

ANALYSEZ LES VENTES D'UNE LIBRAIRIE AVEC PYTHON



Mame Diarra **DABO**

15/05/2023



Lapage

PROBLÉMATIQUE

Consultante Data Analyst



Lapage

L'objectif est « faire le point après deux ans d'exercice, et pouvoir analyser ses points forts, ses points faibles, les comportements clients »

Mission: Demandes de Antoine

- Une analyse des différents indicateurs de vente ;

Mission: Demandes de Julie

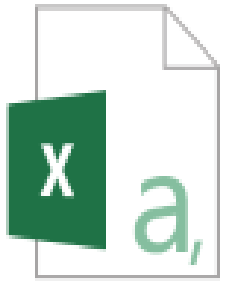
- Une analyse plus ciblée sur les clients.



PLAN

- Importer le jeu de données
- Analyse exploratoire
- Nettoyages des données
- Répondre aux demandes

LE JEU DE DONNÉES



customers.csv



products.csv



transactions.csv



- *Fichier customers : (8623, 3)*
- *Fichier products : (3287, 3)*
- *Fichier transactions : (679532, 4)*

- Nombre de lignes et colonnes :
 - => les types de données
 - => Vérifier les doublons
 - => Vérifier les valeurs manquantes

ANALYSE EXPLORATOIRE

Statistique descriptive des données

Transactions

	count	unique	top	freq
id_prod	679532	3267	1_369	2252
date	679532	679371	test_2021-03-01 02:30:02.237413	13
session_id	679532	342316	s_0	200
client_id	679532	8602	c_1609	25488

Produits

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
id_prod	3287	3287	0_1421	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN
price	3287.0	NaN	NaN	NaN	21.856641	29.847908	-1.0	6.99	13.06	22.99	300.0
categ	3287.0	NaN	NaN	NaN	0.370246	0.615387	0.0	0.0	0.0	1.0	2.0

Clients

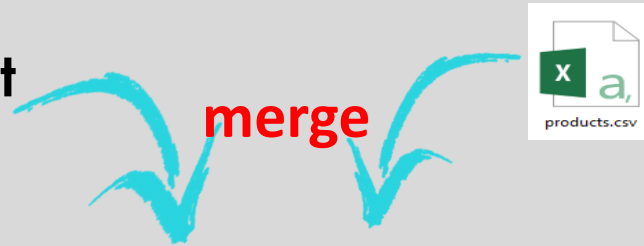
	count	unique	top	freq	mean	std	min	25%	50%	75%	max
client_id	8623	8623	c_4410	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN
sex	8623	2	f	4491	NaN	NaN	NaN	NaN	NaN	NaN	NaN
birth	8623.0	NaN	NaN	NaN	1978.280877	16.919535	1929.0	1966.0	1979.0	1992.0	2004.0

NETTOYAGES DES DONNÉES



Un fichier en fonction des transactions par client : 679353 transactions clients
7 colonnes

- **Résultat**



Un fichier en fonction des transactions par produit : 679375 transactions produits en ligne
9 colonnes

NETTOYAGES DES DONNÉES

Un fichier en fonction des transactions par produit : 679375 transactions produits en ligne
9 colonnes



Gestion des valeurs

- **manquantes :**
 - => la valeur la plus fréquente par catégories
 - => la valeur médiane par variables numériques
- **aberrantes :**
 - => aucun outliers semblent être erronés



