

# Rapport d'Analyse des Données de Vente au Détail en Ligne II(Online Retail II)

Une Analyse Exploratoire des Performances de Vente, des Clients  
et des Annulations

**Ibrahim**  
Data Analyst

## *Objectif du Projet :*

Ce rapport examine le dataset Online Retail II pour extraire des informations exploitables sur les tendances de vente, le comportement des clients et les annulations, démontrant des compétences analytiques et de visualisation pour une présentation de portfolio.

*Outils : Python (Pandas, Matplotlib)*

Janvier 2025

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Aperçu du Projet . . . . .	2
1.2	Résumé du Dataset . . . . .	2
<b>2</b>	<b>Résumé de la Méthodologie</b>	<b>3</b>
2.1	Prétraitement des Données : . . . . .	3
2.2	Segmentation . . . . .	3
2.3	Feature Engineering . . . . .	4
2.4	Outils . . . . .	4
2.5	Méthodes Analytiques: . . . . .	4
<b>3</b>	<b>Résultats et Insights</b>	<b>4</b>
3.1	Performances des Ventes . . . . .	4
3.1.1	Vente au détail . . . . .	4
3.1.2	Vente en gros . . . . .	11
3.2	Analyse des Clients . . . . .	17
3.2.1	Vente au détail . . . . .	17
3.2.2	Vente en gros . . . . .	22
3.3	Analyse des Annulations . . . . .	26
3.3.1	Vente au Détail . . . . .	26
3.3.2	Ventes en gros . . . . .	35
<b>4</b>	<b>Recommandations</b>	<b>44</b>
4.1	Ajustements de la Stratégie de Vente . . . . .	44
4.2	Segmentation des Clients et Fidélisation . . . . .	44
4.3	Réduction des Annulations . . . . .	44
4.4	Améliorations Opérationnelles et Analytiques . . . . .	45
<b>5</b>	<b>Limites de l'Analyse</b>	<b>45</b>
<b>6</b>	<b>Conclusion</b>	<b>45</b>
<b>7</b>	<b>Références</b>	<b>46</b>

# 1 Introduction

## 1.1 Aperçu du Projet

Ce projet analyse les performances de vente et le comportement des clients en utilisant le dataset Online Retail II, qui contient deux ans de données transactionnelles provenant d'un détaillant en ligne basé au Royaume-Uni. L'objectif est de découvrir des informations sur les tendances de vente, le comportement des clients et les annulations, afin de soutenir la prise de décision basée sur les données.

## 1.2 Résumé du Dataset

Online Retail II (Vente au Détail en Ligne II) contient des données transactionnelles d'un détaillant en ligne basé au Royaume-Uni, offrant une vue d'ensemble complète des activités de vente sur une période de deux ans. Voici un aperçu détaillé :

- **Période:** Décembre 2009 – Décembre 2011
- **Portée :** Ce jeu de données enregistre principalement des transactions de détail, mais inclut également certaines transactions en gros, capturant divers scénarios de vente allant des clients individuels aux commandes à grande échelle.
- **Attributs Clés:**
  - **Invoice:** Identifiant unique pour chaque transaction, représentant une vente ou une annulation (par exemple, "536365" pour une vente ou "C536366" pour une annulation).
  - **StockCode:** Un identifiant à cinq caractères pour les produits, qui peut inclure des préfixes ou des suffixes (par exemple, "85123A").
  - **Description:** Une description textuelle du produit (par exemple, "BOUGEOIR EN FORME DE CŒUR BLANC SUSPENDU").
  - **Quantity:** Nombre d'unités par produit dans chaque transaction.
  - **Unit Price:** Prix par unité en GBP.
  - **InvoiceDate:** Date et heure de la transaction.
  - **CustomerID:** Un identifiant unique à cinq chiffres pour chaque client.
  - **Country:** Localisation géographique du client.
- **Objectif de l'Analyse:** Ce rapport vise à explorer:
  - **Performances des Ventes:**
    - \* Modèles de revenus, tendances et taux de croissance pour les transactions de détail et en gros.
    - \* Identifier les produits et les régions les plus performants.
  - **Comportement des Clients:** performances clés, fidélisation et segmentation.
  - **Annulations :** Fréquence et impact des annulations de transactions.
- **Pertinence:** Ce jeu de données est précieux pour examiner les performances commerciales, l'engagement client et les défis opérationnels dans un contexte de vente en ligne.

## 2 Résumé de la Méthodologie

### 2.1 Prétraitement des Données :

Le jeu de données a subi un processus de nettoyage rigoureux pour garantir la précision et la pertinence de l'analyse:

- **Suppression des Duplicatas:** Identification et suppression des doublons pour éviter le double comptage des transactions.
- **Gestion des Valeurs Manquantes:**
  - Les lignes avec des prix de produit manquants ou nuls ont été traitées de manière appropriée, garantissant la cohérence dans les calculs de revenus.
  - Les lignes avec plusieurs valeurs critiques manquantes (par exemple, ID Client et Détails du Produit manquants) ont été supprimées, car elles offraient une valeur analytique limitée.
- **Filtrage des Données Non Transactionnelles:** Exclusion des lignes ne représentant pas de transactions réelles, telles que les entrées administratives, les ajustements ou les données de remplissage.
- **Détection des Valeurs Aberrantes:** Revue des valeurs aberrantes dans les quantités et les prix (par exemple, valeurs excessivement élevées ou négatives) et traitement par segmentation ou exclusion.
- **Assurer la Cohérence:** Standardisation des formats de données pour les attributs clés comme les dates et les identifiants afin de faciliter une segmentation et une analyse fluides.

### 2.2 Segmentation

Le jeu de données a été segmenté pour permettre une analyse adaptée des différents types de transactions et des comportements des clients:

- **Transactions de Détail:** Les transactions de détail régulières, constituant le segment le plus important et représentant des achats individuels plus petits, ont été isolées pour capturer les comportements d'achat typiques des consommateurs.
- **Annulations de Détail:** Les annulations dans les transactions de détail ont été identifiées séparément, permettant l'analyse des tendances d'annulation et de leur effet sur les ventes et la fidélisation des clients.
- **Commandes en Gros:** Les transactions impliquant de grandes quantités (par exemple, plus de 50 unités par produit) ont été classées comme commandes en gros, constituant un petit segment, fournissant des informations sur les modèles d'achat à volume élevé.
- **Annulations en Gros:** Les annulations associées aux commandes en gros ont été identifiées pour analyser leur impact sur les revenus et le comportement des clients, en se concentrant sur les tendances spécifiques aux transactions à grande échelle.

## 2.3 Feature Engineering

Des métriques clés ont été dérivées, telles que le revenu ( $Quantity \times Unit Price$ ), pour mesurer les revenus des transactions. De plus, les données ont été organisées en agrégations temporelles, incluant des résumés mensuels, trimestriels et annuels, afin de permettre une analyse des tendances et des insights temporels plus approfondis.

## 2.4 Outils

- **Excel:** Utilisé pour l'inspection initiale des données et l'exploration.
- **Python (Pandas, Matplotlib):** Employé pour le nettoyage avancé des données, l'analyse, et la création de visualisations pertinentes.

## 2.5 Méthodes Analytiques:

- **Analyse des Tendances:** Calcul de plusieurs indicateurs clés de performance (KPI), suivi des tendances de vente (mensuelles, trimestrielles) et des comportements des clients.
- **Analyse de Segmentation:** Examen des ventes régulières, des commandes en gros et de leurs annulations correspondantes.
- **Visualisation:** Utilisation de divers graphiques (barres, lignes, cartes hiérarchiques, camemberts) pour présenter clairement les insights.

# 3 Résultats et Insights

## 3.1 Performances des Ventes

### 3.1.1 Vente au détail

Cette section se concentre sur les ventes au détail avec une quantité inférieure à 50 unités par produit, excluant les achats en gros ( $Quantité \geq 50$ ) et les annulations, fournissant des informations sur les tendances et la performance des ventes au détail.

#### Vue d'ensemble des ventes au détail:

Couvrant la période de décembre 2009 à décembre 2011, nous avons :

- **Revenu total :** £14,283,212.07
- **Nombre total de transactions :** 970,147
- **Nombre total de commandes (Facture) :** 39,287

#### Quelques indicateurs généraux:

- **Valeur moyenne des commandes (AOV) :** £363.56, ce qui représente le revenu moyen par commande. La valeur élevée de l'AOV suggère que les clients font des investissements importants par commande.

- **Taille moyenne du panier (ABS)** : 24.69 articles, ce qui indique le nombre moyen de produits par transaction. L'ABS élevé reflète les achats multiples effectués par les clients, un bundling efficace et des commandes volumineuses.

### Analyse temporelle:

#### MoM:

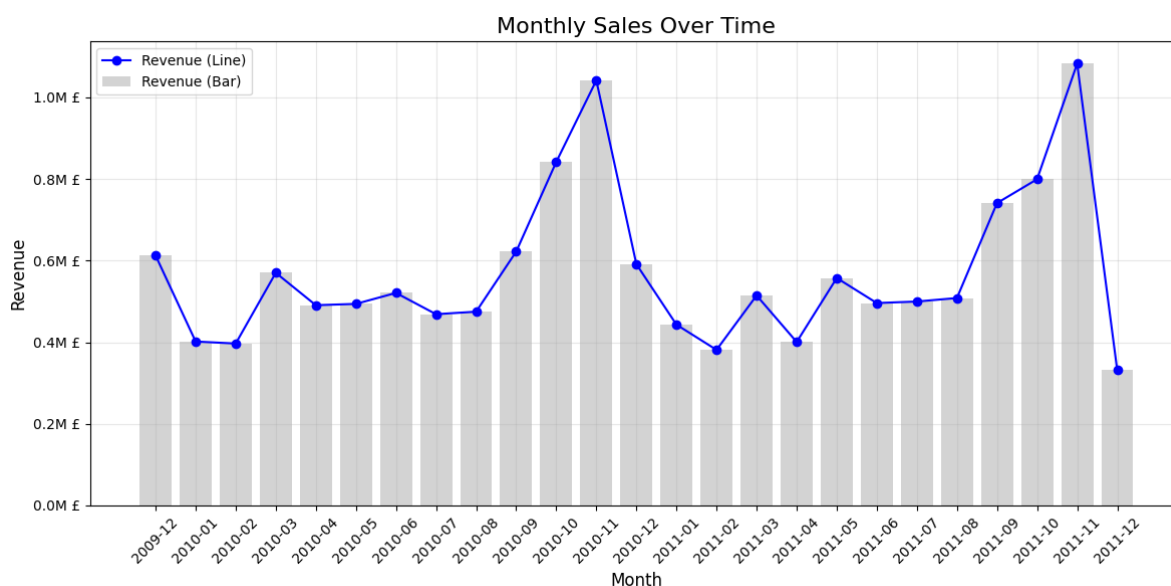


Figure 1: Ventes mensuelles au fil du temps

Les données de vente présentent des fluctuations saisonnières claires, avec un pic notable en novembre probablement dû à la saison des fêtes à venir, suivi d'une baisse dans les mois suivants, suggérant la présence d'une saisonnalité. De plus, la tendance générale montre une diminution progressive des ventes après novembre. De même, la croissance des ventes montre une progression positive régulière de juillet à novembre, suivie d'une forte baisse en décembre, ce qui soutient l'hypothèse d'un effet saisonnier et d'une tendance sous-jacente possible.

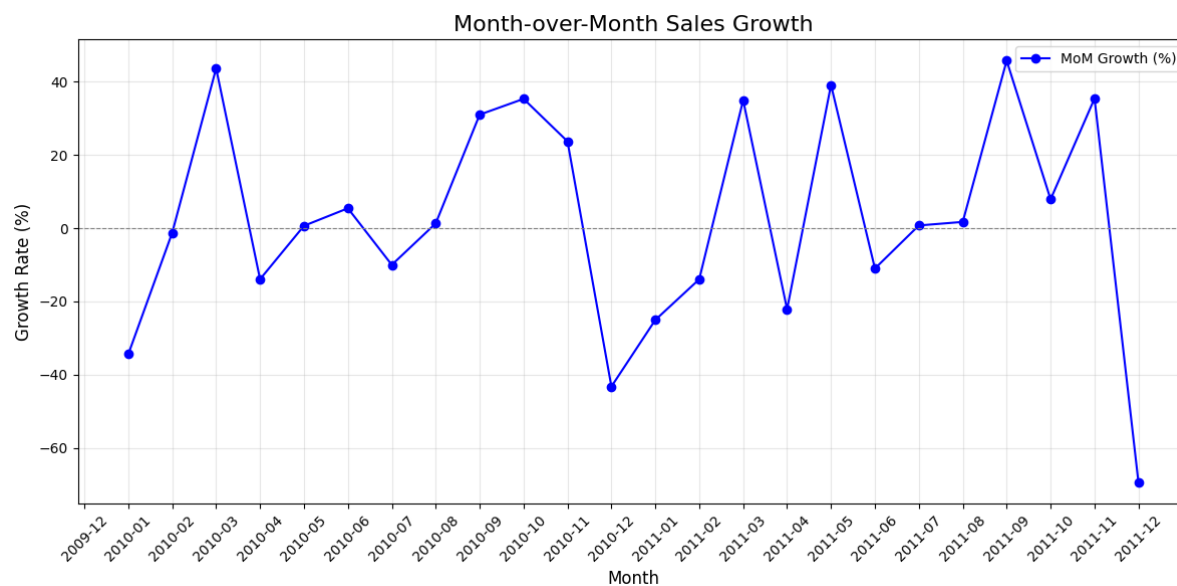


Figure 2: Croissance des ventes mois après mois

Cependant, après cette baisse, les ventes se stabilisent entre mars et août, indiquant une période de stabilité relative après la saison des fêtes.

**QoQ:**

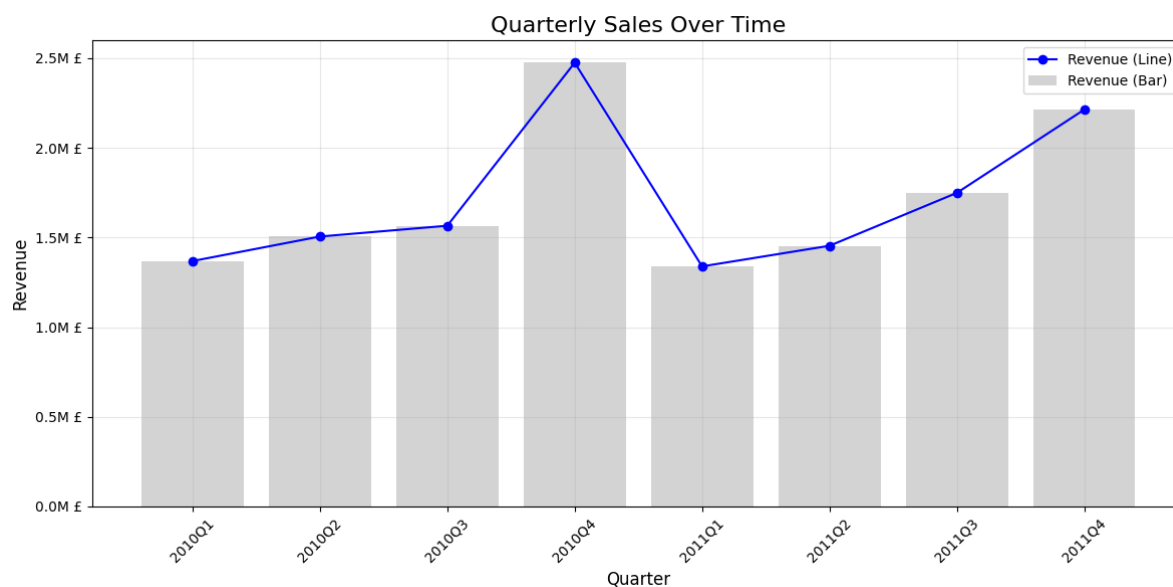


Figure 3: Ventes trimestrielles au fil du temps

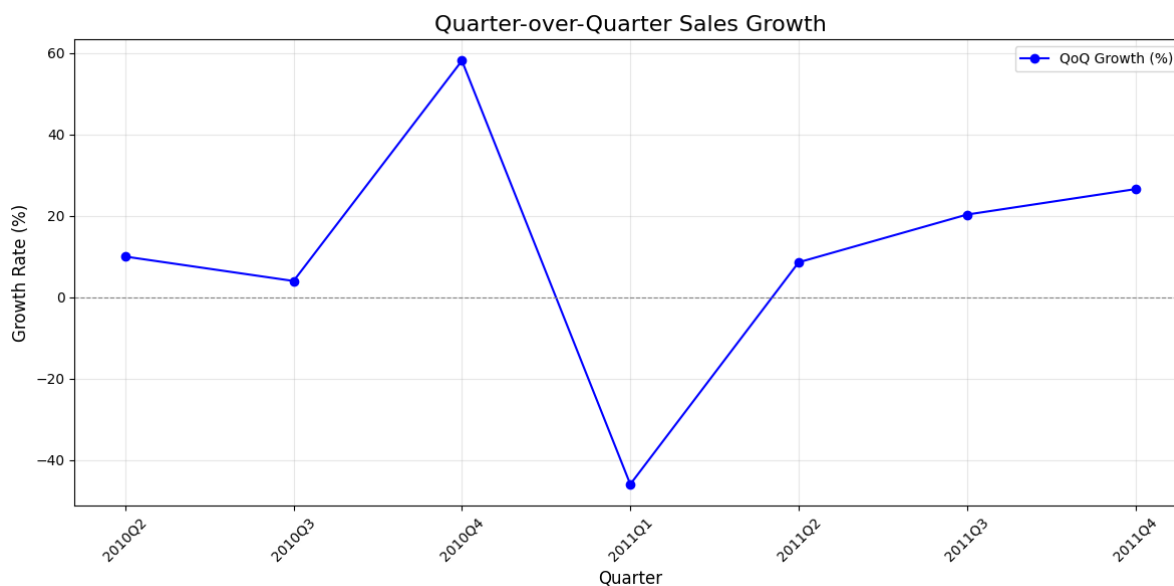


Figure 4: Croissance des ventes trimestre après trimestre

L'analyse trimestre après trimestre soutient l'analyse mensuelle précédente, renforçant le modèle de forte croissance des ventes au T4, en particulier au T4 2010, suivi d'une baisse au T1 2011. Cela suggère que le pic des ventes au T4 est probablement dû à des facteurs saisonniers, similaires aux tendances observées dans les données mensuelles. La croissance dans les trimestres suivants (T2 et T3 des deux années) reflète un retour à des chiffres de ventes plus stables, mais positifs, avec la baisse du T1 2011 indiquant un ralentissement après les fêtes.

