

Projet 4 - Analysez les ventes de votre entreprise

Projet 4 - Analysez les ventes de votre entreprise

1^{er} Partie – Nettoyage

Projet 4 - Analysez les ventes de votre entreprise

1^{er} Partie – Nettoyage

Visualisation des données

Tableau csv	Nombre de lignes	Valeurs Null	Commentaires
Customers	8 623	Aucune	Beaucoup de clients nés en 2004
Products	3 287	Aucune	Prix à '-1'
Transactions	337 016	Aucune	Date avec le mot Test

- Aucun valeur manquante mais des tests ont été effectués sur la base de données principale que nous allons nettoyer

Projet 4 - Analysez les ventes de votre entreprise

1^{er} Partie – Nettoyage

Visualisation des données

Tableau csv	id_product	session_id	client_id	Nombre d'individus
Customers	-	-	ct_0 ct_1	2
Products	t_0	-	-	1
Transactions	t_0	s_0	ct_0 ct_1	200

- 200 tests ont été effectués avec 1 produit - t_0, une session - s_0, et deux clients ct_1 et ct_2

Projet 4 - Analysez les ventes de votre entreprise

1^{er} Partie – Nettoyage

Visualisation des ventes en Octobre – 2/10/21 au 26/10/21

Catégorie	Septembre	Octobre	Novembre
Categ 0	23 112	18 751	14 578
Categ 1	9 268	1 666	12 316
Categ 2	874	1 160	1 373

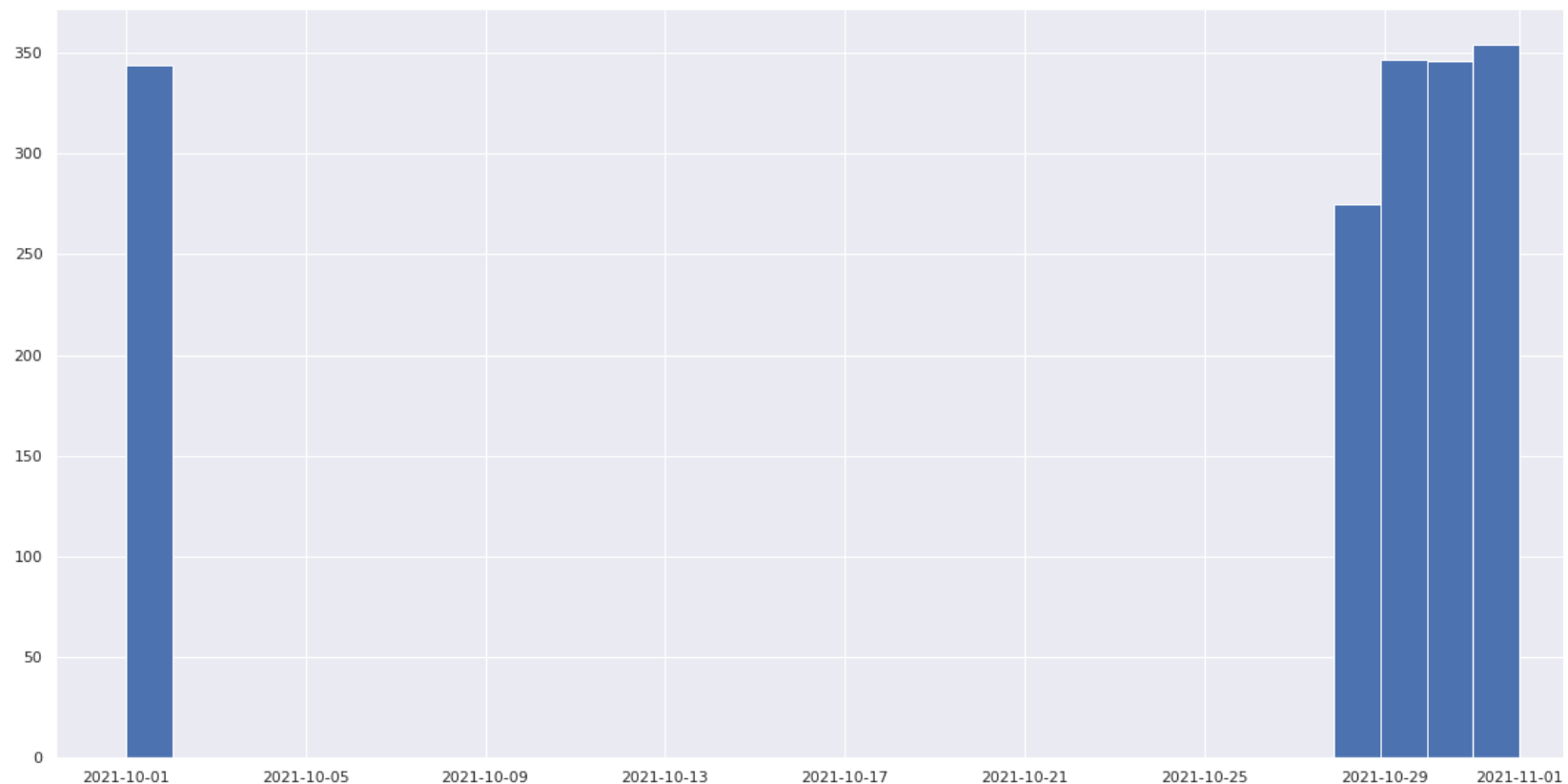
- La variation des ventes de la categ 1 au mois d'Octobre est très importante. Il n'y a aucune vente entre 2 Octobre 2021 et le 26 Octobre 2021.

