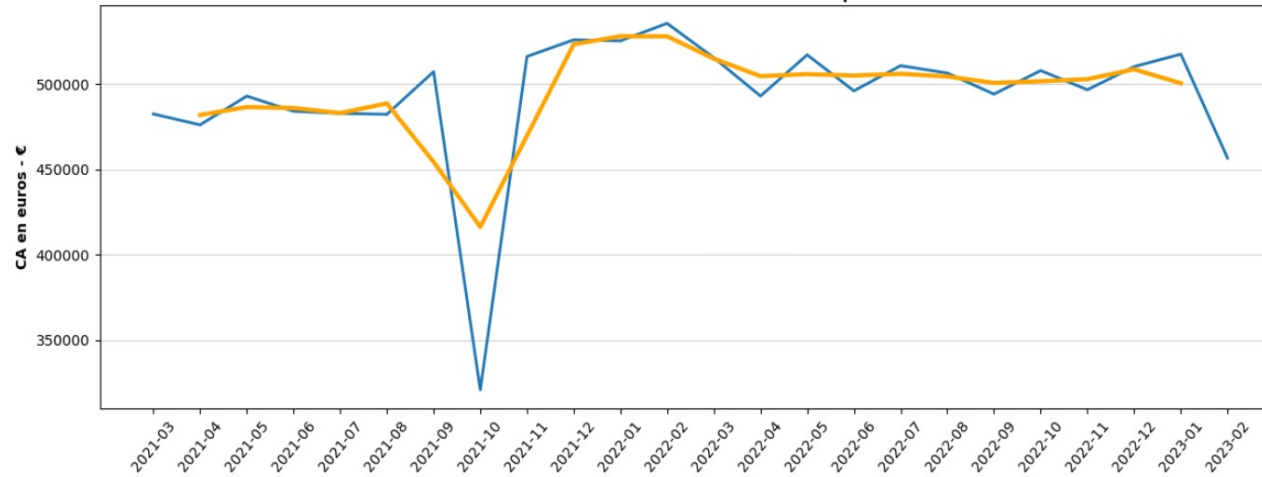
- Beaucoup de réf aux prix les moins chers
- Peu de réf parmi les plus chères