Features engineering

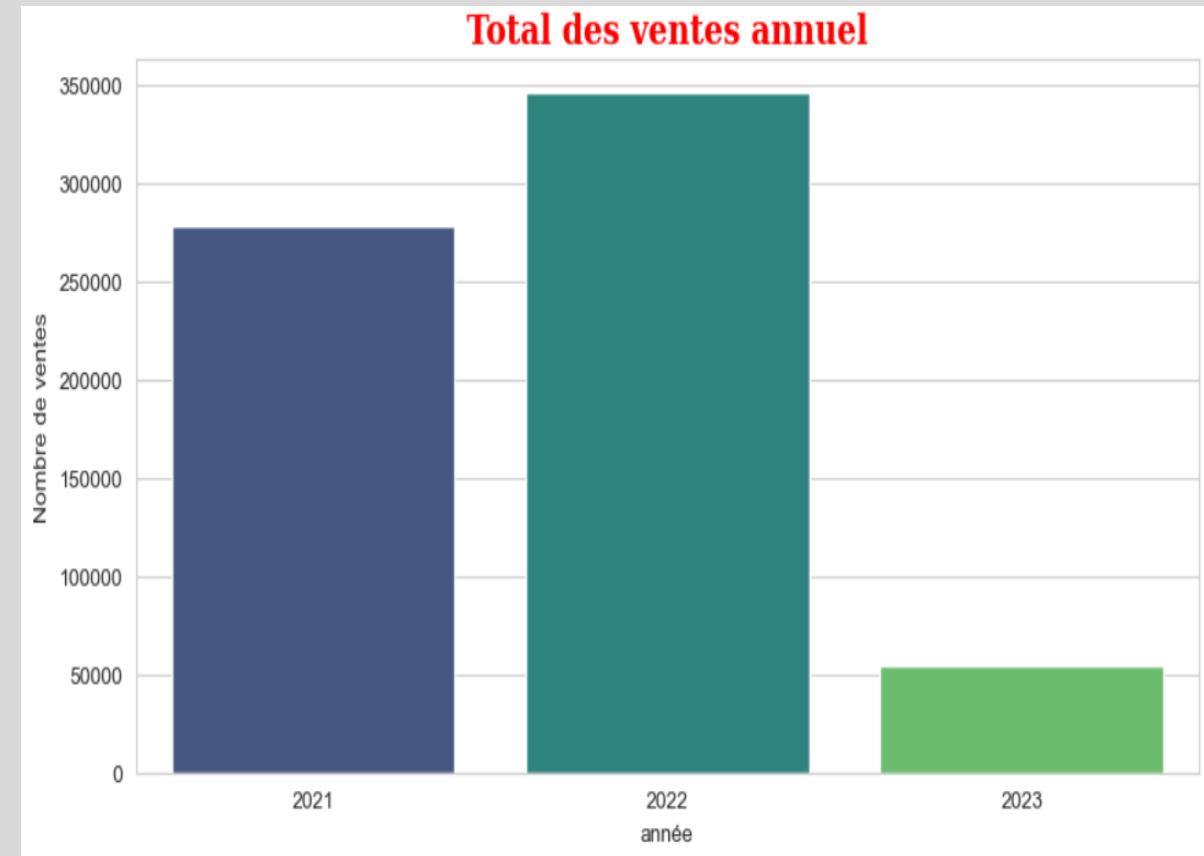
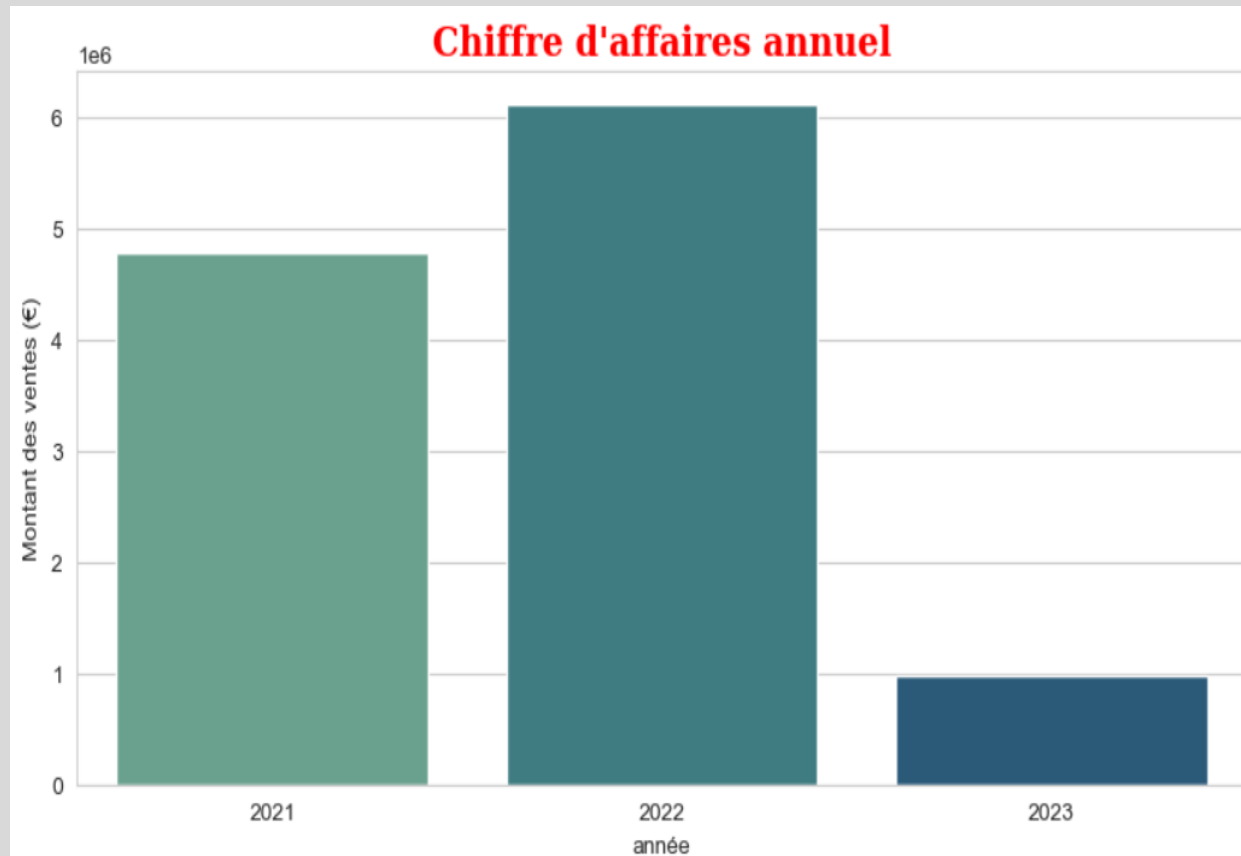
Fichier final : 679111 transactions en lignes