Projet 4 - Analysez les ventes de votre entreprise

1^{er} Partie – Nettoyage

Visualisation des ventes en Octobre – 2/10/21 au 26/10/21

Vente de livres de categ 1 pendant le mois d'Octobre



Projet 4 - Analysez les ventes de votre entreprise

1^{er} Partie – Nettoyage

Création de la Data Frame

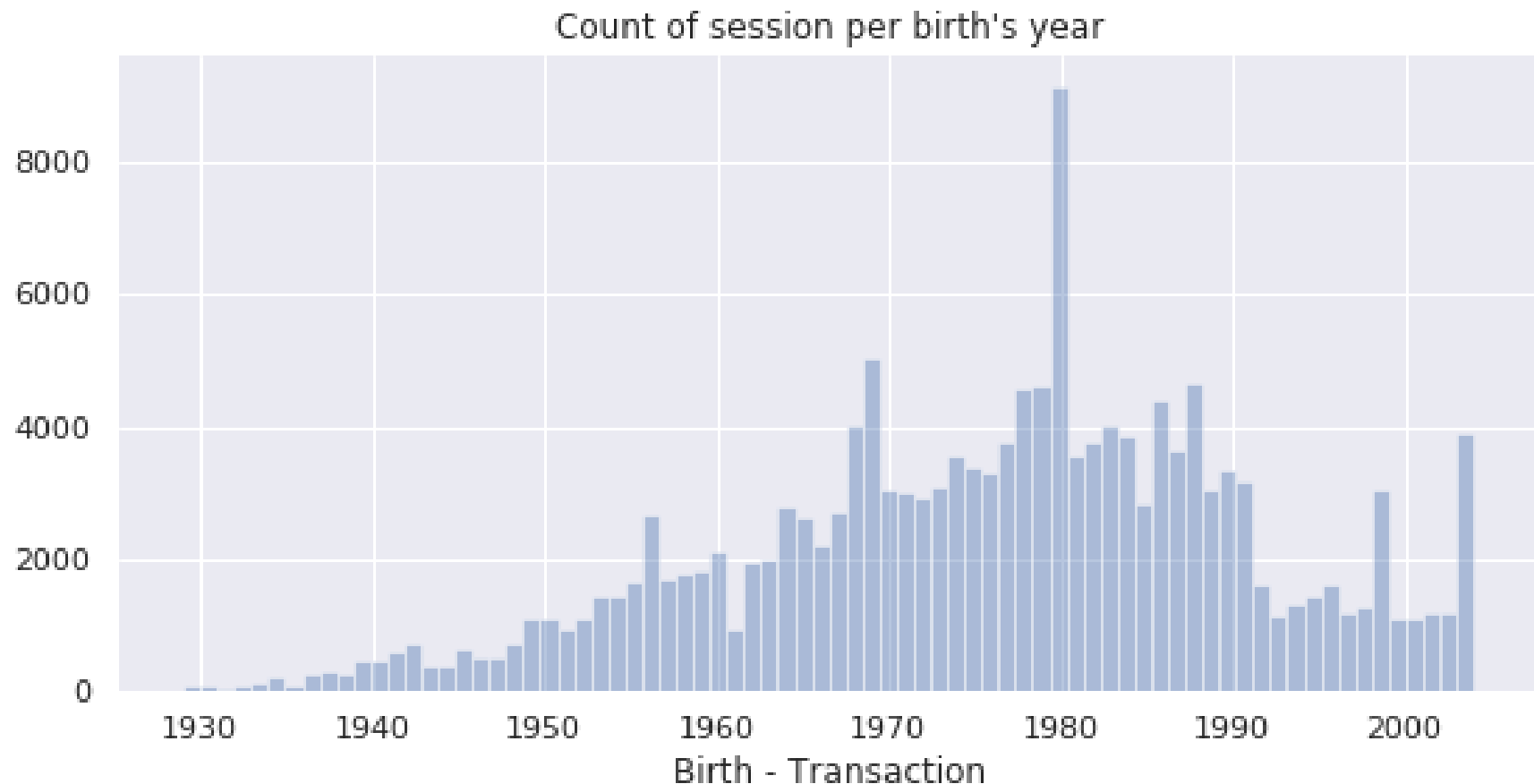
Colonne	Fichier csv Nombres	Data Frame Nombre	Différence	Commentaires
Customers	8 623	8 600	-23	2 tests et 21 non acheteurs
Products	3 287	3 264	-23	1 test et 22 non achetés
Transactions	337 016	336 713	-303	200 tests et 103 de 0_2245

- Il existe un produit dans le fichier Transactions qui n'est pas référencé dans le le fichier Products : 0_2245 avec 103 transactions

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

L'age



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

L'age

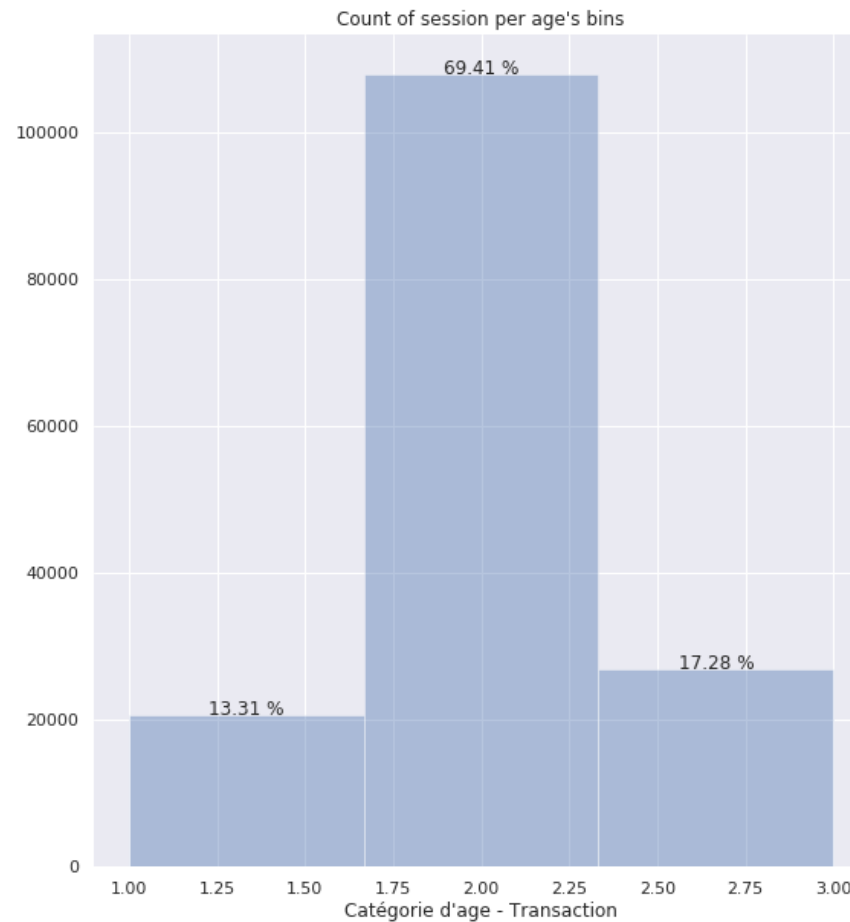
Age	Catégorie	Nombre	Pourcentage
18 à 30 ans	1	20 699	13,31 %
31 à 59 ans	2	107 953	69.41 %
60 à 99 ans	3	26 870	17.28 %

- Nous avons un grand nombre de transactions pour les personnes nées en 1980. Nous avons découvert qu'il existait 4 gros clients, qui seront abordés par la suite.

