Reducing Returns in Fashion & Electronics E-Commerce:

A Clustering-Based Framework for Identifying High-Risk Orders and Products
A Design Science Research at PwC

TIL Thesis

Quirine Japikse

Delft University of Technology

TU Delft

Reducing Returns in Fashion & Electronics E-Commerce:

A Clustering-Based Framework for Identifying High-Risk Orders and Products
A Design Science Research at PwC

by

Quirine Japikse

Studentnumber:	5040930
Supervisor CEG:	Dr. J. M. Vleugel
Supervisor TPM:	R. I. J. Dobbe
Supervisor PwC:	E. Youx
Project Duration:	September, 2025 - February, 2026
Faculty:	Faculty of Civil Engineering, Delft

Cover: https://www.pexels.com/photo/brown-cardboard-boxes-inside-a-delivery-van-7843987/
Style: TU Delft Report Style, with modifications by Daan Zwaneveld

TU Delft

Preface

My interest in logistics and sustainability has long been shaped by observing the growing phenomenon of product returns in e-commerce. I have often seen and heard people send back items they never truly needed or wanted, and I have read numerous articles about retailers struggling to manage the overwhelming volume of return packages, sometimes resorting to simply discarding them, which is costly both economically and environmentally. This project was carried out during my internship at PwC, which provided access to valuable data and expert insights essential to the success of this research.

I would like to express my sincere gratitude to my supervisors, Jaap Vleugel, Roel Dobbe, and Enzo Youx, for their guidance, feedback, and encouragement throughout this process. Their expertise and support were invaluable at every stage of the project.

I hope that the findings of this thesis will contribute to more sustainable and efficient return management practices in e-commerce, benefiting both retailers and the environment.

Quirine Japikse
Delft, February 2026

Summary

Introduction The rapid growth of e-commerce has led to an increase in product returns, making them a challenge for retailers. High return rates impose significant economic and environmental burdens. Moreover, operational inefficiencies further worsen these issues, as returned items often require costly manual processing and may end up discarded. Returns are driven by multiple factors, including impulse buying, free-shipping thresholds, product pricing, biased reviews, demographic influences, and payment methods such as “after pay”. Return policies play a dual role: they influence customer trust and sales while shaping return behavior. Liberal policies boost sales but increase returns, whereas strict policies reduce returns but risk harming customer satisfaction and loyalty. Effective return management involves multiple stakeholders: customers, e-commerce platforms, warehouse operators, IT and data teams, sustainability and compliance units, strategists, and logistics providers. Although most challenges arise in retailers’ warehouses, the retailer ultimately carries the financial and operational burden and is therefore the main problem owner. Because the involved teams each have their own priorities, effective collaboration is essential to tackle the issue and align improvements across operations, customer experience, and broader organizational goals.

Related work Recent research on return logistics optimization focuses on improving efficiency, reducing costs, and promoting sustainability, with AI and machine learning playing a central role. Studies show that ML models can accurately predict return volumes, enabling better planning and decision-making across areas such as inventory and pricing. Research on reverse logistics optimization highlights the importance of efficient network design, often combining traditional and AI-based methods to improve resource allocation and reduce environmental impact. Machine learning models are widely used to predict return rates, timing, and reasons. Conceptual frameworks, such as AI-based recommendation systems and product lifecycle tracking, further expand the field. Despite these advances, most studies focus on technical performance and overlook stakeholder involvement and practical implementation challenges.

In addition to data-driven methods, there are proposals to reduce returns or make handling more sustainable. By for example, systematically identifying and eliminating waste. Customer-centric innovations, such as customer-to-customer (C2C) return logistics, aim to reduce transport and packaging waste, although scalability and operational risks remain concerns. Other approaches include immediate return options to facilitate resale, personalized return policies to reduce fraud, and virtual-reality webrooms for pre-purchase fitting. However, this is limited to fashion and cannot fully replicate the physical product experience. Such strategies are often difficult to scale and frequently overlook the integration of data analytics in their design.

Research questions The literature reveals a persistent gap that the research question aims to address: a lack of comprehensive, stakeholder-informed frameworks that combine data analytics with process mapping and practical interventions. The research focuses on bridging technical innovation with operational feasibility and active stakeholder engagement. Leading to the following research question:

How can a reduction of returned goods in fashion- & electronic e-commerce be achieved by designing a framework that clusters high- and low-risk orders and products?

To address the main research question, several sub-questions were formulated.

Sub-questions

1. Which factors are shown by literature and historical data to significantly affect the return rate?

2. How are retailers currently handling returns, and what limitations exist in these approaches?
3. What is the aim and structure of a return flow optimizing framework, and how can it help the retailer optimize their system?
4. What orders and products are high or low risk, what are their characteristics, and what strategies can help reduce this risk?

Methodology This research applies the Design Science Research (DSR) methodology, which systematically combines literature review, expert interviews, and data analysis to develop and evaluate a practical framework for reducing e-commerce returns. A literature review is conducted to identify the state-of-the-art factors employed in data-driven return models. In addition, the literature is examined to gain insight into current practices in returns management. Finally, existing strategies for reducing product returns are systematically reviewed. Interviews will be used to confirm and expand on findings from the literature on associated features and current practices.

This study uses datasets containing product and order-level information. The dataset was preprocessed by merging information, identifying outliers, and imputing missing values. Furthermore, categories were aggregated into broader groups, mitigating sparsity. This dataset will be analyzed using statistical methods in Python, such as correlation and chi-square tests.

Lastly, clustering methods group observations that are similar to each other and different from other clusters, enabling the discovery of hidden patterns. Three different clustering methods are tested and compared: K-prototyping, CAVE, and Latent Class Clustering (LCC). They represent three different approaches to clustering mixed-type data. The K-prototypes algorithm is a distance-based method that combines numerical and categorical dissimilarities. CAVE further refines the framework by weighting features according to variance and entropy, thereby improving cluster interpretability and supporting categorical data. In contrast, LCC is a probabilistic model that assigns observations to clusters based on posterior class membership probabilities, offering greater flexibility in capturing complex variable dependencies and often achieving superior accuracy when categorical variables dominate.

Key findings The literature identifies a broad set of features associated with product returns, including product attributes, customer characteristics, and order-level factors. Prior studies consistently show that product type, customer data, and basket composition influence return risk, yet no consensus exists on an optimal combination of these. Karl's (2025) systematic review highlights that many potential feature combinations remain unexplored, and the specific configuration used in this study has not previously been documented. These findings serve as an input for the data analysis in Chapter 4. The dataset used in this research encompasses a broad range of consumer products typically sold in the intimate wellness and personal care segment. The analysis showed significant correlations between being returned and product price, order price, product quantity, and the number of items in the order. Notably, against the findings of Urbanke et al. (2015), bracketing showed a negative correlation with returns. Furthermore, a chi-square test examined significant differences in categories in product size, category, color, and delivery modes. The time of day was not significantly correlated with the outcome in correlation tests, but showed significant differences when grouped into part-of-day categories. The correlations all show weak associations, suggesting that no single feature can explain returns.

The literature on return logistics management offers a valuable conceptual distinction between preventive and curative dimensions. While the curative dimension provides a detailed account of operational activities, these studies tend to treat the process as a linear sequence of actions. A gap persists in the limited attention to the allocation of responsibilities among stakeholders throughout the return process. Moreover, the balance between automation and manual intervention remains largely unaddressed, despite its practical significance for efficiency and scalability. This concept of activities in return logistics serves as a basis for understanding current handling. Interviews revealed that most retailers handle returns using a combination of manual processes and basic rules, with limited use of data-driven strategies or automation. Limitations of this approach include high return volumes and manual handling leading to high operational costs. Furthermore, there is underutilization of return data for process improvement. The actions, responsible stakeholders, and level automation are visualized in Figure 5.1. Figure E.1 illustrates that, within the retailer, no single party holds clear ownership of the problem;

instead, various teams pursue misaligned objectives and do not collectively assume responsibility for the issue of returns.

The conceptual framework is grounded in the principal limitation: the high volume of returns. To address the gap concerning the lack of practical, stakeholder-oriented interventions, the first requirement is that the model outputs must be interpretable. It should enable stakeholders to understand the factors driving returns and the rationale behind the grouping of products or orders. In addition, the framework must use the available data effectively, accommodate both categorical and numerical variables, and remain scalable to handle the large datasets characteristic of e-commerce environments. The overarching aim of the framework is to reduce returns by identifying high-risk products and orders. Clustering techniques are used to detect patterns within these groups, after which strategies are formulated around the resulting insights and subsequently validated with stakeholders. This iterative process is illustrated in Figure 6.2. The framework supports the retailer in moving from the current fragmented return process to a more coordinated and data-driven approach. It introduces a new stakeholder, the “Mediator”, responsible for aligning teams, validating strategic decisions, and closing the gap between operational issues and strategic action. By integrating data-based strategies, the optimized process shifts return handling from reactive to proactive, ultimately reducing return volumes.

The current literature focuses on various strategies to reduce returns. Taking product information, operational improvements, return policy, and marketing strategies into account. A combination of clear product information, customer-focused tools, and carefully balanced return policies can reduce returns. However, most strategies still focus more on customer behavior than on retailer decision-making. The strategies serve as input for addressing sub-question 4. Across both order-level and product-level analyses, the K-Proto and CAVE algorithms consistently cluster observations based on numerical features, particularly order size, order value, and product price. The results reveal a clear pattern: larger orders and higher-value orders carry a higher return risk, and mid-priced products tend to produce the highest return rates at the product level. In contrast, LCC identifies clusters predominantly based on shipping carriers and product categories, revealing patterns not detected by the distance-based algorithms. At the order level, COLISSIMO and INPOS emerge as low-risk carriers, while DHLDE consistently forms high-risk clusters. This indicates that carrier-specific effects likely reflect country-level operational differences. At the product level, K-Proto and CAVE give few differences in return rates. LCC, however, reveals a distinctly high-risk product cluster dominated by Category F, in which mid-priced products, especially in black and red and sizes S–XL, display disproportionately elevated return rates. Subclustering confirms that higher prices within this group further increase risk, while “One Size” products remain consistently low risk. Integrating order-level features into the product-level clusters shows that the highest-risk product clusters coincide with large, high-value orders. In contrast, low-risk clusters typically involve small or medium-sized orders with either low or very high total values. This finding indicates bracketing behavior or strategic basket filling. Carrier differences become less influential once product characteristics are taken into account.Lastly, across all clustering methods and clusters, orders placed during nighttime hours exhibit a higher likelihood of being returned.

Suggested strategies are based on clusters and literature. The clusters show that items with specific sizes are more likely to prompt requests for strategies to improve product descriptions. This can be done by improving size charts and visual information. Additionally, offering “One Size” alternatives where feasible could decrease returns. As Cluster 2 shows high-risk products, stricter return conditions for Cluster 2 can reduce returns. For marketing and sales strategies, avoid promotions on high-risk items; market low-risk products instead. Expert validation confirmed product information improvements as highly feasible. Policy changes require further internal evaluation, and logistics-related interventions depend on country-specific constraints. Marketing strategies and free shipping adjustments were considered actionable. These suggested strategies all come with considerable trade-offs. Overall, the highlighted trade-offs show a recurring tension between reducing return risks and preserving sales performance and customer satisfaction. Achieving an optimal balance requires a clear understanding of which items fall into high-risk clusters and close collaboration among stakeholders across the retailer’s organization.

Discussion This research confirms several return drivers established in prior research while also challenging and refining existing assumptions. Consistent with Mishra & Dutta and Cui et al., operational factors such as shipping method, delivery mode, carrier choice, and product attributes significantly influence return rates. Price-related findings partially align with Asdecker & Karl, who associate higher values with elevated return risk; however, this study reveals a more nuanced pattern in which medium-priced items return more frequently than both low- and high-priced products, highlighting the context-dependence of price categorization. The weak negative relationship between order quantity and returns contradicts earlier claims and appears too small to represent meaningful behavioral patterns. Beyond return drivers, the study extends the literature by mapping stakeholder responsibilities and automation levels within the returns process, an area overlooked in prior work by Frei and Stevenson, who describe return activities without identifying ownership. The findings show a fragmented organizational structure with no clear problem owner, hindering coordinated return-reduction efforts. Furthermore, existing frameworks in the literature address either methodological optimization or process improvements, but none integrate data analytics, operational insights, and stakeholder perspectives into a single approach. The study closes this gap by demonstrating how high- and low-risk profiles can inform targeted strategies, while also reaffirming the trade-offs between reducing returns and maintaining customer satisfaction, previously noted by Duong et al. but largely absent from broader discussions on return-reduction methods.

This research contributes to the existing literature in several ways. It clarifies how responsibilities are distributed across stakeholders in the return process and shows that differing goals among internal teams make coordinated return-reduction efforts challenging. Additionally, the study introduces a stakeholder-informed framework that integrates data analytics with practical, operational interventions. The findings suggest that an independent function can help align these teams and develop data-driven strategies that balance their competing objectives. Finally, it demonstrates that Latent Class Clustering (LCC) is the most suitable method for the data and objectives of the cluster analysis in the framework.

The study is subject to several important limitations that influence the interpretation and generalizability of its findings. The dataset lacked demographic variables and return-reason information, restricting the analysis to product- and order-level attributes and preventing deeper insight into customer-specific return behavior. The absence of qualitative data, including customer perceptions and satisfaction, limits the ability to understand the underlying causes behind returns. Furthermore, the five-month observation window introduces potential seasonal bias and the risk of incomplete return capture. Methodologically, the clustering techniques exhibited structural challenges: distance-based methods such as K-prototypes and CAVE were dominated by numerical behavior, while LCC, although more effective for categorical pattern recognition, was computationally intensive and therefore less practical for large-scale use. Finally, because the analysis was conducted using data from a single retailer without external validation across other firms, markets, or product categories, the broader applicability of the conclusions remains constrained. My own perspective also shaped the direction of this research. By placing responsibility primarily on the retailer rather than the customer, the study focused on strategies that require organizational rather than behavioral change. This stems from the view that many customer behaviors often labeled “problematic,” such as bracketing or strategic basket filling, are in fact enabled or even incentivized by retailers through generous return policies, free shipping thresholds, and marketing practices designed to maximize sales. As a result, the strategies developed in this research focus on adjustments within the retailer’s sphere of influence rather than on modifying customer behavior. This reflects the underlying assumption that retailers have both greater control over return-related processes and a responsibility to design systems that minimize avoidable returns.

Conclusion This thesis developed a structured framework to reduce return volumes by systematically identifying high- and low-risk product and order groups. The characteristics of these segments form the foundation for designing targeted, evidence-based strategies to address the underlying drivers of returns. The application of multiple clustering techniques revealed that different methods highlight different determinants of risk. By applying this framework, retailers can derive data-driven interventions that are both operationally relevant and aligned with stakeholder perspectives, thereby supporting more effective, informed return management practices.

Future research and recommendations Future research should focus on expanding the analytical foundations and practical applicability of this study. Collecting and integrating customer-provided return reasons would offer crucial insights into the underlying drivers of return behavior and enable more targeted interventions. Extending the observation period is equally important to capture seasonal dynamics and ensure that return patterns are assessed over a representative time horizon. Methodologically, it is useful to explore whether fewer categorical variables or more computationally efficient clustering techniques can still produce comparable insights. This would help improve scalability without compromising analytical quality. Moreover, future work should place greater emphasis on customer experiences with newly implemented strategies, incorporating systematic feedback collection to strengthen the behavioral dimension of return management. Testing and refining the framework across different industries, product types, and logistical contexts will be essential to enhance its generalizability and robustness. Further investigation into operational factors, such as delivery quality, customer perceptions, and carrier-specific processes, particularly in cross-country contexts, may clarify elevated return risks associated with specific carriers. Finally, future studies assess the long-term financial and environmental outcomes of proposed interventions through mathematical modeling and develop real-time identification techniques to assist customers in making more informed purchase decisions.

The framework proposed in this study can be applied in practice. In addition, the research identifies several targeted strategies to reduce e-commerce returns. A key measure is improving product information and fit guidance to help customers make more accurate purchasing decisions. This improvement can be achieved through clearer descriptions, enhanced size charts, and customer fit feedback. Return policies can also be differentiated by risk level, applying stricter conditions to high-risk products while keeping policies customer-friendly for low-risk items. Operational interventions during high-risk ordering hours, incentivizing low-risk shipping methods, and avoiding aggressive promotions for high-risk products can further help mitigate returns. Technical tools such as live order flaggers and post-purchase confirmations for large or high-value orders may prevent unnecessary returns, while adjusting free-shipping thresholds can discourage risky order compositions. The findings also highlight the importance of understanding carrier-specific return patterns and emphasize the need to appoint a clear problem owner within the organization who can coordinate cross-functional collaboration and manage the trade-offs involved in implementing these strategies effectively.

Contents

List of Figures

List of Tables

1 Introduction

The rapid expansion of e-commerce in recent years has led to a significant surge in consumer returns. According to the National Retail Federation and Appriss Retail (2023), e-commerce continues to grow at double-digit rates, making returns a critical area of concern [37]. With online sales peaking, returns have become a dominant segment of retail logistics, making returns a highly relevant field of research [1, 15]. The COVID-19 pandemic acted as a major catalyst for global e-commerce growth, with its share of total retail sales rising from 15% in 2019 to 21% in 2021. This upward trend is expected to persist, with the global e-commerce market projected to grow from $6.3 trillion in 2023 to $8.1 trillion by 2026, accounting for 24% of all retail sales [44]. However, the gap between the virtual shopping experience and the physical reality of products often leads to inevitable returns. In Germany alone, 24% of parcels were returned in 2022 [39]. This amounts to 530 million parcels and 1.3 billion items. Similar consumer behavior is observed across Europe, where return rates for fashion items can reach up to 60% [39]. This high volume of returns, along with the anticipated rise, calls for more insights into how to handle them in the future.

1.1. Problem definition

From both economic and ecological perspectives, the return trend represents a substantial burden. For retailers, the high volume of returns makes it increasingly difficult to maintain positive margins. At the same time, the environmental impact is considerable: in Germany alone, returns generated approximately 795,000 tonnes of CO_2 emissions in 2021 [20], equivalent to 5.3 billion kilometers driven by car. The financial cost is equally striking, with average return-related expenses (including transport and processing) estimated at €6.95 per shipment, translating to an additional €3.68 billion in costs for German companies in 2022 alone [39]. The growing volume of returns poses a significant challenge for businesses of all sizes. Many returned items are still in usable condition, yet they often get lost in transit, misrouted in warehouses, or overlooked in inventory systems [16]. Even when products do make it back to the retailer, the manual effort required to inspect, reprocess, and restock them is time-consuming and costly, frequently resulting in unnecessary disposal. This not only drives up operational expenses but also contributes to environmental waste [35, 5]. This growing challenge has led to the exploration of predictive tools and data-driven strategies. Asdecker and Karl (2018) highlight the potential of big data analytics for forecasting consumer returns, noting that even relatively simple models, such as binary logistic regression, can yield effective results [4]. If companies have a better understanding of risks and returns, they can identify mitigating strategies, making it easier to streamline logistics and reduce both financial losses and ecological impact. There are many reasons why customers return items to online retailers. Understanding these can help reduce return rates and the challenges they create, as stated above.

1.1.1. Reasons for returns

Multiple factors contribute to the high volume of product returns in e-commerce. Firstly, marketing strategies that encourage spontaneous purchases, such as limited-time offers or heavy discounts, can boost sales. However, they are also associated with elevated return rates. Research shows that impulse buys are significantly more likely to be returned. The use of coupons also contributes to impulsive buying, which leads to more returns. Therefore, it is crucial to balance marketing initiatives with return policy design, as firms must carefully navigate the trade-off between promotional effectiveness and operational efficiency. Shipping practices also play a significant role in influencing product returns. For instance, free shipping promotions are often applied to products that are difficult for consumers to evaluate before purchase, increasing the likelihood of returns [20]. Furthermore, minimum order value requirements are associated with a slight increase in return rates among existing customers, without affecting the purchase incidence or value. Moreover, threshold-based free shipping policies contribute to a higher incidence of strategic returns, as customers are more inclined to order additional items with the intention of returning some [29]. Beyond these factors, customer reviews and product pricing also play an essential role in return behavior. Empirical evidence suggests that a higher number of unbiased customer reviews correlates with lower return rates, whereas a high proportion of biased reviews tends to increase them. In terms of pricing, more expensive products are less likely to be returned compared to lower-priced items [20]. Additionally, demographic factors also influence return behavior. Makkonen et al. (2021) report that women and younger consumers are more likely to return products. Their study also reveals that customers who choose “after pay” payment methods are more likely to return products, consistent with the findings of Gry et al. (2024) [31, 20]. The paper by Gry et al. shows that payment methods matter: orders paid by invoice tend to have higher return rates, likely because they lower perceived commitment at checkout [20]. While some returns are acceptable, others are undesirable and can be minimized to reduce overall return rates. The objective should not be to eliminate returns; achieving a 0% return rate is neither feasible nor desirable, but rather to target behaviors that unnecessarily inflate returns. Examples include bracketing, impulse purchases followed by quick returns, and strategic basket-filling to qualify for free shipping with partial returns. Conversely, acceptable returns, such as those from damaged items, should serve as the ideal baseline for return rates. For instance, if 25% of returns stem from bracketing, this behavior warrants targeted interventions. However, if 10% of returns occur due to damaged deliveries, efforts should focus on improving packaging and handling rather than changing consumer behavior. Despite knowledge of these factors individually, it remains unclear how they jointly influence returns and how strategies and policies can be most effective.

1.1.2. Strategies to reduce returns

Return policies provide guidance on when a customer can return an item and what costs are incurred. These policies are not only operational tools but also strategies for e-tailers to impact sales, costs, customer satisfaction, loyalty, and brand reputation. Because of this, policies should balance between being liberal and strict. Liberal policies will increase return rates, whereas strict policies will lower sales. It is essential to have strategies that do not reduce your sales more than they increase your returns. In the past, several strategies have been used to reduce returns. For example, Zalando and Wehkamp started using return fees again to lower returns. Furthermore, Amazon has 18 different return policies per product category, making it more applicable to the specific product. A well-designed return policy can increase sales by building customer confidence. When shoppers know that returning a product is hassle-free, they feel safer making a purchase. This lower perceived risk often translates into higher return rates and larger basket sizes, as customers are willing to “buy now and decide later” [10]. Research by Duong et al. (2025) confirms that lenient return policies serve as a pre-purchase signal of trustworthiness and quality, encouraging customers to buy more and try new products. Satisfied customers are more likely to become repeat buyers and recommend the retailer to others, creating loyalty and brand reputation. If a retailer makes returns very difficult or costly, customers may hesitate to buy at all. For example, Duong et al. found that while stricter rules do reduce return rates, they also target genuine customers with legitimate returns, decreasing overall satisfaction and future purchase intent. In other words, shoppers reward convenience and may punish a brand perceived as overly rigid or unfriendly to returns. Thus, a too strict policy can hurt customer retention and lifetime value, offsetting the savings from fewer returns [10]. Retailers describe this as the return policy leniency dilemma: how to be generous enough to win customers’ trust and business without being so generous that returns erode profits. Shoppers might go to a competitor if they know that returning an order will be a hassle or expensive. When defining strategies, it is important to consider this and choose strategies that do not decrease customer satisfaction. While many return-reduction strategies predominantly target consumers and their behavioral patterns, e-tailers also make internal strategic decisions that can influence return rates. This raises the question of which firm-level strategies can effectively reduce product returns independent of consumer behavior.

1.1.3. Stakeholder analysis

In the context of return logistics, it is crucial to clearly identify the stakeholders involved, as they each experience and shape the problem from different perspectives. The first stakeholder identified is the customer. After making a purchase, the customer ultimately decides whether to keep or return the product. Their purchasing and return behavior are key drivers of return volumes. Impulse buying, payment method, and expectations about product quality directly influence return rates. As seen in the sections above, their behavior is influenced by marketing and return policies. Furthermore, customer satisfaction is strongly tied to smooth return handling, clear communication, quick refunds, and transparent policies. This makes them an essential priority for any strategy aimed at reducing the high volume of returns and improving return flow.

E-commerce platforms and retailers constitute a primary stakeholder group, bearing the direct financial and operational consequences of elevated return rates. This includes transportation costs, restocking efforts, and losses from unsellable merchandise. They exercise control over return policies and must carefully balance customer-centric practices with operational efficiency and profitability. Decisions regarding free shipping, discount structures, and promotional campaigns influence not only marketing outcomes but also logistical complexity. Teams within these companies may have different needs as well.

Warehouse operators and inventory managers represent another critical stakeholder group. They handle the reception, inspection, reprocessing, and reintegration of returned goods into inventory. Inefficient processes can result in unnecessary waste or mismanaged inventory. Any interventions involving process changes should be reviewed and approved by warehouse operators to ensure feasibility and alignment with operational realities.

IT and data management teams play a role in monitoring and integrating technological solutions within operational platforms. Robust and well-structured data is essential for implementing models and analytics. Similarly, sustainability and compliance teams focus on mitigating the environmental impact of returns by tracking CO2 emissions, promoting reuse and recycling, and ensuring alignment with regulatory frameworks and corporate sustainability objectives. These teams rely on precise return data to design strategies that reduce waste and ecological harm. Finally, strategic management teams utilize return-related insights to form decisions on pricing, marketing campaigns, and policy adjustments. Their aim is to optimize customer experience while having profitability and long-term competitiveness. This indicates that the sales, marketing, and finance teams are also identified as stakeholders in this issue, although their level of interest and influence appear relatively low.

Postal carriers and third-party logistics providers play a pivotal operational role by managing the transportation of returned products. Their efficiency directly impacts turnaround times, operational costs, and overall customer satisfaction. Additionally, selecting a specific postal carrier can affect returns. The operational practices employed during the outbound journey significantly shape the condition and presentation of the package upon arrival, thereby influencing customer perception and satisfaction.

Together, these stakeholders form a complex network of actors whose needs, constraints, and objectives must be balanced when designing a solution for optimizing the return flow. Understanding their roles and perspectives ensures that an artifact addresses not just operational efficiency but also customer satisfaction, environmental impact, and strategic value. An overview of the stakeholders is shown in Figure 1.1 below.

1.1.4. Conclusion

In conclusion, the rapid growth of e-commerce has led to an increase in product returns, posing a challenge for retailers. High return rates impose significant economic and environmental burdens. Moreover, operational inefficiencies further exacerbate these issues, as returned items often require costly

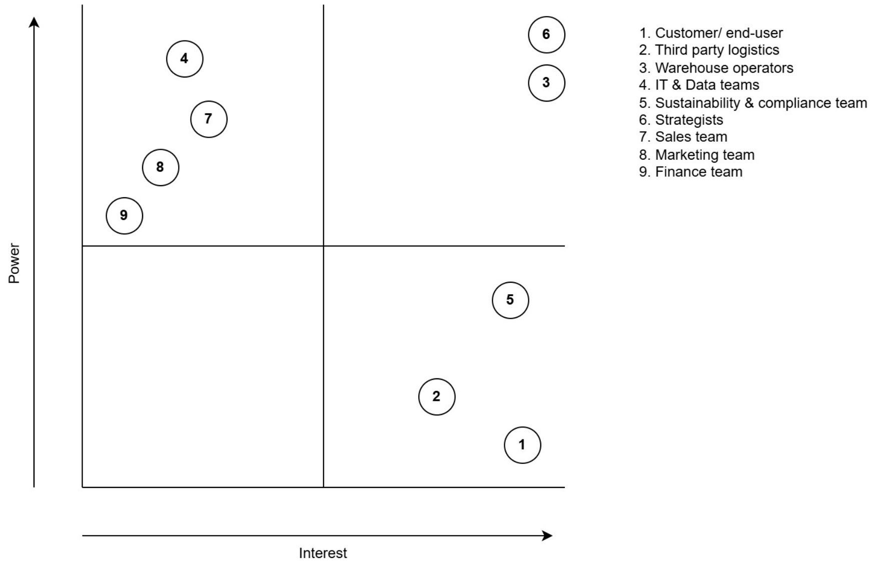


Figure 1.1: Power-interest grid of stakeholders

manual processing and may end up discarded.

Returns are driven by multiple factors, including marketing that encourages impulse buying, free shipping thresholds, product pricing, biased reviews, demographic influences, and payment methods such as “after pay” or invoice. While some returns (e.g., damaged goods) are unavoidable, others, such as bracketing, impulse-driven purchases, and strategic basket-filling, can be mitigated through targeted interventions. While the individual effects of these factors are well-documented, their combined influence warrants further investigation.

Return policies play a dual role: they influence customer trust and sales while shaping return behavior. Liberal policies boost sales but increase returns, whereas strict policies reduce returns but risk harming customer satisfaction and loyalty. Retailers face the “return policy leniency dilemma”, requiring strategies that balance profitability with customer experience.

Effective return management involves multiple stakeholders: customers, e-commerce platforms, warehouse operators, IT and data teams, sustainability and compliance units, strategists, and logistics providers. Each group has distinct priorities, from operational efficiency and cost control to environmental impact and customer satisfaction. Designing solutions for return optimization demands an integrated approach that aligns these interests while using data-driven strategies to reduce return losses and ecological harm.

1.2. Related work

This section discusses the relevant body of work in the field, outlining both data-driven approaches and process-oriented methods examined in prior research. It concludes by identifying the research gap that emerges from this literature.

1.2.1. Data-driven approaches on improving return logistics

The optimization of return logistics has been widely studied, with research increasingly emphasizing efficiency, cost reduction, and sustainability. A recurring theme in the literature is the application of AI and machine learning (ML) as technical enablers for process improvement. For example, Gry et al. (2024) propose an AI-based recommendation system to optimize handling time by routing returns to the most appropriate sales channel, aiming to reduce both economic losses and environmental impact [20]. Niederlaender et al. (2024) extend this perspective by evaluating alternative ML algorithms for return forecasting and introducing regularized target encoding to handle high-cardinality variables, achieving predictive accuracies up to 0.86 [35]. Mishra et al. further enrich the discussion by identifying additional explanatory variables, such as total order value and Cash-on-Delivery fee, as critical predictors of return behavior [32]. Accurate prediction of return volumes and reasons is identified as a crucial first step in improving return logistics. ML algorithms can detect patterns in large datasets and produce forecasts that enable proactive rather than reactive planning, thereby improving the economic, environmental, and social performance of reverse logistics [20, 35, 32, 27]. Niederlaender et al. explore alternative ML algorithms for return forecasting, while Mishra et al. investigate additional explanatory variables influencing return behavior [32]. These studies collectively underscore the value of data-driven approaches for proactive planning in reverse logistics, enabling more informed decisions on inventory management, pricing, and promotional strategies.

Karl (2024) identifies three primary methodological streams: ML algorithms, statistical techniques, and conceptual frameworks [27]. Among these, ML approaches are the most prevalent, with models such as Random Forest, Support Vector Machines, Neural Networks, and Gradient Boosting [20]. These are methods widely used to predict return propensities, volumes, timing, and reasons, as well as to detect fraudulent behavior or “bracketing” practices. These models leverage a rich set of features, including product attributes (e.g., price, size, color, category), customer behavior profiles, order characteristics, transaction data, and even qualitative information such as product reviews or sentiment analysis. While ML dominates the field, classic statistical methods, such as logistic regression or probit models, remain valuable for identifying key predictors and providing interpretable insights into return drivers [20]. Time-series forecasting techniques (e.g., ARIMA, Holt-Winters smoothing) are occasionally applied to predict aggregate return volumes over time [27]. Beyond algorithmic methods, several conceptual frameworks have been proposed, including AI-based recommendation systems for “Second Life” planning, scalable cloud-based platforms enabling real-time return prediction, and advanced graph-based approaches such as HyperGo and HyGraph for modeling customer-product relationships. All this literature together illustrates that there are different ways to predict and analyze returns. These studies have examined various methods for analyzing returns data; however, none involve stakeholders in the design process. Karl systematically categorizes predictors of product returns and highlights research gaps, such as the need to study return timing and to classify predictors across 18 categories. In contrast, Gry et al. (2024) propose an AI-based system for efficiently routing inevitable returns to the most suitable sales channel, emphasizing economic and ecological value. While Karl focuses on how returns are predicted and Gry et al. on post-prediction handling, both overlook stakeholder involvement and practical implementation.

Another area of research concerns the optimization of reverse logistics networks. Designing an efficient network is essential for managing the uncertainty of timing, quantity, and quality of returned products. Approaches combining traditional optimization techniques with AI-based forecasting have been shown to improve resource allocation, network design, and cost-effectiveness [20]. Vehicle routing optimization is critical, as returned products can have multiple destinations, such as repair facilities, resale channels, remanufacturing centers, or recycling locations. Intelligent routing and selection of collection points can significantly enhance efficiency and reduce environmental impact. However, while the technical sophistication of these models continues to advance, most studies remain focused on algorithmic performance, paying less attention to their practical integration into organizational workflows and to stakeholder involvement in their design and implementation. Issues such as data quality, system interoperability, and the need for continuous updates in enterprise platforms are frequently acknowledged but rarely addressed in detail.

Other papers suggest the practical application of return forecasts. Unpredictable return volumes force firms to over-allocate resources, leading to excess labor costs, inefficient warehouse utilization, and mismanaged inventory. The application of forecasting models enables data-driven resource allocation, thereby lowering unnecessary expenses [15]. While Frei focuses on the societal and environmental consequences, Adenuga focuses more on the human factor, introducing AI-driven workforce forecasting and Digital Twin technology to optimize staff scheduling and warehouse capacity, especially during peak periods or crises such as pandemics. This approach complements technical forecasting by ensuring that human resources are dynamically aligned with predicted return flows [2].

Furthermore, forecasting facilitates reducing “avoidable” returns and supports strategic planning for product reuse, resale, or recycling, aligning with sustainability goals [10]. The practical applications of return forecasting extend to multiple operational domains. For instance, weekly return forecasts can be used to verify handling and storage capacity requirements. Predictable return flows also enable faster reintegration of products into stock, reducing the need for over-ordering. Moreover, forecast insights allow the design and testing of differentiated return policies, such as implementing stricter rules for product categories with high return risk while maintaining leniency for loyal, low-risk customers [48]. Secondary consequences of using predictive models include increased customer trust and satisfaction. Ineffective return handling, often evidenced by delays in refunds, exchanges, or stock replenishment, diminishes customer loyalty. Predictive models address these issues by ensuring timely refund processing and faster restocking, thereby improving customer retention.

From a systems perspective, this research could further examine the integration of AI-based forecasting models into enterprise resource planning (ERP) systems. Such integration would allow predictions to automatically trigger workflows, including warehouse space reservation and the scheduling of pick-up routes. Lastly, return forecasts can be linked to sustainability monitoring tools, such as CO2 calculators, to assess the environmental impact of various return policies. This provides e-tailers with a proactive means of aligning operational decisions with ecological and regulatory requirements [38].

Finally, the literature highlights the critical role of data and information management. Accurate, integrated, and high-quality data on products, orders, and customers is essential for reliable return predictions and process optimization [3]. Research emphasizes the need for advanced data integration techniques, real-time monitoring, and AI-driven validation tools to ensure data consistency across systems. Moreover, AI solutions must be integrated into existing enterprise systems, such as ERP systems. Continuous updates of replenishment times within ERP systems are necessary to maintain stock accuracy and prevent returns caused by supply chain deficiencies [32]. In the paper of Gry et al., it is recommended to gain further practical insights into the integration of the daily system [20]. Alzoubi (2025) adopts a broad perspective, examining the integration of AI in reverse logistics within the context of the circular economy. The work identifies that different reverse logistics tasks, such as mechanical sorting versus analytical decision-making, require tailored forms of AI, and introduces the concept of "product passports" to track the full lifecycle of goods for responsible recycling and disposal. Alzoubi also raises important ethical considerations, including algorithmic bias and data privacy in consumer analytics. Despite their differences, all three underscore the necessity of robust data management and the integration of AI into enterprise systems.

In conclusion, while the field has improved greatly in modeling and operational efficiency, future research should focus on bridging the gap between technical innovation and practical implementation. This includes addressing stakeholder involvement, system integration, and ethical challenges. The predictive models are not designed for stakeholder use or for finding strategies to address their most significant challenge in return logistics. This gap highlights the need for a practical, innovative tool to help e-commerce companies optimize their return flows.

1.2.2. Process-oriented approaches on improvement strategies

In addition to data-driven approaches, several studies emphasize lean, process-oriented methods to improve efficiency. Frei et al. advocate for a lean supply chain strategy, in which waste is systematically identified and eliminated [16]. Their work introduces a taxonomy of return codes that enables retailers to trace the causes of high return rates and implement targeted countermeasures. By mapping product information and communication flows, their research enhances understanding of how packaging, transport, and decision-making processes affect the overall efficiency of return logistics. Emerging strategies also explore more customer-centric approaches. Customer-to-customer (C2C) return logistics has been proposed as an innovative concept in which returned items are directly shipped to the next customer. This reduces transportation steps, packaging waste, and lead times, and can be incentivized through discounts, loyalty points, or sustainability messaging [14]. However, this approach would have operational risks and is hard to scale. Zennaro et al. propose a strategy that shortens the waiting time for consumers after ordering and gives them the option to return the product immediately if they do not want it, making it easier to resell the returned product [50]. In addition, personalized return policies are being investigated to reduce opportunistic or fraudulent returns while maintaining customer satisfaction [49]. Additionally, Yang et al. (2020) propose the use of a virtual reality webroom for pre-purchase fitting, aiming to reduce product returns. However, this technology cannot fully replicate the tactile experience of handling a physical product, and its applicability is largely limited to the fashion and accessories sectors. Suggested future research by Frei et al. (2022) includes exploring the potential of AI to detect fraudulent returns, as well as examining the benefits and challenges of implementing sustainable return strategies and conducting environmental impact assessments.

Summarized, alongside data-driven methods, there are some proposals to reduce returns or handle them more sustainably. Frei et al. advocate for systematically identifying and eliminating waste. Customer-centric innovations, such as customer-to-customer (C2C) return logistics, aim to reduce transport and packaging waste, though scalability and operational risks remain concerns. Other approaches include immediate return options to facilitate resale, personalized return policies to reduce fraud, and virtual-reality webrooms for pre-purchase fitting, though this is limited to fashion and cannot fully replicate the physical product experience. Future research is suggested to explore AI for fraud detection and to assess the sustainability and environmental impact of these emerging strategies.

1.2.3. Conclusion

Recent advances in return logistics research have led to significant progress in predictive modeling, operational efficiency, and sustainability, with AI and machine learning at the forefront. Data-driven methods enable firms to anticipate return volumes, optimize resource allocation, and design more effective inventory and pricing strategies. However, most studies still prioritize algorithmic performance over practical integration, often overlooking challenges related to data quality, system compatibility, and stakeholder involvement. This limits the real-world impact and adoption of these innovations.

At the same time, process-oriented and mapping approaches offer valuable insights into the operational and behavioral drivers of returns. These methods help retailers trace the root causes of high return rates and implement targeted interventions, but they also face scalability and implementation challenges.

Critically, the literature reveals a persistent gap: models and process improvements are rarely designed with end-user needs and organizational realities in mind. Combining data-driven analytics with mapping and process-oriented strategies offers a better understanding of the return problem, enabling the development of practical, stakeholder-informed solutions. Future research should focus on bridging technical innovation with operational feasibility, integrating robust data management, addressing ethical considerations, and actively engaging stakeholders. Only through this integrated approach can e-commerce companies truly optimize their return flows, reduce waste, and enhance both customer satisfaction and sustainability.

1.2.4. Research gap

This chapter reviewed the existing literature on return logistics management and explored strategies for optimizing return flows. Existing research tends to focus either on technical optimization (e.g., predictive modeling, AI/ML applications) or on process mapping and lean interventions, but rarely integrates these approaches in a way that is actionable for practitioners. First, while machine learning and advanced analytics have improved the ability to forecast returns and inform strategic decisions, these models are often developed in isolation from the practical realities of organizational workflows. Key challenges such as data quality, system integration, and the need for continuous updates in enterprise environments are frequently acknowledged but seldom addressed in depth. Moreover, the involvement of stakeholders in the design, implementation, and adoption of these predictive systems is largely overlooked, limiting their real-world effectiveness and scalability.

In these papers, it is unclear how these models will improve return flows, and stakeholders are not incorporated into their design. So, the consideration of multiple stakeholder requirements and the inclusion of iterative, formative evaluation are underrepresented in current studies. Furthermore, the literature tends to focus on strategies to influence consumer behavior, leaving the responsibility of the e-tailers' decision-making behind. These gaps underscore the need for research to develop and evaluate a practical, stakeholder-informed artifact that supports data-driven decision-making and improves the efficiency and sustainability of return logistics. Furthermore, the literature review on predictive features found different combinations of associated features with returns.

Second, process-oriented and mapping approaches provide valuable insights into the operational and behavioral drivers of returns, such as waste identification, customer-centric innovations, and return policy design. However, these strategies often face challenges related to scalability, operational risk, and limited applicability across product categories. Furthermore, many interventions remain focused on influencing customer behavior, with less attention paid to organizational decision-making.

Finally, while a wide range of product, order, and customer attributes have been identified as predictors of returns, there is little consensus on how to combine these features effectively in practice, and few studies systematically evaluate the impact of integrating multiple predictors.

In summary, the literature reveals a persistent gap: there is a lack of comprehensive, stakeholder-informed frameworks that combine data-driven analytics with process mapping and practical interventions. Future research should focus on bridging technical innovation with operational feasibility and active stakeholder engagement. Only through this integrated approach can e-commerce companies develop effective, scalable, and sustainable solutions to optimize return flows and reduce the economic and environmental burden of product returns.

1.3. Scope and research questions

1.3.1. Research objectives

This thesis focuses on designing an artifact that helps companies improve their return flows. The objective is to develop a tool that addresses the core challenges identified in the current return logistics system and aligns with the needs of the key stakeholders. By incorporating their requirements, the tool aims to provide a practical, data-driven solution that optimizes operational efficiency, enhances customer satisfaction, and supports sustainable return management. Based on stakeholder requirements, the objectives of the artifact will be stated. The integration of both data-driven and process approaches requires a clear understanding of the problem's environment. The Design Science Research (DSR) methodology, as outlined by vom Brocke et al., provides an appropriate foundation for developing an artifact. A detailed explanation of its application within this study is presented in Section 2.

1.3.2. Research scope

In this thesis, data provided by a PwC client serve as the empirical foundation for the design and evaluation of the artifact. The dataset originates from a European e-commerce retailer operating in the electronics and fashion sectors and consists of two distinct datasets containing sales-order and return-order information. This research will focus solely on e-commerce retailers, excluding retailers with brick-and-mortar stores.

1.3.3. Research questions

Based on the identified problem, research objective, and scope, the main research question has been formulated as follows:

How can a reduction of returned goods in fashion- & electronic e-commerce be achieved by designing a framework that clusters high- and low-risk orders and products?

To address the main research question in a structured manner, a set of sub-questions was developed.

Sub-questions

1. Which factors are shown by literature and historical data to significantly affect the return rate?
2. How are retailers currently handling returns, and what limitations exist in these approaches?
3. What is the aim and structure of a return flow optimizing framework, and how can it help the retailer optimize their system?
4. What orders and products are high or low risk and what are their characteristics & what strategies can help reduce this risk?