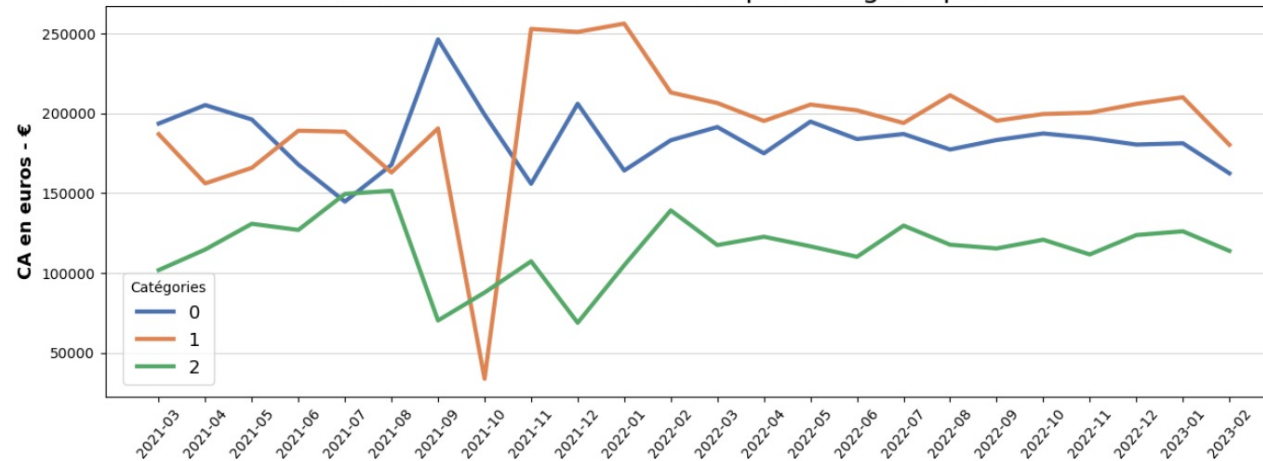


## - INDICATEURS DE VENTE -

Évolution du chiffre d'affaires par mois



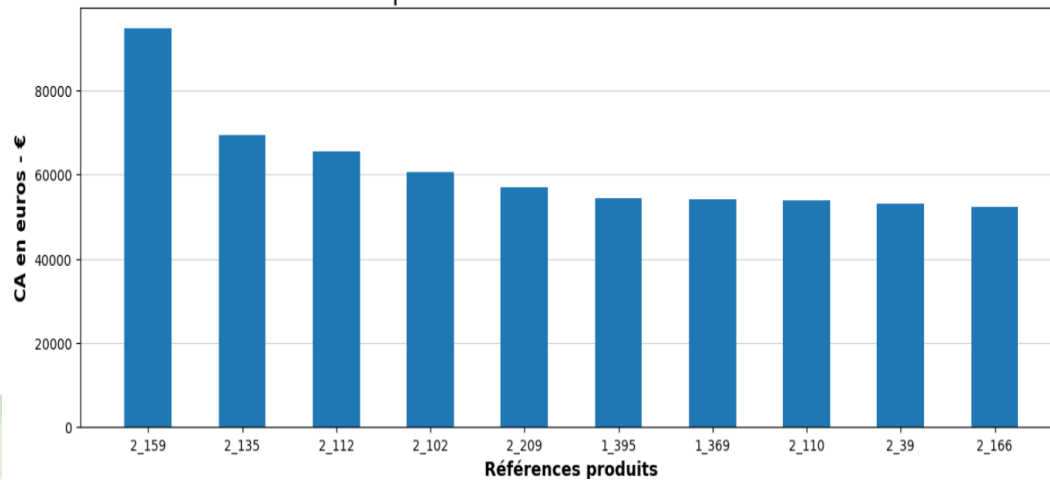
Évolution du chiffre d'affaires par catégorie par mois