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

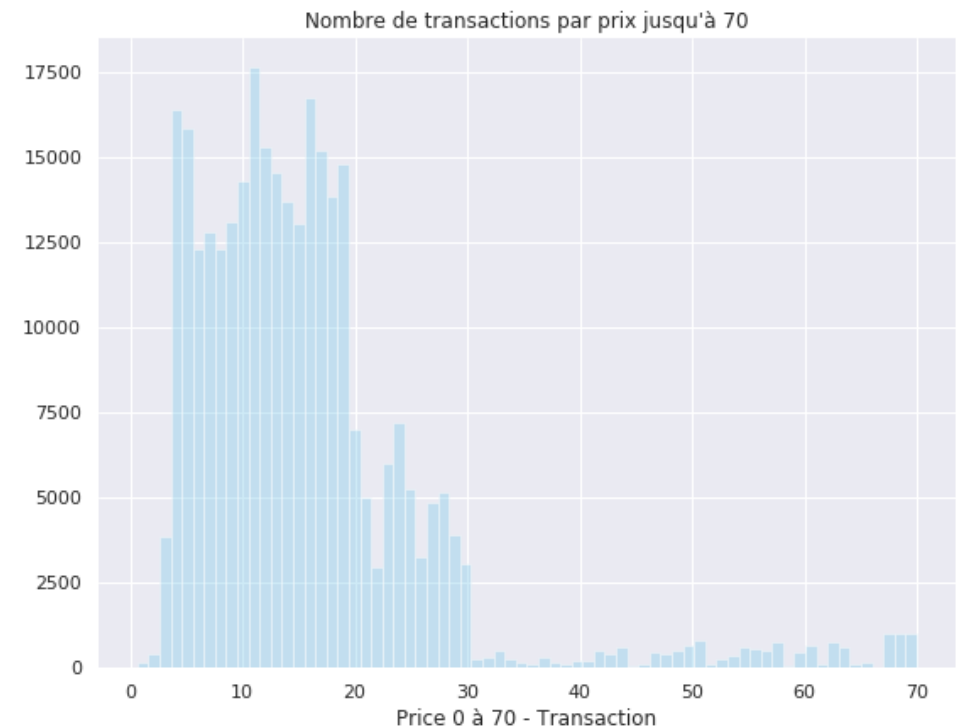
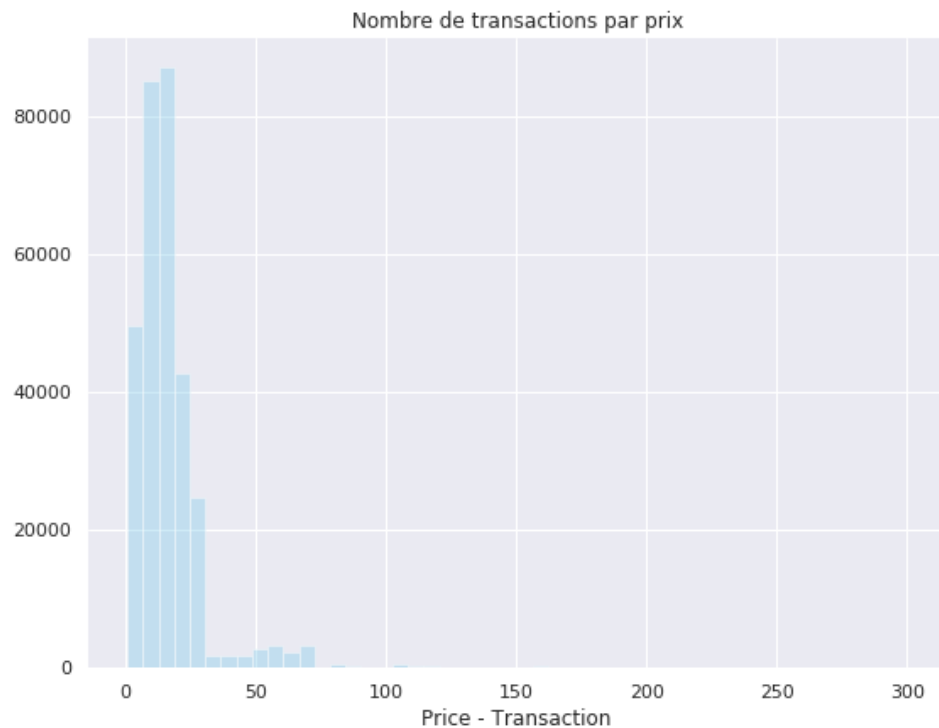
L'age



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les prix



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les prix

Prix	Catégorie	Nombre	Pourcentage
0 à 4,99	1	33 056	10,66 %
5 à 19,99	2	209 903	67,69 %
20 à 300	3	67 131	21,65 %

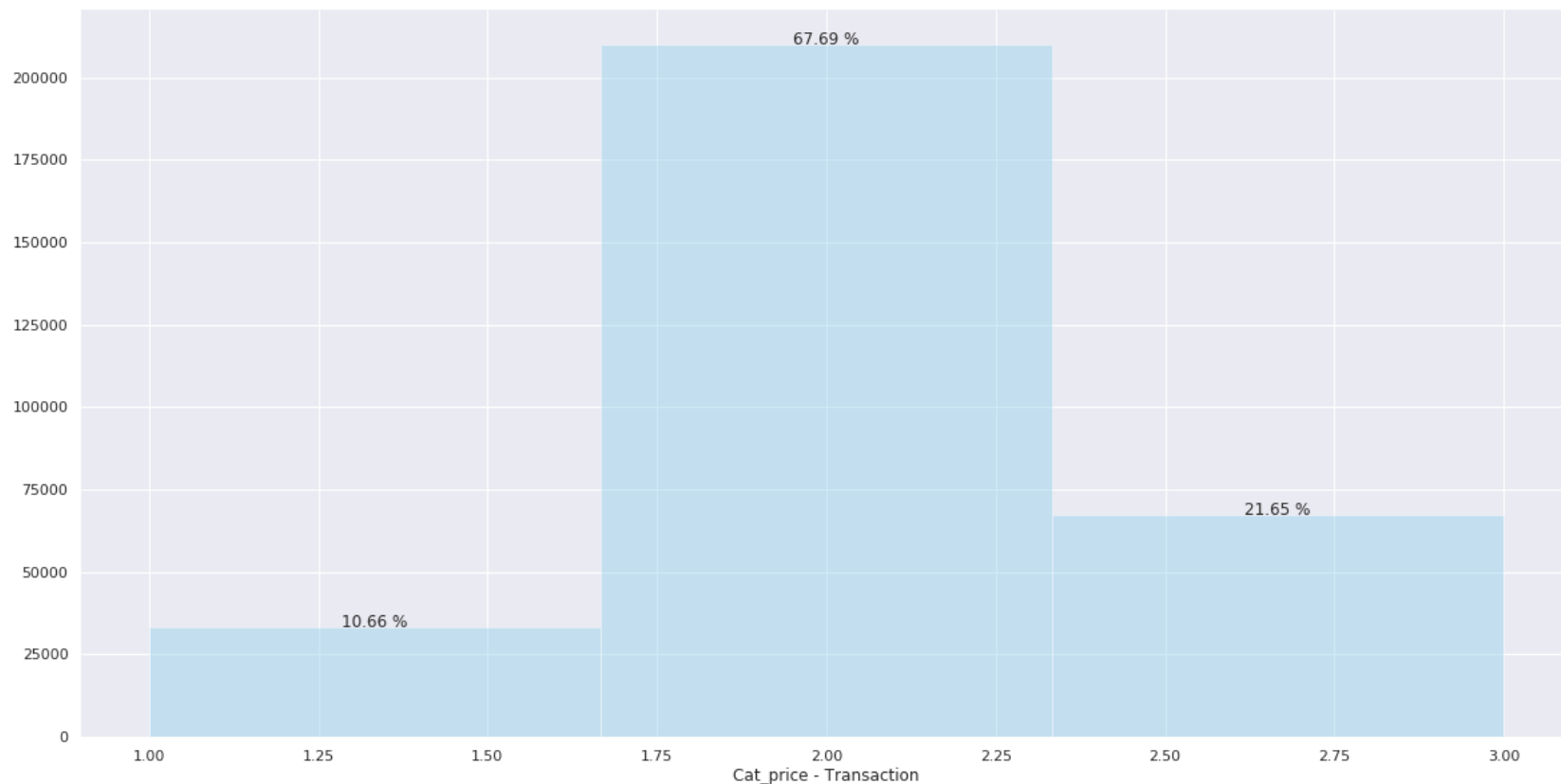
- Nous avons créer 3 catégories de prix en fonction des volumes de transactions.