12 colonnes

Demandes de Antoine

Chiffre d'affaires et total des ventes en ligne annuel

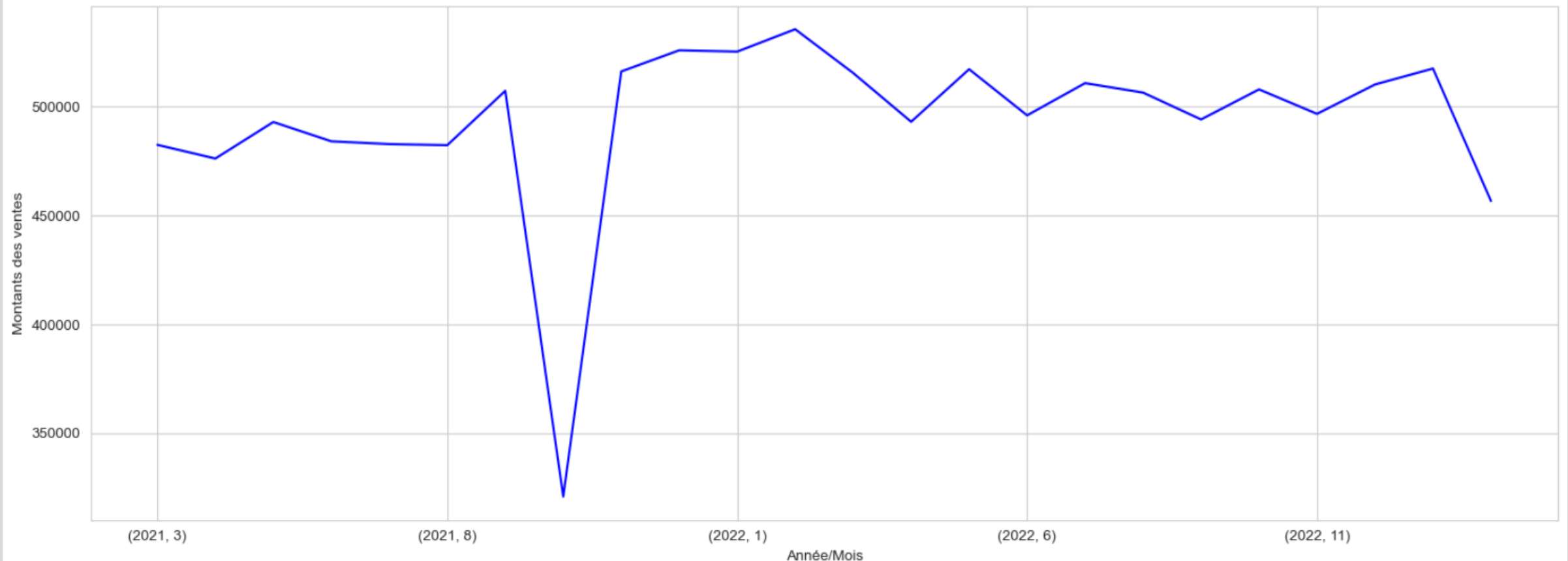
Le total du chiffre d'affaires réalisé des ventes en lignes est : 11855936.47euros
Le nombre total de ventes réalisés en lignes est : 679332 ventes