### Performance régionale :

Le sous-ensemble de données de vente au détail inclut 43 pays, les 10 premiers en termes de revenu généré sont :



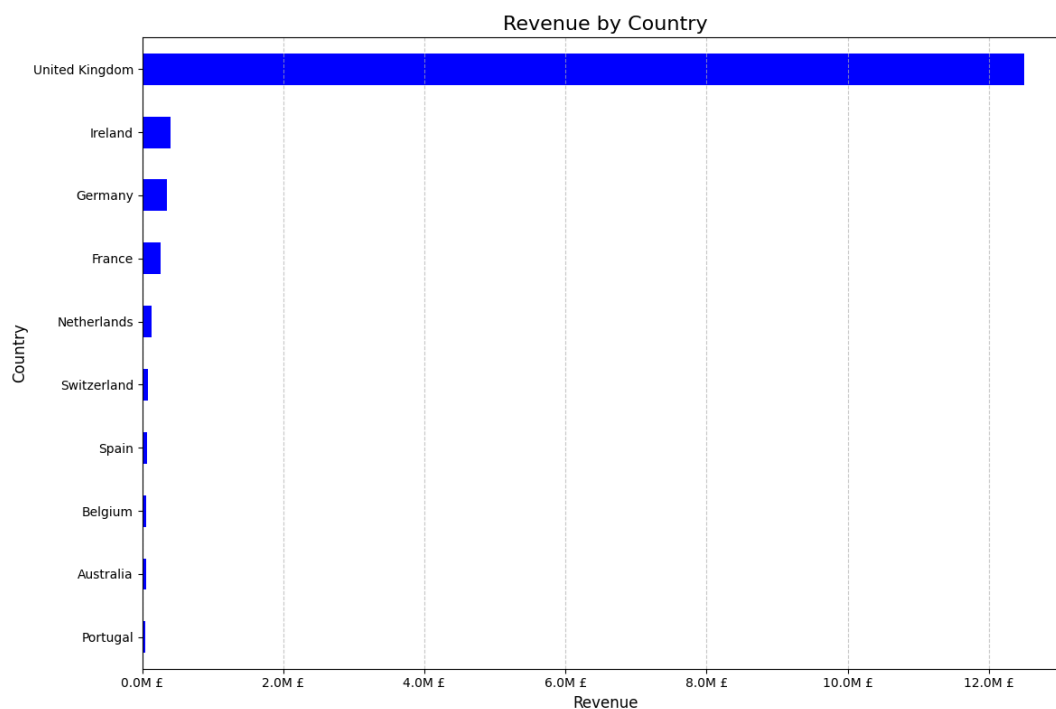


Figure 5: Top 10 des pays par revenu

Le Royaume-Uni domine les revenus avec £12.5M, suivi de l'Irlande (£399K) et de l'Allemagne (£344K). D'autres pays contribuent de petites sommes, mettant en évidence le rôle de leader du Royaume-Uni et les opportunités de croissance potentielles dans des marchés clés comme l'Irlande et l'Allemagne.

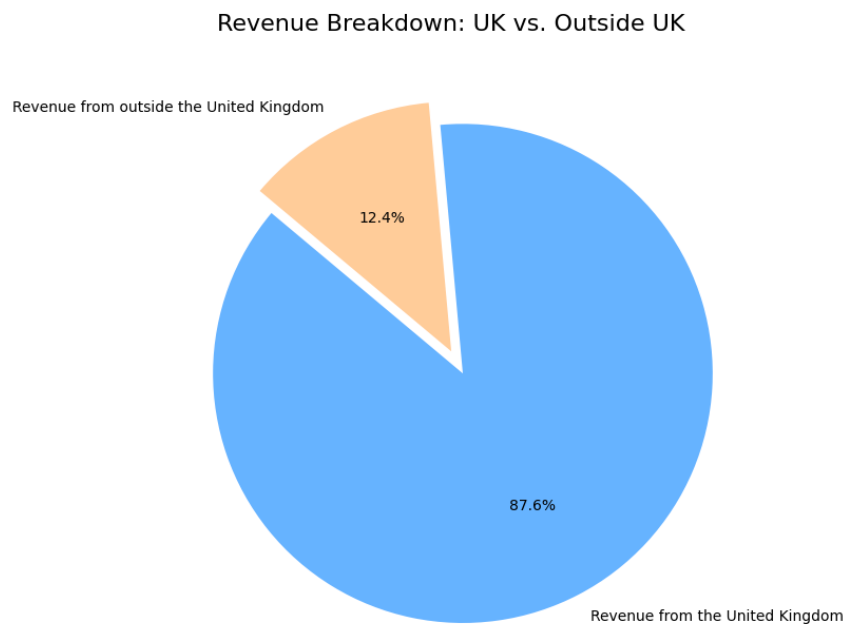


Figure 6: Répartition des revenus : Royaume-Uni vs Hors du Royaume-Uni

Le graphique circulaire illustre la contribution dominante du Royaume-Uni aux revenus totaux de vente au détail, représentant 87,6%. Cela met en lumière le rôle central du Royaume-Uni dans la stimulation des ventes, tandis que toutes les autres régions représentent seulement 12,4%, soulignant une concentration importante de la génération de revenus dans un seul marché.

#### Informations sur les produits :

Le sous-ensemble de données de vente au détail inclut 4883 produits distincts, les 20 premiers en termes de revenu généré sont :

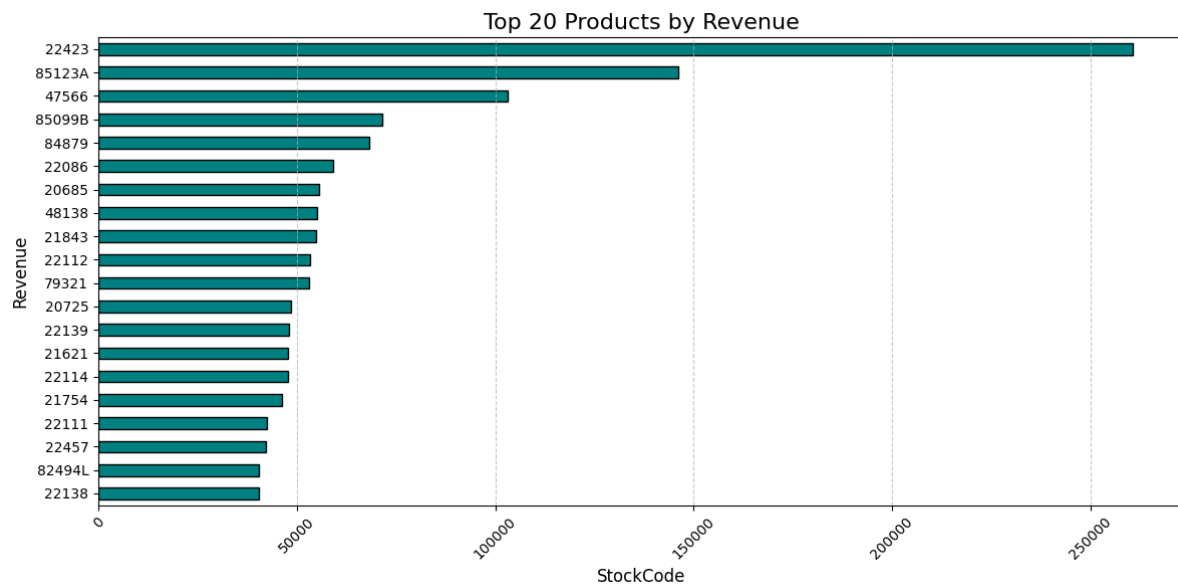


Figure 7: Top 20 des produits par revenu

Revenue Breakdown: Top 20 products vs. The rest of the products

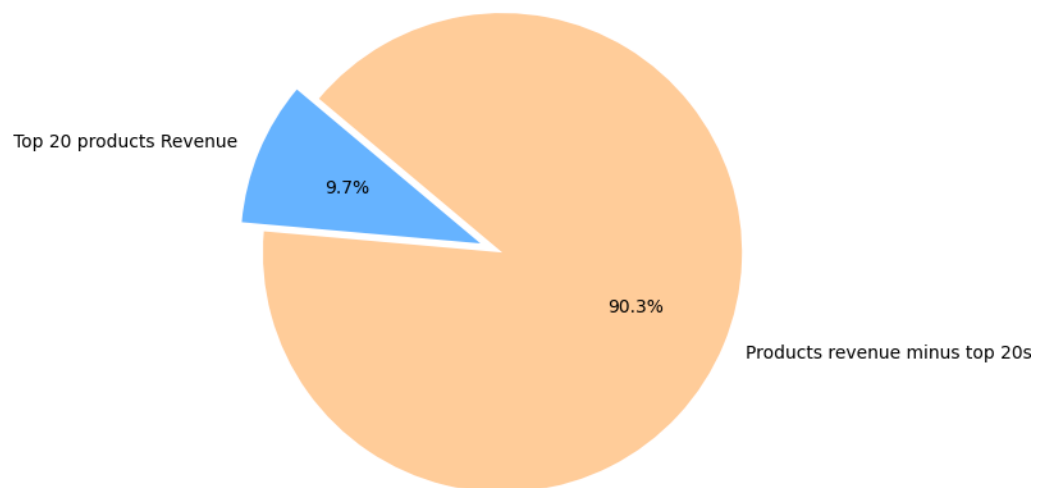


Figure 8: Répartition des revenus : Top 20 des produits vs Le reste des produits

Les 20 produits les plus vendus contribuent de manière significative aux revenus globaux, le produit le plus vendu (22423) générant £260,775.07. Bien qu'ils ne représentent que 0,4% des 4883 produits totaux, ils génèrent 9,7% des revenus totaux. Cela met en évidence leur contribution disproportionnée et leur importance stratégique.

### **Résumé des informations sur les ventes au détail:**

L'analyse met en évidence un fort modèle saisonnier, avec un pic de ventes en décembre, suivi d'une baisse après les fêtes. Cela suggère un effet saisonnier clair, avec des ventes se stabilisant dans les mois suivant les fêtes.

Au niveau régional, le Royaume-Uni représente 87,6% des revenus totaux, ce qui indique son rôle dominant sur le marché, tandis que d'autres régions contribuent à une part plus faible. Cette concentration présente des opportunités de croissance dans d'autres marchés comme l'Irlande et l'Allemagne.

En ce qui concerne la performance des produits, les 20 produits les plus vendus, bien qu'ils ne représentent que 0,4% de l'ensemble de la gamme de produits, génèrent 9,7% des revenus totaux. Cette part disproportionnée souligne leur importance dans la stratégie de produit.

Ces informations suggèrent qu'exploiter la force du Royaume-Uni tout en explorant des expansions dans d'autres régions, parallèlement à l'optimisation du portefeuille de produits, pourrait stimuler une croissance supplémentaire.

### **3.1.2 Vente en gros**

Cette section se concentre sur les ventes en gros avec une quantité supérieure à 50 unités par produit (Quantité > 50), fournissant des informations sur les tendances et la performance des ventes en gros.

#### **Vue d'ensemble des ventes en gros:**

Couvrant la période de décembre 2009 à décembre 2011, nous avons :

- **Revenu total des ventes en gros** : £5,802,651.11
- **Nombre total de transactions en gros** : 24 350
- **Nombre total de commandes incluant des transactions en gros (Facture)**  
: 8 504

Le revenu total, combinant les ventes au détail et en gros, est de £20,085,863.18, dont 28.9% provient des ventes en gros. Bien que les transactions individuelles en gros soient moins nombreuses (24,350 lignes contre 970,147 lignes pour les ventes au détail), elles représentent tout de même une part significative des ventes totales.

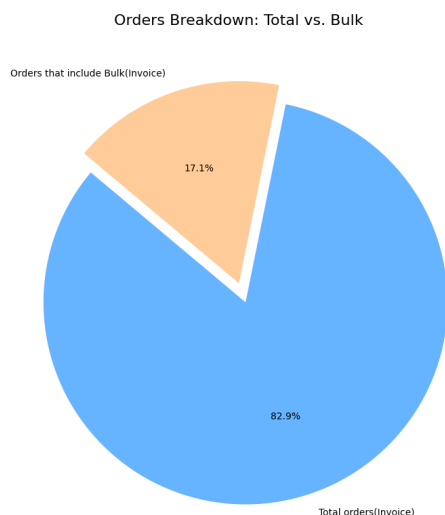


Figure 9: Répartition des commandes :  
Vente au détail vs Vente en gros

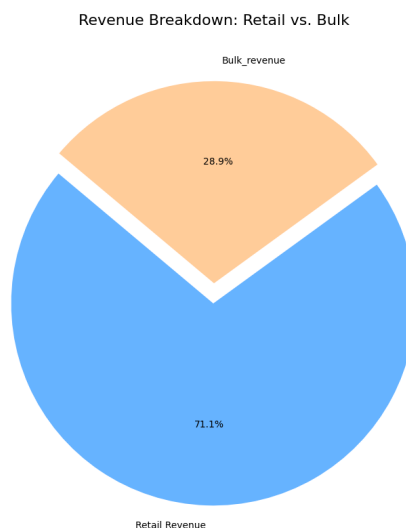


Figure 10: Répartition des revenus :  
Vente au détail vs Vente en gros

Avec 8 504 commandes incluant des achats en gros sur 41 297 commandes totales, le volume des transactions en gros montre une contribution régulière aux revenus.

#### Quelques indicateurs généraux:

**Valeur moyenne des commandes (AOV) :** £682.34. La valeur moyenne des commandes pour les achats en gros de £682.34 reflète la tendance à des dépenses plus élevées, typiques des transactions en gros, où les clients investissent davantage par commande en raison de l'achat de plus grandes quantités. C'est une caractéristique normale des achats en gros.

## Analyse temporelle:

## MoM:

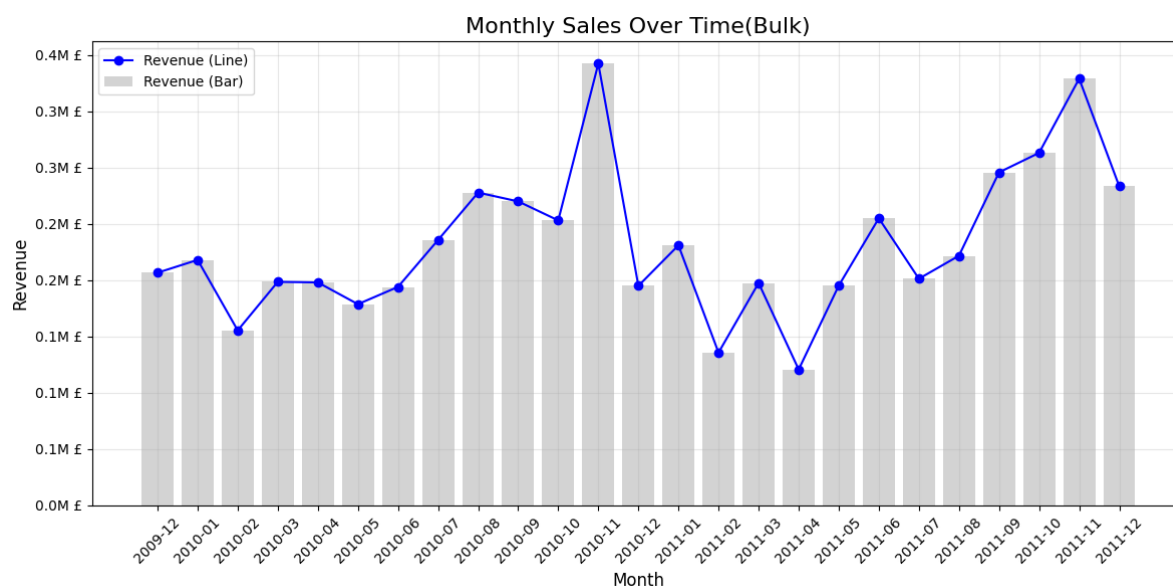


Figure 11: Ventes mensuelles en gros au fil du temps

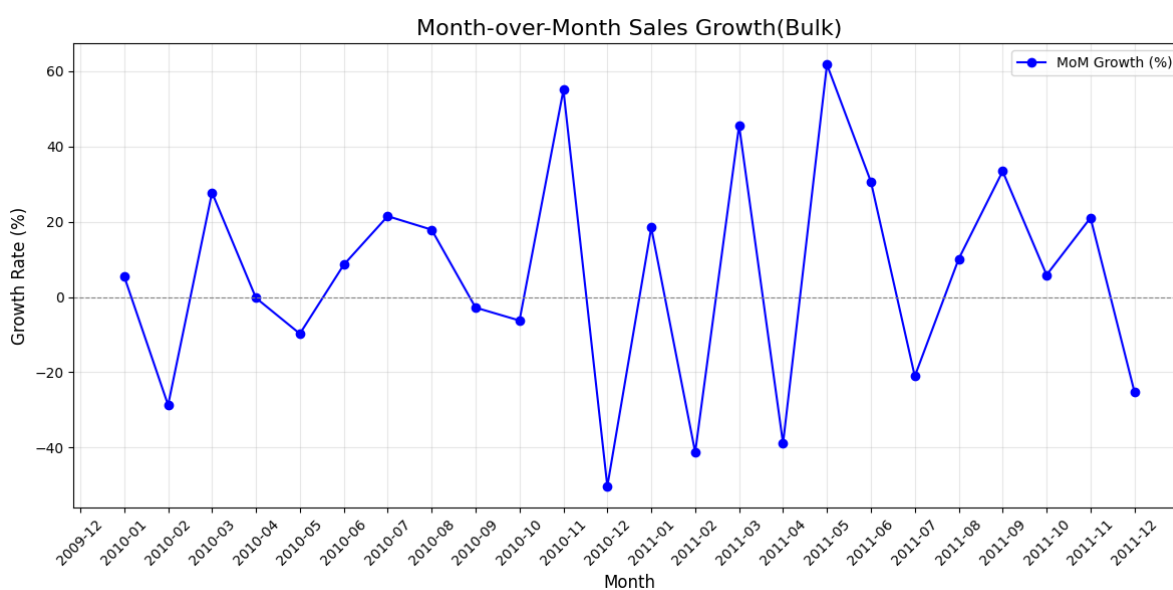


Figure 12: Croissance des ventes mensuelles en gros au fil du temps

QoQ:

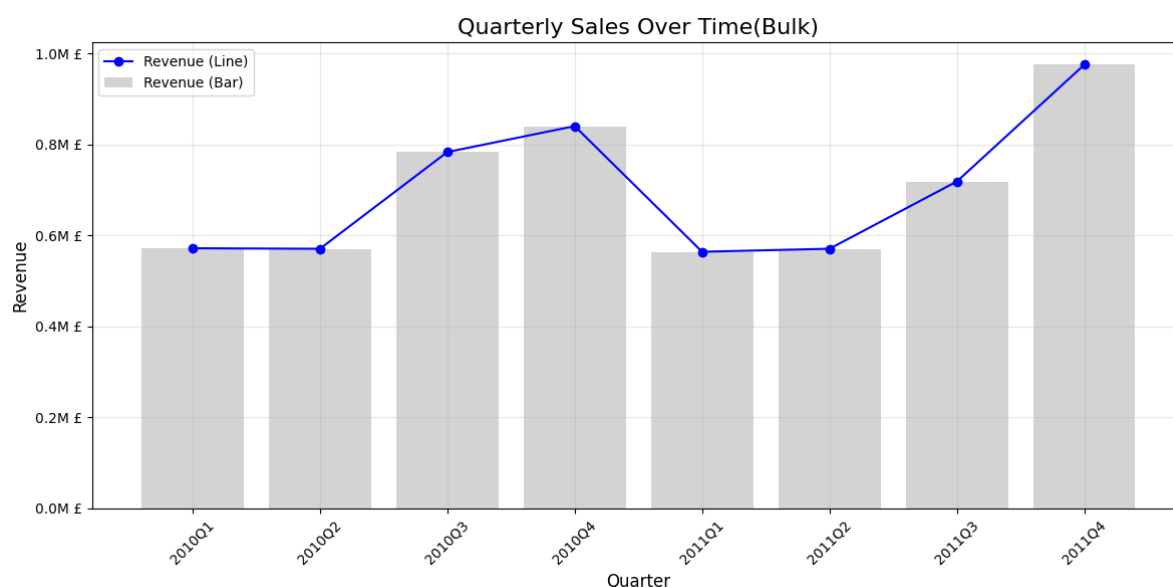


Figure 13: Ventes trimestrielles en gros au fil du temps

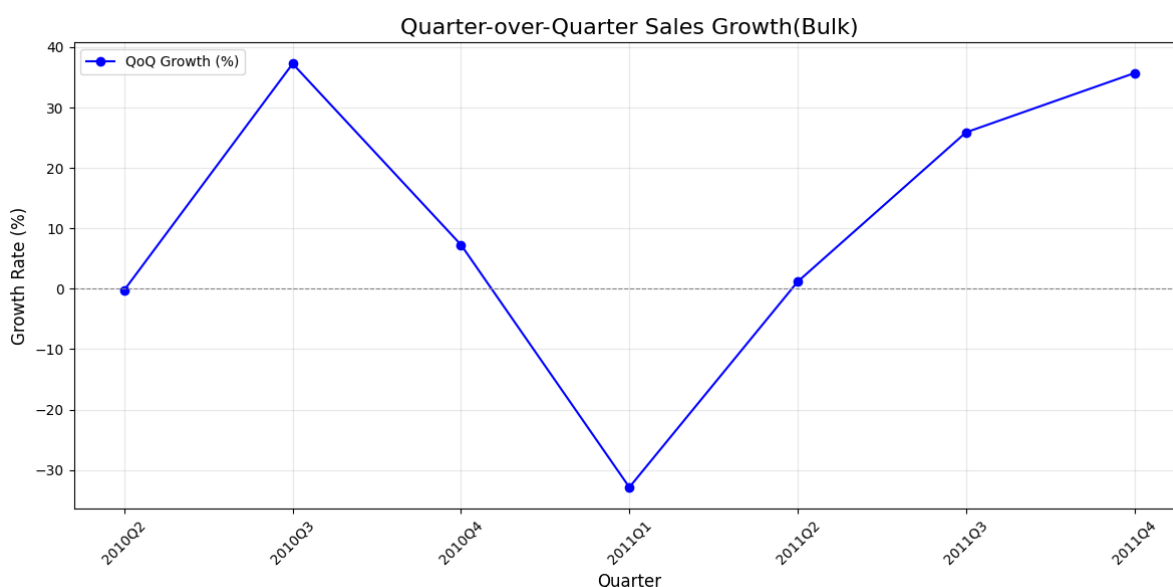


Figure 14: Croissance des ventes trimestrielles en gros au fil du temps

Les ventes en gros montrent des tendances similaires à celles des ventes au détail dans les analyses MoM et QoQ. En novembre (T4 2010), les ventes en gros connaissent une forte augmentation, suivie d'une baisse en décembre (T1 2011). De mars (T1 2011) à août (T3 2011), les ventes en gros se stabilisent, ce qui reflète un retour à une performance plus régulière. Ces tendances indiquent que, comme les ventes au détail, les transactions en gros sont saisonnières et présentent une demande stable après le pic des fêtes.

### Performance régionale:

Le sous-ensemble de données sur les ventes en gros comprend 34 pays distincts, les 10 premiers en termes de revenus générés sont :

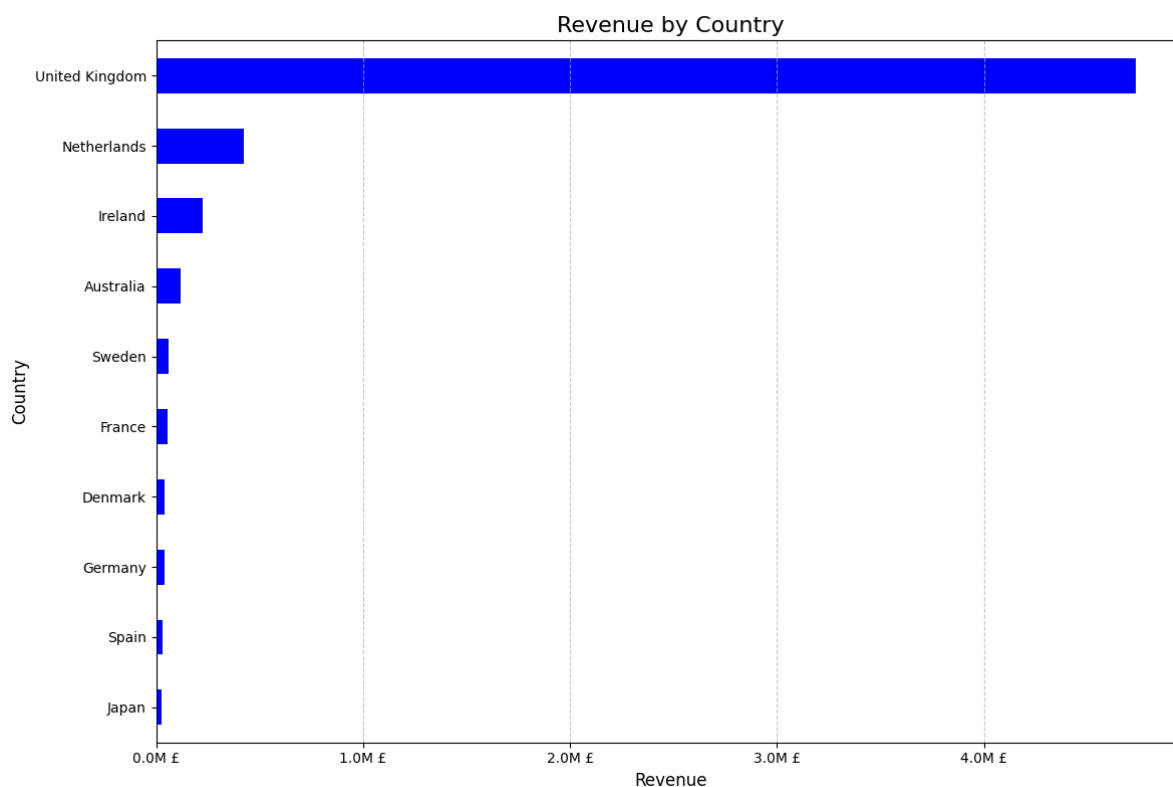


Figure 15: Top 10 des revenus par pays (Vente en gros)

Tout comme les ventes au détail, le Royaume-Uni domine les ventes en gros, avec l'Irlande et l'Allemagne également dans les premières positions. Cependant, les ventes en gros montrent une contribution plus forte des Pays-Bas, de l'Australie et de la Suède, contrairement aux ventes au détail où la France joue un rôle plus important. Cela suggère des demandes de marché différentes entre les deux catégories.

### Informations sur les produits:

Le sous-ensemble de données des ventes en gros comprend 2 691 produits distincts, les 20 premiers en termes de revenus générés sont :



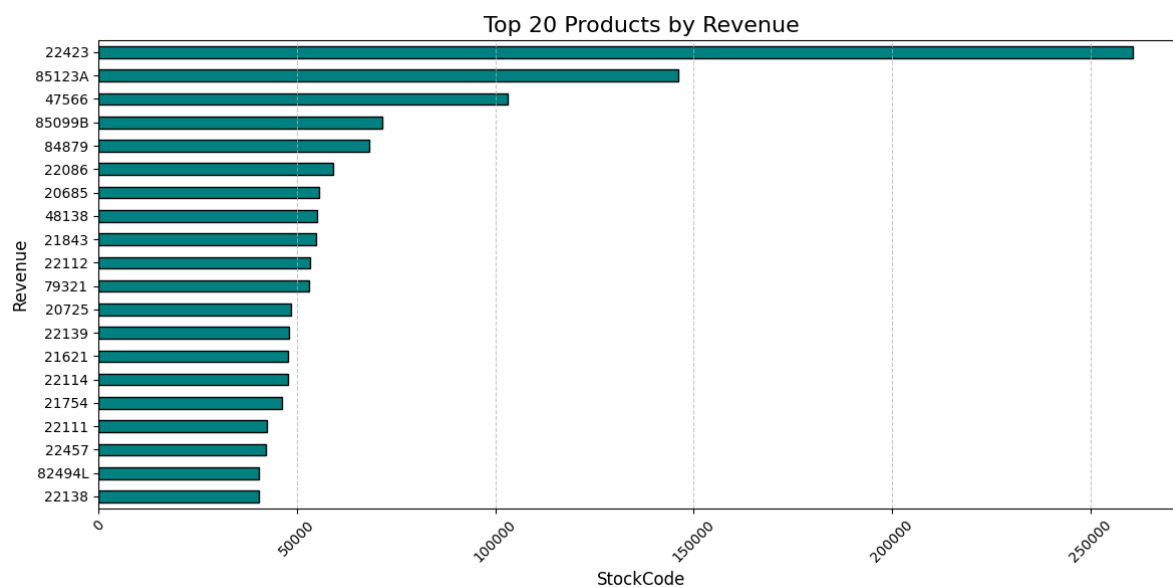


Figure 16: Top 20 des produits par revenu (Vente en gros)

Revenue Breakdown: Top 20 products vs. The rest of the products(Bulk)

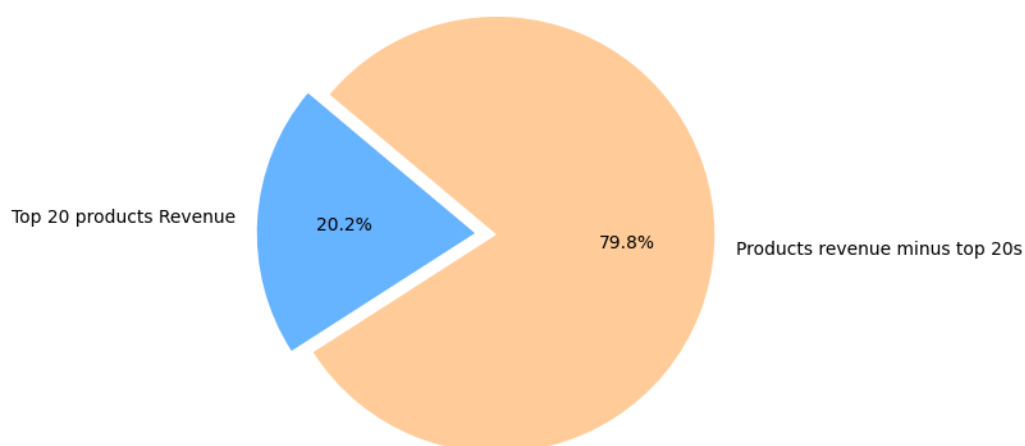


Figure 17: Répartition des revenus : Top 20 des produits vs Le reste des produits (Vente en gros)

Tout comme pour les ventes au détail, les 20 produits les plus vendus contribuent de manière significative aux revenus globaux, le produit le plus vendu (23843) générant £168,469.60

Les 20 produits les plus vendus, représentant seulement 0.7% des 2,691 produits totaux, génèrent 20.2% des revenus totaux. Cela renforce leur rôle essentiel dans la génération de revenus et souligne l'importance stratégique de se concentrer sur les produits à forte performance.

### **Résumé des informations sur les ventes en gros:**

Les ventes en gros, bien que plus petites en volume (24,350 transactions contre 970,147 pour les ventes au détail), contribuent de manière significative aux revenus totaux, représentant 28,9% du total de £20,085,863.18. Les transactions en gros montrent une valeur élevée, avec une valeur moyenne des commandes (AOV) de £682.34, ce qui reflète des achats plus importants typiques de cette catégorie.

Les tendances saisonnières dans les ventes en gros s'alignent étroitement avec celles des ventes au détail, avec un pic en novembre (T4 2010), suivi d'une stabilisation au premier semestre 2011. La performance régionale est dominée de manière similaire par le Royaume-Uni, bien que les ventes en gros connaissent une contribution plus forte des Pays-Bas, de l'Australie et de la Suède, contrairement aux ventes au détail qui reposent davantage sur la France.

La concentration des produits est évidente dans les deux catégories ; les 20 produits les plus vendus en gros, représentant seulement 0,7% des offres, génèrent 20,2% des revenus en gros, ce qui reflète l'importance disproportionnée des produits à forte performance dans les ventes au détail également. Ces tendances soulignent l'importance stratégique des ventes en gros comme source régulière et à forte valeur de revenus.

## **3.2 Analyse des Clients**

Cette section analyse le comportement des clients au sein du jeu de données en examinant des métriques clés telles que la segmentation des clients, la fréquence d'achat, la durée de vie des clients, la valeur à vie des clients, le taux de désabonnement et le taux de fidélisation. Ces informations nous aident à comprendre la fidélité des clients, les tendances d'achat et la viabilité globale de l'entreprise.

### **3.2.1 Vente au détail**

Cette sous-section se concentre sur les clients de détail avec une quantité de moins de 50 unités par produit vendu, à l'exception des achats en gros (Quantité  $\geq 50$ ) et des annulations.

#### **Aperçu de l'Analyse des Clients:**

Couvrant la période de décembre 2009 à décembre 2011, nous avons :

- **Total des Clients Identifiés:** 5 758 clients

- **Nombre total de Transactions avec un client identifié:** 743 119 transactions
- **Transactions par Clients Non Identifiés:** 227 028 transactions
- **Total des Commandes(Factures):** 39 287 commandes
- **Commandes par Clients Non Identifiés:** 4 293 commandes

#### Clients Identifiés vs Non Identifiés :

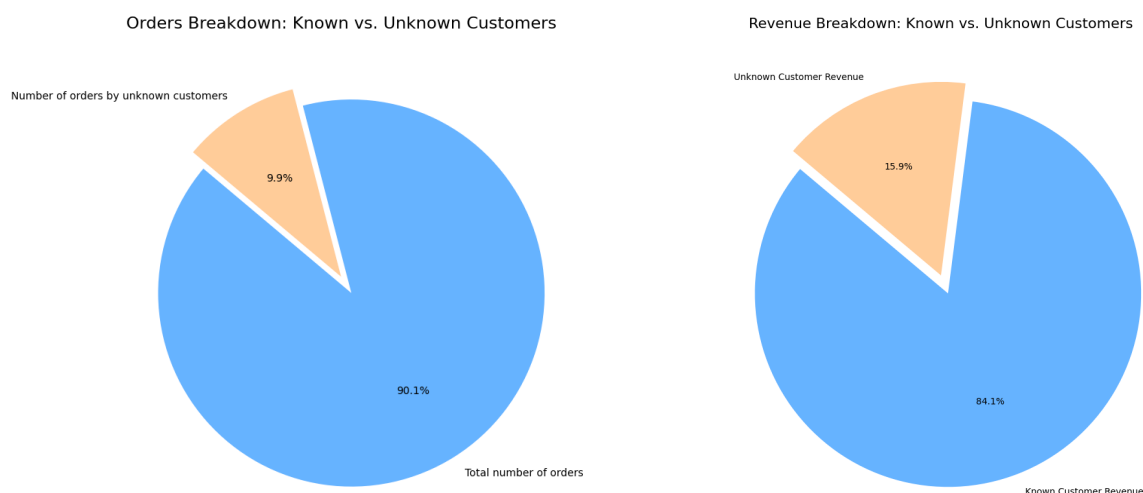


Figure 18: Répartition des Commandes : Clients Identifiés vs Non Identifiés

Figure 19: Répartition des Revenus : Clients Identifiés vs Non Identifiés

Le fait que 15,9% des revenus totaux de détail et 9,9% des commandes proviennent de clients non identifiés soulève une préoccupation concernant la proportion importante de revenus issus de clients non identifiables. Cette lacune souligne l'importance d'améliorer les processus d'identification des clients, car comprendre ces clients pourrait révéler des informations précieuses et des opportunités pour des stratégies ciblées et l'établissement de relations.