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les prix

Nombre de transactions par catégorie de prix



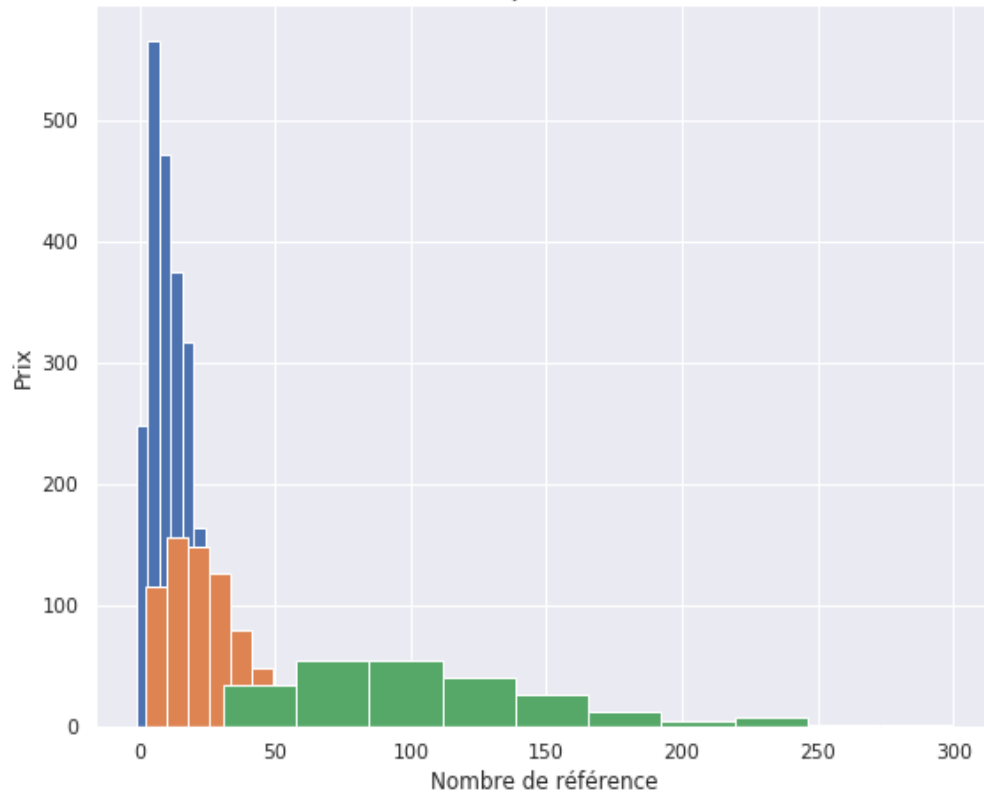
Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

