

LIBRAIRIE APOSTROPHE

Analysez des ventes

DONNÉES

les ventes

la liste des clients

la liste des produits



nettoyage

analyse descriptive

analyse de corrélation

R/RStudio Cloud

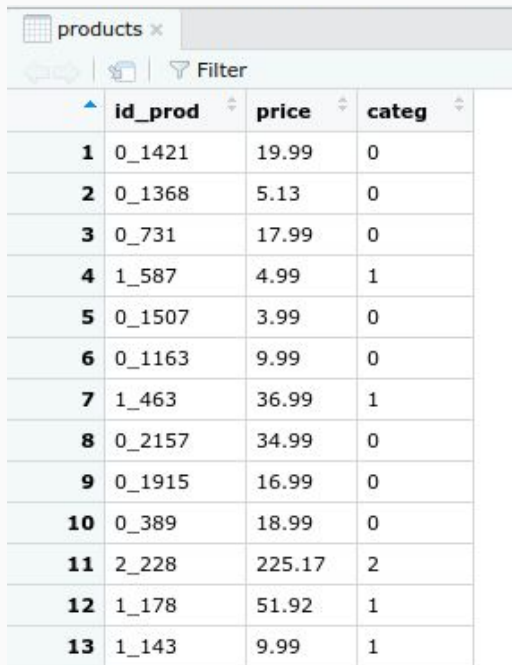
CUSTOMERS

customers x			
Filter			
	client_id	sex	birth
1	c_4410	f	1967
2	c_7839	f	1975
3	c_1699	f	1984
4	c_5961	f	1962
5	c_5320	m	1943
6	c_415	m	1993
7	c_285	f	1967
8	c_160	f	1978
9	c_6446	m	1971

- **client_id**
- **sex**
- **birth**

8623 obs

PRODUCTS



	id_prod	price	categ
1	0_1421	19.99	0
2	0_1368	5.13	0
3	0_731	17.99	0
4	1_587	4.99	1
5	0_1507	3.99	0
6	0_1163	9.99	0
7	1_463	36.99	1
8	0_2157	34.99	0
9	0_1915	16.99	0
10	0_389	18.99	0
11	2_228	225.17	2
12	1_178	51.92	1
13	1_143	9.99	1

- id_prod
- price
- categ

3287 obs

TRANSACTIONS

transactions x				
Filter				
	id_prod	date	session_id	client_id
1	0_1483	2021-04-10 18:37:28	s_18746	c_4450
2	2_226	2022-02-03 01:55:53	s_159142	c_277
3	1_374	2021-09-23 15:13:46	s_94290	c_4270
4	0_2186	2021-10-17 03:27:18	s_105936	c_4597
5	0_1351	2021-07-17 20:34:25	s_63642	c_1242
6	0_1085	2021-09-15 05:47:48	s_90139	c_2526
7	0_1508	2021-08-29 05:39:01	s_82100	c_5799
8	0_1627	2021-09-19 14:54:52	s_92294	c_1422
9	0_1460	2022-02-03 07:13:22	s_159252	c_2207

- id_prod
- date
- session_id
- client_id

337 016 obs

NETTOYAGE : CUSTOMERS

- Vérification valeurs manquantes
- Suppression clients tests (ct_ & c_)

NETTOYAGE : CUSTOMERS

Var1	Freq
c_1 : 1	Min. : 1.00
c_10 : 1	1st Qu.: 13.00
c_100 : 1	Median : 24.00
c_1000 : 1	Mean : 36.66
c_1001 : 1	3rd Qu.: 44.00
c_1002 : 1	Max. : 11839.00
(Other): 8592	

	Var1	Freq
678	c_1609	11839
6337	c_6714	4189
2724	c_3454	3139
4388	c_4958	2457
7791	c_8026	184
1773	c_2595	179
7120	c_7421	178

CA : 410 085.4

7.48%

Action -> Suppression clients pro des transactions

NETTOYAGE : PRODUCTS

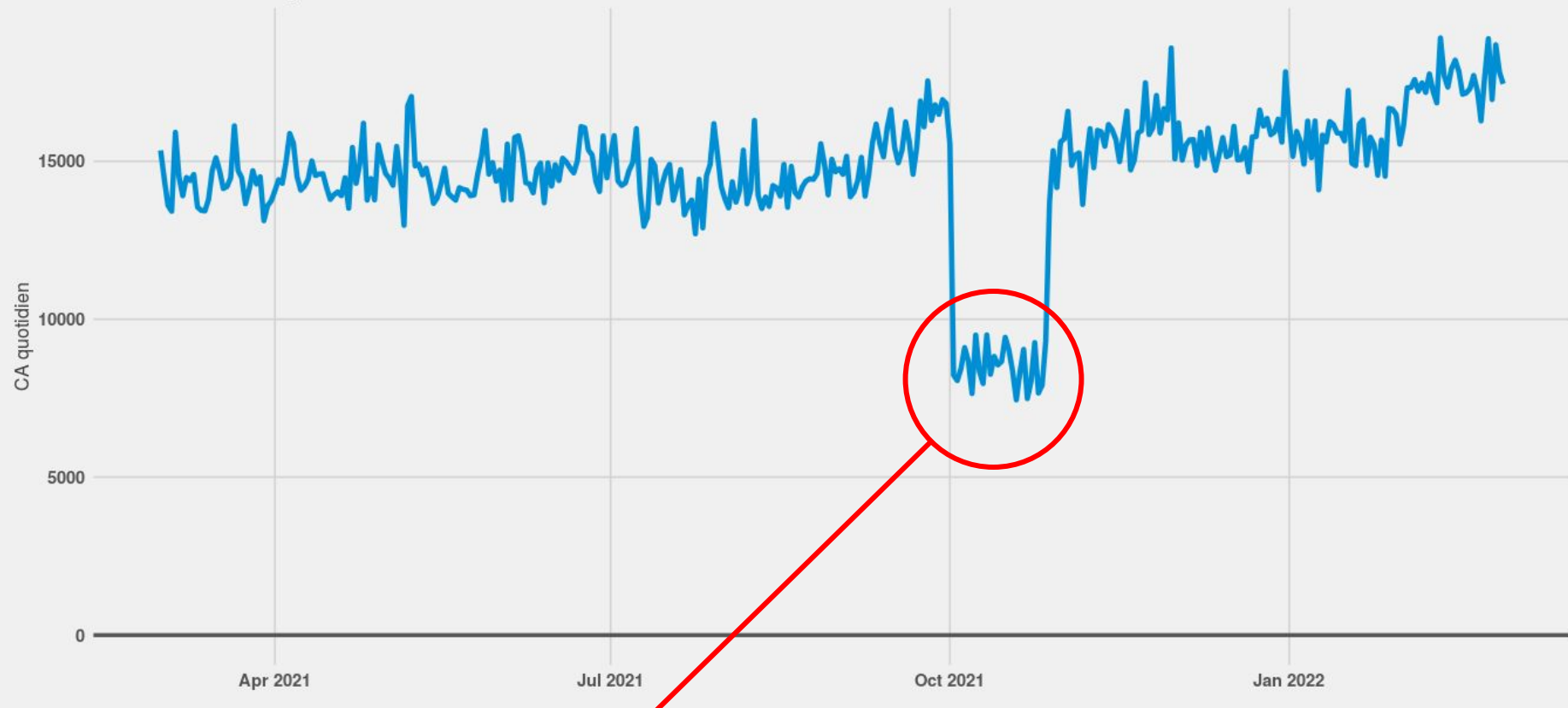
- Vérification valeurs manquantes
- Vérification formats
- Suppression produit test (T_0)

NETTOYAGE : TRANSACTIONS

- Vérification valeurs manquantes
- Vérification formats
- Suppression clients tests (ct_ & c_)
- Suppression produit test (T_0)
- Vérification listes produits et clients
- Imputation par moyenne catégorie 0 pour id_prod = 0_2245

CHIFFRE D'AFFAIRE

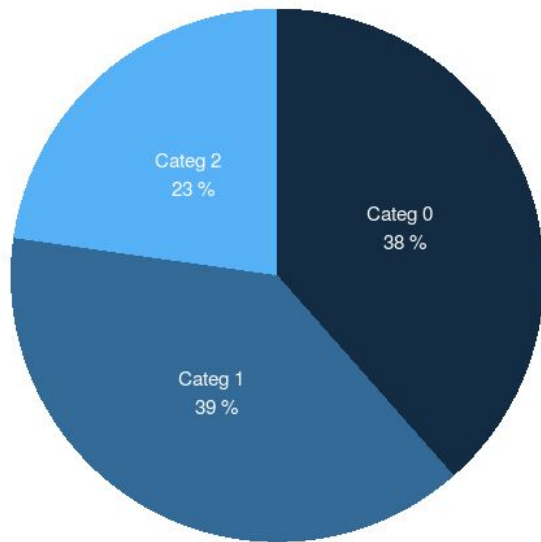
Evolution CA quotidien sur l'année



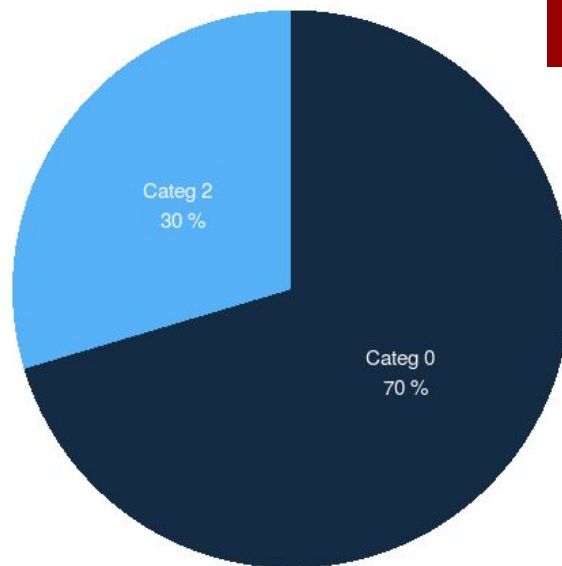
CA octobre anormal

OCTOBRE

Répartition CA selon catégorie sur année



Répartition CA selon catégorie sur octobre

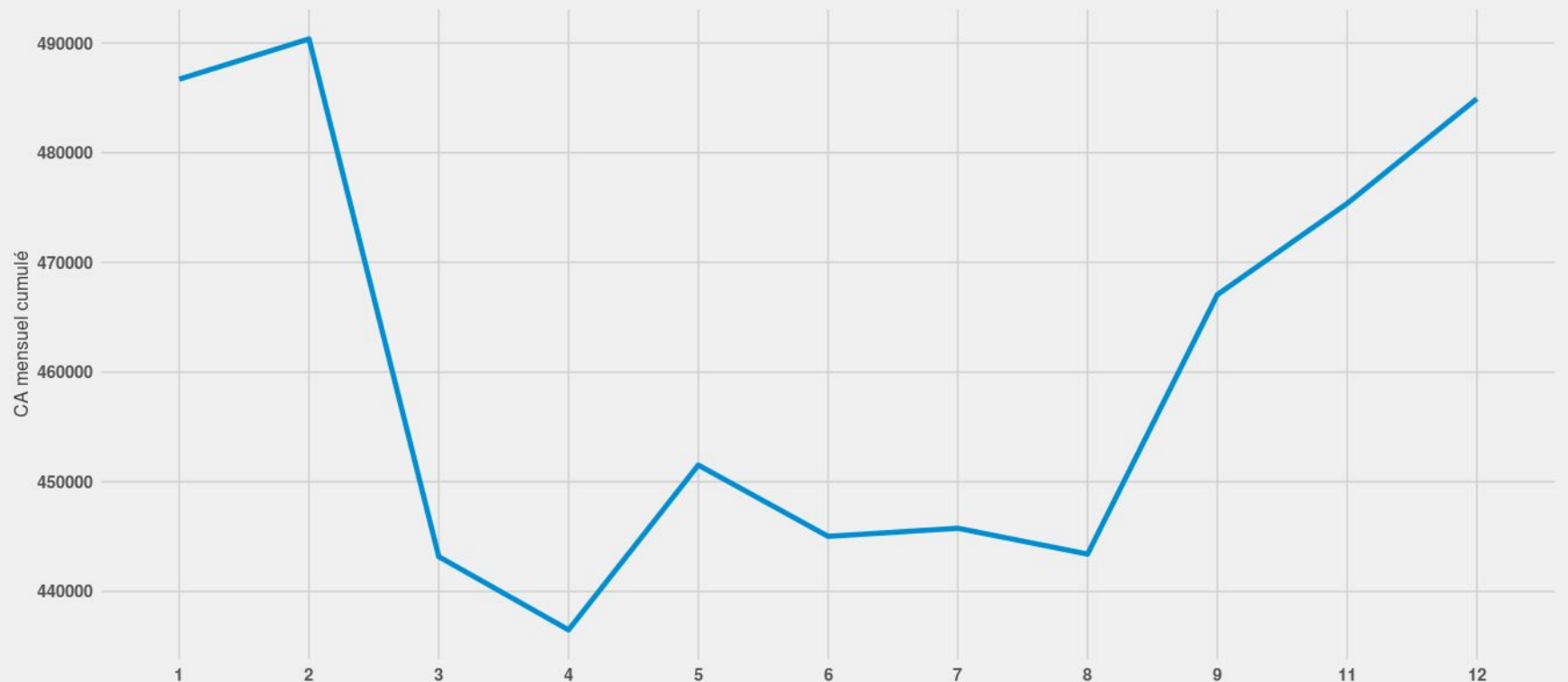


Categ 1 = 0 !