Demandes de Antoine

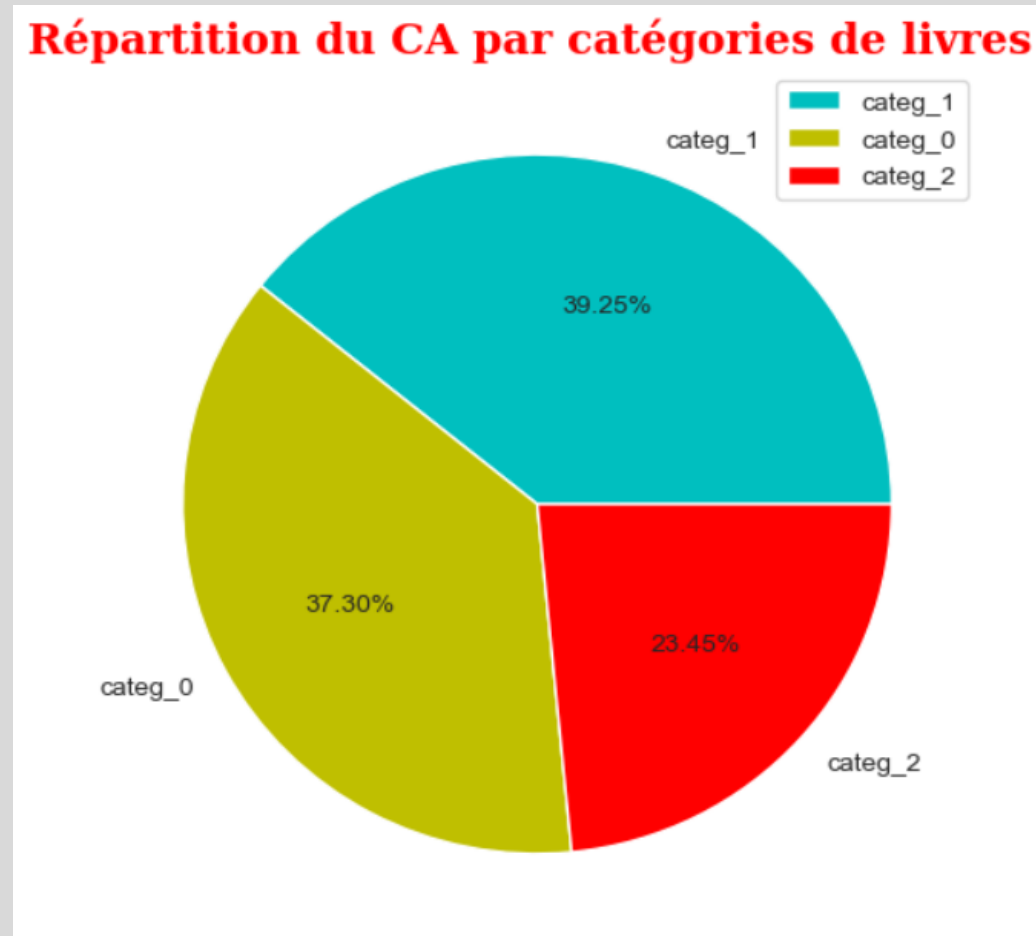
Évolution du chiffre d'affaires

Evolution du chiffre d'affaires année - mois



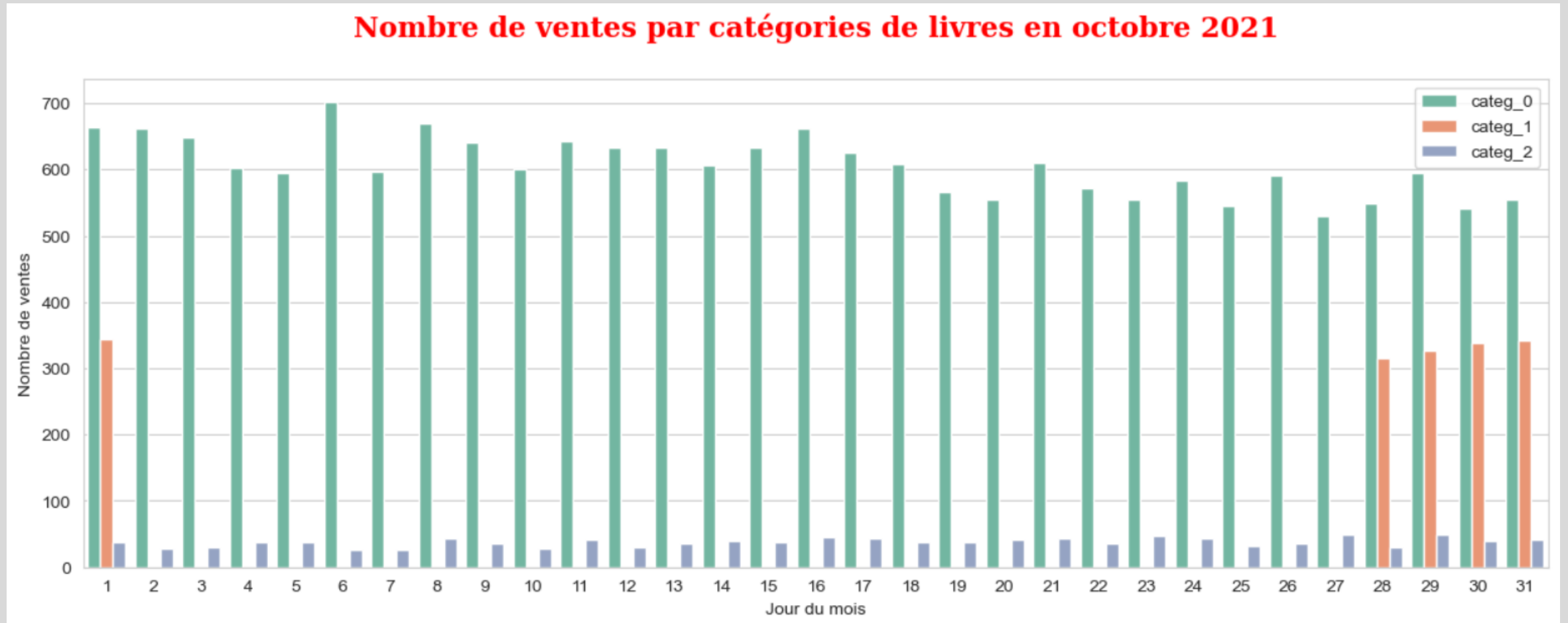
Demandes de Antoine

Les proportions du chiffre d'affaires par catégorie de livres



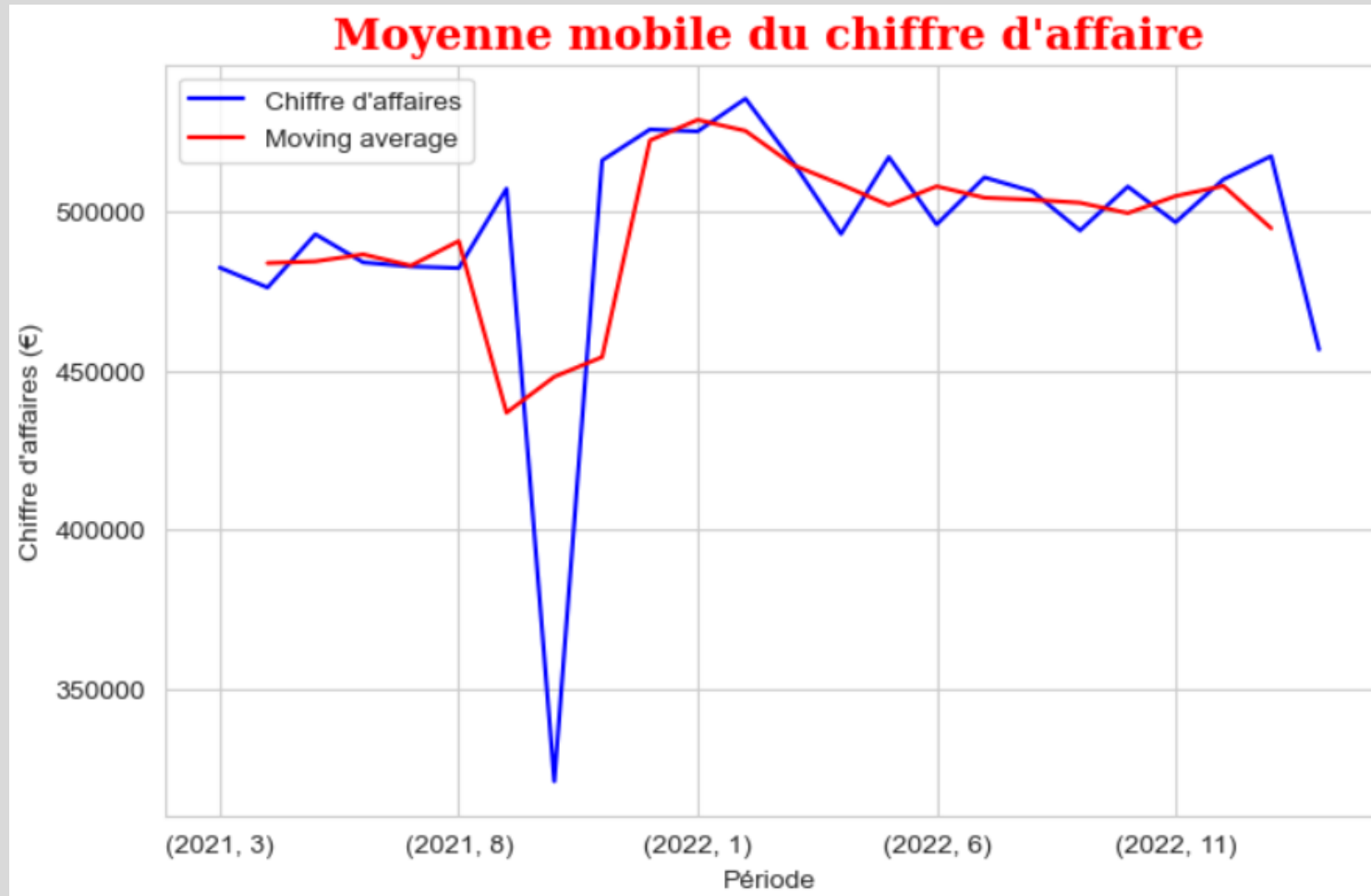
Demandes de Antoine

Les ventes mensuel par catégories de livres au mois d'octobre 2021