#### Top Clients:

Les 20 premiers clients en termes de revenus générés sont :

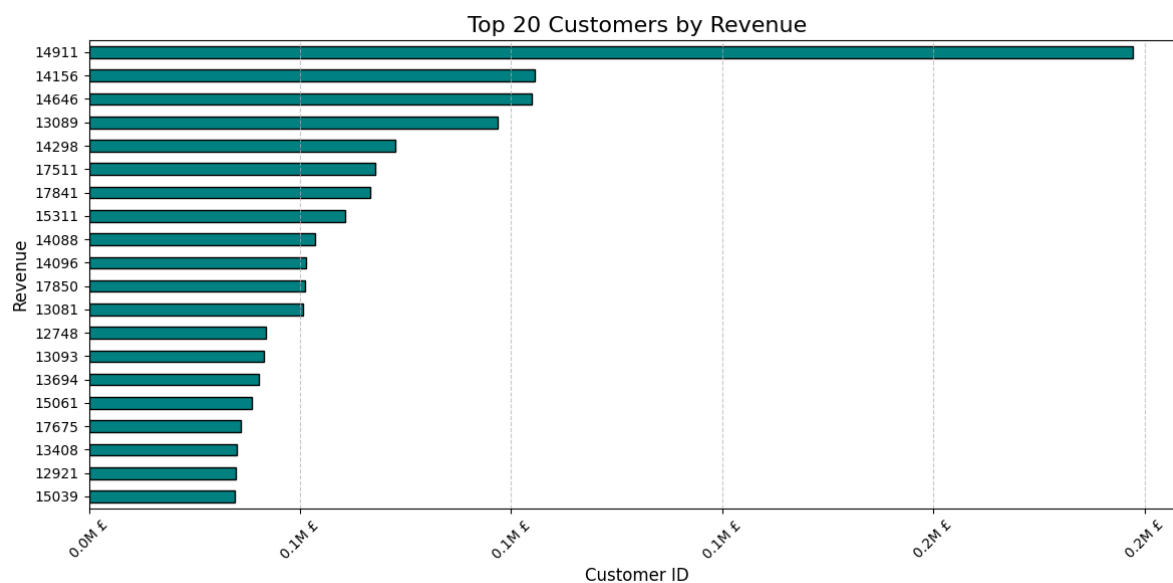


Figure 20: Top 20 Clients par Revenus

## Revenue Breakdown: Top 20 customers vs. The rest of the customers

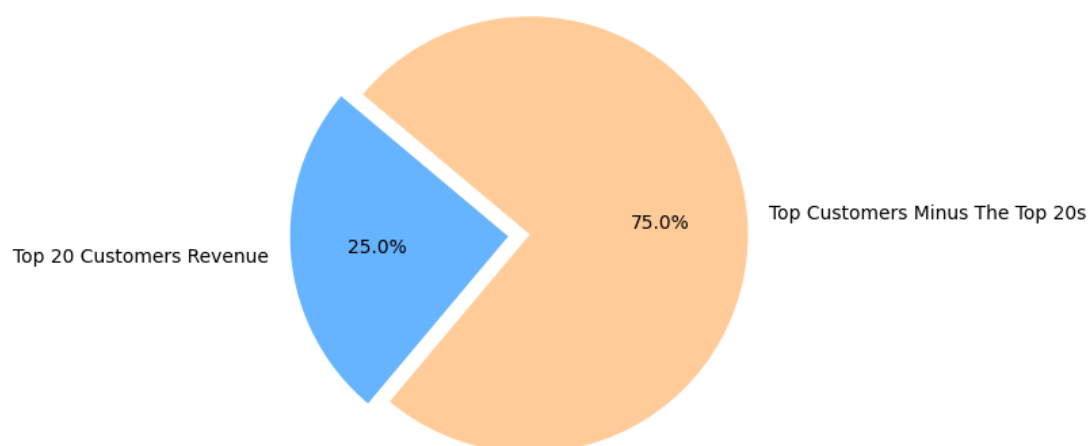


Figure 21: Répartition des Revenus : Top 20 Clients vs Le reste des Clients

Les 20 premiers clients génèrent collectivement une portion importante des revenus totaux (25%), soulignant leur rôle clé dans la croissance de l'entreprise.

### Fréquence d'Achat:

Avec une fréquence d'achat moyenne de **6,82** sur la période de décembre 2009 à décembre 2011, les clients effectuent près de 7 achats ou environ 3,41 achats par an en moyenne.

### Durée de Vie des Clients:

La durée de vie moyenne des clients est de **0,747** ans, ce qui indique qu'en moyenne, les clients s'engagent avec l'entreprise pendant environ 9 mois avant que leur activité d'achat se diminue ou cesse.

### Valeur à Vie du Client (CLV):

Une valeur à vie du client (CLV) de **£1 853,44** indique qu'en moyenne, chaque client génère £1 853,44 de revenus tout au long de leur relation avec l'entreprise.

### Taux de Désabonnement et de Fidélisation:

#### MoM:

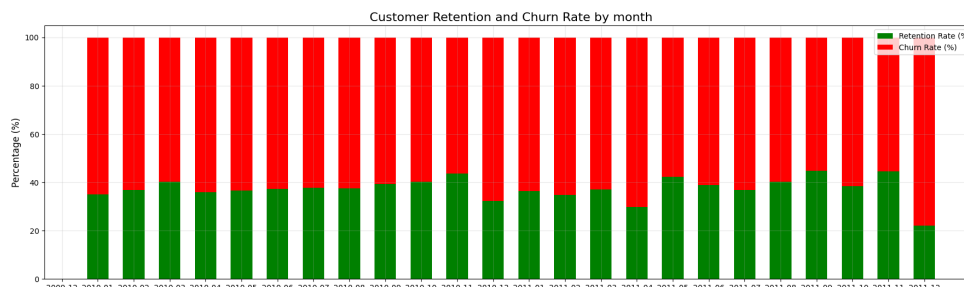


Figure 22: Taux de Fidélisation et Taux de Désabonnement par mois

L'analyse du taux de fidélisation révèle des fluctuations au cours de la période observée, avec des pics notables (par exemple, 44,7% en septembre 2011) et des baisses (par exemple, 22,0% en décembre 2011). Bien que certaines variabilités soient présentes, une baisse de la fidélisation vers la fin de la période n'est pas rare pour les entreprises de commerce en ligne, notamment pendant les transitions saisonnières ou après les périodes de forte consommation, comme le montrent les premières informations sur les ventes mois après mois. Le taux de désabonnement, qui est le complément du taux de fidélisation, met en évidence le nombre de clients perdus chaque mois.

QoQ:

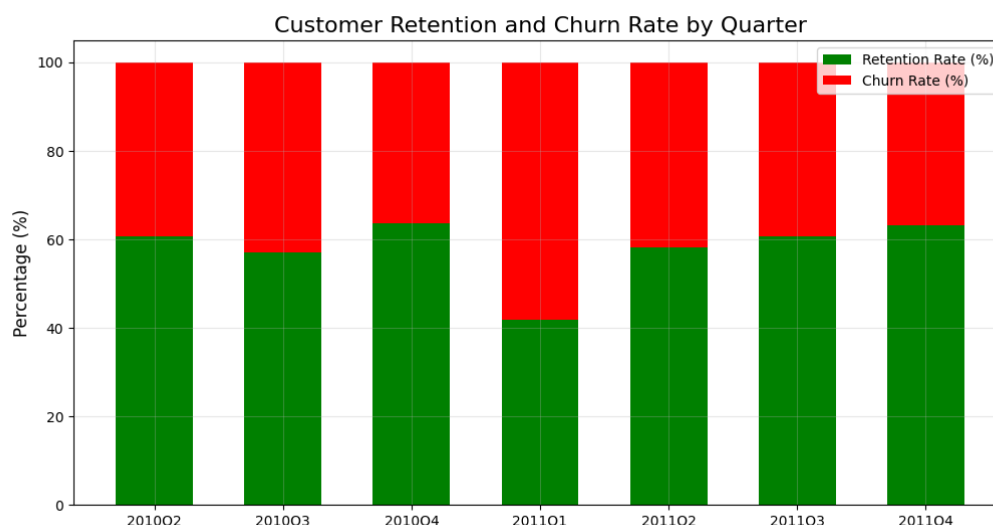


Figure 23: Taux de Fidélisation et Taux de Désabonnement par trimestre

Les taux de fidélisation et de désabonnement trimestriels montrent plus de stabilité par rapport aux données mensuelles, ce qui reflète les comportements des clients à plus long terme. Notamment, les taux de fidélisation sont plus élevés chaque trimestre, en particulier au T4 de 2010 et 2011, ce qui suggère une plus grande fidélité des clients pendant les pics saisonniers. La baisse de la fidélisation au T1 de 2011 indique une déconnexion typique des clients après la saison.

Dans l'ensemble, les données trimestrielles soulignent l'importance des tendances à long terme dans la compréhension de la fidélisation des clients, avec un engagement plus constant chaque trimestre par rapport aux fluctuations mensuelles.

### Résumé des Informations Clients:

L'analyse des clients révèle des informations importantes sur le comportement des clients. Les clients identifiés, au total 5 758, ont généré la majorité des transactions. Cependant, les clients non identifiés représentaient 15,9% des revenus et 9,9% des commandes. Cela souligne la nécessité d'améliorer les processus d'identification des clients. Les 20 premiers clients, responsables d'un quart des revenus de détail, jouent un rôle essentiel dans la croissance de l'entreprise.

Avec une fréquence d'achat moyenne de 6,82, bien au-dessus de la norme de l'industrie, les clients montrent un engagement fort. Cependant, la durée de vie moyenne du client de 9 mois et une CLV de £1 853,44 suggèrent qu'il est crucial de retenir les clients à long terme pour maximiser leur valeur.

Les taux de fidélisation fluctuent mensuellement, avec certains mois tombant en dessous de 30%, bien que des taux de fidélisation supérieurs à ce seuil soient considérés comme solides pour le commerce en ligne. Les données trimestrielles sur la fidélisation

montrent plus de stabilité et une plus grande fidélité des clients, en particulier pendant les pics saisonniers, ce qui souligne l'importance des stratégies de fidélisation des clients à long terme.

### 3.2.2 Vente en gros

Cette sous-section se concentre sur les clients de gros, ayant acheté plus de 50 unités par produit vendu.

#### Aperçu de l'Analyse des Clients:

Pour la période de décembre 2009 à décembre 2011, nous avons :

- **Total des Clients De Gros Identifiés:** 1 971 Clients
- **Nombre total de Transactions De Gros avec un client identifié:** 23 013 transactions
- **Transactions De Gros par Des Clients Non Identifiés:** 1 337 transactions
- **Commandes De Gros par Des Clients Non Identifiés (Factures incluant des Commandes De Gros):** 696 commandes

#### Clients Identifiés vs Non Identifiés :

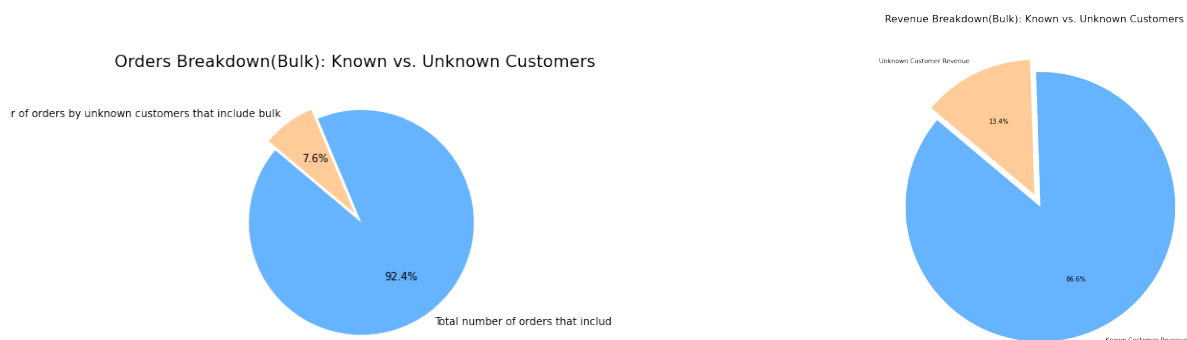


Figure 25: Répartition des Revenus (Gros): Clients Identifiés vs Non Identifiés

Figure 24: Répartition des Commandes (Gros): Clients Identifiés vs Non Identifiés

Dans le segment de gros, 13,4% des revenus totaux et 7,6% des commandes proviennent de clients non identifiés, ce qui constitue une proportion plus faible que dans le détail. Bien que la problématique soit moins marquée dans le gros, elle met en évidence un écart notable dans l'identification des clients. S'attaquer à cette question pourrait offrir des opportunités pour mieux comprendre et engager ces clients non identifiés, ce

qui permettrait potentiellement de découvrir des stratégies pour augmenter les revenus et renforcer les relations commerciales dans le segment de gros.

### Top Clients:

Les 20 premiers clients en termes de revenus générés par le gros sont :

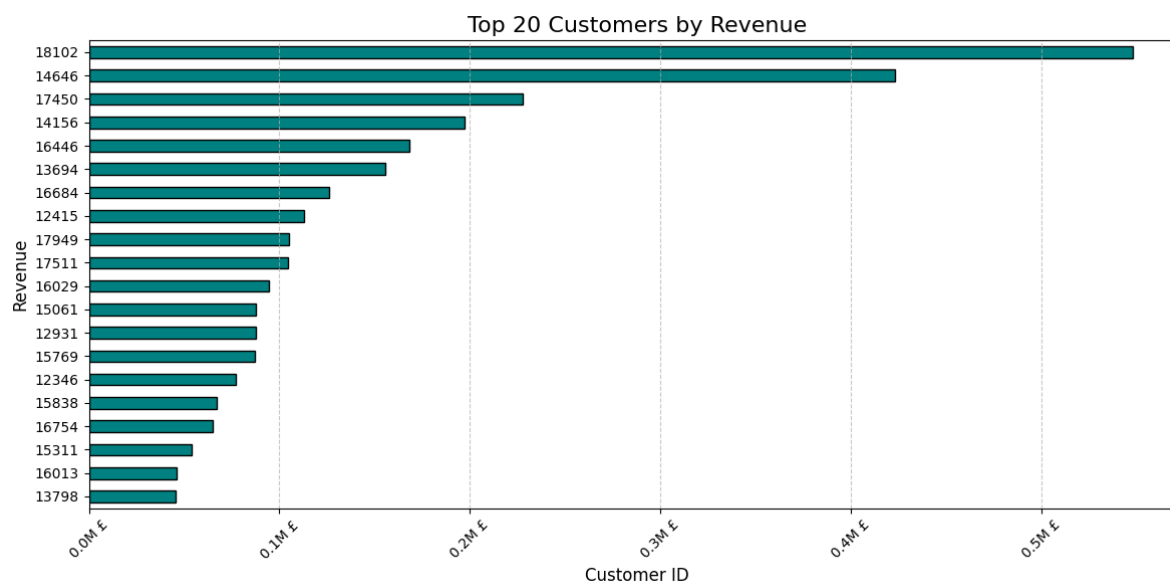


Figure 26: Top 20 Clients par Revenus (Gros)

Revenue Breakdown(Bulk): Top 20 customers vs. The rest of the customer

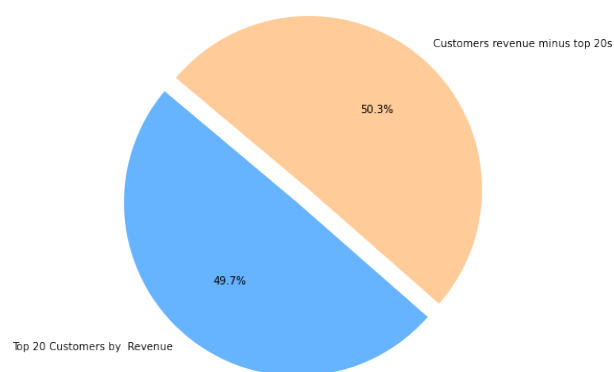


Figure 27: Répartition des Revenus (Gros) : Top 20 Clients vs Le Reste des Clients

Dans le segment détail, les 20 premiers clients contribuent de manière significative aux revenus, mais ne dominent pas l'ensemble des affaires, ce qui met en lumière une base de clients plus diversifiée. En revanche, le segment bulk dépend fortement de ses 20 premiers



clients, qui représentent près de la moitié (49,7%) des revenus totaux. Cela indique une structure de revenus plus concentrée dans les opérations bulk, soulignant l'importance cruciale de fidéliser et entretenir ces clients clés pour maintenir la stabilité des revenus.

### Fréquence des Achats:

Dans le segment de gros, la fréquence moyenne des achats est de **4,31**, ce qui est nettement plus faible par rapport à celle du segment détail, **6,82**. Cela indique que les clients de gros interagissent moins fréquemment avec l'entreprise durant la période observée.

### Durée de Vie du Client (Lifespan):

Les clients de gros ont une durée de vie moyenne plus courte d'environ **0,477 ans** (un peu moins de 6 mois), contre **0,747 ans** pour les clients de détail (environ 9 mois).

### Valeur à Vie du Client (CLV):

La valeur à vie du client (CLV) dans le segment de gros est de **£1 404,90**, inférieure à celle du segment détail de **£1 853,44**. Cela indique que les clients de gros génèrent moins de revenus au cours de leur relation avec l'entreprise.

