

Table des matières

01

Nettoyage des fichiers

02

Analyses bivariées et univariées

03

Tests de correlations





NETTOYAGE DES FICHIERS

Produit N°0_2245
rajouté à la table
'products'.

Total = 3287

3

Nouveau
tableau('tab') créée à
partir de des 3 tables
('transaction',
'customers',
'products')

Suppression des ligne dédiées aux tests, dont valeur inférieur à -1.

4

Ajout de colonnes permettant une meilleure visibilité.

2



Produits les plus vendus

id_prod

1_369: 2252

1_417: 2189

1_414: 2180

1_498: 2128

1_425:2096

Produits les moins vendus

id_prod

0_1151: 1

0_1284: 1

0_1379: 1

0_1498: 1

0_1539: 1

Top 5 clients

client_id

c_1609: 324033.35

c_4958: 289760.34

c_6714: 153662.70

c_3454: 113669.82

c_3263: 5276.87

Flop 5 clients

client_id

c_8351: 6.31

c_8140: 8.30

c_8114: 9.98

c_240 : 11.06

c_4648: 11.20

Moyenne d'achat selon le sexe

categ sex

0 f 44.257550

h 43.543249

1 f 48.362318

h 46.923506

2 f 26.571255

h 25.387031

CA selon le genre:

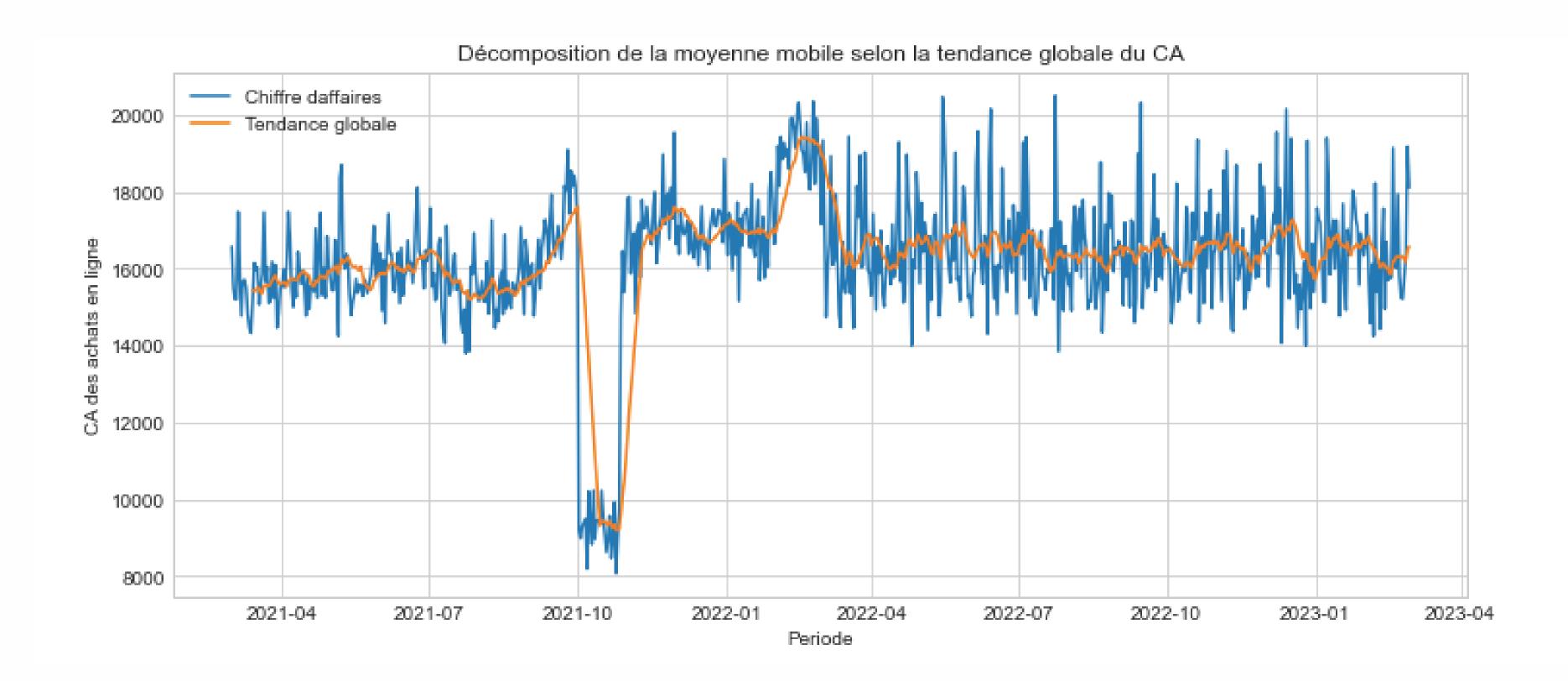
f: 49,4%

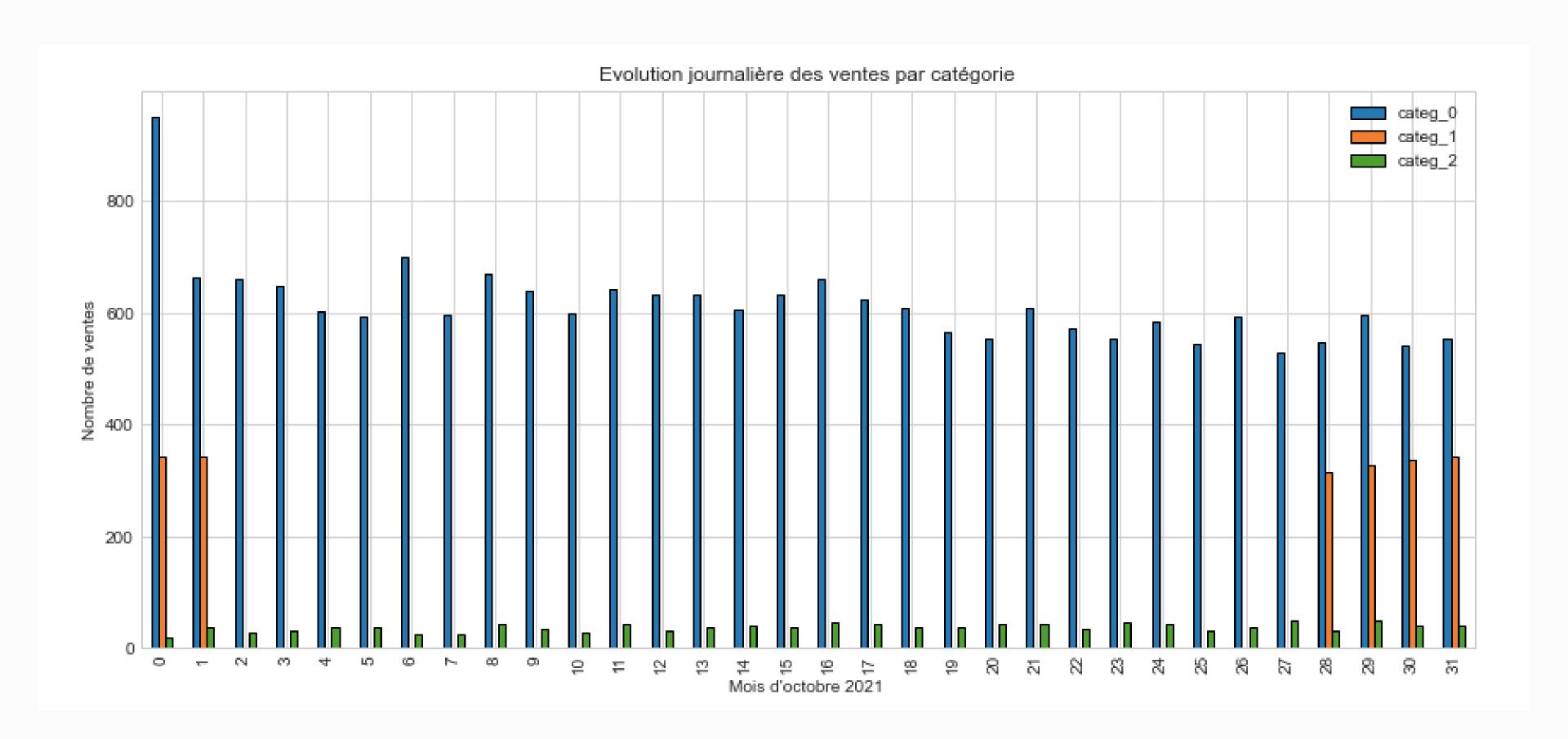