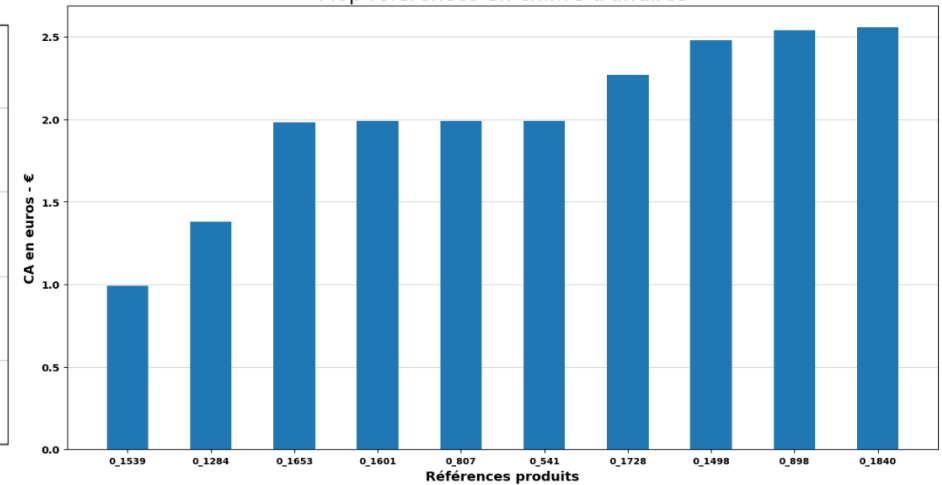
Anomalie catégorie 1  
Octobre 2021

## - INDICATEURS DE VENTE -

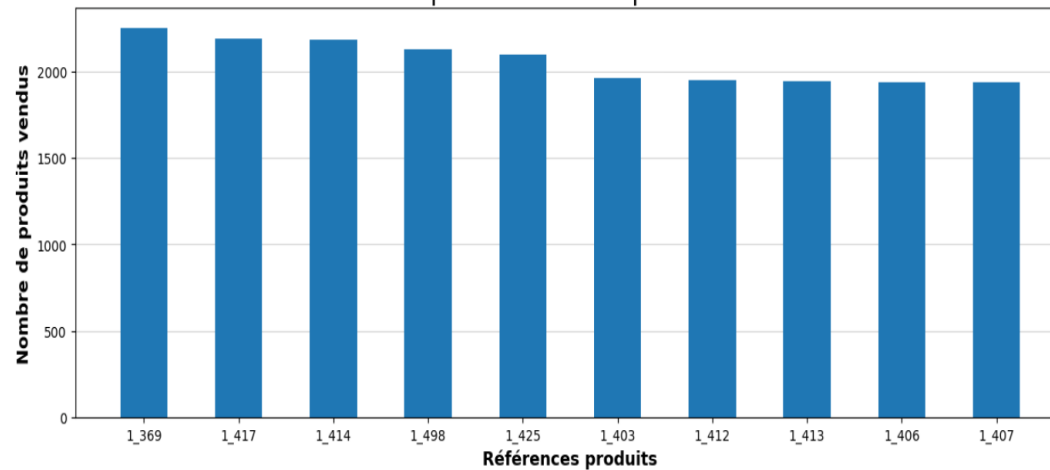
Top références en chiffre d'affaires



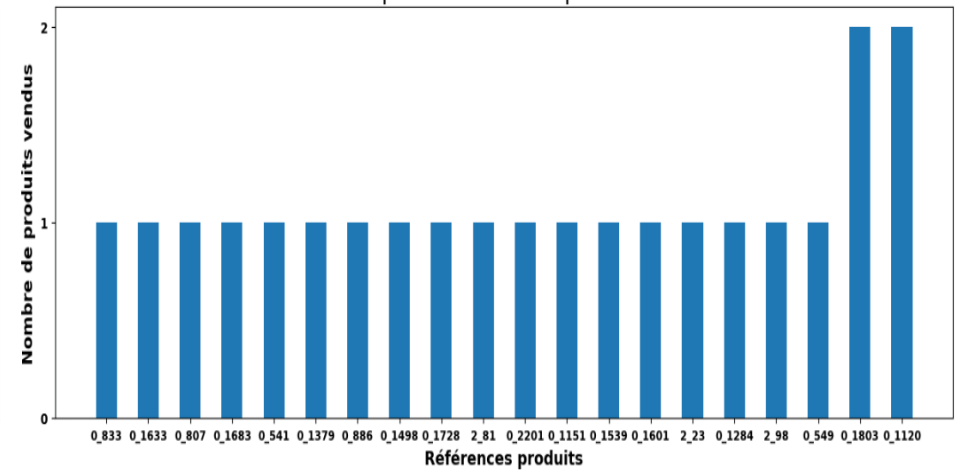
Flop références en chiffre d'affaires



Top références en quantités



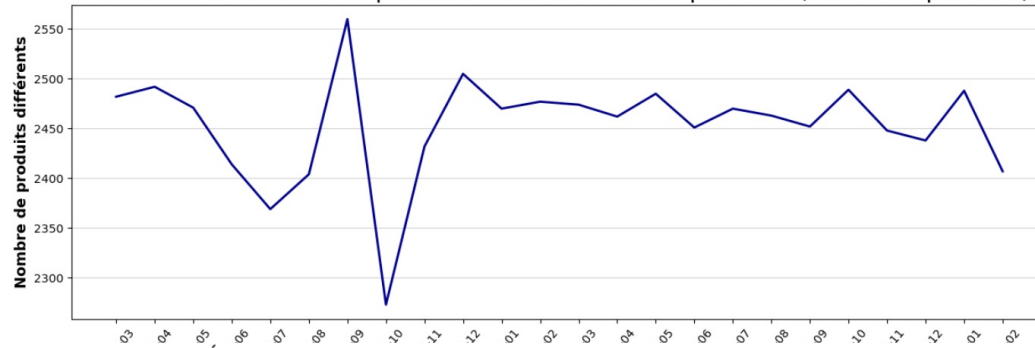