1.4. Report outline

This chapter has introduced the problem context, highlighted the relevance of product returns within the e-commerce sector, and outlined the overarching research objectives and corresponding research questions. Chapter 2 then presents the methodological foundation of the study, describing the Design

Science Research framework and detailing the data collection and analytical procedures employed. Chapter 3 provides a comprehensive review of the existing literature, examining the drivers of product returns, current return management practices, and documented strategies to reduce return volumes. Chapter 4 reports the statistical analysis of product- and order-level attributes associated with return behavior, establishing the empirical basis for subsequent modeling. Chapter 5 complements these findings by examining current return-handling practices, drawing on insights from expert interviews and literature. Chapter 6 introduces the design of the proposed framework for identifying high- and low-risk product and order segments. Chapter 7 then presents the case study of the framework, showing the clustering results, interpreting the identified patterns, and formulating targeted return-reduction strategies, followed by their validation through stakeholder input. Finally, Chapter 8 synthesizes the main findings, reflects on the study in relation to existing literature and its methodological and practical limitations, and concludes with recommendations for both practice and future research.

2 Methodology

This chapter outlines the methodological approach adopted in this study. Guided by the Design Science Research (DSR) framework, the research process comprised four main steps: conducting a literature review, carrying out interviews, preparing and analysing the dataset, and performing clustering analyses to identify patterns relevant to returns. The specific methods applied in each step are described in this chapter.

2.1. Design Science Research framework

This study will use the Design Science Research (DSR) method to address the research questions. This method is a problem-solving paradigm that seeks to enhance human knowledge by creating innovative artifacts [8]. DSR aims to generate knowledge of how things can and should be constructed or arranged (i.e., designed), usually by human agency, to achieve a desired set of goals; this knowledge is referred to as design knowledge (DK). The methodology in this research is based on vom Brocke's work. There are three different categories within this framework: environment, design, and knowledge base. Within this environmental context, the problem space is explored. The needs are subsequently identified and assessed in relation to the organization's strategic orientation, structural configuration, cultural characteristics, and established work processes. The knowledge base comprises theories, methods, and instruments for designing artifacts within the solution space. DSR analyses the academic knowledge base and assesses the extent to which design knowledge is already available to solve a problem of interest. The design category is based on both the environment and the knowledge base. The design activities comprise "build" and "evaluate" activities, typically carried out in multiple iterations.

2.1.1. DSR Process

There are six activities within the cycle of this method. The process begins with problem identification and motivation, in which the research problem is defined and the value of addressing it is established. Justifying the relevance of a solution serves both to motivate the researcher and to help the audience understand the importance of the problem, which requires insight into its current state and broader significance. The second activity is to define the objectives for the solution, derived from the problem definition and the knowledge base. These objectives can be quantitative or qualitative and describe how a new artifact is expected to support problem-solving. The third activity is design and development, in which the artifact is created by determining its desired functionality, specifying its architecture, and constructing the actual artifact. This is followed by a demonstration in which the artifact is applied to one or more instances of the problem to demonstrate its practical utility. The fifth activity is evaluation, which measures how well the artifact supports the solution by comparing the defined objectives with observed results in context. Based on these outcomes, the process can iterate through the design and development stage to improve the artifact or continue to the final step. Evaluation within this method can be carried out at different stages of the process, each focusing on specific aspects of the research and artifact.

2.1.2. Application of DSR framework

The application of this methodological approach is directly aligned with the problem statement outlined in Chapter 1. Guided by this framework, the research is structured around one central research question, supported by several sub-questions that collectively inform the iterative cycle of the design science research process. The primary objective is to develop a framework that enables e-commerce platforms to optimize their return flows. Accordingly, the main research question is formulated as follows: *“How can a reduction of returned goods in fashion- & electronic e-commerce be achieved by designing a framework that clusters high- and low-risk orders and products?”* This overarching question is further elaborated through a series of sub-questions designed to address specific aspects of the research problem.

The first sub-question is *“Which factors are shown by literature and historical data to significantly affect the return rate?”* As outlined in the problem definition, the e-commerce sector faces major challenges in handling returns and managing the high volume of returned products. To develop a solution, it is essential to define the problem clearly and to obtain a comprehensive understanding of the current system and the reasons for the high return rate. This sub-question concerns the environmental space within the DSR framework. To address it, a literature review will be conducted alongside expert interviews and data analysis. This approach allows to map the important factors that affect the return rate. By knowing these factors from literature, experts, and data, a well-defined framework can be designed.

The second sub-question addresses current practices and limitations in return handling. This component of the research seeks to explain the existing processes and the constraints inherent in the approaches. Specifically, it involves a thorough examination of the operational procedures, supporting technologies, and infrastructures currently employed in return logistics. By systematically mapping these elements, the research aims to clearly visualize the problem space and identify the principal limitations of current practices. Addressing this sub-question will facilitate the identification of key bottlenecks within the system. To this end, qualitative data will be collected through interviews with supply chain and retail experts, providing critical insights into both operational realities and areas for improvement.

To address the third sub-question, the objectives and structure of the return flow optimizing framework are defined by integrating insights from literature, expert interviews, and data analysis. The framework’s goal is to reduce return volumes by identifying high- and low-risk products and orders, and by supporting the development of targeted strategies. This aligns with the design phase of the DSR methodology.

The fourth sub-question: *“What orders and products are high or low risk, what are their characteristics, and what strategies can help reduce this risk?”* focuses on the analysis of clustering results. In this phase, the output of the clustering analysis will be systematically examined to identify and label clusters as high- or low-risk based on their cluster return rates. The cluster characteristics serve as the foundation for developing targeted strategies to mitigate return risk. The proposed strategies are subsequently validated through interviews with supply chain and retail experts, thereby incorporating stakeholder feedback into the research process. The methodological approach applied in this phase, grounded in Design Science Research (DSR), is illustrated in Figure 2.1.

Table 2.1: Overview of Research Questions and Methods

Research question	**Method**
Which factors are shown by literature and historical data to significantly affect the return rate?	Exploratory data analysis, literature review
How are retailers currently handling returns, and what limitations exist in these approaches?	Interview supply chain experts
What is the aim and structure of a return flow optimizing framework, and how can it help the retailer optimize their system?	Designing conceptual framework
What orders and products are high or low risk, what are their characteristics, & what strategies can help reduce this risk?	Clustering Analysis & Interview experts

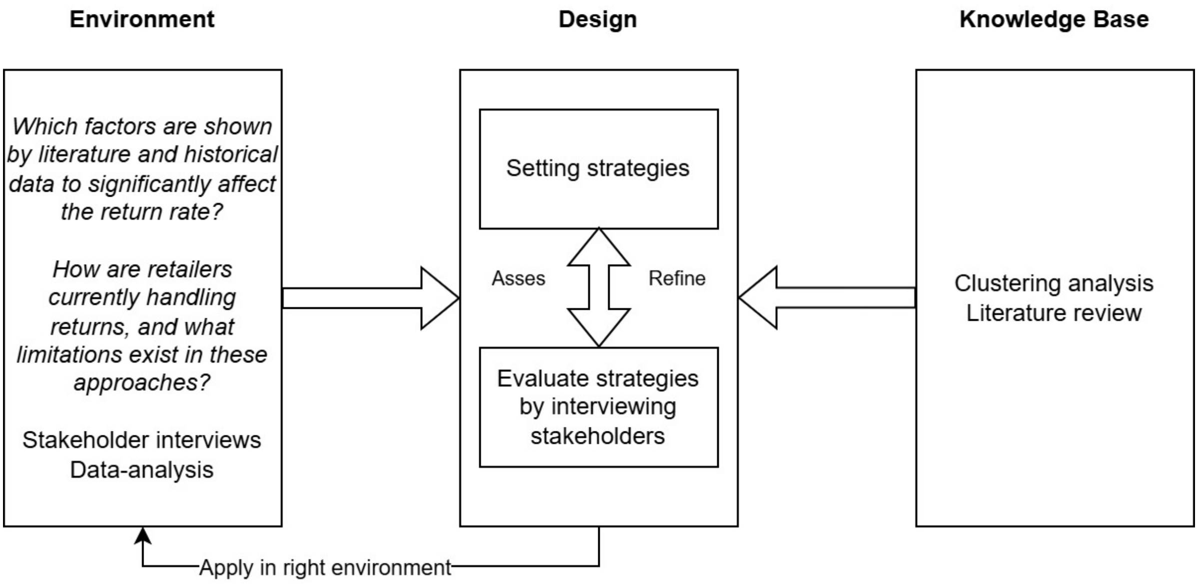


Figure 2.1: Methodology based on DSR framework

2.2. Literature review

A literature review aims to collect and analyze existing research in a transparent, structured, and reproducible manner. The process involves four key stages. First, the review's purpose and scope are clearly defined. Next, search strategies are applied to locate studies relevant to the topic. These studies are then assessed and filtered based on predetermined criteria. Finally, the selected studies are analyzed to draw meaningful conclusions [43].

2.2.1. Purpose of the literature review

The literature review aims to explore three different things. The questions below are part of the sub-questions and should be answered by finding this in the literature.

1. Which factors are shown to be associated with returns?
2. How is return management organized currently?
3. What strategies exist in the literature to reduce returns?

Firstly, it is necessary to get insights into which factors are affecting the return rate. This way, these data variables can be incorporated into the data analysis. Additionally, it is important to understand how e-commerce companies operate, how they handle returns, and the strategies and policies they use.

2.2.2. Search techniques

A comprehensive search was conducted to identify studies aligned with the purpose and scope of this research, as defined in the section above. The majority of sources were retrieved via Scopus, using the search strings listed in Table 2.2. These keywords were carefully selected to ensure the review's replicability. The search was restricted to publications from the past ten years, starting in 2015. Table 2.2 also displays the number of studies initially identified. After screening titles and abstracts, only the articles shown in the third row of the table were considered relevant to the research objectives. To further refine the selection, backward and forward snowballing techniques were applied, resulting in the final set of articles, as detailed in Table 2.3.

Sources found

By using the search string above, several papers were found. This section shows the results of the papers used and how they were found.

Part of literature review	Search string	Papers found
Factors that influence the return rate	“factors” OR “features” OR “variables” AND “affecting” OR “influencing” OR “relationship” AND “returns” OR “return rate” AND “e-commerce” AND “forecasting” OR““predicting”	(Karl, 2025) snowballing from this systematic literature review. [27]
Current return management	“Return management” OR “return policy” OR “return strategies” AND “e-commerce” OR “online retail” AND “Europe”	• (Stevenson et al. 2024) (snowballing from this systematic literature review [45] • & (Frei et al. 2022) [16] • (Frei et al. 2019) [17]
Strategies to reduce returns	“Return reduction” AND “strategies” OR “strategy” OR “interventions” AND “returns” AND e-commerce” OR “online shopping” OR “online retail”	(Duong et al. 2025) snowballing from this systematic literature review [11]

Table 2.2: Literature search scopes, search strings, and their yields

Papers found	How it was found
Karl (2024) [27]	Search string
Urbanke et al. (2015) [28]	Snowballing from Karl (2024)
Asdecker & Karl (2018) [4]	Snowballing from Karl (2024)
Cui et al. (2020) [9]	Snowballing from Karl (2024)
Shang et al. (2020) [42]	Snowballing from Karl (2024)
Imran & Amin (2020) [24]	Snowballing from Karl (2024)
Hofmann et al. (2020) [21]	Snowballing from Karl (2024)
Mishra & Dutta (2024) [32]	Search string
Niederlaender et al. (2025) [36]	Search string
Makkonen et al. (2021) [31]	Snowballing from Niederlaender et al. (2025)
Duong et al. (2025) [11]	Search string
Rezaei et al. (2018) [40]	Search string

Table 2.3: literature found on influential factors for return rate

Papers found	How it was found
Stevenson et al. 2024 [45]	Search string
Frei et al. 2022 [16]	Search string
Frei et al. 2019 [17]	Search string
Zennaro et al. 2021 [50]	Snowballing from Stevenson et al. 2024
Walsh et al. 2014 [46]	Snowballing from Stevenson et al. 2024
Wang et al. 2017 [47]	Snowballing from Stevenson et al. 2024

Table 2.4: literature found on current return management

Papers found	How it was found
Duong et al. 2025 [11]	Search string
Duong et al. 2022 [12]	Snowballing from Duong et al. 2025
Zhang et al. 2022 [51]	Snowballing from Duong et al. 2025
Gathke et al. 2021 [18]	Snowballing from Duong et al. 2025
Baldi et al. 2024 [6]	Snowballing from
Janakiraman et al. 2016 [25]	Snowballing from Duong et al. 2025
Bechwati et al. 2005 [7]	Snowballing from Duong et al. 2025
Mollenkopf et al. 2011 [34]	Snowballing from Duong et al. 2025
Walsh et al. 2014 [46]	Snowballing from Stevenson et al. 2024 (see Table 2.4)

Table 2.5: literature found on strategies to reduce returns

2.3. Interviews

Given the wide array of strategies associated with return logistics, it is essential to systematically clarify these approaches and assess whether any constitute best practices. The interviews conducted as part of this research are designed not only to explain current practices and the use of models or data-driven solutions, but also to explore their practical implications within organizations. Furthermore, these interviews serve to validate proposed strategies through direct engagement with relevant stakeholders.

Interviews will be conducted with supply chain and retail experts to address questions about current return handling and data collection. In a subsequent phase, additional interviews will focus on validating strategies aimed at reducing return rates. The chosen interview format is semi-structured, employing a flexible guide of open-ended questions to facilitate the collection of in-depth qualitative data. This approach combines systematic coverage of predetermined topics with the ability to pursue emergent themes, ask follow-up questions, and adjust the discussion sequence as needed. Such flexibility is particularly well-suited for investigating the complexities of return logistics strategies, the adoption of analytical models, and their organizational implications.

The insights gained from these interviews will inform process mapping, the identification of key factors, and the recognition of principal challenges within return logistics. Access to contact information for supply chain and retail experts is, therefore, a prerequisite for the successful execution of this research component.

2.4. Data analysis

In this analysis, potential factors influencing the return rate were systematically examined to assess the presence of significant relationships. A comprehensive methodological approach was employed, incorporating correlation analyses for numerical variables and chi-square tests for categorical variables to identify statistically significant differences in return rates across groups. Data pre-processing procedures included rigorous assessment of data quality and the identification and treatment of missing values and outliers.

2.4.1. Data needed

The analysis requires high-quality data to develop a reliable clustering algorithm. Although several potential features were initially considered, this study focuses specifically on order-level and product-level variables. Strategies are derived from these levels to prevent discriminatory outcomes that could arise from incorporating customer-level features. General attributes are excluded, as they do not yield actionable insights into the underlying drivers of return-rate variation. Potential risks include incomplete datasets and delays in data availability; these are mitigated through systematic data validation procedures.

Used datasets

This section provides an overview of the available data used in this study. The dataset employed in the analysis is derived from sales orders and return orders retrieved from the Finance & Operations inventory system of a retail company active in e-commerce in Europe. These data are proprietary and not publicly accessible. For each sales order, the variables listed in table 2.6 are available and form the basis of the analyses conducted in this research.

Pre-processing data set The data used in this study were extracted from the Finance & Operations (F&O) system and are available in two primary datasets. The first dataset contains order-level information, structured into a header and corresponding order lines. The header includes details such as shipping carrier, shipment method, total order value, number of items in the order, and the order date and time. The order-line data provide product-specific attributes, including product price, size, color, category, and the quantity of each item purchased. An equivalent structure exists for the return dataset, which mirrors the sales data for returned items.

To construct a unified analytical dataset, sales order-level information was first merged with return order-level information using the webshopreference as the unique key. A binary return indicator was subsequently created, coded as 1 when a product or order appeared in the return dataset and 0 otherwise. The same merging procedure was applied at the product level, resulting in two final datasets: one at the order level and one at the product level.

Both datasets were then systematically explored to assess the nature and quality of the available information. Only variables relevant to order- and product-level analyses were retained. Non-essential identifiers, such as order IDs, were removed, with the exception of the webshopreference, which remained necessary for linking sales and return records. Duplicate entries were identified and removed. Date- and time-related variables were converted to standardized formats, textual fields were normalized, and categorical variables were harmonized to ensure consistent representation across records. Missing values were subsequently examined. Observations without a webshop reference were excluded, as they did not correspond to online transactions and therefore fell outside the scope of this e-commerce-focused study. For product-level data, missing product sizes were imputed as “unknown,” acknowledging that size may not be applicable for certain product categories. Product categories were completed using product descriptions and item identifiers. Outliers were also evaluated to determine whether they represented plausible values or needed to be removed.

Given the high cardinality of variables such as categories, sizes, and colors, these attributes were aggregated into broader groups to mitigate sparsity and prevent categories represented by only a few observations from disproportionately influencing the analysis. The retail company maintains 137 distinct product categories, which are subsequently aggregated into 12 broader categorical groups. The full pre-processing procedure yielded a cleaned, structured dataset containing the final set of features used in this research. Finally, it should be noted that the data were exported on 3 October 2025 and cover a period of 6 months.

Order level information	**Product information**
Delivery mode (carrier combined with method)	Product category (12 categories A-L)
Shipment method (Standard, Pickup, Express)	Product color (24 colors)
Shipping carrier (PostNL, DHL, DHLDE, InPost, SEUR)	Product size (XS, S, M, L, XL, One Size, Sheets)
Order creation time (date and time)	Product price (numerical)
Number of items in order (numerical)	Number of products (numerical)
Total order amount (numerical)	
Returned Yes/No (Binary)	Returned Yes/No (Binary)

Table 2.6: Used dataset

2.4.2. Analysis tools

The data analysis and clustering procedures are conducted using a combination of statistical and machine-learning techniques. All analyses are performed in Python due to its flexibility and its extensive ecosystem of libraries for data processing, modelling, and evaluation. Data cleaning, transformation, and feature engineering are carried out using libraries such as pandas and numpy, ensuring that the dataset is consistent, properly structured, and suitable for subsequent analytical steps.

For the clustering component of this research, multiple algorithms are examined, including k-prototypes, CAVE, and Latent Class Clustering (LCC). These methods are selected to accommodate the mixed-type nature of the data and to enable a comparative assessment of clustering performance across different methodological approaches.

2.4.3. Clustering analysis

Clustering is an unsupervised machine learning technique that groups similar data points into clusters based on shared characteristics, without relying on labeled data. Its objective is to ensure that points within the same cluster are more similar to each other than to those in different clusters, enabling the discovery of natural groupings and hidden patterns in complex datasets. In the context of product returns, literature and interviews highlight that the primary challenge is the high volume of returns. Clustering analysis can uncover patterns in large datasets, allowing clusters to be classified as “high risk” or “low risk.” By examining high-risk clusters in detail, targeted interventions can be developed to reduce returns. Additionally, comparing common characteristics between high-risk and low-risk clusters provides valuable insights for improving return management strategies. Clustering offers interpretability and efficiency: after grouping products and orders, return rates can be calculated per cluster to identify high-risk segments. Furthermore, modern clustering algorithms can handle both numerical and categorical variables effectively, while maintaining a relatively low computational load [23]. This makes it a useful algorithm for exploring high- and low-risk orders and products. Further exploration of these clusters can provide insights into which feature combinations lead to higher or lower risk.

2.4.4. K-Prototyping

The most commonly used clustering algorithm for mixed data types is k-prototypes, which combines k-means (for numerical variables) and k-modes (for categorical variables) by employing Euclidean distance for the continuous part and simple matching for the categorical part of each object.

The k-prototypes algorithm begins with an initialization phase in which the number of clusters k is selected. This is typically done using the elbow method: the algorithm is run for various values of k, and for each run, the total clustering cost (the sum of numerical and categorical distances from each point to its cluster prototype) is recorded. These costs are plotted against the number of clusters, and the “elbow” point, where increasing k yields diminishing returns in cost reduction, indicates the optimal number of clusters. This point represents a good balance between model complexity and fit.

Once the optimal k is chosen, the clustering process begins. For each data point, the dissimilarity to all cluster prototypes is calculated, and the data point is assigned to the cluster with the lowest dissimilarity. The cluster prototypes are then updated: for numerical features, the mean is used; for categorical features, the mode is used. This assignment and update process repeats until cluster assignments stabilize or a maximum number of iterations is reached.

2.4.5. CAVE clustering

However, the standard k-prototypes algorithm has several limitations, particularly in its treatment of categorical variables and the equal weighting of all features. Recent literature has proposed several enhancements to address these shortcomings. One notable improvement is the weighted centroid method, which represents categorical centroids using the full distribution of category frequencies within a cluster, rather than collapsing to a single mode. This approach, as suggested by Dutt et al. (2024), provides a more nuanced and accurate representation of categorical centers, especially in clusters with diverse category distributions [13]. Another refinement is the incorporation of Goodall similarity, which assigns higher weights to rare categories and encodes relationships between variables, thereby capturing more subtle patterns in the data. While this method increases computational complexity, it can yield more meaningful clusters in heterogeneous datasets. Another significant improvement is the adoption of the CAVE algorithm (Clustering Algorithm based on Variance and Entropy), which explicitly balances the contributions of numerical and categorical attributes using variance and entropy measures, respectively [13, 26]. Unlike the K-prototypes algorithm, which counts mismatches for categorical variables, CAVE uses entropy to quantify the degree of category overlap within each cluster. Low entropy indicates a cluster dominated by a single category, while high entropy suggests a more heterogeneous composition. For numerical variables, CAVE minimizes within-cluster variance, similar to k-means. This dual approach ensures that clusters are both numerically compact and categorically homogeneous, and it enables automatic assessment of feature importance. According to Ji et al. (2013), CAVE produces more accurate and interpretable clusters than traditional k-prototypes in mixed-type datasets [26]. Additionally, CAVE provides feature importance scores, enabling practitioners to identify which variables most strongly influence cluster formation.

In this algorithm, a different weight calculation is applied. For numerical features, the variance of each column is calculated. Variance indicates the degree of spread in the data for that feature. Features with higher variance receive a higher weight because they contribute more to distinguishing between data points. The weights are normalized so that they sum to 1 (or 0.5 if combined with categorical features). A feature with little variation contributes less to differentiating clusters. To calculate the distance for categorical features, the feature’s entropy is used. A feature with high entropy (many different values, evenly distributed) receives a lower weight because it is less distinctive. The greater the variation, the more important the feature.

The weights for numerical and categorical features are normalized separately so that both groups together sum to 1 (usually 0.5 each). They are then combined into a single weight vector. The paper by Ji et al. builds on the CAVE algorithm from Hsu & Chen (2007) and introduces an improved method for determining the "significance" (weight) of features when clustering mixed data [22] [26]. The paper states that features that contribute more to distinguishing clusters (i.e., with lower within-cluster distances or higher discrimination) should receive higher weights. Conceptually, this approach is consistent with the principle advanced by Ji et al., which builds on the CAVE algorithm by Hsu & Chen (2007) [26], [22].

2.4.6. Latent class clustering analysis

In addition to distance-based approaches, the literature also mentions probabilistic approaches. Recent work by Ghattas and Sanchez San-Benito (2025) shows a comprehensive comparative study of state-of-the-art clustering algorithms for mixed-type data. It further informs the selection of appropriate methods for practical applications. Latent Class Clustering, in contrast with K-Prototyping and CAVE, is a probabilistic approach. Their findings confirm that k-prototypes remains a widely used baseline due to its simplicity and scalability, but also highlight its limitations, particularly in scenarios with high cluster overlap or complex variable dependencies. Notably, LCC consistently outperformed these methods in terms of clustering accuracy, especially when clusters were spherical or when categorical variables dominated the dataset. Convex k-means and k-prototypes are recommended for datasets with a higher proportion of continuous variables and for cases with imbalanced cluster sizes [19].

Latent Class Cluster (LCC) is a model-based probabilistic clustering method that aims to reveal groups with similar characteristics. In this case, showing orders and product groups that have high or low risk of being returned [33]. The core principle of LCC is that a discrete latent variable explains the observed relationships among a set of indicators. Once this latent class variable is accounted for, these associations become negligible or statistically insignificant [30]. These clusters are similar based on certain observed characteristics, which are the factors identified for sub-question 1. The goal is to find a model with the smallest number of latent classes, which can adequately describe the associations between the indicators. In LCC, individual datapoints are assigned to different clusters with a certain probability of belonging to that cluster, rather than the deterministic approach where the distance is minimized, and a datapoint is assigned to one cluster. This is one of the main advantages of LCC compared to deterministic clustering. To select the optimal number of clusters for latent class clustering, models with varying cluster counts were compared using Bayesian Information Criterion (BIC), log-likelihood, and convergence stability. While log-likelihood improves with more clusters, BIC was prioritized as it penalizes complexity and identifies the point of diminishing returns. Iteration counts were also considered to ensure model stability [30].

3
Literature review

This section explores the existing literature on return logistics in e-commerce, with a focus on three key areas: the features associated with product returns, the current handling of product returns, and strategies developed to reduce return rates. By examining these dimensions, the review aims to provide a comprehensive foundation for the research presented in this thesis.

3.1. Associated features with returns

In Karl's (2025) systematic review, an overview of the predictors used in several return forecasting models is provided [27]. The paper synthesizes findings from 25 studies; the corresponding overview and feature inventory are presented in Appendix C.2. Table C.2 summarizes the predictors previously employed in return-prediction models reported in the literature. The following paragraph elaborates on the empirical and theoretical insights regarding factors associated with product returns.

Historical sales volumes provide useful signals for forecasting return patterns, as they capture behavioral dynamics such as impulse purchasing. Shang et al. (2020) demonstrate that models incorporating time-stamped historical return data can effectively predict future return quantities [42]. Similarly, Cui et al. (2020) show that product type significantly influences return volume, indicating that past return behavior is predictive of future returns [9].

In many papers, there are references to product attributes. Among product attributes, price is one of the most common. Some papers also include price discounts [27]. Price seems to increase returns [4]. According to Imran and Amin (2020), promotions seem to yield higher returns, which explains the behavior of impulse purchases [24]. According to Cui et al. (2020), product type significantly impacts return rates; however, this depends on the industry [9]. The study by Samorani et al. shows that brand perception influences return decisions [41]. Lastly, within the product attributes, the order and return history of a product can be considered a factor in increasing the risk of returns for that product [21]. The return history of product categories can, in line with this, also influence the return rate, as the study of Mishra & Dutta shows that the product category is strongly associated with returns [32].

Some papers also include customer attributes. In the paper of Asdecker & Karl, it is seen that gender and age can also influence the number of returns. Females are more likely to return products, and younger customers are less likely to return products quickly [4] [31]. This is also supported by Makkonen et al. (2021) [31]. Furthermore, customer features, such as return history, are more important than other product or transaction profiles [27]. Next to the order and return history of a product, the order and return history of a customer is also associated with return risks [41, 28].

Asdecker also suggests basket interactions with the return rate [4]. For example, a higher item count in a basket can lead to a higher return rate, as seen in the paper by Urbanke et al. [28]. This variable can explain bracketing behavior. Behavior, where consumers order more of the same in different sizes or colours and intend to send back the others. The paper of Niederlaender et al. dives deeper into basket composition. They explore the influence of the basket having, for example, several items in the same category or subcategory. This behavior will likely increase return rates [36]. The variable payment methods are also shown in different papers to be an important feature. Prepaid products are less likely to be returned [24]. In line with this, the study by Urbanke et al. (2015) assumes that a customer's payment method is associated with returns. The study of Mishra & Dutta (2024) also includes a lot of variables at the basket level. It uses the order date, order quantity, and the prices for the entire order. Furthermore, the shipment and payment methods are seen to influence the return rate [32]. The total order amount and the after-pay method show a strong association with return incidences. The study by Makkonen et al. (2021) shows that an order paid by invoice is more likely to be returned than orders paid by other methods [31]. This is supported by Mishra & Dutta (2024), who state that the cash-on-delivery payment method leads to higher return rates [32]. Appendix C.2 shows a table to compare the combination of variables used in different models. The table is adapted from Karl et al. (2025). It becomes evident that the specific combination of predictors employed in this study has not previously been documented in the literature.

Conclusion In summary, the literature identifies a wide range of features for product returns, spanning product, customer, and basket attributes, as well as payment and shipment methods. While individual studies highlight the importance of factors such as price, product type, customer demographics, and basket composition, there is no consensus on the optimal combination of features used for identifying return risks. The comprehensive overview by Karl (2025) reveals that many possible combinations of variables remain unexplored in existing models. The findings from this literature review inform the first sub-question, as the identified attributes are subsequently examined and statistically evaluated for their associations with returns.

3.2. Return logistics management

Return logistics are usually handled by the company itself. Processing returns can be a costly activity due to handling and transport costs. Furthermore, the returned products have lost value due to ageing, use, and technical changes. Thus, return logistics management must be well-organized. Return management can be broadly divided into two main categories: preventive and curative, encompassing four key tasks. The first task, returns prevention, aims to discourage product returns altogether by implementing measures such as return fees, stricter requirements, or incentives to keep products, thereby reducing unnecessary return flows. The second task, returns avoidance, focuses on proactively addressing the root causes of returns through product quality improvements, better product information, and campaigns that encourage deliberate purchasing decisions. Within literature, these two tasks, return prevention and avoidance, form the preventive dimension, which has been a primary focus in literature [45]. This dimension also includes returns promotion, which encourages the return of products with a positive net value by informing customers about collection and recycling programs and offering incentives. The second dimension is the curative dimension, which consists of effective returns processing. This aims to maximize value recovery from returned products through resale, refurbishing, material recovery, or donation. Despite their potential to enhance circularity and sustainability, returns promotion and effective processing remain underrepresented in the literature [45]. This highlights an opportunity to develop more comprehensive frameworks that integrate both preventive and curative approaches for well-organized return management.

A return flow encompasses a sequence of activities initiated when customers send back purchased goods. Effective management of this flow requires that returned items are properly registered, assessed, and, where necessary, repaired or refurbished before being reintegrated into inventory. The decision to resell, reuse, or dispose of an item depends on its condition and quality, to minimize value loss for the company. Return management, therefore, involves the planning, monitoring, and control of all return-related activities, ensuring both an environmentally responsible process for customers and cost efficiency for retailers. Within the broader context of reverse logistics, several sub-areas are considered: product retrieval (collection and transport), product inspection, product recovery (including recycling, reprocessing, and repair), inventory management, waste management, and reintegration into the forward supply chain [45]. The study by Frei et al in 2022 describes the following return process. The generic process synthesized from case studies begins with purchase and includes various return entry points: carrier, postal services, parcel shops, or drop-boxes. Returned items may follow diverse exit routes: resale, recycling, donation, manufacturer return, or disposal, depending on condition and value. Secure storage and transport practices, such as product cages and anti-theft measures, are critical during transfers to Returns Centers (RCs). At the RC, items undergo inspection, barcode scanning, and refund processing. Outcomes range from reintegration into stock (full price or discounted), minor refurbishment, or routing to secondary markets (outlets, charities, jobbers, auctions). Low-value items often default to the simplest disposal route, with sustainability considerations largely absent. Contractual arrangements with third-party distribution centers influence whether returns are processed for speed or value recovery [16]. The return logistics flow based on the literature is visualized in Figure 3.1 [45].

Reverse logistics processes include return authorization, transportation, acceptance of returned products, disposition, and information management. Often, there are five main product disposition activities: destruction, recycling, refurbishing, remanufacturing, and repackaging. Each action influenced economic performance, operational responsiveness, and service quality differently. To support reverse logistics operations, it is essential to determine the value and necessity of product reconditioning [50]. Another study highlights practices such as 'buy online, return in store.' Returned goods can be processed either in-store or via a postal carrier; however, items returned in-store are subsequently sent to a return distribution center. The process surrounding in-store returns is excluded from the current research, which focuses solely on online retailers. Products returned to the warehouse are returned to the manufacturer, restocked, sent for recycling, or sent to waste [17].

The study of Frei et al. (2022) also mentions several kinds of waste. The seven types of waste in return logistics can be categorized as follows: (1) over-processing of returned goods, (2) excessive inventory costs, (3) unnecessary transportation, (4) avoidable movement of personnel handling returns, (5) delays caused by poorly integrated processes, (6) defects in returned goods requiring additional handling, and (7) inefficient use of storage space by returned items. Ultimately, the processing of returned products aims to maximize value creation by directing goods toward their most optimal subsequent use, whether through resale, refurbishment, material recovery, donation, or responsible disposal.

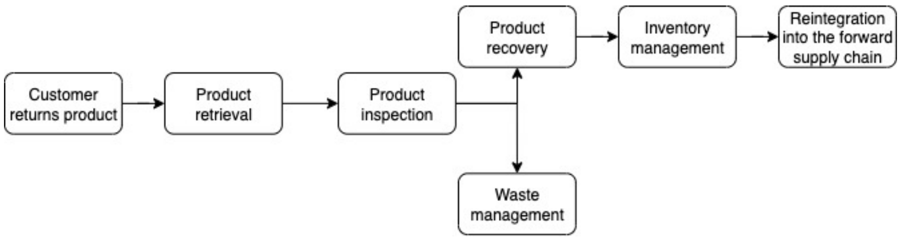


Figure 3.1: Return logistics flow based on literature [45, 50, 17, 16]

Conclusion The literature on return logistics management, as exemplified by Stevenson et al., offers a valuable conceptual distinction between preventive and curative dimensions. While the curative dimension provides a detailed account of operational activities, these studies tend to treat the process as a linear sequence of actions. Frei et al. (2022) further enrich this perspective by categorizing the types of waste generated during returns and highlighting the associated inefficiencies and sustainability challenges. However, a critical gap persists across these works: limited attention is paid to the allocation of responsibilities among stakeholders throughout the return process. The literature rarely specifies which stakeholders are accountable for particular tasks or decision points. Moreover, the balance between automation and manual intervention remains largely unaddressed, despite its practical significance for efficiency and scalability. As a result, while existing research maps the "what" of return logistics management, it falls short in clarifying the "who" and "how," limiting its utility for designing actionable, stakeholder-informed improvements. This conclusion serves as the foundation for the broader system context, within which a more comprehensive process flow is developed, incorporating both manual and automated components and stakeholder responsibilities.

3.3. Strategies to reduce returns

This paragraph highlights common strategies for reducing returns. Existing research has largely concentrated on managing product return volumes to address operational challenges within closed-loop supply chains [11]. However, a 2022 study claims that prioritizing management efficiency over prevention is a key reason why return reduction remains unresolved in practice. A deeper understanding of product return drivers can enable actionable countermeasures, an area that has received limited attention from marketing, operations, and information management perspectives [12].

Attributes related strategies This paragraph dives into strategies related to product features. Packaging emerges as a primary extrinsic factor significantly triggering returns, especially in categories such as electronics, home appliances, and groceries [11]. Consequently, the authors recommend that sellers invest in higher-quality, protective packaging, such as waterproof boxes or absorbent materials, and ensure gentle handling during storage and delivery to minimize damage and negative first impressions. Customer service is shown to have a significantly positive impact on product returns in the fashion category. Due to high product fit uncertainty, effective and responsive communication is essential to address customer concerns and reduce unnecessary returns. The study suggests that online fashion retailers should prioritize customer support and logistics services to enhance satisfaction. Product knowledge and feature clarity are also crucial. The analysis reveals that unknown or poorly understood primary features have a strong negative impact on experience products (e.g., books, home appliances), while secondary features affect both fashion and electronics. Increasing consumer knowledge of core and secondary features through detailed descriptions, guides, or interactive technologies can reduce return rates by aligning expectations with the product's actual attributes.

Retailers increasingly recognize the value of customer-based instruments in reducing returns, without negatively impacting sales. Customer-based instruments are designed to influence shoppers before and during the purchase process by providing comprehensive product information, thereby helping customers make better choices and avoid returns due to misfit. Examples include virtual try-ons, avatars, customer reviews, detailed product descriptions, and size guides. Advanced tools are not yet widely implemented, due to high costs or technical challenges. Easier implementations, such as size guides, detailed product information, and customer reviews, can also help customers make the right choices when ordering [46].

Product usage performance is identified as a strong predictor of returns across all categories. However, the likelihood of returns decreases when reviews emphasize reliability and when customers are prompted to reflect on both performance and reliability. Thus, sellers are encouraged to ask feedback specifically on these aspects, reminding buyers of the product's core value and usage satisfaction. Durability is essential for electronics and home appliances. To manage expectations and reduce returns, manufacturers and retailers should clearly specify the estimated product lifespan and provide transparent warranty policies [11].

Return policy strategies Return policy design is another critical factor in reducing returns. The research of Duong et al. (2025) demonstrates that lenient return policies, while potentially boosting sales, can also lead to higher return rates and increased return abuse. For electronics, in particular, stricter policies that increase the effort required to return (e.g., requiring original tags, receipts, or packaging) and limit the scope of eligible returns are effective in reducing both legitimate and abusive returns [11].

Another study shows that culture significantly moderates the impact of return policy design on product return behavior. Specifically, effort restrictiveness, such as requiring more steps to return an item, effectively reduces product returns among Western customers, but does not have the same effect in Eastern markets. Additionally, perceived customer-oriented norms (e.g., competitive intensity and customer-centric practices) increase product returns among Eastern consumers, whereas they have little impact on Western consumers [18]. For Western markets, implementing more effortful return processes can help reduce return rates; however, this approach may also negatively affect repurchase intentions. To balance operational efficiency and customer satisfaction, the authors recommend leveraging technology and outsourcing reverse logistics to specialized firms, thereby reducing both returns and associated costs [18]. The effortful returns are also discussed in the paper by Baldi et al. (2024). It highlights that returns policies are no longer just an operational necessity but a strategic lever in shaping consumer perceptions and behaviors. Retailers are increasingly adopting more liberal and flexible returns policies, such as extended return windows, simplified processes, and multiple refund options. They not only reduce perceived risk for consumers but also drive higher purchase intent and loyalty. The trend towards one-size-fits-all, consumer-friendly returns policies reflects a broader shift in the industry, where the ease and speed of returns processing are directly linked to increased customer retention, sales growth, and higher return volumes [6].

Lenient or strict return policies are also highlighted in the paper by Janakiraman et al. (2016). It is believed that leniency increases purchases more than returns, despite the substantial expense associated with product returns. It suggests varying five factors (monetary, time, effort, scope, and exchange leniency) in return policies to impose restrictions that dissuade returns or to offer leniency that encourages purchase. A shorter return time period and a full refund versus a longer return time period and an 80% refund influence purchase and return decisions differently. The first policy is more likely to affect purchases, whereas the second is more likely to affect returns. Lower effort return policies also seem to stimulate purchases effectively. If retailers aim to reduce returns, they should implement more restrictive policies by limiting which products are eligible for return, extending the return deadlines, and allowing greater flexibility in exchange options. This combination has proven more effective at discouraging unnecessary returns. Allowing returns on products purchased during sale periods increases returns, suggesting that enforcing restrictions on more price-sensitive consumer segments would reduce returns. Exchange leniency significantly reduces product returns compared to policies that allow exchanges only. This may be because consumers with minor issues are more likely to opt for an exchange when the option is clearly presented, rather than initiating a full return [25]. However, a study by Zhang et al. (2023) suggests that tightening returns policies, such as limiting returns to exchanges or imposing shipping fees, can reduce return rates and fraud. Still, it may also create barriers for honest customers and negatively impact sales. As a result, most retailers prefer frictionless returns to support customer satisfaction and retention. Targeted adjustments, such as shortening the returns window and strictly enforcing refund procedures, can deter abuse with minimal impact on genuine customers. Clear communication of return policies at the point of purchase and consistent staff enforcement further support these measures and help protect the interests of honest consumers [51].

The study by Walsh et al. (2014) also highlights the effortful return policy; they note this as a procedure-related product return management system, a non-monetary strategy that influences the likelihood of returns after a purchase. These include making the return process less transparent, refusing service to frequent returners, and communicating to customers the undesirability of returns. Such “hassle costs”, like requiring customers to print their own return documents, make returns less attractive. Unlike monetary instruments, procedural instruments are widely used by retailers, who often manually review excessive returns and communicate directly with customers to understand the reasons. Limiting return channels will also require more effort to return ordered goods [46].

This study of Walsh et al. (2014) also shows other instruments to prevent returns and thus reduce the volume of returns. They suggest monetary, procedural, and customer-based instruments. Monetary instruments are mechanisms that provide a monetary incentive to the customer to keep the ordered products or reduce the risk of purchase for the consumer. Restocking fees and money-back guarantees are designed to influence consumer return behavior by either deterring returns or reducing purchase risk. While restocking fees serve as a punitive measure, money-back guarantees aim to increase consumer confidence and spending. However, qualitative findings indicate that most retailers are hesitant to implement these instruments due to regulatory constraints, skepticism about their effectiveness, and concerns over increased operational costs. Additional incentives, including discounts and shipping savings, are also viewed cautiously, as administrative expenses may offset potential savings. Walsh also mentions procedural instruments, including optimizing shipping times and ensuring safe packaging to reduce the number of returns [46].

Marketing strategies Results in a paper by Bechwati et al. (2005) show that when retailers use strategies like limited-time offers, customers don’t have enough time to convince themselves they made the right choice. As a result, these rushed decisions often lead to higher product return rates. Thus, using strategies that give consumers the time to generate cognitive responses will decrease the product returns [7]. This is supported by findings in the paper by Mollenkopf et al. (2011). This paper suggests aligning more between marketing and operations teams to find strategies that fit the market and the rretailer’scatalog [34].

Conclusion In conclusion, the current literature focuses on several strategies to reduce returns. Taking product information, operational, return policy, and marketing strategies into account. From considering packaging and shipping methods to having consumer-based instruments. Strategies that arise are clearly communicating the pproduct'sfeatures and lifespan. The trade-off between lenient and strict return policies is discussed in multiple papers, considering five aspects of leniency. Strategic efforts predominantly focus on modifying consumer behavior by adjusting the leniency of return policies. This observation highlights a potential gap in the literature: while consumer-focused interventions are well-explored, there is limited research on strategies that target organizational behavior and internal decision-making processes. The strategies and their corresponding hypotheses are summarized in Table C.3 and serve as input for addressing sub-question 4. The subsequent cluster analysis examines whether the hypothesized patterns can be empirically observed within the data and uses these strategies to build upon.

4
Features associated with return rates

In this chapter, the first sub-question will be addressed, identifying the factors that significantly affect the return rate. First, the conclusion of the literature review will be presented, followed by data analysis.

4.1. Associated features based on literature

To form data-driven strategies for reducing returns, it is important to know which features are associated with returns. Features are variables that can explain the returns on products or orders. In this case, the research focuses on which features are associated with the return rate. Return logistics rely on multiple data types, including customer, product, and order information. The literature review in section 3.1 shows possible features within the return logistics that could be associated with returns.

The existing literature identifies a broad spectrum of determinants associated with product returns, encompassing product-specific, customer-related, and basket-level attributes, as well as payment and shipment characteristics. Although individual studies emphasize the importance of factors such as price, product type, customer demographics, and basket composition, the field lacks consensus on the most effective combination of features for predicting return risk. As highlighted in Karl's (2025) comprehensive review, a substantial number of potentially informative variable combinations remain unexplored in current modeling approaches. Table 4.1 presents the variables derived from the literature review.

In the data analysis, associations between product- and order-level features and return outcomes were statistically assessed. General attributes commonly used in the literature were excluded, as they are unlikely to explain the drivers of returns for specific products or orders. Customer-level variables were omitted as well, given that interviews revealed such information is often not collected by retailers and may introduce discriminatory biases, which would conflict with the objective of developing internally actionable return-reduction strategies. Interviews further indicated that retailers typically record order-level attributes such as order value, number of items, order date and time, shipment method, and delivery method, as well as product-level attributes including product price, category, and dimensions. These variables, therefore, form the basis of the data analysis.

The dataset used in this research encompasses a broad range of consumer products typically sold in the intimate wellness and personal care segment. These items are categorised into twelve overarching product groups, ranging from fashion and apparel (Category F) to wellness products such as lubricants, massage items, and care articles (Category H). Additional categories include electronic or non-electronic personal devices (Categories I and L), accessories, and miscellaneous lifestyle goods (Category J). Each category groups products based on shared functional characteristics or usage contexts, enabling a structured analysis of return patterns across distinct product types. Importantly, while individual items within these categories may vary considerably in form and purpose, the analysis focuses solely on their product attributes such as size, color, and price level, rather than the specific nature of the products themselves.

