

Analyse des ventes de notre librairie

Les données

Le fichier ‘customer’:

- identifiant client
- genre
- année de naissance

Le fichier ‘product’:

- identifiant produit
- prix
- catégorie

Le fichier ‘transaction’:

- identifiant produit
- date de session
- identifiant de session
- identifiant client

Respect du RGPD: Aucune donnée personnelle présente dans ces fichiers

La préparation des données

Les fichiers ont fait l'objet de très peu de transformations

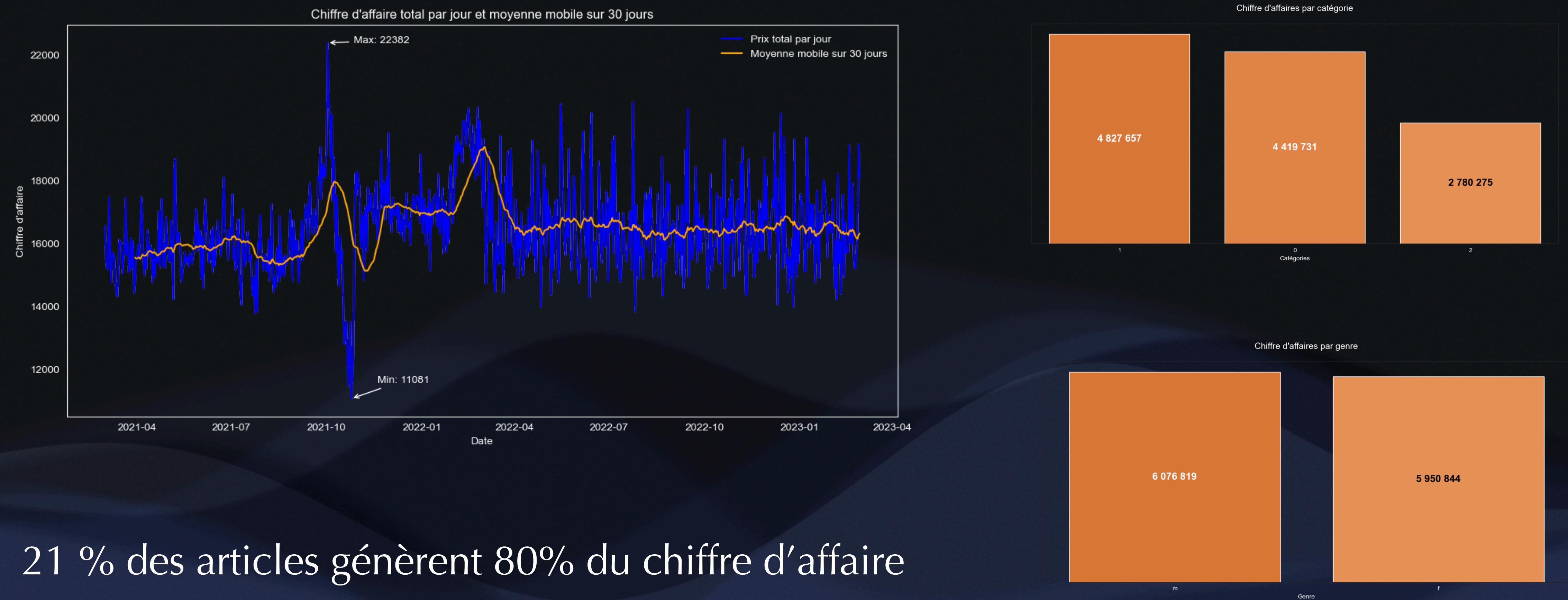
- id_prod renommée prod_id du fichier 'product' pour plus de cohérence avec les autres noms de colonnes.
- Suppression des lignes entièrement vides du fichier 'transaction'.
- id_prod renommée prod_id du fichier 'transaction' pour plus de cohérence avec les autres noms de colonnes.
- Transformation du type de la colonne 'date' en date du fichier 'Transaction' (initialement au format texte).

Limite de l'étude

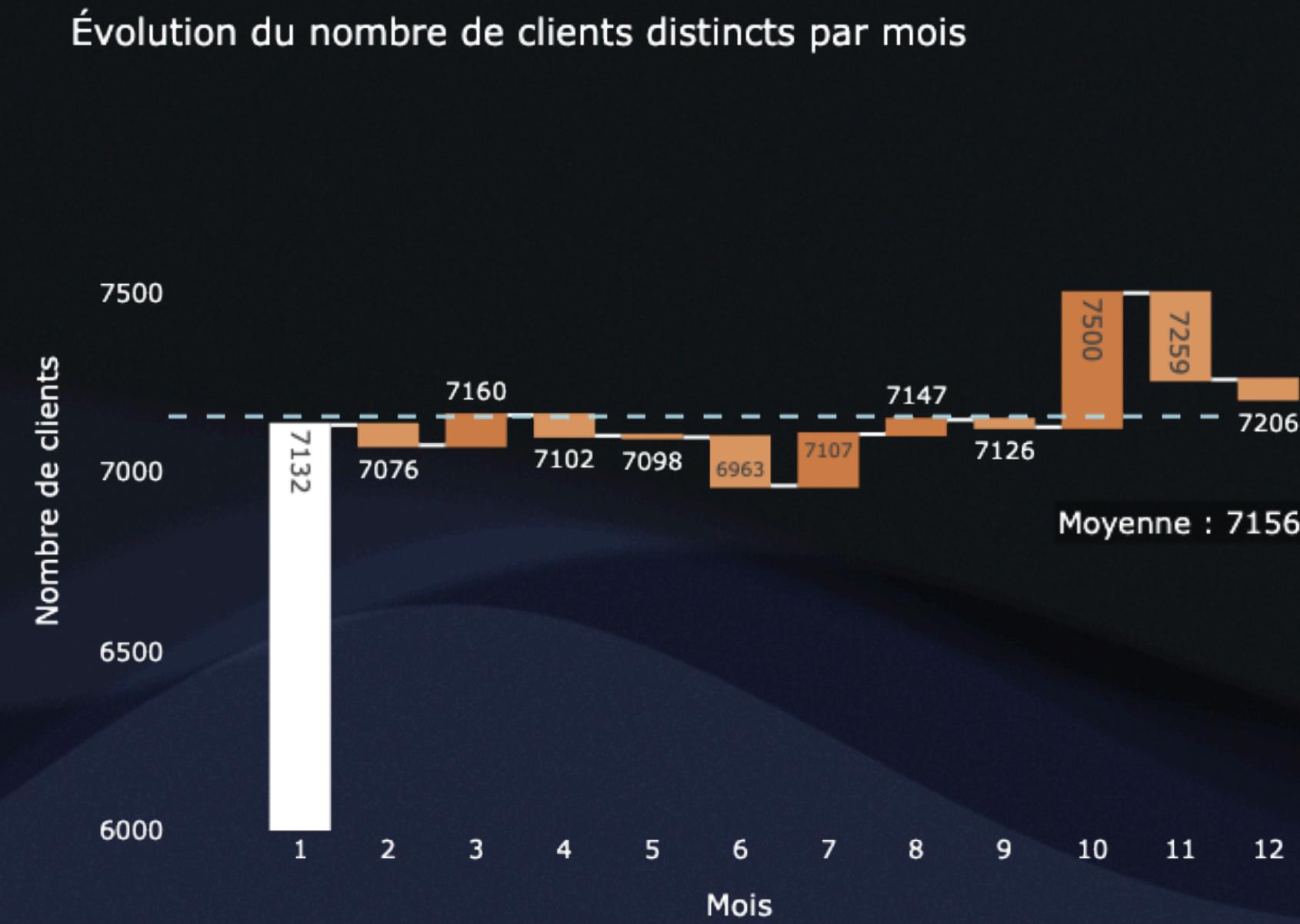
- L'âge des clients n'a pu être calculé précisément, seule l'année de naissance est disponible.
- Aucune indication sur les clients BtoB, ils ont du être estimés.
- Pas de pondération de l'âge par le nombre d'individus.