Flop références en quantités



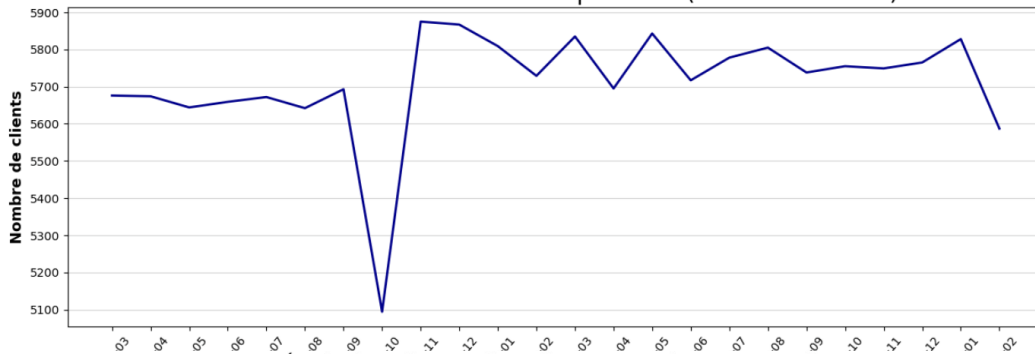
# LES DONNÉES

## - INDICATEURS DE VENTE -

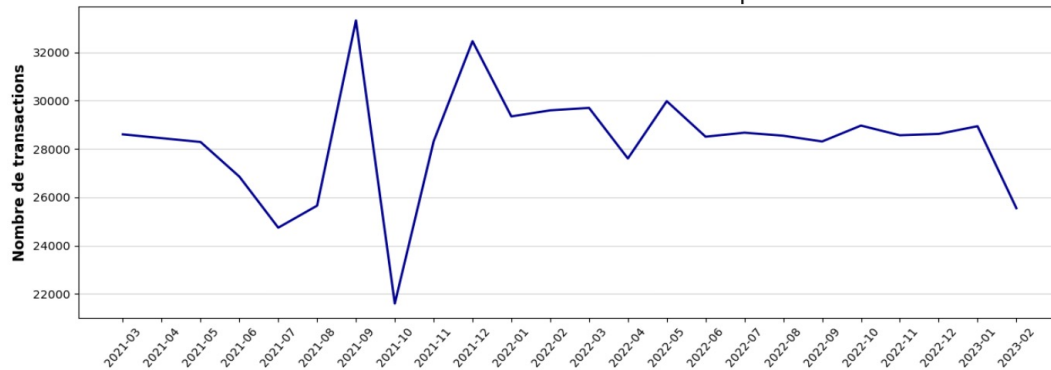
Evolution du nombre de produits différents vendus par mois (sur 3 286 produits)



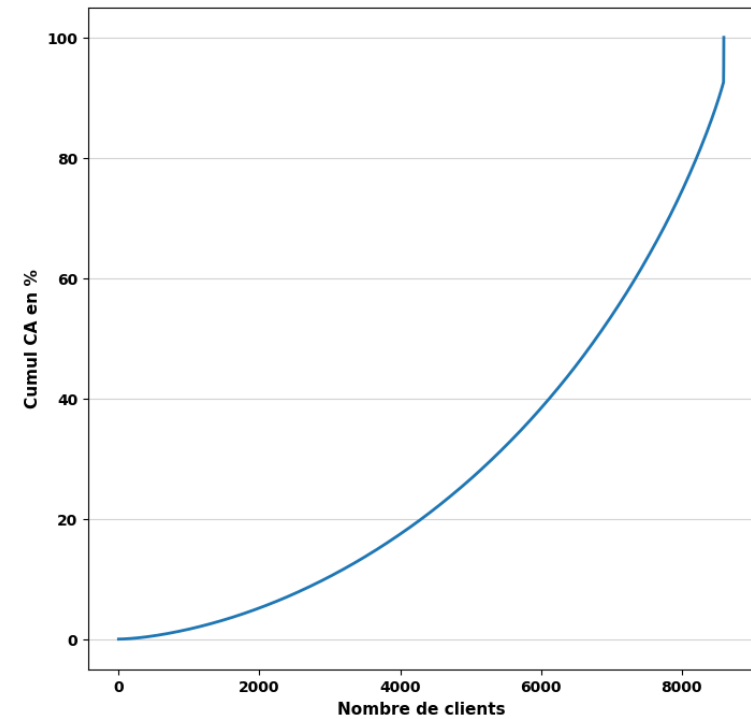
Evolution du nombre de clients par mois (sur 8 621 clients)