Table 4.1: Overview of Features for Model Development from Literature

General features	Customer-level features	Product-level features	Order-level features
Total sales volume	Age	**Price**	**Total order price**
Historical return volume	Gender	**Product category**	**Item count**
	Past return rate	**Product dimensions**	**Order hour**
		Brand perception	**Shipment method**
		Order & Return history	**Delivery carrier**
			Discounts & coupons
			Payment method

4.2. Data analysis of potential associated features

This section presents the results of the data analysis aimed at identifying potential features for inclusion in the clustering model. For categorical variables, return rates were computed for each category, after which chi-square tests were applied to determine whether observed differences were statistically significant rather than attributable to random variation. For numerical variables, correlation analyses were performed to assess the strength and significance of their associations with return outcomes.

4.2.1. Results on data analysis on order level data

The first analysis examines order-level variables, including total order value, the number of products ordered, order timestamp, delivery mode, shipping carrier, and shipment method.

Descriptives The overall order-level return rate is 5.02%, with the majority of returned orders (81.36%) representing full-order returns. These results are presented in Table D.1 in Appendix D.

Return Rate by Shipping Method A significant variation in return rates was observed across different shipment methods. Standard shipping methods show a return rate of 4.27%, compared to 4.57% for deliveries to a pickup point. The other two shipment methods have been used just a few times and are negligible. Figure D.4 illustrates these differences, while Table D.2 demonstrates the limited usage frequency of the other shipment methods.

Return Rate by Hour of the Day Return rates also varied depending on the hour at which the order was placed, as can be seen in section D.1.3. The highest return rates were observed for orders placed during early morning hours (e.g., 02:00 with 8.18%, 01:00 with 7.48%, and 00:00 with 7.36%). In contrast, orders placed during the late morning and afternoon generally had lower return rates, with the lowest observed at 16:00 (4.38%) (see Table D.5).

In the binned analysis, hours of the day were segmented into four time periods: night, morning, afternoon, and evening. The results reveal noticeable differences in return rates across these time segments. In Figure D.6, these differences are visualized. A chi-square test confirms that these variations are statistically significant, indicating that the time of day when an order is placed is associated with differences in return rates, as presented in Table D.4. This phenomenon is also observed in the product-level analysis (Figure D.14). This pattern may reflect differences in customer behavior or decision-making processes across different times of day.

Correlation analysis To identify numerical factors associated with return likelihood, a correlation analysis was conducted. The results indicate weak but statistically significant positive associations between the number of items initially ordered and return occurrence (r=0.055, p <0.001), as well as between initial order amount and return occurrence (r=0.053, p <0.001), suggesting that larger and higher-value orders are marginally more likely to be returned. By contrast, the hour of purchase showed no significant association with returns (p=0.6041), suggesting that, when modeled as a continuous variable, order timing is not associated with returns. The corresponding correlation heatmap (Figure D.7) and results table (Table D.3) are provided in Appendix D.1.4.

Chi-square analysis To assess the influence of categorical variables on return rates, chi-square tests were performed. This statistical method evaluates whether there is a significant association between each categorical variable and the occurrence of a return. A significant result indicates that the variable is not independent of return behavior and thus should be considered in clustering analysis. The results show that the categorical variables are statistically significantly associated with whether an order is returned. Specifically, delivery mode, shipping carrier, and shipment method all show highly significant relationships with return occurrence (p< 0.001). This indicates that the way in which an order is delivered, as well as the specific carrier and service used, can influence the likelihood of a return. The table D.4 supporting this can be found in appendix D.

In addition to categorical variables, binned numerical variables were also tested for association with return behavior using chi-square tests. These include the initial number of items ordered, the order amount, and the order hour. All three variables show highly significant relationships with return occurrence (p < 0.001), as summarized in Table D.4. For example, order amount (χ^2 = 2247.11, df = 3) and initial number of items ordered (χ^2 = 1516.96, df = 4) exhibit strong associations, suggesting that larger orders and higher item counts are linked to different return patterns. Similarly, order hour (χ^2 = 315.42, df = 3) indicates that the time of day when an order is placed may also influence return likelihood. However, this contrasts with the correlation result, which was close to zero, suggesting that while categorical differences exist between time blocks, there is no strong linear relationship across the full range of hours.

Conclusion In conclusion, the analysis indicates that approximately 5% of all orders are returned, with 81.36% of these being full-order returns. While the return rate by hour of the day shows noticeable differences across time blocks, correlation suggests no strong linear relationship between order hour and return likelihood. In contrast, the chi-square test confirms statistically significant differences between time-of-day categories, highlighting that certain periods are associated with higher return rates. Furthermore, both the initial order amount and the number of products in an order exhibit a strong and significant relationship with return behavior. Finally, categorical factors such as delivery mode, shipping carrier, and carrier service also demonstrate highly significant associations with return occurrence, underscoring the importance of operational and logistical variables.

4.2.2. Results on data analysis on product level data

This part of the data analysis provides detailed product-level information, capturing key characteristics that define each product. Specifically, it includes attributes such as product category, color, and size, which allow analysis of product variations. In addition to these descriptive features, the analysis incorporates pricing information for each product, enabling evaluation of price-related trends. Furthermore, it records the number of items ordered per product, offering insights into bracketing behavior.

Descriptives The analysis indicates that 6.84% of all products sold are returned, presented in Table D.5. In the following chapters, the return rates are examined across different product categories. Statistical tests are applied to determine whether significant differences exist between groups and their respective return rates.

Return rate by color As shown in Table D.6, colors such as Red (11.41%) and Cream (9.60%) exhibit notably higher return rates compared to the overall average of 6.84%. Conversely, colors like Copper (3.55%) and Lavender (5.13%) have relatively low return rates. Figure D.11 visually reinforces these differences, highlighting the disproportionate return rates for specific colors. As shown, the color “Sand” exhibits a return rate of 33.3%; however, given that this estimate is based on only three orders, the observed percentage is not meaningful and should be interpreted with caution.

Return rate by product category Differences in return rates across product categories are evident in the analysis. As illustrated in Figure D.12, certain categories exhibit notably high return rates. For example, Category F reaches 18%, whereas others remain well below the overall average of approximately 6.84%. This steep decline from high- to low-return categories suggests that product type is a strong determinant of customer return decisions.

Return rate by product size As shown in Figure D.13, return rates vary across sizes, with all sizes being over average except for One Size and the only sheet size. These findings suggest that sizing plays a critical role in customer satisfaction, potentially due to fit issues or unclear size guidelines.

Return rate by product quantity As illustrated in Figure D.14, return rates tend to decrease as product quantity increases. Orders with a single item exhibit the highest return rates, while larger orders show progressively lower rates. This pattern suggests that customers who purchase multiple items may be more confident in their choices or less likely to return individual products. The relationship between order quantity and return rate was analyzed using a chi-square test for binned numerical variables, seen in D.4.

Return rate by sales price Figure D.14 shows that return rates vary across price segments, with higher-priced products generally exhibiting slightly higher return rates compared to lower-priced items. This trend may reflect increased customer expectations for more expensive purchases.

Correlation analysis To further explore the relationship between numerical variables and product returns, correlations were calculated. As shown in Table D.7 and Figures D.15, all examined variables exhibit statistically significant correlations with return status, although the effect sizes are relatively small. Product price shows the strongest positive correlation (r = 0.0134, p < 0.001), indicating that higher-priced items are slightly more likely to be returned. In contrast, ordered quantity of product(r = -0.0086, p < 0.001) and order hour (r = -0.0020,p< 0.05) display weak negative correlations, implying that larger order quantities and certain times of day are linked to slightly lower return rates. While these correlations are statistically significant, their small magnitudes suggest that these variables alone do not strongly associate with returns, but may contribute when combined with other factors in a multivariate analysis.

Chi-square analysis To assess whether product- and order-level characteristics are associated with return behavior, chi-square tests were conducted for categorical variables as well as for numerically binned variables. As shown in Table D.8, all examined variables exhibit highly significant relationships with return occurrence (p<0.001). For product color, a chi-square statistic of 8034.43 with a p-value < 0.0001 was yielded. This indicates a significant difference in return rates across color groups. For the product category, the statistic was 42,017.67, with a p-value < 0.0001, confirming differences in return rates across categories. Product size also shows significance with a chi-square statistic of 33426.37 and p-value < 0.0001.

The numerical variables were binned into groups and analysed using the chi-square test as well. The test yielded a chi-square statistic of 137.55 with a p-value < 0.0001, confirming a significant association between the number of items ordered and the likelihood of returns. Lastly, to determine whether sales price influences return behavior, a chi-square test was performed on binned price ranges. The test yielded a chi-square statistic of 1703.32 (p < 0.0001), indicating a significant relationship between price and return rates.

Conclusion The analysis demonstrates that a combination of product characteristics influences product returns. Overall, 6.84% of all sold products are returned, but this rate varies significantly across attributes such as color, category, and size, as confirmed by chi-square tests. Additionally, binned numerical variables such as order quantity, sales price, product quantity, and order hour also exhibit significant relationships with return rates, though their effect sizes are smaller. Correlation analysis supports these findings, revealing weak but statistically significant associations between return status and variables like price and order quantity. Interestingly, the results indicate that orders containing multiple units of the same item exhibit lower return rates. This is an outcome that contrasts with expectations from the literature, which would expect elevated returns due to bracketing behavior [28]. Taken together, these results highlight that return behavior is multifaceted, driven primarily by product-specific factors but also influenced by purchase context.

4.3. Conclusion

The comprehensive analysis at both order and product levels demonstrates that return behavior is influenced by a combination of product-specific characteristics and transactional factors, with significant variation across attributes. At the order level, approximately 5.02% of all orders are returned. Shipping method, delivery mode, and shipment method are highly significantly associated with return likelihood ($p < 0.001$), indicating that operational factors play a role in customer satisfaction. While return rates differ by hour of the day, correlation analysis suggests that order time is not a strong feature when treated as a continuous variable. However, chi-square tests confirm that time-of-day segments (night, morning, afternoon, evening) exhibit significant differences in return rates.

At the product level, the overall return rate is 6.84%, but this varies substantially across product categories, size, and color, as confirmed by chi-square tests. Product category, size, and color all show significant differences across groups, highlighting that intrinsic product features drive returns. Certain categories and larger sizes exhibit disproportionately high return rates, suggesting issues related to fit or product expectations. Additionally, colors such as Red and Cream show higher-than-average return rates, while others, such as Copper and Lavender, show lower rates. Beyond categorical attributes, binned numerical variables such as product price, product quantity, and order hour also show significant relationships with return likelihood ($p < 0.001$). Higher-priced products tend to have slightly higher return rates, as confirmed by chi-square tests and correlation analysis ($r = 0.0103$), possibly reflecting increased customer expectations for premium items. Similarly, product price shows a positive correlation with returns ($r = 0.0129$), while order quantity shows a weak negative correlation ($r = -0.0088$), indicating that larger orders are marginally less likely to be returned. Order hour also shows a small but significant effect ($r = -0.0018$, $p < 0.05$), suggesting behavioral differences across purchase timing.

Although these correlations are statistically significant, their effect sizes are small, meaning that no single variable strongly associates with return behavior on its own. Instead, return patterns are multifactorial, shaped by a combination of product attributes, pricing, and order context. These insights provide a strong foundation for clustering and practical interventions. Ultimately, understanding these drivers enables more accurate clustering and supports strategies that enhance customer satisfaction while minimizing operational costs. The variables influencing return behavior are presented in Figure 4.1.

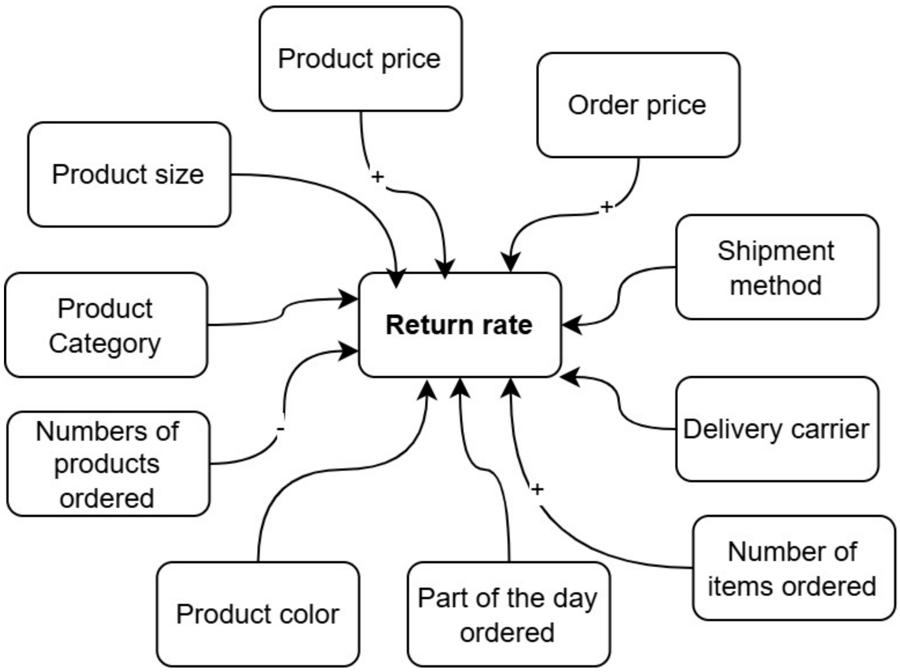


Figure 4.1: Variables influencing returns

5

Current situation handling returns

This chapter describes the current procedures for handling product returns, answering sub-question 2. The process mapping is based on insights derived from the literature, supplemented with findings from the data analysis and expert interviews. Furthermore, key challenges and limitations inherent in the current process are discussed. Summaries of the interviews are provided in Appendix B.

5.1. Current handling of returns based on expert interviews

Return process The return process in e-commerce is typically initiated by the customer, who can register a return via the website, app, customer service, or in-store. Most retailers provide a return label, and customers are often required to use the original packaging, especially for fragile or high-value items. Returns can be dropped off at a designated point or picked up at home. Once received, returns are processed either in-house or by third-party logistics providers, depending on the retailer's setup. Upon arrival at the warehouse or return center, items are inspected to determine their condition and compliance with return policies (e.g., within the allowed return period, unopened, undamaged). Based on this assessment, products are either restocked, refurbished, recycled, or discarded. For low-margin or low-value items, it is often not economically viable to process the return, and these may be scrapped, or customers may receive a refund without returning the product ("credit only" returns). The decision to restock, repair, or dispose of a product is primarily based on its condition and value. For example, opened personal items are almost always discarded for hygiene reasons, while high-value items may be refurbished or resold. Some retailers have developed second-hand or outlet channels for returned goods, but this is not yet common practice. Retailers employ several strategies to minimize returns. Charging a small fee for returns has proven effective in reducing return volumes, as it encourages customers to reconsider before sending items back. Free in-store returns are often promoted to drive foot traffic and encourage additional purchases. However, these strategies must balance customer satisfaction, operational efficiency, and sustainability concerns. While data-driven approaches are widely used in outbound logistics (e.g., demand forecasting, inventory management), their application in return logistics is still limited. Most retailers do not use advanced models to predict return likelihood or optimize return flows. Instead, return volumes are often estimated using simple historical averages (e.g., "we expect X% of sales to be returned after Y weeks"), rather than dynamic, data-driven forecasts. Return reasons are sometimes recorded, but systematic analysis and use of this data for process improvement is rare. In many cases, large amounts of return data are collected but not actively used to optimize operations or reduce future returns. There is broad agreement among experts that the available data remain underutilized, particularly with respect to linking return reasons to product quality issues or customer behaviour.

Figure 5.1 provides a schematic overview of the current return process in e-commerce, based on the literature and as described by experts. The diagram shows how returns are initiated, processed, and either restocked, refurbished, or discarded, depending on product condition and value. It also visualizes the main stakeholders and indicates the extent to which automation is applied within the process.

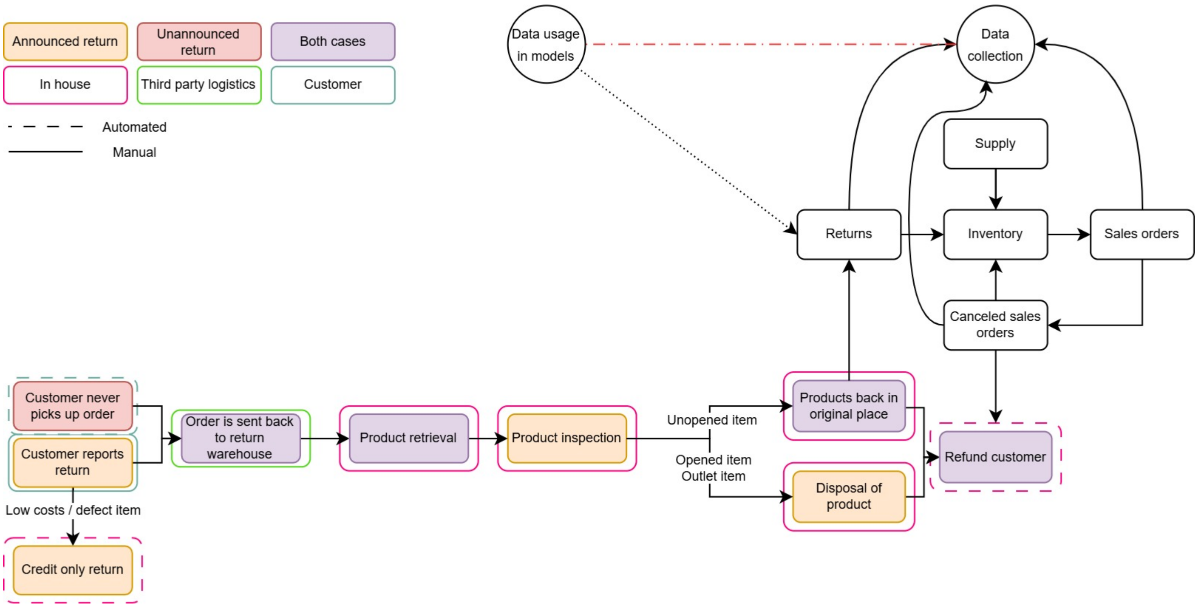


Figure 5.1: Overview of the return process, stakeholders, decision points

The return logistics process begins with inventory management, which is interconnected with sales orders and returns. Returns occur in two main scenarios. First, when a customer never picks up an order, the process is largely automated: the order is sent back to the return warehouse, products are retrieved without inspection since the package is unopened, and items are returned to inventory. The customer receives a full refund automatically. Second, when a customer reports a return, there are two sub-cases: for low-cost or defective items, a credit-only return is processed without the product being sent back, while for standard returns, the order is shipped to the return warehouse via third-party logistics. Here, products are inspected, unopened items return to inventory, while opened or outlet items are disposed of, and the customer is refunded accordingly. Throughout the process, significant data is collected on returns, canceled sales orders, and inventory movements, yet models or analyses based on this data are rarely implemented, representing a missed opportunity to optimize return flow. The figure highlights manual versus automated steps using different border styles, underscoring that limited automation remains a constraint. Additionally, the process reflects a trade-off between economic efficiency and customer satisfaction, while sustainability considerations, such as refurbishment or recycling, are seldom integrated, leading to frequent product disposal.

Operational challenges Experts identified several operational challenges and limitations in the return process. The key challenge addressed is the high volumes of returns. Because of the physical handling of returned items, high volumes make it highly labor-intensive and costly, particularly in large warehouses with fixed storage locations. This manual process can be inefficient, and optimizing the placement of frequently returned items could help improve operational efficiency. Decision-making regarding whether to restock, refurbish, or discard returned products is often based on simple rules, such as restocking unopened items and discarding opened ones, rather than on data-driven assessments. Another significant limitation is the lack of integration between return processes and existing warehouse or ERP systems. Most systems are primarily designed for outbound logistics and are not optimized for handling returns, resulting in additional manual work and inefficiencies. Furthermore, there is little use of automation or artificial intelligence in the return process. Although AI can assist with fraud detection, product assessment, and process optimization, these technologies are not yet widely implemented in practice. Returns also introduce unpredictable inbound flows, which can disrupt inventory management and warehouse planning. This unpredictability makes it challenging to maintain efficient operations. Finally, while some retailers do consider sustainability in their return policies, economic considerations usually take precedence. Returns are rarely truly sustainable due to the additional handling and transport required, which increases costs and the environmental impact.

5.2. Results on data analysis

The analysis of the return process provides several important insights that complement the operational flow described earlier. The process begins with the customer's decision to keep or return a product, after which returned items are sent back to the warehouse through third-party logistics providers, as illustrated in Figure D.2. Once initiated, each return moves through multiple stages (open, delivered to the warehouse, inspected, invoiced, refunded, and finally closed). These stages emphasize the need for accurate tracking and timely execution to maintain customer satisfaction and operational efficiency. Most returns are processed promptly after delivery and inspection, with refunds issued without significant delay. Figures D.9 and D.9 reveal further patterns. Returns are concentrated in specific geographic regions, with a few cities and countries dominating the top 15 return addresses. Similarly, return locations exhibit a skewed distribution, with a small number of facilities accounting for the majority of returned items. Regarding disposition, most products are approved for reimbursement, while a smaller share is rejected or processed differently (Figure f, D.9). Only a minor fraction of returns is scrapped due to damaged items, packaging issues, or defects, as shown in Figure h of D.9. Suggesting improvement for this disposition choice. Delivery modes vary, but a limited number of carriers account for most shipments, suggesting reliance on preferred logistics partners. Status-related variables confirm that nearly all returns are marked as “closed” and “arrived,” suggesting efficient completion once the process starts. Return reasons are diverse, but “RegretOrder” is the most frequent cause. Finally, Figure g of D.9 shows that while some packages go missing within the warehouse, the majority are successfully routed to a designated return station.

5.3. Conclusion

This section outlines the current return process in e-commerce, with Figure 3.1 serving as a basis from the literature. The literature shows the “what” of the current process, but not the “who” and “how”. Expert interviews revealed that most retailers handle returns using a combination of manual processes and basic rules, with limited use of data-driven strategies or automation. There is significant potential for improvement through better data use, automation, and the integration of models into operational systems. The customer is involved in the first part of the return process, after which the third-party carrier takes on the transport to the warehouse. The data show that specific carriers are preferred. When the package arrives at the warehouse, the retailer takes responsibility for retrieval, inspection, and disposition. Limitations of this approach include high return volumes, manual handling, limited system integration, and underutilization of return data for process improvement.

6

Structure of an optimizing framework

This chapter addresses sub-question 3 by outlining the purpose and structure of the proposed framework and discussing how its implementation can support the retailer in optimizing its return management system.

6.1. Design framework

This section outlines the design of the proposed optimization framework. The design requirements are derived from the conclusions presented in Chapter 5. Building on these requirements, a conceptual framework is developed that addresses the identified key challenge.

6.1.1. Design requirements

The framework must be designed to satisfy several key requirements to ensure its practical relevance and usability for stakeholders in the e-commerce domain. First, interpretability is essential: to address the gap identified in the literature, the framework must produce outputs that are readily understandable by diverse stakeholder groups, enabling them to identify the factors driving returns and to justify why particular products or orders are grouped together. Such transparency is critical for developing targeted, evidence-based strategies to reduce return rates. Second, the method should make effective use of available data by incorporating both numerical and categorical features, particularly given that substantial return-related information is collected but remains underutilized, as highlighted in the interviews. Finally, scalability is required, as the framework must be capable of handling the large-scale datasets typical of e-commerce operations. The following requirements therefore guide the design of the method:

- *The outputs should be easy to interpret by stakeholders: Results must be explainable, so users can understand which factors drive returns*
- *The model must make effective use of the available data, especially since currently much return data is collected but underutilized.*
- *It should be able handle both numerical and categorical features.*
- *The model should be scalable to handle large datasets typical in e-commerce.*

6.1.2. Conceptual framework

Figure 6.1 presents the key objective as depicted in the problem tree. Building on this visualisation and the overarching goal, a conceptual framework has been developed. This framework provides the basis for addressing the corresponding research questions.

The key issue identified through interviews and exploratory data analysis is the high volume of returns. The objective, as shown in the problem tree, is to reduce these returns. This can be achieved by targeting both high-risk items and high-risk orders. To determine what constitutes high-risk items and orders, a cluster analysis can be conducted. By clustering orders and products, high- and low-return-

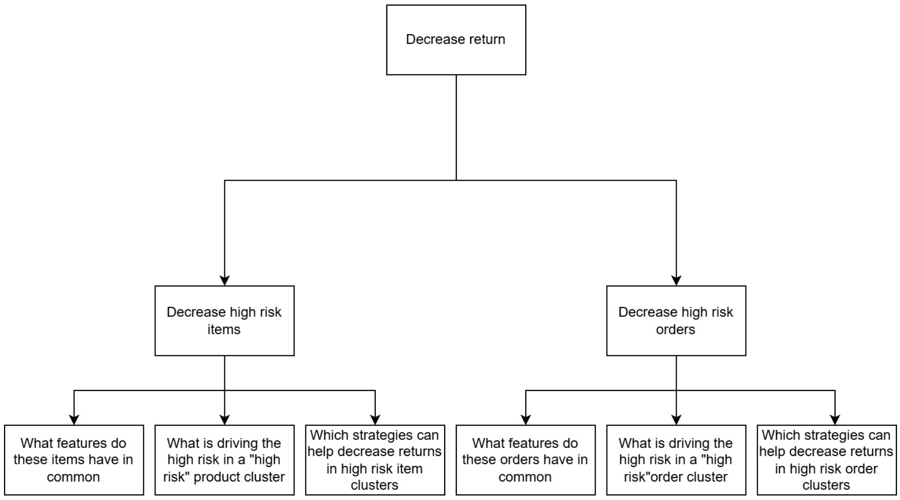


Figure 6.1: Problem tree and questions to ask to solve this problem

risk segments were identified. Within these clusters, it is possible to uncover the features and drivers associated with elevated risk, enabling deeper analysis. Patterns and combinations can then be examined to inform strategies aimed at reducing high-risk orders and products. These strategies will be validated through stakeholder interviews. The conceptual framework in the Figure below illustrates this approach, using the features identified in Chapter 4 as input data for the clustering analysis.

The conceptual framework is illustrated in Figure 6.2. The process begins with collecting order data, which includes product attributes such as color, size, and category. This data serves as the input for clustering analysis. An appropriate algorithm is selected to fit the characteristics of the collected data. After performing the clustering analysis, the resulting clusters are visualized and examined to identify patterns. The primary focus is on distinguishing high-risk clusters from low-risk clusters and analyzing their respective characteristics. Based on these insights, strategies are developed to address the identified risks. These strategies are then validated with stakeholders to ensure their effectiveness and relevance. The process is iterative; after a period of implementation, the framework is revisited to assess and refine the strategies based on new data and outcomes.

Implementing the return flow optimization framework provides retailers with a structured, data-driven approach to identifying strategies to reduce return volumes. The framework uses clustering analysis to identify high- and low-risk products and orders, enabling the development of targeted strategies for return reduction. By reducing returns, inventory management becomes more streamlined, with fewer disruptions and less complexity in stock handling. This leads to a faster and more efficient returns process, reducing operational costs and freeing up resources that can be reinvested in process automation or other value-adding activities. Furthermore, customers are less likely to encounter issues that prompt returns, resulting in higher customer satisfaction and fewer errors in the returns process. The iterative nature of the framework ensures that strategies are regularly reviewed and refined based on updated data and stakeholder feedback. The framework will help the retailer to go from the current situation, as described in Chapter 5, to a new, improved process. The process is depicted in greater detail, with an expanded set of stakeholders, in the Figure included in Appendix E. As illustrated in Figure E.1, the current in-house responsibilities are dispersed across multiple teams within the retailer. At present, warehouse operators primarily experience the consequences of the issue but lack the authority to undertake strategic actions to address it. To mitigate this gap, the proposed return process introduces a new stakeholder responsible for collecting feedback and reassessing previously made strategic decisions. This newly added "Mediator" assumes the role of problem owner. The optimized return process also demonstrates more extensive use of collected data, enabling additional pathways for returned products based on their risk classification and incorporating greater levels of automation. Overall, this framework shifts the process from a reactive to a proactive approach, restructures governance and strategic decision-making, and ultimately reduces returns and inventory disruptions, as shown in Figure E.3.

6.2. Conclusion

This chapter presented the design of the proposed optimization framework, outlining its underlying requirements, conceptual structure, and intended practical application. Grounded in the challenges identified in the current return-handling process, the framework emphasizes interpretability, effective use of available data, compatibility with mixed data types, and scalability to accommodate large e-commerce datasets. By integrating clustering analysis with stakeholder validation, the framework offers a systematic, iterative approach to identifying high- and low-risk products and orders, thereby enabling the development of targeted, evidence-based strategies to reduce return volumes. Ultimately, the framework provides retailers with a structured approach to improving operational efficiency, enhancing customer satisfaction, and informing continuous process refinement through emerging data and feedback.

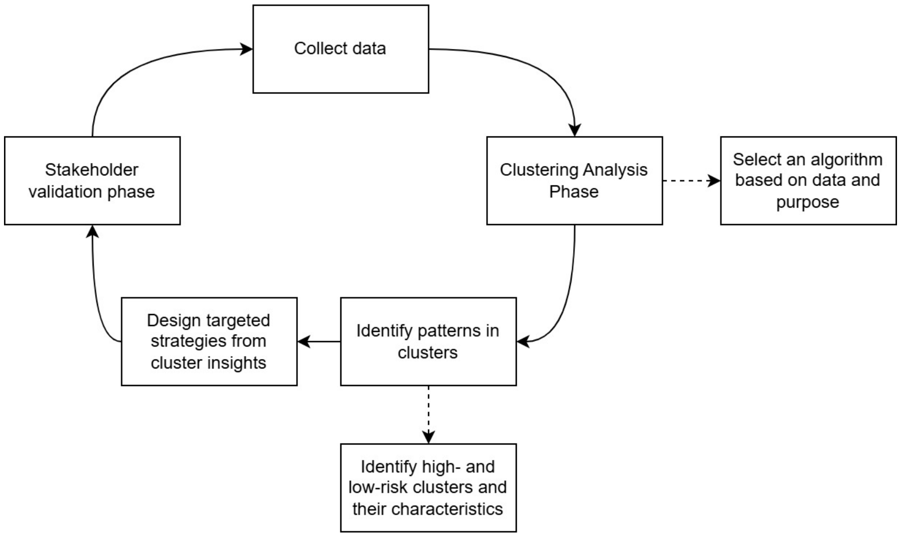


Figure 6.2: Framework finding high-risk orders and products

7

High & low risk clusters and strategies

This chapter addresses sub-question 4 by building on insights from the preceding chapters and applying the framework introduced in Chapter 6. Using the variables identified and statistically tested in Chapter 4, this chapter conducts the clustering analysis, identifies patterns, and develops corresponding strategies. Cluster characteristics inform these strategies, the literature reviewed in Section 3.3, and stakeholder input. The chapter proceeds by first analysing clusters derived from order-level features, then analysing product-level clusters, and finally examining a high-risk product cluster in depth. Order-level features are applied at the product level as well to enable pattern comparison across both analytical layers. Finally, this chapter proposes strategies that are subsequently evaluated and validated by supply chain experts.

7.1. Clustering on Order level

This section provides a more detailed examination of the order-level features. The cluster analysis aims to identify underlying patterns within the order data, including variables such as delivery mode, shipping carrier, and shipment method. In addition to shipment-related attributes, the analysis incorporates order value and the number of items per order to capture a more comprehensive understanding of order characteristics.

7.1.1. K-prototyping

Figure 7.1 illustrates the results of applying the elbow method. The elbow method is used to decide on the optimal number of clusters in the algorithm.

Subsequently, the K-prototypes algorithm is applied, assigning all orders to specific clusters. This chapter provides an overview of cluster characteristics, including summary statistics, cluster sizes, and return rates. Each cluster will be examined, with an appropriate risk label assigned. Furthermore, the clustering methodology and the quality of the resulting clusters will be critically evaluated.

The PCA visualization in Figure 7.2 shows the clustering results of the K-Prototypes algorithm projected onto two principal components. While the clusters are not perfectly separated in this two-dimensional space, there are some visible divisions. The points form distinct diagonal bands along PCA Component 1, suggesting an underlying structure in the data. Cluster 0 (red) appears more distributed and includes several outliers extending far along PCA Component 2. This may represent observations with atypical characteristics. Other clusters, such as those in orange and green, are more compact and concentrated near lower values of PCA Component 2, indicating less variability. Despite the overlap, these diagonal patterns imply that the algorithm captured meaningful differences between clusters, even though full separation may require a higher-dimensional space.

The K-Prototyping algorithm yields a wide range of cluster sizes, from 0.1% to 48.0%. The associated return rates span from 3.82% to 10.26%. An overview of the clusters is given in Table G.1.

Firstly, an important observation is that the value and size of the orders grow with the risk of a cluster.

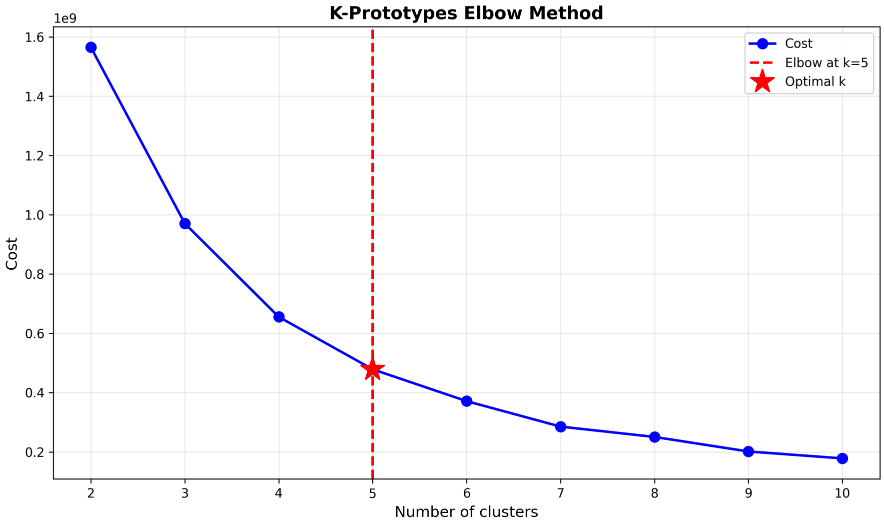


Figure 7.1: Elbow method for deciding the number of clusters orders

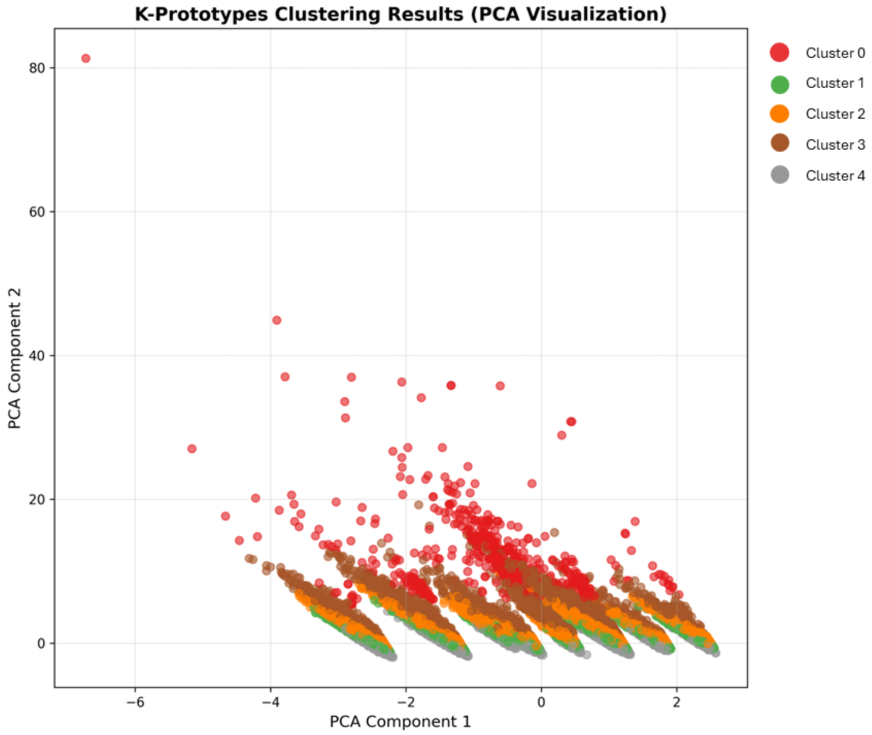


Figure 7.2: PCA visualization of clusters - K-prototyping - Orders

Cluster 4 has a low risk label (RR: 3.82%), whereas Cluster 0 has a high risk label (RR: 10.26%). Clusters 2 and 3 are also high risk, with the second- and third-largest order sizes and values. Within the high-risk clusters, DHLDE increases risk, whereas PostNL-Stan decreases it. Across the low- and medium-risk clusters, PostNL-Standard reduces risk more than PostNL-PU.

The distribution of purchase time is the same across clusters, indicating that this feature was not used during clustering. However, the return rates during the evening and night are above average across all clusters.

In conclusion, this clustering method results in a large imbalance in cluster sizes (the smallest cluster is 0.1%), which may limit the visibility of patterns. The clusters seem to be separated by price and initial number of items, as seen in Figures G.8a and G.8b, with lower amounts and smaller orders being associated with lower risk. The clusters show the same distribution in time of day, which is also seen in Figure G.2.

7.1.2. CAVE algorithm

The optimal number of clusters was determined using the elbow method, consistent with the approach applied in the previous method. Based on the clustering cost curve, the point of inflection indicates that the optimal value of K for this algorithm is 5.

The CAVE algorithm yields clusters ranging from 1.6% to 29.7% of the orders. The return rates range from 3.58% to 8.37%, with a base return rate of 5.02%. The overview of the clusters is shown in Table G.2. Figure 7.3 shows a PCA visualization. From the figure, it is evident that the clusters are arranged in long diagonal bands along PCA Component 1, indicating correlations among features and a structured distribution. While there is noticeable overlap between clusters, some separation is evident, particularly in the concentration of points within distinct color regions. The overlap within clusters suggests that they are not linearly separable in two dimensions.

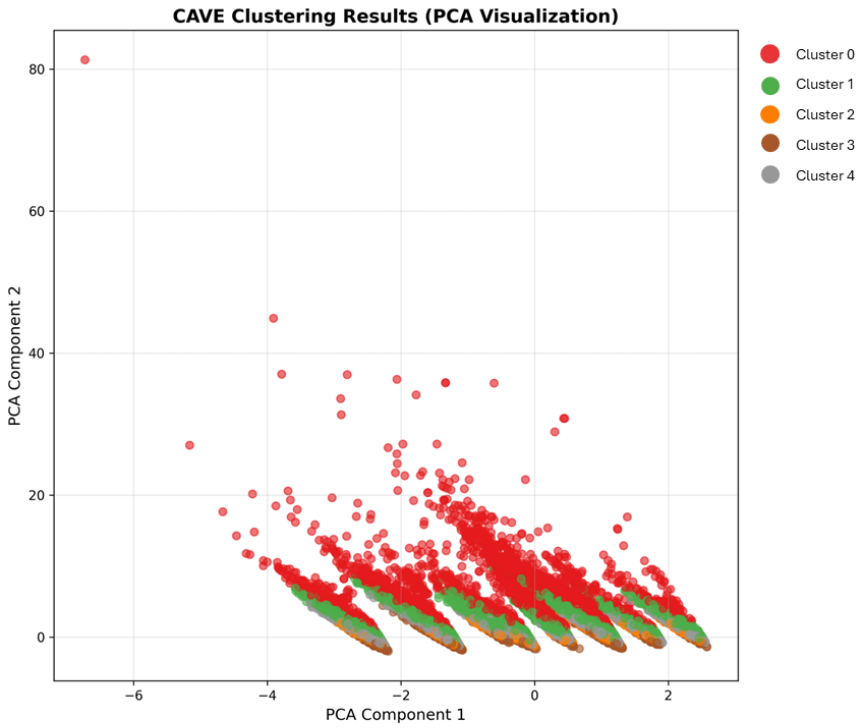


Figure 7.3: PCA Visualization of clusters - CAVE - Orders

This algorithm shows the same patterns as K-Prototyping. Firstly, an important observation is that the value and size of the orders grow with the risk of a cluster. Cluster 3 has a low risk label (RR: 3.58%), whereas Cluster 0 has a high risk label (RR: 8.37%). With prices being 26.88 (± 7.27) and 428.03 (± 228.19) respectively. Cluster 1 also shows a high risk and has an average order value of 166.88 (±39.79). Within the high-risk clusters, DHLDE increases risk, whereas PostNL-Stan decreases it. Across the low- and medium-risk clusters, PostNL-Standard reduces risk more than PostNL-PU.

In conclusion, the CAVE algorithm partitions the dataset into five clusters with return rates ranging from 3.61% to 8.50%. The scatterplots in Figures G.15c and G.15d reveal clear cluster separation based on numerical features, suggesting that order size and order value are the primary drivers of the clustering structure. As illustrated in Figure G.10, the time-of-day distribution is similar across all clusters, indicating that purchase timing does not drive clustering algorithms. This method shows many similarities to the K-prototyping method, suggesting little improvement over it, except for yielding more evenly distributed observations across clusters.

7.1.3. Latent Class Clustering

To select the optimal number of clusters for latent class clustering, models with varying cluster counts were compared using Bayesian Information Criterion (BIC), log-likelihood, and convergence stability. The five-cluster solution was chosen as it achieved the lowest BIC while maintaining reasonable fit and convergence (see Table 7.1).

K	BIC	Log-Likelihood	Iterations	Return rate range
2	$8.53x10^6$	$-4.27x10^6$	11	4.88% - 5.23%
3	$8.15x10^6$	$-4.07x10^6$	15	4.51% - 6.77%
4	$7.66x10^6$	$-3.83x10^6$	47	2.44% - 9.80%
5	$7.54x10^6$	$-3.77x10^6$	**30**	**1.93% - 7.93%**

Table 7.1: Choosing optimal K

In Figure 7.4, the distribution of the clusters can be seen. Clusters are partially overlapping, but a clear gradient along PCA Component 1 is observed. The clusters seem to form elongated bands rather than tight, spherical groups, suggesting that the underlying data has strong directional variance. Most points are concentrated between PCA Component 1: -6 to +2 and PCA Component 2: 0 to 20, forming a diagonal pattern. There is a long tail of points extending upward along PCA Component 2, which is unusual and likely represents outliers or a small subgroup with very different characteristics. The orange cluster occupies the far right and appears largest and densest, suggesting it represents the majority of observations. The red cluster is on the far left and is relatively compact. Green and brown clusters are in the middle, overlapping a lot, which may indicate similar profiles and thus less distinct separation. A few points (brown/orange) are far above the main diagonal, which could indicate extreme values in original features.

