



PROJET DE VISUALISATION DE DONNEES AVEC POWER BI

Analyse des retours produits via la visualisation



Nous sommes des étudiants en Master 1 de Data Science à l'université à Félix Houphouët Boigny de Cocody. Dans le cadre du projet, nous avons analysé les données d'une entreprise de vente de produits électroniques afin de comprendre les raisons derrière les retours produits.

Noms et Prénoms des membres :

KOFFI Kouamé Jean-Baptiste
KOUASSI Kouadio Prosper
YAO Ayéménés Marcel

Problématique :

Quels sont les facteurs qui influencent le retour des produits par les clients ?

Contexte

Une entreprise de vente de produits électroniques en ligne souhaite comprendre les raisons derrière les retours produits. Grâce à un jeu de données transactionnel contenant plus de 10 000 lignes, vous êtes chargé(e) de mener une analyse exploratoire et visuelle pour identifier les facteurs influençant les retours clients.

Jeu de données : synthetic_sales_data.csv

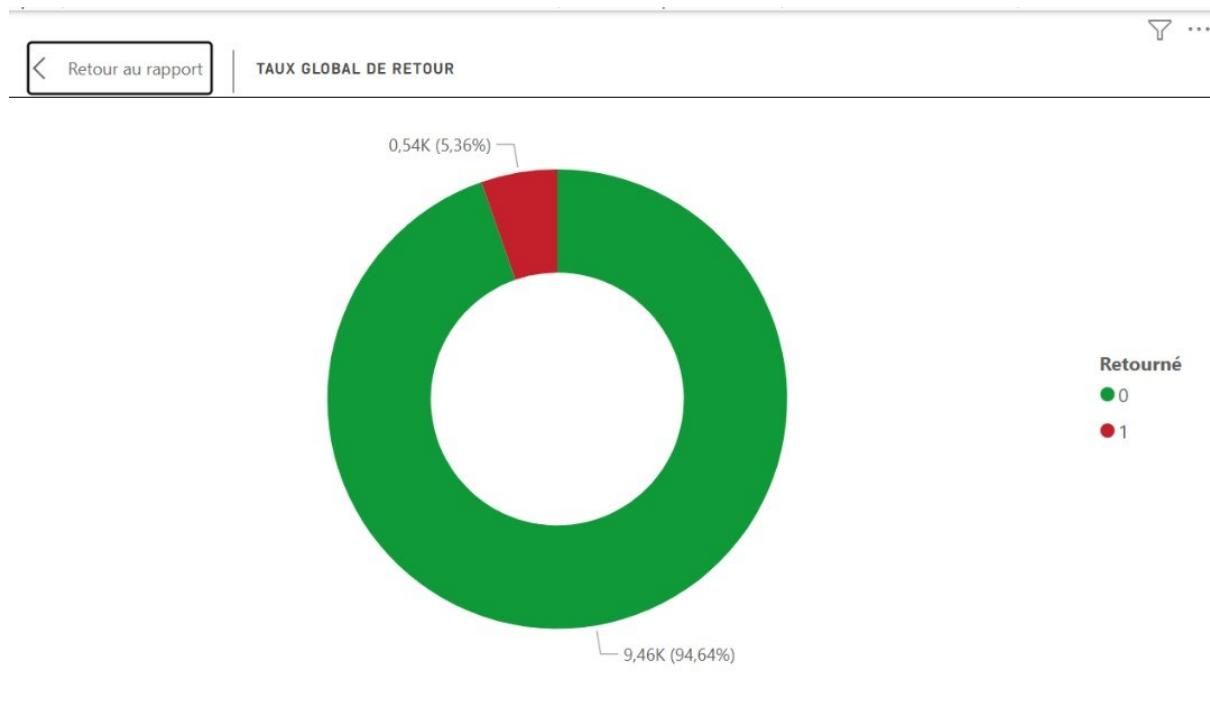
Chaque ligne représente une transaction. Les colonnes principales sont :

- TransactionID, CustomerID
- Product, Category
- PurchaseDate
- Price, Quantity
- PaymentMethod
- CustomerRegion
- Returned (binaire : 1 = produit retourné, 0 = conservé)
- DeliveryDays : délai de livraison
- SatisfactionRating : note de satisfaction client (1 à 5)

RESULTATS DES ANALYSES

PARTIE 1 : EXPLORATION INITIALE

1. Taux global de retour



Analyse du graphique – Taux global de retour

Ce graphique en anneau illustre la répartition des produits retournés par rapport à l'ensemble des produits vendus.

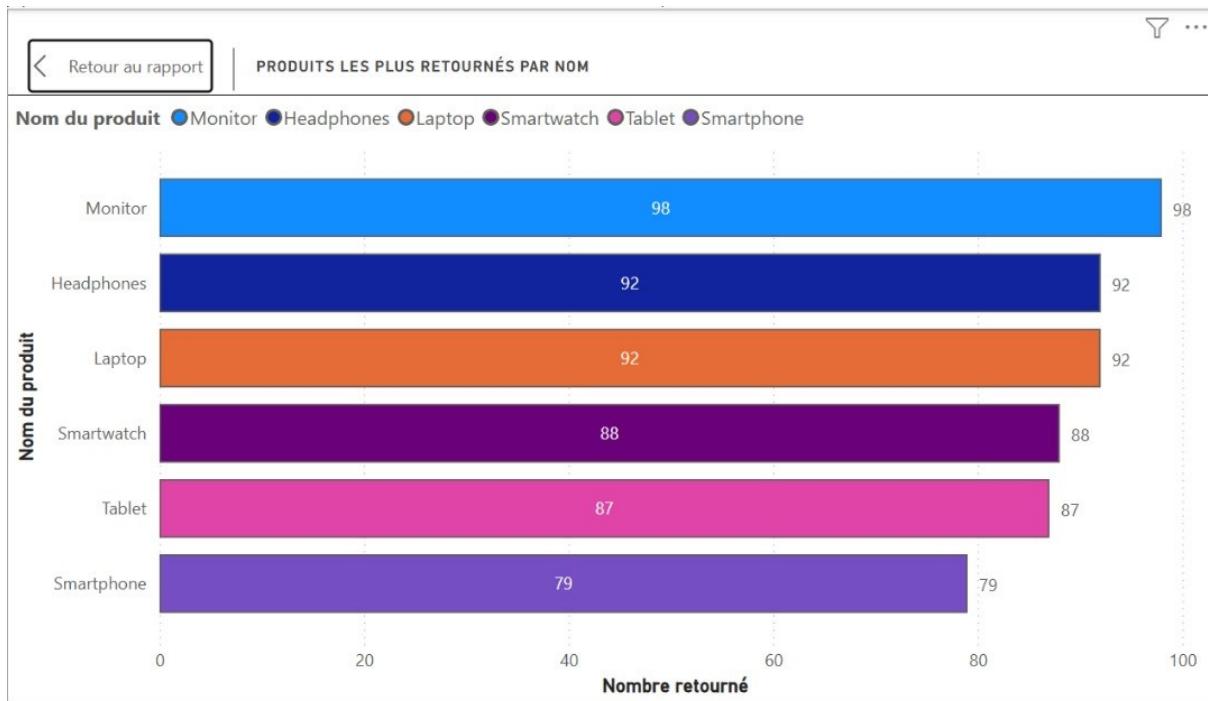
On observe que **94,64 %** des commandes **n'ont pas été retournées** (9 464 unités), tandis que seulement **5,36 %** ont été **retournées** (536 unités).