Que nous disent les données?

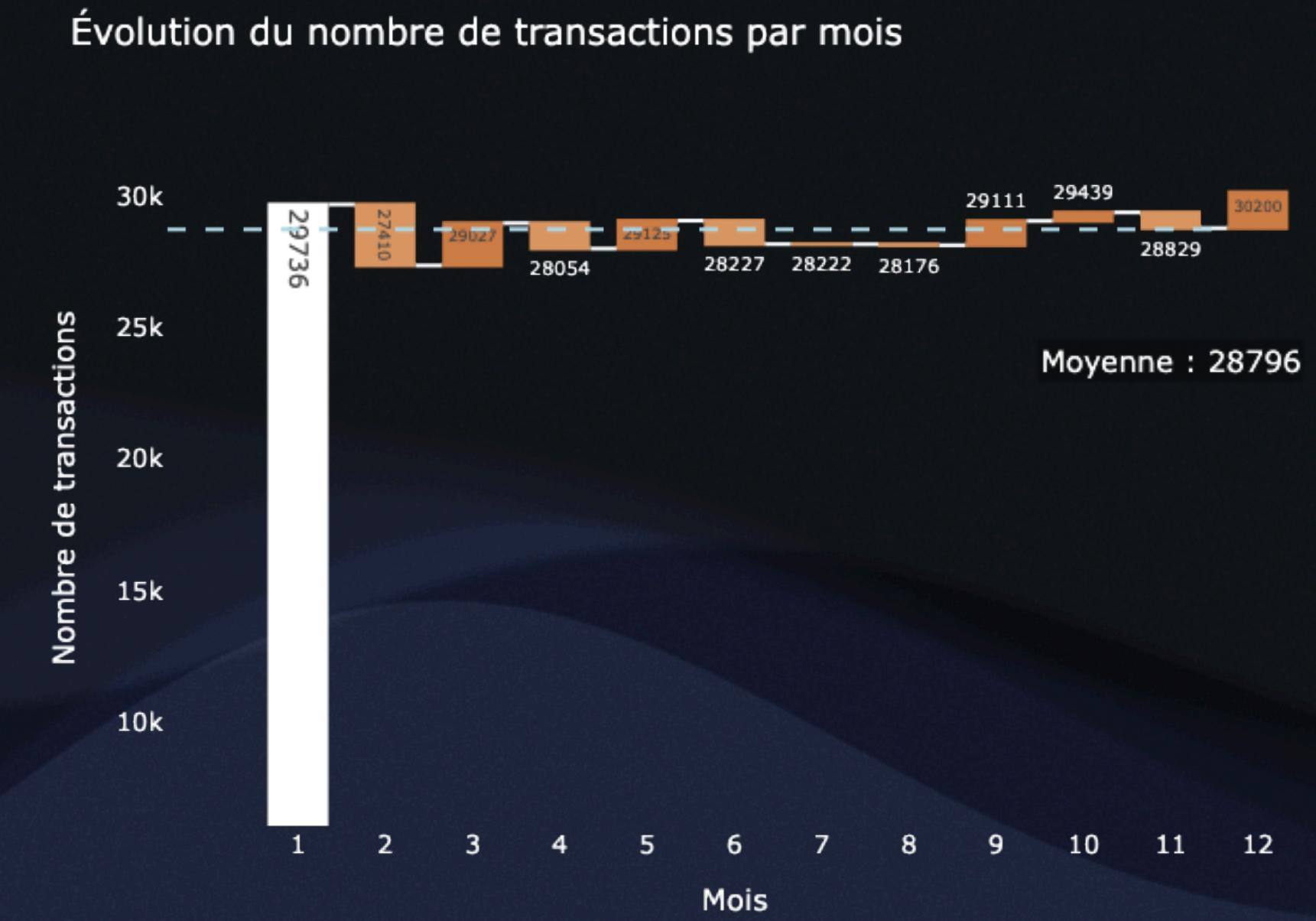
Analyse du chiffre d'affaire



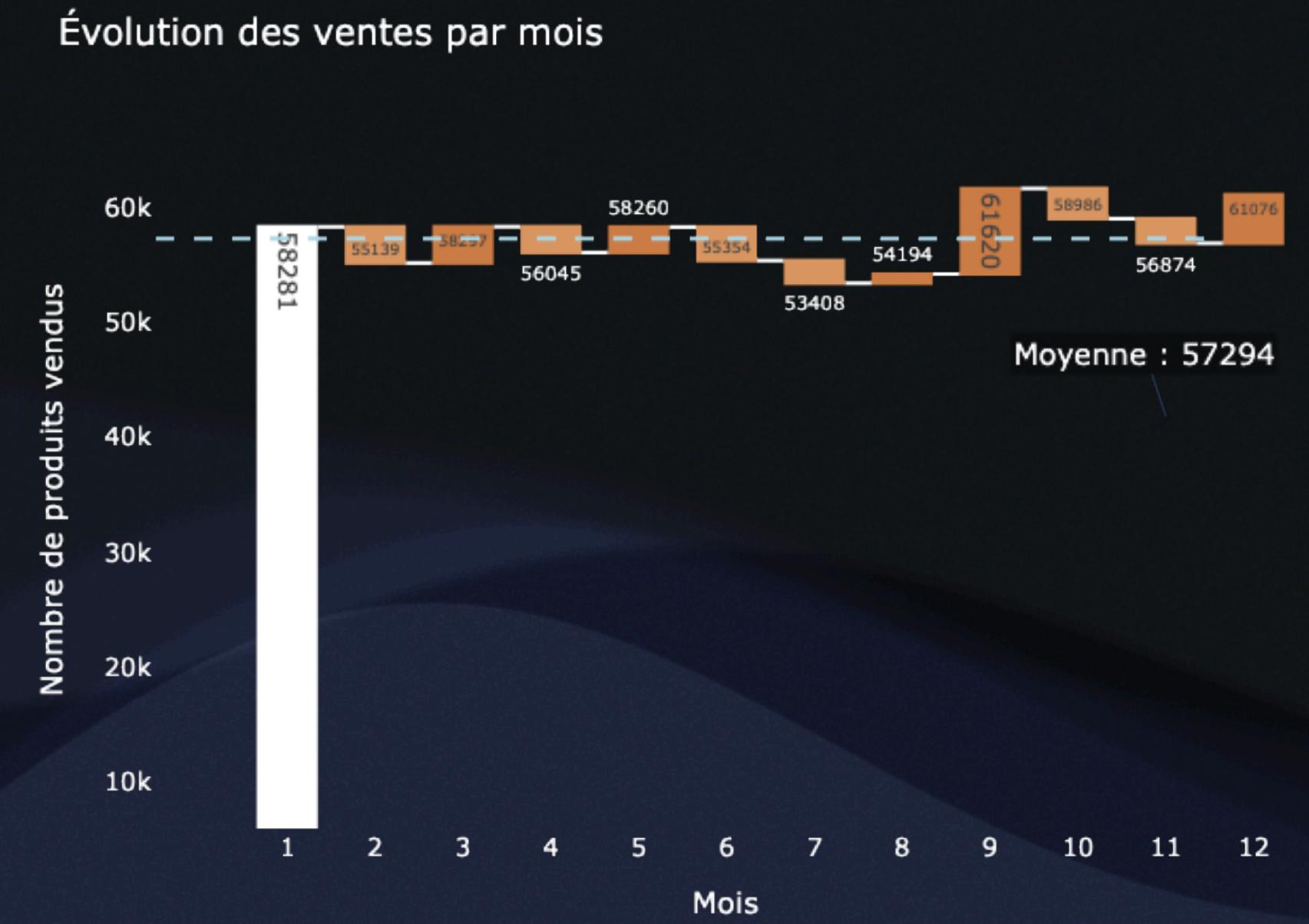
Analyse du nombre de clients