Evolution du nombre de transactions par mois



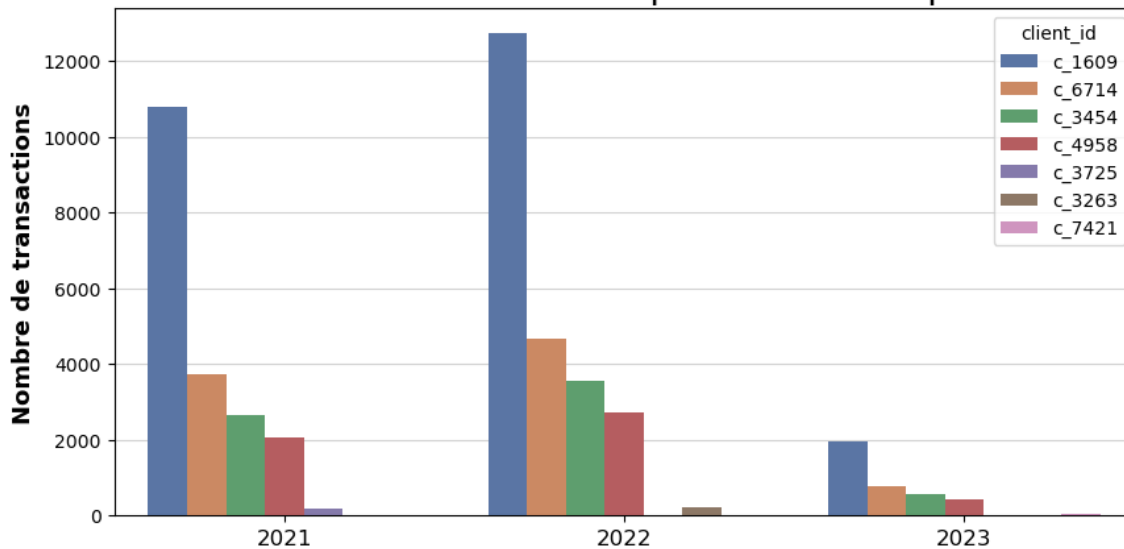
Courbe de Lorenz





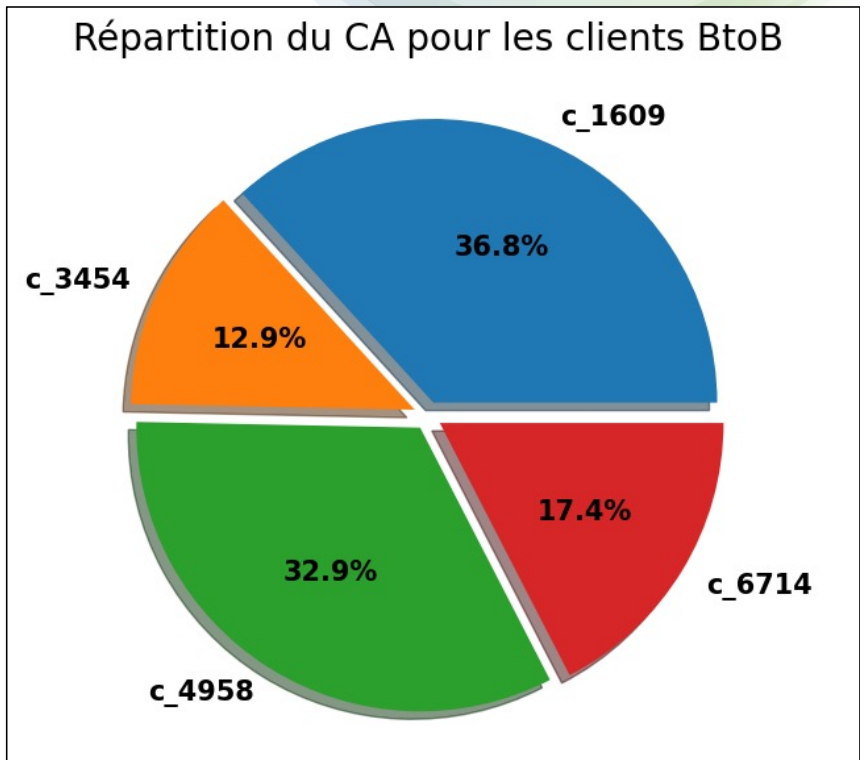
# LES DONNÉES - ANALYSES CLIENTS -

Nombre de transactions : Top 5 des clients par année



➤ 4 clients identiques ressortent chaque année