Demandes de Antoine

Décomposition moyenne mobile

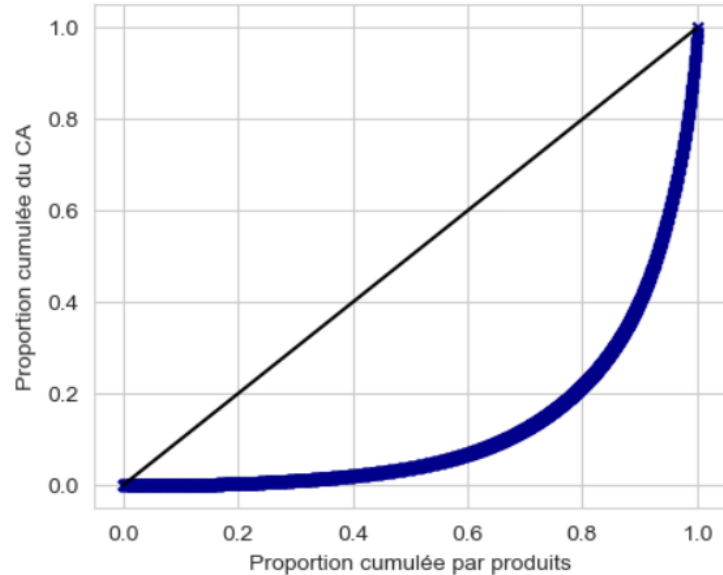


Demandes de Antoine

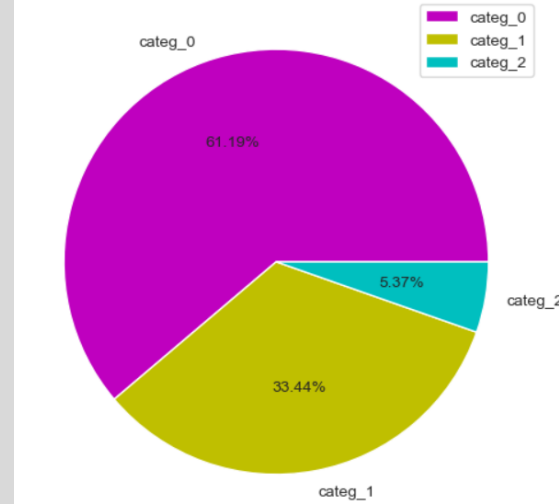
Zoom sur les références et la répartition par catégories de produits

Incide de Gini : 0.743

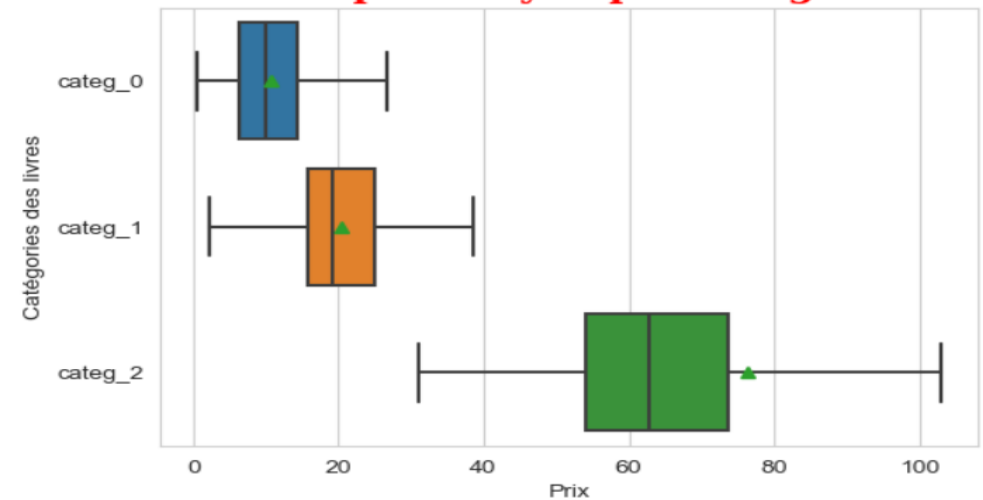
Répartition du chiffre d'affaires entre les produits