### Taux de Désabonnement et de Fidélisation::

#### MoM :

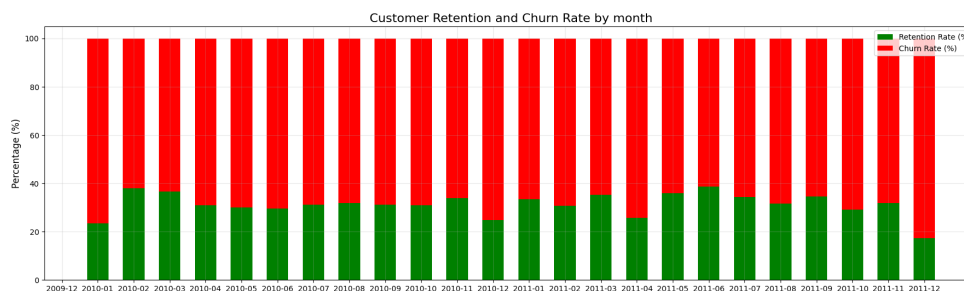


Figure 28: Taux de Désabonnement et de Fidélisation par mois (Gros)

Les taux de rétention bulk varient de 17,24% à 38,71%, avec des taux de churn atteignant jusqu'à 82,76%. Comparé aux taux de rétention plus élevés et stables du détail (40-50%), le segment bulk montre une plus grande volatilité et une base de clients moins fidèle, comme l'illustre le graphique en barres empilées.

QoQ :

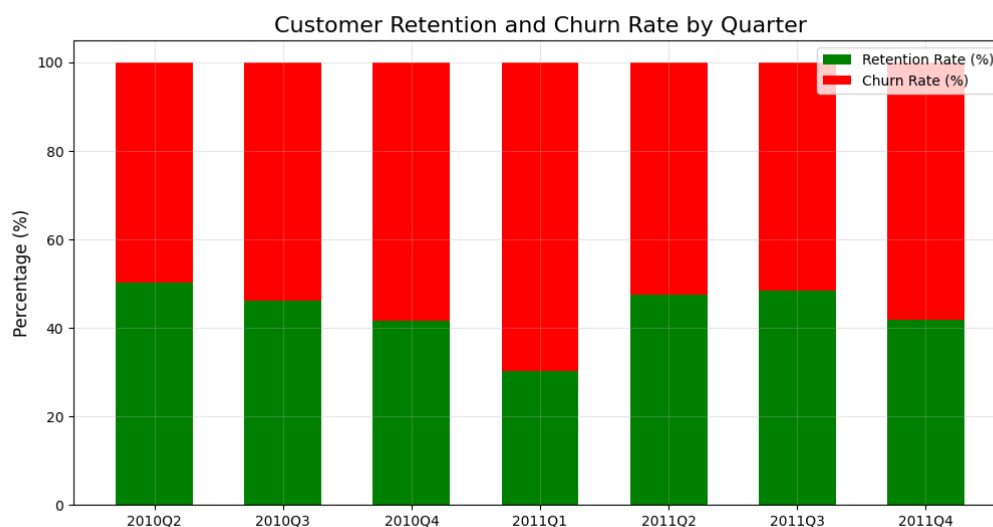


Figure 29: Taux de Désabonnement et de Fidélisation par trimestre (Gros)

Comparés aux chiffres mensuels, les taux de fidélisation trimestriels du segment Gros (de 30,29% à 50,25%) et les taux de Désabonnement (de 49,75% à 69,71%) montrent une plus grande stabilité, atténuant les fluctuations mensuelles. En revanche, les taux de fidélisation trimestriels du détail sont constamment plus élevés (de 47,95% à 62,93%), ce qui reflète une base de clients plus fidèle, tandis que les taux de désabonnement sont plus faibles (de 37,07% à 52,05%), mettant en évidence un engagement plus fort des clients dans le segment détail. Cette comparaison souligne la stabilité relative des métriques du segment de gros au niveau trimestriel, mais aussi l'importance d'une dynamique de relation client plus solide dans le segment de détail.

### Résumé des Insights Clients :

Le segment des clients bulk présente des caractéristiques uniques par rapport au détail, avec 1 971 clients identifiés générant 23 013 transactions durant la période observée. Les clients non identifiés représentent 13,4% des revenus, une proportion plus faible que dans le détail, mais suffisamment significative pour justifier des efforts d'amélioration de l'identification des clients. Les revenus du segment bulk sont fortement concentrés, avec près de 50% générés par les 20 premiers clients, ce qui met en évidence la nécessité cruciale de fidéliser et de nourrir ces comptes clés.

Les clients bulk interagissent moins fréquemment que les clients de détail, avec une fréquence d'achat plus faible (4,31 contre 6,82) et une durée de vie moyenne plus courte (6 mois contre 9 mois). Les taux de rétention dans le segment bulk sont également plus faibles et plus volatils, variant de 17,24% à 38,71%, contre 40-50% pour le détail. Cependant, les métriques trimestrielles montrent une plus grande stabilité des taux de rétention et de churn dans le segment bulk, en atténuant les fluctuations mensuelles.

Malgré ces défis, les clients bulk restent un segment précieux, avec une valeur à vie de £1 404,90. Des efforts stratégiques pour améliorer la fidélité, élargir l'engagement et réduire la dépendance des revenus vis-à-vis des clients clés pourraient améliorer de manière significative les performances à long terme.

### 3.3 Analyse des Annulations

Cette section se penche sur l'analyse des annulations dans les ensembles de données de vente au détail et en gros, en examinant les taux d'annulation, les tendances au fil du temps et leur impact sur les ventes. Les informations sur les produits annulés et les clients responsables des taux d'annulation les plus élevés fourniront un contexte précieux pour améliorer les stratégies opérationnelles.

#### 3.3.1 Vente au Détail

Cette sous-section se concentre sur les annulations dans le secteur de la vente au détail, avec une quantité inférieure à 50 unités par produit, en offrant des perspectives sur les tendances des annulations et leur impact.

##### **Vue d'ensemble des Annulations en Vente au Détail:**

Couvrant la période de décembre 2009 à décembre 2011, nous avons :

- **Perte de Revenus Totale des Annulations en Vente au Détail** : £256,097.23
- **Nombre de Transactions Annulées en Vente au Détail** : 17 219
- **Nombre de Commandes Annulées en Vente au Détail (Factures)** : 7 216

**Annulations vs Ventes :**

Invoice Breakdown: Sold orders vs. Cancelled orders

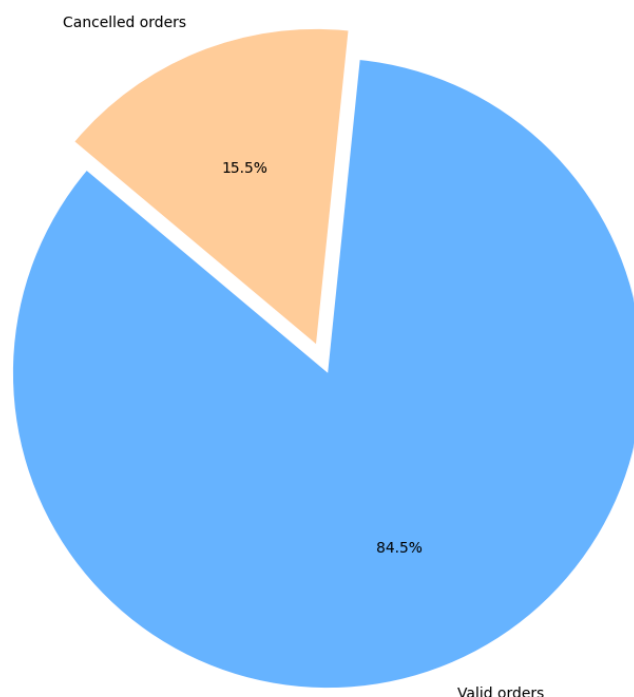


Figure 30: Répartition des Factures : Commandes VENDUES vs Commandes ANNULÉES

Bien que le taux brut d'annulation puisse sembler notable, il est important de considérer le contexte. Le nombre moyen d'articles par commande annulée est significativement plus bas que celui des commandes en vente au détail, ce qui suggère une dynamique transactionnelle différente. Cette observation invite à une analyse plus approfondie de la répartition des annulations et de leur influence sur la performance globale des ventes.

**Taux d'Annulation :** Le taux brut d'annulation de 15.51% fournit une première perspective sur l'impact des annulations ; cependant, il ne capture pas pleinement les dynamiques transactionnelles en raison de la différence dans le nombre moyen d'articles par commande entre les commandes annulées et les ventes au détail. Les commandes annulées contiennent généralement moins d'articles, avec une moyenne de 2.39 articles par annulation contre 24.69 articles par commande en vente au détail.

Pour ajuster cette différence et offrir un taux d'annulation plus précis, un facteur de correction a été appliqué en multipliant le taux brut par le rapport du nombre moyen d'articles par commande annulée au nombre moyen d'articles par commande en vente au détail. Cet ajustement assure une représentation plus significative des annulations dans le contexte des ventes globales, offrant ainsi une vue plus claire de l'impact réel sur les revenus et le comportement des clients.

$$\text{Taux d'Annulation Ajusté} = \text{Taux d'Annulation} \times \left( \frac{\text{Transaction Moyenne par Commande Annulée}}{\text{Transaction Moyenne par Commande en Vente au Détail}} \right)$$

Breakdown(adjusted to the avrage basket size): Sold orders vs. Cancell

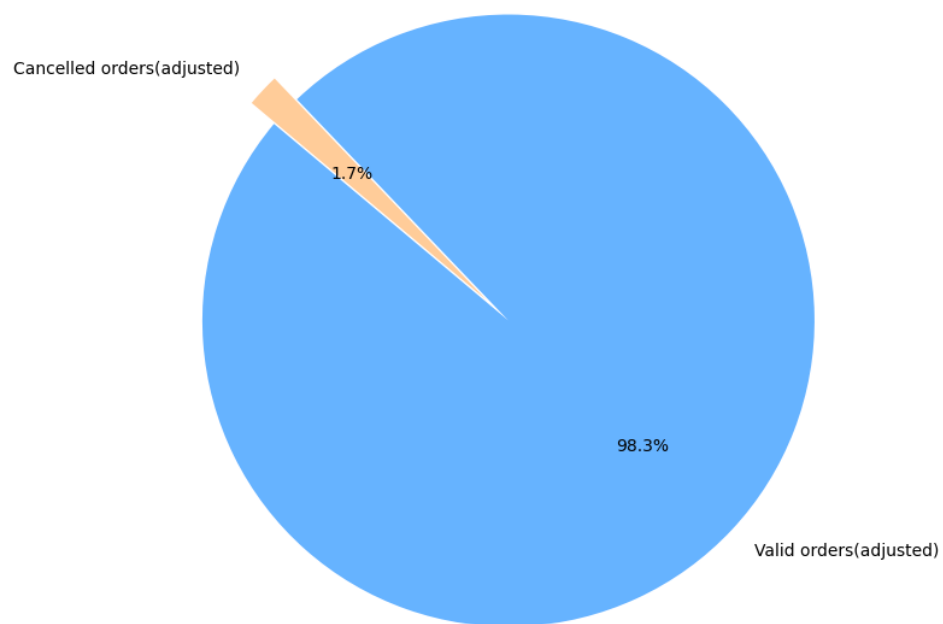


Figure 31: Répartition des Factures (ajustée à la taille moyenne du panier) : Commandes VENDUES vs Commandes ANNULÉES

## Annulations au Fil du Temps:

## MOM:

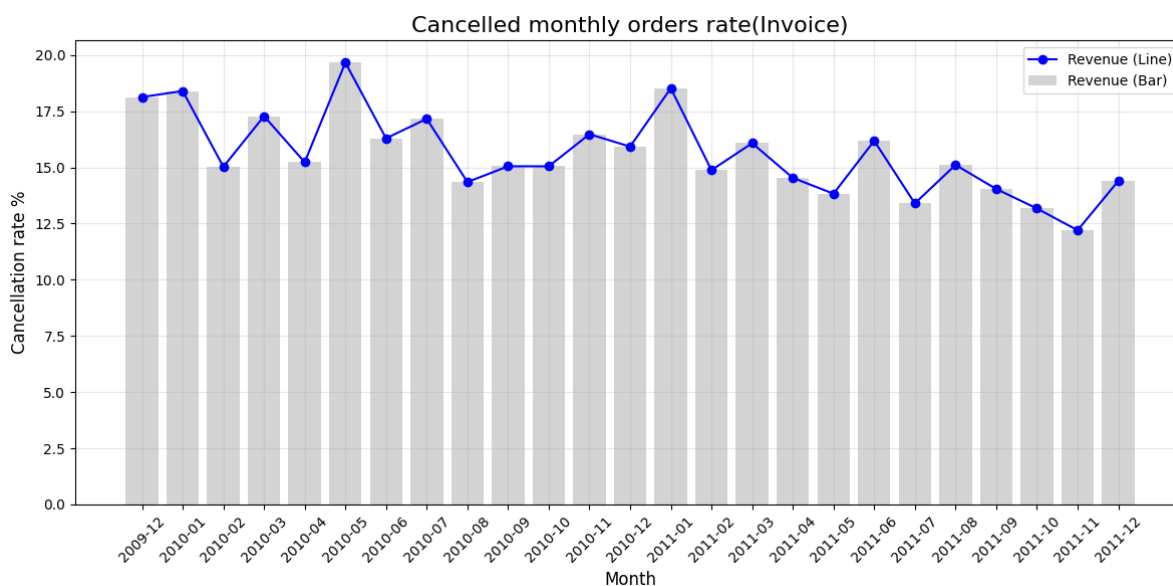


Figure 32: Taux des commandes annulées mensuelles (Factures)

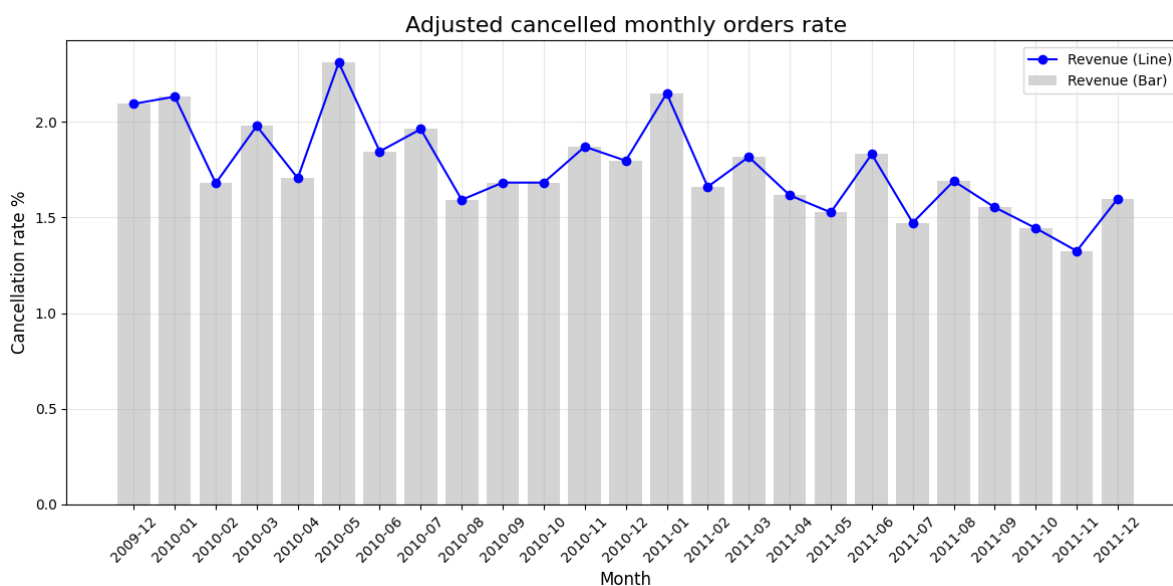


Figure 33: Taux des commandes annulées mensuelles ajustées (Factures)

Les taux mensuels d'annulation reflètent des fluctuations dans le pourcentage de commandes annulées, avec une lente diminution globale au fil du temps, un pic notable en mai 2010 à 19.66% et un point bas en novembre 2011 à 12.20%. En ajustant les taux, le modèle d'annulation suit la même tendance, avec des taux nettement plus bas.