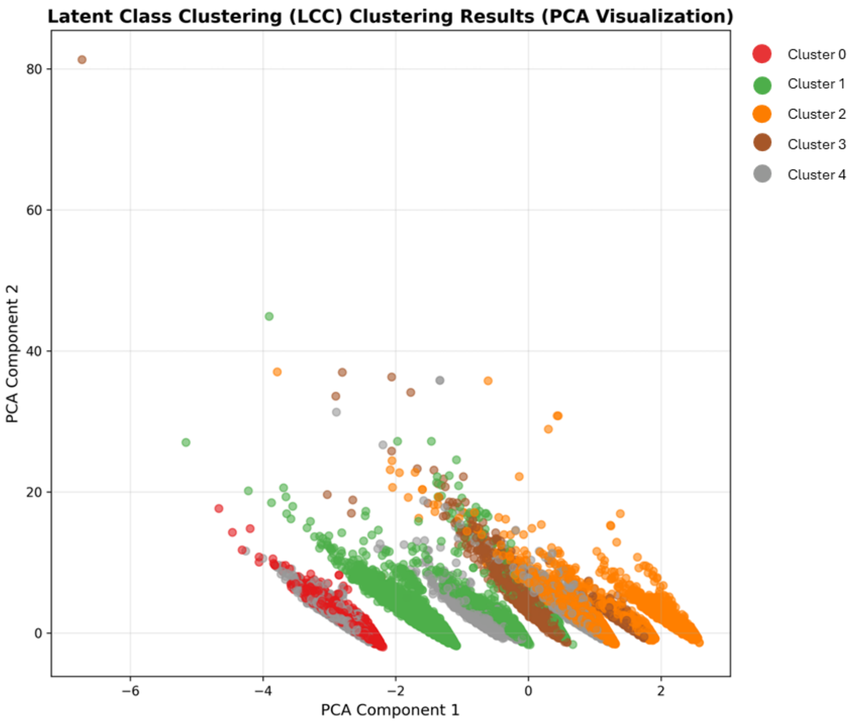


Figure 7.4: PCA visualization of clusters - LCC - Orders

Latent class clustering yields clusters ranging from 3.2% to 40.7% of the orders. The return rates range from 1.93% to 7.93%, with a base return rate of 5.02%. The overview of the clusters is shown in Table G.3.

Cluster 0 accounts for 9.8% of all orders and has a return rate of 1.93%, making it the lowest-risk segment. It is entirely dominated by COLISSIMO Standard (100%, RR 1.93%), which explains the very low return rate; this is visualized in Figure G.16. Orders are mid-sized (3.18 ± 1.65 items) with moderate value (€64.37 ± 45.87). Cluster 1 is a large segment (24.4%) and shows a high return rate of 7.00%. It is dominated by DHLDE Standard (84.7%, RR 7.59%) and GLS Standard (13.2%, RR 4.13%), with DHLDE driving risk upward. Orders are slightly larger (3.83 ± 1.97 items) and higher in value (€81.37 ± 74.54) than Cluster 0. Cluster 2 is the largest segment (40.7%) with a return rate of 3.19%, well below average and thus a low-risk cluster. It is dominated by PostNL Standard (94.5%, RR 3.20%). Orders are mid-sized (3.13 ± 1.65 items) with moderate value (€65.99 ± 52.54). Cluster 3 is the smallest segment (3.2%) with a return rate of 2.81%, indicating a low-risk cluster. It is characterized by INPOST Pickup (87.8%, RR 2.57%) and SEUR Pickup (12.2%, RR 4.53%). Orders are mid-sized (3.26 ± 1.55 items) but notably high in value (€210.84 ± 171.89), indicating premium purchases. Cluster 4 accounts for 21.8% of orders and has the highest return rate (7.93%), marking it as a high-risk cluster. It is dominated by PostNL Pickup (72.9%, RR 5.33%). Orders are mid-sized (3.23 ± 1.67 items) with moderate value (€67.36 ± 52.95). The time-of-day distribution is the same across clusters, with evening and night showing an above-average return rate. The overview supporting this is shown in table G.3.

In conclusion, this method, unlike the others, clusters more closely around the shipping carrier and method. This is supported by Figures G.20 and G.21. Across clusters, the distribution by time of day remains consistent, with afternoons and evenings being most common, as observed in previous analyses and seen in Figure G.17. Among carriers, INPOST and COLISSIMO emerge as the most reliable, associated with lower return rates. The Figures G.25 and G.24 illustrate that even though less prevalent than in the other two methods, there is still a distinction in order value and size across clusters.

7.1.4. Conclusion

When comparing the three clustering methods, the K-prototype algorithm produces the widest range in cluster sizes. LCC yields more balanced partitions. Extremely large clusters can obscure patterns and reduce interpretability. Both CAVE and K-prototype emphasize order value and size as key drivers, whereas LCC places greater weight on carrier and delivery method differences. Across all methods, the distribution by time of day remains consistent, with no significant impact on a cluster's risk label. However, night observations increase the overall risk. The K-protoyping and CAVE algorithms both show that risk increases proportionally as both the order value and the order size grow. In LCC, there is also a distinction between order sizes and values, but it is less distinct. DHLDE shows increased risk in all methods. PostNL-Standard seems to reduce risk compared with PostNL-Pickup, but PostNL in general seems a low-risk carrier. Additional cluster analysis at the product level is required, as order-level patterns alone are insufficient to develop effective risk-mitigation strategies; understanding how risks vary across specific products or categories is essential for actionable decision-making. The clusters of the different methods are summarized in Table 7.2.

Method	Cluster	Ass. data-points	Percentage of total	Return rate	Cluster description	Risk label
K-Proto	0	848	0.1%	10.3%	PostNL, Standard, High prices and items	High
	1	225266	37.4%	5.6%	PostNL-Stan, avg number of items and price	Medium +
	2	74585	12.4%	7.3%	PostNL & DHLDE, Standard, avg number of items and price	High
	3	12243	2%	8.2%	INPOST & PostNL, PU & Standard, High amount & ordered items	High
	4	288695	48%	3.8%	PostNL, Standard, low amount and items	Low
CAVE	0	9439	1.6%	8.4%	INPOS-PU, PostNL-Stand, Large orders, High Value	High
	1	72199	12%	7.4%	PostNL-Stan, DHLDE-Stand Large & High value	High
	2	178868	29.7%	4.4%	PostNL-Stan/PU, Average and middle value	Medium -
	3	169882	28.2%	3.6%	DHLDE-Stan, Small, Low value orders	Low
	4	171249	28.5%	5.9%	PostNL-Stan, Mid & Mid value orders	Medium +
LCC	0	58839	9.8%	1.9%	COLLISSIMO-Stan, average order size and value	Low
	1	146745	24.4%	7%	DHLDE-Stand, Moderately big order size and value	High
	2	245112	40.7%	3.2%	PostNL-Stan, average order size and value	Low
	3	19494	3.2%	2.8%	INPOS-PU, average order size, high value	Low
	4	131447	21.8%	7.9%	PostNL-PU, average order size and value	High

Note: Overall return rate: 5.02%

Table 7.2: Summary of clusters for different methods - orders

7.2. Clustering on Product Level

This section provides a more detailed examination of the product-level features. The cluster analysis aims to identify underlying patterns in the product data, including variables such as product category, color, size, price, and the quantity ordered.

7.2.1. K-prototyping

Figure 7.5 illustrates the results of applying the elbow method. The elbow method is used to decide on the optimal number of clusters in the algorithm. The optimal in this case is four clusters.

Using K-prototyping, four clusters are identified, ranging from 2.3% to 63.8%. 63.8% is a disproportionately high share that may obscure meaningful patterns within the data. The return rates range from 6.24 to 8.38%, with a base return rate of 6.84%. A comprehensive overview of the cluster characteristics is presented in Table G.4.

Figure 7.6 shows the PCA of this clustering. In this case, the clusters show a substantial overlap,

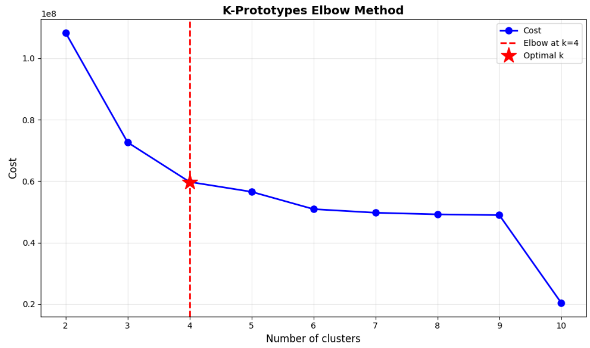


Figure 7.5: Elbow method for deciding the number of clusters

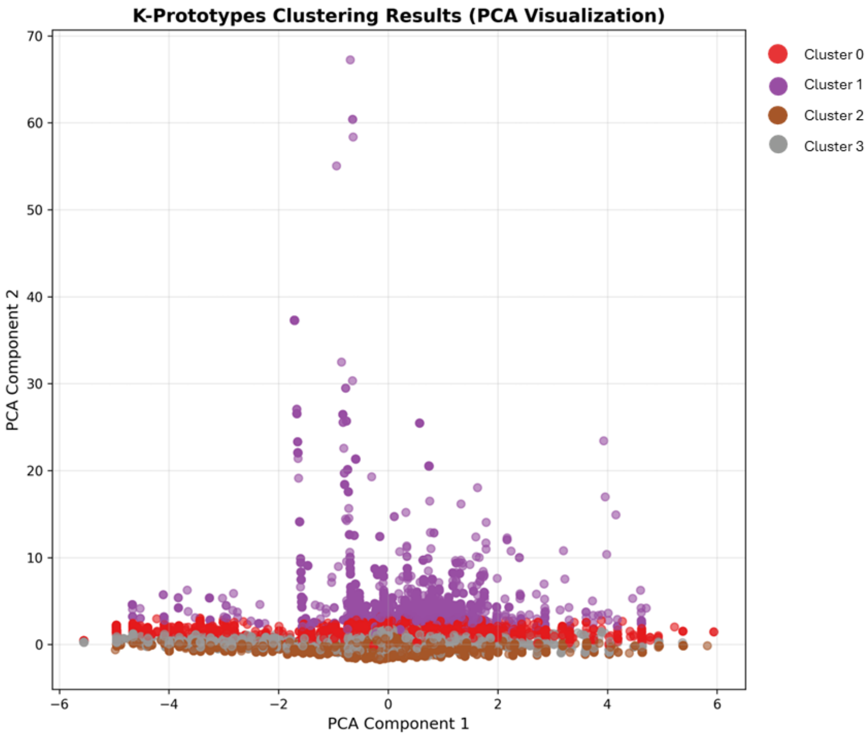


Figure 7.6: PCA visualization of clusters - K-prototyping - Products

especially around the PCA-1 origin, indicating that the variance captured by PCA-1/2 is not aligned with the K-prototypes decision boundaries. One notable observation is that Cluster 1 (purple) shows a wider spread along PCA Component 2, reflecting its higher internal variability and heterogeneous composition. Overall, the plot highlights the imbalance in the clusters and partial overlap.

Cluster 0 (Medium- Risk, 8.1%) exhibits a slightly below-average return rate of 6.54% and is characterized by dominance of Categories L and I, with Category I showing above-average returns (6.93%). Black is the most frequent color (24.5%, return rate 8.22%). Products fall within a mid-range price level (€24.68 ± 5.54). Cluster 1 (Medium+ Risk, 2.3%) has an above-average return rate of 7.22% and is heavily composed of Category L (71.9%), followed by I (12.7%). Color distribution is varied, with black (23.6%) and pink (16.3%) returning at similar levels to the cluster average. Cluster 1 contains the highest-priced products (€59.74 ± 34.14). Cluster 2 (Medium- Risk, 63.8%) demonstrates a below-average return rate of 6.24% and is dominated by Category H and L. Unknown colors are prevalent and associated with lower return rates (5.57%). This cluster represents the lowest-priced products (€2.37 ± 1.61). Cluster 3 (High-Risk, 25.8%) has the highest return rate (8.38%) and is marked by Categories L and F, with Category F showing an exceptionally high return rate (21.61%). Black is the dominant color with a return rate of 12.01%. Products show moderate pricing (€9.47 ± 2.78).

In conclusion, the K-prototypes clustering reveals substantial imbalance among clusters, with Cluster 2 dominating (63.8%) and Cluster 1 being very small (2.3%), which may limit pattern visibility. The clusters predominantly exhibit medium risk, which complicates the development of targeted strategies because they do not show clear or contrasting distinctions. Price segmentation is clear: Cluster 2 contains low-priced products, Cluster 1 high-priced items, and Clusters 3 and 0 fall in between. Supported by Figures G.32a and G.32b in appendix G. This suggests price is a key driver in the clustering output. The high-risk cluster (3) shows average prices and products in Cat. L and F. The prevalent color and size are black and unknown; however, a conclusion cannot be drawn from this, as this appears in multiple clusters. Time-of-day distributions appear consistent across clusters, suggesting that this variable did not substantially influence the clustering process. Nonetheless, orders placed during the night consistently exhibit elevated return rates in all clusters representing a higher-risk segment.

7.2.2. CAVE algorithm

The optimal number of clusters was determined using the elbow method, consistent with the approach applied in previous experiments. Based on the clustering cost curve, the point of inflection indicates that the optimal value of K for this algorithm is 4.

Using the CAVE algorithm, four clusters were identified, spanning proportions from 3.0% to 47.8%. The corresponding return rates range from 6.04% to 7.96%, relative to a baseline return rate of 6.84%, indicating only a modest degree of variation. An overview of these results is provided in Table G.5.

Figure 7.7 shows a PCA visualization, which reveals that most data points are concentrated near the origin along PCA Component 2, forming a dense horizontal band, while PCA Component 1 spans a wider range from approximately -6 to +6. This indicates that the first component explains more of the variance across clusters than the second. Cluster overlap is substantial, suggesting that clusters are not linearly separable in two dimensions and that the clustering structure relies on higher-dimensional relationships. Notably, Cluster 3 (shown in grey) exhibits significant dispersion and includes extreme outliers with PCA Component 2 values exceeding 60, suggesting products with atypical characteristics or anomalies. Other clusters appear more compact and closer to the origin, indicating lower variability within those groups.

Cluster 1 dominates, comprising 47.8% of all observations, and exhibits a return rate of 6.04%, slightly below the overall average. This cluster contains low-priced items from Cat. H, with mostly unknown colors, suggesting that colors do not affect this product group. Furthermore, Cluster 0 accounts for 34.9% of the data and has a return rate of 7.45%, indicating a slightly higher risk segment. Items in this cluster fall in the low-average priced black items from Cat. L and H. Cluster 2 contains 14.3% of the purchased items and has a return rate of 7.96%, the highest risk cluster in this method. It contains Black colored products from Cat. L and I with average prices. Cluster 3 is the smallest group (3.0%) and has a slightly higher return rate of 7.06%, containing high-priced items from Cat. L with unknown colors or black. Color analysis indicates that black is consistently prevalent across clusters and correlates with elevated return rates, particularly in Cluster 2 (RR: 11.21%). Unknown colors dominate Cluster 1 (54.0%), showing below-average return rates, suggesting limited influence of color in this cluster. Size remains largely unspecified across all clusters (>84%), making it hard to draw conclusions. Temporal purchase patterns are similar across clusters, with the afternoon and evening periods accounting for the majority of transactions, suggesting that time of day was not a primary clustering driver. However, across all clusters, purchases made during the night show a higher risk than the cluster average.

In conclusion, the method yields only moderate dispersion in return rates, limiting its usefulness for strategic differentiation, as no distinctly high- or low-risk clusters emerge. The clustering outcome is predominantly driven by the numerical feature, product price, resulting in pronounced price-based segmentation: Cluster 1 comprises the lowest-priced products (€ 1.60 ± 0.98), whereas Cluster 3 contains the highest-priced items (€54.13 ± 31.40). Clusters 0 and 2 fall within intermediate price ranges (€6.52 and €18.59, respectively). This is supported by Figure G.40a. No meaningful conclusions can be drawn regarding product sizes, as size distributions remain highly diffuse across clusters. Similar patterns were observed in the K-prototyping results.

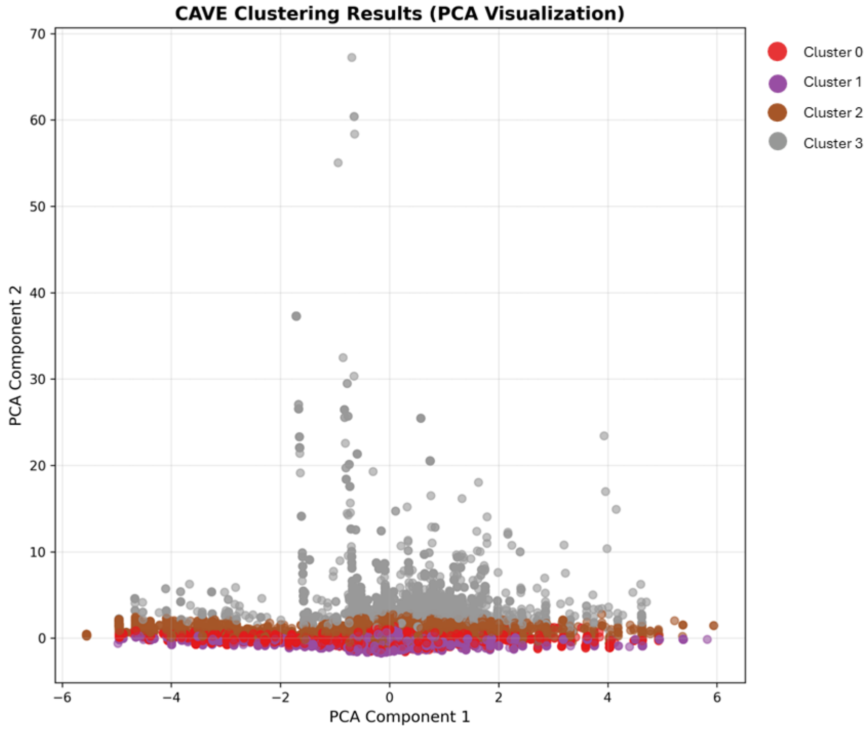


Figure 7.7: PCA visualization of clusters - CAVE - Products

7.2.3. Latent Class Clustering

To determine the optimal number of clusters for LCC, several models with varying cluster counts were estimated, and their performance was evaluated using BIC, log-likelihood, and the number of iterations required for convergence. For clustering the products, the optimal number of clusters for this model is 5 (Table 7.3).

K	BIC	Log-Likelihood	Iterations	Return rate range
2	$2.13x10^7$	$-1.07x10^7$	100	5.41% - 7.73%
3	$2.07x10^7$	$-1.04x10^7$	100	5.35% - 9.06%
4	$2.00x10^7$	$-1.00x10^7$	100	5.40% - 18.90%
5	$1.97x10^7$	$-9.87x10^6$	**100**	**5.36% - 18.63%**

Table 7.3: Choosing optimal K

Using LCC yields five clusters, with the smallest containing 8.3% of all ordered products and the largest containing 36.0%. The return rate ranges between 5.36% and 18.63%. An overview of the different cluster characteristics is presented in Table G.6.

For this method, PCA is visualized in the Figure below (Fig. 7.8). The plot shows that most data points are concentrated near the origin along PCA Component 2, forming a dense horizontal band, while PCA Component 1 spans a range from approximately -6 to +6. This indicates that the first component explains more variance than the second. Cluster separation is limited in this projection, as substantial overlap among clusters suggests that the clustering structure relies on higher-dimensional relationships rather than linear separability in two dimensions. Cluster 2 (Orange) dominates the left side of the plot and appears more compact, while Clusters 0 and 3 (Red and Brown) are scattered across the center and right regions. A few extreme outliers extend far along PCA Component 2, particularly in Cluster 0, which may represent observations with unique or atypical feature combinations.

Cluster 0 represents 12.9% of all products and has a return rate of 6.40%, slightly below the overall average, giving it the risk label of Medium -. It is dominated by Categories I (46.6%) and E (41.0%), with Category E showing a higher return rate (7.16%). Color distribution is led by beige (27.9%) and black

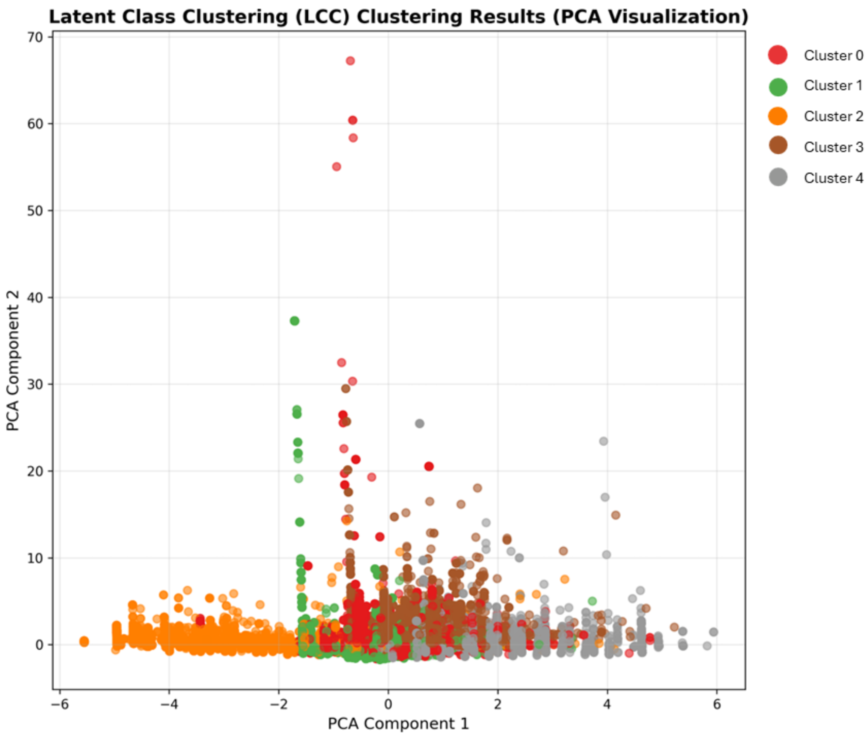


Figure 7.8: PCA visualization of clusters - LCC - Products

(21.1%), with beige having an above-average return rate (7.15%). Size information is not applicable in this cluster. Cluster 1 accounts for 19.3% of products, has a return rate of 6.24%, and thus has a risk label of Medium -. It is primarily associated with Categories B (33.3%), K (27.5%), and C (25.4%). Black is the dominant color (55.8%) with a return rate of 6.52%, while blue and silver follow with lower shares. This cluster contains the lowest-priced products (€4.23 ± 11.40) and an average quantity of 1 unit. Cluster 2 is the smallest (8.3%) and exhibits the highest return rate (18.63%), indicating a high-risk cluster. It is dominated by Category F (97.8%), which drives the high return rate. Black (79.9%) and red (11.8%) are the most frequent colors. Size distribution is more detailed here, with small (27.6%), large (20.7%), and extra-large (11.3%) sizes all showing very high return rates (up to 21.66%). Price levels are mid-range (€7.35 ± 4.96), and the average quantity is 1.02 units. Cluster 3 comprises 23.6% of products and has a return rate of 5.68%, the second-lowest among clusters. Category L, dominates it and Pink and black are the leading colors, both showing moderate return rates. Price levels are relatively high (€13.48 ± 14.75), and the average quantity remains stable at 1.01 units. Cluster 4 is the largest, covering 36.0% of products, and has the lowest return rate at 5.36%, indicating a low-risk segment. It is dominated by Category H (72.2%). Color is almost entirely unknown (98.7%), suggesting limited influence of color on returns in this cluster. Size is also unknown for nearly all products. This cluster contains the lowest-priced products (€3.49 ± 5.88) and the highest average ordered quantity (1.06 units).

In conclusion, Latent Class Clustering provides a more distinct segmentation than other clustering methods, with clearer separation between high- and low-risk clusters. Notably, Cluster 2, the smallest segment (8.3%), exhibits the highest return rate (18.63%), indicating a concentrated, high-risk group. This cluster primarily consists of products from Category F, predominantly black in color, and priced at an intermediate level. Size plays a critical role within this cluster, as specified sizes consistently show above-average return rates. The scatter plots in Appendix G illustrate that LCC achieves a more pronounced differentiation across product categories, colors, and sizes, while also accounting for price variations (see Figures G.49, G.50, G.51, and G.48). Conversely, time-of-day does not appear to influence clustering outcomes, as purchase distributions remain similar across all clusters (Figure G.41). Given that this method produces the most pronounced distinctions in risk levels across clusters, the high-risk cluster identified through this approach is selected for more detailed examination.

Deeper analysis within Clusters and products

When examining the high-risk cluster (Cluster 2 in LCC) in greater detail, additional patterns emerge. This cluster was further segmented using the LCC method, and the summary of these subclusters is presented in Table 7.4, based on the overview in Table G.8. All descriptions and risk labels are relative to the averages within Cluster 2.

As shown in the table, Subclusters 1 and 4 exhibit the highest risk levels within Cluster 2. Both share common characteristics: higher-priced products and sizes ranging from S to XL are more prone to returns compared to One Size items. Across all subclusters, black-colored items appear consistently, though not disproportionately in any single cluster. High-priced items in Cluster 2 range from €8.37 (±5.18) to €12.26 (±6.10). Products with relatively high prices in Category F are at greater risk of being returned if they feature the color Black or Red and come in sizes S, M, L, or XL.

Cluster	Ass. data-points	Percentage of total	Return rate	Cluster description	Risk label
2.0	24054	18.5%	13.84%	Low priced, 1 item, black, One Size	Low
2.1	**21202**	**16.3%**	**21.31%**	**High priced, 1,5 item, Black & white, M & XL**	**High**
2.2	20202	15.5%	15.51%	Low priced, 1 item, Black & white, S & One Size	Low
2.3	20178	15.5%	18.04%	Average priced, Black & Red, S&L, 1 item	Average
2.4	**44412**	**34.2%**	**21.63%**	**High priced, 1 item, Black & Red, S & L**	**High**

Note: Overall return rate: 18.6%

Table 7.4: Summary of clusters of Cluster 2 - products

After clustering product information based on Category, color, size, time of day, price, and quantity, order-level data was integrated into the dataset for further analysis. Exploring order characteristics across clusters reveals that Cluster 2, the high-risk cluster, primarily consists of large orders with high total values. PostNL accounts for the largest share of shipments in this cluster; however, DHLDE, the second-largest carrier, significantly contributes to the elevated return rate through its standard shipping method. Cluster 2 includes approximately 130,000 products linked to 68,000 unique orders, indicating that multiple items from the same order are often returned. While Clusters 1 and 3 also feature high-value orders, their return rates remain above average and closely align with their second most frequent order type: medium + value orders. A broader overview is presented in Table G.9.

Cluster	Unique Orders	Size of Order	Ordervalue	Delivery mode	Total products	Return rate
0	165.156	Medium	Mid +	PostNL-Stan	201866	6.4%
1	193.156	Large	High	PostNL-Stan	302270	6.2%
2	68.923	Large	High	PostNL-Stan	130048	18.6%
3	279.929	Medium	High	PostNL-Stan	369593	5.7%
4	337.474	Medium	Low	PostNL-Stan	564007	5.4%

Note: Overall return rate: 6.84%
Order Sizes Ranges: Small order (1-2); Medium order (3-4); Large order (5+)
Order value ranges: Low: €0.01 - €49.94; Medium -: €49.94 - €72.96; Medium +: €72.96 - €111.35; High: €111.35+

Table 7.5: Order level information in clusters based on product information

7.2.4. Conclusion

In this section, the three clustering methods are compared. K-prototypes and CAVE algorithms produce clusters of similar sizes and exhibit comparable return rate ranges, indicating limited differentiation between these approaches. In contrast, Latent Class Clustering achieves a more balanced segmentation, clearly distinguishing between low-risk and high-risk clusters. The first two methods seem to cluster more on numerical factors, whereas the LCC method divides the Categories across clusters more evenly. A summary of cluster characteristics and key metrics is presented in Table 7.6. The LCC method is further applied for creating subclusters in cluster 2. This shows that sizes other than “One Size” have a higher risk in Cluster 2. Furthermore, Black and Red have an increased risk. As the final part of the analysis, the product order features were analysed. Cluster 2 showed mostly large orders and high order values, suggesting increased risk. Cluster 3 shows high-value orders with fewer items, suggesting this combination poses less risk.

Method	**Cluster**	**Ass. datapoints**	**Percentage of total**	**Return rate**	**Cluster description**	**Risk label**
K-Proto	0	127258	8.1%	6.5%	Mid-range price, Black, Cat. L & I	Medium -
	1	35310	2.3%	7.2%	High priced, Cat. L & I, Black	Medium +
	2	1000487	63.8%	6.2%	Low-priced, Cat. H & L, color unknown	Medium -
	3	404729	25.8%	8.4%	Moderate prices, Cat. L & F, black	High
CAVE	0	547884	34.9%	7.5%	Middle priced, Cat L, black, unknown size	Medium +
	1	749463	47.8%	6.0%	Low priced, Cat H, unknown color and size	Medium -
	2	223976	14.3%	7.9%	Medium-high priced, Cat L, black and unknown colors, unknown size	High
	3	46461	3.0%	7.1%	High priced, mostly Cat L, Black, Unknown sized	Medium +
LCC	0	201866	12.9%	6.4%	Medium-high prices, Cat I & E, beige, unknown sizes,	Medium -
	1	302270	19.3%	6.2%	Low priced, Cat B, Black colored, unknown sizes	Medium -
	2	130048	8.3%	18.6%	Medium priced, Cat. F, black & red, diff sizes	High
	3	369593	23.6%	5.7%	High priced, Cat. L, Black, Blue & Pink	Low
	4	564007	36%	5.4%	Low priced, Cat. H, sizes and colors unknown	Low

Note: Overall return rate: 6.84%

Table 7.6: Summary of clusters - products

7.3. Suggested strategies

In this section, strategies will be developed based on findings from the literature in Section 3.3 and on the characteristics derived from the cluster analyses. The most important outcomes of the cluster analyses are highlighted in Table 7.7.

High Risk	**Low Risk**
Product level features	
Cat. F	Cat. H
Medium priced (7.35 ± 4.96)	Low / High priced (3.49 (± 5.88); 13.48 (± 14.75))
Black & Red	Unknown
Sizes different than "One Size"	Unknown / "One Size"
Order level features	
Large orders (5+)	Small & Medium orders (1-2; 3-4)
DHLDE	COLISSIMO, INPOS
Night & Evening	Morning & Afternoon
PostNL-PU (increases risk in average/low risk clusters)	PostNL-Stan
Increased risk in Cluster 2 product level LCC	
Black & Red	White, Pink, Purple & Multi
Highest prices among cluster 2 (8.37 (±5.18); 12.26 (±6.10))	Lower prices: 4.17 (±0.80) - 6.20 (±0.95)

Table 7.7: Summary high & low risk characteristics

7.3.1. Strategies

To target these high- and low-risk products, the following strategies to reduce return volumes are considered. The strategies are based on some constraints and assumptions. A key assumption in this research is that retailers have the ability and willingness to adjust their operational processes, marketing strategies, and return policies. This places the responsibility for return reduction largely within the organisation, assuming that customer behavior is a response to retailer-designed systems. The framework assumes that internal teams can collaborate effectively around shared data insights. However, the interviews showed that teams often have conflicting incentives. Strategies that require coordination across departments depend on organisational readiness and governance structures that may not yet exist.

Product Information & Fit As demonstrated by the cluster analysis, products in which size is a relevant factor are at higher risk of return. To address this, improvements can be made to product descriptions, images, and size guides. Providing detailed measurements can help customers make more informed choices. Additionally, incorporating "fit feedback" from previous buyers, particularly for Category F items, through customer reviews can help set realistic expectations for future customers and further reduce the likelihood of returns.

Another finding from the analysis is that "One Size" products have a lower return rate in Cluster 2. Where possible, adjusting products to a "One Size" format rather than offering multiple sizes can help reduce returns, as it minimizes the need for customers to order multiple sizes to find the best fit. However, there are Several trade-offs to consider when applying these strategies. Simplifying the assortment can help reduce returns by limiting size and style options. Still, it may also weaken revenue potential by offering customers less variety and reducing the appeal for those seeking specific fits or aesthetics. Likewise, providing more detailed fit information can reduce uncertainty and improve purchase accuracy, yet it risks overwhelming customers or lowering conversion if the complexity deters quick decision-making. Finally, moving toward "One Size" standardization can streamline production and reduce size-related returns. Still, it inevitably shrinks the addressable market by excluding specific customers, potentially harming overall sales and customer satisfaction.

Policy leniency & strictness The literature highlights that the leniency or strictness of return policies significantly influences return rates. Several studies recommend tailoring return policies to specific products, categories, or risk clusters, rather than applying a one-size-fits-all approach. In practice, this means implementing stricter return conditions for high-risk products or orders, such as those in Category F within Cluster 2, while maintaining more flexible policies for low-risk segments. For high-risk clusters, policy adjustments could include shortening the return window or introducing restocking fees. The cluster analysis in this thesis supports these recommendations: products in Cluster 2, especially those with higher prices, non-One Size sizing, and shipped via DHLDE, show a markedly higher return risk. Introducing restocking fees for large orders, which are prevalent in this cluster, could further reduce returns, as large, high-value orders are more likely to include items intended for return. These targeted policy changes align with both academic recommendations and the empirical findings from the clustering analysis.

Stricter, risk-based return policies offer a targeted approach to reducing return volumes, yet they introduce important strategic trade-offs. While shortening the return window or introducing restocking fees for high-risk segments can effectively discourage opportunistic or excessive returning behavior, these measures may simultaneously decrease customer satisfaction and reduce purchase likelihood for items already associated with higher return uncertainty. Products in Cluster 2, particularly those with defined sizing, higher price points, or shipments via DHLDE, exhibit a significantly elevated return risk, supporting the rationale for stricter policy interventions. However, imposing stricter conditions on these items risks creating perceptions of reduced fairness or increased friction in the purchasing process, potentially deterring customers from engaging with affected product categories. In this way, the benefits of reduced operational burden must be carefully weighed against potential negative impacts on customer trust, conversion rates, and long-term loyalty.

Operational improvements The cluster analysis yields two important insights for return reduction strategies on the operational side. The clusters show a high distinction between shipping carriers. The use of DHLDE as a shipping carrier is consistently associated with a higher likelihood of returns, whereas COLISSIMO is associated with a lower return risk. As both carriers operate in different countries, this pattern may reflect geographic variation. Further investigation is required to clarify why DHLDE contributes to elevated return rates and COLISSIMO to lower rates, with potential explanations including differences in delivery quality, customer expectations, or handling practices.

Furthermore, the Order Level Clustering shows that in Low and Medium risk clusters, PostNL's standard shipment has a lower risk than PostNL-PU. The retailer could incentivize the use of PostNL-Standard.

Secondly, the analysis reveals that, across all clusters, orders placed during the evening and night consistently have above-average return rates. Although the total number of orders during these hours is relatively small, this pattern suggests a higher return rate for late-night purchases. As a targeted intervention, retailers could temporarily close the webshop during high-risk hours to reduce returns from this segment, or implement additional prompts or reminders to encourage more deliberate purchasing decisions during these hours. These strategies align with literature recommendations to tailor interventions to high-risk segments and leverage operational levers, such as carrier selection and purchase timing, to reduce overall return volumes.

Marketing & Sales Aggressive promotions on high-risk products, such as heavy discounting or limited-time offers, can drive impulse purchases and, in turn, increase return rates. To mitigate this, avoid using such promotional tactics for products identified as high risk through clustering analysis, particularly Category F items in Cluster 2. Instead, allowing customers more time to consider their purchases can help reduce impulsive buying and lower returns in these segments. For marketing efforts, focus on low-risk products in other categories, or, if Category F must be promoted, target items that are One Size or in colors like pink or blue, rather than black or red, as these attributes are associated with lower return rates. The cluster analysis confirms that targeted offers and promotions can be used more confidently for low-risk segments, such as those found in Clusters 3 and 4.

Additionally, flagging large, high-value orders for post-purchase confirmation, such as sending a follow-up message to verify the customer's selection, can help prevent bracketing and unnecessary returns. This can reduce the number of large orders in high-risk clusters, which are more likely to be returned.

Offering bundle discounts only when all items are kept further discourages partial returns. A technical improvement could be the implementation of a live flagger, which alerts customers when they are ordering many items and offers assistance, rather than allowing high-risk orders to proceed unchecked.

Currently, incentives like threshold-based free shipping are commonly used to encourage customers to order more. However, the cluster analysis shows that large orders are associated with increased return risk. Lowering the free shipping threshold could reduce the incentive to place larger orders, thereby decreasing returns. Since low-risk orders are typically medium-sized, setting the free shipping threshold around this order value may help reduce returns without negatively impacting sales. Further research is recommended to determine the optimal threshold that balances sales and return reduction.

Implementing marketing and promotional strategies targeted at high- and low-risk product segments involves several trade-offs that must be carefully weighed. Avoiding aggressive promotions for high-risk items can help reduce impulsive purchases that typically lead to higher return rates, as indicated by the clustering analysis. However, limiting promotional activity in these categories may also suppress short-term sales volumes and reduce the visibility of products that rely on discounts to gain competitive advantage. Similarly, introducing post-purchase confirmation steps or live flagging systems for large, high-value orders can help prevent bracketing behavior and reduce returns. Still, these interventions add friction to the customer journey and risk lowering conversion rates, particularly among confident or experienced shoppers. Adjusting free-shipping thresholds to discourage large, high-risk orders presents another trade-off. While it may decrease return-prone order patterns, it can also weaken a key incentive that drives overall cart value and customer motivation to complete purchases. Thus, while these strategies align with the empirical patterns identified in the cluster analysis, especially the heightened return risk associated with large orders and specific product attributes, they require balancing the benefits of reduced returns against potential declines in customer satisfaction, purchase intent, and revenue performance.

Customer perspective While the proposed strategies are grounded in data-driven insights, their effectiveness also depends on how customers perceive and experience them. Considering the customer perspective is essential, as return-reduction measures that introduce friction or reduce flexibility may unintentionally harm satisfaction, trust, and long-term loyalty.

Product information and fit–related interventions generally align well with customer expectations. Enhanced size guides, clearer product descriptions, and fit feedback improve decision-making and reduce uncertainty, which customers typically view positively. However, customers may experience information overload when product pages become excessively detailed, potentially increasing cognitive effort and discouraging purchases. Likewise, a shift toward “One Size” items may simplify choices but reduce inclusivity and can alienate customers whose fit needs fall outside the standardized range.

Risk-based return-policy adjustments have more complex implications. Stricter conditions for high-risk items, such as shorter return windows or restocking fees, may effectively discourage opportunistic returns. Still, customers may also perceive these measures as unfair or punitive, especially when applied to categories already associated with fit uncertainty. Such perceptions can reduce purchase confidence, increase pre-purchase hesitation, and damage retailer credibility. Transparency and clear communication are, therefore, critical for maintaining customer trust. Furthermore, hassle costs are experienced as more expensive than, for example, monetary punishments.

Operational interventions, such as encouraging specific carriers or adjusting webshop availability during high-risk hours, may also influence the customer experience. Although improved shipping reliability can enhance satisfaction, restricting purchase hours or adding friction to late-night shopping may be perceived as intrusive or inconvenient. Customers increasingly expect 24/7 availability and seamless purchasing; deviations from this norm require careful justification and communication. Marketing and sales adjustments also create trade-offs from the customer’s perspective. Avoiding aggressive promotions on high-risk items reduces impulse buying but may diminish perceived value, particularly for price-sensitive customers. Additional confirmation steps or live flagging systems may help reduce bracketing, but can be experienced as interruptions or questioning of customer intent. Similarly, lowering free-shipping thresholds may weaken a key incentive that customers rely on to justify online purchases.

Overall, the customer perspective underscores the need for strategies that balance return reduction with a positive shopping experience. Interventions should be designed to minimize friction, maintain transparency, and ensure that differentiated conditions are perceived as reasonable and justified. Retailers must therefore consider not only the operational and financial benefits of each strategy but also their potential effects on customer behavior, satisfaction, and long-term loyalty.

7.3.2. Strategy validation

The proposed strategies were reviewed and assessed with supply chain experts. The strategies were evaluated for feasibility and alignment with insights from the clustering analysis. Enhancing product information was validated as a highly practical and impactful approach, as experts emphasized that improving clarity in product descriptions directly reduces avoidable returns. Extra measures could be given on the model's size. Policy adjustments were identified as potential interventions; however, experts advised that further internal research is needed to ensure these measures do not negatively affect customer satisfaction or operational stability. One expert noted that monetary punishment is more customer-friendly than a hassle-based punishment. Operational strategies require refinement, as the incentive for PostNL is ineffective given DHLDE's limited operations in Germany, underscoring the importance of country-specific considerations in both order patterns and return flows.

Experts also stressed several broader practical constraints. First, strategies targeting high-risk ordering hours (evening and night) were deemed impractical, as retailers cannot influence customer ordering behavior at specific times. Marketing and sales initiatives, however, were confirmed as relevant and actionable, and can now be tailored according to product-level risk patterns. Adjustments to the free-shipping threshold should take both order value and number of items into account. Flagging unusually large orders for manual confirmation was considered resource-intensive and difficult to operationalize, particularly given current staffing levels and existing system limitations.

Furthermore, the interviews highlighted that operational feasibility depends strongly on workload distribution and warehouse capacity. This validation process and the underlying considerations are documented in the interview summary in Appendix B.2.

7.4. Conclusion

Taken together, K-prototyping creates a large imbalance in cluster sizes at both the order and product levels. This may limit the visibility of patterns. The CAVE algorithm improves this balance slightly, but the smallest cluster still accounts for only a small part of the observations. Both clustering methods seem to separate based on the numerical variables, with the numerical variables having a larger influence on the calculated weight than the Category differences. At the product level, no distinctly high- or low-risk clusters emerge, which limits its usefulness for strategic differentiation. LCC seems to separate more on categorical features, such as shipping carrier and method at the order level, and product Category at the product level. LCC shows a more distinct segmentation than the other clustering methods, providing clearer separation between high- and low-risk clusters.

At the order level, smaller orders and lower amounts are associated with lower risk, whereas larger orders with higher values increase risk. All cluster methods show a similar distribution in time of day, with the return rate in the evening and night having a higher risk. Among carriers, INPOST and COLISSIMO emerge as the most reliable, associated with lower return rates. On the contrary, DHLDE seems to increase the risk of returns. PostNL-Standard seems to reduce risk compared with PostNL-Pickup, but PostNL in general seems a low-risk carrier.

On the product level, k-prototyping and CAVE show the average prices and products in the Category. L and F have an increased risk. LCC clusters show that the high-risk cluster primarily consists of products from Category F, predominantly black in color, and priced at an intermediate level. Size plays a critical role within this cluster, as specified sizes consistently show above-average return rates. The LCC method is further applied for creating subclusters in cluster 2. This shows that sizes other than "One Size" have a higher risk in the high-risk cluster. Furthermore, Black and Red have an increased risk. As the final part of the analysis, the product order features were analyzed. The high-risk cluster showed mostly large orders and high order values, suggesting increased risk. Cluster 3, a low-risk cluster, shows high-value orders with fewer items, suggesting this combination poses less risk.

Lastly, a potential strategy to reduce return risk is to improve product fit information. Furthermore, further investigation into policy leniency, operational improvements, and marketing & sales strategies is suggested.

These strategies always come with a trade-off. To avoid aggressive promotions for high-risk products, it is necessary to find a balance between reducing impulse-driven purchases and return rates, and lowering short-term sales while avoiding price-sensitive customers who might switch to competitors offering stronger discounts. On the other hand, targeting low-risk items improves conversion and maintains low return levels, yet limits promotional visibility for high-risk categories, potentially reducing demand and competitiveness in those product groups. Flagging large, high-value orders for post-purchase confirmation helps prevent bracketing and unnecessary returns. Still, it adds friction to the buying process, which may frustrate customers, reduce conversion, or push them toward competitors that offer more seamless experiences—offering bundle discounts only when all items are kept. This discourages partial returns and aligns incentives with retailer goals, but may feel restrictive to customers and reduce their willingness to buy bundles at all. Implementing a live flagger for risky orders reduces excessive quantity orders and encourages more deliberate purchasing, but increases intervention in the customer journey, which may be perceived as intrusive and reduce shopping satisfaction. Lowering the free-shipping threshold decreases incentives for large, return-prone orders and lowers return rates, but can reduce average order value and weaken a key competitive advantage in online retail (free shipping). Overall, the trade-off is mostly about reducing return risks, but at the cost of reduced sales and customer satisfaction. The right balance in this can only be found by identifying which items are at high risk and working together, drawing on input from different stakeholders within the retailer.