OCTOBRE

- Catégorie 1 = 739 références
- Suppression des transactions du mois d'octobre

Evolution CA mensuel sur l'année

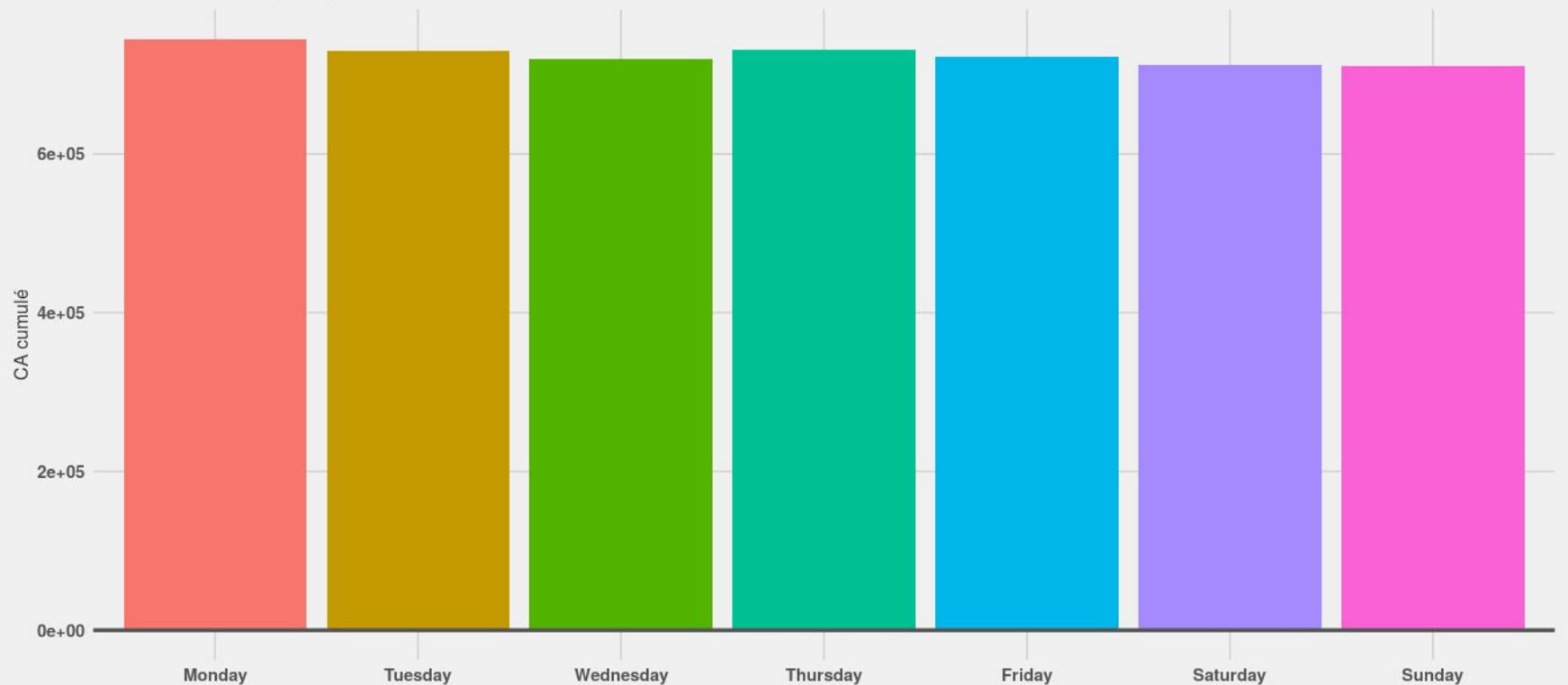


460 889.1

CA mensuel moyen

Saisonnalité : sept - fév

Evolution CA par jour de la semaine

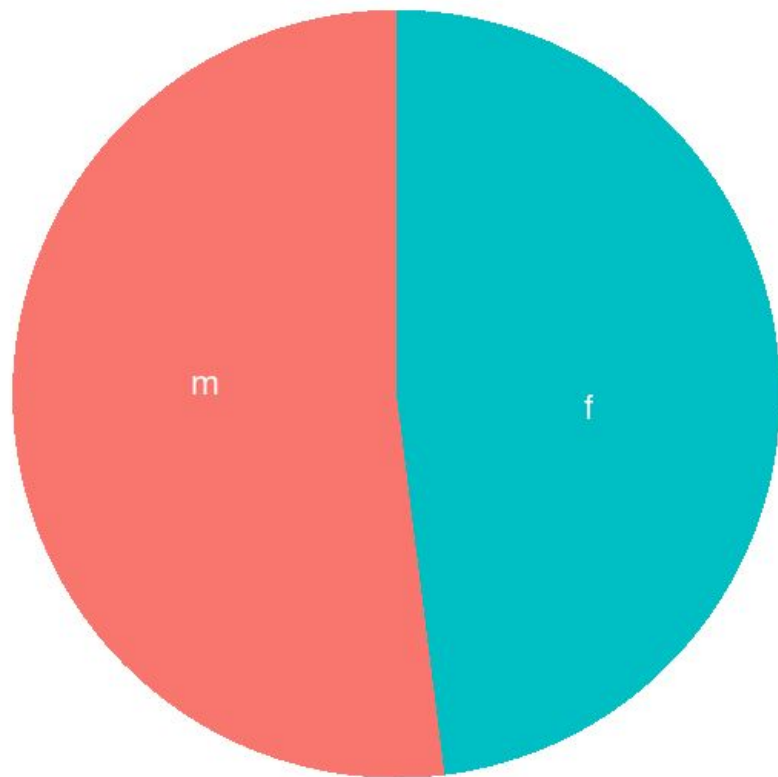


78,2838

CA journalier moyen

CA journalier équilibré

Répartition CA selon Sexe



Panier médian

25.99

Panier moyen

34.48

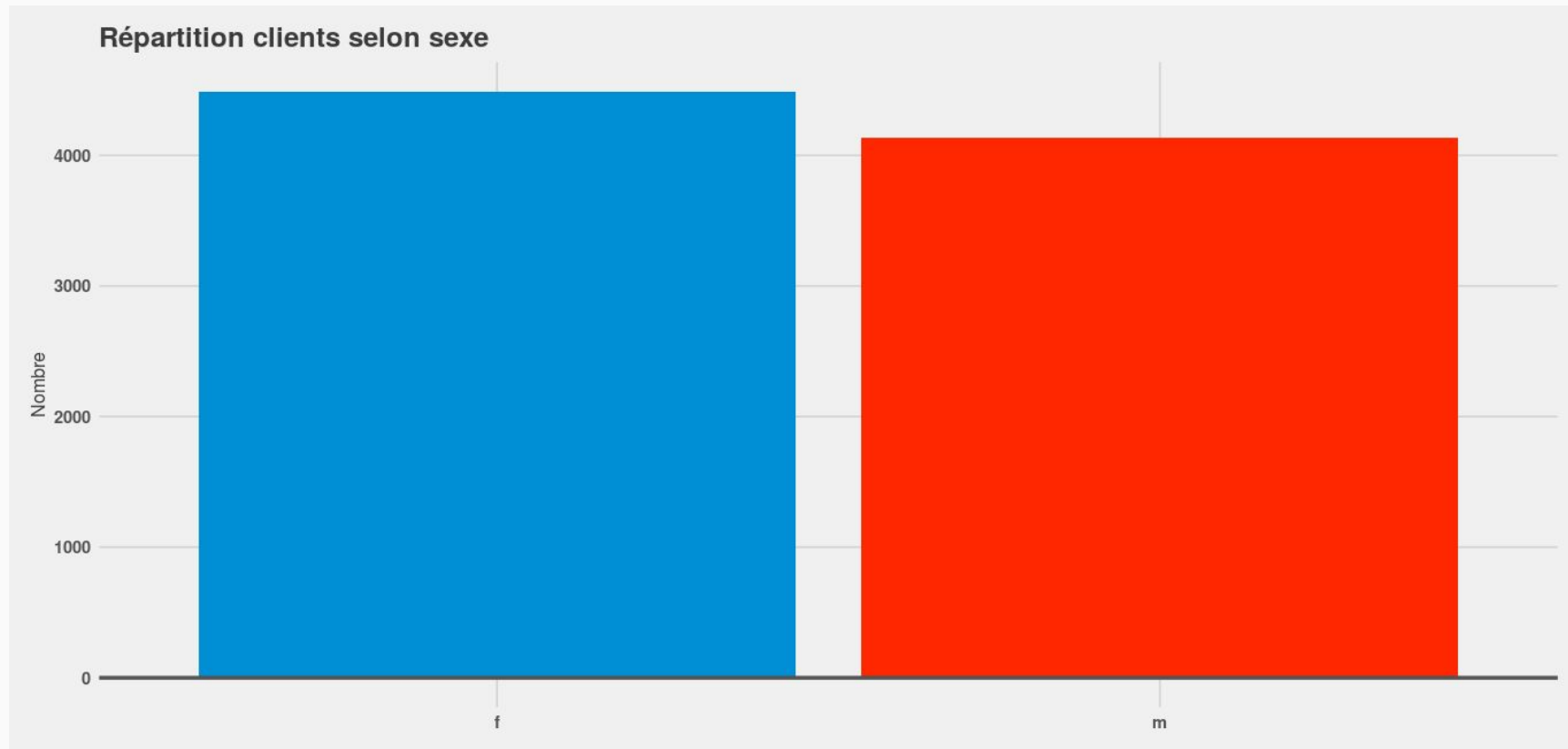
CA par sexe équilibré

CLIENTS