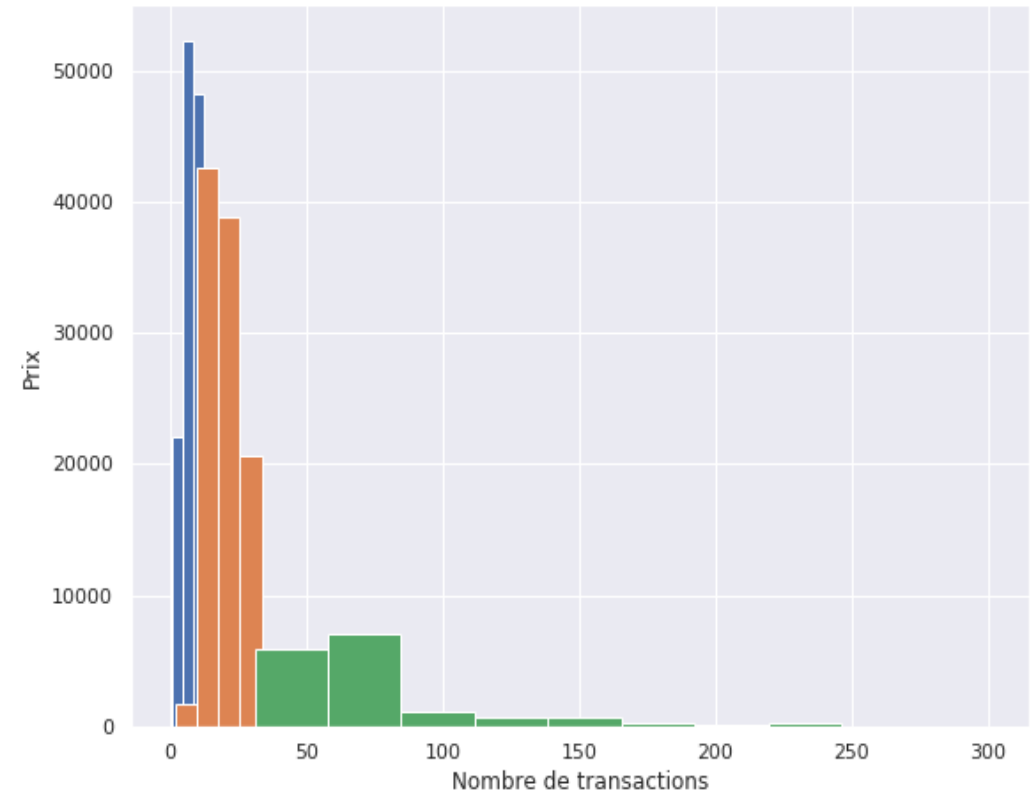
Les catégories de livres

Distributions des catégories de produit par prix

Fichier product.csv



Data Frame

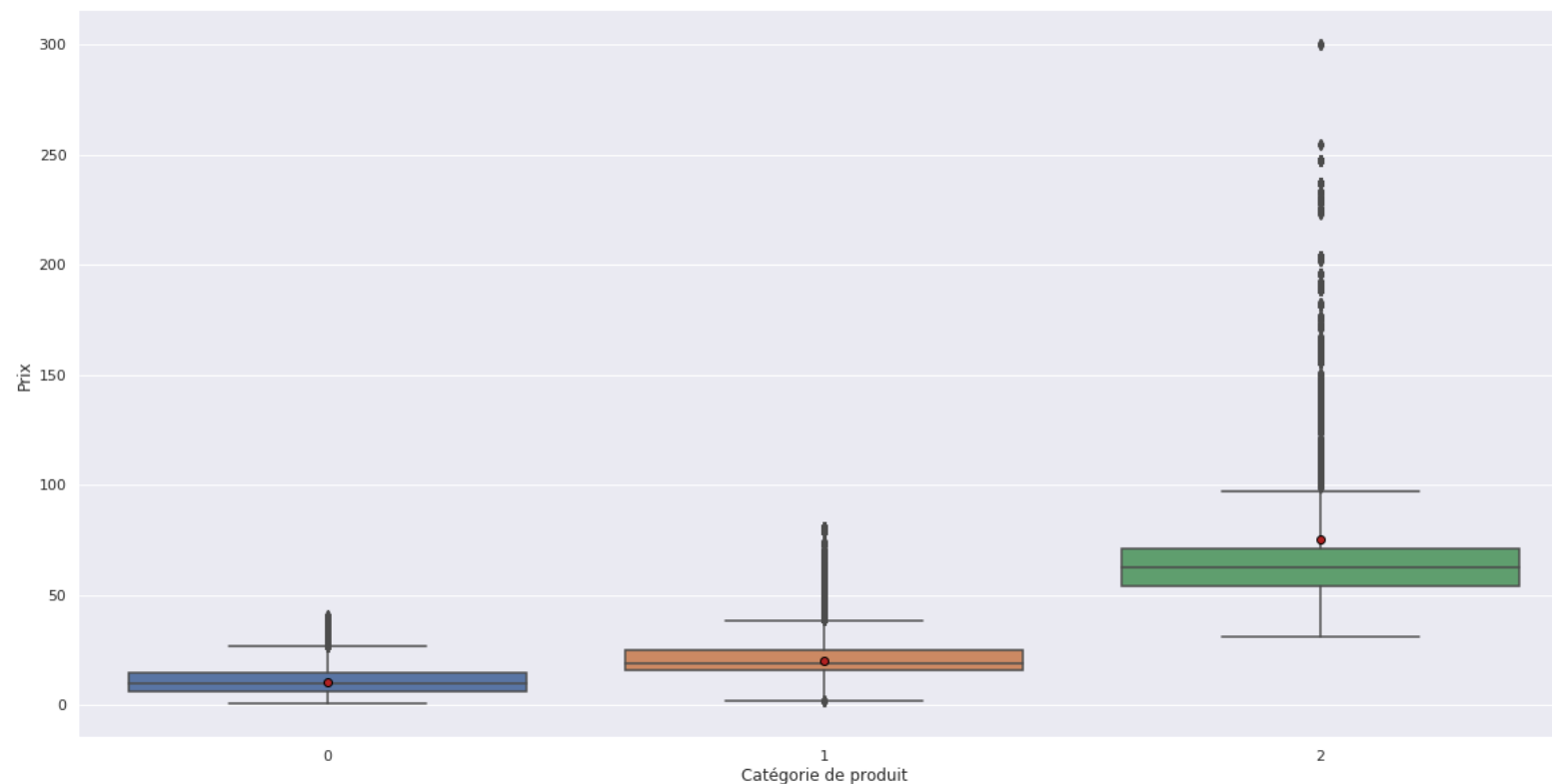


Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les catégories de livres

Répartition du prix en fonction de la catégorie du livre



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les catégories de livres

Catégorie	Moyenne Product.csv	Moyenne Data Frame	Ecart-type Product.csv	Ecart-Type Data Frame
0	11,73	10,65	7,57	4,93
1	25,53	20,48	15,42	7,57
2	108,36	75,23	49,56	39,18

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les catégories de livres

Type de livres	Catégorie	Nombre	Pourcentage
Livre de poche	0	187 017	60,31 %
Grand Format	1	106 768	34,43 %
Livre Encyclopédie	2	16 305	05,26 %

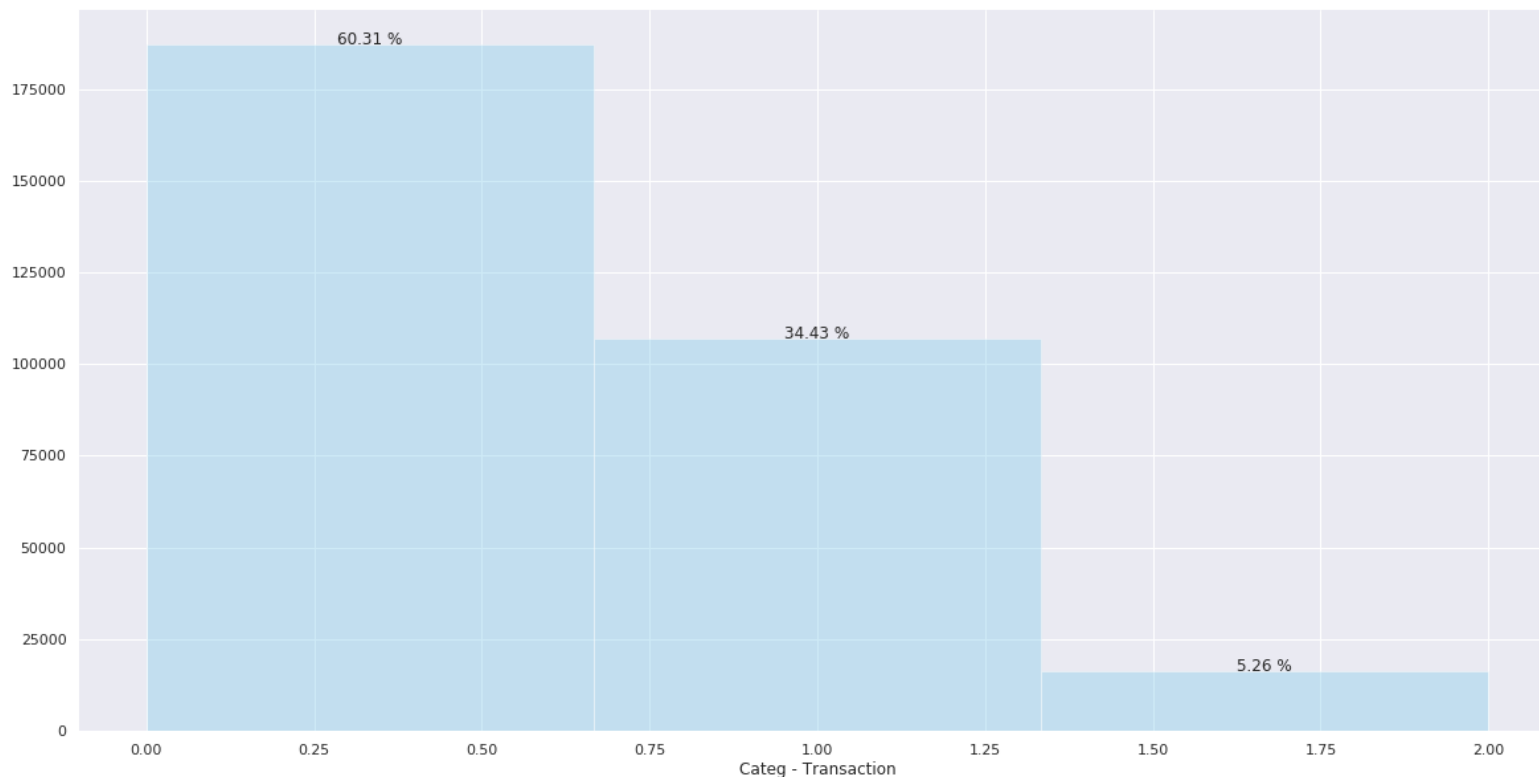
- Nous avons 3 Catégories de Livres réparties selon leur prix en fonction d'un format supposée : Livre de poche, Livre Grand Format et Livre type encyclopédie

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les catégories de livres

Nombre de transactions par catégorie de livre



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les Meilleurs Clients

Clients	Age	Montant Total	Transactions	Prix Livre Moyen
c_1609	41 ans	148 351	11 641	12,74
c_6714	53 ans	68 427	4 116	16,63
c_3454	52 ans	52 016	3 095	16,81
c_4958	22 ans	135 919	2 433	55,87