These strategies inevitably involve trade-offs that require careful consideration. Avoiding aggressive promotions for high-risk products helps reduce impulse-driven purchases and subsequent returns. Yet, it may also lower short-term sales and increase the likelihood that price-sensitive customers turn to competitors offering deeper discounts. Conversely, directing promotional efforts toward low-risk items supports higher conversion rates and helps maintain low return levels. However, this approach reduces promotional visibility for high-risk categories, potentially diminishing demand in those segments. As a result, competitiveness may weaken for high-risk product groups, even when these items are potentially high-running products. Introducing post-purchase confirmation for large, high-value orders can effectively limit bracketing and prevent unnecessary returns. However, this added friction in the purchasing process may frustrate customers or deter them from completing their order. Similarly, bundle discounts that apply only when all items are kept can discourage partial returns and align customer incentives with retailer objectives, but may feel overly restrictive and reduce customers' willingness to purchase bundles. Implementing a live flagger to alert customers when placing unusually large or risky orders can encourage more deliberate purchasing behavior. Yet, such interventions may be perceived as intrusive and adversely impact the shopping experience. Finally, lowering the free-shipping threshold reduces the incentive to place large, return-prone orders, but risks decreasing average order value and weakening a key competitive advantage. Overall, these strategies highlight a recurring tension between reducing return risks and preserving sales performance and customer satisfaction. Achieving an optimal balance requires a clear understanding of which items fall into high-risk clusters and close collaboration among stakeholders across the retailer's organization.

8

Discussion & Conclusion

This chapter synthesizes the study's findings by addressing the main research question and its sub-questions. It further discusses the implications and finally reflects on the study's limitations. Lastly, it will give recommendations and suggestions for future research.

8.1. Discussion

This study was guided by the objective of reducing return volumes within the e-commerce operations. To achieve this, a Design Science Research approach was employed to systematically define and analyze the underlying problem. The central research question formulated to address this objective is: *"How can a reduction of returned goods in fashion- and electronic e-commerce be achieved by designing a framework that clusters high- and low-risk orders and products?"*. Four sub-questions were developed to support this aim. This section presents the main findings derived from the research.

8.1.1. Main findings

Sub-question 1: Influential factors

The first sub-question is: *Which factors are shown by literature and historical data to significantly affect the return rate?* Existing research identifies product, customer, and order attributes as key drivers of returns. However, interviews with retail experts reveal that many retailers do not collect detailed customer data, such as gender, age, or return history, due to privacy considerations. As a result, the analyses in this study focus on product and order attributes and assess their significance. The analysis demonstrates that product returns are driven by a combination of product-specific attributes and transactional factors rather than a single dominant feature. At the order level, approximately 5% of all orders are returned. Operational variables such as shipping method, delivery mode, and shipping carrier exhibit significant differences in return rates across categories. This is supported by the literature of Mishra & Dutta. Price-related variables also matter: higher-priced items and higher order values correlate positively with returns. This finding supports Asdecker & Karl's findings, suggesting that high-value items and orders have a higher return risk. Order quantity, on the other hand, shows a weak negative correlation, indicating that customers buying multiple items are slightly less likely to return, which contradicts the literature of Asdecker & Karl. Although order quantity shows a statistically significant negative correlation with returns, the effect size is extremely small. This indicates that the direction of the association is unstable and may reflect random variation rather than a meaningful pattern. Furthermore, the analysis reveals that return rates vary significantly across different time-of-day segments. Although such temporal patterns have not been documented in existing literature, the observed differences indicate variation between these groups. At the product level, the overall return rate is 6.84%, yet substantial variation is observed across product categories, sizes, and colors. Larger sizes and specific colors exhibit disproportionately high return rates, suggesting potential issues related to fit, product expectations, or the accuracy of product representation. These findings are consistent with those reported by Cui et al. (2020), who also identified product attributes as significant determinants of returns.

Sub-question 2: Current approach and limitations

The second sub-question is as follows: *How are retailers currently handling returns, and what limitations exist in these approaches?* Retailers currently handle returns primarily through manual, in-house processes. Customers initiate returns via online platforms or customer service, after which products are inspected and either restocked, refurbished, or discarded depending on their condition and value. Although some automation is in place, most activities are done manually, making return processing labor-intensive and leading to high operational costs, particularly given the substantial volume of returns. Integration between return processes and warehouse or ERP systems is often limited, resulting in inefficiencies and additional manual work. While retailers collect considerable return data, this information is not utilized for analytics or process optimization. Instead, most rely on historical averages rather than dynamic, data-driven strategies. Preventive strategies, such as return fees or improved product information, are less frequently implemented, with most efforts focused on reactive handling. The involvement of various stakeholders, including logistics providers, warehouse managers, IT teams, and sustainability officers, further complicates process improvement. Economic efficiency generally takes precedence over sustainability, and the environmental impact of returns is not adequately addressed. Figure 5.1 provides an overview of the various workflows involved in the returns process. It adds information on stakeholder responsibility and level of automation to the state-of-the-art literature. This representation is derived from insights obtained through the literature and expert interviews, ensuring that both theoretical and practical perspectives are reflected. Although some retailers outsource the entire returns process to third-party service providers, such cases appear to be relatively uncommon in practice and were therefore not incorporated into the scope of this study. The main limitations of the current approach include high return volumes, high costs, and labor intensity. Additionally, underutilization of data and limited adoption of automation or preventive measures underscore the need for data-driven strategies to reduce returns. The studies conducted by Frei et al. (2019, 2022) and Stevenson et al. describe return-related actions but do not specify the responsible stakeholders. The present research seeks to address this gap by identifying stakeholder responsibilities; however, determining a clear problem owner remains difficult, as stakeholders pursue divergent primary objectives. Consequently, although the retailer is nominally the problem owner, no specific internal actor assumes responsibility for developing a structural solution.

Sub-question 3: Aim and structure of the framework

The first two sub-questions establish the context and problem environment for the DSR approach. This sub-question introduces the design component of the research framework. The sub-question is as follows: *What is the aim and structure of a return flow optimizing framework, and how can it help the retailer optimize their system?* The primary aim of the return optimizing framework is to reduce the volume of returns in e-commerce by identifying and addressing high-risk products and orders. Achieving this objective requires a systematic approach that enables retailers to distinguish between high- and low-risk segments and implement targeted strategies to reduce returns.

The framework is designed to meet several essential requirements. First, it must be interpretable, providing clear, explainable results so stakeholders can understand which factors drive returns and why certain products or orders are classified as high or low risk. Second, the framework should make effective use of available data, leveraging both numerical and categorical features, as much of the data collected is currently underutilized. Third, scalability is crucial, as the model must handle large datasets typical of e-commerce environments.

The structure of the framework consists of the following iterative steps (see Figure 6.2):

1. **Data Collection:** Gather order and product data
2. **Clustering Analysis:** Groups orders and products into clusters based on shared features.
3. **Pattern Identification** Visualize and analyze the resulting clusters to identify patterns.
4. **Strategy Development:** Develop targeted strategies to address the risks identified in high-risk clusters.
5. **Stakeholder Validation:** Validate the proposed strategies with relevant stakeholders.
6. **Iteration and Review:** The process is iterative; after a period of implementation, the framework should be revisited.

This approach ensures that the framework remains adaptive and responsive to changing return patterns and business needs. By systematically applying these steps, e-commerce retailers can continuously improve their return management processes, reducing operational costs.

Implementing this framework enables retailers to develop effective strategies for reducing returns. By reducing return volumes, inventory management becomes more streamlined, with less disruption and complexity in stock handling. The returns process itself becomes faster and more efficient, reducing operational costs and freeing up resources for further process automation. Additionally, customer satisfaction is likely to improve, as customers are less likely to encounter issues that prompt returns. This reduces the likelihood of errors in the returns process and helps maintain a positive customer experience. This framework is intended to support stakeholders involved in return logistics in developing strategies to reduce return volumes. However, its effectiveness has thus far been evaluated using only a single case study within this research, which limits the extent to which broader conclusions can be drawn.

This type of framework has not previously been proposed in the literature. While studies by Niederlaender et al. (2024) and Karl evaluate different methodological approaches to optimization, they do not incorporate stakeholder perspectives. Likewise, Zennaro and Frei focus on process-oriented improvements, but neither integrates data analytics, process insights, and stakeholder considerations into a single approach.

Sub-question 4: High-risk orders and products

The fourth question to answer is: *What orders and products are high or low risk, what are their characteristics, and what strategies can help reduce this risk?* The identification of high- and low-risk orders and products is achieved through clustering analysis, which groups orders and products based on shared characteristics such as product category, color, size, price, order value, and shipping method. The analysis reveals that certain clusters consistently exhibit higher return rates, while others exhibit lower risk.

High Risk High-risk products and orders characterized by belonging to product category F, which show return rates as high as 18.6%. Furthermore, products having attributes such as sizes other than "One Size" and colors like black and red are associated with higher return rates. Also, being part of large, high-value orders has increased the risk of a product being returned. Having the shipping carrier DHLDE is also associated with an elevated return risk. Lastly, orders placed during evening and night hours show consistently above-average return rates. Note that a higher product price does not necessarily increase the risk of returns. In fact, the outcomes in this case indicate that medium-priced products are more likely to be returned than both low- and high-priced products, which contradicts the findings of Asdecker & Karl. It is important to note, however, that the labels "low-priced" and "high-priced" products in this study are defined relative to the price distributions within each (sub)cluster. Consequently, a product classified as high-priced within a subcluster may not correspond to the high-priced category in the overarching dataset. Because these classifications are inherently context-dependent and thus subjective, differences in operationalization of price categories may explain why the results reported by Asdecker and Karl diverge from those of the present study.

Low Risk Low-risk products and orders are typically in Category H and have attributes such as "One Size" or sizes where fit is less of an issue. It is more associated with colors such as pink or blue. These products are more often in medium-sized orders with moderate order values. This finding aligns with those of Cui et al., which highlight that different product categories entail different risks. Lastly, PostNL is associated with a lower return risk. This observation is consistent with Mishra & Dutta's findings, which highlight that the shipping carrier and method significantly influence return risk.

Strategies The suggested strategies for implementation are based on the characteristics identified in the cluster analysis. Firstly, improving product information and fit guidance is essential, especially for products where size is a key factor; detailed descriptions, accurate size guides, and customer fit feedback can help customers make better choices and lower return rates. Furthermore, offering more "One Size" options where feasible can also reduce returns. Return policy design should be tailored to risk profiles, with stricter conditions, such as shorter return windows or restocking fees, applied to high-risk products and orders, while maintaining flexibility for low-risk segments. Operational improvements include incentivizing the use of lower-risk shipping carriers, if multiple carriers are available in a country. Additionally, considering interventions for orders placed during high-risk times, such as evening and night hours, where closing the webshop overnight is not preferred. In marketing and sales, aggressive promotions should be avoided for high-risk products, and post-purchase confirmations or live flaggers can help prevent unnecessary returns from large, high-value orders. Adjusting free shipping thresholds to discourage large, high-risk orders may also be effective. These strategies are validated through interviews. The trade-offs of these strategies reveal a recurring tension between reducing return risk and preserving sales performance and customer satisfaction. Achieving an optimal balance requires a clear understanding of which items fall into high-risk clusters and close collaboration among stakeholders across the retail organization. This trade-off was previously discussed by Duong et al. (2025) in the context of return-policy leniency and strictness, but is largely overlooked in other studies. The present findings show that this tension extends beyond policy design and also appears in other types of return-reduction strategies, such as those related to the information provided to customers.

Integrated Findings and Contributions to the Literature

Taken together, the results of the four sub-questions show that a reduction of returned goods in fashion- and electronic e-commerce can be achieved by applying a stakeholder-informed, data-driven framework that identifies high- and low-risk products and orders and translates these insights into targeted interventions. The analysis of influential factors demonstrates that returns are driven by a combination of product attributes, transactional characteristics, and logistical conditions rather than by a single variable. Understanding these patterns enables more precise segmentation. The evaluation of current practices highlights that retailers predominantly rely on manual handling, limited automation, and underutilized return data, reinforcing the need for a structured analytical approach. The developed framework addresses this gap by clustering products and orders, uncovering risk-driven patterns, and connecting these patterns to practical strategies that stakeholders can validate. Finally, the cluster results, especially those derived from LCC, show clear distinctions between high- and low-risk segments, enabling the design of tailored strategies, such as improved product information, differentiated return policies, operational adjustments, and marketing interventions. Together, these insights confirm that the proposed clustering-based framework provides an effective and actionable method for reducing return volumes in e-commerce.

This research makes several contributions to the existing literature on e-commerce return management, data-driven decision support, and clustering methodologies. First, it bridges the gap between technically oriented return-optimizing models and the practical realities of organizational return handling. Whereas previous studies frequently focus on algorithmic performance or process improvement in isolation, this thesis introduces a stakeholder-informed, interpretable framework that integrates data analytics with operational feasibility. By incorporating expert insights into the design of the artifact, the study demonstrates how analytical outcomes can be translated into interventions that align with organizational constraints, an element largely absent in prior work. Second, the study provides new empirical evidence on the interaction of product- and order-level attributes in shaping return behavior. It identifies variations across product categories, sizes, color, shipment methods, and order compositions, and presents findings that nuance or challenge earlier assumptions. For example, the elevated risk associated with mid-priced items and the weaker-than-expected relationship between multi-item orders and bracketing. These insights contribute to ongoing efforts to refine features associated with return risk. Third, this work offers a methodological contribution by systematically comparing three clustering approaches on large-scale, mixed-type e-commerce data. The results show that LCC provides the clearest segmentation of high-risk groups, advancing methodological guidance for future research. Finally, the study connects cluster-derived risk profiles to validated, actionable strategies to reduce returns, thereby demonstrating how data-driven segmentation can inform differentiated return policies, product information improvements, operational adjustments, and marketing decisions. This strengthens the link between analytical research and practical return-management interventions.

8.1.2. Implications

The findings of this thesis have several important implications for both research and practice. This study aimed to fill a research gap by developing a comprehensive, stakeholder-informed framework that combines data-driven analytics, process mapping, and practical interventions. The clustering-based framework developed in this study advances the literature on return management by demonstrating that data-driven segmentation, leveraging both product- and order-level features, can reveal actionable patterns in return risk. This approach deepens understanding of the complex drivers of returns. It demonstrates the value of integrating operational data with stakeholder perspectives, supporting the preventive management of returns emphasized in the literature. The framework also illustrates that meaningful reductions in returns can be achieved through careful data analysis and stakeholder involvement, even before implementing complex models or automation. This ensures that interventions are focused on actual problems identified in the data, rather than being retrofitted to pre-existing solutions. In addition, this research addresses a gap in the literature by providing detailed insights into the automation level and stakeholder responsibilities of current return process activities, which had not been explicitly described before.

Furthermore, the findings indicate that the choice of clustering method substantially influences the resulting cluster structures. Although the CAVE algorithm was expected to yield better outcomes than K-prototyping, both methods produced similar results, with numerical features remaining the primary determinants of cluster formation. This outcome diverges from the expectations set by Ji et al. (2013) [26]. The cluster outputs of these methods showed little difference in return rates across clusters, making it hard to build strategies on. In contrast, the LCC method demonstrated the clearest differentiation between high- and low-risk clusters in e-commerce data, suggesting it is the most suitable clustering approach for this dataset. Lastly, the clustering at the product level provided the best insights for strategizing, with the addition of order-level features later to the analysis. This makes clustering based solely on order-level features redundant.

From a practical standpoint, the framework provides e-commerce retailers with a systematic method to identify high- and low-risk products and orders, enabling the development of targeted strategies to reduce returns. By implementing these strategies, retailers can streamline inventory management, reduce operational costs, and improve customer satisfaction. The iterative nature of the framework ensures that interventions remain relevant as consumer behavior and market conditions evolve, supporting continuous improvement in return management processes. The study also highlights which stakeholders are involved in the return process and identifies who is responsible for each step. The findings show that different stakeholder groups within a retail organization often have conflicting goals, which makes it unclear who ultimately owns the return-reduction problem. This lack of a clear problem owner means that no single actor is responsible for aligning these differing objectives and translating them into coherent organization-wide strategies. When internal roles and responsibilities are too fragmented to support this coordination, an external party may be needed to bring these perspectives together and guide the development of effective return-management strategies. Furthermore, the research underscores the economic and environmental benefits of reducing returns. Lower return volumes not only decrease waste and associated emissions but also free up resources for investment in automation and sustainability initiatives. Policymakers and industry leaders may draw on these insights to encourage best practices in reverse logistics and promote more sustainable consumption patterns.

Currently, e-commerce returns contribute significantly to environmental pollution. This research aims to support a reduction in unnecessary product returns, addressing not only the transportation-related emissions but also the waste of raw materials and time associated with processing returned goods. By providing retailers with data-driven strategies to minimize avoidable returns, this work seeks to reduce both the economic and environmental costs inherent in the current returns process.

8.1.3. Limitations

Several limitations emerged, offering direction for further research and improvement. These constraints primarily concern data availability, the performance and interpretability of the clustering methods, and the extent to which the findings can be generalized to other contexts.

Limitations in data

The dataset used in this study does not include customer demographic information, such as age or gender. Consequently, it is not possible to analyze whether certain customer groups are at a higher risk of returning products. This limits the depth of the research, as relevant patterns associated with customer characteristics remain unexamined. As a result, all analyses focus on the order and product levels, omitting potential explanatory variables identified as significant in other studies.

A major limitation of this study is the absence of detailed information on return reasons within the dataset. After clustering the data points, the return reasons could have been examined within each cluster to determine why some clusters exhibit high or low return rates (independent of explanatory variables, based on reasons provided by customers themselves). Specific return reasons may indicate structural issues in the ordering or delivery process, or issues with product quality or fit. Due to the lack of such data in this research, it is not possible to analyze these relationships or provide targeted recommendations based on these insights. Future research would benefit from collecting and integrating return reasons to develop a more comprehensive understanding of the drivers behind product returns and to design more effective interventions.

The analyses depend on the quality, completeness, and time span of the provided data. In this study, data quality was verified, and missing values and outliers were addressed. Nevertheless, several limitations remain: the dataset covers only five months of transactions, while the return period is one month. Consequently, returns for purchases made in the final month of the dataset but processed after the observation period may not be included in the analysis. This could lead to an underestimation of the actual number of returns during that period. Additionally, errors, inconsistencies, or seasonal effects may influence the results. There is also a risk that certain product groups or product categories are under- or overrepresented, potentially introducing bias into the findings. Future research should aim for a longer observation period to ensure that returns within the full return window are consistently captured.

Limitations in clustering

In the clustering methods, some limitations arose. The results indicate that the k-prototypes algorithm is predominantly driven by numerical features, reflecting the strong influence of the Euclidean distance component in the dissimilarity measure. At the same time, categorical attributes receive comparatively less weight in the clustering process. Although the CAVE algorithm seeks to address the limitations of k-prototyping by introducing variance- and entropy-based feature weighting, the clustering outcomes in this study nonetheless indicate that numerical attributes continue to dominate the partitioning process. In contrast, the LCC method uses a probabilistic modeling approach, assigning each observation a probability of membership across all latent classes. Because it relies on probability distributions rather than distance-based criteria, it is particularly effective in identifying patterns within categorical variables. It is not subject to the numerical dominance inherent in Euclidean-based clustering methods. LCC is a powerful technique for uncovering complex patterns, but performing these analyses is highly time-intensive on standard hardware. This may restrict its practical applicability, particularly for companies without access to advanced IT infrastructure. It is advisable to exploit optimizations or alternative methods that require less computational power. In this study, processing approximately 1.5 million observations took about 1 day, illustrating the substantial computational burden of large-scale data analysis. Although the alternative clustering methods applied were somewhat less time-intensive, they still required substantial computational resources, particularly when handling numerous categorical variables and high-dimensional feature spaces. Investigating whether fewer categories or alternative, more efficient methods could yield comparable results may be worthwhile. Lastly, the use of advanced clustering techniques and numerous variables introduces a risk of overfitting: the model may adapt too closely to the specific dataset, resulting in poorer performance on new data.

Practical limitations

This research primarily focuses on quantitative data, such as order and product characteristics, return rates, and logistical variables. As a result, qualitative aspects, such as customer satisfaction, perceptions of the return process, and underlying motives for returning products, are largely excluded. This means that important explanations for return behavior, such as frustration over unclear product information, negative delivery experiences, or dissatisfaction with return policies, are not incorporated into the analysis. Similarly, insights from customer feedback, reviews, or consumer interviews are absent, even though they could provide a valuable understanding of the deeper causes of returns. Consequently, the study offers insight mainly into what happens (which products and orders are returned), but less into why customers decide to return items. This omission may lead to missed opportunities for improving the return process, customer communication, or product presentation. Furthermore, the lack of qualitative customer insights could result in proposed interventions that do not fully align with customer expectations, ultimately limiting the effectiveness of return-reduction strategies. Future research should place greater emphasis on customer experiences with newly implemented strategies. Systematically collecting and analyzing customer feedback on the return process could also yield valuable insights. Combining such qualitative data with quantitative analyses would create a more comprehensive understanding. My own perspective also shaped the direction of this research. By placing responsibility primarily on the retailer rather than the customer, the study focused on strategies that require organizational rather than behavioral change. This stems from the view that many customer behaviors often labeled “problematic,” such as bracketing or strategic basket filling, are in fact enabled or incentivized by retailers through generous return policies, free shipping thresholds, and marketing practices designed to maximize sales. As a result, the strategies developed in this research focus on adjustments within the retailer’s sphere of influence rather than on modifying customer behavior. This reflects the underlying assumption that retailers have both greater control over return-related processes and a responsibility to design systems to minimize avoidable returns.

Generalizability

The findings of this study are derived from a single case study, and any attempts to generalize them to other contexts should be approached with caution. Although the findings of this case study have limited generalizability, the framework presented in Chapter 6 offers a broader structure that can be applied across different contexts. This framework serves as a conceptual foundation for future research aimed at testing and refining its applicability in other sectors, product profiles, or logistical processes. By validating and adapting the framework in diverse environments, subsequent studies can enhance its robustness and practical relevance.

8.2. Conclusion

This study aims to reduce return volumes within the e-commerce operations by designing and evaluating a framework capable of identifying high- and low-risk orders and products. Using a Design Science Research approach, the work systematically analyzed the underlying problem, reviewed influencing factors, assessed current return-handling practices, and constructed an interpretable, scalable framework informed by both data analytics and stakeholder perspectives. Across the four sub-questions, the findings show that a combination of product attributes and order-level variables, including shipping carrier, delivery mode, order value, and order time, drives returns. Current return-management practices rely heavily on manual, reactive processes with limited use of available data, highlighting a clear opportunity for data-driven decision support. The proposed framework addresses this gap by integrating clustering analysis, pattern identification, and stakeholder validation into an iterative process that enables targeted strategy development. Applying the framework to the case study reveals distinct high- and low-risk product and order profiles. It provides actionable strategies to reduce the likelihood of returns through tailored policy design, improved product information, operational adjustments, and marketing interventions. While the findings demonstrate the framework’s potential to support retailers in reducing return volumes and improving operational efficiency, its validation in a single case limits generalisability. Future applications across broader contexts are therefore needed to further substantiate its robustness and practical value. This section further outlines the recommendations arising from the study, followed by potential directions for future research.

8.2.1. Recommendations

The thesis identifies several targeted strategies to reduce product returns in e-commerce. First, improving product information and fit guidance, such as detailed descriptions, images, and size guides, can help customers make better choices, especially for products where sizing is critical. Incorporating customer fit feedback and focusing on “One Size” products, where possible, further reduces return risk. Second, return policies should be tailored to product risk profiles, with stricter conditions like shorter return windows or restocking fees for high-risk items, while maintaining flexibility for low-risk segments. Operational improvements addressing high-risk hours, such as closing the webshop or adding prompts to encourage more deliberate purchasing. Additionally, incentivizing low-risk shipment methods could also decrease returns. In marketing and sales, aggressive promotions for high-risk products should be avoided, and targeted offers should focus on low-risk segments. Technical solutions, such as live flaggers and post-purchase confirmations for large orders, can help prevent unnecessary returns. Finally, adjusting free shipping thresholds to discourage large, high-risk orders may further reduce returns, and ongoing research is recommended to optimize these interventions. Additionally, examining why certain carriers are consistently associated with higher or lower return risks may yield further insights for developing targeted return-reduction strategies, particularly in light of cross-country operational differences. In conclusion, it is essential to designate a responsible actor within the retail organization who can coordinate cross-functional collaboration and facilitate the necessary trade-offs associated with the proposed return-reduction strategies.

8.2.2. Future research

Future research should address several areas to strengthen the understanding of return behavior and improve practical interventions. Collecting and integrating return reasons would provide deeper insights into the drivers behind product returns and support the development of more effective strategies. Extending the observation period is essential to capture seasonal variations and ensure that all returns within the full return window are included. It would also be valuable to examine whether reducing the number of categories or applying alternative, more efficient clustering methods can produce comparable results with lower computational demands.

Further studies should place greater emphasis on customer experiences with newly implemented strategies. Systematic collection and analysis of customer feedback on the return process could yield important insights. Combining qualitative data with quantitative analyses would create a more comprehensive understanding of return behavior and enable the design of customer-oriented interventions.

The framework presented in this study provides a conceptual foundation for future research. Testing and refining its applicability across different sectors, product profiles, and logistical processes will enhance its robustness and practical relevance. Additional research is needed to investigate the specific causes of high return rates within certain product categories. This requires a deeper analysis of product characteristics, customer behavior, and external factors to develop targeted interventions.

The analysis also indicates that using DHLDE as a shipping carrier is associated with a higher return rate. The underlying causes were not examined in this study. Future research should explore this phenomenon by assessing delivery quality, customer perception, and operational processes. Since DHLDE is linked to orders from Germany, examining cross-country differences could provide further insights.

Finally, future studies should investigate the causes of high-risk return clusters and evaluate the long-term effects of specific interventions, for example, by building mathematical models to calculate financial and environmental consequences. A last proposal for future research is to develop real-time techniques to identify high-risk orders and assist customers in making better purchase decisions.

References

[1] H. Abdulla, M. Ketzenberg, and J. D. Abbey. “Taking stock of consumer returns: A review and classification of the literature”. In: *Journal of Operations Management* 65.6 (July 2019), pp. 560–605. ISSN: 1873-1317. DOI: 10.1002/joom.1047.

[2] T. Adenuga, A. T. Ayobami, and F. C. Okolo. “AI-Driven Workforce Forecasting for Peak Planning and Disruption Resilience in Global Logistics and Supply Networks”. In: *International Journal of Multidisciplinary Research and Growth Evaluation* 2.2 (2020), pp. 71–87. ISSN: 2582-7138. DOI: 10.54660/.ijmrge.2020.1.2.71-87.

[3] Y. I. Alzoubi. “AI-Powered Reverse Logistics: A Pathway to Sustainable Supply Chains”. In: *Advanced Network Technologies and Intelligent Computing*. Cham: Springer Nature Switzerland, 2025, pp. 384–402. ISBN: 978-3-031-83793-7. DOI: 10.1007/978-3-031-83793-7_25.

[4] B. Asdecker and D. Karl. “Big data analytics in returns management – Are complex techniques necessary to forecast consumer returns properly?” In: CARMA2018. Universitat Politècnica València, July 2018. DOI: 10.4995/carma2018.2018.8303.

[5] B. Asdecker and D. Karl. “SHEDDING SOME LIGHT ON THE REVERSE PART OF E-COMMERCE: A SYSTEMATIC LOOK INTO THE BLACK BOX OF CONSUMER RETURNS IN GERMANY”. In: *European Journal of Management* 22.1 (Mar. 2022), pp. 59–81. ISSN: 2378-8526. DOI: 10.18374/ejm-22-1.4.

[6] B. Baldi et al. “Consumer-Centric Supply Chain Management: A Literature Review, Framework, and Research Agenda”. In: *Journal of Business Logistics* 45.4 (Oct. 2024). ISSN: 2158-1592. DOI: 10.1111/jbl.12399.

[7] N. N. Bechwati and W. S. Siegal. “The Impact of the Prechoice Process on Product Returns”. In: *Journal of Marketing Research* 42.3 (2005), pp. 358–367. DOI: 10.1509/jmkr.2005.42.3.358.

[8] J. vom Brocke, A. Hevner, and A. Maedche. “Introduction to Design Science Research”. In: *Design Science Research. Cases*. Springer International Publishing, 2020, pp. 1–13. ISBN: 9783030467814. DOI: 10.1007/978-3-030-46781-4_1.

[9] H. Cui, S. Rajagopalan, and A. R. Ward. “Predicting product return volume using machine learning methods”. In: *European Journal of Operational Research* 281.3 (2020). Featured Cluster: Business Analytics: Defining the field and identifying a research agenda, pp. 612–627. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2019.05.046.

[10] Q. H. Duong et al. “Striking the right balance: Customising return policy leniency for managing customer online return proclivity and satisfaction”. In: *Journal of Retailing and Consumer Services* 85 (2025), p. 104315. ISSN: 0969-6989. DOI: https://doi.org/10.1016/j.jretconser.2025.104315.

[11] Q. H. Duong et al. “Understanding and predicting online product return behavior: An interpretable machine learning approach”. In: *International Journal of Production Economics* 280 (2025), p. 109499. ISSN: 0925-5273. DOI: https://doi.org/10.1016/j.ijpe.2024.109499.

[12] Q. H. Duong et al. “Understanding product returns: A systematic literature review using machine learning and bibliometric analysis”. In: *International Journal of Production Economics* 243 (2022), p. 108340. ISSN: 0925-5273. DOI: https://doi.org/10.1016/j.ijpe.2021.108340.

[13] A. Dutt et al. “Partition-Based Clustering Algorithms Applied to Mixed Data for Educational Data Mining: A Survey From 1971 to 2024”. In: *IEEE Access* PP (Jan. 2024), pp. 1–1. DOI: 10.1109/ACCESS.2024.3496929.

[14] A. S. Eruguz et al. “Customer-to-customer returns logistics: Can it mitigate the negative impact of online returns?” In: *Omega* 128 (Oct. 2024), p. 103127. ISSN: 0305-0483. DOI: 10.1016/j.omega.2024.103127.

[15] R. Frei, L. Jack, and S. Brown. “Product returns: a growing problem for business, society and environment”. In: *International Journal of Operations & Production Management* 40.10 (June 2020), pp. 1613–1621. ISSN: 0144-3577. DOI: 10.1108/ijopm-02-2020-0083.

[16] R. Frei, L. Jack, and S. Krzyzaniak. “Mapping Product Returns Processes in Multichannel Retailing: Challenges and Opportunities”. In: *Sustainability* 14.3 (Jan. 2022), p. 1382. ISSN: 2071-1050. DOI: 10.3390/su14031382.

[17] R. Frei, S. Krzyzaniak, and L. Jack. “Sustainable Reverse Supply Chains for Retail Product Returns”. In: *Sustainable Development Goals and Sustainable Supply Chains in the Post-global Economy*. Cham: Springer International Publishing, 2019, pp. 159–182. ISBN: 978-3-030-15066-2. DOI: 10.1007/978-3-030-15066-2_9.

[18] J. Gäthke, K. Gelbrich, and S. Chen. “A Cross-National Service Strategy to Manage Product Returns: E-Tailers’ Return Policies and the Legitimating Role of the Institutional Environment”. In: *Journal of Service Research* 25.3 (2022), pp. 402–421. DOI: 10.1177/1094670521989440.

[19] B. Ghattas and A. Sanchez San-Benito. “Clustering Approaches for Mixed-Type Data: A Comparative Study”. In: *Journal of Probability and Statistics* 2025.1 (Jan. 2025). ISSN: 1687-9538. DOI: 10.1155/jpas/2242100.

[20] S. Gry et al. “A Conceptual Approach for an AI-Based Recommendation System for Handling Returns in Fashion E-Commerce”. In: *Smart Business Technologies*. Springer Nature Switzerland, 2024, pp. 1–23. ISBN: 978-3-031-67904-9. DOI: 10.1007/978-3-031-67904-9_1.

[21] A. Hofmann et al. “An Industry-Agnostic Approach for the Prediction of Return Shipments”. In: *Proceedings of the 26th Americas Conference on Information Systems, AMCIS 2020*. Aug. 2020. URL: https://www.researchgate.net/publication/342820571_An_Industry-Agnostic_Approach_for_the_Prediction_of_Return_Shipments.

[22] C. Hsu and Y. Chen. “Mining of mixed data with application to catalog marketing”. In: *Expert Systems with Applications* 32.1 (2007), pp. 12–23. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2005.11.017.

[23] H. Ibrahim Hayatu, A. Mohammed, and A. Barroon Isma’eel. “Big Data Clustering Techniques: Recent Advances and Survey”. In: *Machine Learning and Data Mining for Emerging Trend in Cyber Dynamics*. Springer International Publishing, 2021, pp. 57–79. ISBN: 9783030662882. DOI: 10.1007/978-3-030-66288-2_3.

[24] A. Al Imran and M. Nur Amin. “Predicting the Return of Orders in the E-Tail Industry Accompanying with Model Interpretation”. In: *Procedia Computer Science* 176 (2020). Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020, pp. 1170–1179. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2020.09.113.

[25] N. Janakiraman, H. A. Syrdal, and R. Freling. “The Effect of Return Policy Leniency on Consumer Purchase and Return Decisions: A Meta-analytic Review”. In: *Journal of Retailing* 92.2 (2016), pp. 226–235. ISSN: 0022-4359. DOI: https://doi.org/10.1016/j.jretai.2015.11.002.

[26] J. Ji et al. “An improved k-prototypes clustering algorithm for mixed numeric and categorical data”. In: *Neurocomputing* 120 (2013). Image Feature Detection and Description, pp. 590–596. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2013.04.011.

[27] D. Karl. “Forecasting e-commerce consumer returns: a systematic literature review”. In: *Management Review Quarterly* (May 2024). ISSN: 2198-1639. DOI: 10.1007/s11301-024-00436-x.

[28] J. Kranz, P. Urbanke, and L. M. Kolbe. “Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction”. In: *Proceedings of the International Conference on Information Systems (ICIS)*. Fort Worth, Texas, USA: Association for Information Systems (AIS), 2015. URL: https://dblp.org/rec/conf/icis/UrbankeKK15.

[29] A. Lepthien and M. Clement. “Shipping fee schedules and return behavior”. In: *Marketing Letters* 30.2 (Apr. 2019), pp. 151–165. ISSN: 1573-059X. DOI: 10.1007/s11002-019-09486-8.

[30] J. Magidson and J. Vermunt. “Latent class models”. In: Sage, Jan. 2004, pp. 175–198. URL: https://www.researchgate.net/publication/276323261_Latent_class_models.

[31] M. Makkonen, L. Frank, and T. Kemppainen. “The Effects of Consumer Demographics and Payment Method Preference on Product Return Frequency and Reasons in Online Shopping”. In: *34th Bled eConference Digital Support from Crisis to Progressive Change: Conference Proceedings*. University of Maribor Press, 2021, pp. 563–575. DOI: 10.18690/978-961-286-485-9.40.

[32] A. Mishra and P. Dutta. “Return management in e-commerce firms: A machine learning approach to predict product returns and examine variables influencing returns”. In: *Journal of Cleaner Production* 477 (Oct. 2024), p. 143802. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2024.143802.

[33] E. Molin, P. Mokhtarian, and M. Kroesen. “Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers”. In: *Transportation Research Part A: Policy and Practice* 83 (2016), pp. 14–29. ISSN: 0965-8564. DOI: https://doi.org/10.1016/j.tra.2015.11.001.

[34] D. A. Mollenkopf, R. Frankel, and I. Russo. “Creating value through returns management: Exploring the marketing–operations interface”. In: *Journal of Operations Management* 29.5 (2011). Special Issue on Field Research in Operations and Supply Chain Management, pp. 391–403. ISSN: 0272-6963. DOI: https://doi.org/10.1016/j.jom.2010.11.004.

[35] M. Niederlaender et al. “Garment Returns Prediction for AI-Based Processing and Waste Reduction in E-Commerce”. In: *Proceedings of the 16th International Conference on Agents and Artificial Intelligence*. SCITEPRESS - Science and Technology Publications, 2024, pp. 156–164. DOI: 10.5220/0012321300003636.

[36] M. Niederlaender et al. “Towards Waste Reduction in E-Commerce: A Comparative Analysis of Machine Learning Algorithms and Optimisation Techniques for Garment Returns Prediction with Feature Importance Evaluation”. In: *SN Computer Science* 6.5 (May 2025). ISSN: 2661-8907. DOI: 10.1007/s42979-025-03944-z.

[37] NRF. *Consumer Returns in the Retail Industry*. 2022. URL: https://nrf.com/research/2022-consumer-returns-retail-industry (visited on 09/08/2025).

[38] J. Ojadi et al. “Big Data Analytics and AI for Optimizing Supply Chain Sustainability and Reducing Greenhouse Gas Emissions in Logistics and Transportation”. In: *International Journal of Multidisciplinary Research and Growth Evaluation* 05 (Jan. 2024), pp. 1536–1548. DOI: 10.54660/.IJMRGE.2024.5.1.1536-1548.

[39] Forschungsgruppe Retourenmanagement. *Ergebnisse des Europäischen Retourentachos veröffentlicht*. 2022. URL: https://www.retourenforschung.de/info-ergebnisse-des-europaeischen-retourentachos-veroeffentlicht.html (visited on 09/08/2025).

[40] M. Rezaei, I. Cribben, and M. Samorani. “A clustering-based feature selection method for automatically generated relational attributes”. In: *Annals of Operations Research* 303.1–2 (Apr. 2018), pp. 233–263. ISSN: 1572-9338. DOI: 10.1007/s10479-018-2830-2.

[41] M. Samorani, F. Ahmed, and O. R. Zaïane. “An application to product returns”. In: *2016 IEEE International Conference on Big Data (Big Data)*. Dec. 2016, pp. 1454–1463. DOI: 10.1109/BigData.2016.7840753.

[42] G. Shang et al. “Using transactions data to improve consumer returns forecasting”. In: *J. Oper. Manag.* 66.3 (Dec. 2019), pp. 326–348. ISSN: 0272-6963. DOI: 10.1002/joom.1071.

[43] H. Snyder. “Literature review as a research methodology: An overview and guidelines”. In: *Journal of Business Research* 104 (2019), pp. 333–339. ISSN: 0148-2963. DOI: https://doi.org/10.1016/j.jbusres.2019.07.039.

[44] K. Snyder. *35 Top E-Commerce Statistics*. 2024. URL: https://www.forbes.com/advisor/business/ecommerce-statistics/ (visited on 09/08/2025).

[45] A. B. Stevenson and J. Rieck. “Investigating Returns Management across E-Commerce Sectors and Countries: Trends, Perspectives, and Future Research”. In: *Logistics* 8.3 (Aug. 2024), p. 82. ISSN: 2305-6290. DOI: 10.3390/logistics8030082.

[46] G. Walsh et al. “Preventive Product returns Management Systems-a Review and Model.” In: *ECIS*. 2014. URL: https://www.researchgate.net/publication/263083214_Preventive_Product_Returns_Management_Systems_A_Review_and_a_Model.

[47] J. Wang et al. “A bibliometric analysis of reverse logistics research (1992-2015) and opportunities for future research”. In: *International Journal of Physical Distribution & Logistics Management* 47.8 (Sept. 2017), pp. 666–687. ISSN: 0960-0035. DOI: 10.1108/IJPDLM-10-2016-0299.

[48] T. Wang et al. “How Supply Chain Complexity Drives Inventory Record Inaccuracy: Empirical Evidence from Cross-border E-commerce”. In: *SSRN Electronic Journal* (2022). ISSN: 1556-5068. DOI: 10.2139/ssrn.4275704.

[49] G. Yang, G. Ji, and K. H. Tan. “Impact of artificial intelligence adoption on online returns policies”. In: *Annals of Operations Research* 308.1–2 (Apr. 2020), pp. 703–726. ISSN: 1572-9338. DOI: 10.1007/s10479-020-03602-y.

[50] I. Zennaro et al. “Implementing E-Commerce from Logistic Perspective: Literature Review and Methodological Framework”. In: *Sustainability* 14.2 (2022). ISSN: 2071-1050. DOI: 10.3390/su14020911.

[51] D. Zhang et al. “Understanding fraudulent returns and mitigation strategies in multichannel retailing”. In: *Journal of Retailing and Consumer Services* 70 (2023), p. 103145. ISSN: 0969-6989. DOI: https://doi.org/10.1016/j.jretconser.2022.103145.

AI Statement In this research, AI tools were used for rewriting and editing texts, helping with writing and debugging code (Copilot), and generating spoken texts when injury limited my typing. Interpretation of results and formulation of conclusions were conducted independently by the author. All content was critically reviewed and approved by the author to ensure academic integrity and originality.

A

Scientific paper

Reducing Product Returns in Fashion & Electronics E-Commerce: A Clustering-Based Framework for Identifying High-Risk Orders and Products

Quirine Japikse*
*Delft University of Technology, Netherlands
(Dated: February 9, 2026)

Abstract The rapid growth of e-commerce has led to escalating product return volumes, generating substantial economic costs and environmental impact. Existing research largely focuses on either advanced predictive modelling or process optimisation, yet it commonly overlooks the role of stakeholders in designing actionable return reduction strategies. This study introduces a practical and interpretable framework that clusters products and orders features, finding the return risk to support targeted interventions. Using a Design Science Research (DSR) approach, the framework integrates expert interviews, a structured literature review, and extensive data analysis. Statistical tests reveal significant return rate variation across product attributes (category, color, size, price) and order characteristics (order value, quantity, shipping carrier). Three clustering techniques, K-Prototypes, CAVE, and LCC, were evaluated, with LCC providing the most distinctive segmentation of high-risk groups. High-risk products were predominantly found in the Fashion Category, particularly in black and red variants and in sizes other than "One Size". Furthermore, high-risk orders were associated with large, high-value purchases shipped via DHLDE. The resulting framework enables retailers to implement validated strategies such as improved product information, category-specific return policies, and targeted marketing adjustments. Overall, the approach offers a scalable, stakeholder-aligned foundation for reducing return flows.

Keywords: Return reduction, Clustering analysis, Design Science Research

I. INTRODUCTION

The rapid expansion of e-commerce has fundamentally transformed global retail markets, accompanied by a substantial surge in consumer returns. Online retail has continued to grow at double-digit rates in recent years, making returns an increasingly critical area of concern for both researchers and industry practitioners [1]. Returns now constitute a dominant component of retail logistics, positioning them as a central topic within operations and supply chain management research [2, 3]. Germany provides one of the clearest examples of this development. In 2022, 24% of all parcels, corresponding to 530 million packages and 1.3 billion individual items, were returned [4]. The processing of returns in Germany alone generated approximately 795.000 tonnes of CO2 emissions in 2021, equivalent to an estimated 5.3 billion kilometers driven by car [5]. Financially, returns are equally costly: per parcel processing and transportation expenses average €6.95, generating an additional €3.68 billion in costs for German companies in 2022 alone [4].

Operationally, the handling of returned products remains labor-intensive and prone to inefficiencies. Returned items may be lost during transit, misrouted inside warehouses, or overlooked in inventory systems, often resulting in unnecessary disposal [6–8]. Even when products are successfully reintegrated into the supply chain, inspection, reprocessing, and restocking require considerable manual effort, limiting scalability and driving up costs. These challenges have prompted growing interest in predictive analytics and data-driven decision-support systems. Research shows that even relatively simple statistical methods, such as binary logistic regression, can effectively forecast return probabilities, enabling retailers to implement proactive mitigation strategies [9].