QoQ:

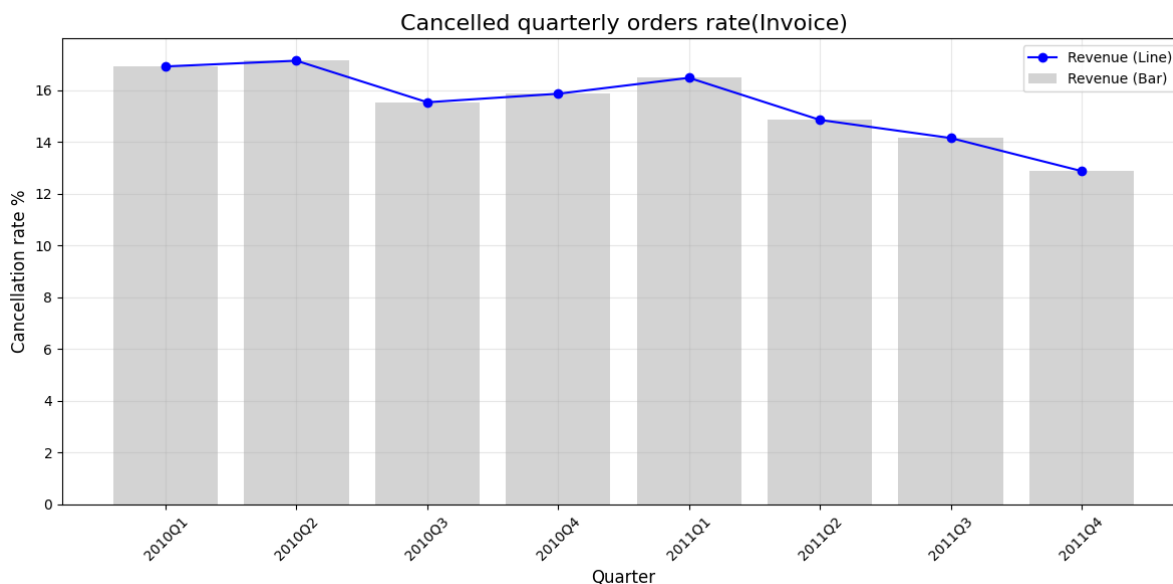


Figure 34: Taux des commandes annulées trimestrielles (Factures)

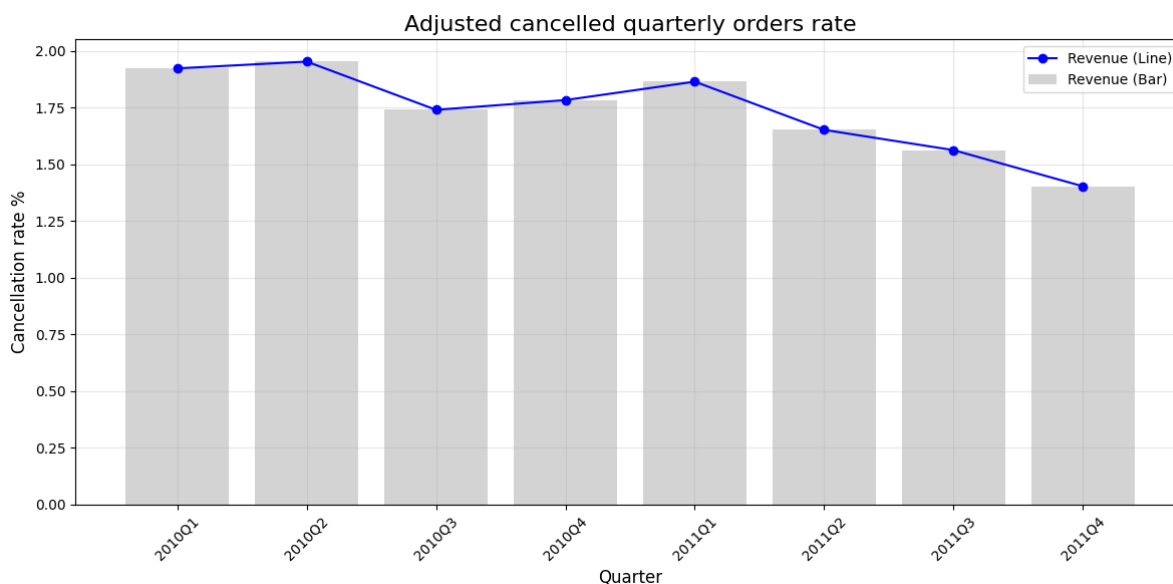


Figure 35: Taux des commandes annulées trimestrielles ajustées (Factures)

Tout comme les annulations mensuelles, les taux d'annulation trimestriels montrent une tendance généralement décroissante, passant de 16.91% au T1 2010 à 12.87% au T4 2011, avec un pic au T2 2010 à 17.13%. Cela indique que les annulations étaient plus

fréquentes dans les trimestres antérieurs, diminuant progressivement vers la fin de l'année 2011. En ajustant les taux, le modèle d'annulation suit la même tendance avec des taux considérablement plus bas. La baisse des taux bruts et ajustés suggère une amélioration progressive du comportement d'achat des clients au fil du temps.

### Produits les Plus Annulés:

Les 20 produits les plus annulés sont :

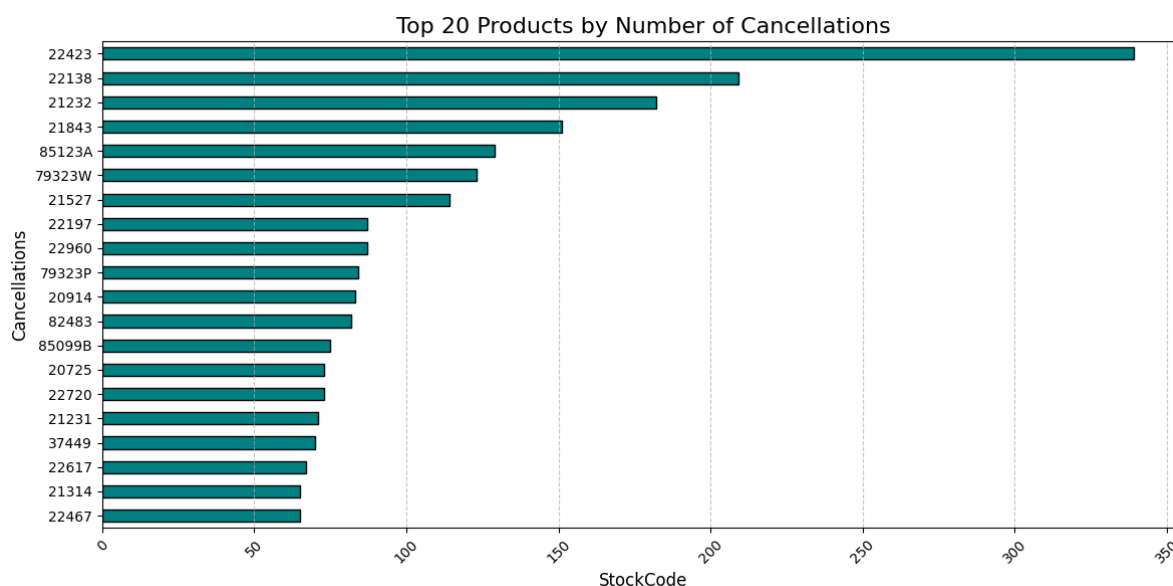


Figure 36: Top 20 Produits par Nombre d'Annulations

Tandis que les produits ayant les taux d'annulation les plus élevés sont : Certaines

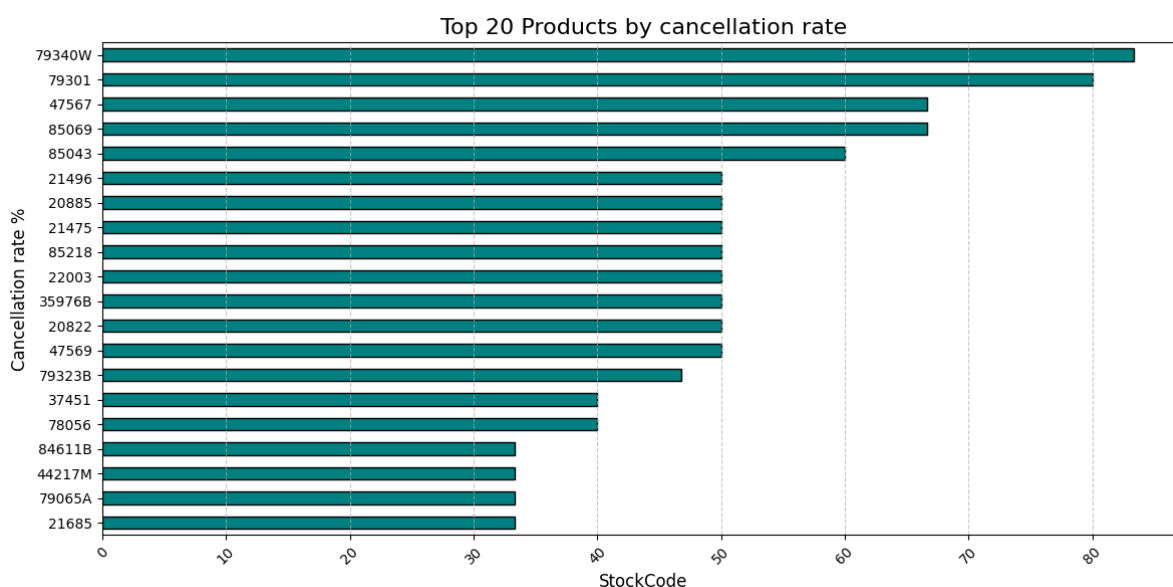


Figure 37: Top 20 Produits par Taux d'Annulation



articles comme 79340W ou 79301 (taux d'annulation de 83.33% et 80%) montrent des taux disproportionnellement élevés, signalant des problèmes potentiels avec ces produits.

### Clients les Plus Annulants:

Les 20 clients ayant annulé le plus sont :

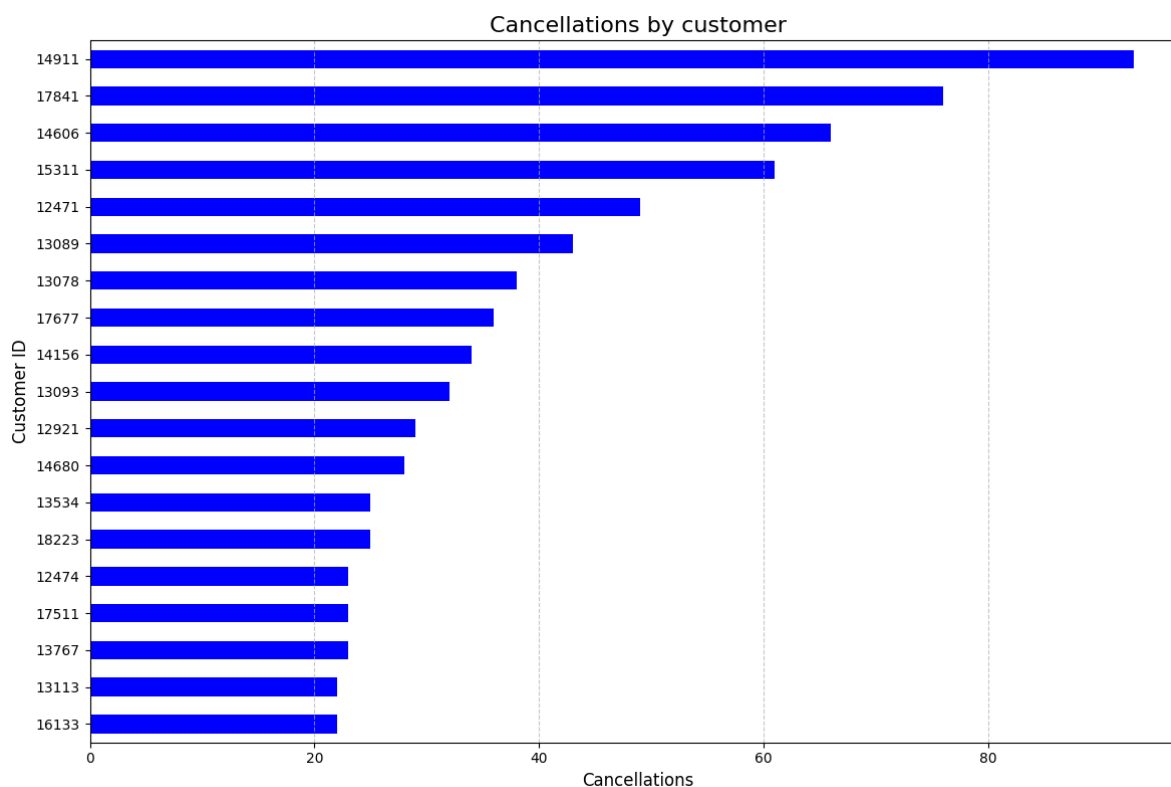


Figure 38: Top 20 Annulations par Client

Tandis que les clients avec les taux d'annulation les plus élevés sont :

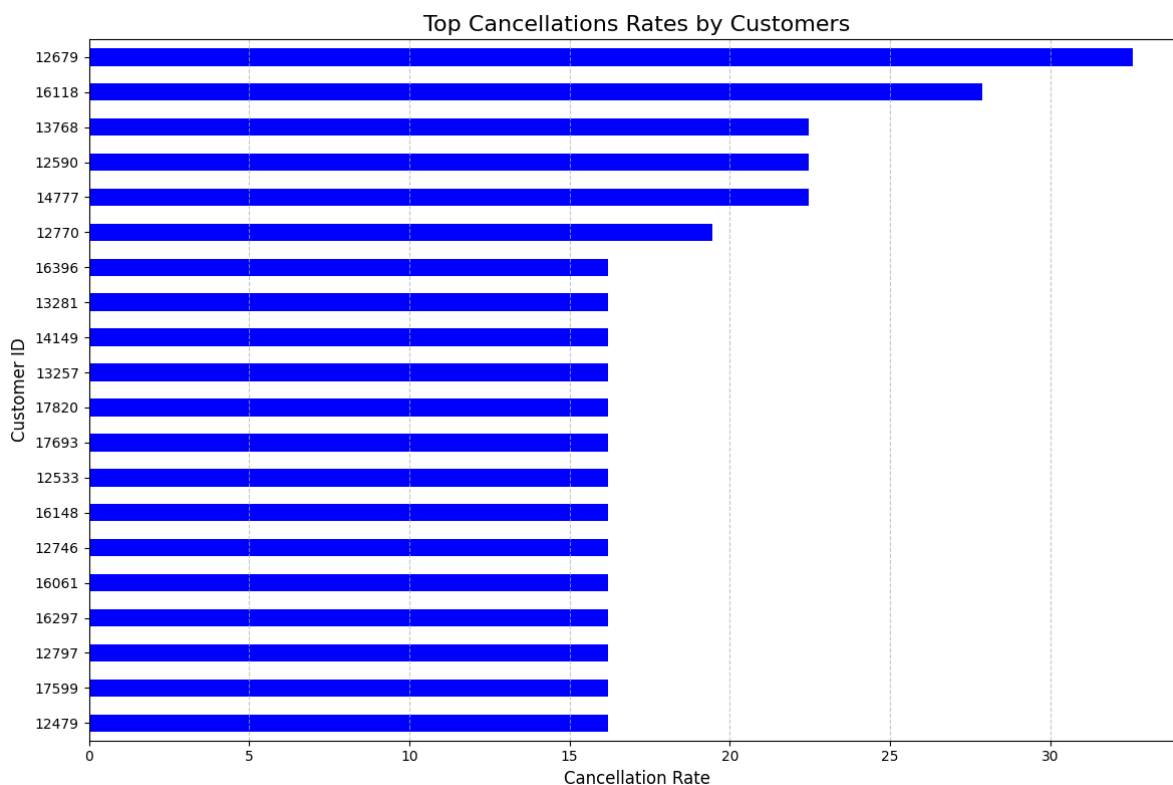


Figure 39: Top 20 Taux d'Annulations par Client

Les 20 clients les plus annulants montrent que quelques clients, comme le Client ID 14911, ont un nombre élevé d'annulations. Certains clients avec un taux élevé d'annulations, comme le Client ID 12679, ont également des taux d'annulation remarquablement élevés, ce qui indique des problèmes potentiels à résoudre pour réduire les annulations.

#### **Perte de Revenus due aux Annulations:**

La perte totale de revenus due aux annulations en vente au détail s'élève à £256,097.23, ce qui représente 1.8% du revenu total de la vente au détail. Bien que cela semble une proportion modeste, cela souligne l'impact que les annulations peuvent avoir sur la performance globale des ventes.

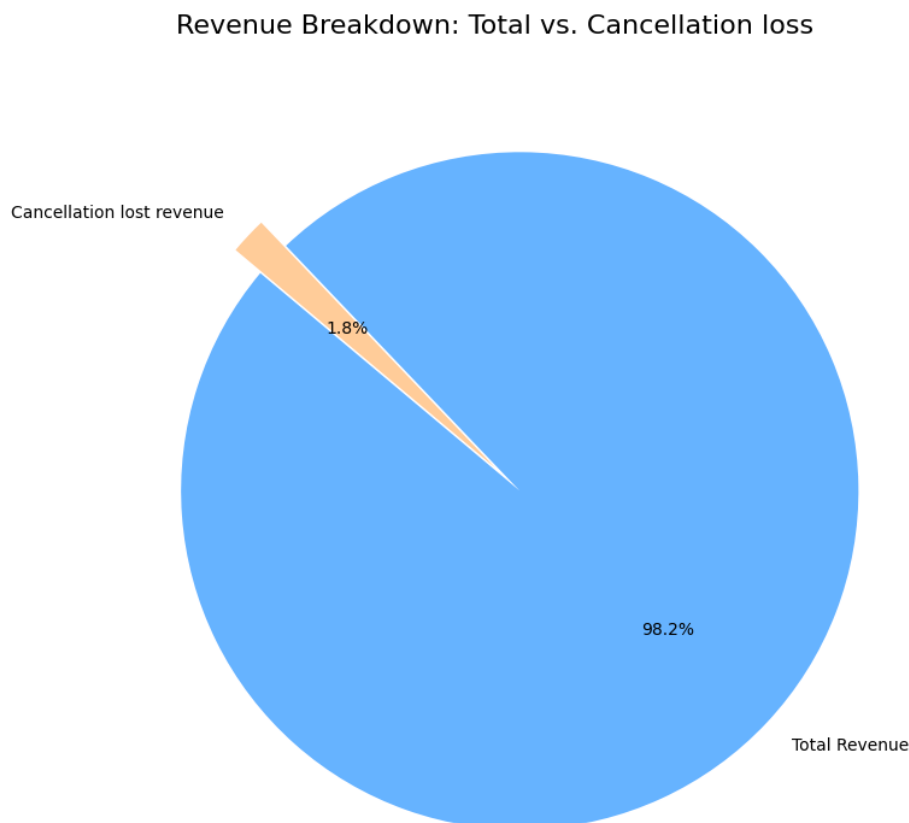


Figure 40: Répartition des Revenus : Total vs Perte liée aux Annulations

### Résumé des Résultats sur les Annulations en Vente au Détail:

L'analyse des annulations dans le secteur de la vente au détail révèle une perte de revenu raisonnable, avec un total de 256 097,23 £ perdu à cause de 17 219 transactions annulées sur une période de deux ans. Bien que le taux brut d'annulation soit de 15,51%, il est ajusté pour tenir compte de la taille moyenne plus petite des paniers des commandes annulées (2,39 articles) par rapport aux commandes de vente au détail (24,69 articles). Cet ajustement, combiné au ratio de perte de revenu, offre une image plus précise des effets réels des annulations.