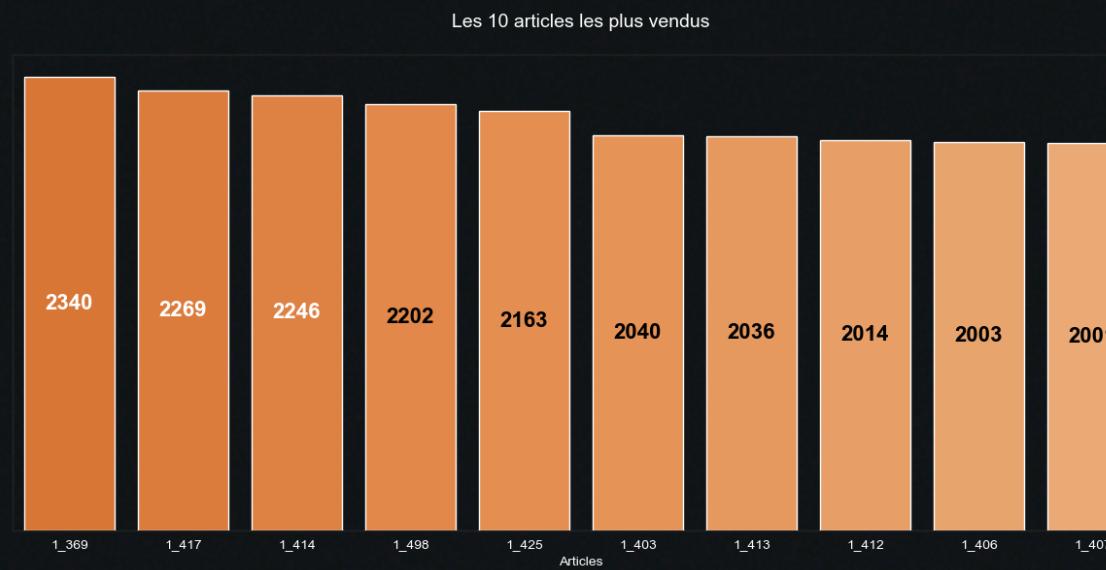
Analyse du nombre de transactions



Analyse du nombre de ventes

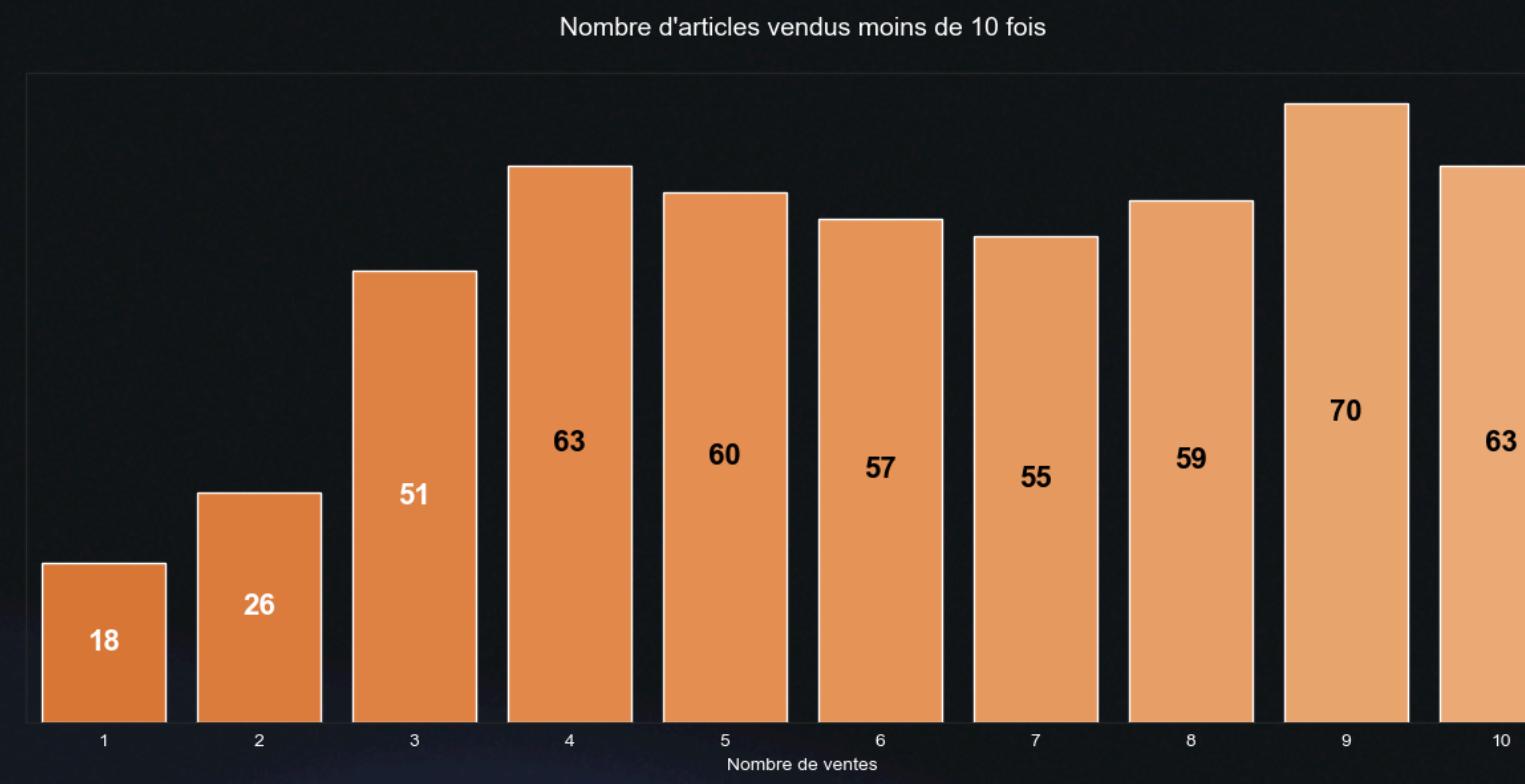
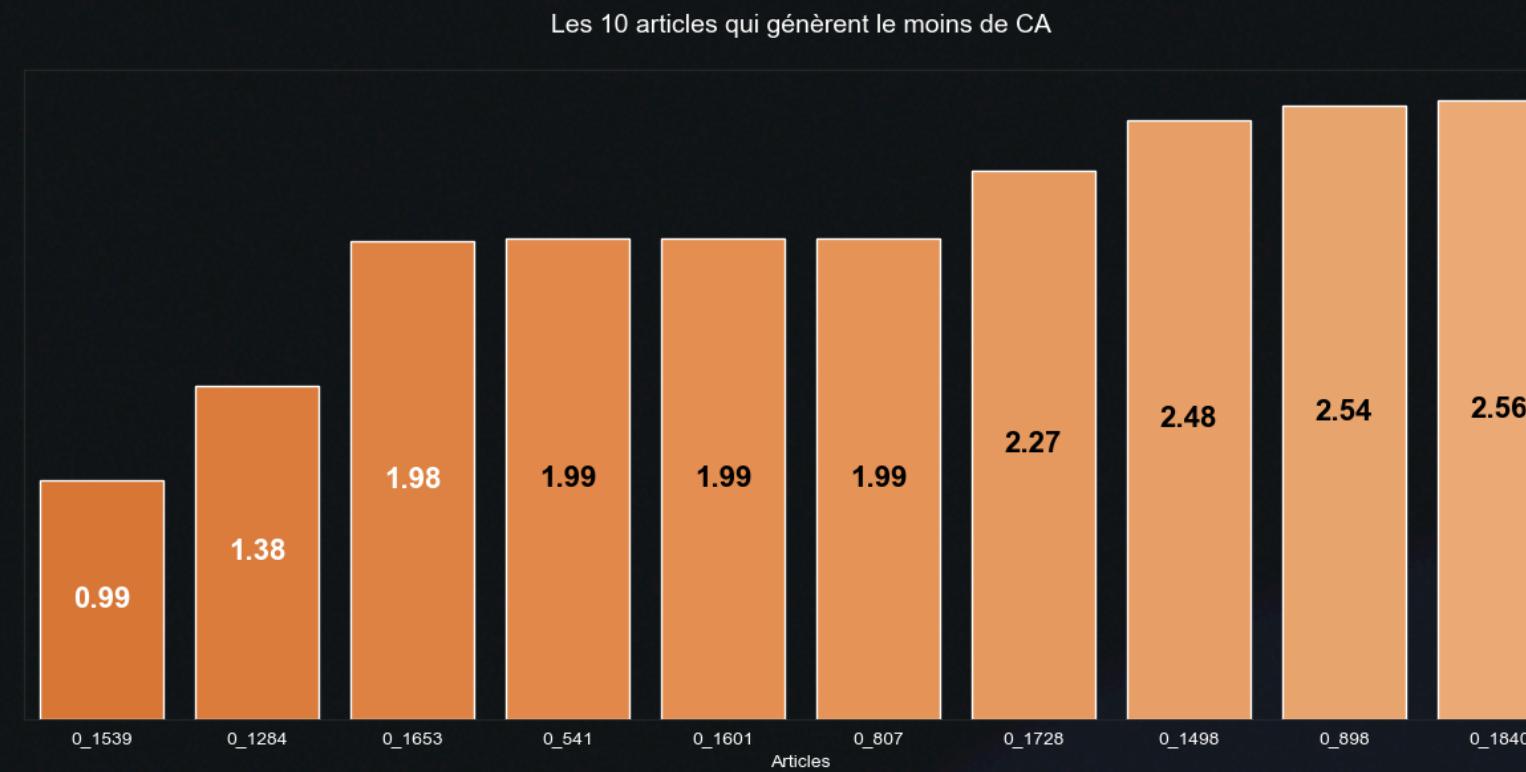