8621

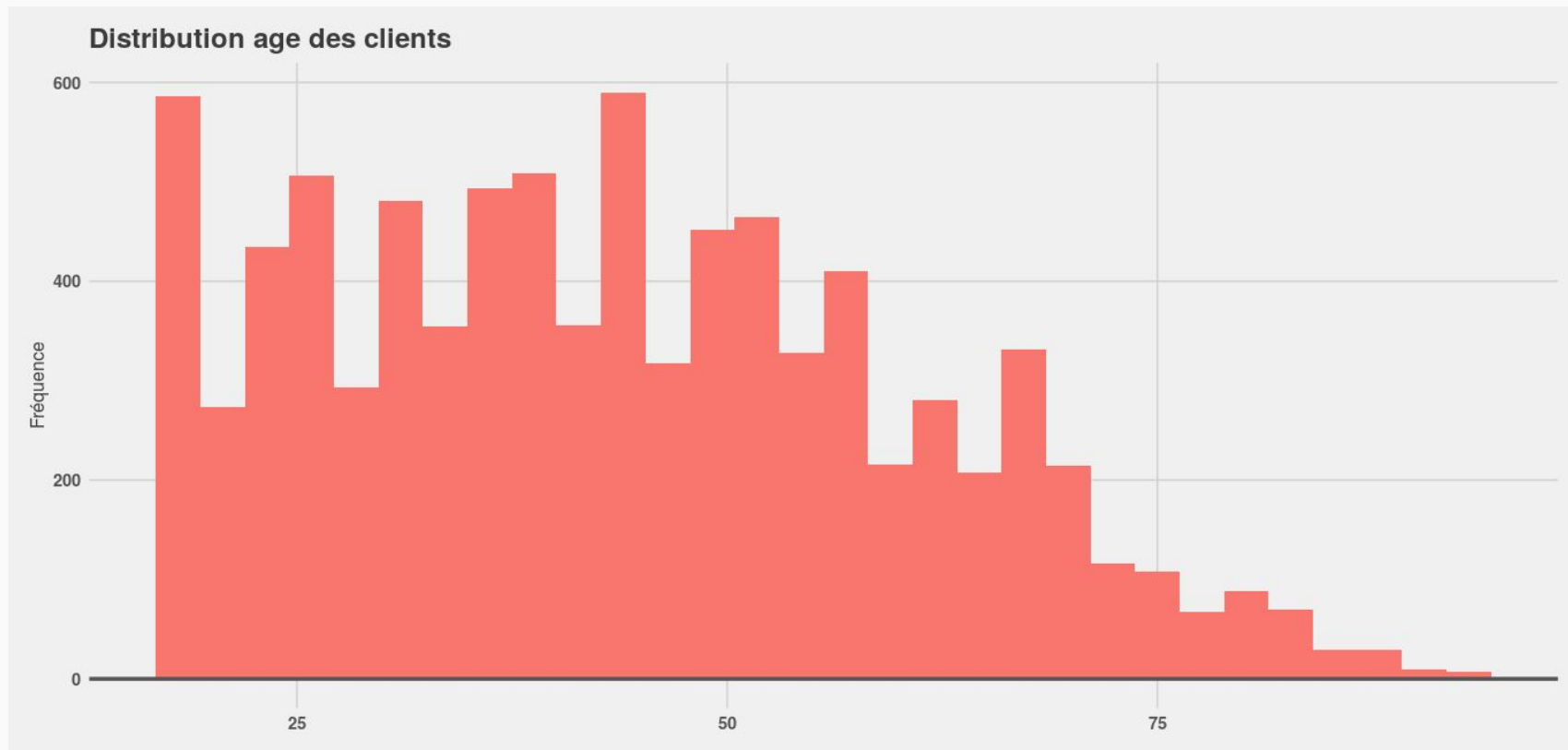
Clients

Répartition équilibré sexe clients



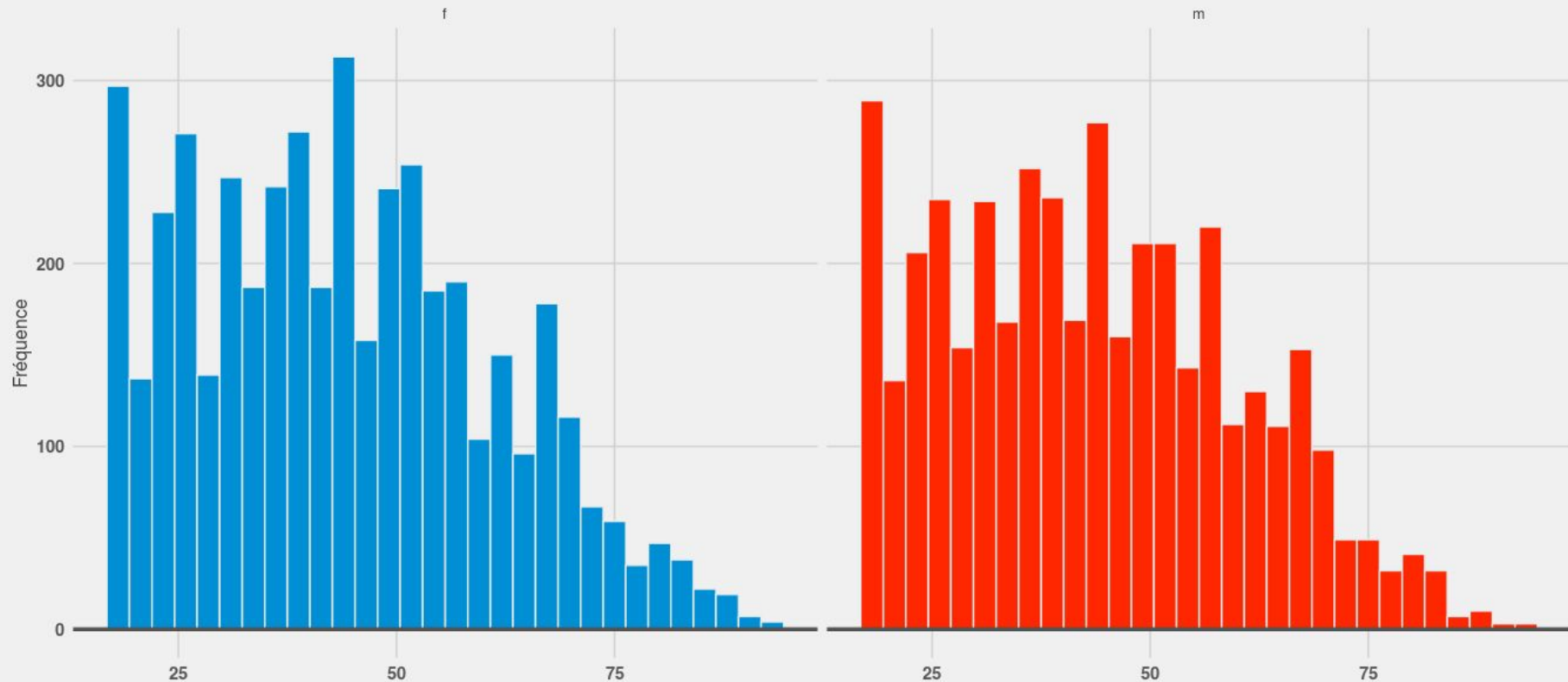
43 ans

Age moyen



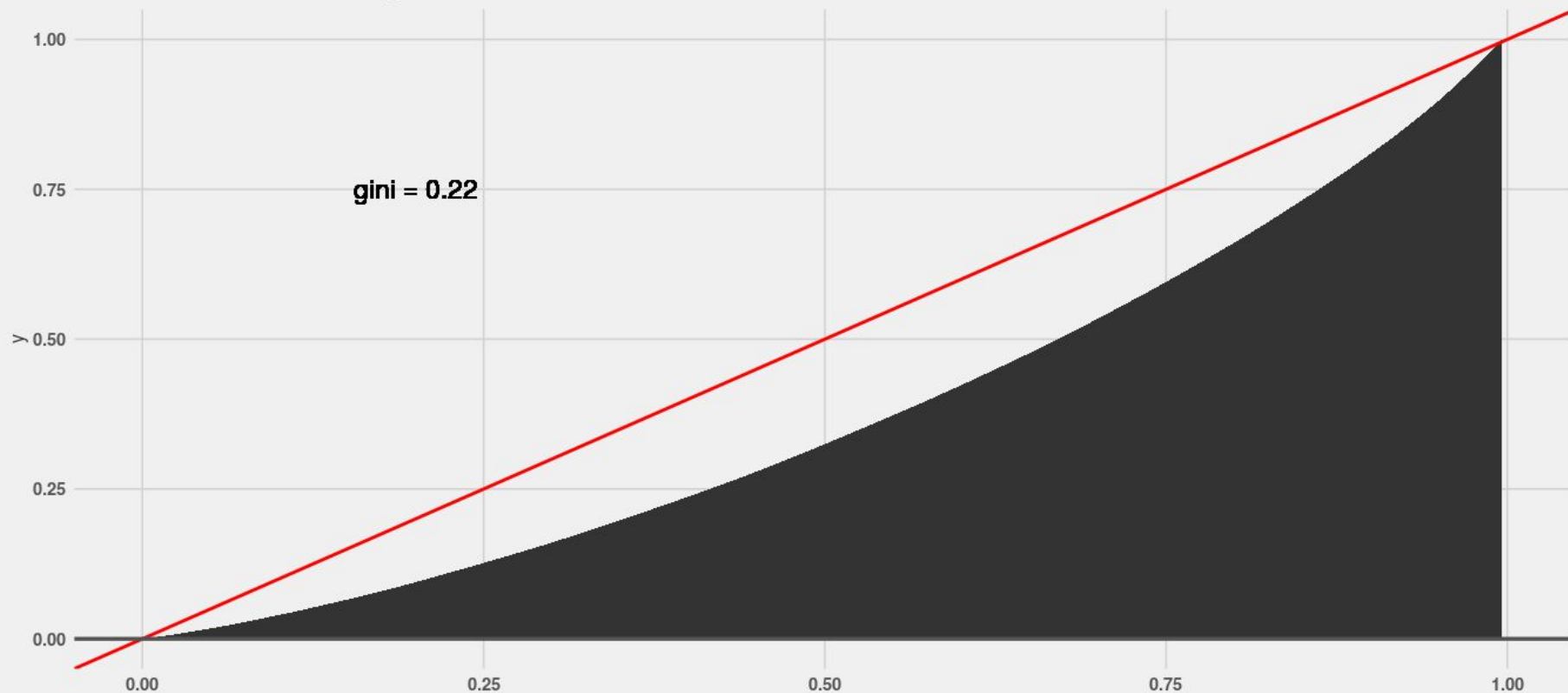
Distribution asymétrique et multimodale

Distribution age des clients selon de sexe



Distribution âge similaire

Lorenz concentration age clients



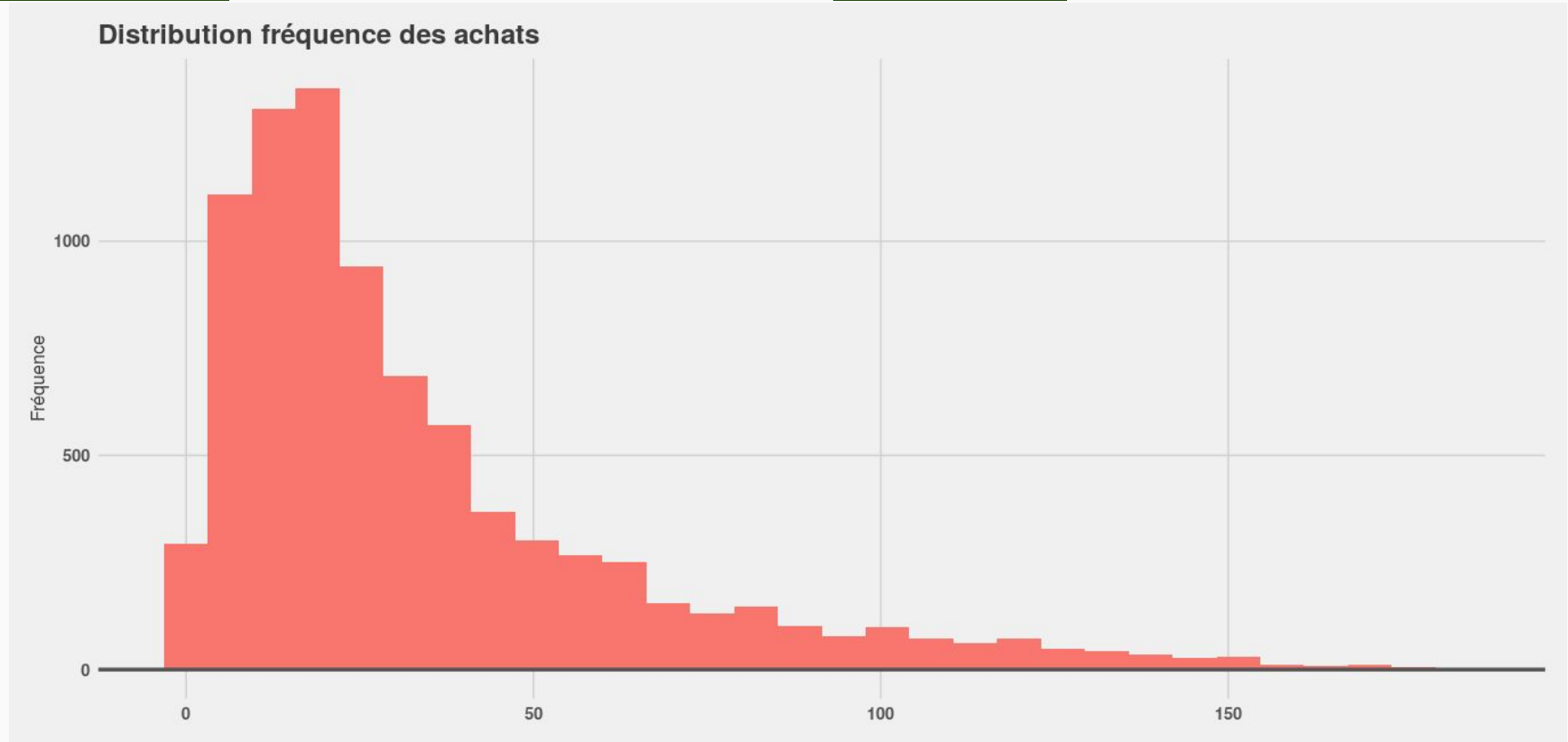
Faible concentration âges clients

24

Nombre d'achat médian par année

0.5%

D'acheteurs unique