Interprétation : Ce faible taux de retour constitue un **indicateur positif de la satisfaction client et de la qualité du produit**. En effet, un taux de retour inférieur à 10 % est généralement considéré comme performant, selon les standards du commerce de détail. Cela suggère que la majorité des clients conservent le produit après achat, ce qui témoigne probablement de :

- La conformité du produit aux attentes,
- Une bonne qualité perçue,
- Un processus de vente efficace.

2. Produits les plus retournés

2.1 Produits les plus retournés par nom de produit



Analyse du graphique Produits les plus retournés par nom :

1. Lecture du graphique :

Ce graphique à barres présente le nombre total de retours pour chaque type de produit électronique vendu par l'entreprise. Sur l'axe vertical, on retrouve les noms des produits ; sur l'axe horizontal, le nombre total de produits retournés par les clients.

Points clés :

1. Produits les plus retournés

- **Monitor** (écran d'ordinateur) est le produit le plus retourné, avec presque 100 retours.
- **Headphones** (casques audio) et **Laptop** (ordinateurs portables) arrivent ensuite, avec 92 retours chacun.
- **Smartwatch** (montres connectées) : 88 retours.
- **Tablet** (tablettes) : 87 retours.
- **Smartphone** : 79 retours.

2. Part relative

- À titre d'exemple, les Headphones représentent **92** retours, soit **17,16 %** de l'ensemble des produits retournés.

Interprétation / pistes d'analyse

- **Monitor en tête** : Les écrans sont les plus retournés. Cela peut s'expliquer par des problèmes de compatibilité, de taille, de défauts à la livraison, ou des attentes non satisfaites.
- **Casques audio et laptops** : Ces produits sont également fortement retournés. Cela peut indiquer des problèmes de qualité, de performance ou de satisfaction client sur ces catégories.
- **Smartphones et tablettes** : Sont un peu moins concernés par les retours, mais restent tout de même significatifs.

Homogénéité

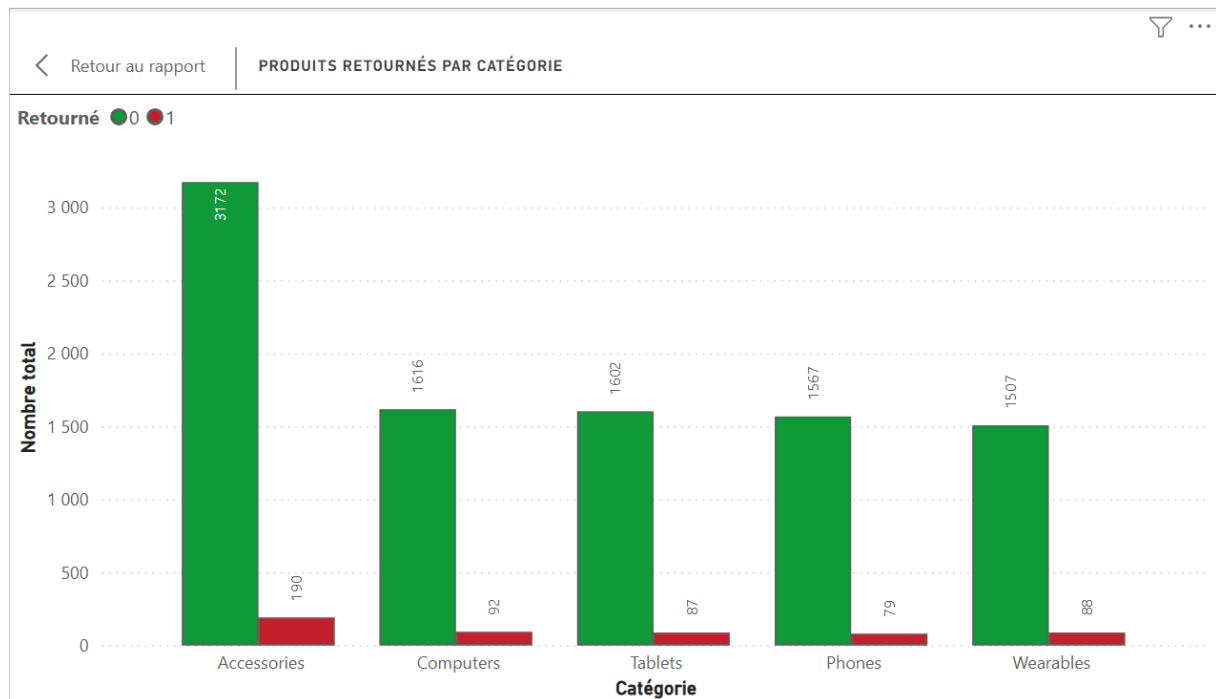
Aucun produit ne domine outrageusement. L'ensemble des familles de produits électroniques connaît un taux de retour important, ce qui peut révéler un enjeu global de satisfaction ou de qualité.

En résumé :

Les produits les plus retournés sont les écrans, suivis des casques audios et des ordinateurs portables. Cela suggère des axes d'amélioration possibles sur la qualité, le descriptif produit, ou le service après-vente pour ces catégories.