Répartition par catégories de livres



Distribution du prix moyen par catégories de livres



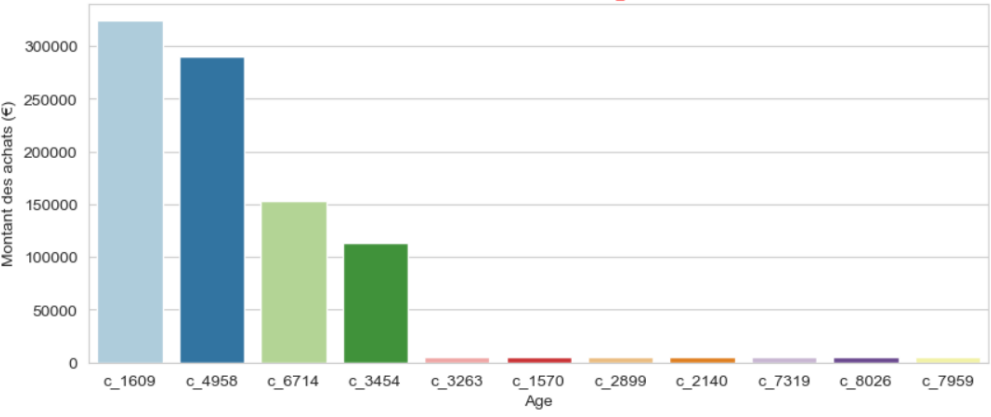
Demandes de Antoine

Analyse sur le profil de nos clients

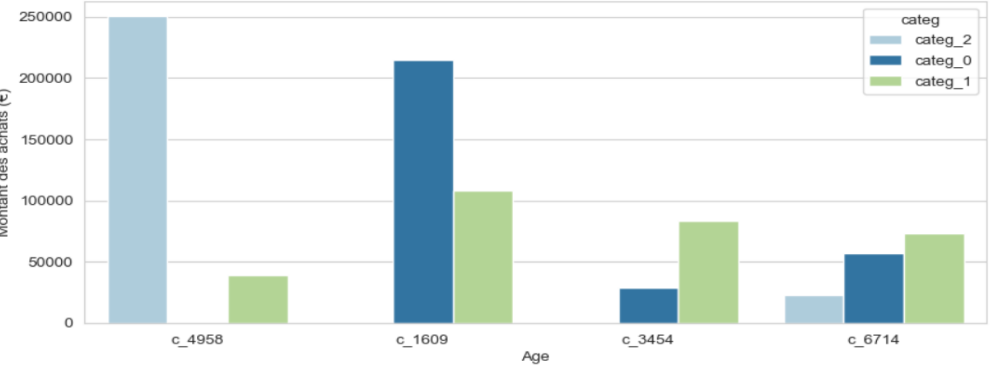
Le montant moyen d'achats par client est de : 1378.597263953488