A wide range of behavioral, commercial, and logistical factors contribute to high return rates. Impulse buying driven by marketing tactics, such as flash sales and heavy discounting, increases the likelihood of returns, while shipping-related policies like free-shipping thresholds and minimum-order requirements encourage consumers to intentionally over-order with the expectation of returning part of the purchase [5, 10]. Product attributes, such as pricing and review profiles, further influence return behavior: unbiased reviews are associated with lower return rates, while biased reviews and lower-priced items tend to produce higher return frequencies. Demographic variables, including age and gender, also play a role [11]. Payment methods, particularly invoice-based or "after pay" systems, reduce customers' perceived financial commitment at checkout, thereby increasing the likelihood of returns [5].

Return policies form another crucial mechanism shaping both consumer behavior and operational outcomes. Liberal return policies tend to increase customer trust, basket size, and purchase frequency, but inevitably lead to higher return volumes. Conversely, strict policies reduce return rates yet risk damaging customer satisfaction and long-term profitability. Retailers therefore face the well-documented "return policy leniency dilemma", wherein they must balance operational efficiency with customer experience [12]. Finally, return logistics involves multiple stakeholders. Customers, retailers, warehouse operators, IT and data teams, sustainability units, strategic managers, and postal carriers all shape and experience return-related challenges. Addressing the return issue, therefore, requires solutions that integrate behavioral insights, operational feasibility, technological capabilities, and environmental objectives.

A. Related work

Research on improvement in return logistics in e-commerce spans two main streams: data-driven analytical approaches and process-oriented improvement strategies. A substantial part of the literature focuses on using machine learning (ML) and AI-based methods to predict return behavior and optimize reverse logistics operations. Prior studies demonstrate that ML models, such as Random Forests, Support Vector Machines, Neural Networks, and Gradient Boosting, can forecast return likelihood, volumes, and reasons with increasing accuracy [13]. These models typically use product attributes, customer behavior patterns, and order characteristics as predictors. Recent works show that data-driven forecasting enables more efficient resource allocation, proactive staffing, better inventory planning, and scenario-based return policy testing [5, 7, 14]. Beyond forecasting, conceptual frameworks such as AI-based recommendation engines and real-time decision platforms demonstrate the potential to route returns to appropriate resale, repair, or recycling channels [5]. However, these solutions prioritize algorithmic performance and pay limited attention to real-world integration challenges, including data quality, system interoperability, and the operational constraints of e-commerce organizations. Moreover, studies rarely consider stakeholder involvement during design and implementation, limiting the practical adoption and interpretability of AI-driven tools.

Complementing analytics-focused studies, a second stream emphasizes process mapping, lean management, and operational interventions. This work highlights inefficiencies within reverse logistics flows, such as transport waste, redundant handling, and poorly integrated warehouse activities, and proposes measures to eliminate them. Strategies include standardizing return codes, improving packaging quality, enhancing communication flows, and shortening inspection and resale cycles [3]. Recent innovation efforts also explore customer-centric interventions, such as customer-to-customer returns, immediate return options, personalized return policies, and virtual try-on technologies. These approaches aim to reduce unnecessary returns by influencing customer decision-making. Yet, many of these solutions face scalability challenges, rely heavily on consumer adoption, or apply only to specific product categories [15].

Across both research streams, significant progress has been made in forecasting, efficiency, and sustainability. However, the literature consistently reveals a critical gap: existing models and interventions are rarely designed with organizational realities, stakeholder needs, or practical adoption in mind. Predictive models often function as "black boxes" with limited interpretability for practitioners, while process-oriented strategies frequently lack data-driven insight into which products or orders pose the greatest return risk. This study aims to design a comprehensive, stakeholder-informed, and interpretable framework that unites data-driven analysis with practical intervention design. The present study addresses this by the following question: *How can a reduction of returned goods in fashion- & electronic e-commerce be achieved by designing a framework that clusters high- and low-risk orders and products?*

II. METHODOLOGY AND DATA

This study is done using the Design Science Research method. DSR aims to generate knowledge of how things can and should be constructed or arranged to achieve a desired set of goals. There are three different categories within this framework: environment, design, and knowledge base.

A. Literature review

The literature review is designed to examine three core aspects relevant to this study. The sub-questions listed below guide this exploration and are addressed through a systematic analysis of existing work. This approach supports a deeper understanding of the problem context (environment) and provides a foundation for formulating the requirements and objectives of the research framework.

1. Which factors are shown to be associated with returns?
2. How is return management organized currently?
3. What strategies exist in the literature to reduce returns?

B. Interviews

The interviews conducted in this research serve a dual purpose: they aim to explain current organisational practices in return logistics and the use of data-driven or model-based solutions. Additionally, the interviews will be used to find practical implications of these practices within operational environments. Moreover, interviews will be conducted to validate proposed strategies through direct engagement with relevant stakeholders. This ensures that the resulting framework is both context-specific and practically grounded. To this end, interviews are held with supply chain and retail experts to address questions concerning existing return-handling processes, data collection practices, and the utilization of analytical tools within organisations. In a subsequent phase, additional interviews are conducted to assess and refine strategies intended to reduce return rates, ensuring alignment with practitioner experience and feasibility constraints.

A semi-structured interview format is employed, using a flexible guide composed of open-ended questions to facilitate the collection of rich, in-depth qualitative data. This methodological approach balances the systematic exploration of predetermined themes with the adaptability needed to pursue insights, probe relevant follow-up topics, and adjust the sequencing of questions where appropriate. Such flexibility is particularly well-suited to studying the complexities of return logistics processes, the organisational adoption of analytical models, and the broader implications of integrating data-driven decision-support tools into operational workflows.

C. Data analysis

This study relies on high-quality operational data to develop a robust clustering approach, focusing exclusively on order-level and product-level variables to avoid bias and ensure actionable insights. Proprietary sales-order and return-order datasets were extracted from the retailer's Finance & Operations system, containing information such as shipment characteristics, order value, item quantities, and detailed product attributes (e.g., price, size, color, category). Sales and return records were merged using a unique webshopreference, and a binary return indicator was created for both order- and product-level datasets. Data preparation included removal of non-essential identifiers and duplicates, standardization of timestamps and categorical fields, treatment of missing values (with "unknown" imputed where appropriate), and evaluation of outliers. High-cardinality attributes such as product category, size, and color were aggregated into broader groups to reduce sparsity and improve model performance. The resulting cleaned datasets, covering six months of transactions exported on 3 October 2025, form the basis for the clustering analysis.

1. Cluster analysis

Clustering is an unsupervised machine learning technique that identifies natural groupings in data by grouping observations with similar characteristics. In the context of product returns, clustering enables the detection of high- and low-risk segments without relying on labeled data, making it well-suited for exploring complex return patterns. By calculating return rates per cluster, high-risk groups can be isolated and examined to identify the feature combinations that drive elevated return likelihood, thereby supporting the development of targeted return reduction strategies. Modern clustering algorithms can efficiently handle both numerical and categorical variables while remaining computationally scalable [16], making them a practical tool for uncovering hidden structures in large e-commerce datasets and informing data-driven intervention design. Three different methods are explored.

a. K-Prototypes K-prototypes is a clustering method specifically designed for datasets containing both numerical and categorical variables. It integrates the principles of k-means and k-modes by combining Euclidean distance for numerical attributes with a simple matching dissimilarity measure for categorical attributes. This hybrid distance function enables the algorithm to form clusters that reflect similarity across mixed data types. Numerical cluster prototypes are represented by feature means, while categorical prototypes are defined by feature modes, allowing the method to capture central tendencies in both domains. Owing to this formulation, k-prototypes provides an efficient and interpretable approach for uncovering structure in large, heterogeneous datasets, making it well-suited for applications such as identifying return risk patterns in e-commerce.

b. CAVE The CAVE algorithm (Clustering Algorithm based on Variance and Entropy) is designed to improve clustering performance for mixed-type data by explicitly weighting numerical and categorical features according to their discriminative power. Unlike k-prototypes, CAVE balances feature contributions using two principles: variance and entropy. Numerical features with higher variance receive greater weight, reflecting their stronger ability to distinguish between clusters. Categorical features are weighted inversely to their entropy; categories with low entropy (i.e., highly concentrated distributions) are assigned higher importance because they more clearly differentiate cluster membership, while high-entropy features contribute less. This dual weighting framework results in clusters that are both numerically compact and categorically homogeneous, improving interpretability and reducing bias toward dominant variable types. Prior research shows that CAVE yields more accurate and meaningful cluster structures than traditional distance-based approaches for mixed data [17–19].

c. Latent Class Clustering LCC is a model-based probabilistic approach for identifying unobserved subgroups in mixed-type datasets. Unlike distance-based methods such as k-prototypes or CAVE, which assign each observation deterministically to the nearest cluster, LCC assumes that the associations among observed variables are explained by an underlying discrete latent variable. Each latent class represents a probabilistically distinct group, and observations belong to classes with estimated membership probabilities rather than fixed assignments [20, 21]. This formulation enables LCC to capture complex variable dependencies, making it particularly effective in settings where categorical variables dominate or cluster boundaries overlap. These are conditions under which distance-based methods often struggle [22]. Overall, LCC provides a flexible and statistically principled framework for uncovering hidden structure in heterogeneous datasets, offering advantages in accuracy and interpretability over deterministic clustering approaches.

III. LITERATURE REVIEW

To address the reduction of product returns, three key domains require investigation. First, existing literature on factors associated with return behaviour provides essential theoretical grounding. Second, an examination of current return-logistics practices offers insight into operational processes and constraints. Finally, established return reduction strategies are reviewed to identify effective mechanisms on which new, data-driven interventions can be developed.

A. Associated features

Prior research identifies a broad spectrum of features associated with product returns in e-commerce. Karl's (2024) systematic review provides an overview of variables frequently used in return-prediction models and synthesizes insights from 25 empirical studies [13]. These features span product-level, customer-level, and order-level.

1. Product attributes

Product attributes consistently emerge as strong determinants of return likelihood. Price is one of the most recurring predictors, with several studies showing that higher prices and price promotions can increase the probability of returns, often linked to impulse-driven purchases [9, 23]. Product type and category also play a significant role: some categories inherently exhibit higher return rates, and return histories of products or categories can be used as signals of future return risk [14, 24, 25]. Other product-related factors include brand perception, size, colour, and product-specific order histories, all of which have been shown to influence customer decisions to keep or return items [26].

2. Customer attributes

Customer attributes are highlighted in multiple studies, particularly demographics such as gender and age, which have been found to correlate with return frequency [11]. Customer return history is frequently identified as a strong predictor, often outperforming individual product features or transactional factors [13]. Prior research shows that customers with higher historical return behaviour exhibit a higher future propensity to return, making behavioural profiles valuable for risk segmentation [26, 27].

3. Order attributes

Basket and transaction attributes also contribute significantly to return risk. Basket size, item count, and the degree of product similarity within an order are associated with elevated returns, particularly in cases of bracketing behaviour [9, 27, 28]. Payment method is another influential variable: orders paid by invoice, cash-on-delivery, or deferred payment methods ("after pay") are repeatedly shown to have higher return rates than prepaid orders [11, 23]. Additionally, total order value, order date, and shipment characteristics (e.g., shipping carrier, delivery mode) appear as relevant predictors in multiple studies [14]. While the literature highlights a wide variety of influential features, it also reveals a lack of consensus on the optimal combination of variables for identifying return risks.

B. Return logistics management

Return logistics management is a critical operational domain for e-commerce retailers, as returned products often incur substantial handling and transportation costs. The literature broadly distinguishes between preventive and curative dimensions of return management. The curative dimension encompasses the effective processing of returned items. This often involves attention to maximizing value recovery through resale, refurbishment, material extraction, or donation [29].

A return flow typically comprises a sequence of activities initiated when a customer sends back a product. Effective management requires product retrieval, inspection, recovery (e.g., repair, recycling), inventory management, waste management, and reintegration into the forward supply chain [29]. Case studies synthesized by Frei et al. (2022) illustrate a generic process beginning at purchase and spanning multiple return entry points, such as parcel carriers, postal services, or drop-off locations. Returned products may subsequently follow diverse exit routes, including resale, recycling, donation, manufacturer return, or disposal. This depends on the condition and the residual value. A retailer often decides whether returns are processed for speed or value maximization [6]. Return disposition encompasses five primary activities: destroying, recycling, refurbishing, remanufacturing, and repackaging. Each of these activities has different implications for economic performance, service quality, and operational responsiveness [30]. Frei et al. (2022) further identify seven sources of waste within return logistics: over-processing, excessive inventory costs, unnecessary transportation, avoidable personnel movement, process delays, defects requiring rework, and inefficient storage utilization.

This study broadened the framework by adding stakeholder responsibility and level of automation in the return processing actions, presented in Figure 1. The key challenge found in the current return handling process appears to be the high amount of returns, with a low level of automation. This makes the process labor-intensive and costly.

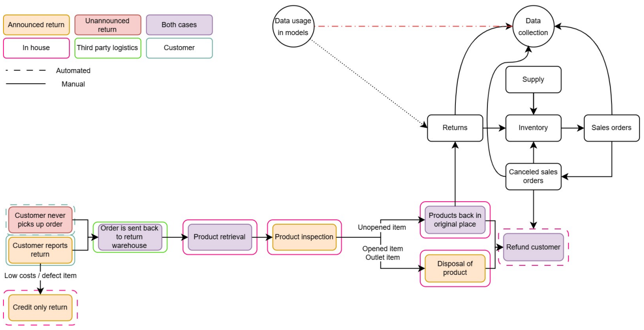


FIG. 1. Return logistics process

C. Reduction strategies

Research on return reduction strategies spans product-related, policy-related, operational, and marketing domains. While most studies aim to address the growing operational burdens of product returns in closed-loop supply chains, recent work suggests that prioritizing efficiency over prevention has limited the effectiveness of return reduction initiatives. A deeper understanding of the behavioural and operational drivers of returns is therefore essential for designing effective interventions [31, 32].

Several studies emphasize strategies targeting product attributes and customer information. Packaging quality emerges as a critical extrinsic factor influencing returns, particularly for electronics, home appliances, and perishables. Improved protective packaging and careful handling can reduce damage-related returns [32]. In fashion, where fit uncertainty is high, customer service quality is shown to significantly reduce returns by addressing ambiguity before purchase. Clear and detailed product information is shown to reduce mismatch-driven returns. Advanced tools, such as avatars and VR, are not yet widely adopted due to technical constraints; however, lower-cost information tools are also consistently associated with reduced return likelihood [33]. Lastly, emphasizing product reliability and usage performance in customer feedback also helps align expectations with reality, particularly for electronics and home appliances [32].

Return policy design plays a central role in shaping both purchasing behaviour and return rates. Lenient policies may increase conversion and customer satisfaction, but also stimulate excessive or fraudulent returns [32]. Stricter policies can effectively reduce return volumes, especially for electronics. Cultural differences further moderate policy effects. Effort-based restrictions reduce returns in Western markets but are less effective in Eastern markets, where customer-oriented norms can even increase return tendencies [34].

A broader body of work analyses the five dimensions of return policy leniency: monetary, time, effort, scope, and exchange. Adjusting these levers can influence both return incidence and purchase intent [35]. For example, exchange leniency reduces returns more effectively than restrictive exchange policies, and restricting returns on sale items can curb high-risk return patterns in price-sensitive segments. More recent work adds nuance by warning that over-restrictive policies may alienate legitimate customers. Targeted enforcement can mitigate abuse without compromising loyalty [36]. Procedural instruments such as increased "hassle costs", limited return channels, and reduced transparency have also been explored as non-monetary ways to discourage returns, though their acceptance varies across markets and retailers [33].

Marketing strategies influence return behaviour through their effect on customer decision processes. Limited-time promotions and other high-pressure tactics can lead to rushed decisions and elevated returns [37]. Aligning marketing and operations functions can help prevent mismatches that lead to unnecessary returns [38]. For example, ensuring that promotions are suitable for the retailer's assortment and logistical capabilities can help reduce returns.

Overall, the literature offers a diverse set of strategies to reduce returns, ranging from improved product information and packaging to policy design, operational adjustments, and marketing interventions.

IV. RESULTS

A. Framework

This study developed a framework combining a data-driven analytical approach with active stakeholder involvement to evaluate its practical relevance. The framework is designed to prioritise interpretability, effective utilisation of available data, compatibility with mixed data types, and scalability for large e-commerce environments. By integrating clustering analysis with iterative stakeholder validation, the framework provides a systematic method for identifying high- and low-risk products and orders, thereby enabling the formulation of targeted, evidence-based strategies to reduce return volumes. The applied framework, visualised in Figure 2, forms the basis for the case study presented in this research.

B. Associated features

Statistical analyses reveal that returns are driven by a combination of order-level and product-level attributes, with several variables showing statistically significant associations with return likelihood. Numerical variables show statistically significant correlations with being returned. Even though the correlations are small, their sig-

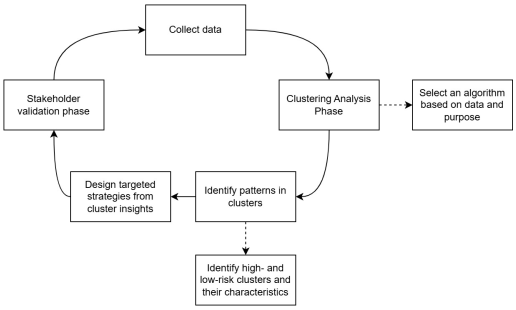


FIG. 2. Data-driven, stakeholder embedded return reduction framework

nificance indicates the influence of these variables. Both order amount ($r = 0.053, p < 0.001$) and the number of items ordered ($r = 0.055, p < 0.001$) show positive associations with returns, implying that larger and higher-value baskets are marginally more likely to be returned. Product price has a positive correlation with returns ($r = 0.0129, p < 0.001$), indicating that higher-priced items are slightly more likely to be returned. In contrast, product quantity shows a weak negative correlation ($r = –0.0086, p < 0.001$), suggesting that items purchased in larger quantities tend to have lower return rates (Table I). This finding diverges from expectations surrounding bracketing behavior, which would typically predict higher return rates for multi-item purchases. Order hour, when treated as a continuous variable, does not significantly correlate with returns. However, when binned into groups that show the time of day, it shows a significant difference in return rates across groups. Furthermore, chi-square tests indicate that delivery mode, shipping carrier, shipment method, product category, size, and color have significant differences in return rates in their categories. Their chi-square and p-values are presented in Table II.

Overall, the combined statistical evidence demonstrates that return outcomes are influenced by both product- and order-specific characteristics. While individual correlations are small in magnitude, chi-square tests consistently confirm strong group-level differences across categorical variables. Together, these findings indicate that returns are influenced by multiple features and best understood through the interaction of intrinsic product attributes, pricing considerations, and logistical conditions. This provides a robust foundation for clustering analysis and targeted intervention design.

C. Clustering analysis

1. *Comparing clustering methods*

Across all clustering approaches, distinct differences emerged in the ability to separate high- and low-risk product segments. Both K-Prototypes and CAVE gen-

TABLE I. Significance of Correlations

Variable	Correlation (r)	p-value	Significance
Product price (€)	0.0129	0.000	***
Product quantity	-0.0086	0.000	***
Items ordered	0.055	0.0000	***
Order amount (€)	0.053	0.000	***
Order hour	0.001	0.6041	

*Significance levels: *** p<0.001, ** p<0.01, * p<0.05*
The variables are tested for correlation vs Returned

TABLE II. Chi-Square Test Results

Variable	Chi-Square	p-value	df
Delivery mode***	85508.93	0.0000	15
Shipping Carrier***	84787.10	0.0000	8
Shipping method***	80960.19	0.0000	4
Product Size***	33426.37	0.0000	7
Product Category***	31084.87	0.0000	12
Product Color***	8034.43	0.0000	24
Order hour (Binned)***	315.42	0.0000	3

*Significance levels: *** p<0.001, ** p<0.01, * p<0.05*

erated clusters that were primarily driven by numerical attributes, particularly product price, resulting in moderately differentiated return rate profiles. In these methods, clusters were dominated by broad price tiers, with substantial overlap in product categories and colors. This limited their ability to isolate specific high-risk product groups, as most clusters fell into medium-risk ranges. In contrast, Latent Class Clustering (LCC) produced the clearest segmentation. LCC identified a distinct high-risk product cluster (Cluster 2) with a return rate of 18.63%, substantially above the baseline. This cluster was almost entirely composed of products from Category F, especially in black and red, and represented by defined sizes (S–XL) rather than “One Size” or unknown sizes. The remaining clusters were more clearly grouped into low- and medium-risk segments, dominated by Categories H, L, I, and E, often containing low-priced items with unknown or non-distinctive colors and sizes. Thus, LCC provided the strongest categorical structure and revealed the most actionable high-risk grouping. The summary of these clustering methods is represented in Table III.

2. *In-depth cluster analysis*

a. Cluster 2 A detailed decomposition of LCC’s Cluster 2 showed substantial internal heterogeneity, motivating a sub-cluster analysis. Category F dominates nearly all observations in all clusters. Across the five subclusters, two emerged as exceptionally high-risk (Subclusters 2.1 and 2.4, with return rates over 20%. Table IV shows an overview of the subclusters. It shows that the relatively high-priced items in cluster 2 increase the likelihood of being returned. These are price levels between approx. €8–12 per unit. Furthermore, Black and red are the most prevalent colors in the high-risk

TABLE III. Summary of clusters for all three clustering methods.

Method	Cluster	Observations	% of total	Return Rate	Description	Risk
K-Proto	0	127258	8.1%	6.5%	Mid-range prices; black items; categories L and I.	Medium -
	1	35310	2.3%	7.2%	High-priced items; categories L and I; mostly black.	Medium +
	2	1000487	63.8%	6.2%	Low-priced items; categories H and L; color unknown.	Medium -
	3	404729	25.8%	8.4%	Moderately priced; categories L and F; black items.	High
CAVE	0	547884	34.9%	7.45%	Middle priced; category L; black; size unknown.	Medium +
	1	749463	47.8%	6.04%	Low priced; category H; color and size unknown.	Medium -
	2	223976	14.3%	7.96%	Medium–high priced; category L; black or unknown colors; size unknown.	High
	3	46461	3.0%	7.06%	High priced; mostly category L; black; size unknown.	Medium +
LCC	0	201866	12.9%	6.4%	Medium–high priced; categories I and E; beige; sizes unknown.	Medium -
	1	302270	19.3%	6.2%	Low priced; category B; black; sizes unknown.	Medium -
	2	130048	8.3%	18.6%	Medium priced; category F; black or red; various sizes.	High
	3	369593	23.6%	5.7%	High priced; category L; black, blue, and pink.	Low
	4	564007	36.0%	5.4%	Low priced; category H; color and size unknown.	Low

Overall return rate: 6.84%

TABLE IV. Summary of clusters of Cluster 2

Cl.	Obs.	% of total	RR	Description	Risk
2.0	24054	18.5%	13.84%	Low priced, 1 item, black, One Size	Low
2.1	**21202**	**16.3%**	**21.31%**	**High priced, 1,5 item, Black & white, M & XL**	**High**
2.2	20202	15.5%	15.51%	Low priced, 1 item, Black & white, S & One Size	Low
2.3	20178	15.5%	18.04%	Average priced, Black & Red, S&L, 1 item	Average
2.4	**44412**	**34.2%**	**21.63%**	**High priced, 1 item, Black & Red, S & L**	**High**

Note: Overall return rate: 18.6%

clusters. Lastly, size was observed to elevate risk in the overarching clusters, but when zooming in on cluster 2, it shows that all sizes different than "One Size" increase risk (The low-risk sub-clusters show "One Size" as prevalent). This segmentation indicates that within the same high-risk cluster, price, color, and especially size serve as additional risk multipliers.

b. Order level features Following the clustering of product-level attributes, order-level characteristics were incorporated to obtain a more comprehensive understanding of return risk patterns. The results indicate that Cluster 2, the high-risk product cluster, is predominantly associated with large orders and high order values. This suggests that high-risk products are frequently purchased within substantial, high-value baskets. Although PostNL accounts for the largest share of shipments in this cluster, the elevated return rate appears to be driven primarily by DHLDE, which consistently contributes to an increased likelihood of returns. While Clusters 1 and 3 also contain a notable proportion of high-value orders, their return rates are only moderately above average and correspond more closely with mid-range order values, rather than displaying the concentrated risk observed in Cluster 2. An overview of the order-level characteristics of the LCC clusters is presented in Table V. The order-size categories applied are: small (1–2 items), medium (3–4 items), and large (5+ items). The order-value ranges are defined as follows: Low (€0.01–€49.94), Medium– (€49.94–€72.96), Medium+ (€72.96–€111.35), and High (€111.35+).

TABLE V. Order level features

Cluster	Order Size	Value	Delivery mode	RR
0	Medium	Mid +	PostNL-Stan	6.4%
1	Large	High	PostNL-Stan	6.2%
2	Large	High	PostNL-Stan	18.6%
3	Medium	High	PostNL-Stan	5.7%
4	Medium	Low	PostNL-Stan	5.4%

3. High and low risk characteristics

Across order- and product-level analyses, a consistent pattern emerges in which return risk is shaped by the interaction of order characteristics, product attributes, and logistical factors. At the order level, smaller and lower-value orders exhibit substantially lower return rates, whereas large, high-value orders are repeatedly associated with increased risk. Although the distribution of purchase time is similar across all clusters, evening and night purchases consistently show a higher return likelihood. Carrier effects are pronounced: INPOST and COLISSIMO appear as reliable low-risk carriers, whereas

TABLE VI. Summary high & low risk characteristics

High Risk	**Low Risk**
Product level features	
Cat. F	Cat. H
Medium priced (7.35 ± 4.96)	Low / High priced (3.49 (±5.88); 13.48 (±14.75))
Black & Red	Unknown
Sizes other than “One Size”	Unknown / “One Size”
Order level features	
Large orders (5+)	Small & Medium orders (1–2; 3–4)
DHLDE	COLISSIMO, INPOS
Night & Evening	Morning & Afternoon
PostNL-PU (increases risk in average/low risk clusters)	PostNL-Stan
Increased risk in Cluster 2 product level LCC	
Black & Red	White, Pink, Purple & Multi
Highest prices among cluster 2 (8.37 (±5.18); 12.26 (±6.10))	Lower prices: 4.17 (±0.80)–6.20 (±0.95)

DHLDE is consistently represented in high-risk clusters. PostNL-Standard performs more favorably than PostNL-Pickup, but PostNL is positioned generally as a lower-risk option. At the product level, both K-Prototypes and CAVE indicate elevated return rates for moderately priced items in Categories L and F. LCC provides a clearer segmentation, isolating a high-risk cluster primarily consisting of Category F products, predominantly in black and red, and priced at intermediate levels. Size-related patterns are particularly salient: items offered in defined sizes (S–XL) exhibit markedly higher return rates than “One Size” products, a finding reinforced by subcluster analysis within the LCC high-risk group. When order-level features are mapped onto product clusters, the high-risk product cluster aligns predominantly with large, high-value orders, whereas lower-risk clusters include high-value orders with fewer items, suggesting that order composition moderates risk differently across product types. These findings are summarized in Table VI.

D. Strategies

The combined statistical and clustering results point to several targeted strategies:

a. Product Information & Fit Because size-dependent items drive returns, especially in Category F, improving size guidance, providing model-specific measurements, and integrating fit-feedback from previous customers could reduce mismatches. For items where feasible, simplifying size variations (e.g., offering “One Size” alternatives) may also reduce fit-driven returns.

b. Differentiated Return Policies The findings support category-specific and risk-specific return policies. For high-risk items (Category F, sized S–XL, black/red), stricter return windows, reduced refund flexibility, or selective restocking fees may discourage over-ordering. Lower-risk items can retain more lenient policies to preserve customer satisfaction.

c. Operational Interventions Evening/night purchases exhibit above-average risk; although retailers cannot control purchase timing, they may introduce additional prompts or friction-reducing information during late-night shopping sessions.

d. Marketing & Sales Adjustments High-risk products should not be subject to heavy discounting or limited-time promotions, as these tactics increase impulsive purchases and subsequent returns. Instead, promotional focus should shift to low-risk product segments. For large orders, identified as a major return driver, retailers may implement confirmation prompts or discourage excessive ordering through bundle-based incentives that reward keeping all items.

V. CONCLUSION AND DISCUSSION

This study provides a systematic examination of return risk patterns in e-commerce using a clustering approach, demonstrating how product- and order-level data can be translated into actionable strategies. Three clustering algorithms: K-Prototypes, CAVE, and Latent Class Clustering (LCC), were evaluated, revealing clear methodological differences with substantial implications for return risk identification.

Across methods, distance-based algorithms (K-Prototypes and CAVE) primarily separated products on the basis of numerical features such as price, yielding clusters with limited variation in return risk. This aligns with prior observations that these methods often struggle when categorical variables dominate or when cluster boundaries are not spherical. Their strong reliance on price and limited ability to separate categorical patterns resulted in medium-risk clusters that lacked clear interpretability or decision relevance. By contrast, LCC produced the most meaningful segmentation, isolating a single, distinct high-risk cluster (Cluster 2) with an 18.63% return rate. This cluster captured a confluence of risk-intensive attributes highlighted in earlier literature: category-specific effects (Category: Fashion), fit-sensitive sizes (S–XL), and prominent color patterns (black and red). The subcluster analysis showed that even within a single high-risk category, price, color, and size amplify return risk, reinforcing the multifactorial nature of return behavior. This supports the theoretical understanding that product returns arise from the interaction of multiple attributes rather than single drivers. Integrating order-level information revealed equally important contextual factors. High-risk products disproportionately appeared in large, high-value orders, which have been linked to both bracketing behavior and increased customer uncertainty. Carrier effects were also pronounced: DHLDE emerged consistently as a high-risk carrier, while

COLISSIMO and INPOST were strongly associated with low-risk clusters. These findings echo prior work identifying logistics quality and carrier reliability as influential determinants of return outcomes. Time of purchase showed consistent but non-deterministic effects. Night-time purchases exhibited higher return rates, though this did not influence cluster segmentation. Collectively, these results confirm that return behavior is multifactorial, shaped by the interaction of intrinsic product attributes, order context, and logistical conditions. The clustering-based framework demonstrates value by revealing how these factors co-occur within meaningful segments. Importantly, the expert validation process highlighted practical constraints that influence strategy adoption. While improvements in product information and category-specific return policies were deemed highly feasible, operational interventions (e.g., altering carrier distribution or restricting ordering hours) require more nuanced, country-specific considerations. Experts also emphasized the need to balance return reduction with customer satisfaction and operational capacity, indicating that strategy implementation must be iterative and context dependent. These findings strengthen the case for stakeholder-aligned, interpretable analytical tools, addressing a long-standing gap in return logistics research where models often lack practical integration. The framework developed here, therefore, contributes not only methodologically but also practically, enabling risk-differentiated decision-making across product design, logistics, and marketing.

However, several limitations shape the interpretation and generalisability of the findings. A first set of constraints arises from data availability. The dataset lacked demographic variables and customer-reported return reasons, preventing deeper behavioral insights and limiting the analysis to product- and order-level characteristics. Although this aligns with fairness and practicality considerations in industry settings, it excludes relevant explanatory factors identified in prior research. Moreover, the dataset covered a five-month period, insufficient to capture seasonal effects. These constraints may have led to an underestimation of returns and reduced model robustness. Second, limitations relate to the clustering methods. K-Prototypes and CAVE produced unbalanced clusters driven largely by numerical variables, limiting their ability to isolate meaningful categorical patterns. Although the CAVE algorithm theoretically weights categorical attributes more effectively, numerical dominance persisted in practice. LCC provided the clearest segmentation but required substantial computational resources and long runtimes, constraining its practical deployment. Across methods, the risk of overfitting remains, given the high number of variables and the use of data from a single retailer. Third, practical limitations concern the absence of qualitative insights. Customer motivations, satisfaction, and experiences with product fit, delivery, or return processes were outside the scope of this study. As a result, the analysis reveals what is returned and under which conditions, but not why. This limits the completeness of strategy development and may reduce alignment between proposed interventions and customer expectations. Finally, the findings are based on data from one European retailer. External validation was not conducted, limiting generalisability across sectors, markets, and supply chain contexts. Testing the framework across different industries, regions and retailers is necessary to confirm its broader relevance.

Future research should therefore focus on integrating return reasons, extending observation periods, combining qualitative and quantitative data, and validating the framework across multiple retailers. In addition, examining country-specific risks could further explain why specific carriers increase or decrease return risk. Lastly, the long-term impact of the strategies should be evaluated. It is important to test the consequences of the strategies for return volumes, sales volumes, customer satisfaction, and environmental outcomes.

[1] NRF, Consumer returns in the retail industry (2022).
[2] H. Abdulla, M. Ketzenberg, and J. D. Abbey, Taking stock of consumer returns: A review and classification of the literature, Journal of Operations Management **65**, 560–605 (2019).
[3] R. Frei, L. Jack, and S. Brown, Product returns: a growing problem for business, society and environment, International Journal of Operations & Production Management **40**, 1613–1621 (2020).
[4] F. Retourenmanagement, Ergebnisse des europäischen retourentachos veröffentlicht (2022).
[5] S. Gry, M. Niederlaender, A. N. Lodi, M. Mutz, and D. Werth, A conceptual approach for an ai-based recommendation system for handling returns in fashion e-commerce, in *Smart Business Technologies* (Springer Nature Switzerland, 2024) pp. 1–23.
[6] R. Frei, L. Jack, and S. Krzyzaniak, Mapping product returns processes in multichannel retailing: Challenges and opportunities, Sustainability **14**, 1382 (2022).
[7] M. Niederlaender, A. Lodi, S. Gry, R. Biswas, and D. Werth, Garment returns prediction for ai-based processing and waste reduction in e-commerce, in *Proceedings of the 16th International Conference on Agents and Artificial Intelligence* (SCITEPRESS - Science and Technology Publications, 2024) p. 156–164.
[8] B. Asdecker and D. Karl, Shedding some light on the reverse part of e-commerce: A systematic look into the black box of consumer returns in germany, European Journal of Management **22**, 59–81 (2022).
[9] B. Asdecker and D. Karl, Big data analytics in returns management – are complex techniques necessary to forecast consumer returns properly? (Universitat Politècnica València, 2018).
[10] A. Lepthien and M. Clement, Shipping fee schedules and

return behavior, Marketing Letters **30**, 151–165 (2019).

[11] M. Makkonen, L. Frank, and T. Kemppainen, The effects of consumer demographics and payment method preference on product return frequency and reasons in online shopping, in *34th Bled eConference Digital Support from Crisis to Progressive Change: Conference Proceedings* (University of Maribor Press, 2021) p. 563–575.

[12] Q. H. Duong, L. Zhou, M. Meng, L. T. A. Dang, and T. D. Nguyen, Striking the right balance: Customising return policy leniency for managing customer online return proclivity and satisfaction, Journal of Retailing and Consumer Services **85**, 104315 (2025).

[13] D. Karl, Forecasting e-commerce consumer returns: a systematic literature review, Management Review Quarterly 10.1007/s11301-024-00436-x (2024).

[14] A. Mishra and P. Dutta, Return management in e-commerce firms: A machine learning approach to predict product returns and examine variables influencing returns, Journal of Cleaner Production **477**, 143802 (2024).

[15] A. S. Eruguz, O. Karabağ, E. Tetteroo, C. van Heijst, W. van den Heuvel, and R. Dekker, Customer-to-customer returns logistics: Can it mitigate the negative impact of online returns?, Omega **128**, 103127 (2024).

[16] H. Ibrahim Hayatu, A. Mohammed, and A. Barroon Isma'eel, Big data clustering techniques: Recent advances and survey, in *Machine Learning and Data Mining for Emerging Trend in Cyber Dynamics* (Springer International Publishing, 2021) p. 57–79.

[17] A. Dutt, M. A. Ismail, T. Herawan, and I. Targio, Partition-based clustering algorithms applied to mixed data for educational data mining: A survey from 1971 to 2024, IEEE Access **PP**, 1 (2024).

[18] J. Ji, T. Bai, C. Zhou, C. Ma, and Z. Wang, An improved k-prototypes clustering algorithm for mixed numeric and categorical data, Neurocomputing **120**, 590 (2013), image Feature Detection and Description.

[19] C. Hsu and Y. Chen, Mining of mixed data with application to catalog marketing, Expert Systems with Applications **32**, 12 (2007).

[20] J. Magidson and J. Vermunt, Latent class models (Sage, 2004) pp. 175–198.

[21] E. Molin, P. Mokhtarian, and M. Kroesen, Multimodal travel groups and attitudes: A latent class cluster analysis of dutch travelers, Transportation Research Part A: Policy and Practice **83**, 14 (2016).

[22] B. Ghattas and A. Sanchez San-Benito, Clustering approaches for mixed-type data: A comparative study, Journal of Probability and Statistics **2025**, 10.1155/jpas/2242100 (2025).

[23] A. A. Imran and M. N. Amin, Predicting the return of orders in the e-tail industry accompanying with model interpretation, Procedia Computer Science **176**, 1170 (2020), knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020.

[24] H. Cui, S. Rajagopalan, and A. R. Ward, Predicting product return volume using machine learning methods, European Journal of Operational Research **281**, 612 (2020), featured Cluster: Business Analytics: Defining the field and identifying a research agenda.

[25] A. Hofmann, F. Gwinner, K. Fuchs, and A. Winkelmann, An industry-agnostic approach for the prediction of return shipments, in *Proceedings of the 26th Americas Conference on Information Systems, AMCIS 2020* (2020).

[26] M. Samorani, F. Ahmed, and O. R. Zaïane, An application to product returns, in *2016 IEEE International Conference on Big Data (Big Data)* (2016) pp. 1454–1463.

[27] J. Kranz, P. Urbanke, and L. M. Kolbe, Predicting product returns in e-commerce: The contribution of mahalanobis feature extraction, in *Proceedings of the International Conference on Information Systems (ICIS)* (Association for Information Systems (AIS), Fort Worth, Texas, USA, 2015).

[28] M. Niederlaender, S. Driouech, S. Gry, A. N. Lodi, R. Biswas, and D. Werth, Towards Waste Reduction in E-Commerce: A Comparative Analysis of Machine Learning Algorithms and Optimisation Techniques for Garment Returns Prediction with Feature Importance Evaluation, SN Computer Science **6**, 10.1007/s42979-025-03944-z (2025).

[29] A. B. Stevenson and J. Rieck, Investigating returns management across e-commerce sectors and countries: Trends, perspectives, and future research, Logistics **8**, 82 (2024).

[30] I. Zennaro, . Finco, M. Calzavara, and A. Persona, Implementing e-commerce from logistic perspective: Literature review and methodological framework, Sustainability **14**, 10.3390/su14020911 (2022).

[31] Q. H. Duong, L. Zhou, M. Meng, T. V. Nguyen, P. Ieromonachou, and D. T. Nguyen, Understanding product returns: A systematic literature review using machine learning and bibliometric analysis, International Journal of Production Economics **243**, 108340 (2022).

[32] Q. H. Duong, L. Zhou, T. Van Nguyen, and M. Meng, Understanding and predicting online product return behavior: An interpretable machine learning approach, International Journal of Production Economics **280**, 109499 (2025).

[33] G. Walsh, M. Möhring, C. Koot, and M. Schaarschmidt, Preventive product returns management systems-a review and model., in *ECIS* (2014).

[34] J. Gäthke, K. Gelbrich, and S. Chen, A cross-national service strategy to manage product returns: E-tailers' return policies and the legitimating role of the institutional environment, Journal of Service Research **25**, 402 (2022).

[35] N. Janakiraman, H. A. Syrdal, and R. Freling, The effect of return policy leniency on consumer purchase and return decisions: A meta-analytic review, Journal of Retailing **92**, 226 (2016).

[36] D. Zhang, R. Frei, P. Senyo, S. Bayer, E. Gerding, G. Wills, and A. Beck, Understanding fraudulent returns and mitigation strategies in multichannel retailing, Journal of Retailing and Consumer Services **70**, 103145 (2023).

[37] N. N. Bechwati and W. S. Siegal, The impact of the prechoice process on product returns, Journal of Marketing Research **42**, 358 (2005).

[38] D. A. Mollenkopf, R. Frankel, and I. Russo, Creating value through returns management: Exploring the marketing–operations interface, Journal of Operations Management **29**, 391 (2011), special Issue on Field Research in Operations and Supply Chain Management.

B

Summary of interviews

B.1. Early interviews

Interview 1: Supply chain and procurement expert The interviewee has 14 years of experience within working in supply chain. The takeaway from this experience is that return logistics is often seen as an add-on, not a differentiator. Logistics is outsourced to third parties. This is expected to happen for returns as well. For B2B, returns can be more complex due to technical components. For simple products (like clothing or games), inspection is straightforward; for technical items, more detailed checks are needed. Third parties handle returns. The process for what happens to returned items (resell, repair, discard) depends on the product. The reason for returns are not always systematically tracked, except for complaints or quality issues, which are generally well documented. Right now, there is no systematic analysis of return reasons or trends per product or season. Returns introduce a new supply flow, but it's unclear how this affects demand planning. Returns are not typically seen as demand, so they don't impact demand planning directly. Outbound logistics uses predictive models and software, but return flows are not well supported by most systems. Returns require inventory and financial adjustments (e.g., VAT, value reassignment), which many systems struggle with. A big bike company sold accessories, which were returned to local offices instead of the central warehouse, causing confusion about what to do with them. The returns were not a priority due to their small share of total revenue.

Interview 2: Supply chain and operations expert The return process can be explained in different phases. The customers can initiate a return via app, website, customer service, or in-store. When returning via app or website it is preferred by customers to have a return label included in the orders. Returns are brought to a drop-off point, picked up at home, or brought to a store. After this, a return will go to a return center or warehouse. This can either be in-house or outsourced. These third party logistics offer different packages. They can handle inspection, repair and sorting. When arrived in the warehouse or return center, there will be inspection on production condition, compliant with return conditions, these differ per product. This determines the next steps, restocking, refurbishing, recycling or discard. An important trade of is the economic value vs the return costs. Some companies will focus on sustainability as well in this trade off, but returns are rarely sustainable due to extra handling. This is why there are some strategies for reducing returns such as charging for returns, limiting free returns, incentivizing in-store drop-off.

Within this return process there is a limited use of AI/automation in returns management. For example, beauty retailers often refund even if a product is used, prioritizing customer satisfaction over fraud prevention. AI could help with fraud detection and process optimization, but is not widely implemented.

The biggest challenges are high operational costs, especially for low-margin or bulky items. Space and process disruption in warehouses due to inbound returns. Manual assessments are common; sometimes items are discarded before full assessment. For fast fashion specifically, returned items may be out of assortment by the time they come back. In this process data like return reasons are sometimes logged, but this is underutilized for process improvement. Predictive models are used for outbound logistics, but rarely for returns. The return model follows the outbound model and sees seasonal trends in returns. Outsourcing partners rely on historical data for planning.

Some brands (e.g., a Swedish children's clothing brand) integrate returns into their business model, offering second-hand sales and repair services. This is a very innovative example.

Interview 3: Retailer expert The interviewee has a background in social sciences and business administration. The expertise lies in program and change management for international clients and has experience of over ten years. This also includes the research on retailing called the retail monitor.This includes research on trends such as last-mile delivery and omnichannel experiences. Return logistics plays an important role in this context: retailers aim to create a seamless experience where online purchases can easily be returned in-store. The return process typically involves external logistics service providers, with inspection taking place in the company's own warehouses and decisions made regarding resale, repair, or disposal. Key challenges include high return volumes, products arriving in poor condition, and complex inventory management. To reduce returns, retailers implement strategies such as charging a small fee or encouraging free in-store returns, with ease of use being crucial. Technologies like AI and predictive models are still rarely applied in return processes but hold potential for forecasting return likelihood and optimization. Sustainability does play a role, but profitability remains the main priority. The future of return logistics lies in better data analysis, size prediction, and omnichannel integration to reduce costs and improve customer satisfaction.