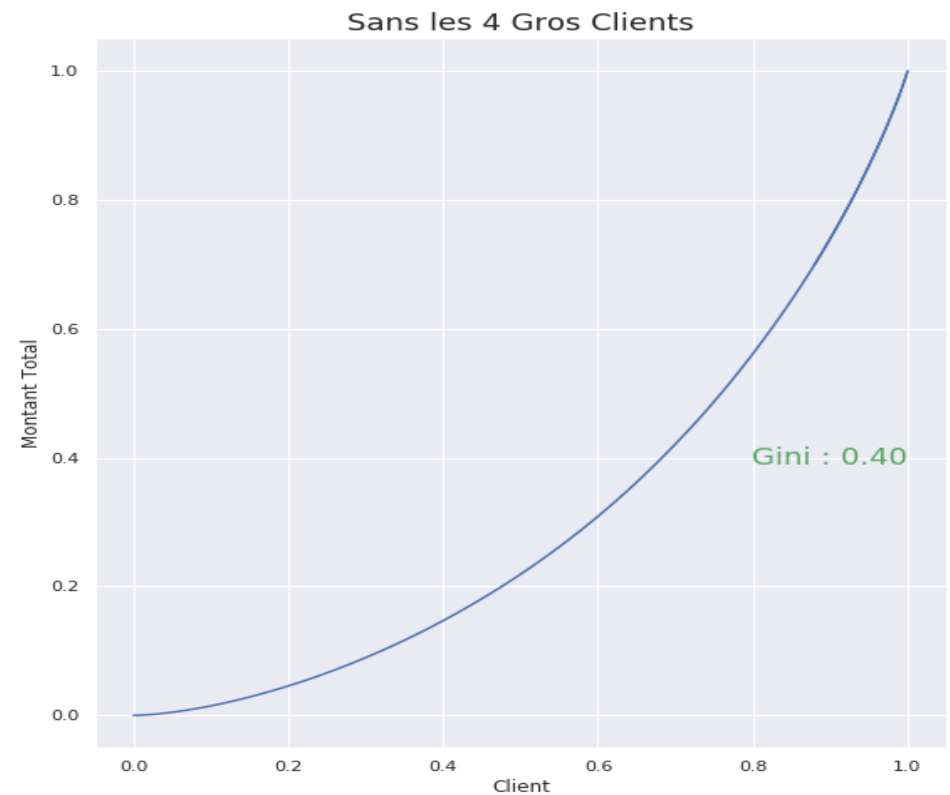
- Nous avons 4 Gros Clients – 3 dans la tranche d'âge 30-59 ans et 1 dans la tranche d'âge des moins de 30 ans. Ils représentent 7,49 % des montants de transactions

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

Les Meilleurs Clients

Courbe de Lorenz en fonction des montants d'achat des clients



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

La catégorie d'âge des moins de 30 ans – Sans les Gros Clients

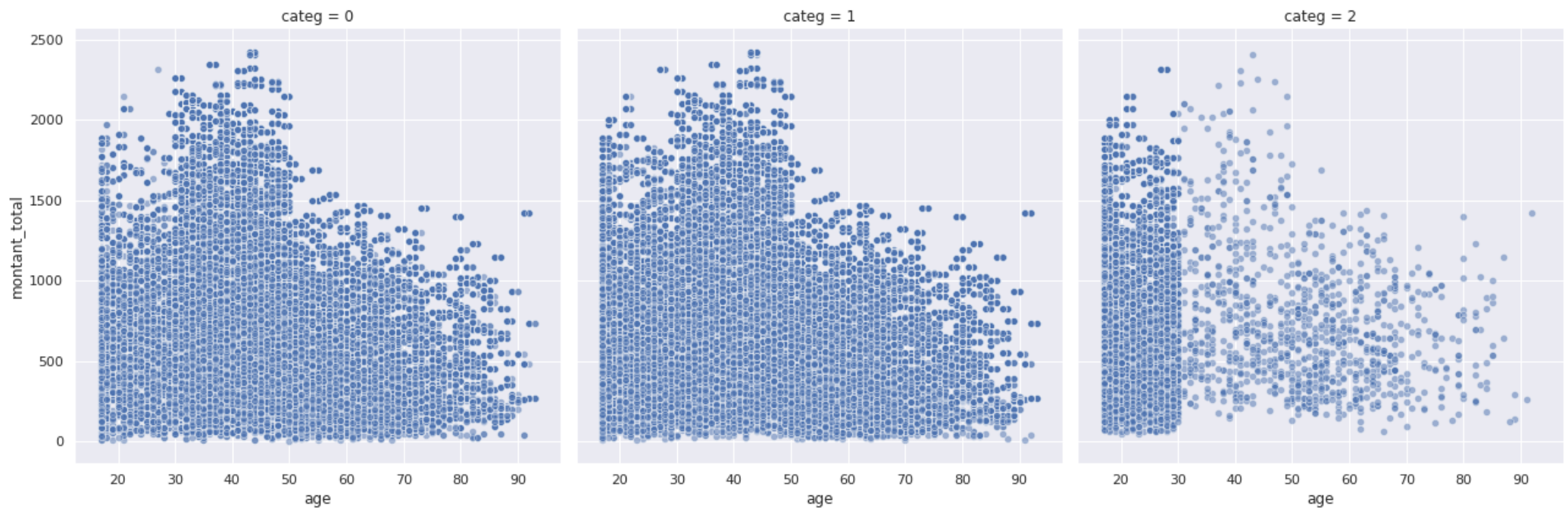
	Categ 0	Categ 1	Categ 2	Montant Total	Nombre Livre
Moyenne				600,13	15,04
Écart-type				405,18	9,22
Pourcentage	04,22 %	12,99 %	90,64 %	26,82 %	11,64 %

- Les moins de 30 ans sont les plus gros acheteurs de livres de catégorie 2, les plus chers. Il ne représente moins d'un quart des achats, en ayant un profil de client identique que la classe d'âge intermédiaire, les plus représentée.

Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

La catégorie d'âge des moins de 30 ans – Sans les Gros Clients



Projet 4 - Analysez les ventes de votre entreprise

2^{ème} Partie – Visualisation des données

La catégorie d'âge intermédiaire – Sans les Gros Clients

	Categ 0	Categ 1	Categ 2	Montant Total	Nombre Livre
Moyenne				626,44	45,08
Écart-type				458,96	35,10
Pourcentage	87,08 %	66,12 %	07,22 %	59,60 %	74,71 %

- La classe d'âge intermédiaire est la plus représentée. Par contre, il représente la majorité des achats pour les livres de catégorie 0 et 1, les moins chers.

Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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



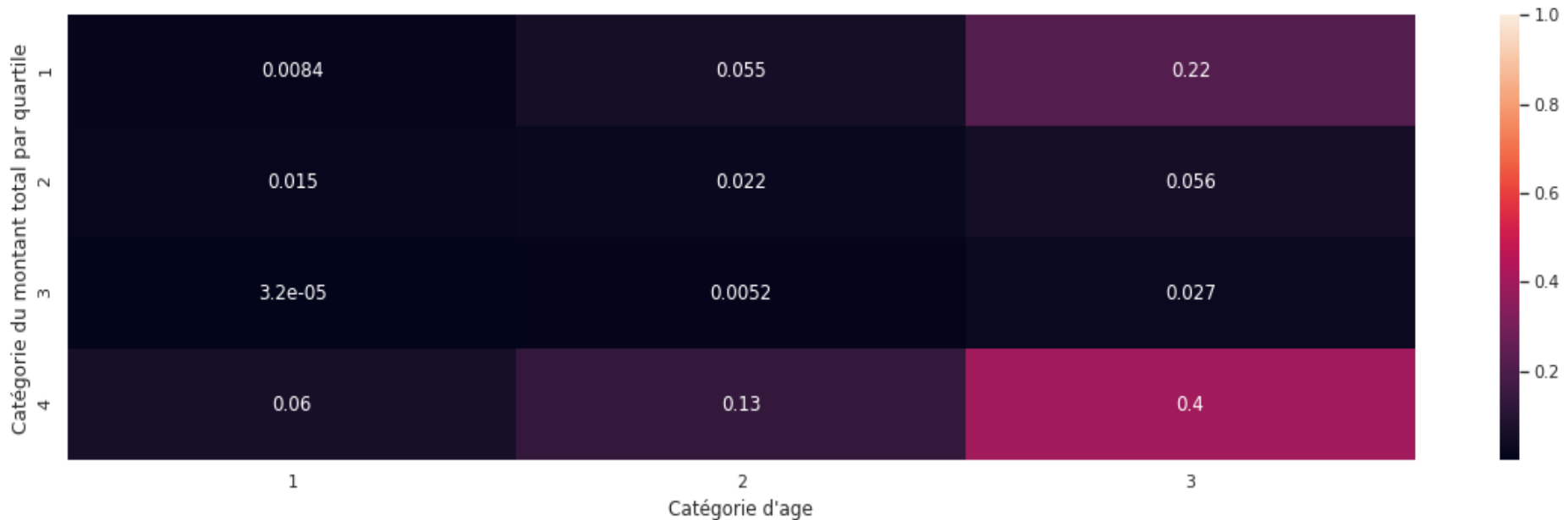
- Nous voyons qu'il n'existe pas de corrélation entre le sexe du client et la catégorie de produit.

Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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

Tableau de contingence entre la catégorie du montant total par quartile et la catégorie d'âge



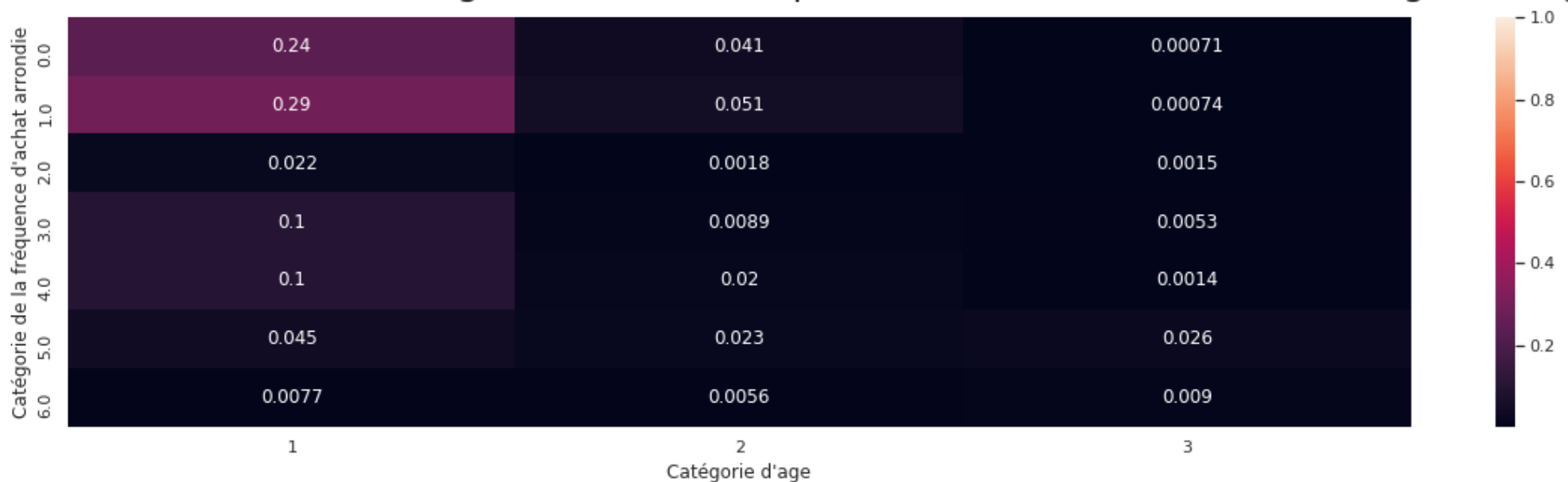
- Il n'y a pas de corrélation franche entre la catégorie du montant d'achat et la catégorie d'âge.

Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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

Tableau de contingence entre la fréquence d'achat arrondie et la catégorie d'âge



- Nous voyons qu'il n'existe pas de corrélation entre la fréquence d'achat et l'âge même si les consommateurs de moins de 30 ans ont une fréquence d'achat en général plus faible que les autres.

Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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

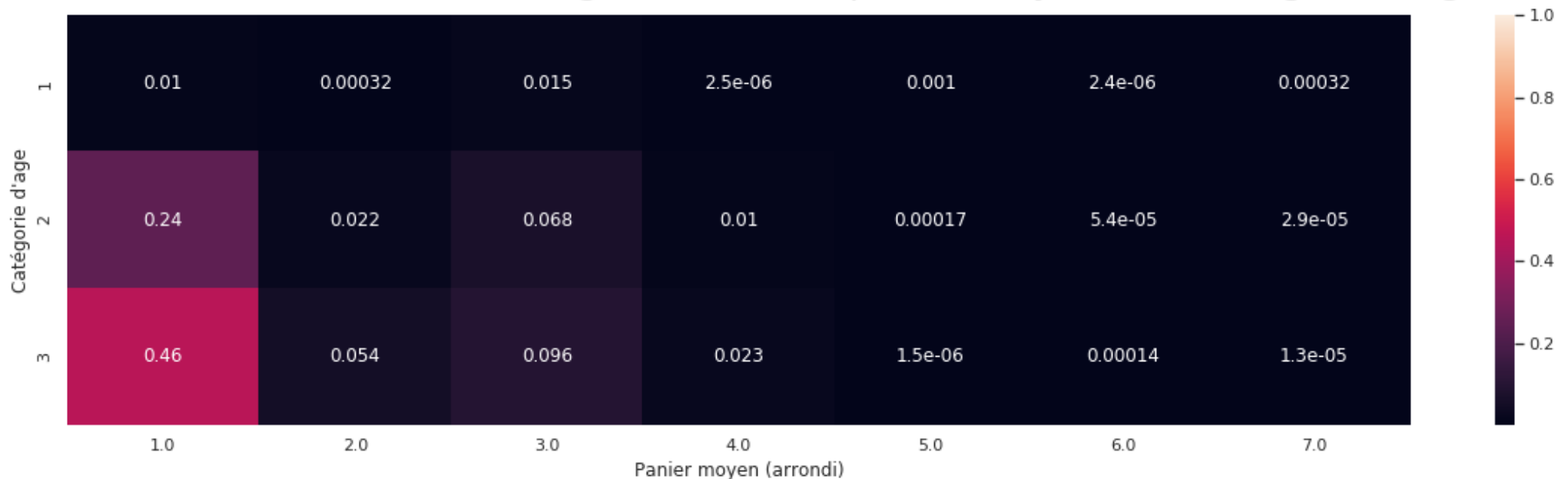


Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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