Les top



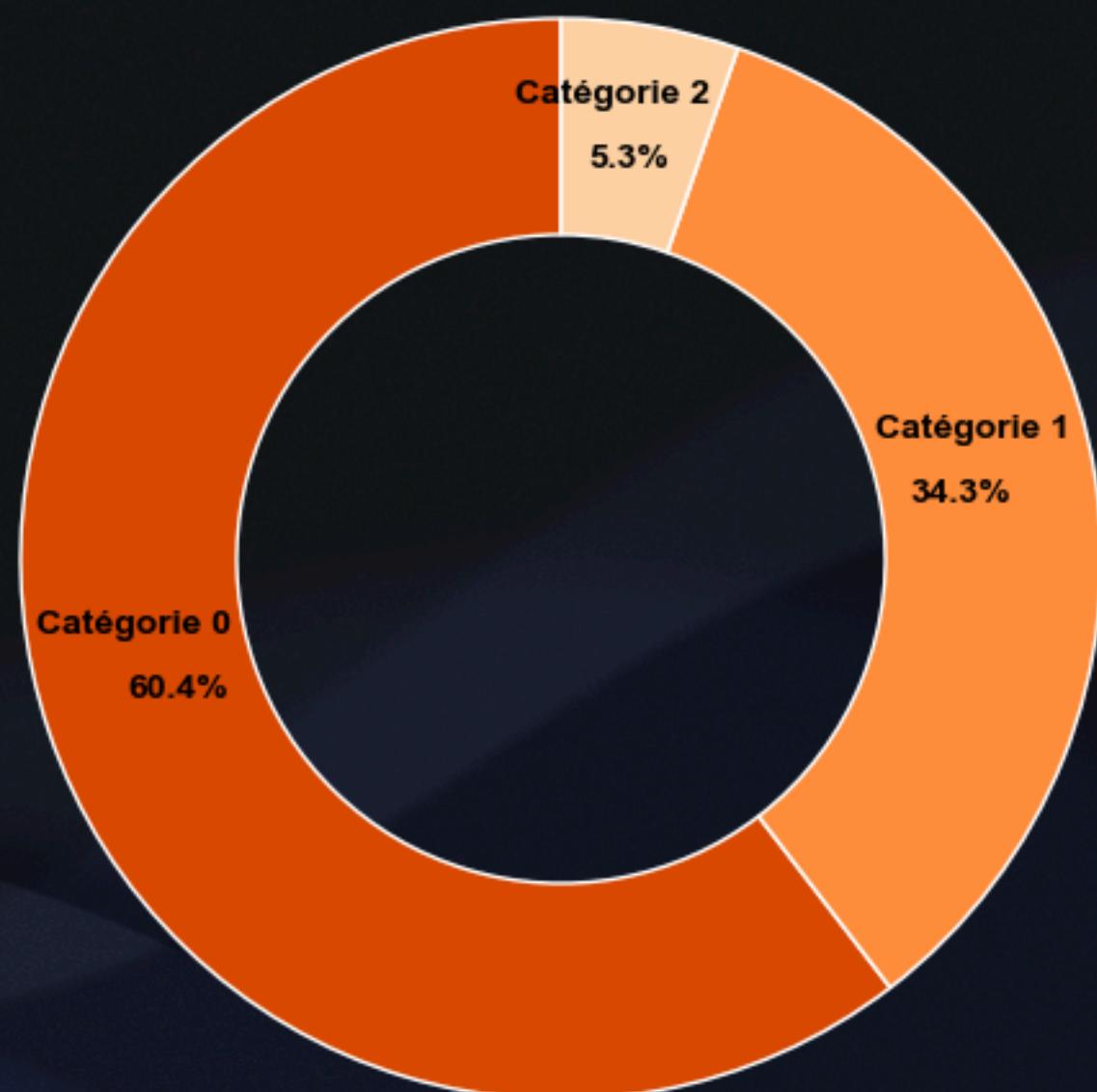
11 articles se trouvent à la fois parmi les articles les plus vendus et parmi les articles qui génèrent le plus de CA