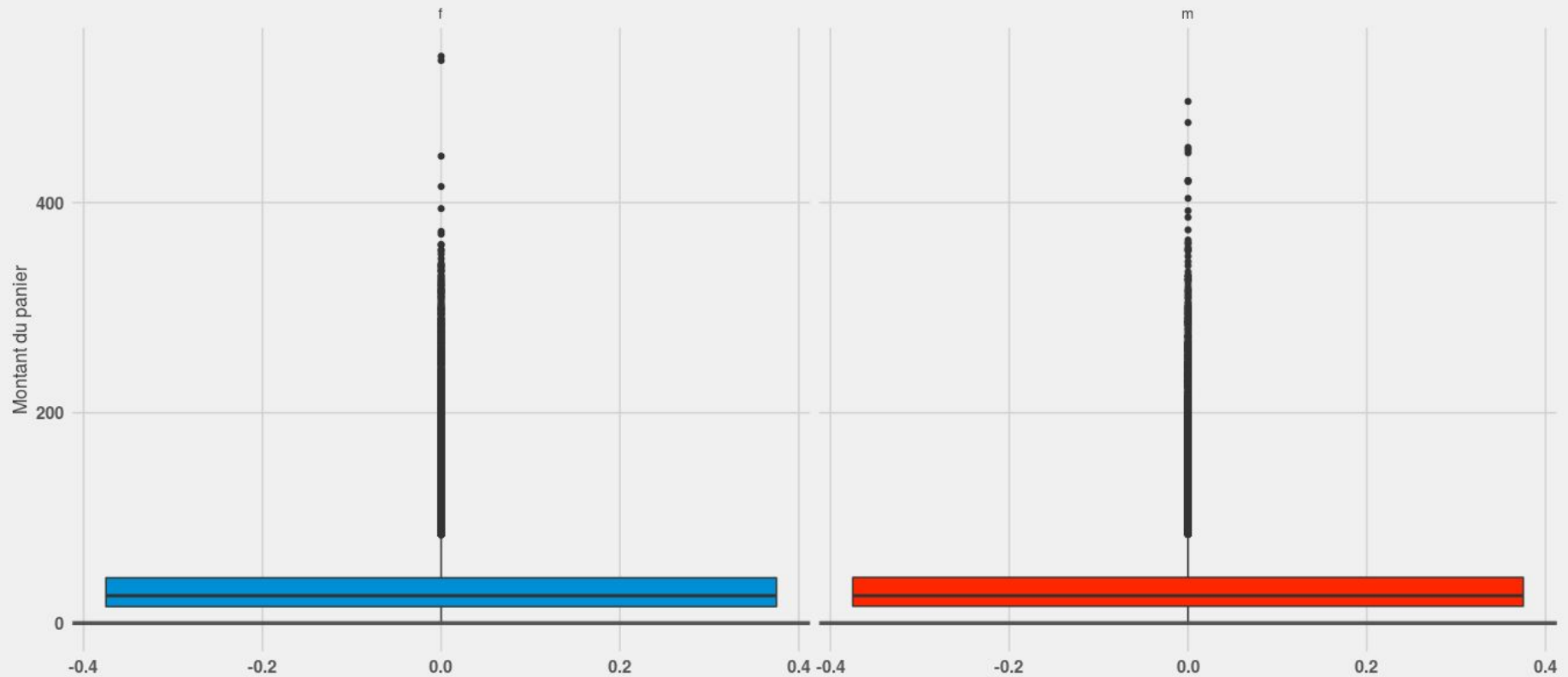
Distribution asymétrique

f



Distribution fréquence achat similaire

Montant du panier moyen selon de sexe



Montant panier moyen similaire

PRODUITS

3287

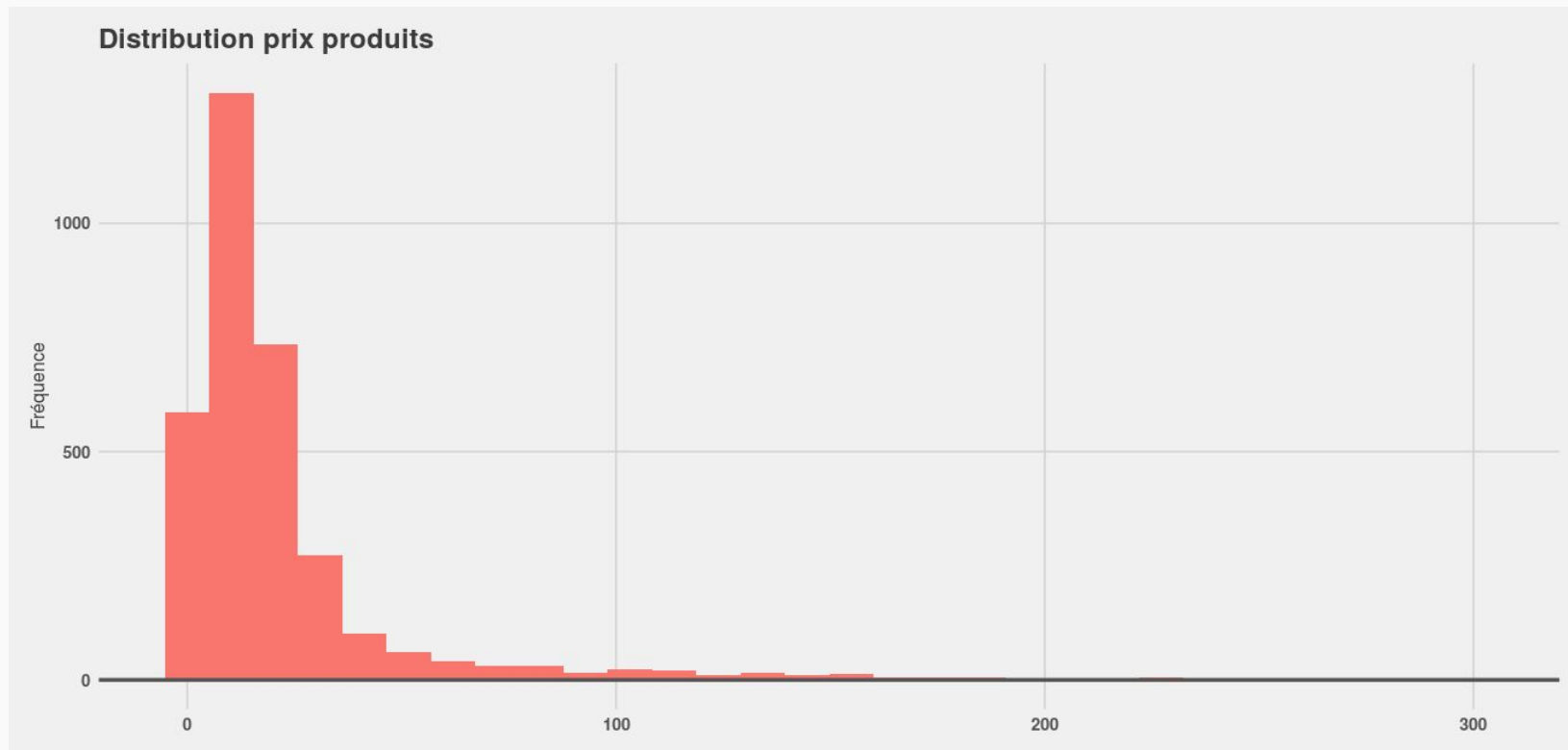
produits

13.09

prix médian

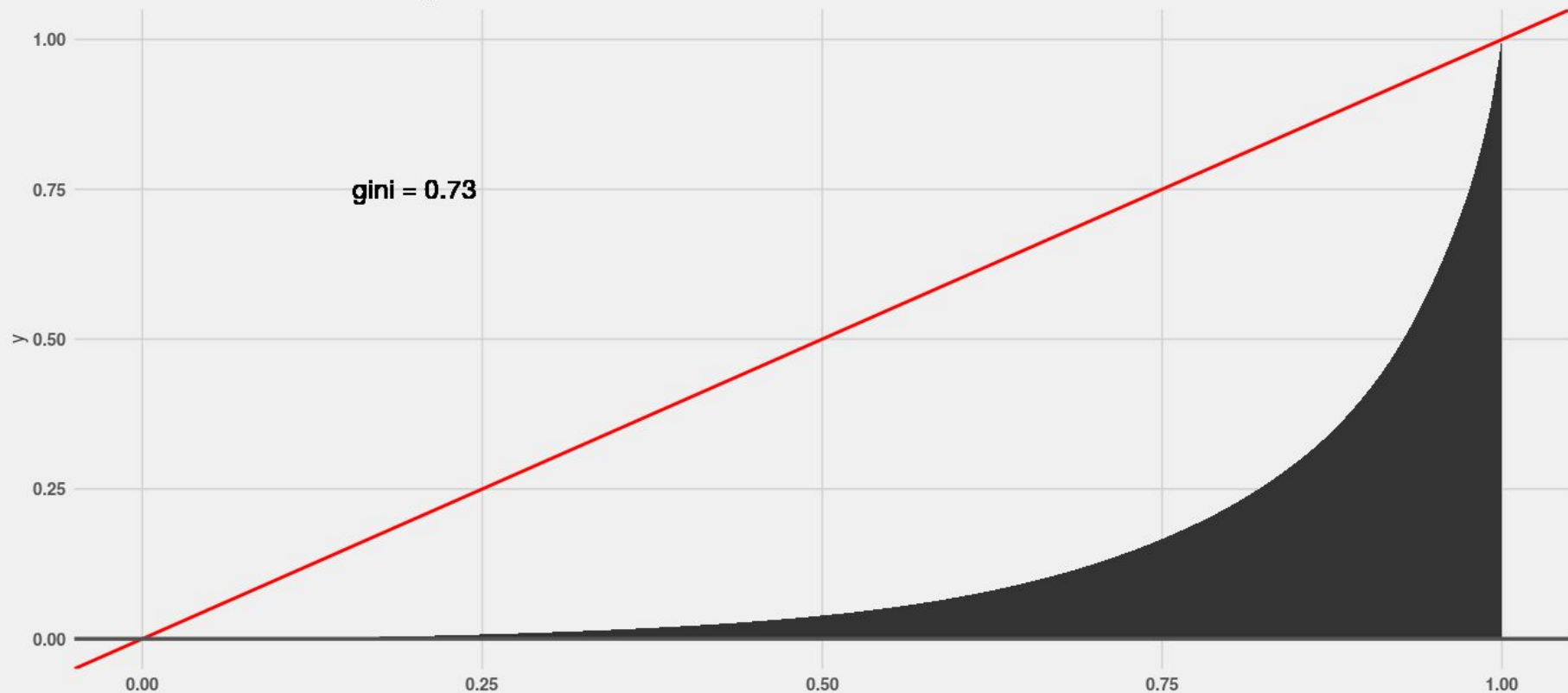
21.86

prix moyen



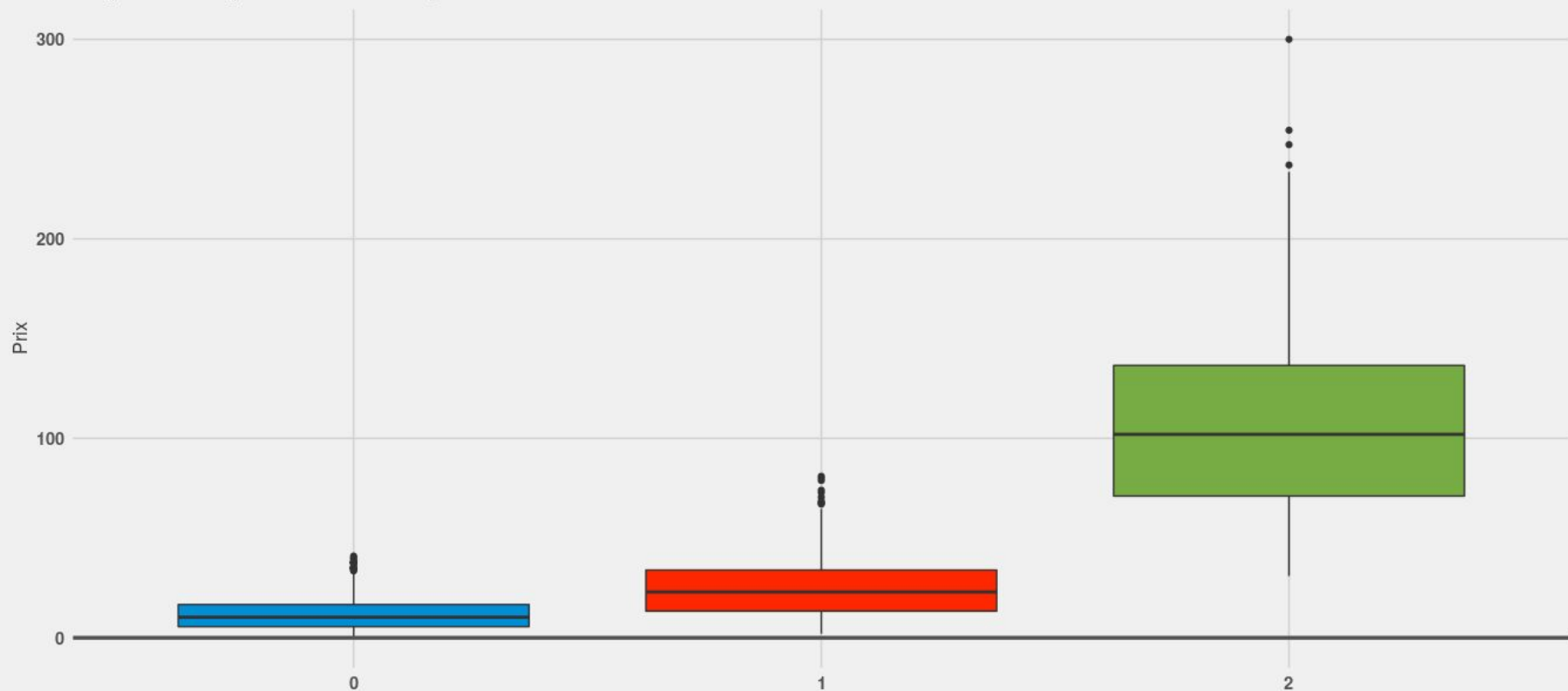
Distribution asymétrique

Lorenz concentration des prix



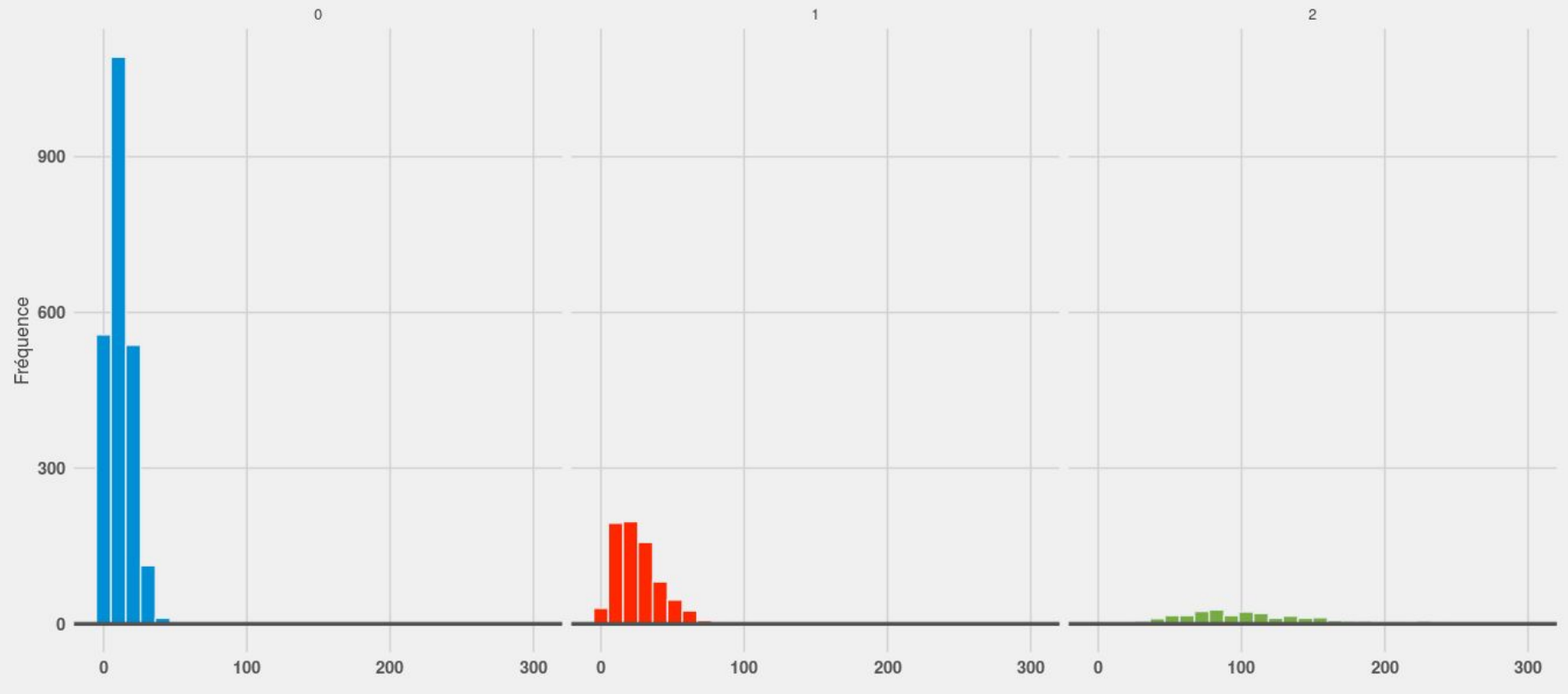
Importante concentration des prix produits