Répartition du CA pour les clients BtoB

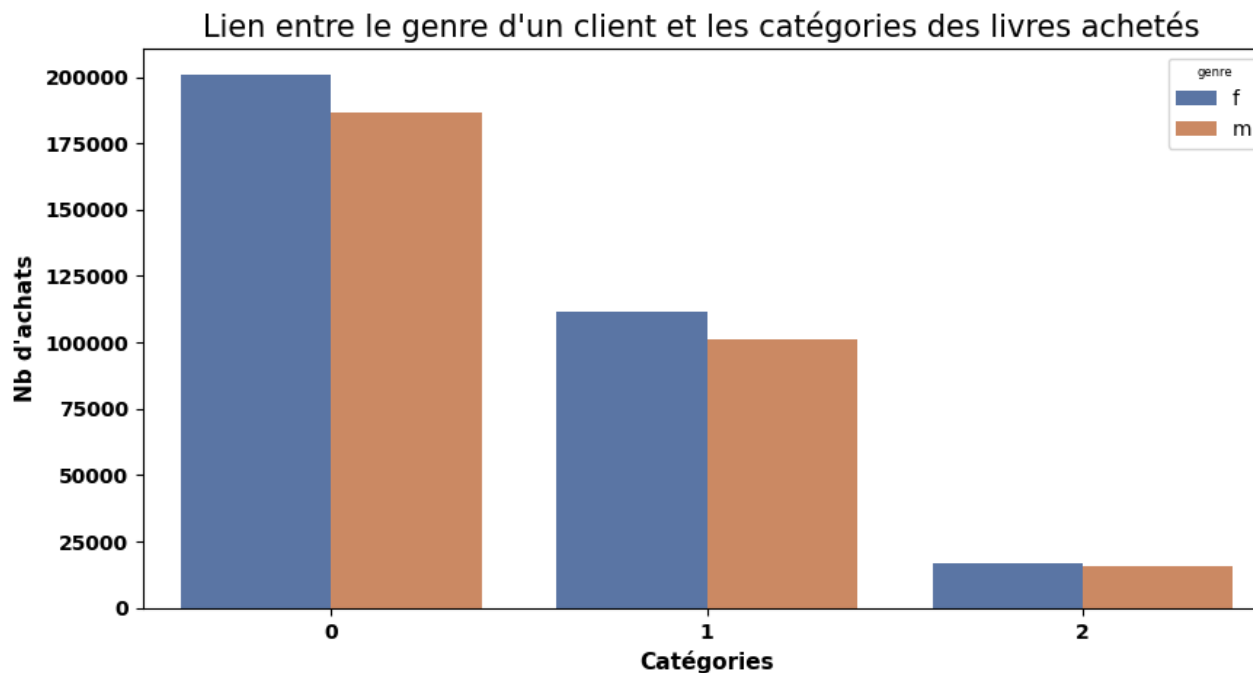






## - ANALYSES CLIENTS -

### - 2 VARIABLES QUALITATIVES – TEST DU $\chi^2$



➤  $H_0$  : p-value > 5% ,  
il n'y a pas dépendance

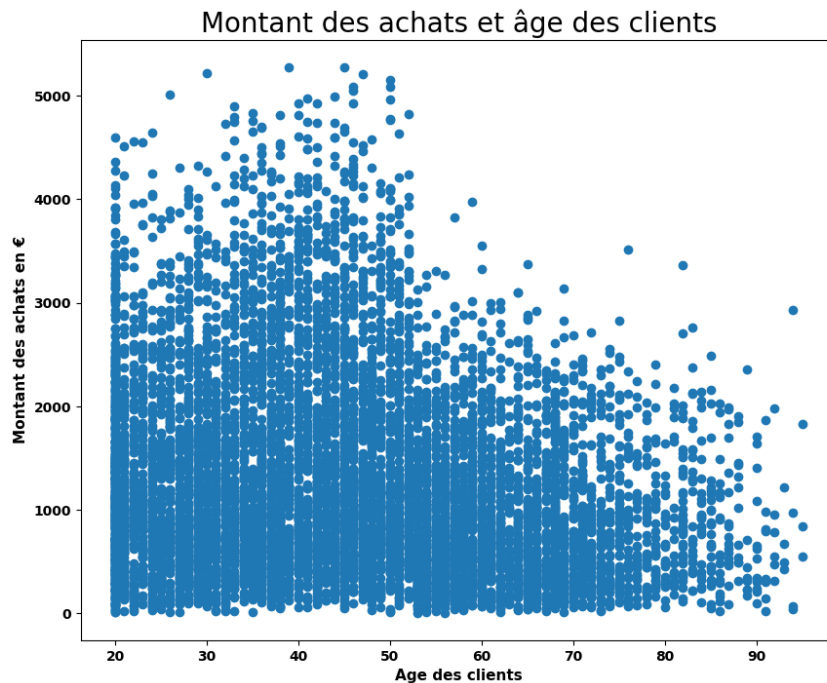
➤  $H_1$  : p-value < 5% ,  
il y a corrélation