Les flop

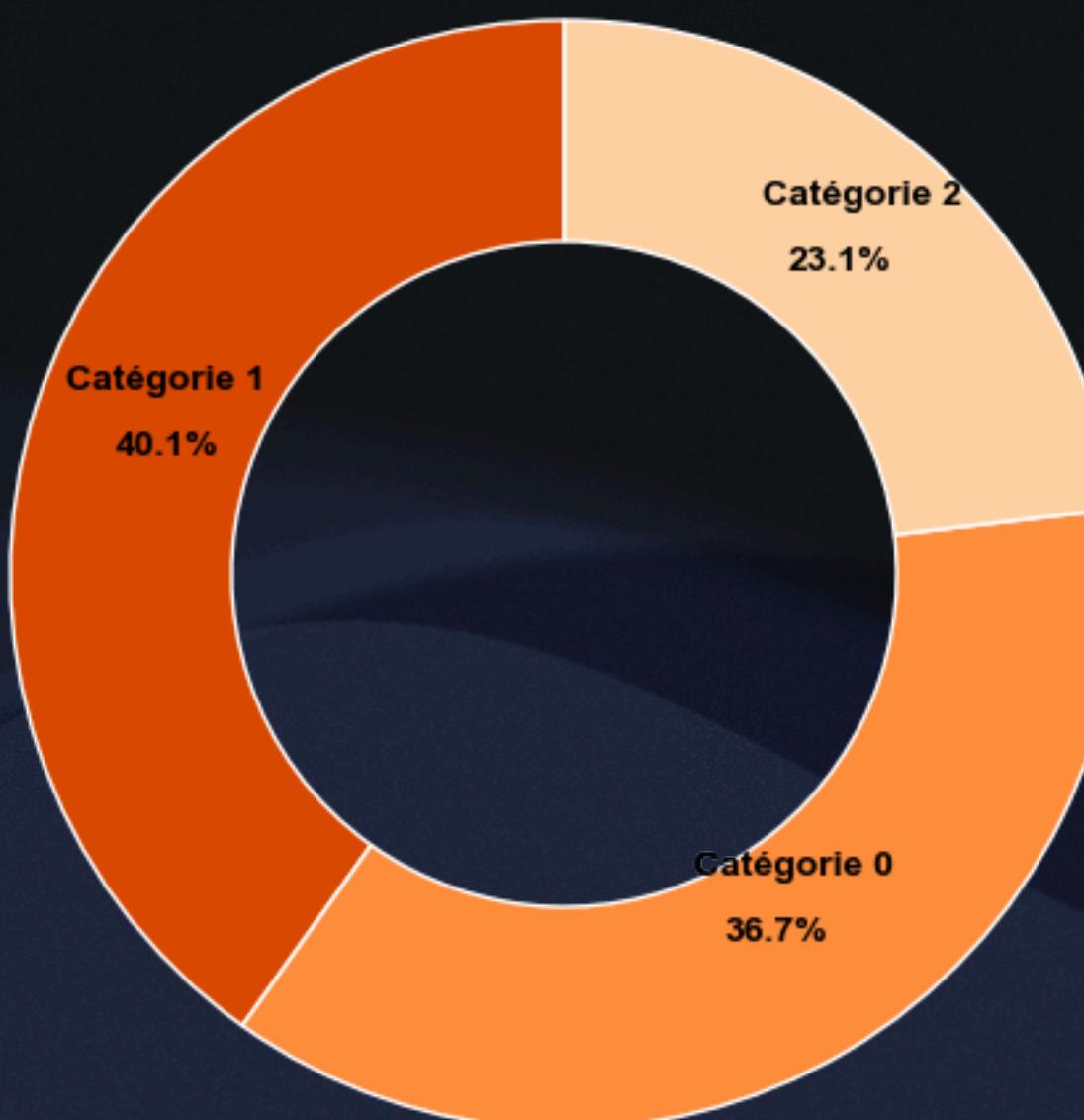


Répartition par catégories

Répartition du nombre de ventes par catégories

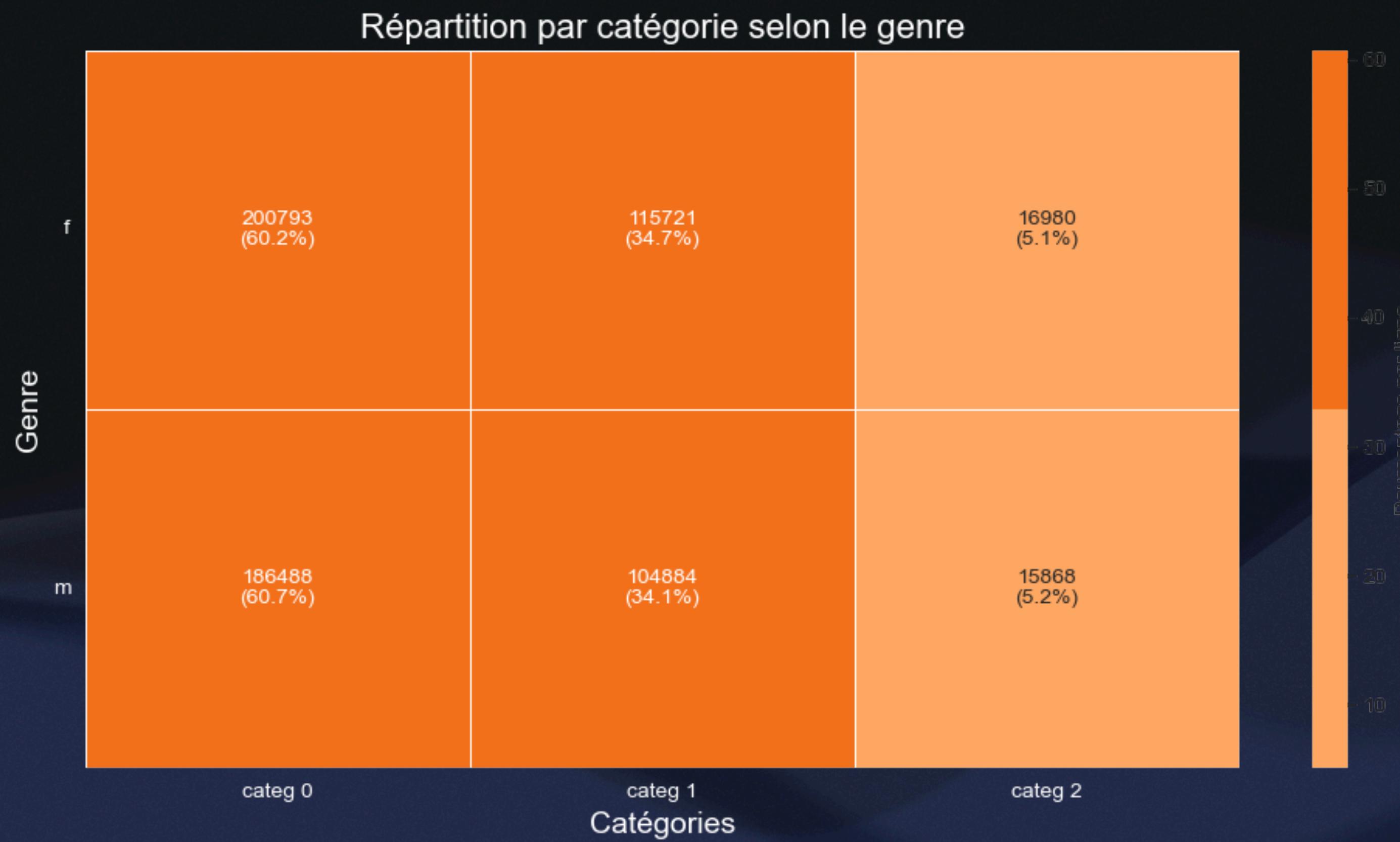


Répartition du chiffre d'affaire par catégories



Les corrélations

Relation entre genres et catégories

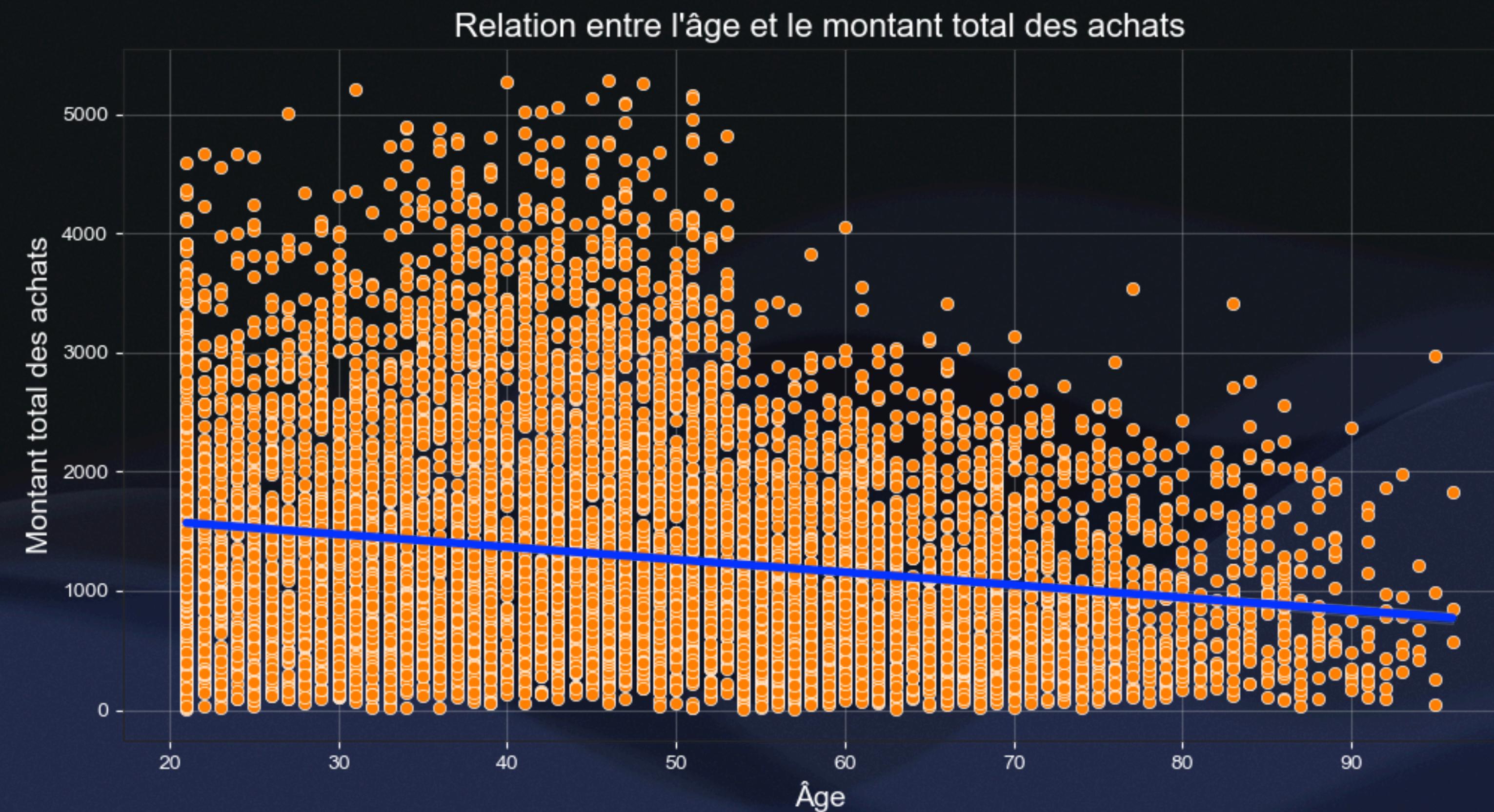