Répartition prix selon catégorie



Importante différence de prix selon catégories produits

Distribution prix produits



Distribution prix par catégorie semble inégale

Categ 0

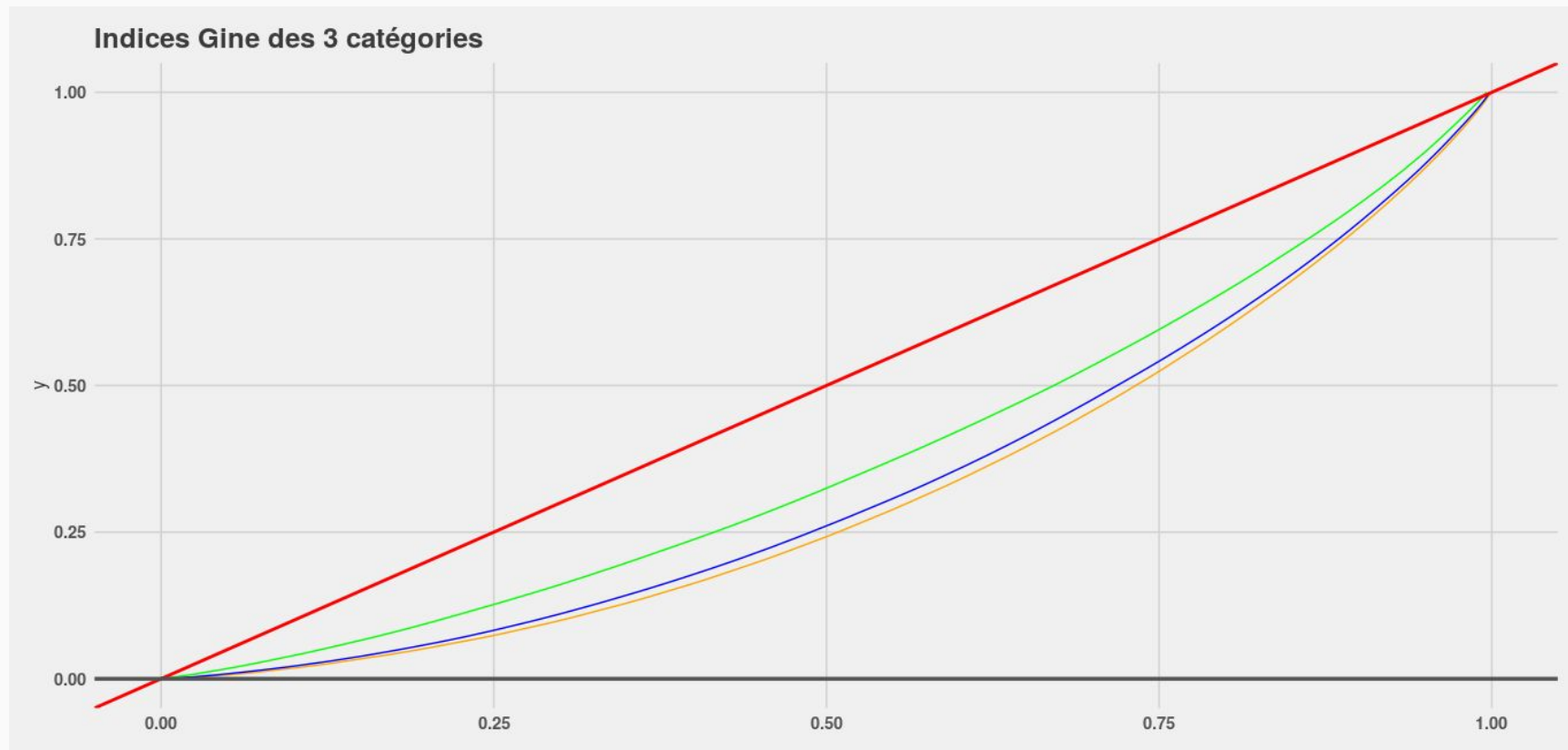
0.35

Categ 1

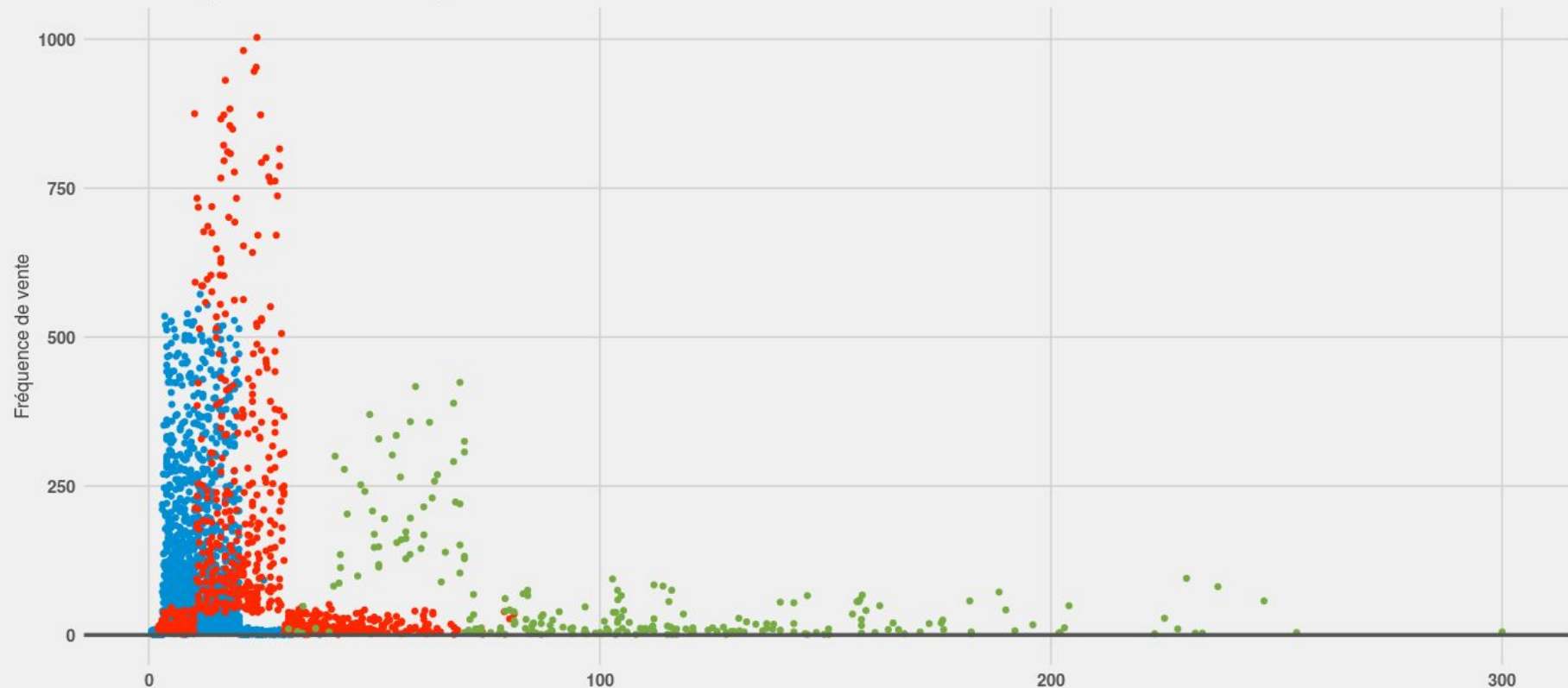
0.33

Categ 2

0.25

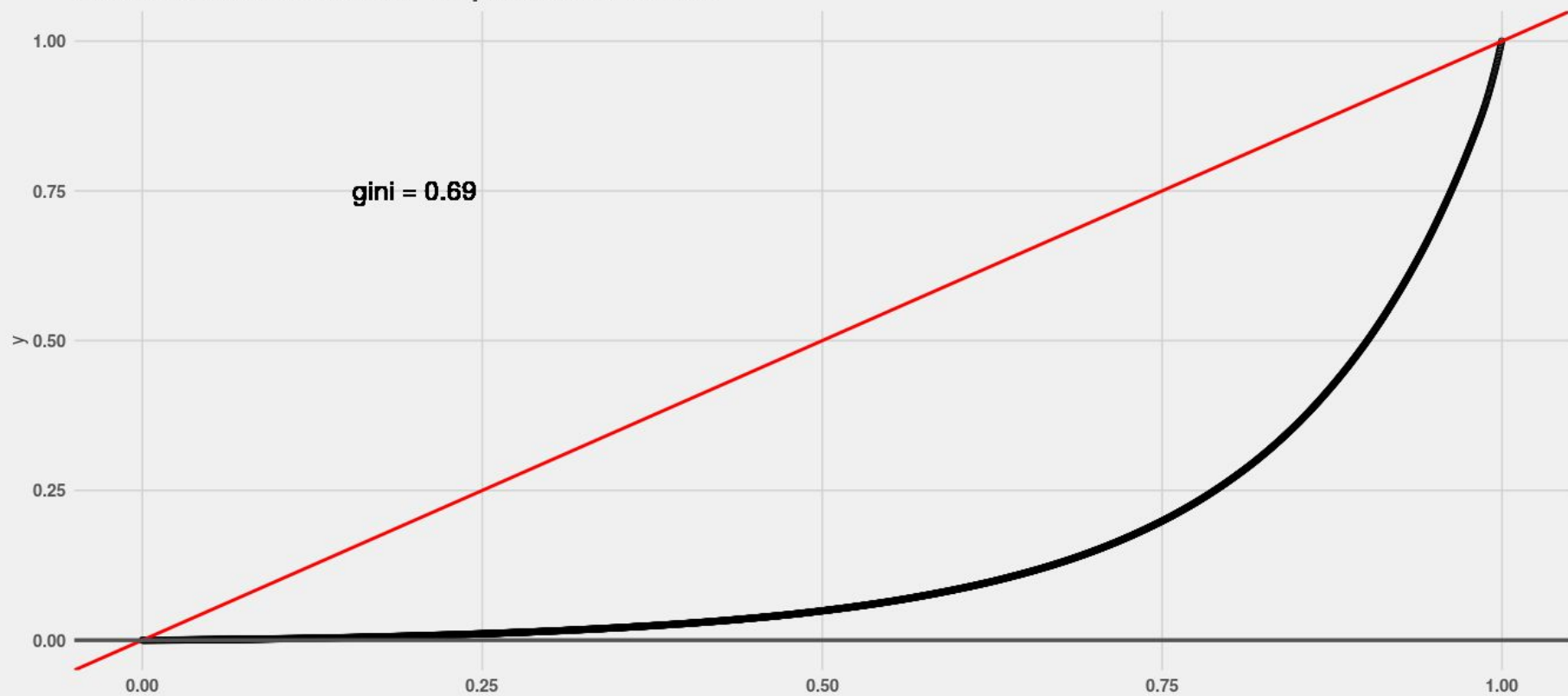


Relation prix ventes et fréquence achat



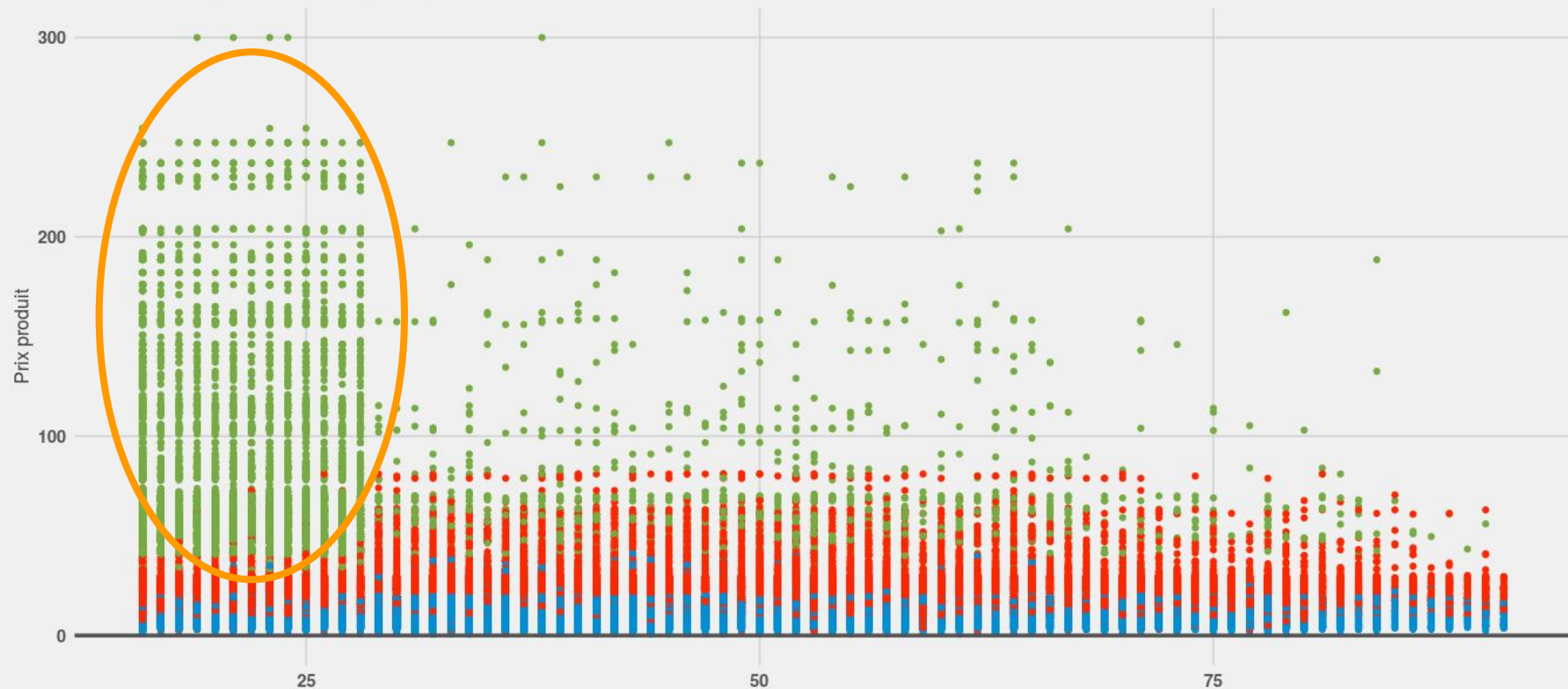
Fréquence ventes / prix semble inégale

Lorenz concentration des fréquences de ventes



Importante concentration des ventes

Relation age client et prix produit



Age / prix semble inégale

	Freq
Min. :	1.00
1st Qu. :	7.00
Median :	25.00
Mean :	90.15
3rd Qu. :	104.00
Max. :	1003.00



TOP 15

Var1	Freq
1_369	1003
1_417	981
1_414	953
1_498	946
1_425	931
1_413	883
1_398	875
1_406	873
1_412	873
1_407	866
1_403	855
1_396	849
1_400	822
1_395	816
1_376	811