Appliquons le test du  $\chi^2$  afin de  
déterminer la p-value : 0.002537

➤ Pour un seuil de 5%, nous  
pouvons dire que les 2  
variables sont corrélées



## 2 VARIABLES QUANTITATIVES TEST DE PEARSON

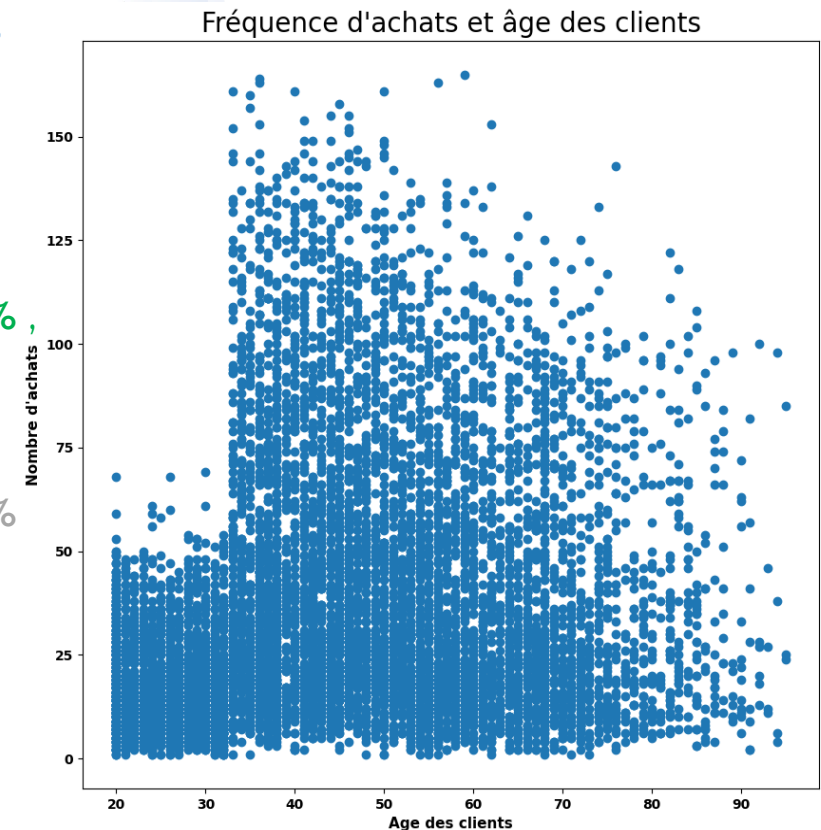


➤  $H_0 : p\text{-value} > 5\%$  ,  
il n'y a pas de  
corrélation linéaire

➤  $H_1 : p\text{-value} < 5\%$   
il y a corrélation  
linéaire

Appliquons le test de Pearson afin de  
déterminer la p-value : 0.1884

➤ Pour un seuil de 5%, nous pouvons  
dire que les 2 variables ne sont pas  
corrélées linéairement



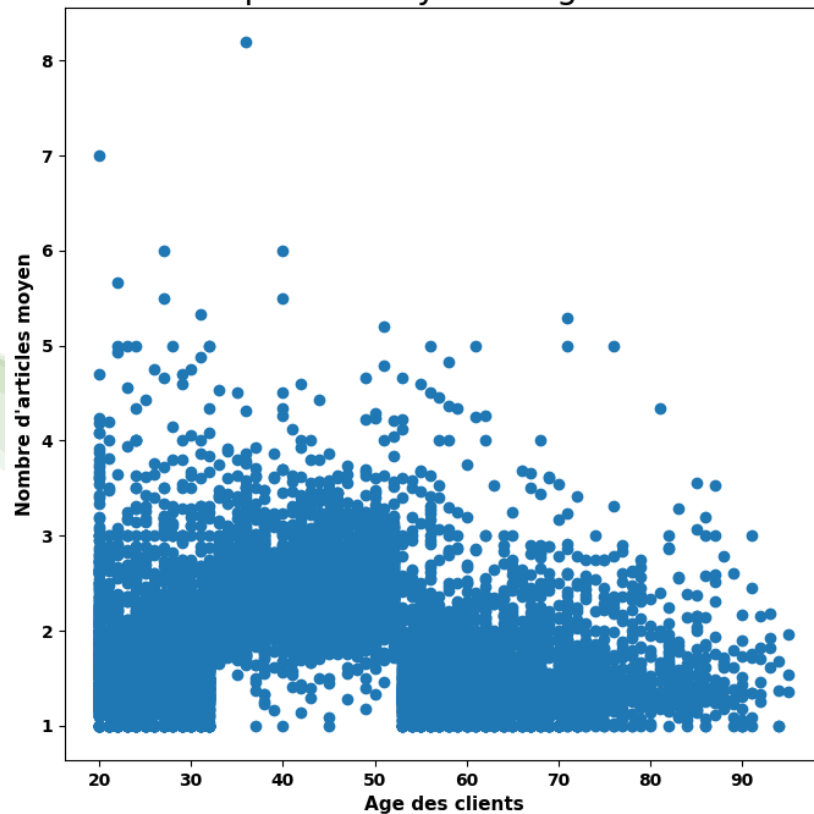
Appliquons le test de Pearson afin de  
déterminer la p-value : 0.1646