m: 50,6%

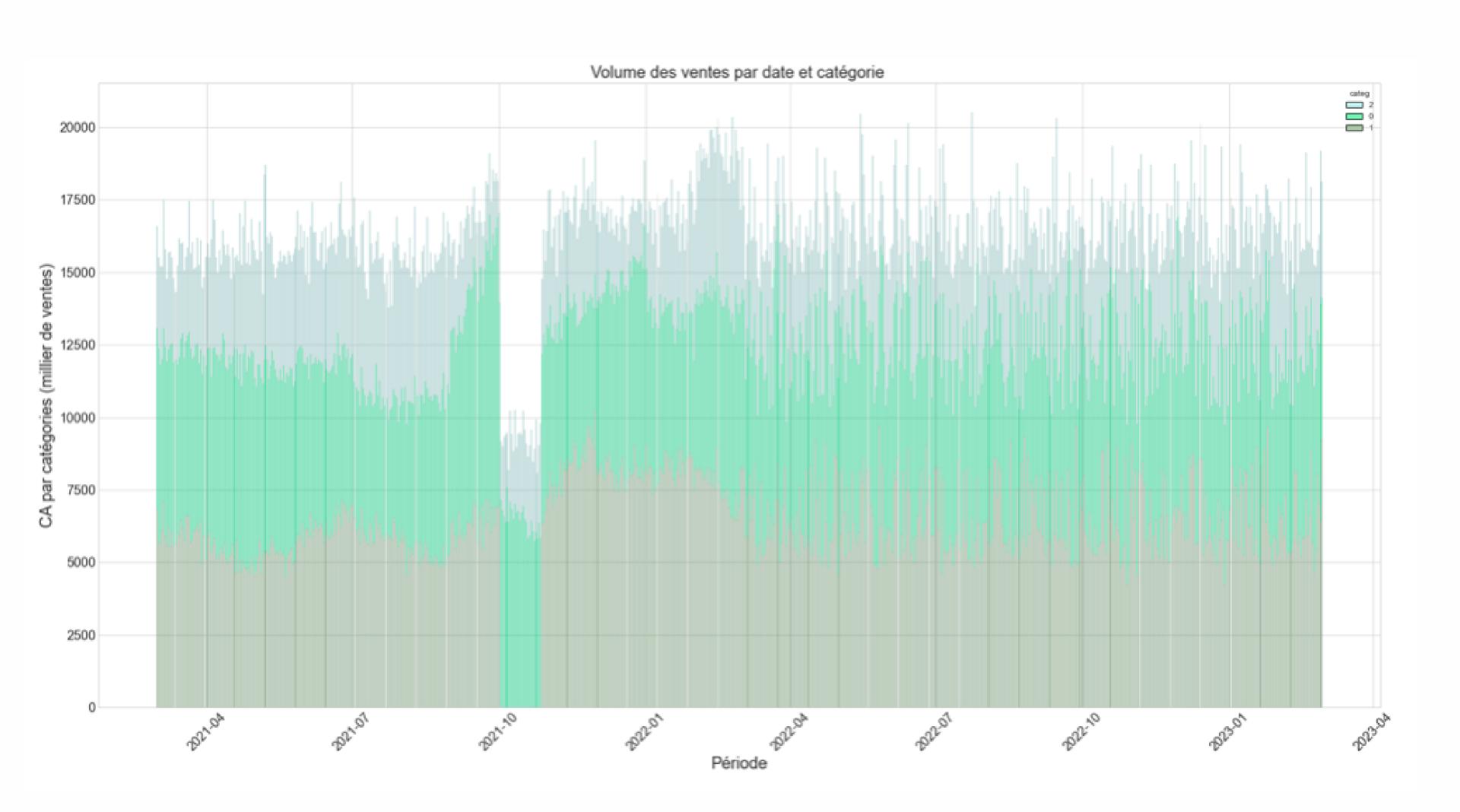
Répartition des genres:

f: 52,1%

h:47,9%





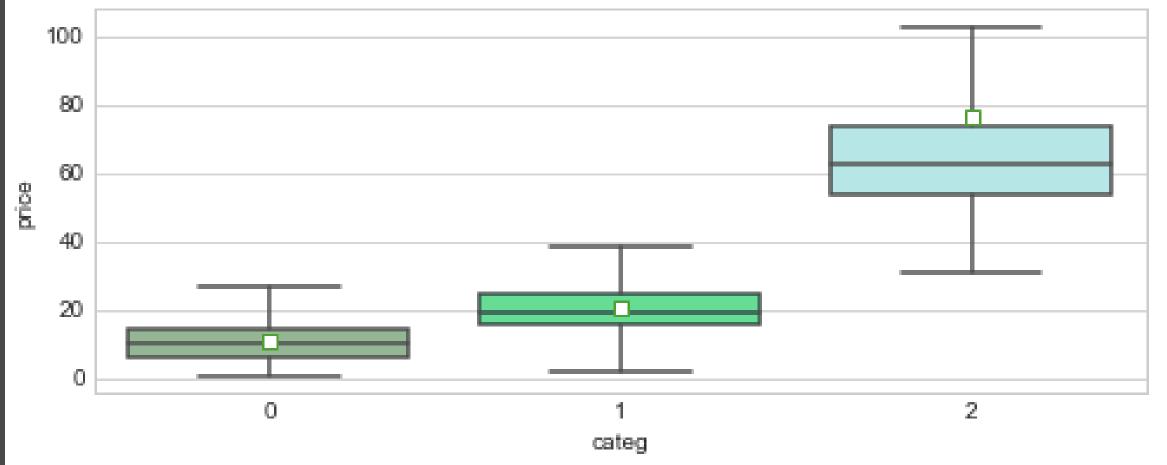




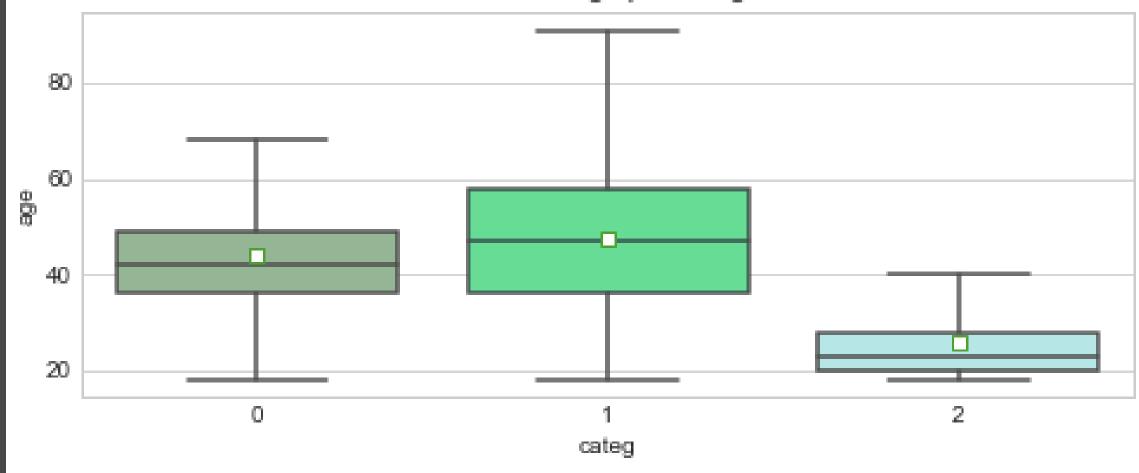
Analyses

- La catégorie 2 est la plus chère
- La catégorie 0 la moins chère
- Les jeunes achètent le plus souvents les articles de la catégorie





