Analyser les ventes d’une librairie avec R ou Python

**Analyses des Indicateurs de ventes**

1 - Le CA du mois d’octobre 2021 dénote clairement donc je suis rentré dans le détail en analysant ainsi le CA par catégorie

2 - Le CA pour la catégorie 1 est très faible comparé aux autres mois donc j’ai continuer mon analyse avec le CA réaliser pour chaque jour du mois d’octobre 2021 pour la catégorie 1

3 - J’en ai conclu qu’il y a eu un soucis technique et que les chiffres ne sont pas remontés dans la base de données, ce qui explique ce grand manque de données.

4 - La moyenne mobile en vert montre une tendance du chiffre à la baisse après avoir connu un pic au début 2022.

5 - Les produits qui générent le plus de CA ne sont pas les mêmes que ceux qui sont le plus vendus. Les produits vendus en plus grande quantité ne sont donc pas forcément les plus chers.

(peu d’informations sont exploitables avec les flops, a priori plusieurs références n’ont été vendues qu’une seule fois.

6 - Une majorité de produits n’est vendu qu’un assez nombre de faible de fois. (dans un objectif d’optimisation du nombre de référence disponible, il pourrait être intéressant de pousser ces analyses sur ce point).

7 - Les catégories 0 et 1 générent un CA assez comparable tandis que la categ 2 est inférieure. La répartition des quantités est par contre très inégale, on peut distinguer que les prix des produits des différentes catégories sont très inégaux.

8 - La catégorie 0 est composée de produits autour de 10€,

La catégorie 1 est composée de produits autour de 20€,

La catégorie 2 est composée de produits bien plus chers.

**Analyses des clients**

1 - La distribution de l’âge des clients est assez similaire pour les femmes et les hommes. Dans les 2 cas, les clients les nombreux sont les jeunes puis ceux ayant un peu plus de 40 ans

2 - 50 % des clients générent 20% du CA. A l’extrémité du graphique on peut observer une zone où la courbe est complètement droite, il s’agit des clients professionnels

3 - Le samedi en fin d’apres midi est une période assez importante pour les ventes. Il y a un pic d’activité un peu plus étonnant le mardi à 23h, peut être les commandes des clients pro?

4 - top 10 meilleurs clients, (présenter)

5 - top 10 clients les moins dépensiers (présenter)

6 - clients atypiques (présenter)

7 - L’écart-type des z-scores parmi les clients atypiques est d’environ 19.75 , ce qui montre qu’il y a une grande diversité même parmi les cas extrêmes. Certains clients sont donc non seulement atypiques, mais carrément hors normes.

**Analyses des corrélations**

1 - Dans cette dernière partie, l’objectif est de vérifier l’existence de corrélations entre le profil des clients et leurs comportements

Le genre des clients est composé de 2 groupes et la catégorie est composée de 3 groupes, le test de Chi² a donc été réalisé pour commencer les différents test.

L’hypothèse H0 a été rejetée, cela signifie que le genre d’un client influe sur les catégories de livres qu’il achète.

3 - (age,turnover etc..) matrice de corrélation

En ce qui concerne l’âge, il est a priori corrélé avec chacunes des variables, notamment avec le panier moyen.

Plus le client est agé, moins son panier est élévé.

4 - (relation entre l’age et les variables)

Pour affiner notre compréhension des relations, on affiche le diagramme de dispersion de chaque variables avec l’âge.

Les graphiques confirment que les résultats de la matrice de corrélations et notamment le fait que le lien entre l’âge et le panier moyen soit le plus marqué.

5 - J’ai donc suivi l’analyse en réalisant un test de pearson

le test d’indépendance est toujours rejeté

j’en déduis que l’âge est corrélé avec chacunes des variables donc qu’il va y avoir une influence sur le comportement des clients.

6 - (répartition des âges par catégorie, boite à moustaches)

Pour pousser un peu plus loin mes analyses j’ai réaliser 3 autres différents test,

test de Bartlett, Kruskal-Wallis et Anova

les résultats des différents test m’ont permis de conclure que l’age impacte les catégories de livres achetés.

FIN