2.2 Produits les plus retournés Par catégorie :

Nous avons utilisé un graphique en barres groupées pour comparer le nombre de retours par catégorie.



Lecture du graphique

- Ce graphique présente, pour chaque catégorie de produits (Accessories, Computers, Tablets, Phones, Wearables), le nombre total de commandes non retournées (barre verte) et le nombre de commandes retournées (barre rouge).
- L'axe vertical indique le nombre total.
- L'indication « Retourné 0/1 » permet de distinguer les produits conservés (0) des retournés (1).

Points clés

1. Accessoires :

- C'est la catégorie la plus vendue, avec 3 362 ventes.
- Elle enregistre aussi le plus grand nombre de retours : 190.
- Cela peut s'expliquer par le fait que les accessoires (câbles, coques, écouteurs, etc.) sont souvent achetés en complément, mais aussi plus susceptibles d'être retournés (problèmes de compatibilité, attentes non satisfaites...).

2. Computers, Tablets, Phones, Wearables :

- Ces catégories présentent des volumes de ventes similaires (entre 1 500 et 1 650).
- Le nombre de retours est également assez proche (entre 80 et 95).
- Cela montre une certaine homogénéité dans le comportement de retour pour ces familles de produits.

Interprétation

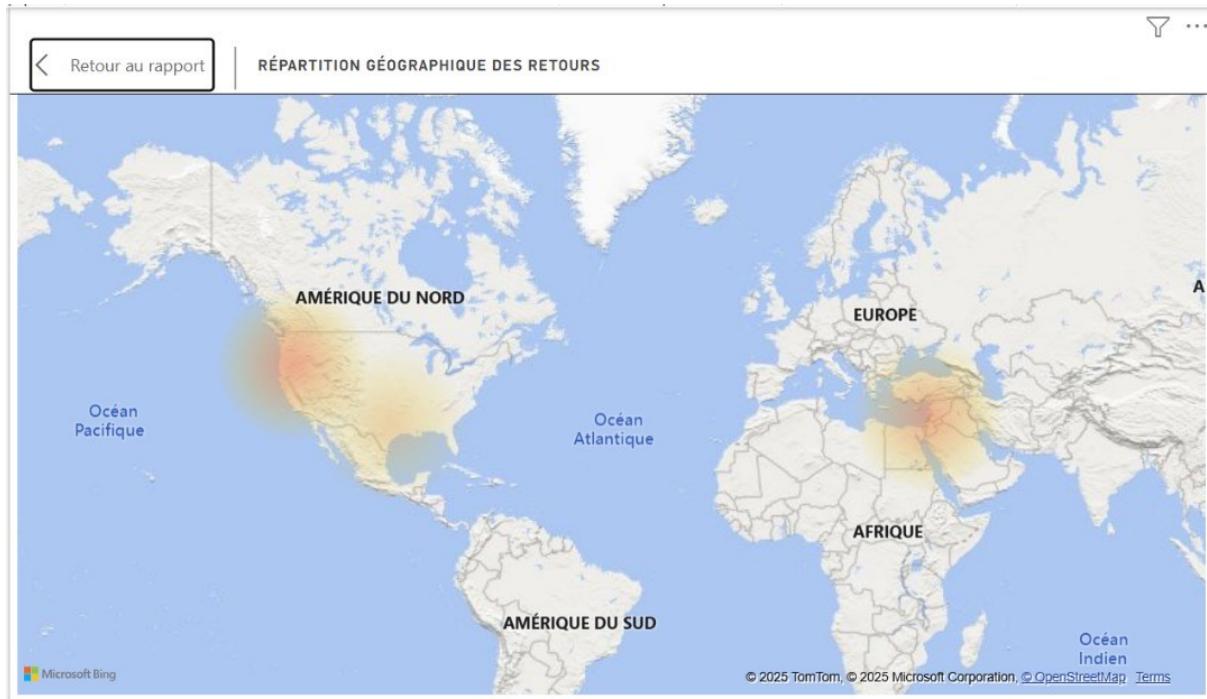
- Les accessoires sont les plus sujets aux retours, en nombre absolu et en proportion. Cela peut révéler un problème de qualité, d'incompatibilité, ou d'attentes mal gérées chez les clients pour cette catégorie.
- Les autres catégories (ordinateurs, tablettes, téléphones, montres connectées) présentent des taux de retour comparables, suggérant que les raisons de retour sont globalement similaires ou que les processus de vente et de suivi sont cohérents.

En résumé :

Les accessoires concentrent le plus grand nombre de retours, ce qui en fait une catégorie prioritaire pour l'amélioration. Les autres familles de produits présentent des taux de retour similaires et modérés. Une analyse plus fine des causes de retour pourrait aider à cibler les actions correctives.

3. Répartition géographique des retours

Nous avons réalisé une carte par région pour visualiser les zones avec un taux élevé de retour.



Points clés du graphique

- **Visualisation :** Il s'agit d'une carte thermique mondiale. Les zones en rouge indiquent une forte concentration de retours produits, les zones jaunes/orangées une concentration modérée.

- Zones les plus concernées :

- Amérique du Nord (Région East) : On observe une importante zone rouge dans le centre-ouest des États-Unis, signalant un nombre particulièrement élevé de retours dans cette région.
- Europe/Moyen-Orient (Région South & North) : Une autre zone de forte intensité (rouge/orange) se situe entre l'Europe centrale et le Moyen-Orient/Nord de l'Afrique.

Interprétations

1. Fort taux de retours en Amérique du Nord :

Cela peut s'expliquer par :

- Une forte part de marché de l'entreprise dans cette région (plus de ventes → plus de retours en nombre absolu).
- Des habitudes de consommation : les clients nord-américains sont souvent plus enclins à retourner les produits, profitant de politiques de retour plus souples.
- Une insatisfaction ou des attentes élevées sur certains produits.

2. Concentration en Europe centrale / Moyen-Orient :

- La présence de zones rouges dans cette région peut traduire une forte activité commerciale et/ou des problématiques logistiques ou de satisfaction particulière à cet endroit.

3. Absence de points chauds ailleurs :

- Cela peut être lié à un volume de ventes plus faible dans d'autres régions, ou à des politiques de retour plus strictes.