Le nombre d'achats moyen par client est de : 78.99209302325582

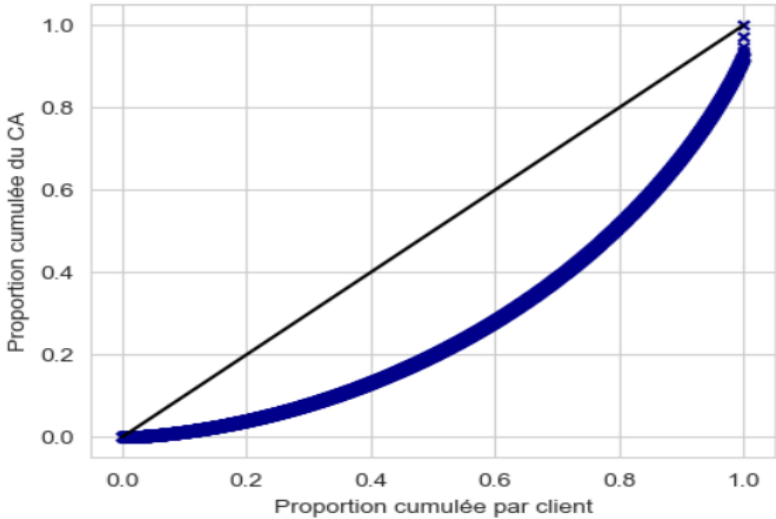
Montant des achats par client



Volume des achats par catégories de livres par client

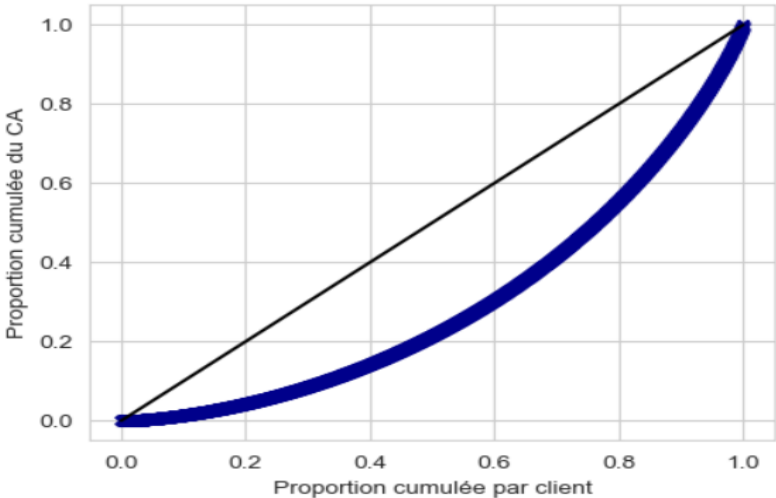


Répartition du chiffre d'affaires entre les clients



Incide de Gini : 0.446

Répartition du chiffre d'affaires entre les clients

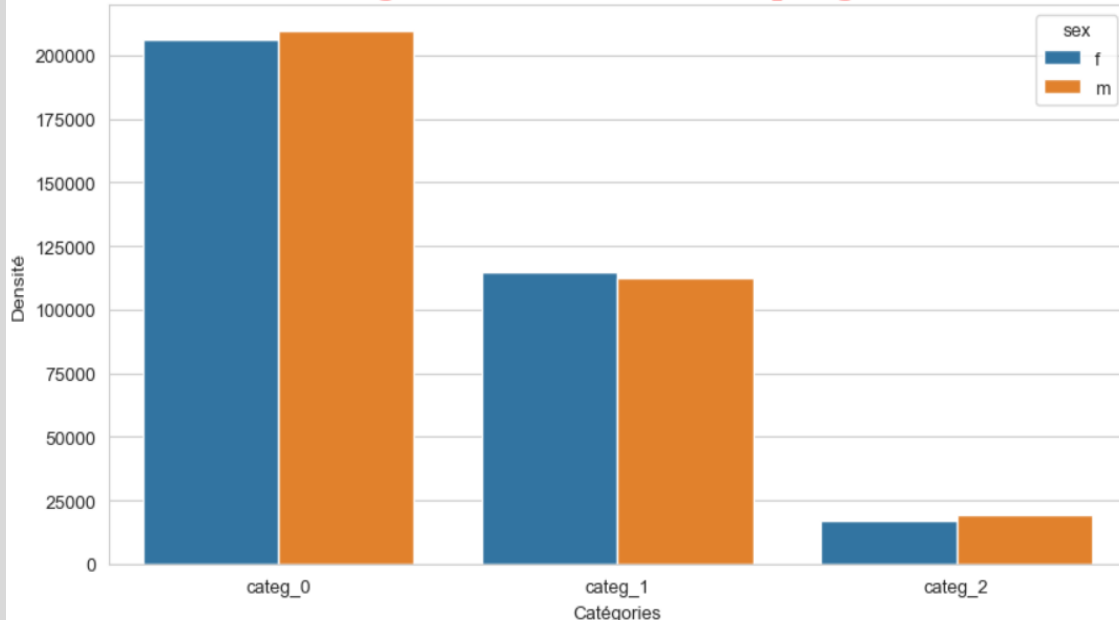


Incide de Gini : 0.403

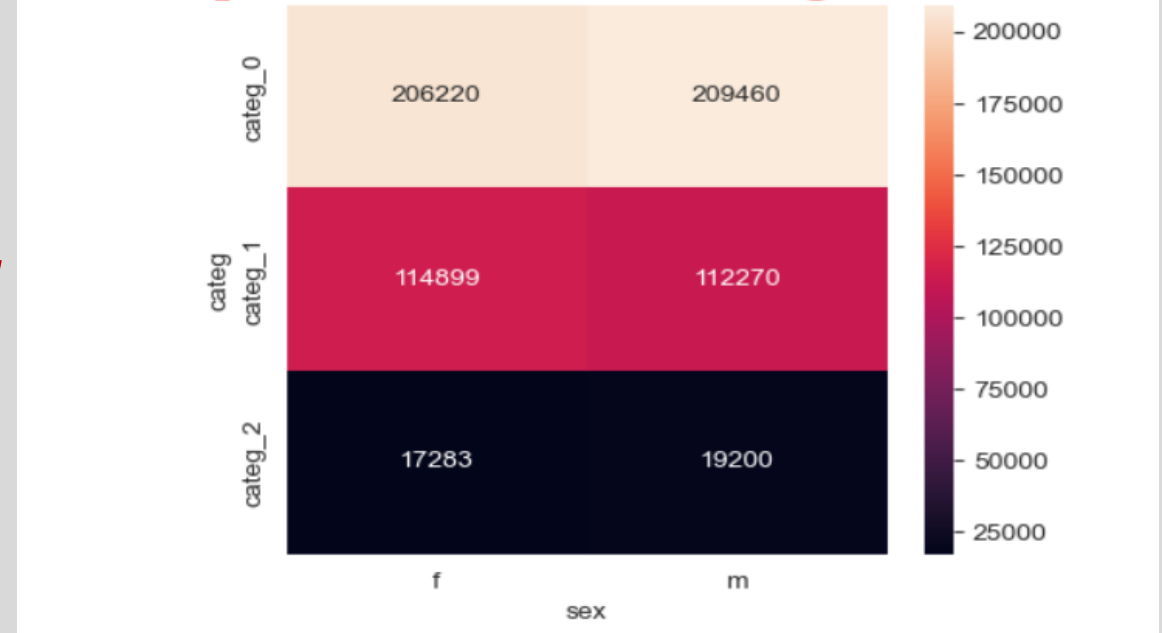
Demandes de Julie: Analyse des comportements de nos clients