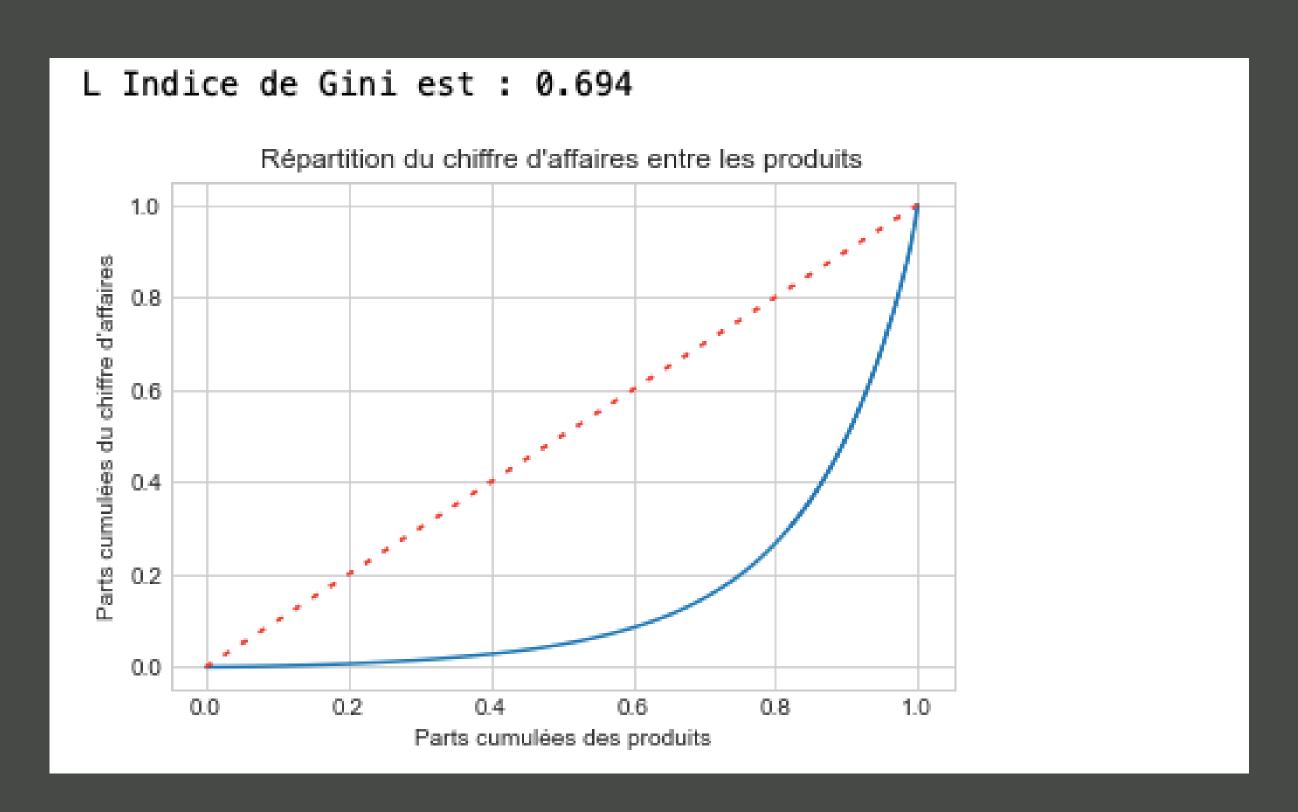
Distribution age par catégorie





Tests de corrélation

L'indice de Gini s'éloigne de 0, en ce sens, on peut dire que le poids des produits dans le chiffre d'affaires global est donc très inégal. Soit, 80% des produits participent à environ 30 % du CA.





Lien entre le sexe et la catégorie : Chi2



Lien entre le panier moyen et l'âge: ANNOVA



Lien entre Fréquence d'achat et l'âge: ANNOVA



Lien entre la catégorie et l'âge: ANNOVA