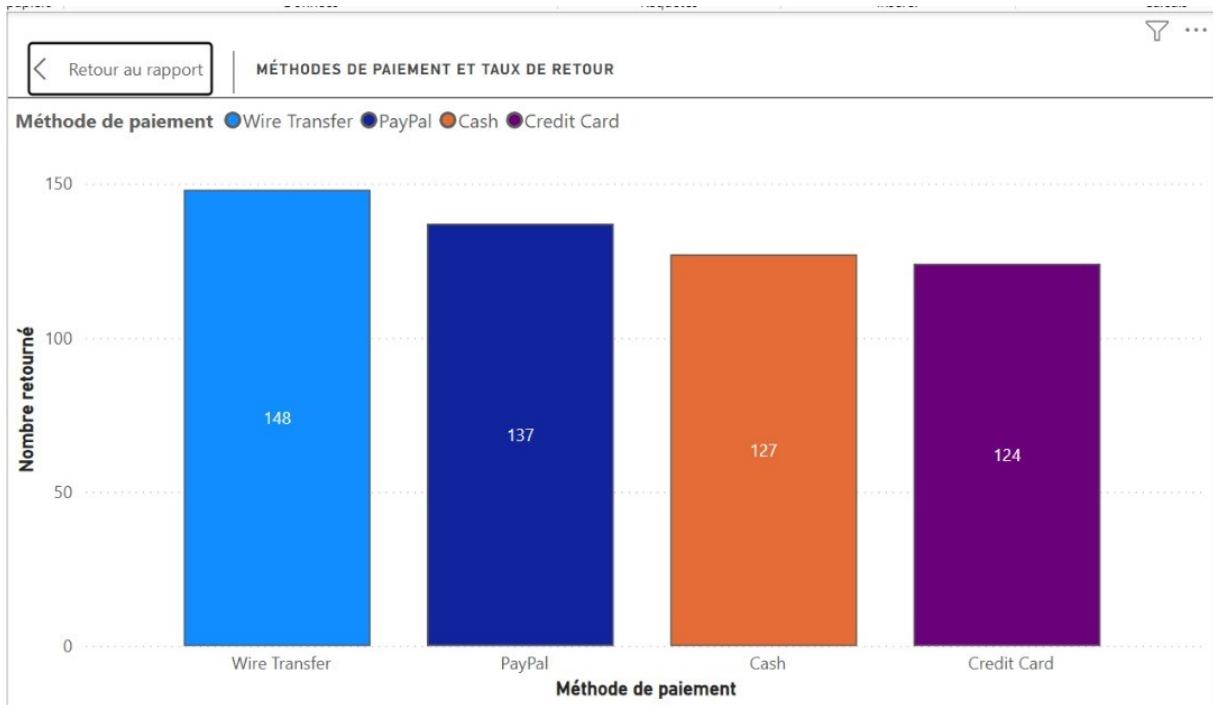
Conclusion synthétique

La répartition géographique des retours met en évidence deux principaux foyers :

- Le centre-ouest des États-Unis
- L'Europe centrale / Moyen-Orient

Cela invite à cibler en priorité l'analyse des facteurs de retour dans ces zones pour réduire les retours produits, par exemple en améliorant la qualité, l'information client, ou la logistique.

4. Méthodes de paiement et taux de retour



Lecture du graphique

Ce graphique à barres présente le nombre de produits retournés en fonction de la méthode de paiement utilisée lors de l'achat.

Les méthodes considérées sont : **Wire Transfer (virement bancaire)**, **PayPal**, **Cash**, **Credit Card (carte de crédit)**.

L'axe vertical indique le nombre de produits retournés pour chaque méthode et l'axe horizontale indique le moyen paiement utilisé.

Points clés

- ✓ Wire Transfer enregistre le plus grand nombre de retours (148).
- ✓ PayPal suit avec 137 retours.
- ✓ Cash : 127 retours.
- ✓ Credit Card : 124 retours.
- ✓ Les écarts entre méthodes ne sont pas énormes, mais le virement bancaire et PayPal ressortent légèrement au-dessus.

Interprétations :

1. Wire Transfer et PayPal en tête

- Les clients ayant payé par virement ou PayPal semblent légèrement plus enclins à retourner leurs produits.
- Cela peut traduire un profil client spécifique (acheteurs prudents, achats à distance, etc.) ou une relation différente à la satisfaction et au service après-vente.

2. Cash et Credit Card

- Les achats en espèces et par carte de crédit génèrent un peu moins de retours.
- Il est possible que les clients qui paient en cash soient plus sûrs de leur choix, ou que la politique de retour diffère selon la méthode de paiement.

3. Écarts modérés

- Globalement, toutes les méthodes de paiement sont concernées par les retours.
- Aucun mode de paiement ne semble totalement à l'abri du phénomène.

Conclusion :

Le virement bancaire et PayPal sont associés au plus grand nombre de retours produits, mais la différence reste modérée par rapport aux autres méthodes. Toutes les méthodes de paiement sont concernées par le phénomène des retours. Un calcul du taux de retour par méthode et une analyse croisée aideraient à affiner l'interprétation et à cibler d'éventuelles actions correctives.

PARTIE 2 : ANALYSE DE CAUSES POSSIBLES

5. Taux de retour par tranche de prix.



Lecture du graphique :

- Ce graphique montre le nombre de produits retournés par tranche de prix (classe de prix).
- Sur l'axe horizontal : les différentes classes de prix (en euros, de [50–180 €] jusqu'à [1869–1999 €]).
- Sur l'axe vertical : le nombre de produits retournés pour chaque tranche de prix.

Points clés

- **Tranches les plus concernées :**
- La tranche [1740–1869 €] est celle qui compte le plus de retours (46), suivie de près par la tranche [50–180 €] (45).
- Plusieurs tranches intermédiaires ([830–960 €], [440–570 €], [1480–1610 €], [700–830 €]) affichent aussi un nombre de retours élevé (autour de 39–41).
- **Tendance globale :**
- Il n'y a pas de relation strictement linéaire entre le prix et le nombre de retours : aussi bien des produits très chers que des produits d'entrée de gamme sont fréquemment retournés.

- Les tranches de prix extrêmes (très bas ou très haut) figurent en haut du classement par nombre de retours.
- Tranches avec moins de retours :
- Certaines tranches, notamment [960–1090 €] (25) et [180–310 €] (26), enregistrent les niveaux les plus bas de retours.