Au fil du temps, le taux d'annulation a fluctué mais a généralement diminué.

Les produits et clients les plus annulés mettent en évidence des domaines spécifiques nécessitant une attention particulière, tels que des produits ayant des taux d'annulation exceptionnellement élevés comme 79340W et 79301. Bien que les annulations représentent 1,8% du revenu total des ventes au détail, traiter les problèmes sous-jacents pourrait aider à limiter davantage les pertes de revenus.

Les produits et clients les plus annulés sont identifiés, et les tendances des annulations révèlent une réduction progressive du taux d'annulation au fil du temps. L'analyse suggère qu'une gestion plus rigoureuse des annulations pourrait améliorer les revenus potentiels.

### 3.3.2 Ventes en gros

Cette sous-section se concentre sur les annulations en gros, avec une quantité de plus de 50 unités par produit vendu, fournissant des informations sur les tendances et l'impact des annulations en gros.

#### Aperçu des Annulations en Gros:

Couvrant la période de décembre 2009 à décembre 2011, nous avons :

- **Revenu perdu dû aux Annulations en Gros:** £456,833.69
- **Nombre de Transactions Annulées en Gros:** 631
- **Nombre de Commandes Annulées en Gros (Facture):** 275

#### Annulations vs. Ventes (Gros):

Invoice Breakdown(Bulk): Valid orders vs. Cancelled orders

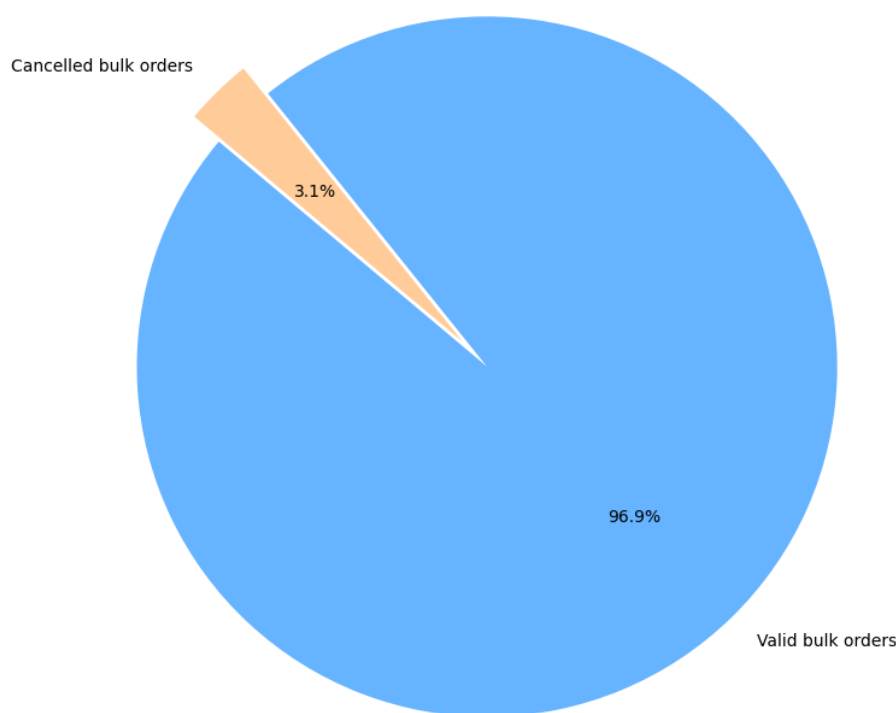


Figure 41: Répartition des Factures (Gros): Commandes vendues vs Commandes annulées

Pour les annulations en gros, il y a eu 275 annulations contre 7,216 dans le secteur de la vente au détail, ce qui indique un volume beaucoup plus faible dans la catégorie des gros. Cependant, en raison des ventes en plus grande quantité, ces annulations peuvent

toujours avoir un impact notable sur le revenu.

### Taux d'annulation:

Les taux d'annulation pour les commandes en gros montrent un modèle différent par rapport aux annulations au détail. Le taux brut d'annulation pour les commandes en gros est de 3.13%, ce qui est inférieur au taux d'annulation au détail de 15.51%. Cependant, lorsqu'on l'ajuste pour la taille des transactions (nombre moyen d'articles par commande), le taux d'annulation en gros chute à 2.29%, contre 1.7%, indiquant un modèle plus stable lorsque l'on considère les différences dans le volume des commandes. Cela suggère que les commandes en gros sont moins impactées par les annulations que les commandes au détail, probablement en raison de tailles de paniers plus grandes ou d'un comportement client différent, mettant en lumière des dynamiques distinctes dans les tendances d'annulation entre les deux catégories.

### Annulations au Fil du Temps:

#### MOM:

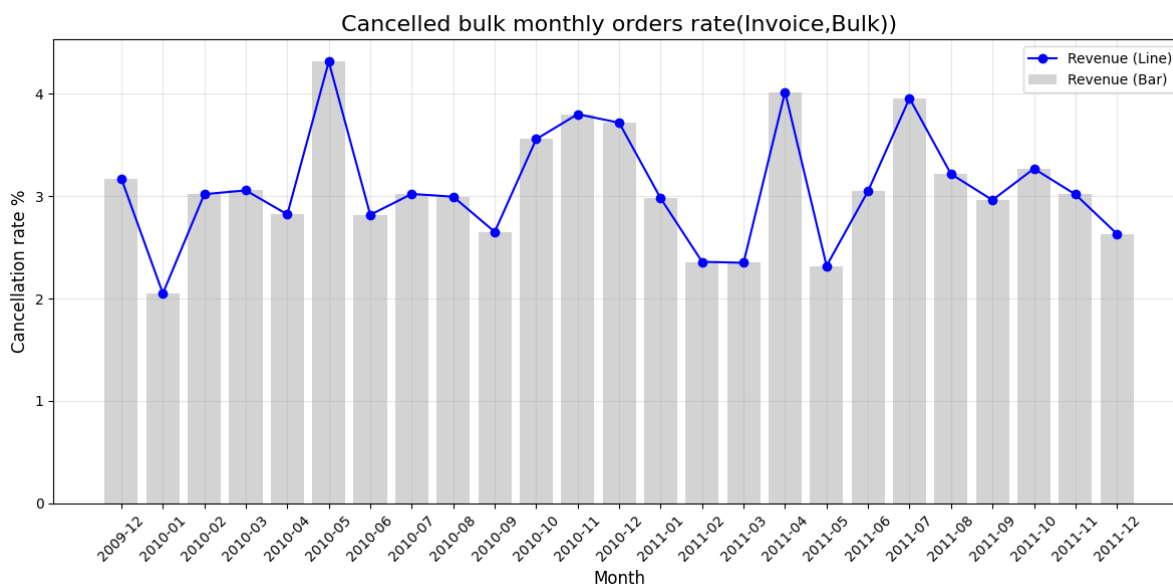


Figure 42: Taux d'annulation des commandes mensuelles (Facture)

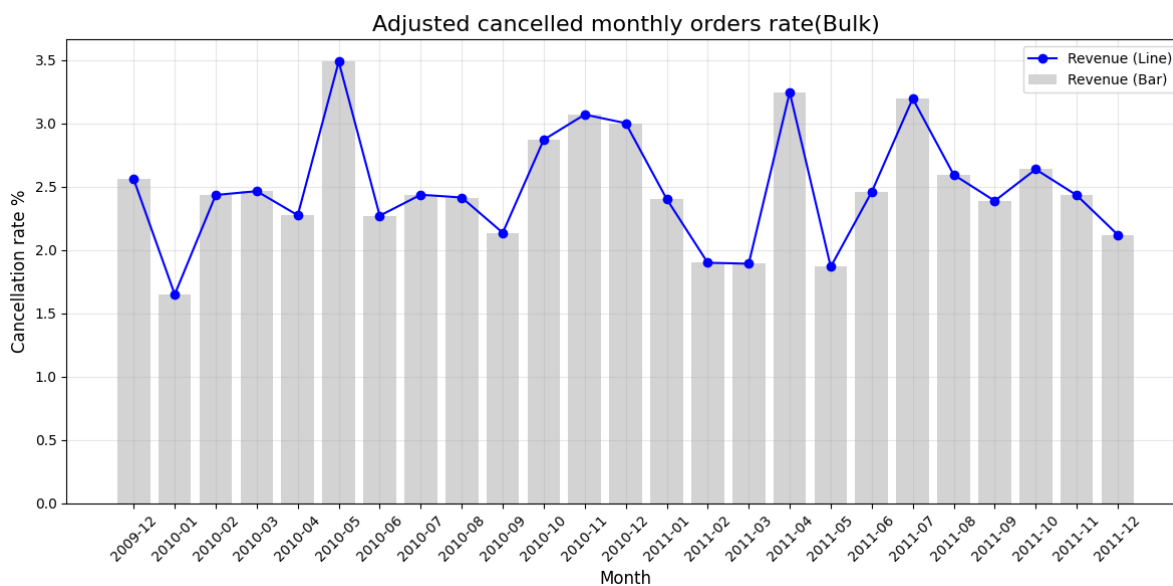


Figure 43: Taux d'annulation ajusté des commandes mensuelles (Facture)

Les taux d'annulation pour les commandes en gros montrent des fluctuations modérées au fil du temps, avec un taux d'annulation brut moyen d'environ 3.13% par mois. Ces taux atteignent un pic en mai 2010 à 4.31% et fluctuent entre 2.35% et 4.02% pendant 2011. Lorsqu'ils sont ajustés en fonction de la taille des transactions, les taux d'annulation diminuent à environ 2.29%, mais les chiffres ajustés ne sont pas aussi drastiques que ceux des annulations au détail, indiquant que les paniers plus petits dans les commandes en gros jouent un rôle important dans la formation de ces chiffres.

### Trimestre par Trimestre (QoQ):

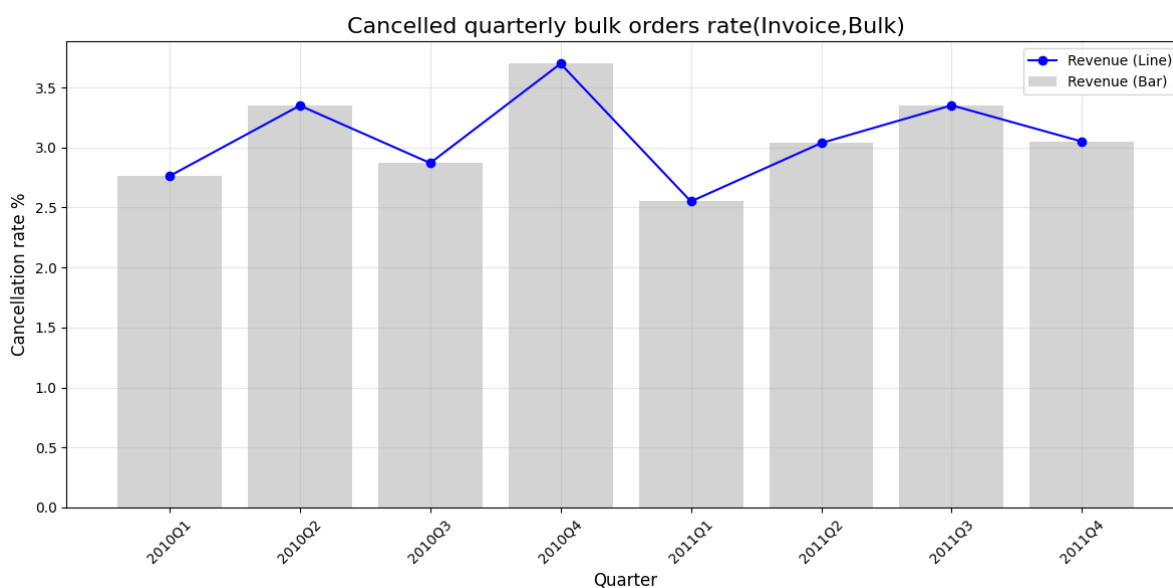


Figure 44: Taux d'annulation des commandes en gros par trimestre (Facture, Gros)

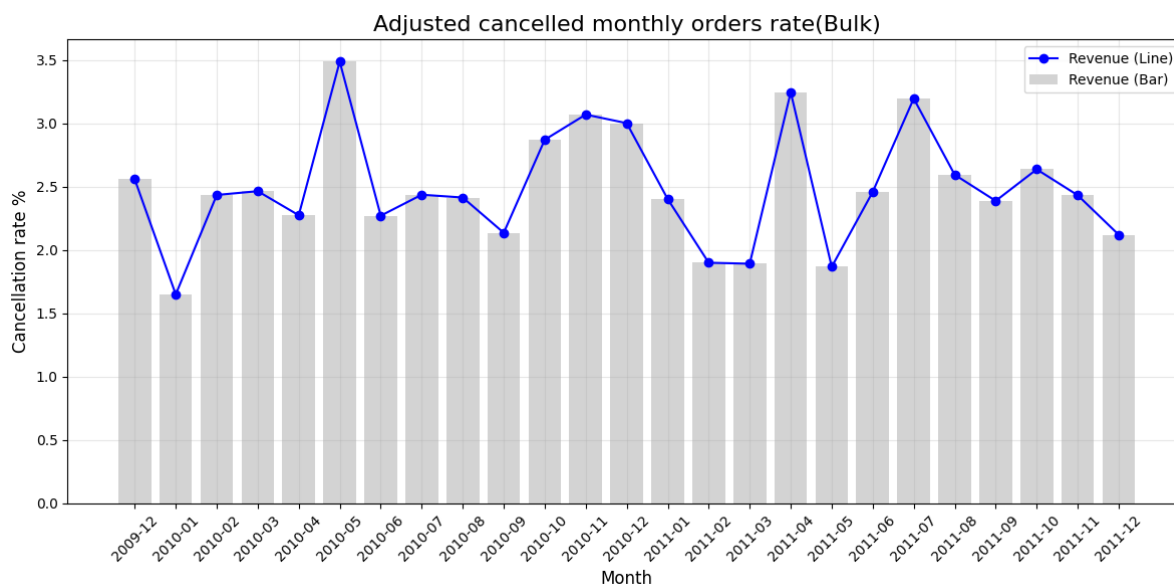


Figure 45: Taux d'annulation ajusté des commandes mensuelles (Facture, Gros)

Les taux d'annulation en gros par trimestre sont stables, allant de 2.55% à 3.70%, avec peu de fluctuations entre les trimestres. Les taux bruts et ajustés sont presque similaires, indiquant peu d'impact de la taille des paniers.

**Produits les Plus Annulés:**

Les 20 produits en gros les plus annulés sont :

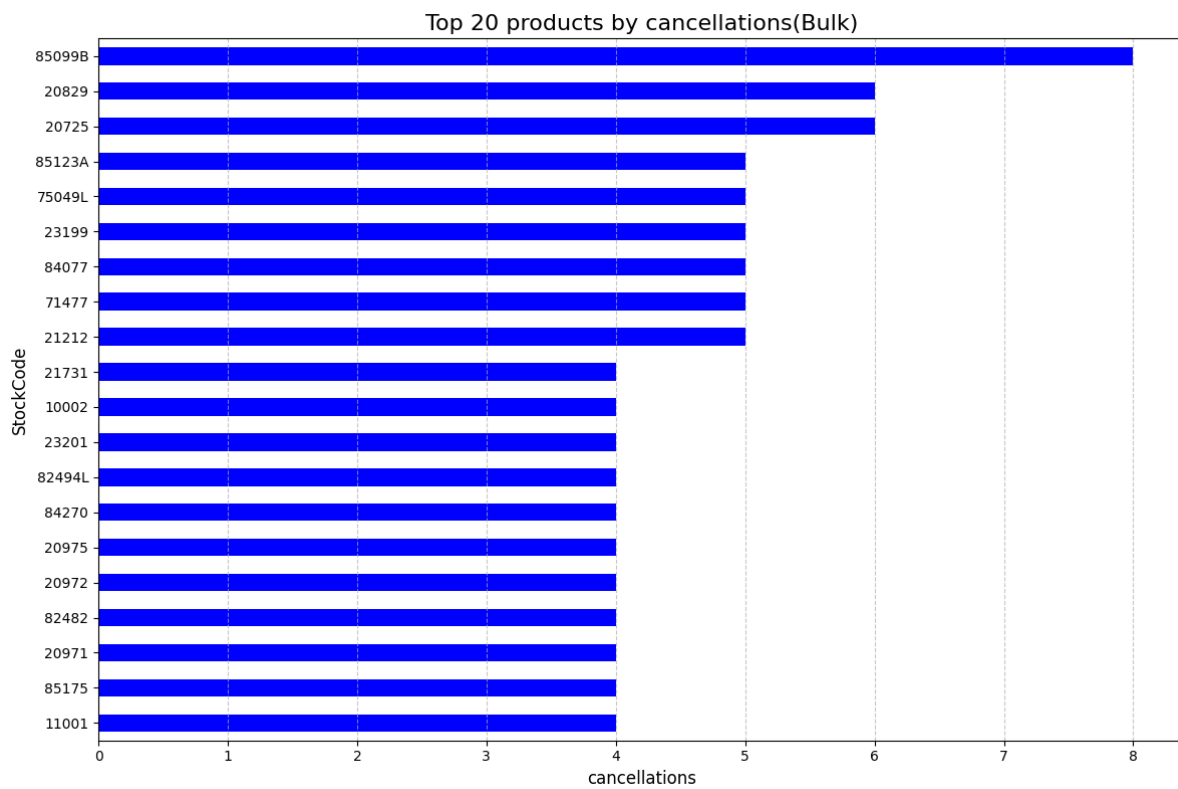


Figure 46: Top 20 Produits en Gros par Nombre d'Annulations



En même temps, les 20 taux d'annulation les plus élevés par produit sont :

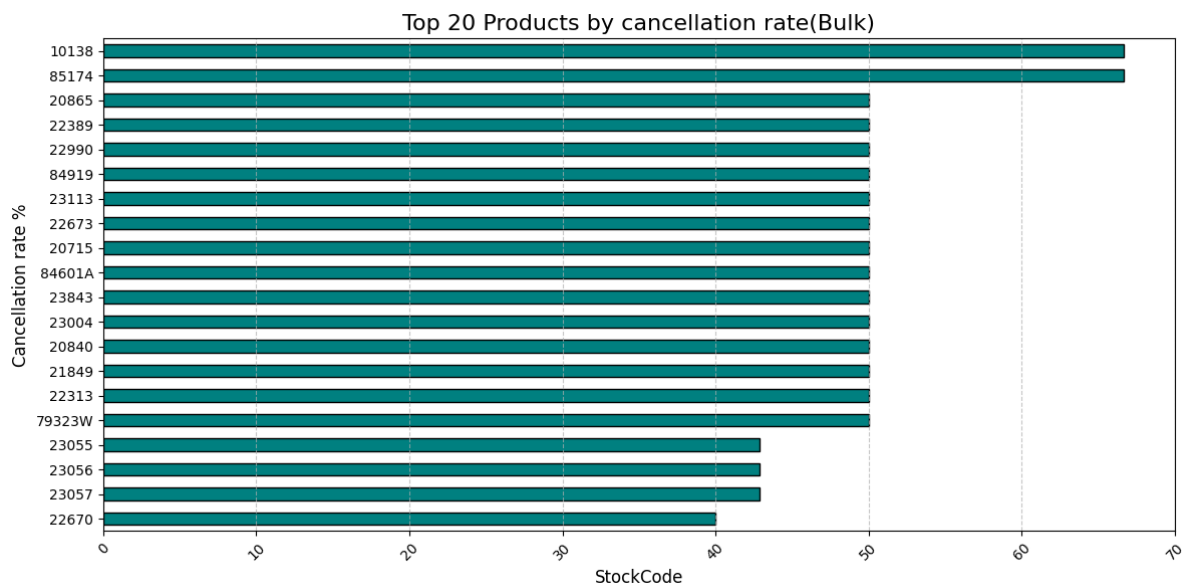


Figure 47: Top 20 Produits par Taux d'Annulation (Gros)

Certains articles comme 10138 ou 85174 (taux d'annulation de 66%) montrent des taux d'annulation élevés, signalant des problèmes potentiels avec ces produits.