1_392	808

- 25% des articles nb commandes ≤ 7
- Soit 800 produits

CLIENT TYPE

Sexe

H/F

Âge

30 - 56 ans

CA annuel

260.95 - 822.36

Fréquence achat

13 - 44 produits / an

Panier moyen

15.81 - 43.28



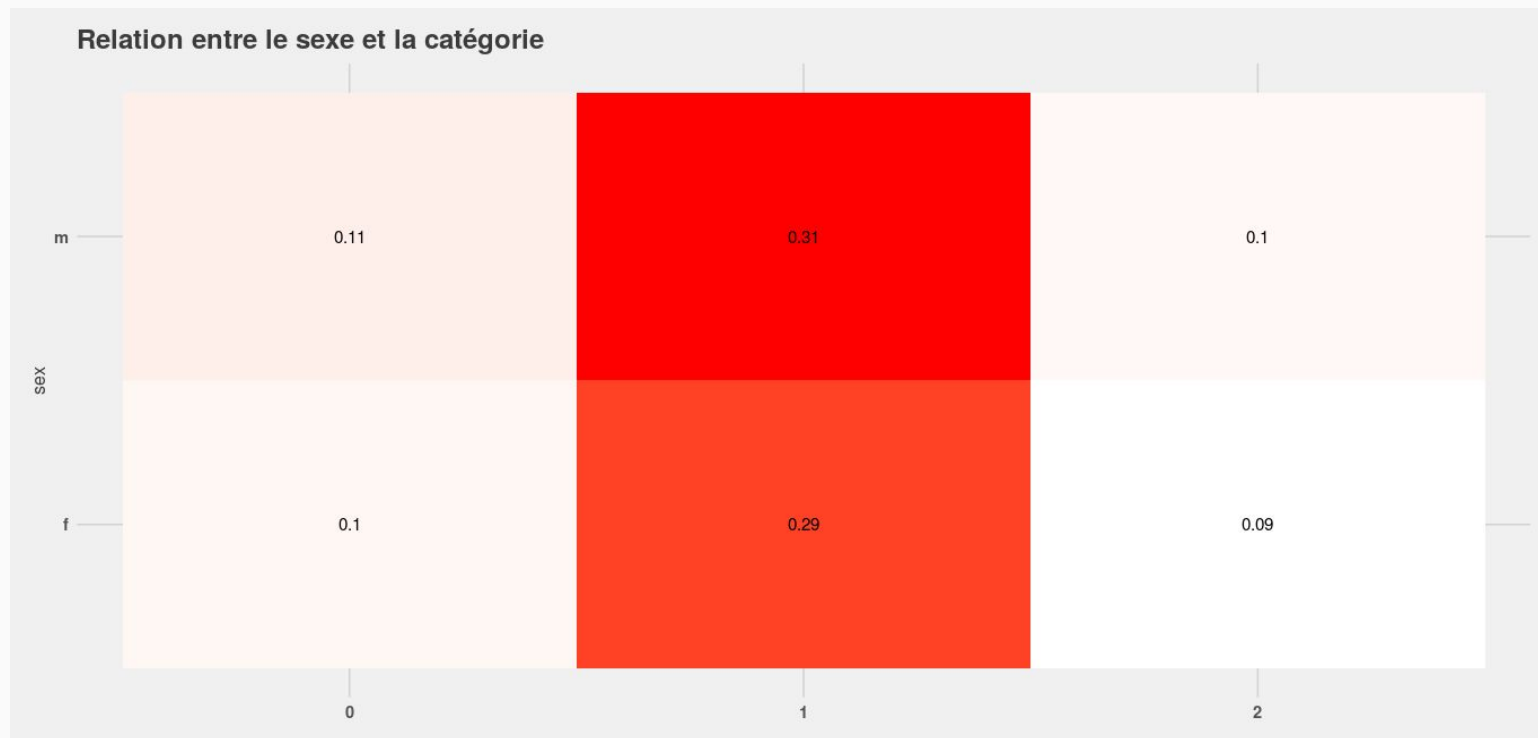
**Indices selon sexe &
catégories**

Relations sexe/âge client et achats

Questions

- **Y a-t-il une corrélation entre le sexe des clients et les catégories de produits achetés ?**
- **Y a-t-il une corrélation entre l'âge des clients et le montant total des achats ?**
- **Y a-t-il une corrélation entre l'âge des clients et la fréquence d'achat ?**
- **Y a-t-il une corrélation entre l'âge des clients et la taille du panier moyen ?**
- **Y a-t-il une corrélation entre l'âge des clients et la les catégories de produits achetés ?**

Y a-t-il une corrélation entre le sexe des clients et les catégories de produits achetés ?