Il existe une relation significative entre le genre et les catégories

Statistique Chi2: 22,66

P valeur: 12.10⁻⁶

Relation entre l'âge et le montant total des achats



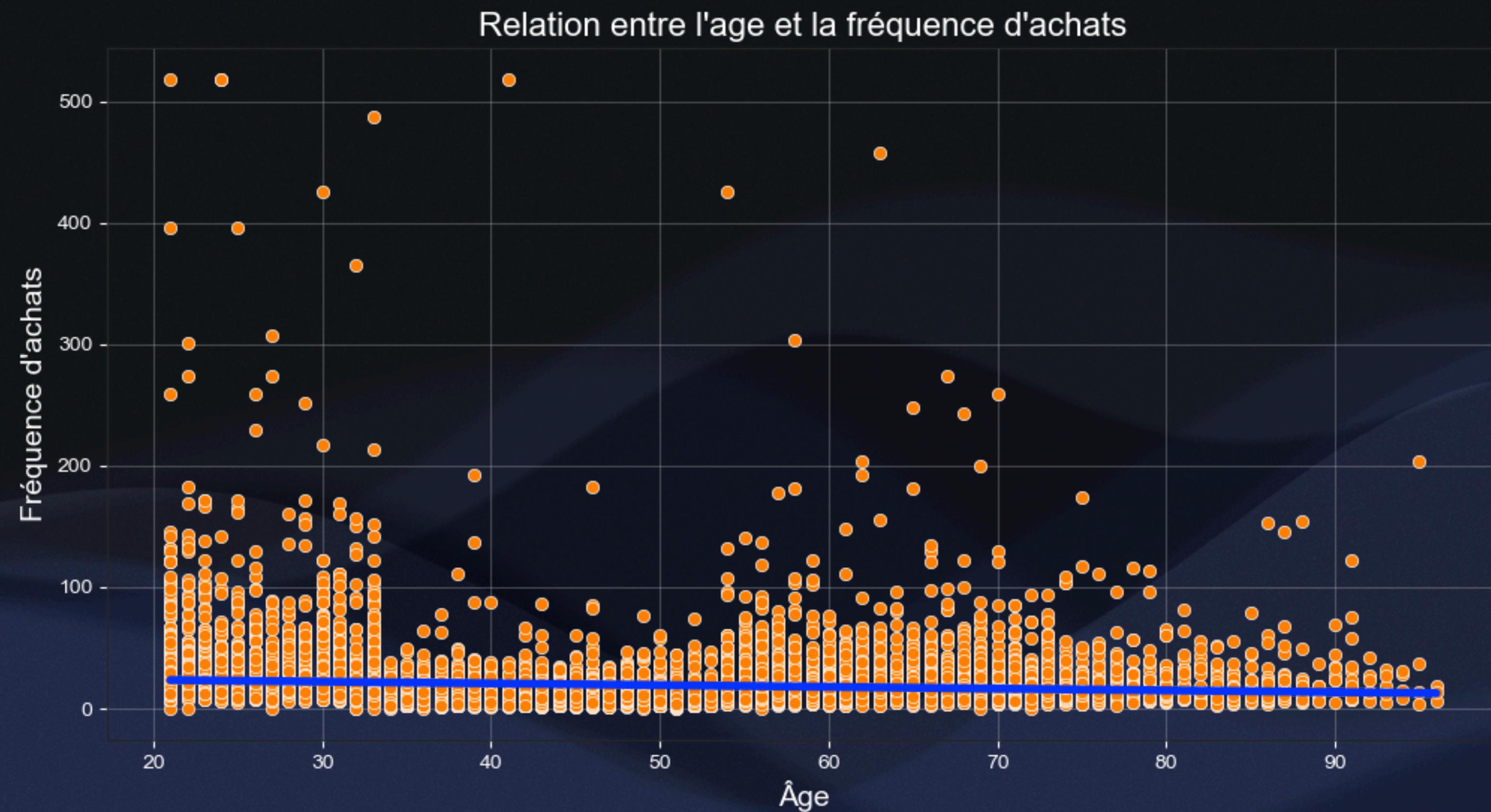
Il existe une faible relation significative entre l'âge et le montant total des achats.