Tableau de contingence entre le panier moyen et la catégorie d'age

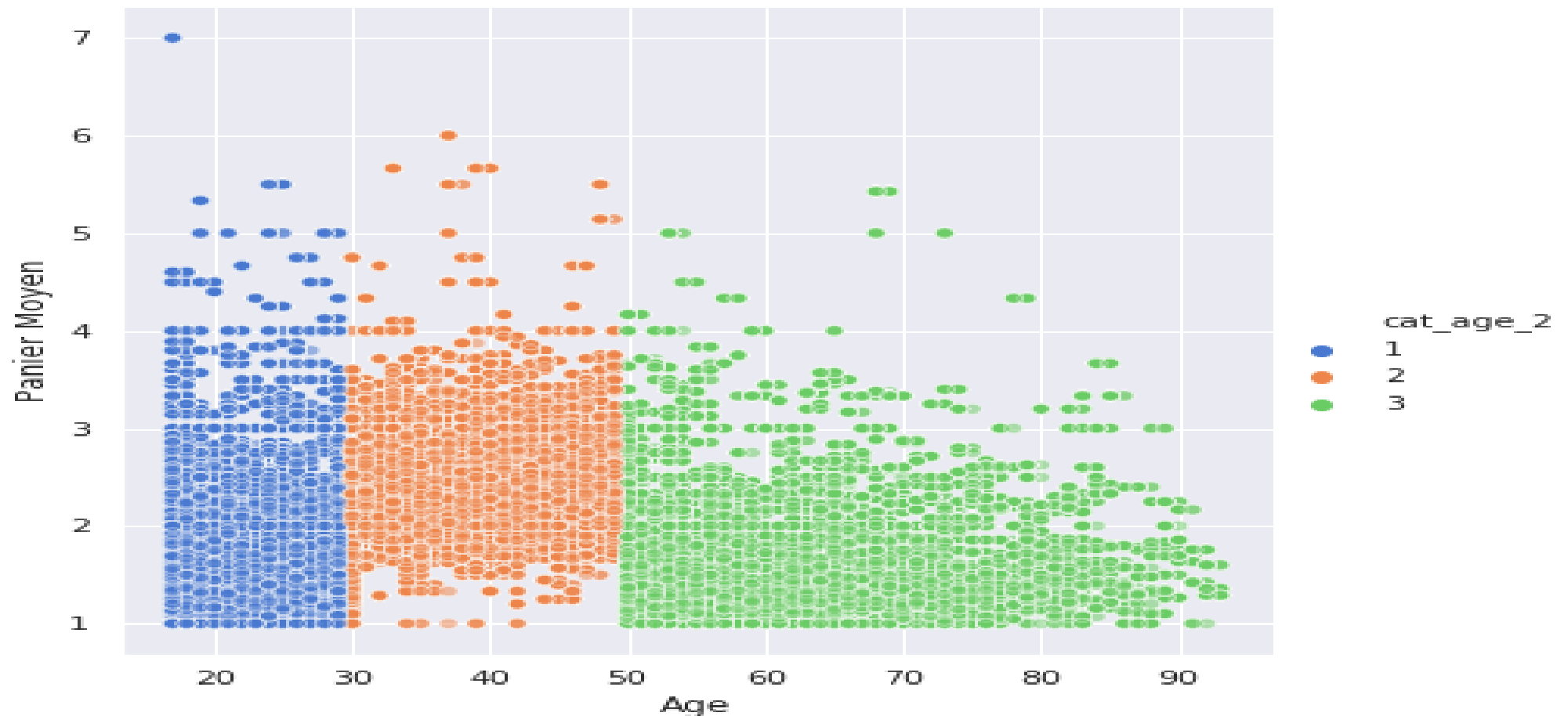


- Il semble exister une corrélation positive entre le nombre d'achat moyen arrondi et les personnes de plus de 50 ans. Ils sont nombreux à n'acheter en moyenne entre 1 et 2 articles par session.

Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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

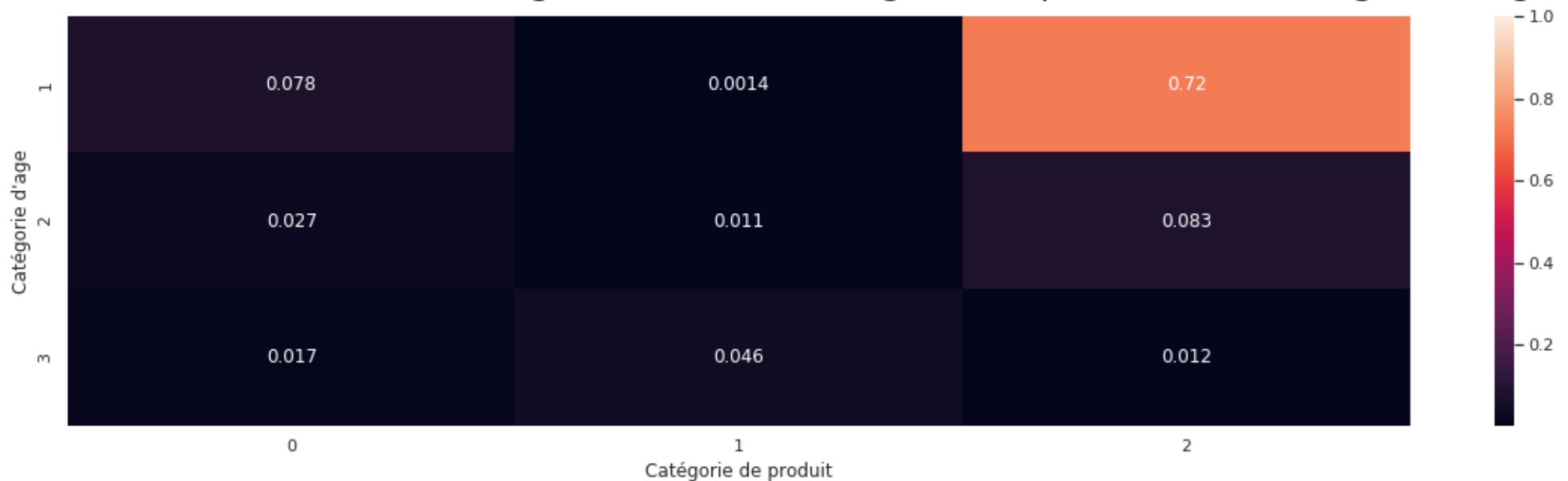


Projet 4 - Analysez les ventes de votre entreprise

3^{ème} Partie – Corrélation

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

Tableau de contingence entre la catégorie de produit et la catégorie d'âge



- Nous montrons qu'il existe une corrélation positive entre la catégorie d'âge des moins de 30 ans et la catégorie de livre de categ 2