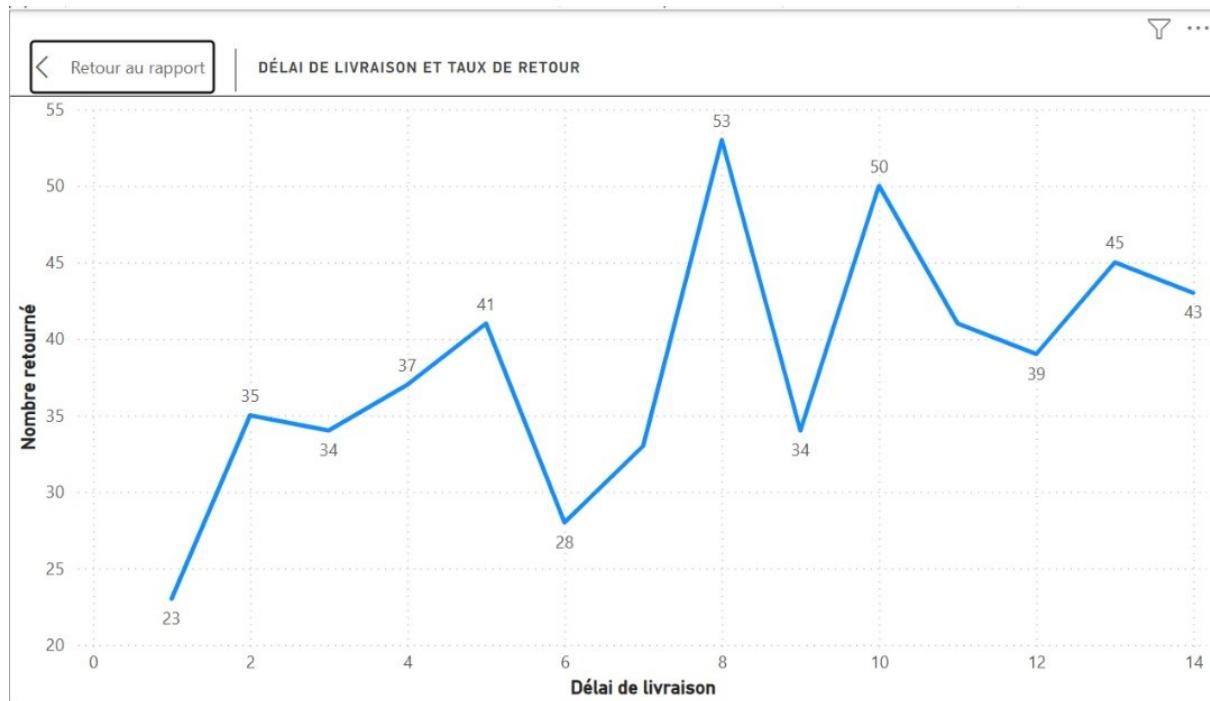
Interprétations

- Produits très chers souvent retournés : Les produits haut de gamme ([1740–1869 €], [1480–1610 €], [1869–1999 €]) sont parmi les plus retournés. Cela peut s'expliquer par des attentes très élevées des clients, une exigence accrue en termes de qualité ou d'adéquation au besoin.
- Entrée de gamme également à risque : Les produits les moins chers ([50–180 €]) sont eux aussi fréquemment retournés, possiblement en raison d'une qualité perçue plus faible, de problèmes de compatibilité ou d'achats impulsifs.
- Pas de lien direct entre prix et retour : Les retours touchent toutes les classes de prix, ce qui suggère que d'autres facteurs (qualité, adéquation au besoin, informations sur le produit, attentes...) sont déterminants.
- Tranches médianes parfois épargnées : Certaines tranches intermédiaires connaissent relativement moins de retours : il serait intéressant d'analyser la nature des produits dans ces tranches.

Conclusion :

Les produits les plus chers et les moins chers sont les plus susceptibles d'être retournés, mais toutes les tranches de prix sont concernées. Cela indique que le prix seul n'explique pas le phénomène des retours, et qu'il faut analyser aussi la qualité, les attentes et les catégories de produits pour mieux comprendre ce comportement client.

6. Délai de livraison et taux de retour.



Lecture du graphique

- Axe horizontal (x) : Délai de livraison (DeliveryDays) en jours, de 1 à 14.
- Axe vertical (y) : Nombre de produits retournés pour chaque délai de livraison.

Points clés

Tendance générale :

- Le nombre de retours a tendance à augmenter avec l'allongement du délai de livraison, même si la relation n'est pas strictement linéaire.

- Retours faibles pour les livraisons très rapides :

- Pour un délai de livraison d'1 jour, le nombre de retours est le plus bas (23).

- Pic de retours :

- Les plus forts pics de retours sont observés pour des délais de 8 jours (53) et 10 jours (50).

- Variabilité :

- Quelques fluctuations : les retours baissent par exemple à 6 jours (28) après avoir augmenté à 4 jours (37) et 5 jours (41).
- Après 8 jours, les retours restent globalement élevés, rarement sous la barre des 39.
- Livraisons longues = plus de retours :
- Les délais de 9 à 14 jours connaissent systématiquement un nombre de retours supérieur à 39.