Les moins de 55 ans dépensent davantage.

Coefficient de corrélation de Spearman: - 0,18

P valeur: $10 \cdot 10^{-67}$

Relation entre l'âge et la fréquence d'achats



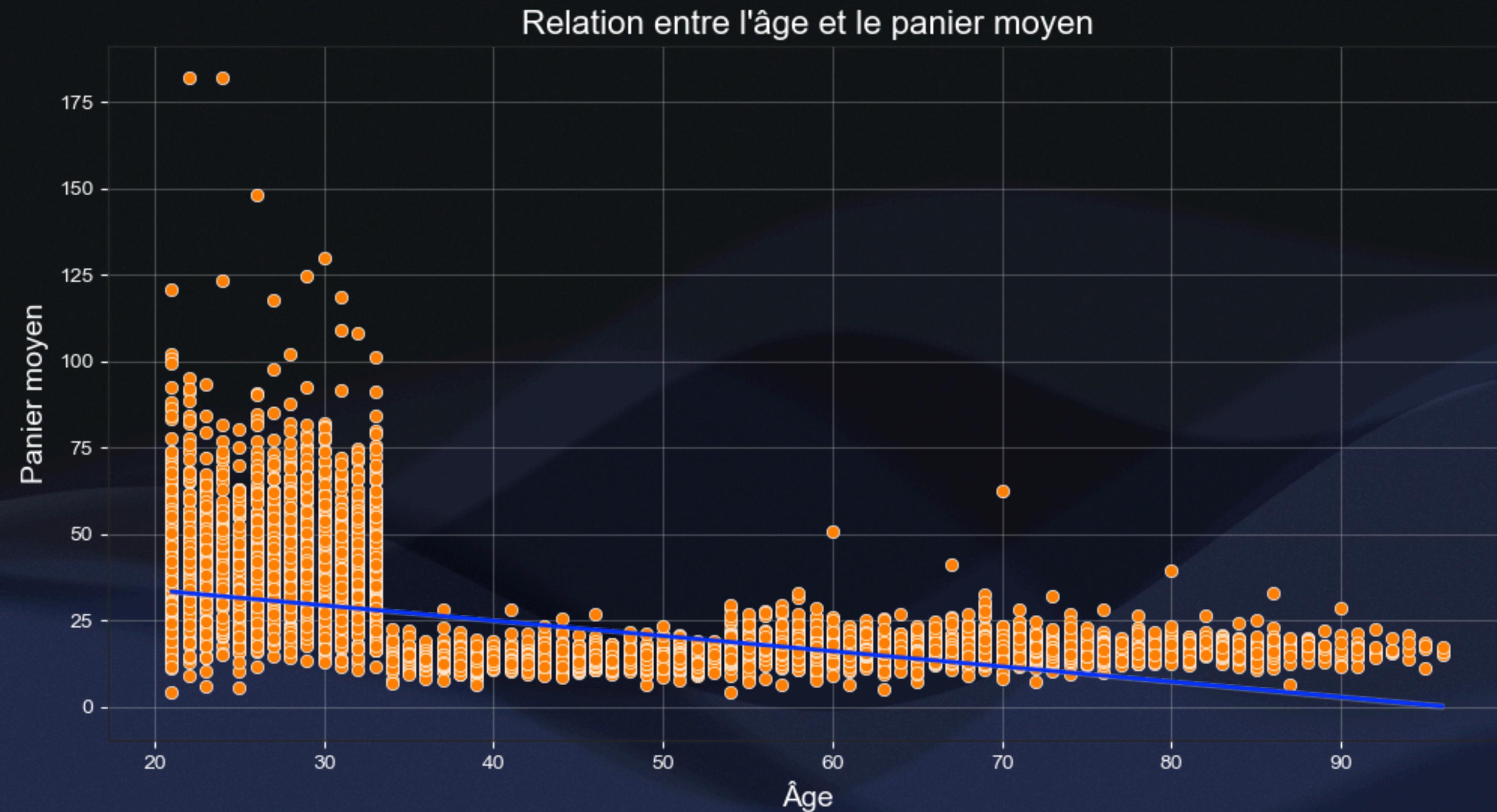
Il existe une faible relation significative entre l'âge et la fréquence d'achats

Les moins de 35 ans ont une fréquence d'achat plus élevées que les plus de 35 ans

Coefficient de corrélation de Spearman: - 0,12

P valeur: 14.10 -29

Relation entre l'âge et le panier moyen



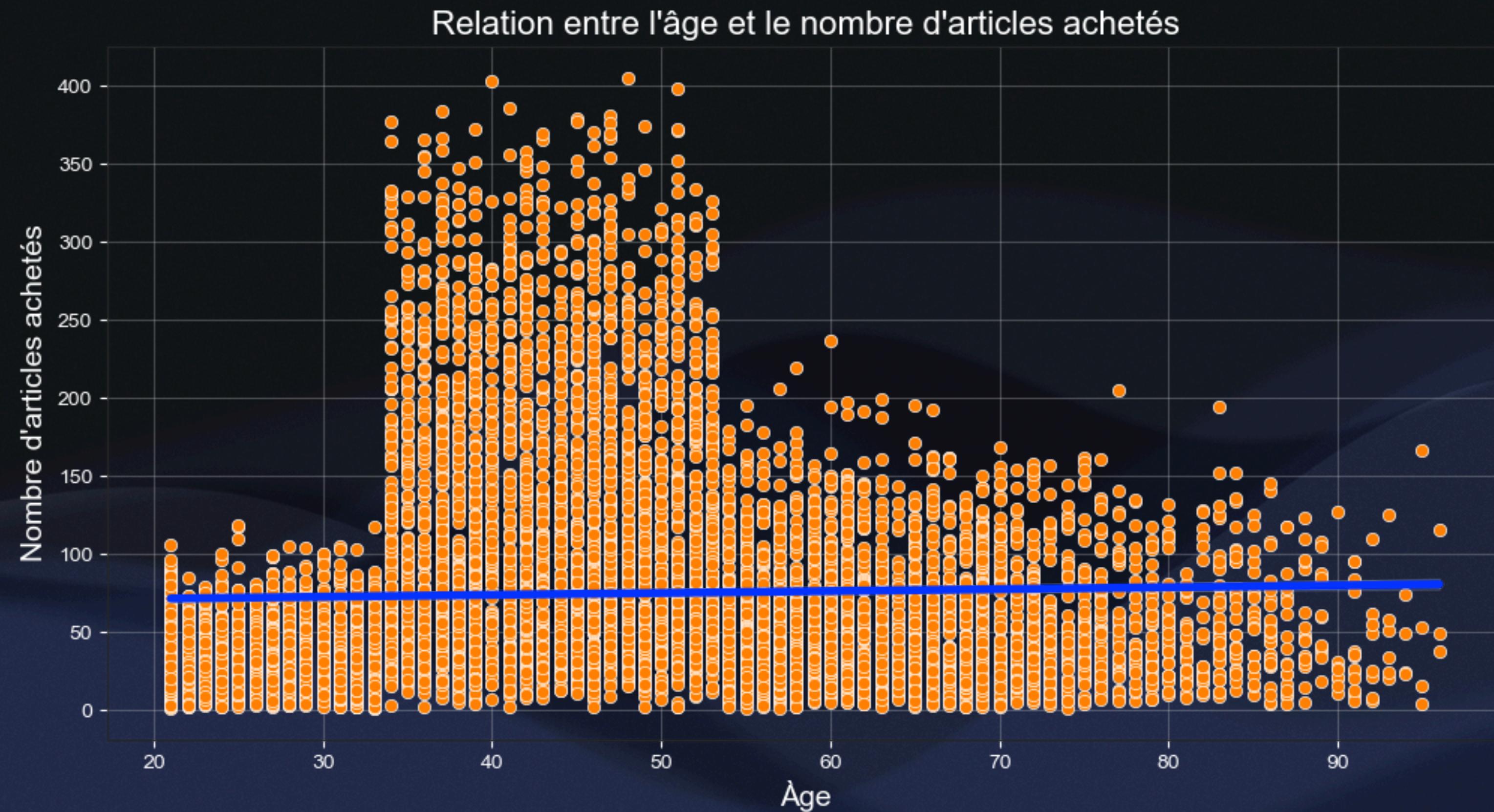
Il existe une faible relation significative entre l'âge et le panier moyen