**Clients les Plus Annulants:**

Les 20 clients les plus annulants en gros sont :

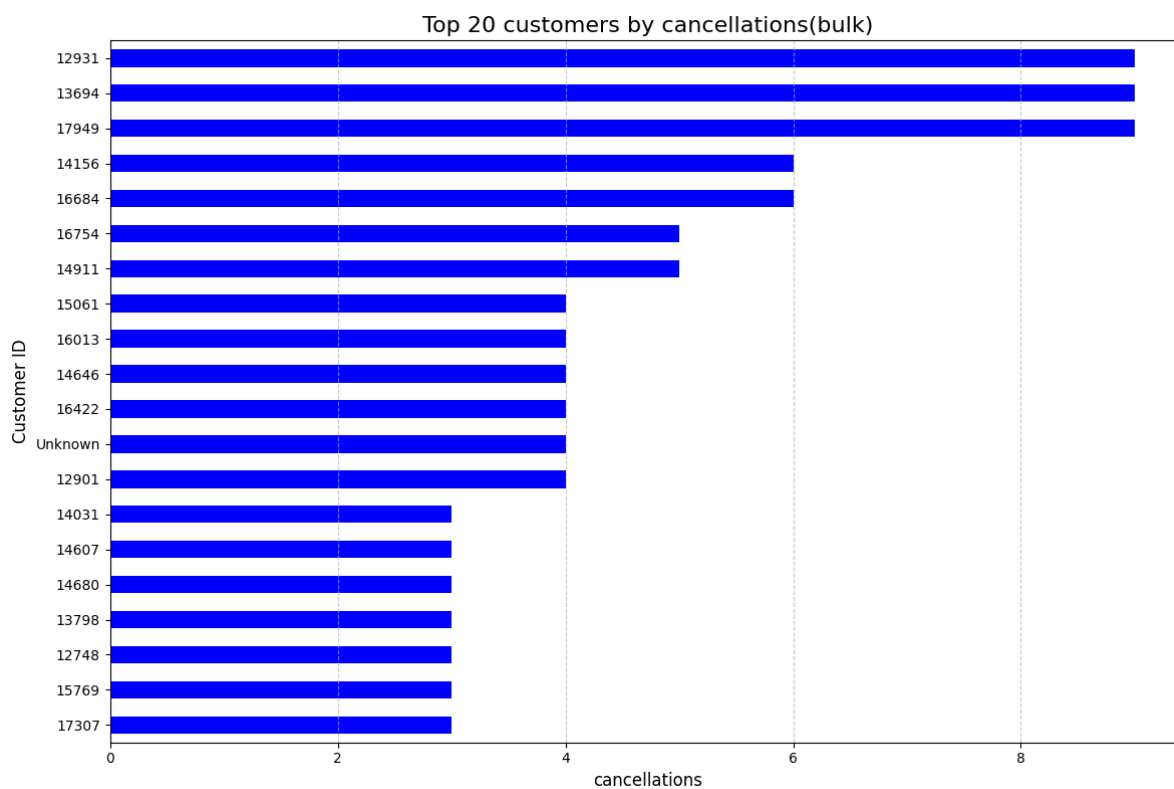


Figure 48: Top 20 Nombre d'Annulations par Client (Gros)

En même temps, les taux d'annulation les plus élevés par client en gros sont :

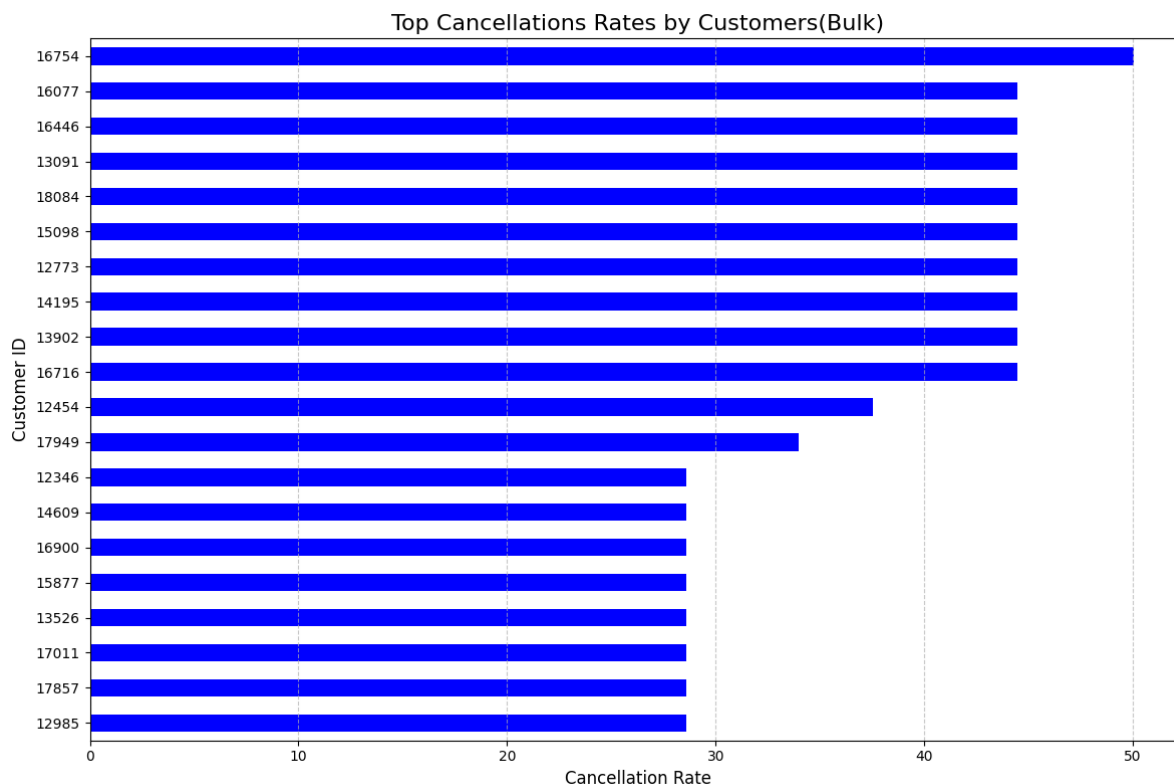
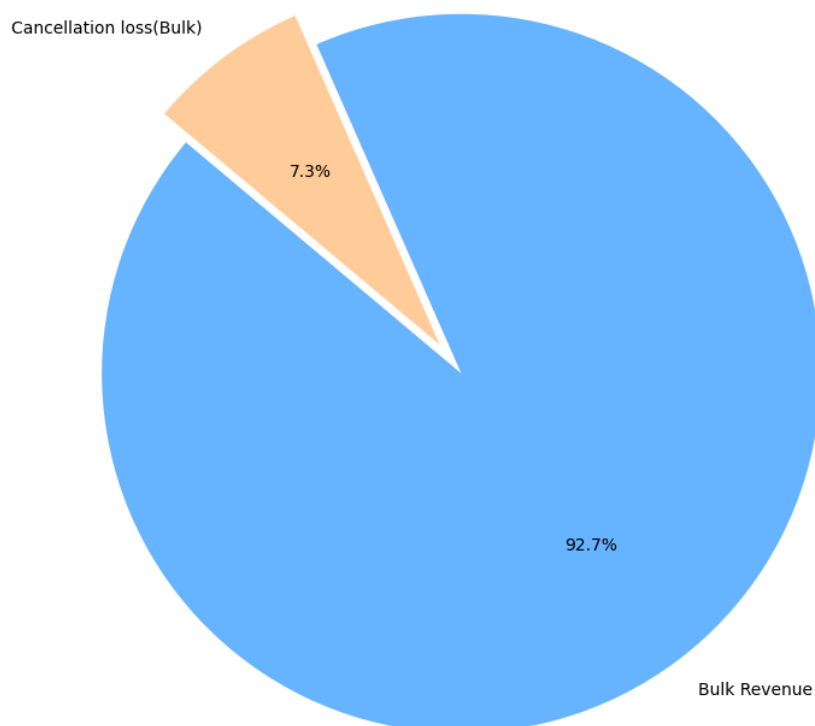


Figure 49: Top 20 Taux d'Annulations par Client (Gros)

Les annulations par clients montrent quelques clients avec des taux remarquablement élevés, surtout ceux au-dessus de 40%, indiquant des annulations fréquentes par certains clients particuliers.

**Perte de Revenu Due aux Annulations:****Revenue Breakdown(Bulk): Sales vs. Cancellation loss****Figure 50: Répartition des Revenus (Gros): Total vs Perte Due aux Annulations**

La perte de revenu due aux annulations en gros s'élève à £456,833.69, soit 7.3% des revenus totaux du gros (£5,802,651.11). Cela indique que, bien que le taux global d'annulation soit relativement faible, il a un impact notable sur le revenu, représentant un coût significatif pour l'entreprise, surtout compte tenu de la valeur moyenne des commandes plus élevée dans les transactions en gros.

**Résumé des Informations sur les Annulations en Gros:**

Les annulations en gros ont entraîné une perte de revenu de £456,833.69, soit 7.3% des revenus totaux du gros de £5,802,651.11. Bien que les annulations en gros soient moins fréquentes (275 annulations contre 7,216 dans le détail), la valeur moyenne des commandes plus élevée signifie que les annulations ont toujours un impact financier substantiel.

Le taux d'annulation en gros est de 3.13% (brut) et de 2.29% (ajusté), inférieur au taux brut au détail de 15.51% mais plus élevé que le taux ajusté. Comparées au détail, les annulations en gros sont moins fréquentes et plus gérables, avec des fluctuations stables et moins de perturbations. Les produits et clients les plus

annulés indiquent que certains articles et individus contribuent à une part plus élevée des annulations.

## 4 Recommandations

Sur la base des informations tirées de l'analyse, cette section fournit des recommandations exploitables adaptées aux tendances observées dans les ventes, les annulations et le comportement des clients. Bien que ces suggestions soient fondées sur les données disponibles, leur mise en œuvre doit tenir compte des processus commerciaux plus larges et du contexte supplémentaire.

### 4.1 Ajustements de la Stratégie de Vente

- **Exploiter les produits à forte performance :** Concentrez les efforts sur les produits ayant le plus grand volume de ventes et les taux d'annulation les plus faibles, tels qu'identifiés dans l'analyse.
- **Surveillance des tendances de vente :** Utilisez les modèles identifiés, comme les variations saisonnières et les tendances émergentes, pour anticiper les défis, optimiser la gestion des stocks et aborder proactivement les problèmes récurrents.
- **Capitaliser sur les ventes en gros :** Étant donné que les ventes en gros présentent des taux d'annulation plus faibles et contribuent de manière significative aux revenus, allouez davantage de ressources pour entretenir les relations avec les acheteurs en gros.

### 4.2 Segmentation des Clients et Fidélisation

- **Adapter les stratégies de segmentation :** Utilisez les segments de clients identifiés (par exemple, les annulateurs fréquents) pour concevoir des stratégies visant à traiter les raisons possibles des annulations et à réduire le taux de désabonnement.
- **Récompenser les meilleurs acheteurs :** Impliquez les meilleurs clients de détail avec des offres exclusives et des programmes de fidélité pour favoriser des relations à long terme.
- **Inciter les acheteurs en gros :** Proposez des incitations sur mesure pour les acheteurs en gros afin d'encourager les commandes répétées et renforcer les partenariats.
- **Activer les clients inconnus :** Mettez en place des initiatives pour convertir les clients inconnus en comptes identifiables.

### 4.3 Réduction des Annulations

- **Traiter les produits à haute annulation :** Examinez les produits ayant les taux d'annulation les plus élevés afin d'identifier les problèmes potentiels tels que des descriptions inexactes, des préoccupations de qualité ou des discordances de prix.

- **Améliorer les processus de commande** : Améliorez les procédures de traitement des commandes pour minimiser les annulations dans les transactions de détail et en gros.

#### 4.4 Améliorations Opérationnelles et Analytiques

- **Surveillance des tendances d'annulation** : Utilisez les modèles identifiés (par exemple, les variations saisonnières, les tendances, les problèmes spécifiques aux produits) pour traiter de manière proactive les problèmes récurrents.
- **Décisions basées sur les données** : Développez des systèmes internes pour mieux suivre et identifier les clients, permettant ainsi une meilleure analyse des données et des stratégies.

### 5 Limites de l'Analyse

1. **Absence d'objectifs spécifiques** : Ce rapport a été réalisé sans objectifs commerciaux prédéfinis ni orientation contextuelle. Les informations fournies sont exploratoires et conçues pour servir de point de départ à des analyses plus approfondies et ciblées.
2. **Absence de contexte externe** : L'analyse est limitée au jeu de données fourni et ne prend pas en compte les tendances du marché externe, les données des concurrents ou les facteurs opérationnels pouvant influencer les résultats.
3. **Profondeur exploratoire vs. contraintes de temps** : Bien que le rapport couvre des domaines clés tels que les tendances de vente, le comportement des clients et les annulations, des analyses plus profondes (par exemple, l'analyse des causes profondes des produits à forte annulation ou la modélisation prédictive des ventes futures) n'ont pas été menées en raison de la nature exploratoire de la tâche et du temps nécessaire pour de telles analyses.
4. **Limitations des données** : Des problèmes tels que l'absence d'identification des clients et des détails spécifiques sur les produits limitent la capacité de rendre certaines informations plus précises. De plus, le jeu de données peut ne pas capturer pleinement les nuances opérationnelles, telles que les retards dans la chaîne d'approvisionnement ou les efforts marketing, qui pourraient avoir un impact sur les tendances.
5. **Analyse statique** : Le rapport est basé sur des données historiques sans outils dynamiques (par exemple, des tableaux de bord) pour une surveillance continue ou une analyse en temps réel.

### 6 Conclusion

Ce rapport a analysé le jeu de données Online Retail II pour découvrir les principales tendances des performances des ventes, du comportement des clients et des annulations. Les résultats ont mis en évidence des différences significatives entre les transactions de détail et en gros, les ventes en gros montrant une stabilité plus élevée et une contribution

importante aux revenus malgré des volumes de commandes plus faibles. Les ventes au détail, quant à elles, ont contribué à la plus grande part des revenus, mais ont également montré une volatilité plus élevée, une saisonnalité et un taux d'annulation plus important.

L'analyse des clients a souligné l'importance de comprendre les segments de clients divers, avec des opportunités pour favoriser la fidélité parmi les meilleurs acheteurs et traiter les annulateurs fréquents. Les tendances des annulations ont fourni des informations précieuses sur les problèmes spécifiques aux produits et générés par les clients, révélant des opportunités exploitables pour l'optimisation des processus.

Sur la base de ces informations, des recommandations ont été proposées pour améliorer les stratégies de vente, réduire les annulations et améliorer la fidélisation des clients. Bien que ces suggestions soient basées sur les données disponibles, un contexte supplémentaire sur les processus opérationnels de l'entreprise pourrait affiner ces stratégies.

Dans l'ensemble, ce rapport souligne la valeur de la prise de décisions basée sur les données pour résoudre les défis opérationnels et identifier les opportunités de croissance. Les analyses futures pourraient explorer des ensembles de données complémentaires ou approfondir des domaines spécifiques de l'entreprise pour fournir des recommandations encore plus ciblées.

## 7 Références

- UCI Machine Learning Repository. (2019). *Ensemble de Données Online Retail II*. Récupéré depuis <https://archive.ics.uci.edu/ml/datasets/Online+Retail+II>
- Les ensembles de données traités utilisés pour cette analyse, y compris les données nettoyées et segmentées, sont disponibles pour examen et reproduction. Accédez au dépôt ici : <https://github.com/username/processed-data>.