Interprétation

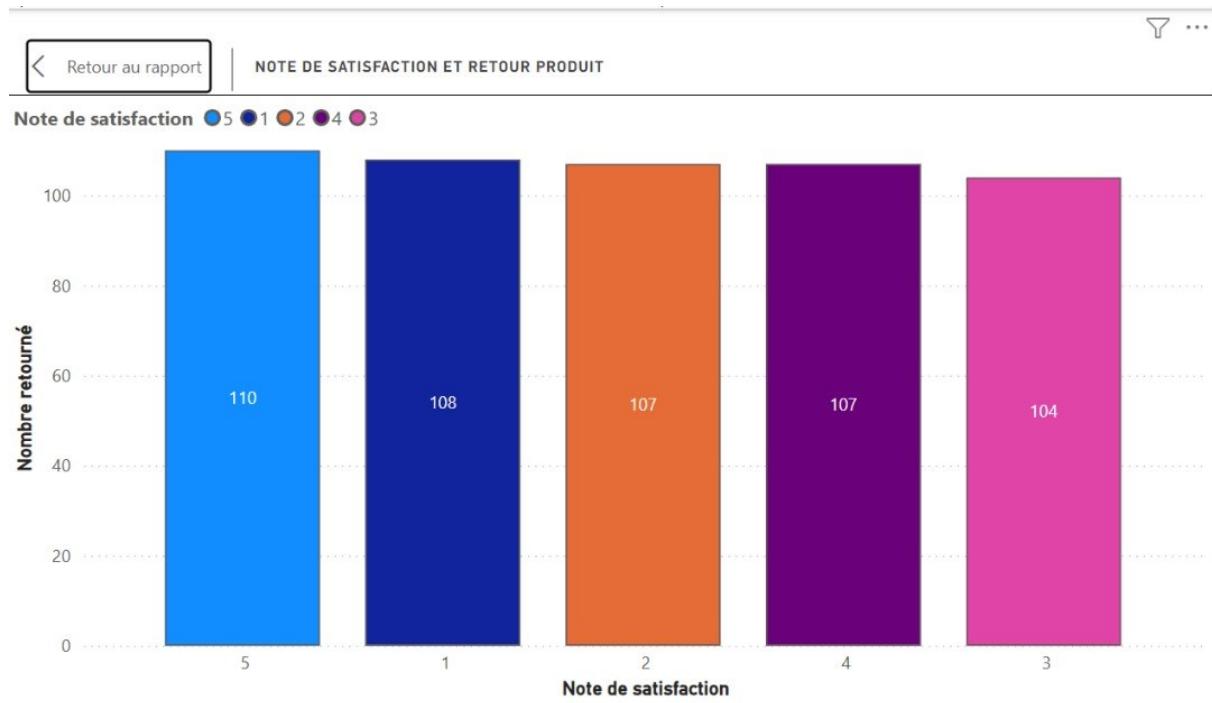
- Plus le délai de livraison s'allonge, plus le produit risque d'être retourné.

- Cela peut s'expliquer par l'impatience des clients, une perte d'intérêt pour le produit, ou une moins bonne expérience globale d'achat.
- Le délai optimal semble être de 1 à 5 jours, où les retours restent relativement bas.
- Les délais longs (8 jours et plus) sont associés à une insatisfaction accrue et donc à plus de retours.
- La variabilité peut aussi indiquer que d'autres facteurs interviennent (type de produit, communication client, motif d'achat...).

Conclusion :

Le graphique met en évidence une corrélation entre l'augmentation du délai de livraison et le taux de retour des produits : plus l'attente est longue, plus les clients sont susceptibles de retourner leur commande. Pour limiter les retours, l'entreprise a donc intérêt à optimiser ses délais de livraison.

7. Note de satisfaction et retour du produit.



Le graphique affiché est un **diagramme en barres groupées** montrant le **nombre de retours produits (axe Y)** en fonction de la **note de satisfaction client (axe X)**. Chaque barre représente une **note de satisfaction (de 1 à 5)**, et la valeur affichée à l'intérieur de chaque barre correspond au **nombre de produits retournés** pour cette note.

Analyse des résultats du graphique

Distribution des retours selon les notes de satisfaction :

Observations clés :

1. Retour élevé même pour les notes élevées :

- Contre-intuitivement, la note de satisfaction **5** a le **plus grand nombre de retours (110)**, ce qui peut sembler paradoxal. Cela suggère que **certains clients retournent des produits même s'ils se disent très satisfaits**, ce qui peut être dû à des erreurs de commande, des cadeaux, ou des incohérences dans les données.

2. Notes basses (1 et 2) également associées à un grand nombre de retours :

- Les retours sont également très fréquents chez les clients ayant attribué des **notes faibles (1 et 2)**, ce qui est plus attendu, car une insatisfaction est souvent un motif de retour.

3. Faible variation globale :

- Les différences de retours entre les notes sont **très faibles** (**variation de seulement 6 unités entre le minimum et le maximum**). Cela indique qu'aucune note ne se distingue fortement en termes de volume de retours.

Interprétation

- **Pas de lien clair entre satisfaction et retour :**
 - Ce graphique **ne montre pas de corrélation directe entre une mauvaise note de satisfaction et un retour du produit**. En fait, les retours sont presque uniformément répartis entre toutes les notes.
- **Hypothèse possible :**
 - Il est probable que la variable SatisfactionRating ne soit pas **indépendante** du retour mais qu'elle soit **renseignée après le retour**, ou qu'elle soit influencée par d'autres facteurs comme :
 - La qualité du service client,
 - Les délais de livraison,
 - Des promotions/compensations.

Recommandations pour analyses complémentaires :

1. **Comparer le taux de retour (%) par note de satisfaction**, plutôt que les **nombres absolus** :

Par exemple : combien de clients sur 100 ayant mis la note 1 ont retourné le produit ?

2. **Croiser la variable "Returned" avec "DeliveryDays" ou "Category"** pour chercher d'autres facteurs explicatifs.
3. **Utiliser un boxplot du DeliveryDays par SatisfactionRating** pour voir si un délai de livraison influence la satisfaction.
4. **Créer une matrice de corrélation** pour détecter les variables les plus corrélées au retour produit.