Les moins de 35 ans ont un panier moyen supérieurs aux plus de 35 ans

Coefficient de corrélation de Spearman: - 0,33

P valeur: $83 \cdot 10^{-213}$

Relation entre l'âge et le nombre d'articles achetés

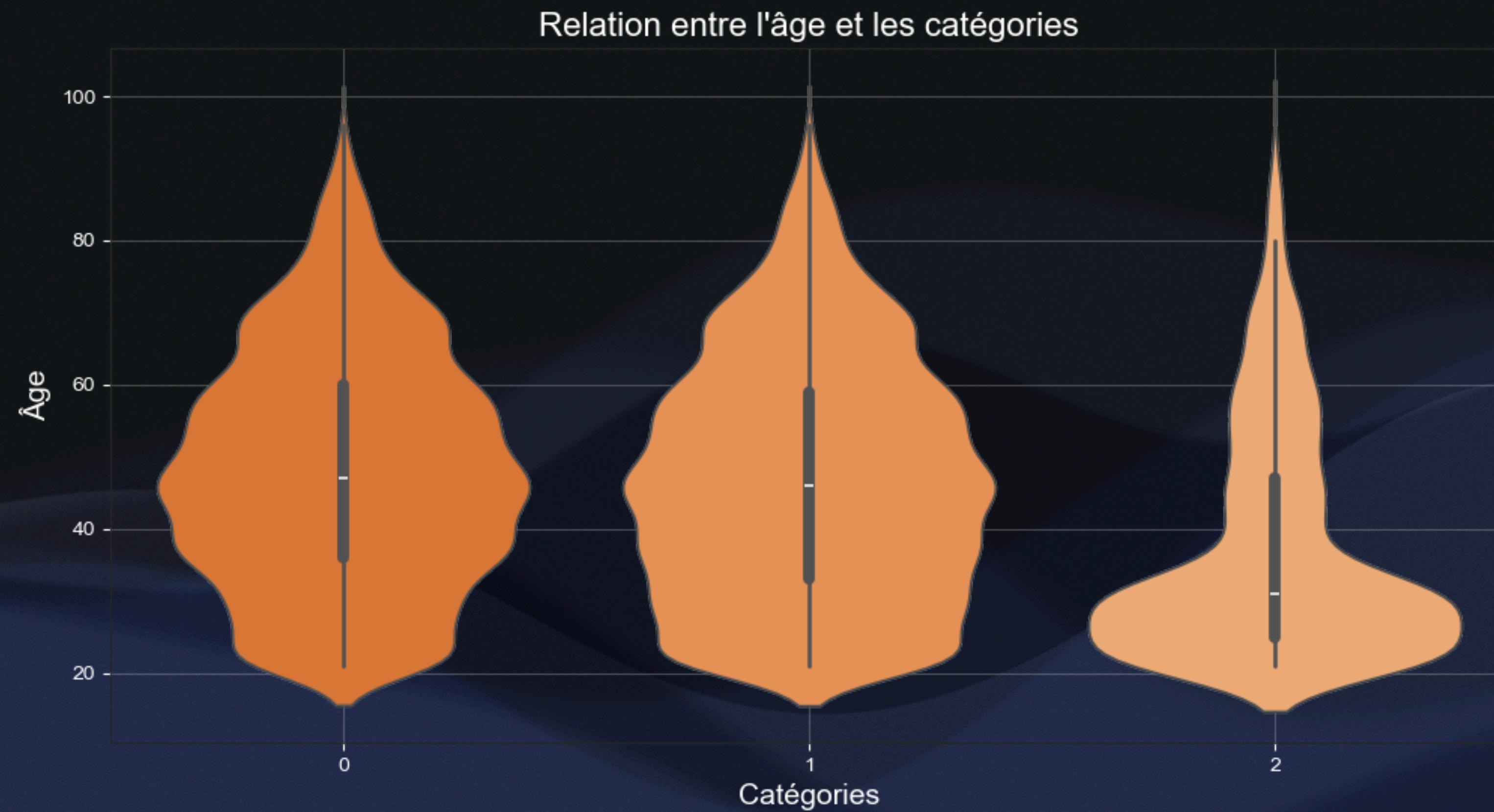


Il existe une faible relation significative entre l'âge et le nombre d'articles achetés
Les 35-55 ans achètent davantage que les autres catégories d'âge

Coefficient de corrélation de Spearman: 0,13

P valeur: $14 \cdot 10^{-33}$

Relation entre l'âge et catégories



Il existe une faible relation inverse significative entre l'âge et les catégories

La catégorie 2 est le choix privilégié des moins de 35 ans

Statistique de Kruskal-Wallis: 1251,5

P valeur: $17 \cdot 10^{-273}$

Conclusion

- ❖ CA et nombre de clients constants. On remarque une légère augmentation d'activité en fin d'année (cumul par mois).
- ❖ Quatre clients se distinguent des autres, très certainement des clients B2B. Ils ont tous les quatre fait de grosses commandes en Septembre 2021 et Février 2023 qu'ils n'ont pas réitérées par la suite (insatisfaction, rentrée scolaire, obtention d'un marché..).
- ❖ De nombreux articles se vendent très peu.
- ❖ La catégorie 0 est la plus vendue, mais c'est la catégorie 1 qui génère le plus de CA.
- ❖ La catégorie 2 se vend peu, mais génère 1/4 du CA. Elle est la préférée des 20-35 ans.
- ❖ L'âge a une influence sur le montant total des achats, la fréquence d'achats, le panier moyen, le nombre d'articles achetés et le choix de la catégorie.

Propositions

- ❖ Promouvoir les moins de 35 ans et la catégorie 2.
- ❖ Promouvoir la vente B2B.
- ❖ Que faire des articles qui génèrent très peu de CA et de ceux qui se vendent très peu? Réaliser une analyse complémentaire et fixer un seuil de rentabilité (valeur stock).
- ❖ 20% des articles génèrent 80% du CA. Analyse des 80% d'articles restants?
- ❖ Définir un seuil minimum de ventes par mois pour décider de conserver un article.
- ❖ Identifier les clients qui génèrent très peu de CA, analyser les articles qu'ils consomment, décider de les conserver ?