➤ Pour un seuil de 5%, nous pouvons  
dire que les 2 variables ne sont pas  
corrélées linéairement



## - ANALYSES CLIENTS — 2 VARIABLES QUANTITATIVES

Taille du panier moyen et âge des clients



➤  $H_0$  :  $p\text{-value} > 5\%$  ,  
il n'y a pas de corrélation linéaire

➤  $H_1$  :  $p\text{-value} < 5\%$   
il y a corrélation linéaire

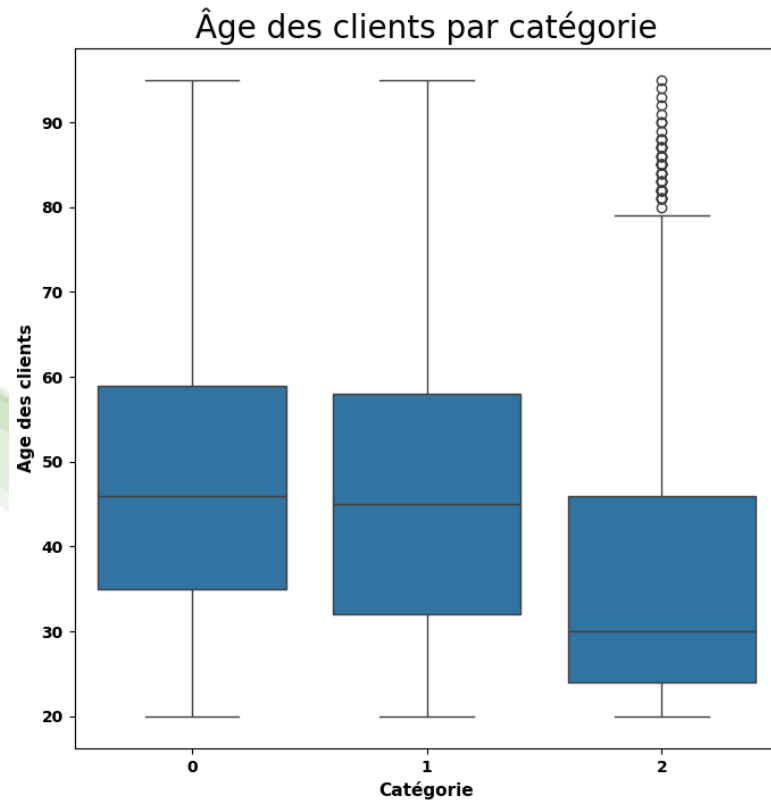
Appliquons le test de Pearson afin de  
déterminer la  $p\text{-value}$  : -0.212

➤ Pour un seuil de 5%, nous pouvons  
dire que les 2 variables ne sont pas  
corrélées linéairement



## - ANALYSES CLIENTS -

### 1 VARIABLE QUANTITATIVE & 1 VARIABLE QUALITATIVE



➤  $H_0$  :  $p\text{-value} > 5\%$  ,  
il n'y a pas dépendance

➤  $H_1$  :  $p\text{-value} < 5\%$  ,  
il y a corrélation

Appliquons le test ANOVA afin de  
déterminer la  $p\text{-value}$  :  $2.06 \times 10^{-187}$

➤ Pour un seuil de 5%, nous  
pouvons dire que les 2  
variables sont corrélées



## CONCLUSION

- PARMi CES ANALYSES -

- ✓ Anomalie dans les données (catégorie 1 – oct 2021)
- ✓ 4 clients BtoB
- ✓ Corrélations en fonction de l'âge du genre des clients