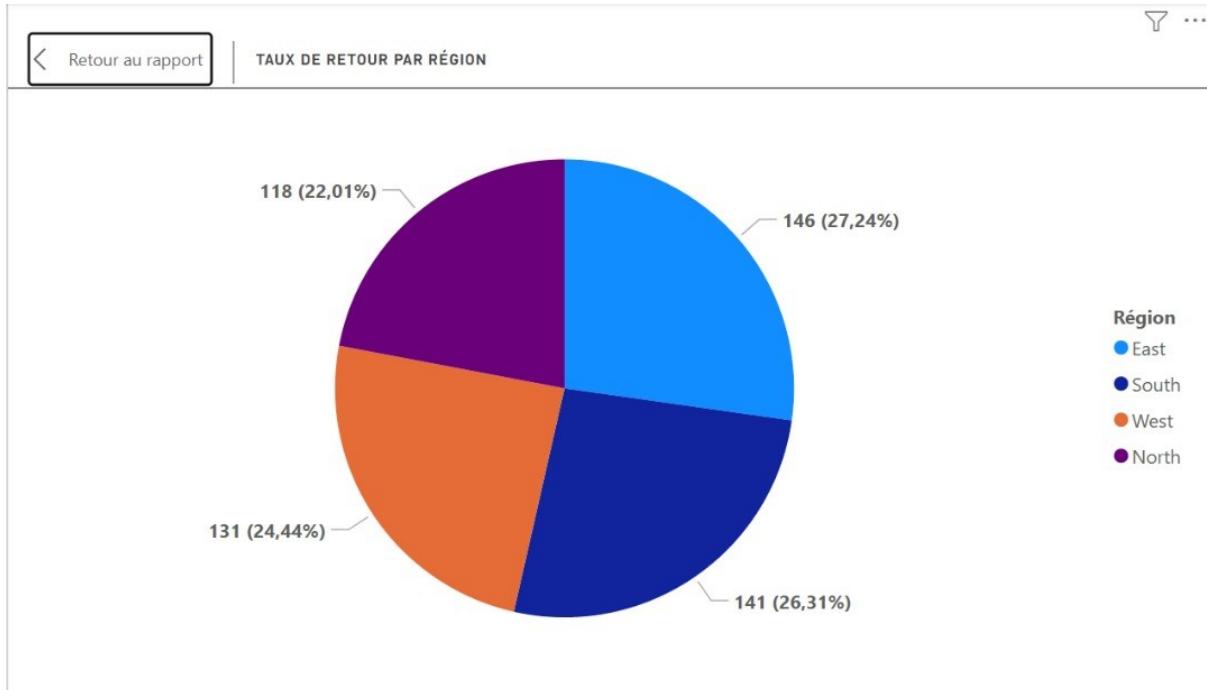
Conclusion

Le graphique fourni montre que les retours sont **presque également répartis entre les différentes notes de satisfaction**, avec une **anomalie potentielle pour les notes élevées**. Cela remet en question l'**utilisation de la satisfaction comme facteur déterminant du retour**, et suggère de poursuivre l'**exploration avec d'autres variables** pour mieux comprendre les causes des retours produits.

Souhaitez-vous que je vous aide à créer un graphique du **taux de retour par note de satisfaction** ou analyser d'autres variables du fichier synthetic_sales_data.csv ?

8.1. Profils clients à risque :

- Représentation du taux de retour par région.



Cette image représente un **diagramme circulaire** montrant le **nombre de retours et pourcentage du total**. Voici une analyse structurée en réponse à la consigne "**Profils clients à risque – Taux de retour par région**".

Observations clés :

1. **Région East = pourcentage le plus élevé (27,24 %) :**
 - Cela représente le **profil client à plus haut risque de retour produit**.
 - Nécessite une **analyse plus fine** : types de produits achetés, délais de livraison, satisfaction, etc.
2. **South suit de près (26,31 %)**, écart très faible avec East :
 - Ces deux régions totalisent ensemble **plus de 53 %** des retours, ce qui est significatif.
 - Il est possible qu'elles concentrent une **plus forte activité commerciale** (plus de ventes), ou qu'il existe un **problème logistique ou produit** spécifique à ces zones.
3. **North = pourcentage le plus bas (22,01 %) :**
 - Les clients de cette région semblent **moins enclins à retourner** leurs produits.

- Cela pourrait indiquer un **meilleur service ou une meilleure adéquation produit-attente**.

Interprétation : Profils à risque

Les **régions East et South** représentent les **profils géographiques clients les plus à risque en matière de retour produit**. Plusieurs hypothèses sont à explorer pour expliquer cela :

- **Problèmes logistiques** : retards de livraison, mauvaise gestion du dernier kilomètre.
- **Produits spécifiques vendus** dans ces régions (certains produits pourraient être plus souvent retournés).
- **Profil client différent** (ex : attentes plus élevées, clients plus exigeants).
- **Différences culturelles** ou comportementales en matière de retour.

Recommandations pour la suite :

1. **Croiser les données par région + délai de livraison + satisfaction** pour mieux comprendre l'origine des retours.
2. **Analyser les types de produits** les plus retournés par région.
3. **Évaluer les méthodes de paiement ou le canal d'achat** (mobile, desktop) utilisés par région.
4. **Calculer le taux de retour relatif**, c'est-à-dire **retours / ventes** dans chaque région, pour corriger l'effet du volume.