Le lien entre le genre d'un client et les catégories de livres achetés

Catégories de livres achetés par genre

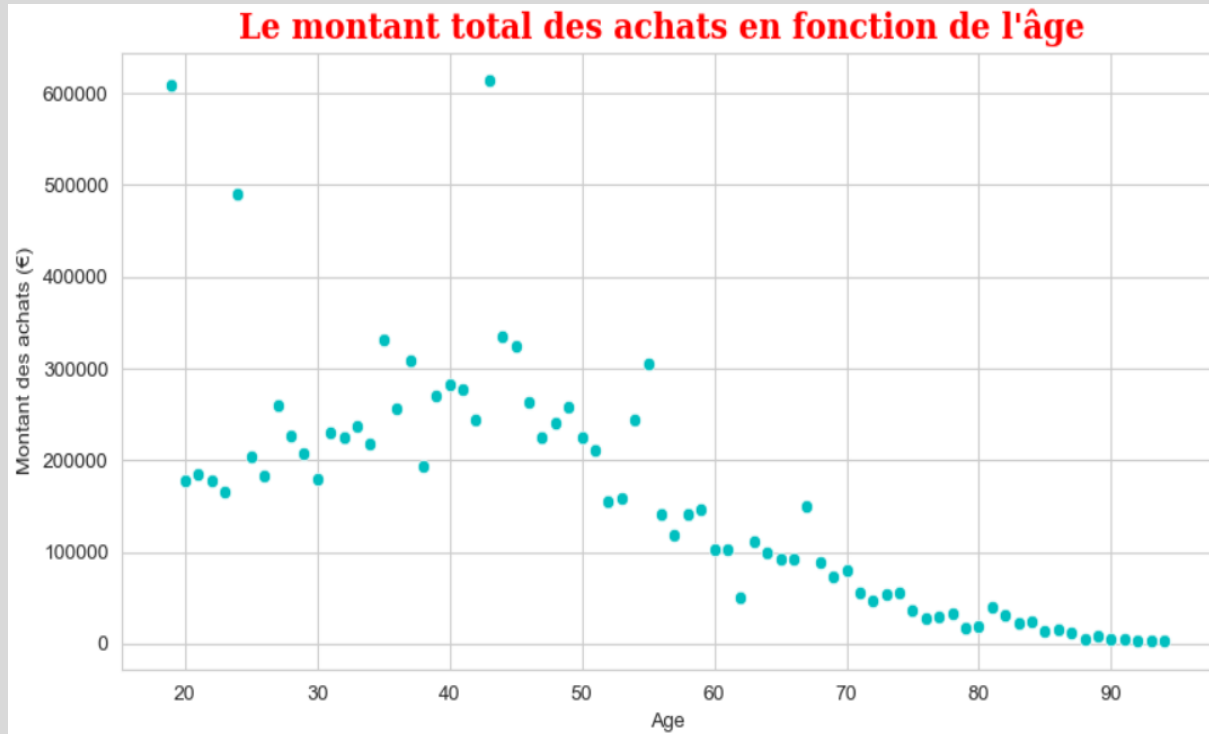


Heatmap entre le sexe et la catégorie de livres



Demandes de Julie: Analyse des comportements de nos clients

Le lien entre l'âge et le montant total des achats

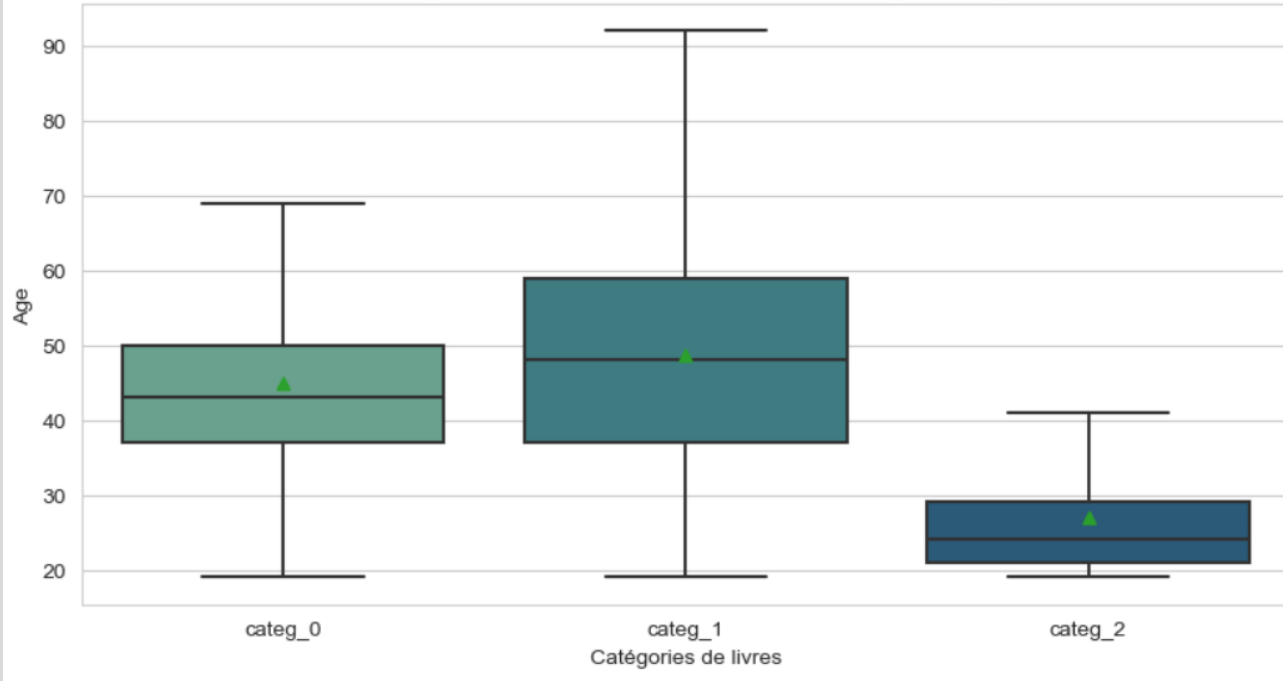


	W	pval	normal
birth	35269.636312	0.0	False
age	35269.636312	0.0	False
price	720071.616552	0.0	False
année	43850.680172	0.0	False
mois	800755.359323	0.0	False
jour_du_mois	587731.545221	0.0	False

Demandes de Julie: Analyse des comportements de nos clients

Le lien entre l'âge et les catégories de livres achetés

Catégories de livres achetés par âge



	W	pval	normal
birth	35269.636312	0.0	False
age	35269.636312	0.0	False

	W	pval	equal_var
levene	26091.73084	0.0	False

BILAN



- Données



- Explore et Analyse



- Réponses