Interview 4: Supply chain expert The company manages returns through two primary workflows: pre-announced returns and unannounced returns. Pre-announced returns occur when customers register their return through the webshop, select the items they wish to send back, and receive a return label to ship the products to the warehouse. In certain cases, particularly for low-value items, the company may opt for a "credit-only" return, where the customer is refunded without sending the product back, as the cost of return shipping would exceed the product's value. Customers can also choose whether they want a replacement item instead of a refund. Historically, gift cards were offered as compensation, but this practice is being phased out in favor of direct refunds or replacements. This is also done with defect items.

Unannounced returns typically involve undelivered packages, such as those not collected from parcel lockers or failed home deliveries, which are automatically sent back to the warehouse. When these packages arrive, they are scanned, and a return order is created based on the original shipping label. This ensures that all items from the original order are accounted for and that the customer receives the correct credit. This functionality is custom-built, as standard Dynamics processes do not support this level of automation. Once returns arrive at the warehouse, they undergo inspection. Items that have been opened or tampered with are marked as scrap and disposed of, while unopened and resellable goods are restocked. There is no price-based decision-making at this stage; the primary criterion is whether the product is in a sellable condition. Exceptions include outlet or inactive items, which are also scrapped to avoid unnecessary storage costs and inventory complexity. The company aims to process returns quickly, minimizing delays in restocking. At this time, there is no return costs for the customer. This company is one of the few in this branche to allow free returns.

One of the biggest operational challenges lies in restocking returned items. With hundreds of packages arriving daily, often totaling 500 or more individual items, placing each product back in its designated location across the warehouse is labor-intensive. The current system uses fixed storage locations, which means that employees often need to traverse the entire warehouse to return items, significantly increasing time and cost. To address this, the company is exploring optimization strategies, such as creating consolidated bins for frequently returned or high-demand items. This approach would allow order pickers to retrieve these items during regular picking routes, reducing walking time and improving efficiency. Additionally, for products with moderate return and sales volumes, grouping them in shared bins may strike a balance between search time and overall operational cost savings.

B.2. Strategy validation interviews

Interview 1: Supply chain expert During the interview, the outputs of the clustering algorithms and the proposed strategies were discussed. The expert confirmed that the clusters aligned with expectations. Products in Category F were anticipated to have the highest return rates, and the sizing issue was also expected and validated by the analysis. A new insight emerged: items labeled “One Size” carry a lower return risk compared to other sizes.

Regarding strategies, the expert emphasized that improving product information is useful and based on solid insights. Policy adjustments were noted as potential interventions, but further internal research is recommended before implementation. For operational improvements, the expert highlighted that DHLDE operates only in Germany, making incentives for PostNL ineffective. This suggests the need for deeper analysis of orders and returns across different countries and environments.

Addressing high-risk hours (evening and night), the idea of halting orders during these times was considered impractical for the retailer. Marketing and sales strategies appear straightforward and can now be informed by product risk levels. Lowering the free shipping threshold should involve a trade-off between order value and item count. Flagging large orders for confirmation was seen as resource-intensive and challenging to implement.

Interview 2 The interview discussed the validation of the strategy by reviewing key cluster characteristics and related insights. The distinction between high-risk and low-risk clusters appeared logical to the interviewee, particularly the finding that sizes other than one size carry higher return risks. After explaining that medium, low, and high prices are defined relative to the retailer’s average price level, the interviewee agreed that this categorization was sensible. Questions were raised about where one-item orders fall within the clusters, and the interviewee found it notable that DHLDE shipments show higher risk, potentially due to lower accuracy rather than delivery speed. A referenced study emphasizing that accuracy is more important than speed for customer satisfaction supported this idea. It also seemed reasonable that sub-clusters with around 1.5 items would involve more risk than those with a single item, and the differing return rates across times of day were considered interesting. Regarding the proposed strategies, improved product descriptions were viewed as consistently valuable. With maybe adding an indication of the model’s size. The strategy “pay extra attention to S&L” was deemed too vague. One-size options may already be applied where feasible. Stricter rules were discussed, with monetary penalties perceived as more customer-friendly than inconvenience-based measures, while shortening the return window may not be legally allowed. Restocking fees could undermine the benefits of a free-shipping threshold, and addressing high-risk hours cannot realistically involve shutting the webshop overnight. Overall, the marketing strategies presented were considered appropriate.

C

Literature review: overview tables

C.1. Overview Literature review

Table C.1: Overview of Literature Review

Paper	Dimension	Data-driven approach	Main Activity	Product Type
Stevenson et al. 2024 [45]	Introduces preventive & reactive	Some of the examined literature use model	Literature review	Fashion; Electronics; Home & Living; Media
Frei et al. 2022 [16]	Reactive	No	Mapping the return process	-
Gry et al. 2024 [20]	Reactive	AI-based recommendation system	Conducting literature review and interviews	Fashion
Karl et al. 2024 [27]	Preventive	Several forecasting models	Systematic literature review; discusses different predicting algorithms	Fashion & Electronics
Mishra et al. 2024 [32]	Preventive	Yes	Comparing the prediction accuracy of ML techniques	Various product types
Niederlander et al. 2024 [35]	Preventive	Yes	Comparing different ML techniques	Fashion
Alzoubi 2025 [3]	-	-	Researching AI integration in return logistics processing	-
Eruguzz et al. 2024 [14]	Reactive	Mathematical model	Customer-to-customer strategy	Fashion
Yang et al. 2020 [49]	Preventive	Yes	Case study of an AI online virtual-reality webroom	Fashion
This thesis	**Preventive strategies**	**Yes, clustering product- & order level information**	**Combining data-driven and process approaches in designing a tool for return flow improvements**	**Electronics & fashion**

C.2. Literature review: associated features with returns

This table (C.2) is an adjusted table from the systematic research of Karl (2025) [27]. It shows the columns which will be used in this research in order to compare whether this combination has been used before. As seen in the table this is not the case. The papers below Fuchs and Lutz (2021) are added by me. These are papers found by the search string and are also compared with this study.

Table C.2: Overview of papers and features used

Paper	Product/Order Price /discounts	Return Attributes (e.g. Reason Codes)	Product Attributes (Category, Brand, Size)	Product order History	Product return History	Basket composition	Order attributes (e.g., Payment)	Order/return timing
This thesis	X	X	X			X	X	X
Hess and Mayhew (1997)	X	X						
Potdar and Rogers (2012)	X	X	X					
Urbanke et al. (2015)			X			X	X	X
Ahmed et al. (2016)	X							
Heilig et al. (2016)	X		X			X		
Ding et al. (2016)								
Fu et al. (2016)	X	X	X	X	X			
Samorani et al. (2016)	X		X					
Drechsler and Lasch (2015)					X			X
Urbanke et al. (2017)	X		X			X	X	X
Li et al. (2018)			X			X		X
Zhu et al. (2018)			X					
Asdecker and Karl (2018)	X							X
Joshi et al. (2018)			X					
Li et al. (2019)	X		X	X	X			
Cui et al. (2020)			X	X	X	X		X
Shang et al. (2020)								X
Imran and Amin 2020						X	X	X
Ketzenberg et al. (2020)	X		X				X	

Paper	**Product/Order Price /discounts**	**Return Attributes (e.g. Reason Codes)**	**Product Attributes (Category, Brand, Size)**	**Product order History**	**Product return History**	**Basket composition**	**Order attributes (e.g., Payment)**	**Order/return timing**
Hofmann et al. (2020)				X	X	X		
John et al. (2020)		X					X	
Rezaei et al. (2021)	X							
Rajasekaran and Priyadarshini (2021)				X	X			X
Sweidan et al. (2020)	X					X		X
Fuchs and Lutz (2021)	X					X		
Mishra and Dutta (2024)					X	X		
Niederlaender (2025)	X					X		
Makkonen et al. (2021)						X		
Duong et al. (2025)	X					X		
Rezaei et al. (2021)	X	X				X		

C.3. Literature review: Strategies

Hypothesis	Strategy	Source
Packaging negatively influences returns	Sellers should improve the design quality and handling of packaging	[11] [46]
Lean return policy leads to more returns and return abuse	More restricted return policies in terms of effort and scope	[11]
Customer service has a positive influence on returns	Online fashion sellers would improve their effective communication service to support customers' doubts and questions	[11]
Unknown primary features have a high negative impact on experience products; Secondary features have negative effect on both	Increase knowledge on the features of a product	[11]
Product usage performance strongly increases return risk across all categories, while emphasizing reliability in reviews and customer interest in both performance and reliability significantly lowers the likelihood of returns	Ask consumers to review product performance	[11]
Durability of a product should be clear for electronics items	Clearly specify estimated product lifespan	[11]
Culture moderates the relationship between effort restrictiveness and product return behavior, such that effort restrictiveness decreases product returns among Western but not among Eastern customers		[18]
Culture moderates the relationship between perceived customer-oriented norms and product return behavior, such that perceived customer-oriented norms increase product returns among Eastern but not among Western customers	For Western markets, making the return process more effortful (e.g., requiring more steps to return an item) can reduce return rates, but this may negatively impact repurchase intentions	[18]
	Use technology and outsourcing to reduce returns and costs	[18]
Consumers' evaluations of each of the five levers that constitute returns policy leniency (monetary, time, effort, scope, exchange) represent a significant antecedent of the perceived value of a returns service	Making a trade-off between strict and lenient return policies	[6]
Policy leniency affects both purchase and returns	Varying in monetary, time, effort, scope and exchange leniency to reach the goal of the retailer	[25]
Allowing returns on products purchased during sale periods increases returns	Enforcing restrictions on more price sensitive segments of consumers	[25] [46]
Push strategies increase the number of returns	Choose marketing strategies that give time to generate cognitive responses	[7] [34]
Lenient return policy leads to more returns and more fraud	Make targeted adjustments in tightening policies	[51]
Increasing the costs of returning will decrease the return volume	Introduce restocking fees or other monetary instruments	[46]
Optimizing the customer satisfaction will lead to less returns	Optimizing shipping times	[46]
Introducing more effort in returning orders will lead to less returns	Limiting return channels	[46]
Returns are due to misfit or not living up to expectations of the customer	Using customer based instruments such as virtual try-ons, avatars, customer reviews and sizing guides can help reduce these kind of returns	[46]

Table C.3: Strategies for return reduction from literature

D

Statistics

D.1. Descriptive analysis order level

D.1.1. Numerical variables

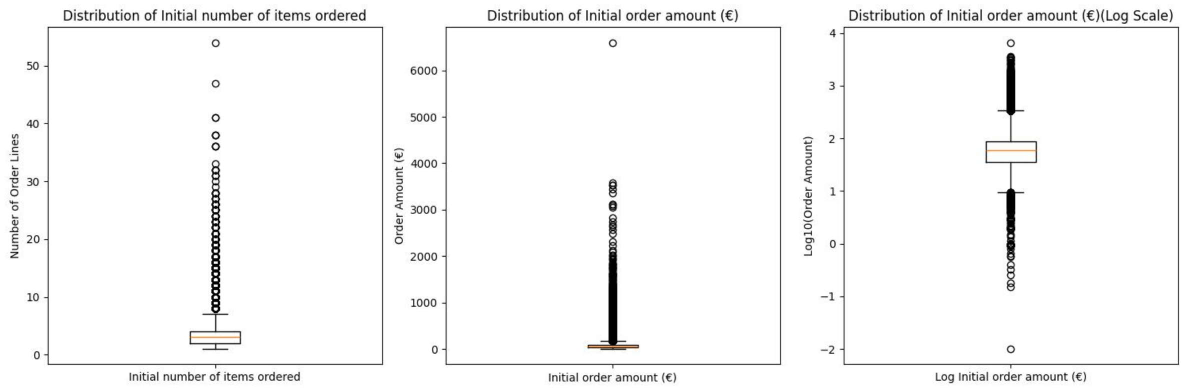
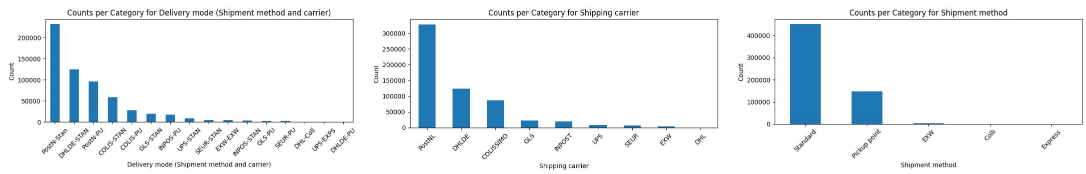


Figure D.1: Distribution of Numerical variables

D.1.2. Categorical variables

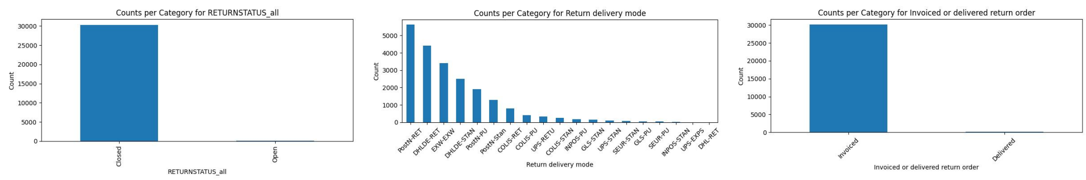


(a) Return delivery mode

(b) Shipping carriers

(c) Shipment method

Figure D.2: Overview of categorical variable distributions (Part 1)

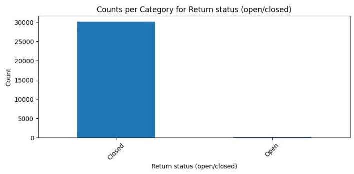


(d) Return statuses

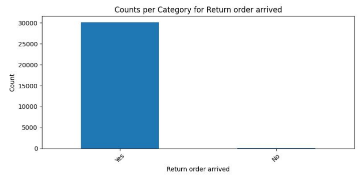
(e) Return delivery modes

(f) Return order statuses

Figure D.2: Overview of categorical variable distributions (Part 2)



(g) Return status closed/open



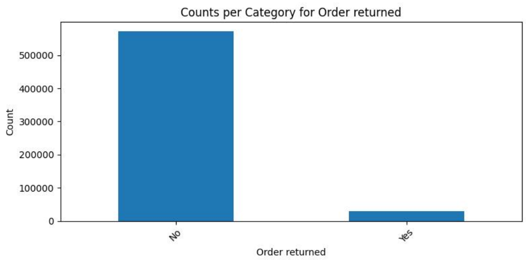
(h) Return arrived

Figure D.2: Overview of categorical variable distributions (Part 3)

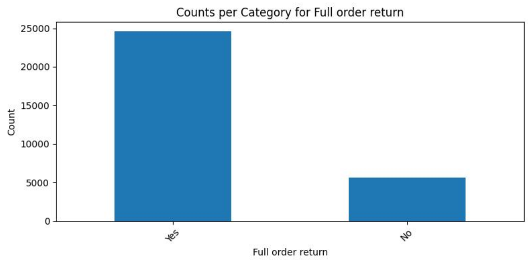
D.1.3. Return rate

Table D.1: Order Return Percentages

Metric	Yes (%)	No (%)
Order Returned	5.02	94.98
Full Order return	81.36	18.64



(a) Number of orders returned



(b) Number of returns with all products returned

Figure D.3: Comparison of return-related metrics

By Shipping method

Table D.2: Return Rate by shipping method

Service ID	Returned (%)	Count
Colli	0.00%	(0/5)
Express	0.00%	(0/4)
Pickup point	4.57%	(6708/146652)
Standard	4.27%	(19239/450691)

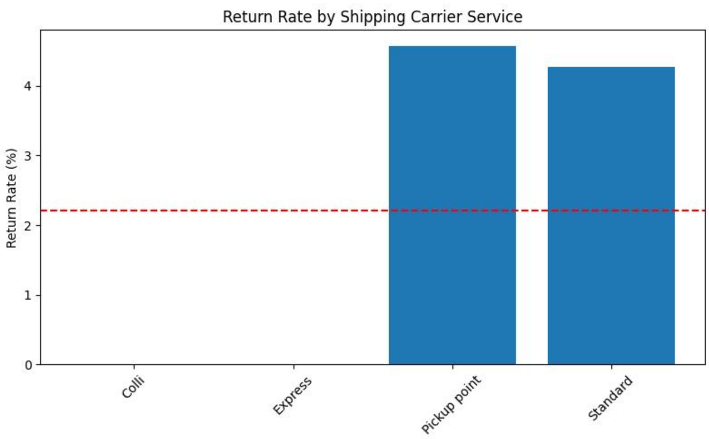


Figure D.4: Return rate by shipping method

By Hour of the day

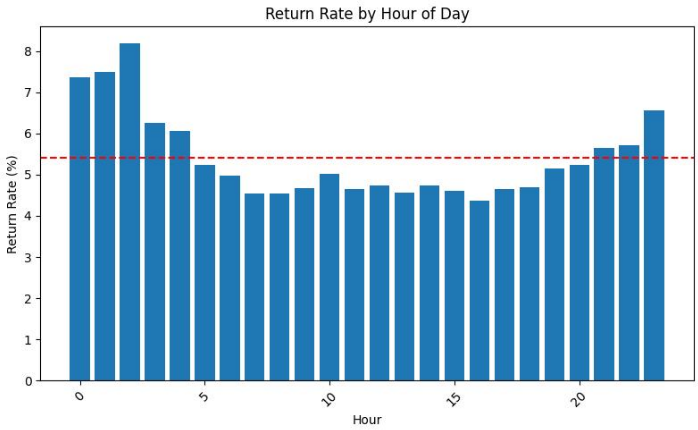


Figure D.5: Return rate by hour of the day

Return rates for binned variables

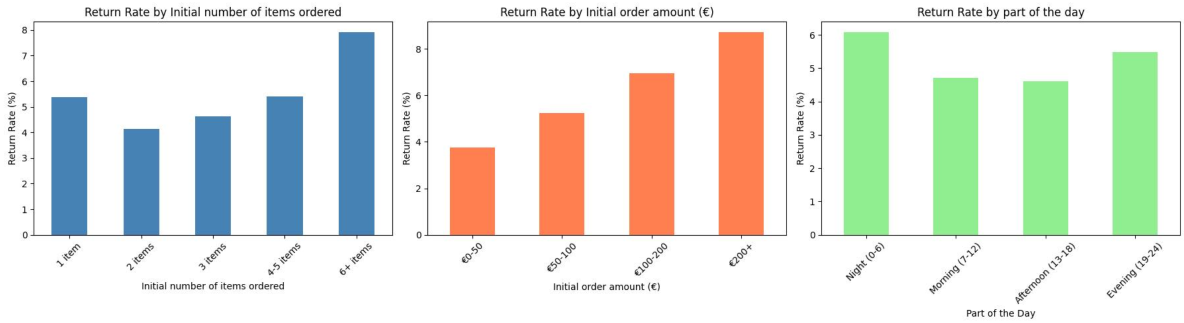


Figure D.6: Binned variables return rates

D.1.4. Correlations

Table D.3: Significance of Correlations

Variable	Correlation (r)	p-value	Significance
Initial number of items ordered vs Returned	0.055	0.0000	***
Initial order amount (€) vs Returned	0.053	0.000	***
Order hour vs Returned	0.001	0.6041	

*Significance levels: *** p<0.001, ** p<0.01, * p<0.05*

D.1.5. Chi-Square test

Table D.4: Chi-Square Test Results for Categorical Variables for relationship with return occurrences

Variable	Chi-Square	p-value	df	Significance
Delivery mode (Shipment method and carrier)	85508.93	0.0000	15	***
Shipping carrier	84787.10	0.0000	8	***
Shipment method	80960.19	0.0000	4	***
Initial number of items ordered (Binned)	1516.96	0.0000	4	***
Initial order amount (€) (Binned)	2247.11	0.0000	3	***
Order hour (Binned)	315.42	0.0000	3	***

*Significance levels: *** p<0.001, ** p<0.01, * p<0.05*

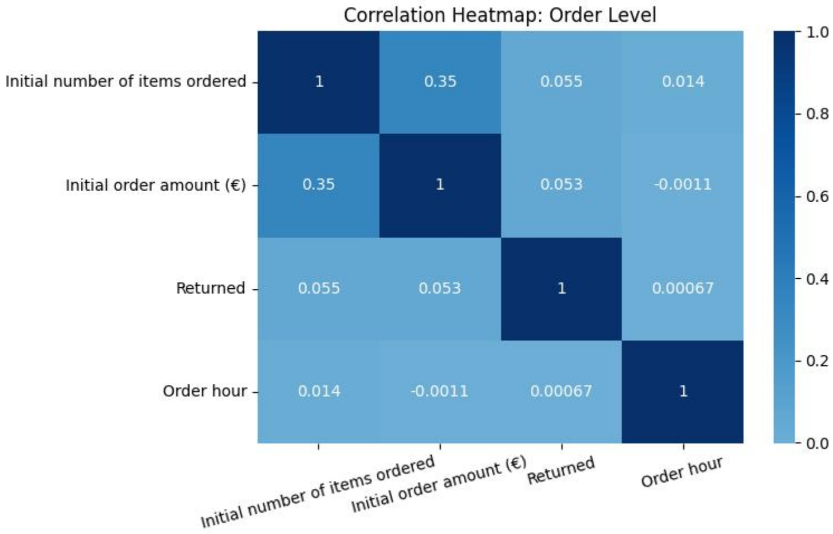


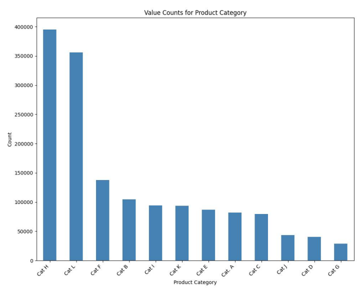
Figure D.7: Correlation heatmap

D.2. Descriptive analysis Product level

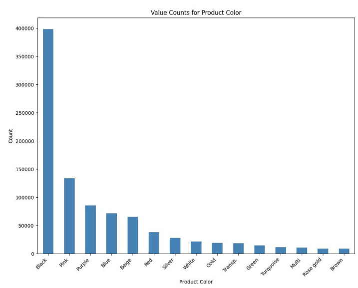
D.2.1. Numerical variables

Figure D.8: Distribution of Numerical variables

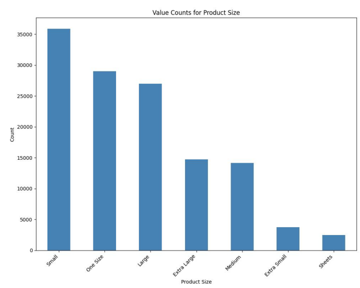
D.2.2. Categorical variables



(a) Product Categories

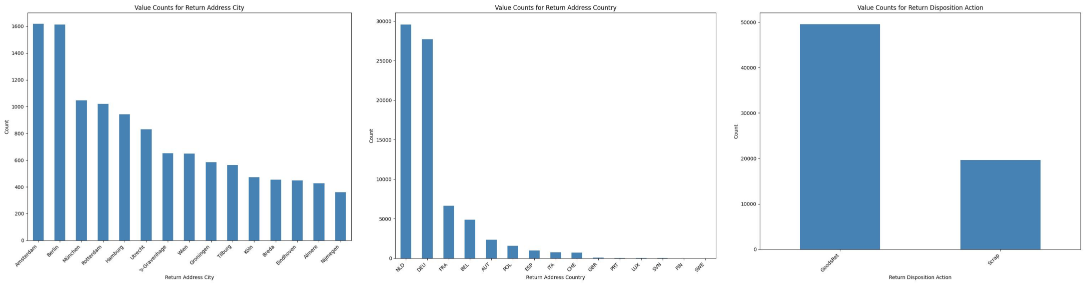


(b) Product colors (Top 15)



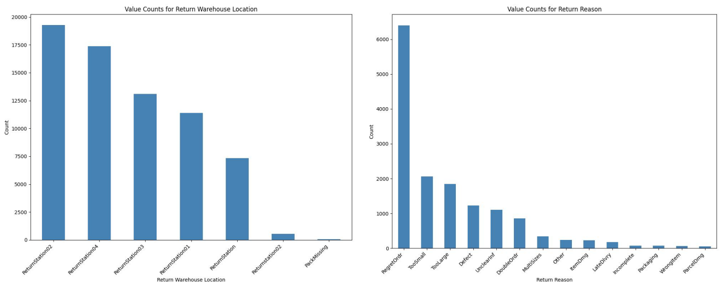
(c) Product sizes

Figure D.9: Overview of categorical variable distributions (Part 1)



(d) Return address city (Top 15) **(e)** Return address country (Top 15) **(f)** Return disposition

Figure D.9: Overview of categorical variable distributions (Part 2)



(g) Return Locations **(h)** Return reasons

Figure D.9: Overview of categorical variable distributions (Part 3)

D.2.3. Return rate

Table D.5: Product Return Percentages

Metric	Yes (%)	No (%)
Product Returned	6.84	93.16

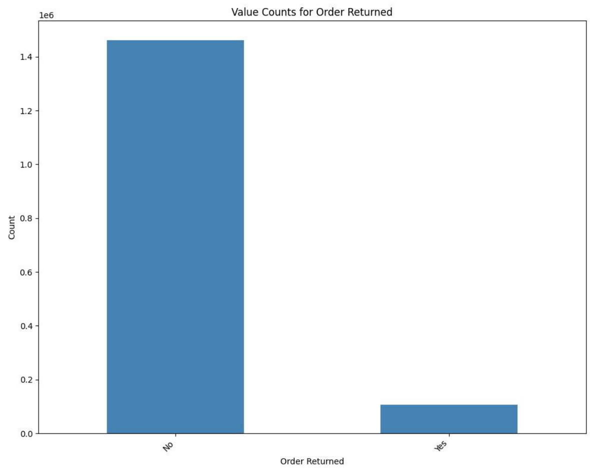


Figure D.10: Returned products counts

By Color

Table D.6: Return Rate by Product Color

(a)

Color	RR (%)	Ret/Total
Beige	7.19%	(4710/65505)
Black	9.40%	(37465/398434)
Blue	5.79%	(4137/71488)
Brown	6.95%	(601/8646)
Copper	3.55%	(5/141)
Cream	9.60%	(99/1031)
Gold	4.94%	(953/19309)
Green	6.45%	(950/14737)
Red	11.41%	(4352/38140)
Rose gold	5.50%	(484/8799)
Sand	33.33%	(1/3)
Silver	6.40%	(1794/28022)

(b)

Color	RR (%)	Ret/Total
Grey	6.65%	(188/2825)
Lavender	5.13%	(250/4870)
Magenta	7.05%	(550/7798)
Multi	5.81%	(614/10574)
Orange	4.56%	(134/2941)
Peach	6.93%	(49/707)
Pink	5.53%	(7384/133487)
Purple	6.33%	(5394/85249)
Transp.	5.99%	(1104/18422)
Turquoise	7.30%	(838/11472)
White	7.88%	(1713/21752)
Yellow	6.01%	(50/832)

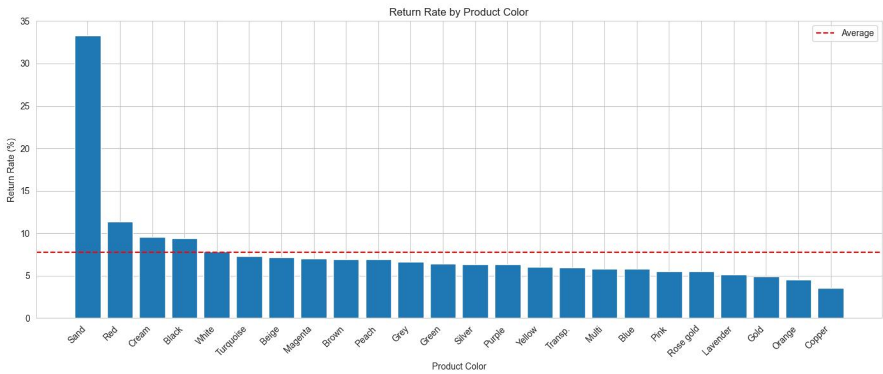


Figure D.11: Return Rate by Color

By Product Category

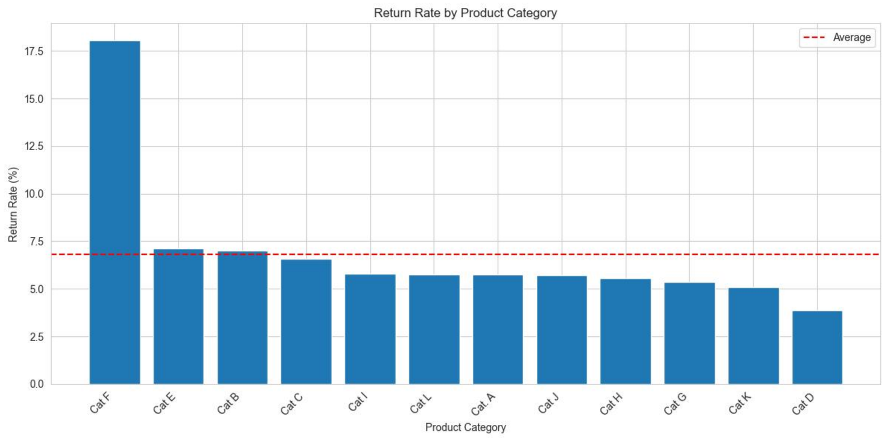


Figure D.12: Return Rate by product category

By Product size

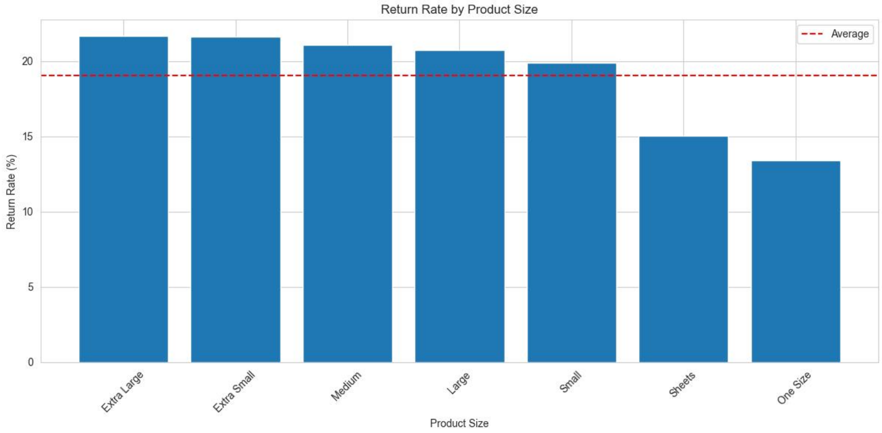


Figure D.13: Return Rate by size

By binned numerical variables

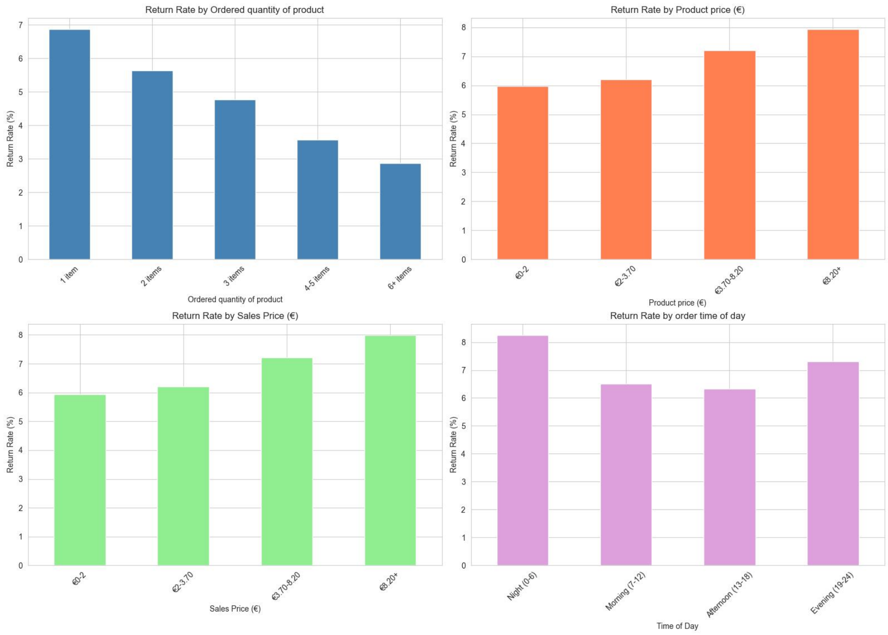


Figure D.14: Return rate of binned variables

D.2.4. Correlations

Table D.7: Significance of Correlations

Variable	Correlation (r)	p-value	Significance
Product price (€) vs Returned	0.0129	0.000	***
Ordered quantity of product vs Returned	-0.0086	0.000	***
Sales Price (€) vs Returned	0.0134	0.000	***
Order hour vs Returned	-0.0020	0.011	*

*Significance levels: *** $p<0.001$, ** $p<0.01$, * $p<0.05$*

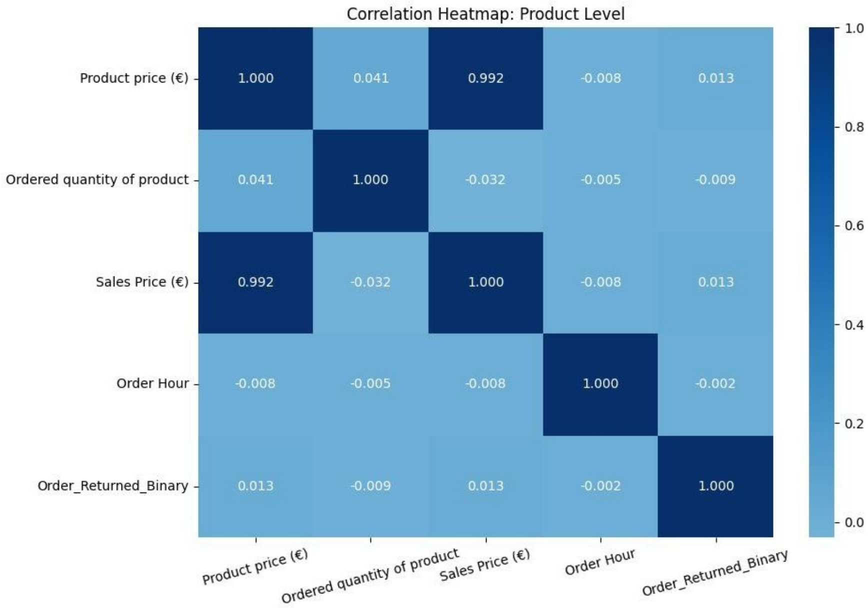


Figure D.15: Correlation Heatmap

D.2.5. Chi-Square test

Table D.8: Chi-Square Test Results for Categorical Variables for relationship with return occurrences

Variable	Chi-Square	p-value	df	Significance
Product Size	33426.37	0.0000	7	***
Product Category	31084.87	0.0000	12	***
Product Color	8034.43	0.0000	24	***
Ordered quantity of product (Binned)	137.55	0.0000	4	***
Product price (€) (Binned)	1597.32	0.0000	3	***
Sales Price (€) (Binned)	1703.32	0.0000	3	***
Order hour (Binned)	872.28	0.0000	3	***

*Significance levels: *** $p<0.001$, ** $p<0.01$, * $p<0.05$*

E Visualized Return processes

E.1. Current situation

Figure E.1 shows the current situation of the return process.

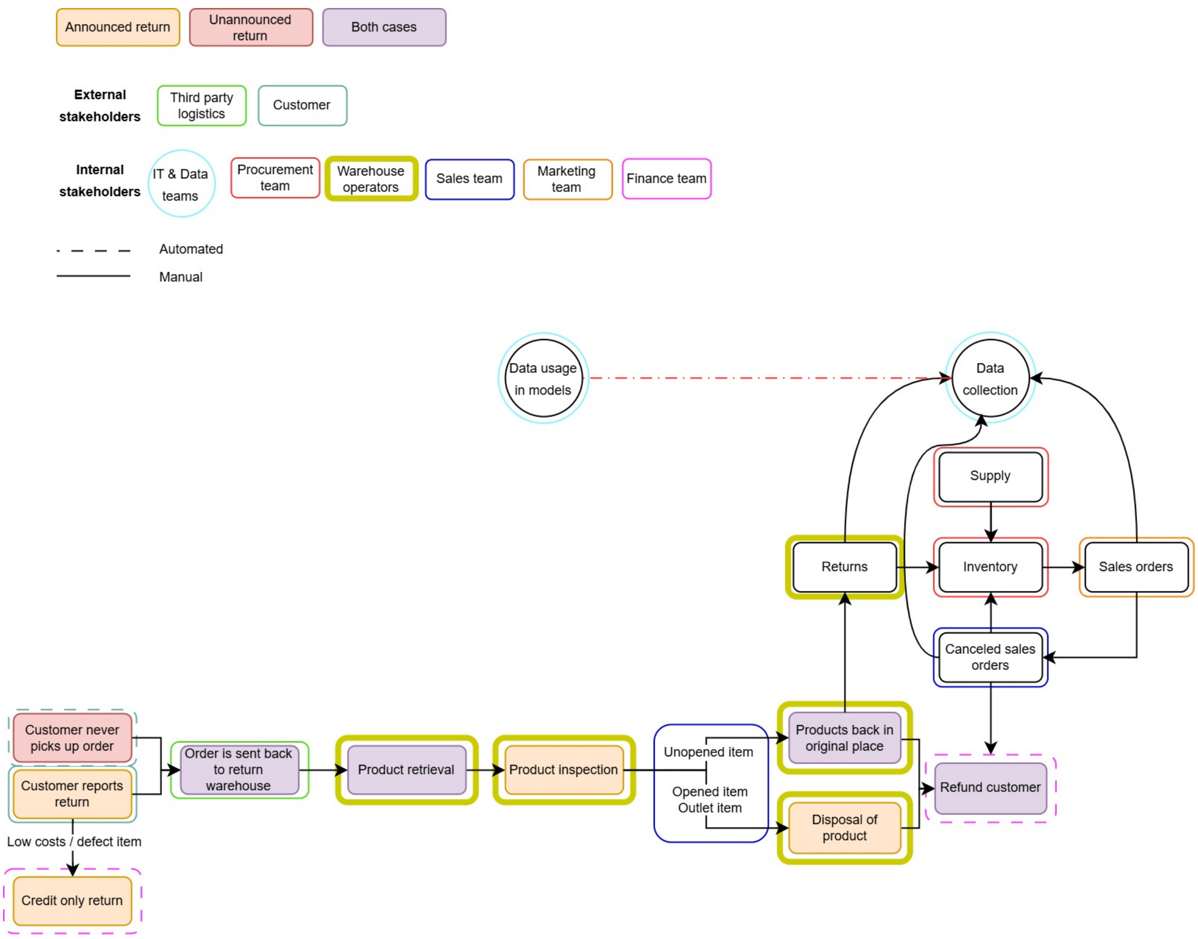


Figure E.1: Current situation – in depth stakeholders

E.2. Improved return process

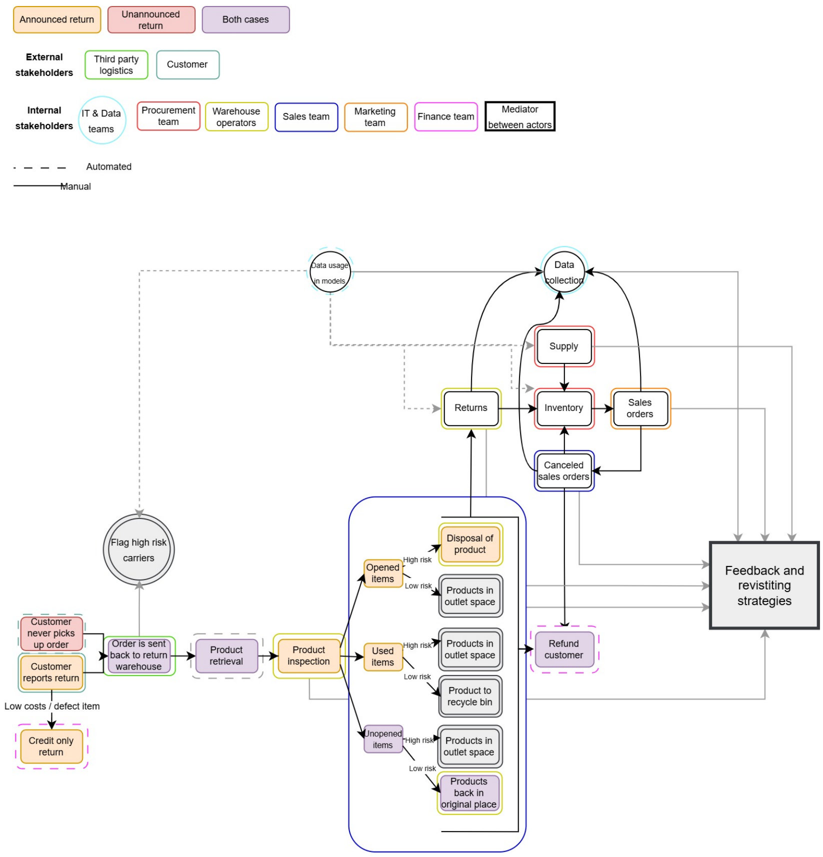


Figure E.2: New return process

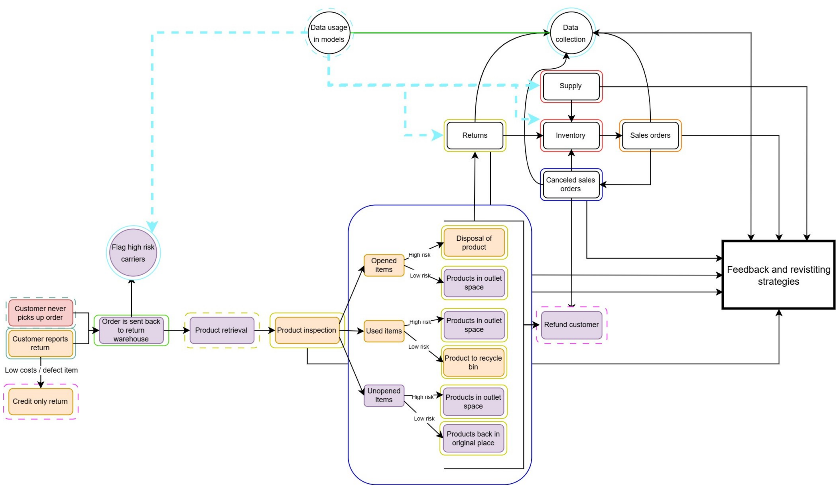


Figure E.3: New return process – stakeholders colored

F

Code

In this chapter the clustering algorithm codes will be shown. In this example the code is used for the product level information dataset.