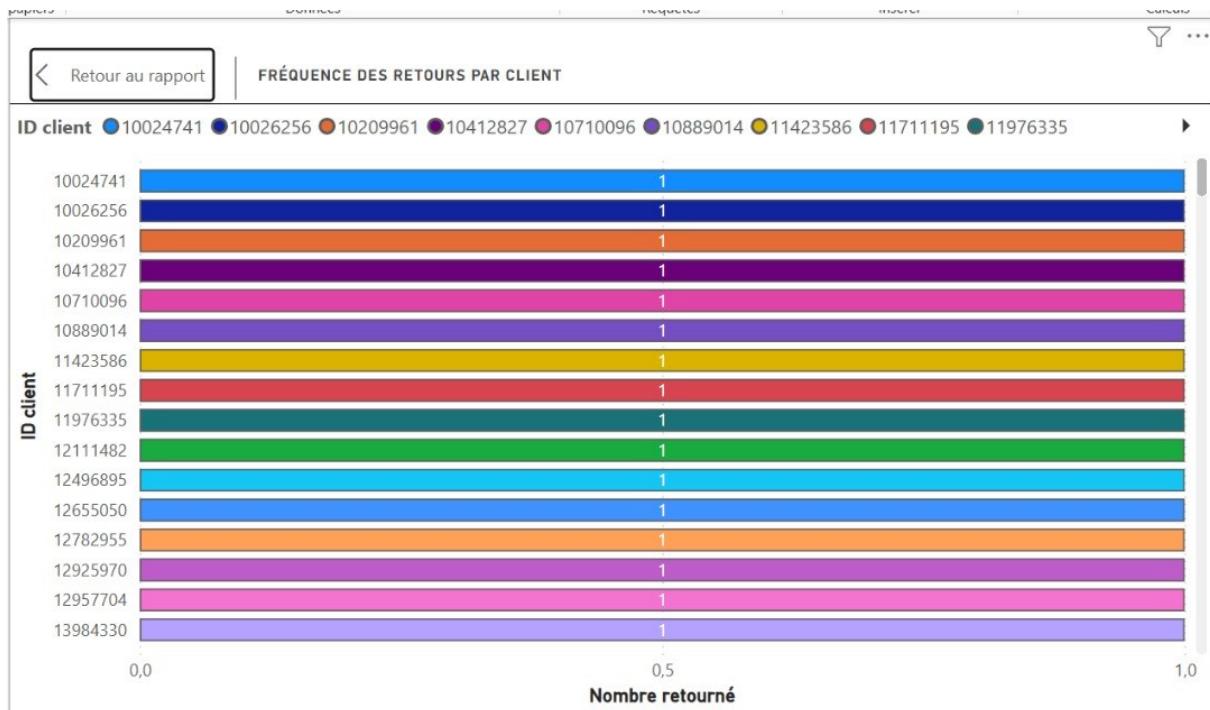
Conclusion

Le graphique met en évidence que les régions **East (27,24 %)** et **South (26,31 %)** concentrent la majorité des retours, ce qui en fait les **principales zones à surveiller**. Ces insights permettent d'orienter des actions ciblées pour **réduire le taux de retour dans ces régions**, comme améliorer la logistique, revoir certaines offres produit ou renforcer le suivi client.

Souhaitez-vous une visualisation ou analyse croisée (par exemple région vs délai de livraison ou satisfaction) pour mieux comprendre ces écarts ?

8.2. Profils clients à risque

- Représentation du taux de retour par région.



Cette image montre la **fréquence des retours par client**, avec un graphique où chaque client a un retour unique.

Analyse :

1. Observation :

- Tous les clients listés dans cette visualisation semblent avoir effectué **exactement un retour**.
- La fréquence est uniformément **1** pour chaque ID client.

8.2 Profils clients à risque

- Analyse de la fréquence des retours par client (CustomerID le plus souvent présent avec Returned = 1).

FRÉQUENCE DES RETOURS PAR CLIENT - TABLE		
ID client	Nombre retourné	% du total général
10024741	1	0,19%
10026256	1	0,19%
10209961	1	0,19%
10412827	1	0,19%
10710096	1	0,19%
10889014	1	0,19%
11423586	1	0,19%
11711195	1	0,19%
11976335	1	0,19%
12111482	1	0,19%
12496895	1	0,19%
12655050	1	0,19%
12782955	1	0,19%
12925970	1	0,19%
12957704	1	0,19%
13984330	1	0,19%
13989423	1	0,19%
14216175	1	0,19%
Total	536	100,00%

Cette image montre la **fréquence des retours par client**, détaillée dans un tableau avec les colonnes **ID client**, **Nombre retourné**, et **% du total général**.

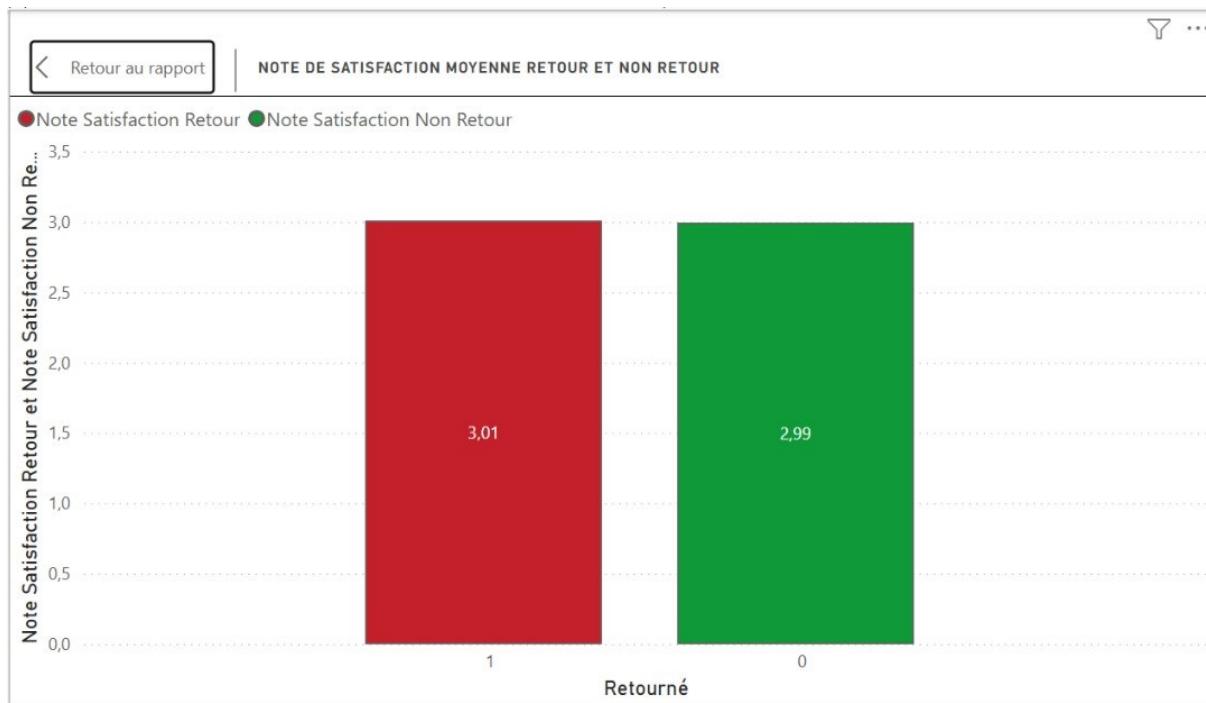
Analyse :

1. Observation :

- Chaque client a effectué **exactement un retour**, représentant **0,19%** du total général.
- Le total des retours est de **536**, soit **100%**

8.3 Profils clients à risque

Création d'un graphique comparant la note de satisfaction moyenne des clients qui retournent des produits vs ceux qui ne retournent pas.



Cette image montre un graphique comparant la **note de satisfaction moyenne** des clients qui retournent des produits (en rouge) à ceux qui ne retournent pas (en vert).

Analyse :

1. Observation :

- La note moyenne des clients ayant retourné des produits est légèrement supérieure (**3,01**) à celle des clients n'ayant pas retourné de produits (**2,99**).
- L'écart entre les deux groupes est très faible, ce qui suggère que le retour des produits n'affecte pas significativement la satisfaction