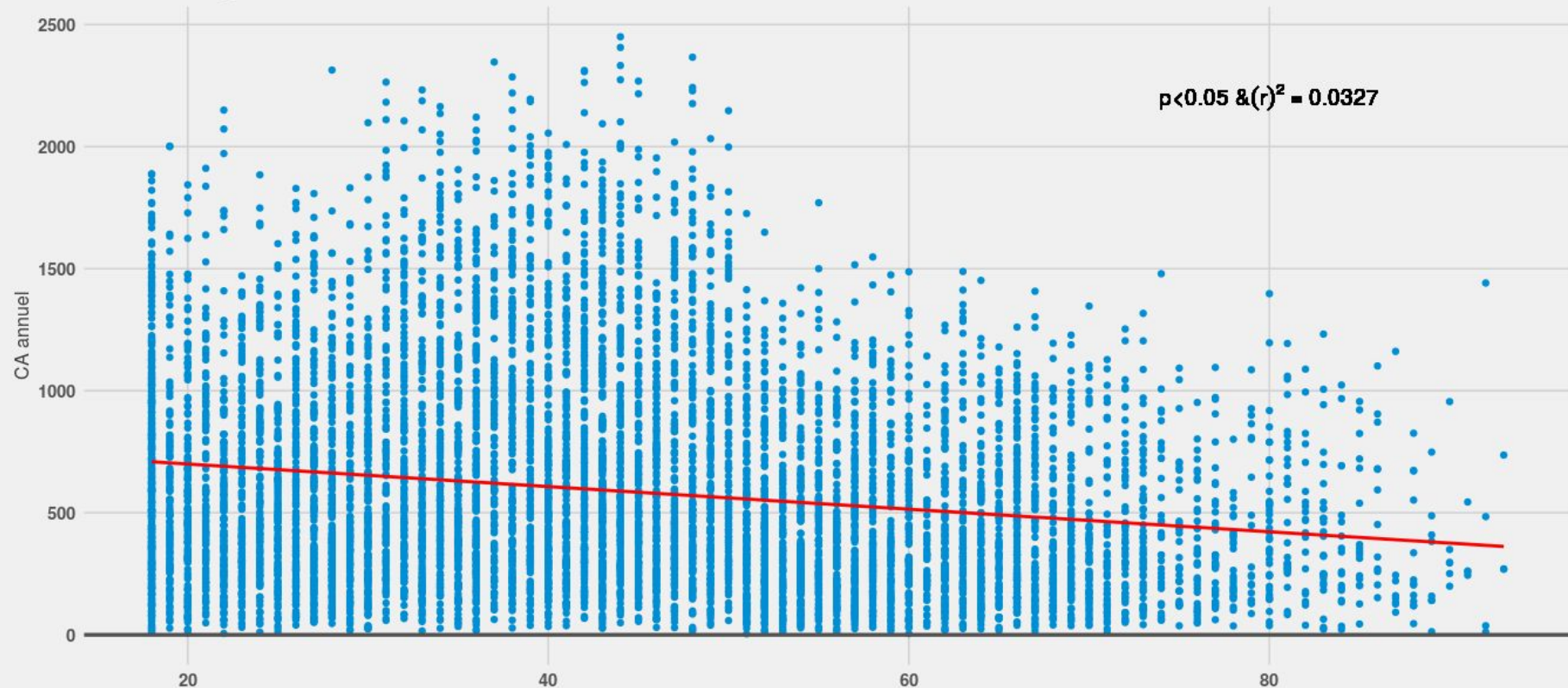
Relation significative et faible

Quantité achats plus importante chez les hommes

X-squared = 79.18
P < 0.05
 $\eta^2 = 5.906479e-06$

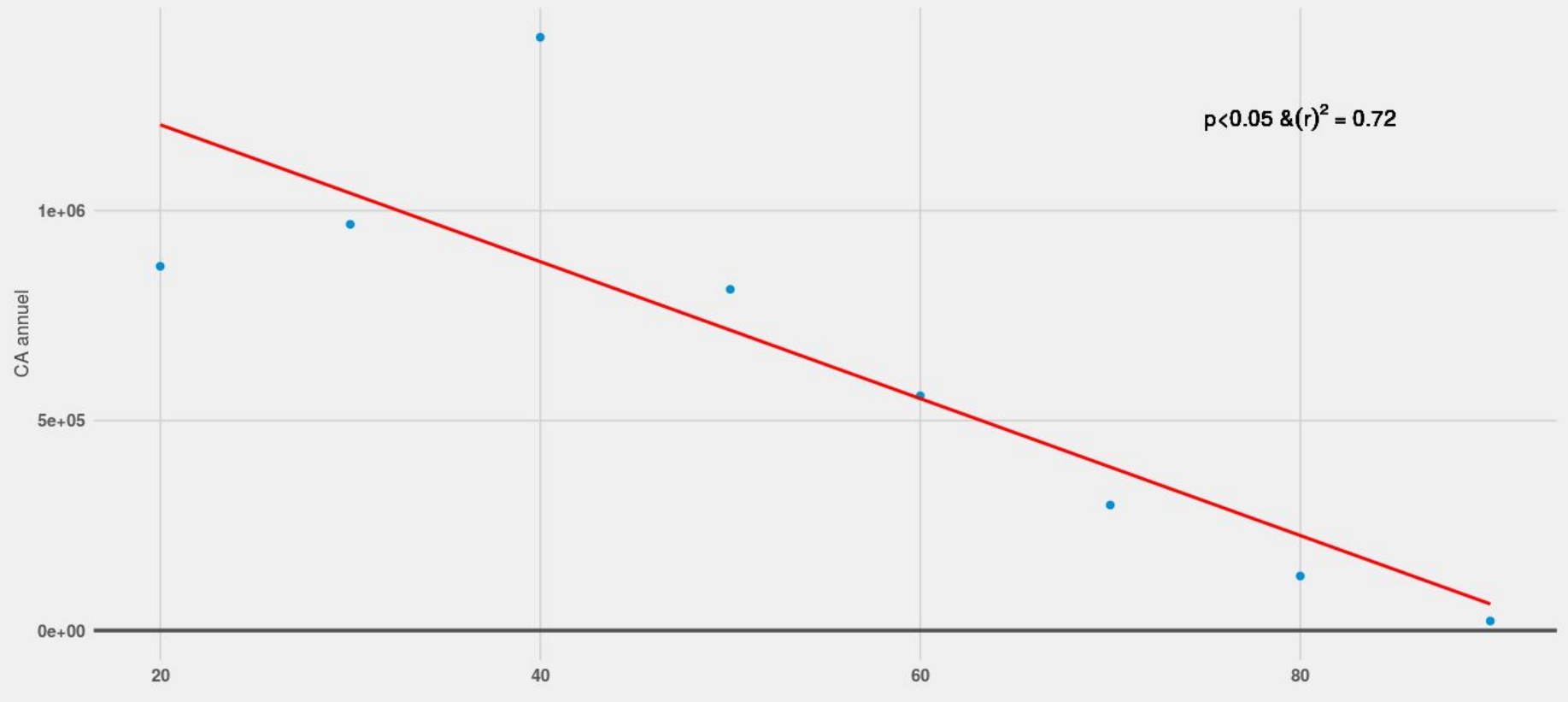
**Y a-t-il une corrélation entre l'âge
des clients et le montant total
des achats ?**

Relation age client et CA annuel



Relation significative et faible

Relation décennie client et CA annuel

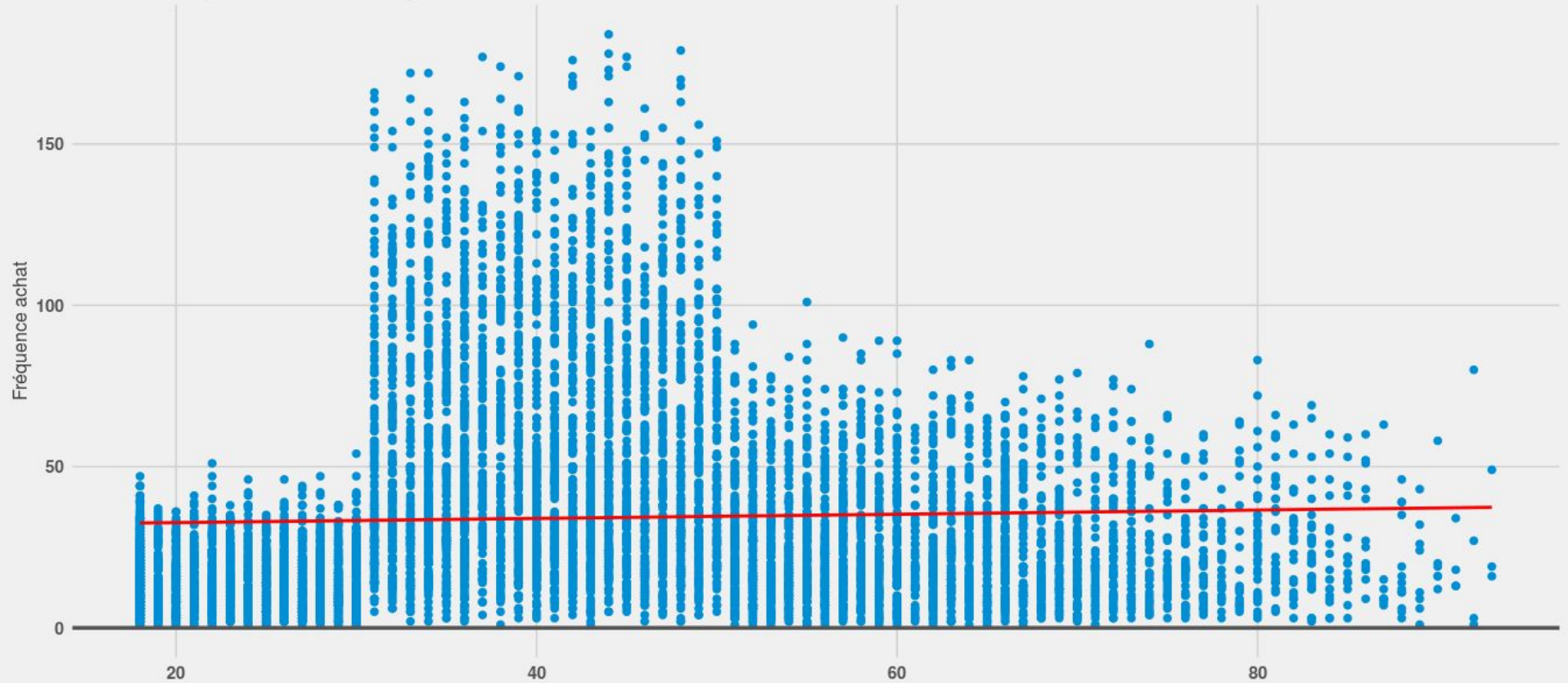


Relation significative et forte

Groupe < 30 ans & >50 ?

Y a-t-il une corrélation entre l'âge des clients et la fréquence d'achat ?

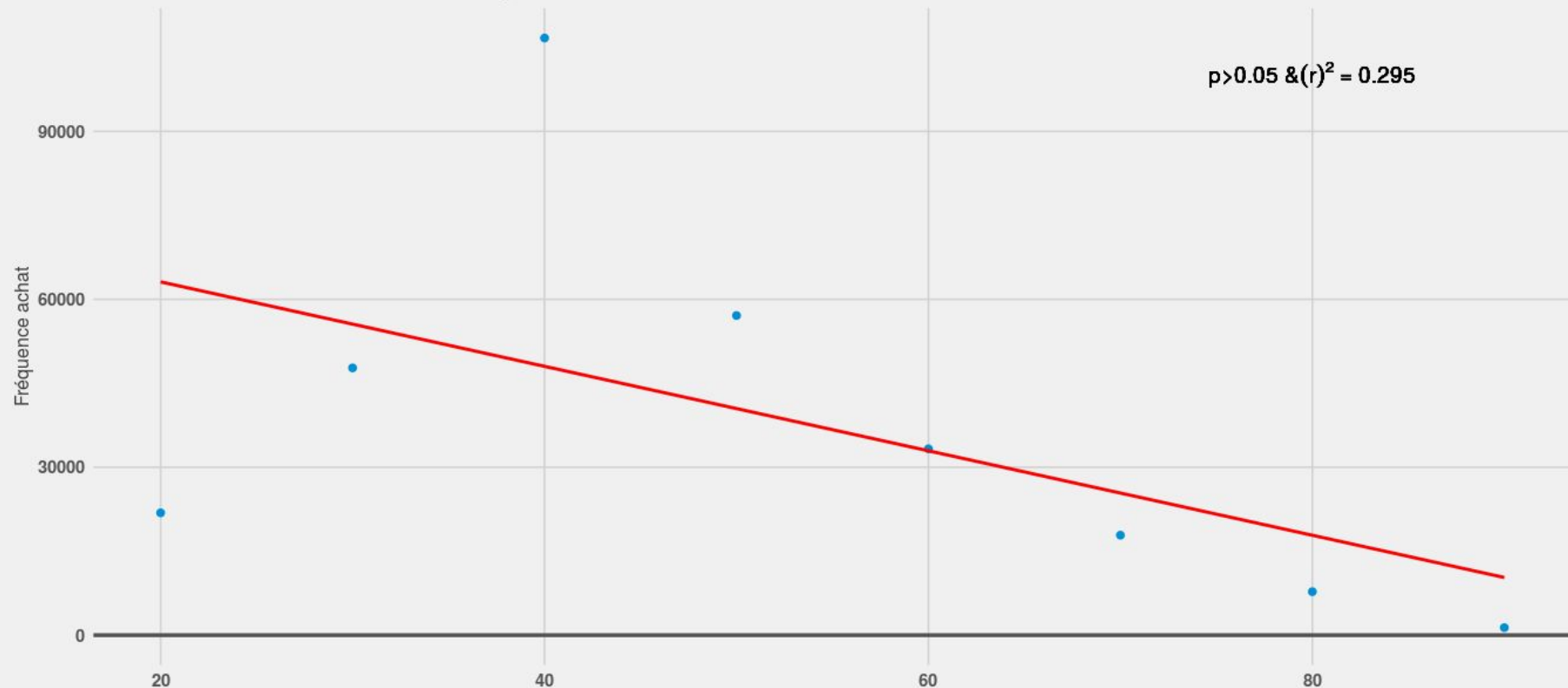
Relation age client et fréquence achat



Relation significative et faible

30-50 ?! & age < 30 ?!

Relation décennie client et fréquence achat

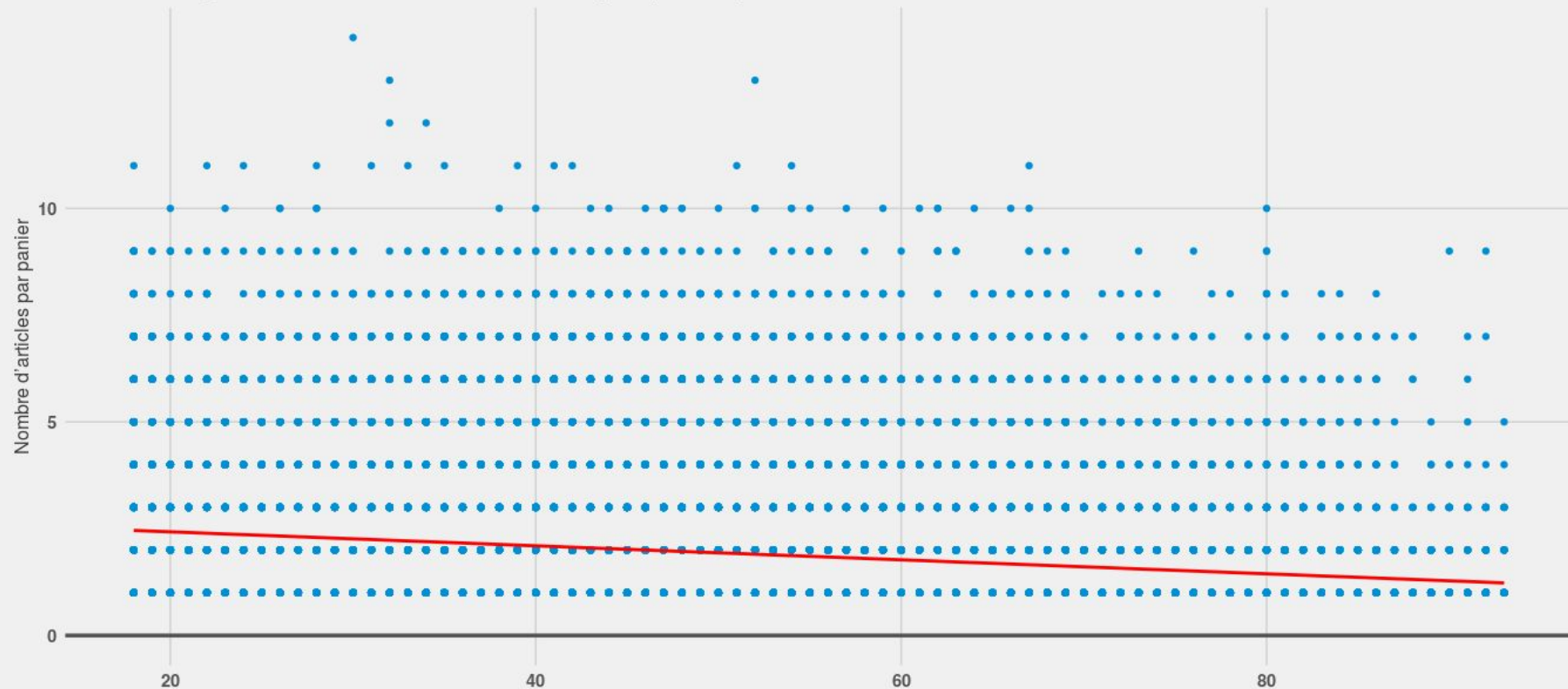


Relation NON significative

30-50 ?! & age < 30 ?!

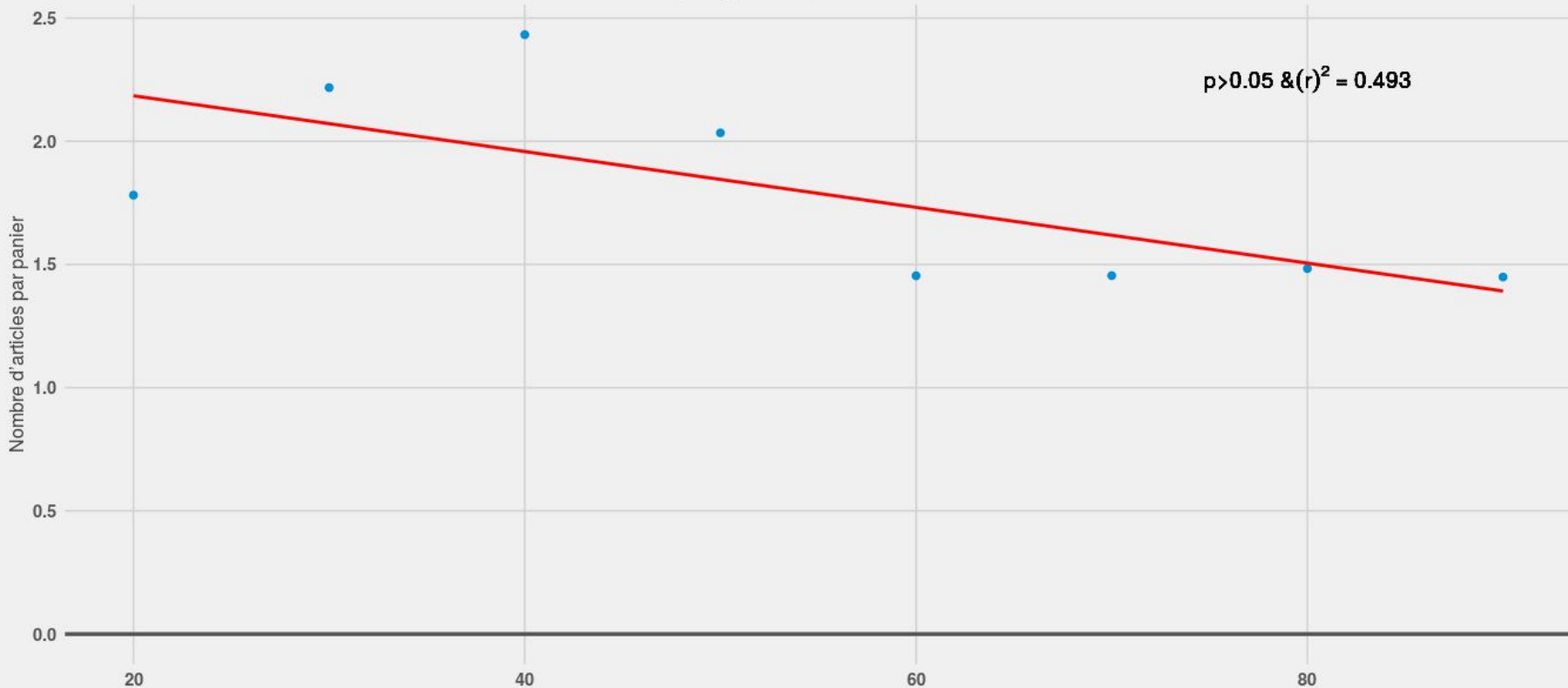
Y a-t-il une corrélation entre l'âge des clients et la taille du panier moyen ?

Relation age client et nombre d'articles par panier)



Relation significative et faible

Relation décennie client et nombre d'articles par panier



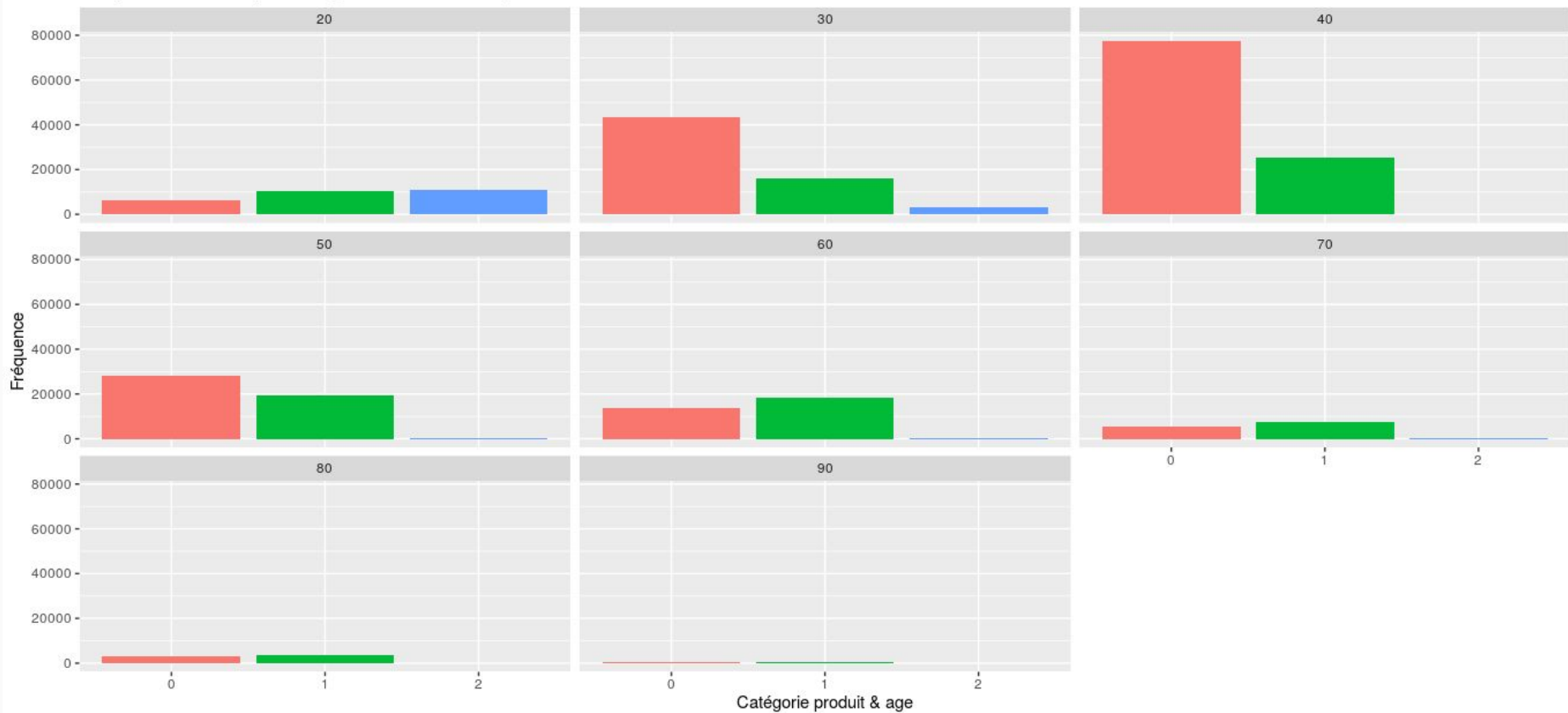
Relation NON significative

chute post 50 ans !!

Segmentation

Y a-t-il une corrélation entre l'âge des clients et la les catégories de produits achetés ?

Fréquence d'achat par catégorie en fonction age



Distribution achats par catégorie semble dépendante de l'âge

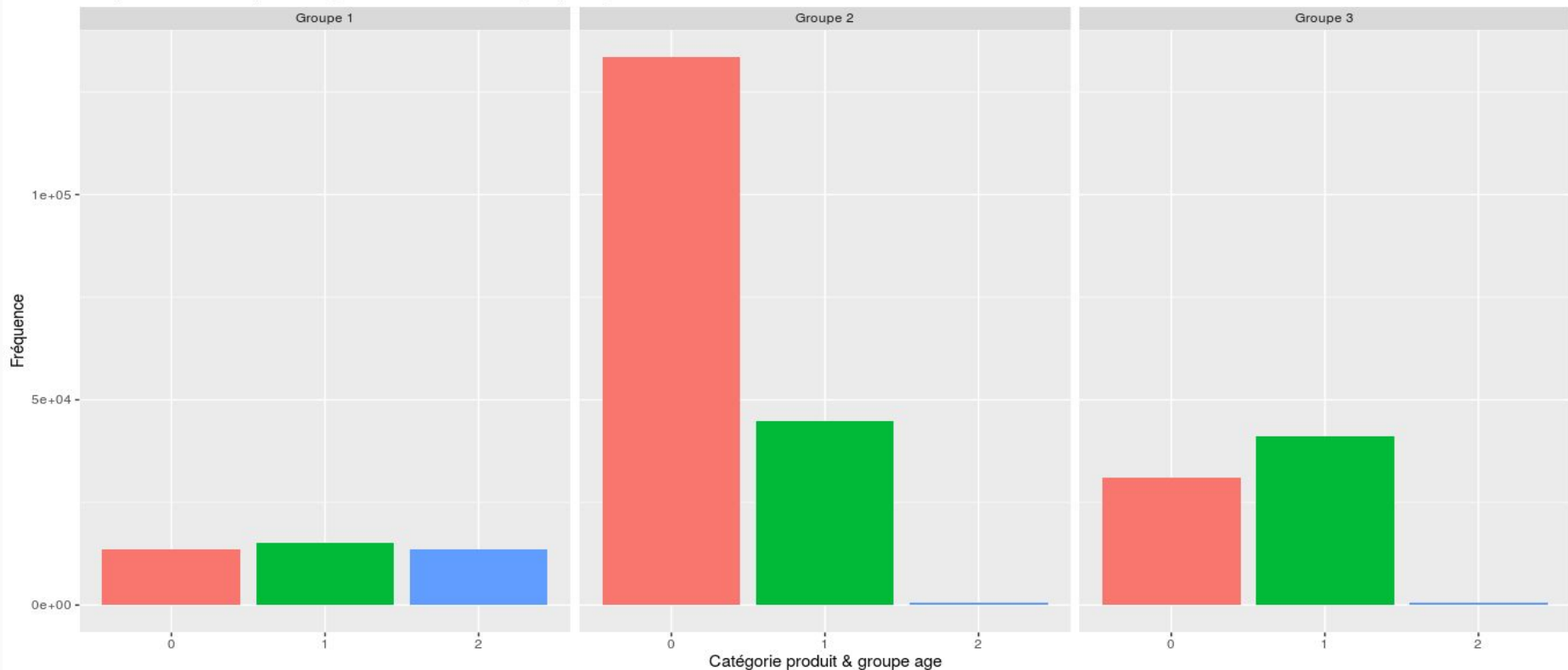
Effet taille ?



X-squared = 98384 / $P < 0.05$ / $\eta^2 = 0.26$

Relation significative et faible, sauf pour 20 ans et categ 2

Fréquence d'achat par catégorie en fonction du groupe age



-30 ans

30-50 ans

+ 50 ans

Distribution achats par catégorie semble dépendante de l'âge

Effet taille ?

Relation groupe age et catégorie

+ 50 ans

Groupe 3

0.04

0.1

0.02

30-50 ans

Groupe 2

0.06

0.04

0.08

-30 ans

Groupe 1

0.06

0

0.6

0

1

2

X-squared = 102633 / $P < 0.05$ / $\eta^2 = 0.20$

Relation significative et faible, sauf pour -30 ans et categ 2

POINTS CLÉS

Y a-t-il une corrélation entre le sexe des clients et les catégories de produits achetés ?	OUI , FAIBLE
Y a-t-il une corrélation entre l'âge des clients et le montant total des achats	OUI, FORTE
Y a-t-il une corrélation entre l'âge des clients et la fréquence d'achat	NON, sauf 30-50 ans
Y a-t-il une corrélation entre l'âge des clients et la taille du panier moyen	INCERTAIN
Y a-t-il une corrélation entre l'âge des clients et la les catégories de produits achetés.	OUI, FAIBLE

OPPORTUNITÉS

1. **Fidélisation client type :**

Cible très présente et à haute fréquence d'achat sur catégorie 0 & 1

Probablement déjà atteinte par com & marketing

OPPORTUNITÉS

2. Rentabilité client -30 ans :

Cible CA important relatif à sa représentation parmi les clients et relation importante dans l'achat catégorie 2

Segmentation com & marketing

OPPORTUNITÉS

3. Diversification client + 50 ans :

Augmentation fréquence et panier moyen (nb articles + faible) via offre groupée, package, etc.

Segmentation com & marketing

MERCI !



@xavbarbier



<https://www.linkedin.com/in/barbierxavier/>



<https://github.com/xavierbarbier/>



contact@xavierbarbier.com