F.1. Elbow method algorithm, defining optimal k

In this code the elbow method is used to find the optimal number of clusters. This is done by testing clusters in the range of 2-11 and finding the costs. Once the optimal number of clusters have been found, this is used in the clustering algorithms.

```python
# import necessary libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.decomposition import PCA
from sklearn.preprocessing import LabelEncoder
from kmodes.kprototypes import KPrototypes
from kneed import KneeLocator

# Load dataset
input_path = "C:\\Users\\qjapikse001\\OneDrive - PwC\\Documents\\Thesis\\Data\\Order- and return data [XYZ]\\Right datasets_cleaned\\ProductLevelFile_clusterAnalysis.csv"
df = pd.read_csv(input_path, low_memory=False)

# Cluster on these variables:
independent_vars = [
    'Product Category',             # Categorical
    'Product Color',                # Categorical
    'Product price €()',            # Numerical
    'Ordered quantity of product',  # Numerical
    'Product Size',                 # Categorical
    'Hour_Binned'                   # Categorical
]

dependant_var = 'Order_Returned_Binary'  # Target variable (not used in clustering)


# Filter dataframe to only include independent variables
df_cluster = df[independent_vars].copy()

# Get the position of categorical columns in the filtered dataframe
catColsPos = [df_cluster.columns.get_loc(col) for col in list(df_cluster.select_dtypes(
    include=['object']).columns)]
print("Categorical Columns:", list(df_cluster.select_dtypes(include=['object']).columns))
print("Categorical Columns Positions:", catColsPos)

# Convert FILTERED dataframe to numpy matrix (only independent vars)
dfMatrix = df_cluster.to_numpy()

```

```
# Choose number of clusters using elbow method
cost = []

for clusters in range(2, 11):
    try:
        kproto = KPrototypes(n_clusters=clusters, init='Cao', n_init=1, verbose=1, random_state=42)
        kproto.fit(dfMatrix, categorical=catColsPos)
        cost.append(kproto.cost_)
        print(f'Clusters: {clusters}, Cost: {kproto.cost_}')
    except:
        break

# Plot the elbow graph with detected elbow point
plt.figure(figsize=(10, 6))
k_values = range(2, 2 + len(cost))
plt.plot(k_values, cost, 'bo-', linewidth=2, markersize=8, label='Cost')
plt.xlabel('Number of clusters', fontsize=12)
plt.ylabel('Cost', fontsize=12)
plt.title('K-Prototypes Elbow Method', fontsize=14, fontweight='bold')
plt.grid(True, alpha=0.3)

# From the elbow graph, choose the optimal number of clusters
# Kneelocator detects maximum curvature = best cost/complexity trade-off
cost_knee = KneeLocator(k_values, cost, curve='convex', direction='decreasing', online=True)
K_cost_knee = cost_knee.elbow
print(f'Detected elbow at k={K_cost_knee}')

# Mark the elbow on the plot
if K_cost_knee:
    elbow_idx = K_cost_knee - 2
    plt.axvline(x=K_cost_knee, color='red', linestyle='--', linewidth=2, label=f'Elbow at k={K_cost_knee}')
    plt.plot(K_cost_knee, cost[elbow_idx], 'r*', markersize=20, label='Optimal k')

plt.legend(fontsize=10)
plt.tight_layout()
plt.savefig('Output Clustering Analysis/kprototypes_elbow.png', dpi=300, bbox_inches='tight')
plt.show()
```

F.2. k-prototype algorithm

Once the optimal number of clusters have been found the clusters will be filled using the following code to follow the algorithm. This algorithm can be used, using a python library. The clusters are added to the original dataset and then analysed, without running this algorithm again.

```
# import necessary libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.decomposition import PCA
from sklearn.preprocessing import LabelEncoder
from kmodes.kprototypes import KPrototypes
from kneed import KneeLocator

# Load dataset
input_path = "C:\\Users\\qjapikse001\\OneDrive - PwC\\Documents\\Thesis\\Data\\Order- and return data [XYZ]\\Right datasets_cleaned\\ProductLevelFile_clusterAnalysis.csv"
df = pd.read_csv(input_path, low_memory=False)

# Cluster on these variables:
independent_vars = [
    'Product Category',              # Categorical
    'Product Color',                 # Categorical
    'Product price €()',             # Numerical
    'Ordered quantity of product',   # Numerical
    'Product Size',                  # Categorical
    'Hour_Binned'                    # Categorical
]

```

```
dependant_var = 'Order_Returned_Binary'  # Target variable (not used in this clustering algorithm)

# Filter dataframe to only include independent variables
df_cluster = df[independent_vars].copy()

# Get the position of categorical columns in the filtered dataframe
catColsPos = [df_cluster.columns.get_loc(col) for col in list(df_cluster.select_dtypes(include=['object']).columns)]
print("Categorical␣Columns:", list(df_cluster.select_dtypes(include=['object']).columns))
print("Categorical␣Columns␣Positions:", catColsPos)

# Convert FILTERED dataframe to numpy matrix (only independent vars)
dfMatrix = df_cluster.to_numpy()

# Optimal number of clusters determined from elbow method
K_optimal = 4  # Set this based on the elbow method result

# Fit K-Prototypes model with optimal number of clusters
kproto = KPrototypes(n_clusters=K_optimal, init='Cao', n_init=5, verbose=1, random_state=42)
clusters = kproto.fit_predict(dfMatrix, categorical=catColsPos)
print("Cluster␣centroids:")
print(kproto.cluster_centroids_)

# Add cluster assignments to the original dataframe
df['KPrototypes_Cluster'] = clusters

#print("Cluster assignments added to the dataframe.")
# Export the dataframe with cluster assignments
output_path = "C:\\Users\\qjapikse001\\OneDrive␣-␣PwC\\Documents\\Thesis\\Data\\Order-␣and␣return␣data␣[XYZ]\\Right␣datasets_cleaned\\ProductLevelFile_clusterAnalysis_withKPrototypes.csv"
df.to_csv(output_path, index=False)
print(f"Data␣with␣cluster␣assignments␣exported␣to␣{output_path}")

# Read new dataframe with clusters for PCA visualization
df_with_clusters = pd.read_csv(output_path, low_memory=False)
```

F.3. CAVE clustering algorithm

```
#-----------------------------------------------------------------------------------
# CAVE Prototyping for Clustering with Mixed Data Types
# Loading libraries
import numpy as np
import pandas as pd
import math
from collections import Counter
#-----------------------------------------------------------------------------------

# ------ Step 1: Load data -----------

df = pd.read_csv #Put your input data here

# Cluster on these variables:
independent_vars = [
    'Delivery␣mode␣(Shipment␣method␣and␣carrier)',           # Categorical
    'Initial␣order␣amount␣€()',                               # Numerical
    'Shipping␣carrier',                                       # Categorical
    'Shipment␣method',                                        # Categorical
    'Hour_Binned'                                             # Categorical
]

# Specify your numeric and categorical column names
num_cols = ['Product␣price␣€()', 'Ordered␣quantity␣of␣product']  # Place numeric columns
cat_cols = ['Product␣Color', 'Product␣Category', 'Product␣Size','Hour_Binned']  # Place categorical columns

# --- Step 2: Compute weights ---
def compute_entropy(series):
    counts = Counter(series)
```

```
    total = sum(counts.values())
    probs = [c / total for c in counts.values()]
    return -sum(p * math.log2(p) for p in probs)

# OPTION A: VARIANCE-BASED (original CAVE - ignores rare but important patterns)
print("\n" + "="*80)
print("WEIGHTING␣METHODS:")
print("="*80 + "\n")

# Variance for numeric features
variances = df[num_cols].var().values
num_weights_variance = variances / variances.sum()

for col, w, v in zip(num_cols, num_weights_variance, variances):
    print(f"␣␣{col}:␣weight={w:.4f},␣variance={v:.2f}")

# Entropy for categorical features
cat_weights = []
for col in cat_cols:
    ent = compute_entropy(df[col])
    max_ent = math.log2(df[col].nunique())
    norm_ent = ent / max_ent if max_ent > 0 else 0
    cat_weights.append(1 - norm_ent)

cat_weights = np.array(cat_weights)

print("\nCATEGORICAL␣FEATURE␣WEIGHTS:")
for col, w in zip(cat_cols, cat_weights):
    print(f"␣␣{col}:␣{w:.4f}")

# Normalize so numeric and categorical blocks each sum to 0.5
num_weights = num_weights / num_weights.sum() * 0.5
cat_weights = cat_weights / cat_weights.sum() * 0.5
weights = np.concatenate([num_weights, cat_weights])

print("\n" + "="*80)
print("FINAL␣NORMALIZED␣WEIGHTS:")
print("="*80)
for col, w in zip(num_cols + cat_cols, weights):
    print(f"{col}:␣{w:.4f}")

# --- Step 3: Simple weighted distance function ---
def cave_distance(row, center_num, center_cat, weights):
    num_dist = ((row[num_cols].values.astype(float) - center_num) ** 2)
    cat_dist = (row[cat_cols].values != center_cat).astype(float)
    w_num = weights[:len(num_cols)]
    w_cat = weights[len(num_cols):]
    return np.sum(w_num * num_dist) + np.sum(w_cat * cat_dist)

# --- Step 4: Basic clustering loop ---
k = 4 # Put optimal k

#Convert to NumPy arrays ONCE (much faster than df.iloc in loop)
data_num = df_sample[num_cols].values.astype(float)  # Numerical data as NumPy array
data_cat = df_sample[cat_cols].values  # Categorical data as NumPy array
w_num = weights[:len(num_cols)]
w_cat = weights[len(num_cols):]

for iteration in range(10):  # simple fixed iterations
    print(f"␣␣Iteration␣{iteration␣+␣1}/10...", end="␣")
    old_assignments = assignments.copy()

    # VECTORIZED ASSIGNMENT: Calculate all distances at once
    # For each cluster, calculate distance to all points
    all_distances = np.zeros((n_samples, k))
    for j in range(k):
        # Numerical distance (vectorized)
        num_diff = data_num - centers_num[j]
        num_dist = np.sum(w_num * (num_diff ** 2), axis=1)

        # Categorical distance (vectorized)
```

```
        cat_dist = np.sum(w_cat * (data_cat != centers_cat[j]), axis=1)

        all_distances[:, j] = num_dist + cat_dist

    # Assign to nearest cluster
    assignments = np.argmin(all_distances, axis=1)

    moves = np.sum(assignments != old_assignments)

    # Update centers
    for j in range(k):
        mask = assignments == j
        if np.sum(mask) > 0:
            centers_num[j] = data_num[mask].mean(axis=0)
            # Categorical: mode of each column
            centers_cat[j] = [pd.Series(data_cat[mask, i]).mode()[0] for i in range(len(
                cat_cols))]

    print(f"{moves:,}␣moves")
    if moves == 0:
        print(f"␣␣␣Converged␣after␣{iteration␣+␣1}␣iterations!")
        break

# Add cluster assignments back to dataframe
df_sample['CAVE_Cluster'] = assignments.astype(int)

print("\nCluster␣sizes:", np.bincount(assignments.astype(int)))

# Export the dataframe with cluster assignments to a new CSV
output_path = "C:\\Users\\qjapikse001\\OneDrive␣-␣PwC\\Documents\\Thesis\\Data\\Order-␣and␣
    return␣data␣[XYZ]\\Right␣datasets_cleaned\\ProductLevelFile_with_CAVE_clusters.csv"
df_sample.to_csv(output_path, index=False)
print(f"\n␣Clustered␣data␣exported␣to:\n␣␣{output_path}")
```

F.4. Latent Clustering Analysis

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder, KBinsDiscretizer
from LatentClassClustering import LatentClassClustering

# Load product-level data

input_path = "C:\\Users\\qjapikse001\\OneDrive␣-␣PwC\\Documents\\Thesis\\Data\\Order-␣and␣
    return␣data␣[XYZ]\\Right␣datasets_cleaned\\
    ProductLevelFile_clusterAnalysis_withKPrototypes.csv"
df = pd.read_csv(input_path, low_memory=False)

# Cluster on these variables:
independent_vars = [
    'Product␣Category',            # Categorical
    'Product␣Color',               # Categorical
    'Product␣price␣€()',           # Numerical
    'Ordered␣quantity␣of␣product', # Numerical
    'Product␣Size',                # Categorical
    'Hour_Binned'                  # Categorical
]

dependant_var = 'Order_Returned_Binary'  # Target variable

print(f"Loaded␣{len(df):,}␣orders\n")
# Keep original dataframe to merge clusters back later
df_original = df.copy()

# Select features for clustering
categorical_vars = [
    'Product␣Category',            # Categorical
    'Product␣Color',               # Categorical
    'Product␣Size',                # Categorical
```

```python
    'Hour_Binned'                  # Categorical
]

numerical_vars = [
    'Product price €()',           # Numerical
    'Ordered quantity of product', # Numerical
]

dependant_var = 'Order_Returned_Binary'

# ==================== PREPARE DATA ====================
# Latent Class Clustering works with CATEGORICAL data only
# Bin the numerical features

# Create a copy for LCC
df_lcc = df[categorical_vars + numerical_vars + [dependant_var]].copy()

# Bin numerical features into categories
n_bins = 5
binner = KBinsDiscretizer(n_bins=n_bins, encode='ordinal', strategy='quantile')

df_lcc['Product_Price_Binned'] = binner.fit_transform(df_lcc[['Product price €()']])
df_lcc['Quantity_Binned'] = binner.fit_transform(df_lcc[['Ordered quantity of product']])

# Drop original numerical columns
df_lcc = df_lcc.drop(columns=numerical_vars)

# Update feature list (all categorical now)
all_features = categorical_vars + ['Product_Price_Binned', 'Quantity_Binned']

print(f"Features for clustering ({len(all_features)} categorical variables):")
for feat in all_features:
    n_categories = df_lcc[feat].nunique()
    print(f"  - {feat}: {n_categories} categories")
print()

# ==================== ENCODE CATEGORICAL VARIABLES ====================
print("Encoding categorical variables...")

# Label encode all features
le_dict = {}
X_encoded = np.zeros((len(df_lcc), len(all_features)), dtype=int)

for idx, feat in enumerate(all_features):
    le = LabelEncoder()
    X_encoded[:, idx] = le.fit_transform(df_lcc[feat].astype(str))
    le_dict[feat] = le
    print(f"   Encoded {feat}: {len(le.classes_)} unique values")

print()

# ==================== FIT LATENT CLASS CLUSTERING ====================
print("="*80)
print("FITTING LATENT CLASS CLUSTERING")
print("="*80)

# Try different numbers of clusters
K_range = range(2, 6)
results = []

for k in K_range:
    print(f"\nFitting LCC with {k} clusters...")
    lcc = LatentClassClustering(
        n_clusters=k,
        max_iter=100,
        tol=1e-4,
        random_state=42
    )

    lcc.fit(X_encoded)

```

```
    # Get cluster assignments
    labels = lcc.predict(X_encoded)

    # Calculate BIC (Bayesian Information Criterion)
    log_likelihood = lcc._compute_log_likelihood(X_encoded)
    n_params = k - 1  # class probabilities
    for j in range(len(all_features)):
        n_params += k * (lcc.n_categories_[j] - 1)  # conditional probabilities

    bic = -2 * log_likelihood + n_params * np.log(len(df_lcc))

    # Calculate return rate per cluster
    df_lcc[f'LCC_Cluster_{k}'] = labels

    # Convert 'Order returned' to binary if needed
    if df_lcc[dependant_var].dtype == 'object':
        return_binary = df_lcc[dependant_var].map({'Yes': 1, 'No': 0})
    else:
        return_binary = df_lcc[dependant_var]

    cluster_return_rates = []
    for cluster_id in range(k):
        mask = labels == cluster_id
        return_rate = return_binary[mask].mean() * 100
        cluster_return_rates.append(return_rate)

    results.append({
        'K': k,
        'BIC': bic,
        'Log_Likelihood': log_likelihood,
        'Iterations': lcc.n_iter_,
        'Return_Rate_Range': f"{min(cluster_return_rates):.2f}% - {max(cluster_return_rates):.2f}%"
    })

    print(f"  Converged in {lcc.n_iter_} iterations")
    print(f"  BIC: {bic:,.2f}")
    print(f"  Return rate range across clusters: {min(cluster_return_rates):.2f}% - {max(cluster_return_rates):.2f}%")

print("\n" + "="*80)
print("MODEL SELECTION SUMMARY")
print("="*80)

results_df = pd.DataFrame(results)
print(results_df.to_string(index=False))

# Select optimal K based on lowest BIC
optimal_k = results_df.loc[results_df['BIC'].idxmin(), 'K']
print(f"\n Optimal K based on BIC: {optimal_k}")

# ==================== ANALYZE OPTIMAL CLUSTERS ====================
print("\n" + "="*80)
print(f"ANALYZING LATENT CLASS CLUSTERS (K={optimal_k})")
print("="*80)

# Fit final model with optimal K
lcc_final = LatentClassClustering(
    n_clusters=optimal_k,
    max_iter=100,
    tol=1e-4,
    random_state=42
)

lcc_final.fit(X_encoded)
df_lcc['LCC_Cluster'] = lcc_final.predict(X_encoded)

# Add cluster to original dataframe
df_original['LCC_Cluster'] = df_lcc['LCC_Cluster']

# Convert 'Order returned' to binary
```

```python
if df_lcc[dependant_var].dtype == 'object':
    df_lcc[dependant_var] = df_lcc[dependant_var].map({'Yes': 1, 'No': 0})

overall_return_rate = df_lcc[dependant_var].mean() * 100
print(f"\nOverall Return Rate: {overall_return_rate:.2f}%\n")

# Cluster characteristics
for cluster_id in range(optimal_k):
    cluster_data = df_lcc[df_lcc['LCC_Cluster'] == cluster_id]
    n_orders = len(cluster_data)
    return_rate = cluster_data[dependant_var].mean() * 100

    print(f"\n{''*80}")
    print(f"CLUSTER {cluster_id} ({n_orders:,} orders, {(n_orders/len(df_lcc)*100):.1f}%)")
    print(f"Return Rate: {return_rate:.2f}% ({return_rate - overall_return_rate:+.2f}% vs baseline)")
    print(f"{''*80}")

    # Show most common values for each categorical feature
    for feat in all_features:
        top_value = cluster_data[feat].mode()[0]

        # Decode if it's one of the original categorical features
        if feat in categorical_vars:
            top_count = (cluster_data[feat] == top_value).sum()
        else:
            top_count = (cluster_data[feat] == top_value).sum()

        pct = (top_count / n_orders) * 100
        print(f"  {feat}: {top_value} ({top_count:,} orders, {pct:.1f}%)")

# =================== SAVE RESULTS ====================
output_path = "C:\\Users\\qjapikse001\\OneDrive - PwC\\Documents\\Thesis\\Data\\Order- and return data [XYZ]\\Right datasets_cleaned\\ProductLevelFile_with_LatentClass.csv"
df_original.to_csv(output_path, index=False)
print(f"\n Saved results with LCC clusters to: {output_path}")

# Save model selection results
results_df.to_csv('Output Clustering Analysis/LatentClass_ModelSelection.csv', index=False)
print(f" Saved model selection summary to: Output Clustering Analysis/LatentClass_ModelSelection.csv")

print("\n" + "="*80)
print("LATENT CLASS CLUSTERING COMPLETE!")
print("="*80)
```

G

Cluster analysis output

G.1. Order Level

G.1.1. K-prototyping

Overview clusters

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Orders (share)	848 (0.1%)	225,266 (37.4%)	74,585 (12.4%)	12,243 (2.0%)	288,695 (48.0%)
Return rate	10.26%	5.61%	7.30%	8.21%	3.82%
Delivery mode (shipment method & carrier)					
INPOS-PU / INPOST	399 (47.1%) — RR: 4.76%	—	—	5,160 (42.1%) — RR: 2.85%	—
PostN-Stan / PostNL	155 (18.3%) — RR: 7.10%	86,951 (38.6%) — RR: 3.77%	21,972 (29.5%) — RR: 5.71%	1,913 (15.6%) — RR: 8.10%	120,763 (41.8%) — RR: 2.25%
DHLDE-STAN / DHLDE	92 (10.8%) — RR: 19.57%	46,924 (20.8%) — RR: 8.59%	18,804 (25.2%) — RR: 11.41%	1,805 (14.7%) — RR: 20.00%	56,628 (19.6%) — RR: 5.07%
PostN-PU / PostNL	—	37,787 (16.8%) — RR: 5.37%	9,049 (12.1%) — RR: 7.29%	—	48,123 (16.7%) — RR: 4.80%
Shipping carrier					
INPOST	554 (65.3%) — RR: 5.42%	—	9,417 (12.6%) — RR: 2.17%	6,280 (51.3%) — RR: 2.63%	—
PostNL	123 (14.5%) — RR: 17.07%	124,738 (55.4%) — RR: 4.26%	31,021 (41.6%) — RR: 6.17%	2,698 (22.0%) — RR: 9.38%	168,886 (58.5%) — RR: 2.98%
DHLDE	92 (10.8%) — RR: 19.57%	46,925 (20.8%) — RR: 8.59%	18,804 (25.2%) — RR: 11.41%	1,805 (14.7%) — RR: 20.00%	56,628 (19.6%) — RR: 5.07%
COLISSIMO	—	32,111 (14.3%) — RR: 2.69%	—	—	45,998 (15.9%) — RR: 1.97%
Shipment method					
Pickup point	440 (51.9%) — RR: 5.68%	54,147 (24.0%) — RR: 4.81%	20,741 (27.8%) — RR: 4.70%	6,259 (51.1%) — RR: 4.20%	65,065 (22.5%) — RR: 4.36%
Standard	392 (46.2%) — RR: 13.52%	169,438 (75.2%) — RR: 4.93%	53,224 (71.4%) — RR: 7.25%	5,840 (47.7%) — RR: 10.29%	221,797 (76.8%) — RR: 2.87%
EXW	12 (1.4%) — RR: 75.00%	1,681 (0.7%) — RR: 99.64%	619 (0.8%) — RR: 99.19%	141 (1.2%) — RR: 100.00%	1,832 (0.6%) — RR: 99.84%
Time-of-day					
Afternoon (13–18)	261 (30.8%) — RR: 13.03%	79,859 (35.5%) — RR: 5.16%	26,408 (35.4%) — RR: 6.69%	4,251 (34.7%) — RR: 7.72%	103,149 (35.7%) — RR: 3.49%
Evening (19–24)	246 (29.0%) — RR: 9.35%	70,662 (31.4%) — RR: 5.97%	22,497 (30.2%) — RR: 7.95%	3,799 (31.0%) — RR: 8.11%	86,186 (29.9%) — RR: 4.32%
Morning (7–12)	237 (27.9%) — RR: 8.44%	56,371 (25.0%) — RR: 5.48%	19,234 (25.8%) — RR: 6.98%	3,061 (25.0%) — RR: 8.04%	74,959 (26.0%) — RR: 3.38%
Night (0–6)	104 (12.3%) — RR: 9.62%	18,374 (8.2%) — RR: 6.56%	6,446 (8.6%) — RR: 8.53%	1,132 (9.2%) — RR: 10.87%	24,401 (8.5%) — RR: 4.86%
Numerical features					
Initial # items	6.85 (± 6.80)	3.61 (± 1.69)	4.19 (± 2.48)	5.63 (± 3.63)	2.79 (± 1.05)
Initial order amount (€)	964.97 (± 436.06)	79.17 (± 15.93)	158.17 (± 33.70)	342.89 (± 86.64)	35.39 (± 12.13)

Table G.1: Overview clusters K-prototyping – orders – with shares and return rates

Distributions across clusters

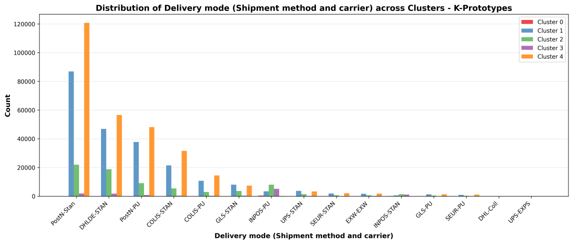


Figure G.1: Distribution of Delivery mode in clusters

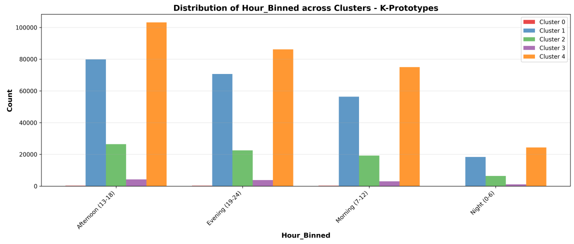


Figure G.2: Distribution of part of the days in clusters

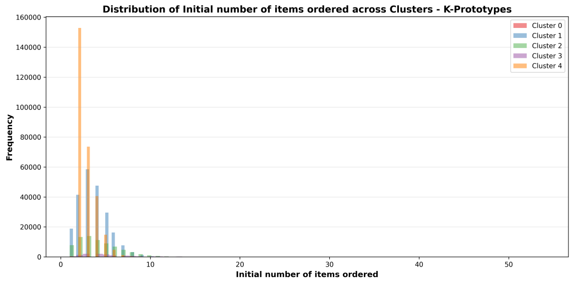


Figure G.3: Distribution of Initial number of items ordered in clusters

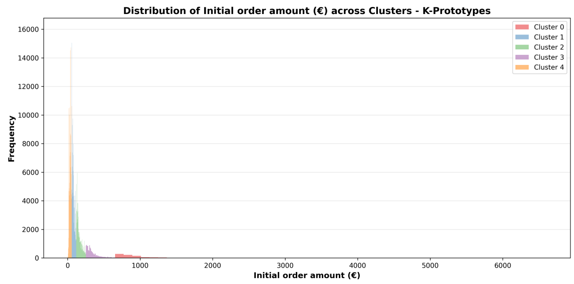


Figure G.4: Distribution of initial order amount in clusters

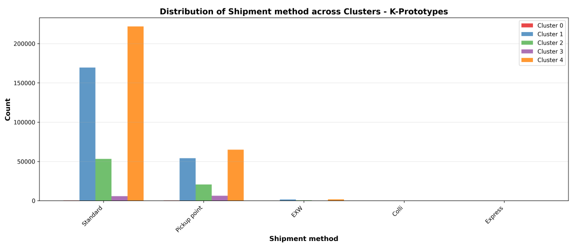


Figure G.5: Distribution of shipment method in clusters

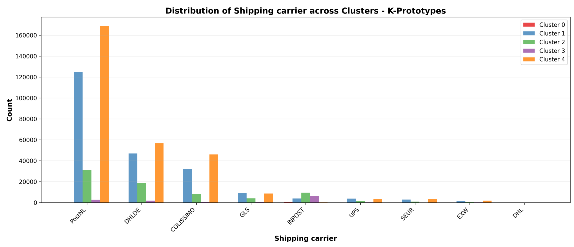
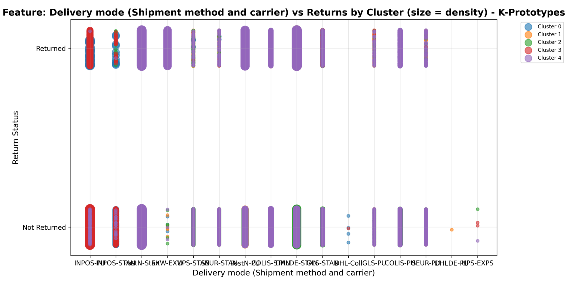
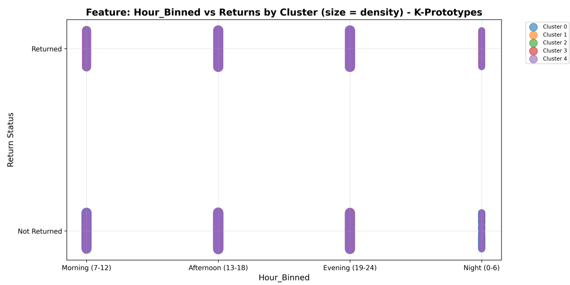


Figure G.6: Distribution of shipping carrier in clusters

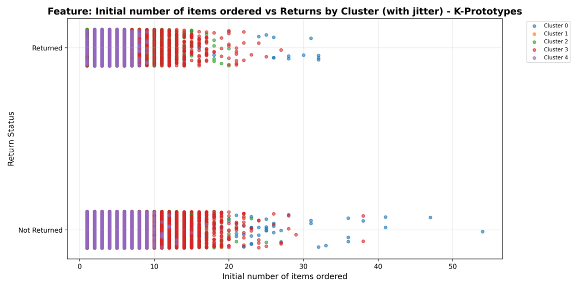
Scatterplots returned / not returned



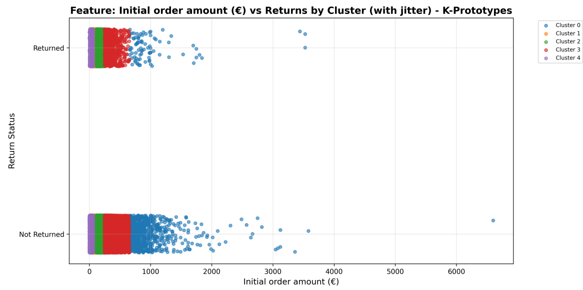
(a) Scatterplot delivery mode



(b) Scatterplot part of the day



(a) Scatterplot initial number of items ordered



(b) Scatterplot initial order amount

G.1.2. CAVE

Overview clusters

Feature	**Cluster 0**	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**
Orders (share)	9,439 (1.6%)	72,199 (12.0%)	178,868 (29.7%)	169,882 (28.2%)	171,249 (28.5%)
Return rate	8.37%	7.44%	4.35%	3.58%	5.94%
Delivery mode (shipment method & carrier)					
INPOS-PU / INPOST	4,482 (47.5%) — RR: 3.15%	—	—	—	—
PostN-Stan / PostNL	1,215 (12.9%) — RR: 9.63%	20,869 (28.9%) — RR: 5.81%	80,004 (44.7%) — RR: 2.84%	71,680 (42.2%) — RR: 1.99%	57,920 (33.8%) — RR: 4.13%
DHLDE-STAN / DHLDE	1,166 (12.4%) — RR: 21.70%	18,164 (25.2%) — RR: 11.80%	31,049 (17.4%) — RR: 6.11%	32,954 (19.4%) — RR: 4.54%	40,920 (23.9%) — RR: 8.89%
INPOS-STAN / INPOST	1,085 (11.5%) — RR: 2.40%	—	—	—	—
PostN-PU / PostNL	—	8,567 (11.9%) — RR: 7.38%	33,930 (19.0%) — RR: 4.77%	27,517 (16.2%) — RR: 4.94%	25,257 (14.7%) — RR: 5.63%
Shipping carrier					
INPOST	5,567 (59.0%) — RR: 3.00%	10,267 (14.2%) — RR: 2.24%	—	—	—
PostNL	1,722 (18.2%) — RR: 11.09%	29,436 (40.8%) — RR: 6.26%	113,934 (63.7%) — RR: 3.41%	99,197 (58.4%) — RR: 2.81%	83,177 (48.6%) — RR: 4.59%
DHLDE	1,166 (12.4%) — RR: 21.70%	18,164 (25.2%) — RR: 11.80%	31,049 (17.4%) — RR: 6.11%	32,954 (19.4%) — RR: 4.54%	40,921 (23.9%) — RR: 8.89%
COLISSIMO	—	—	23,209 (13.0%) — RR: 2.16%	28,032 (16.5%) — RR: 1.87%	27,860 (16.3%) — RR: 2.79%
Shipment method					
Pickup point	5,189 (55.0%) — RR: 4.43%	20,681 (28.6%) — RR: 4.59%	43,104 (24.1%) — RR: 4.45%	37,600 (22.1%) — RR: 4.44%	40,078 (23.4%) — RR: 4.84%
Standard	4,125 (43.7%) — RR: 10.76%	50,907 (70.5%) — RR: 7.50%	134,558 (75.2%) — RR: 3.47%	131,212 (77.2%) — RR: 2.55%	129,889 (75.8%) — RR: 5.36%
EXW	119 (1.3%) — RR: 97.48%	609 (0.8%) — RR: 99.18%	1,206 (0.7%) — RR: 99.75%	1,069 (0.6%) — RR: 99.81%	1,282 (0.7%) — RR: 99.69%
Time-of-day					
Afternoon (13–18)	3,211 (34.0%) — RR: 8.56%	25,545 (35.4%) — RR: 6.76%	63,563 (35.5%) — RR: 3.96%	60,673 (35.7%) — RR: 3.28%	60,936 (35.6%) — RR: 5.47%
Evening (19–24)	2,922 (31.0%) — RR: 8.42%	21,782 (30.2%) — RR: 8.00%	55,639 (31.1%) — RR: 4.85%	50,012 (29.4%) — RR: 4.01%	53,035 (31.0%) — RR: 6.35%
Morning (7–12)	2,404 (25.5%) — RR: 7.36%	18,602 (25.8%) — RR: 7.26%	45,130 (25.2%) — RR: 4.01%	44,550 (26.2%) — RR: 3.15%	43,176 (25.2%) — RR: 5.78%
Night (0–6)	902 (9.6%) — RR: 10.20%	6,270 (8.7%) — RR: 8.77%	14,536 (8.1%) — RR: 5.22%	14,647 (8.6%) — RR: 4.72%	14,102 (8.2%) — RR: 6.96%
Numerical features					
Initial # items	5.73 (± 4.18)	4.31 (± 2.56)	3.25 (± 1.28)	2.54 (± 0.86)	3.66 (± 1.79)
Initial order amount (€)	428.03 (± 228.19)	166.88 (± 39.79)	52.52 (± 8.39)	26.88 (± 7.27)	86.53 (± 15.40)

Table G.2: Overview clusters CAVE — orders — with shares and return rates

Distributions across clusters

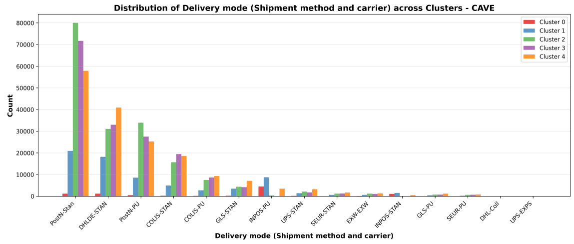


Figure G.9: Distribution of Delivery mode in clusters

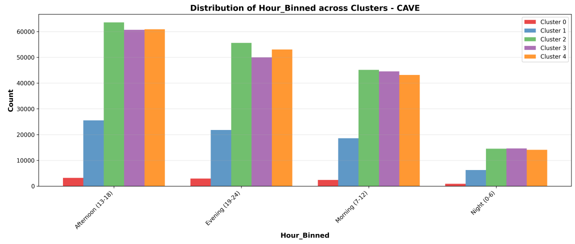


Figure G.10: Distribution of part of the days in clusters

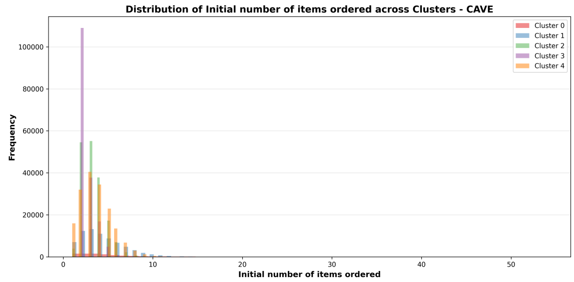


Figure G.11: Distribution of Initial number of items ordered in clusters

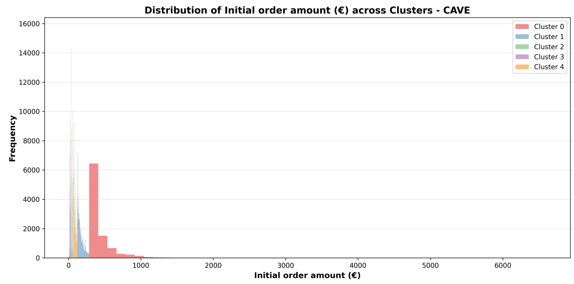


Figure G.12: Distribution of initial order amount in clusters

Figure G.13: Distribution of shipment method in clusters

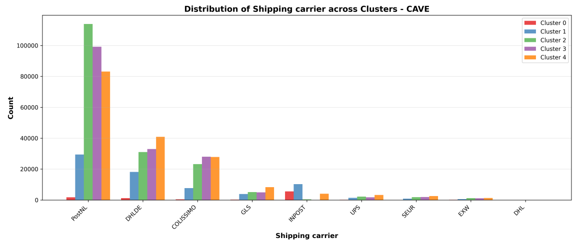
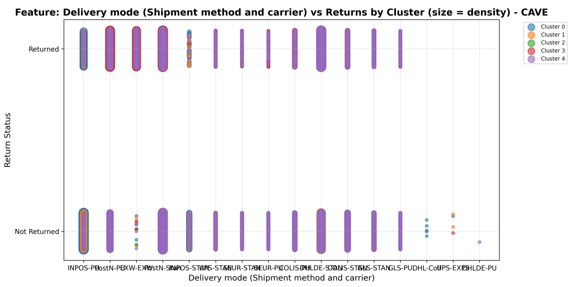


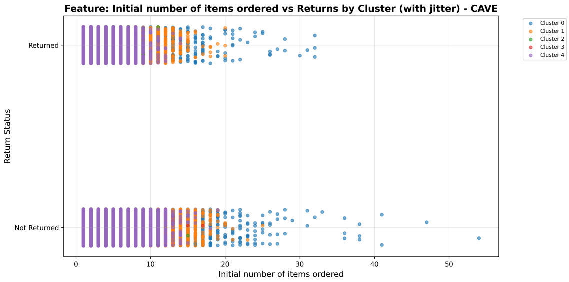
Figure G.14: Distribution of shipping carrier in clusters

Scatterplots returned / not returned

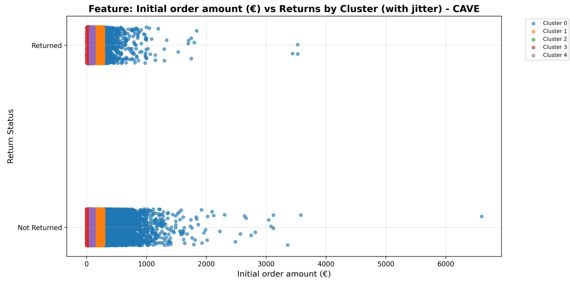


(a) Scatterplot delivery mode

(b) Scatterplot part of the day



(c) Scatterplot initial number of items ordered



(d) Scatterplot initial order amount

G.1.3. LCC

Overview Clusters

Distributions across clusters

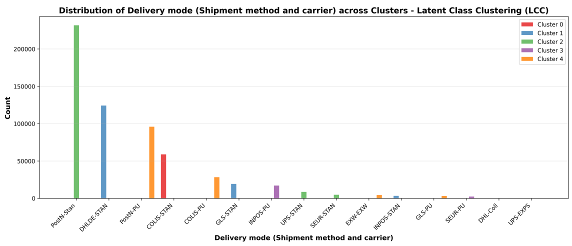


Figure G.16: Distribution of Delivery mode in clusters

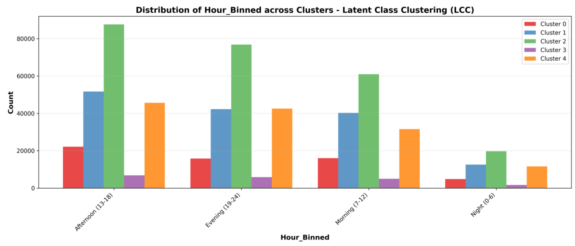


Figure G.17: Distribution of part of the days in clusters

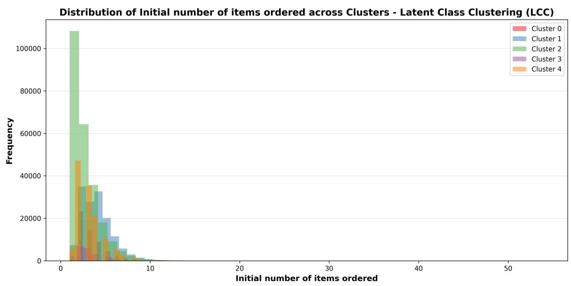


Figure G.18: Distribution of Initial number of items ordered in clusters

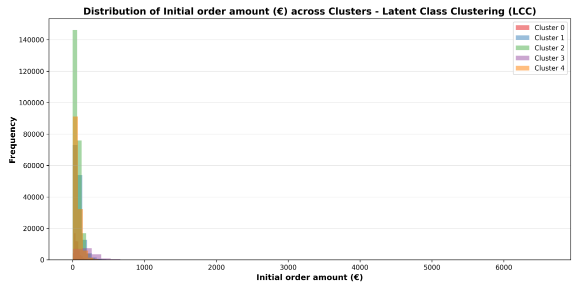


Figure G.19: Distribution of initial order amount in clusters

Figure G.20: Distribution of shipment method in clusters

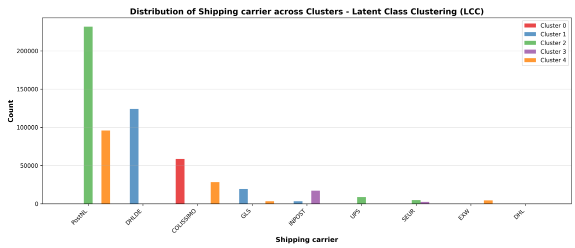


Figure G.21: Distribution of shipping carrier in clusters

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Orders (share)	58,839 (9.8%)	146,745 (24.4%)	245,112 (40.7%)	19,494 (3.2%)	131,447 (21.8%)
Return rate	1.93%	7.00%	3.19%	2.81%	7.93%
Delivery mode (shipment method & carrier)					
COLIS-STAN / COLISSIMO	58,839 (100.0%) — RR: 1.93%	—	—	—	28,295 (21.5%) — RR: 3.21%
DHLDE-STAN / DHLDE	—	124,253 (84.7%) — RR: 7.59%	—	—	—
GLS-STAN / GLS	—	19,327 (13.2%) — RR: 4.13%	—	—	3,089 (2.3%) — RR: 4.73%
INPOS-STAN / INPOST	—	3,164 (2.2%) — RR: 1.74%	—	—	—
PostN-Stan / PostNL	—	—	231,688 (94.5%) — RR: 3.20%	—	—
PostN-PU / PostNL	—	—	—	—	95,778 (72.9%) — RR: 5.33%
INPOS-PU / INPOST	—	—	—	17,106 (87.8%) — RR: 2.57%	—
SEUR-PU / SEUR	—	—	—	2,383 (12.2%) — RR: 4.53%	—
EXW-EXW / EXW	—	—	—	—	4,285 (3.3%) — RR: 99.60%
Shipping carrier					
COLISSIMO	58,839 (100.0%) — RR: 1.93%	—	—	—	28,295 (21.5%) — RR: 3.21%
DHLDE	—	124,254 (84.7%) — RR: 7.59%	—	—	—
GLS	—	19,327 (13.2%) — RR: 4.13%	—	—	3,089 (2.3%) — RR: 4.73%
PostNL	—	—	231,688 (94.5%) — RR: 3.20%	—	95,778 (72.9%) — RR: 5.33%
UPS	—	—	8,677 (3.5%) — RR: 2.96%	—	—
INPOST	—	3,164 (2.2%) — RR: 1.74%	—	17,106 (87.8%) — RR: 2.57%	—
SEUR	—	—	4,747 (1.9%) — RR: 3.16%	2,383 (12.2%) — RR: 4.53%	—
EXW	—	—	—	—	4,285 (3.3%) — RR: 99.60%
Shipment method					
Standard	58,839 (100.0%) — RR: 1.93%	146,744 (100.0%) — RR: 7.00%	245,108 (100.0%) — RR: 3.19%	—	—
Pickup point	—	1 (0.0%) — RR: 0.00%	—	19,489 (100.0%) — RR: 2.81%	127,162 (96.7%) — RR: 4.84%
EXW	—	—	—	—	4,285 (3.3%) — RR: 99.60%
Time-of-day					
Afternoon (13–18)	22,131 (37.6%) — RR: 1.96%	51,663 (35.2%) — RR: 6.82%	87,610 (35.7%) — RR: 2.95%	6,868 (35.2%) — RR: 2.66%	45,656 (34.7%) — RR: 6.84%
Evening (19–24)	15,833 (26.9%) — RR: 2.00%	42,263 (28.8%) — RR: 7.20%	76,809 (31.3%) — RR: 3.47%	5,906 (30.3%) — RR: 2.62%	42,579 (32.4%) — RR: 9.12%
Morning (7–12)	16,016 (27.2%) — RR: 1.84%	40,265 (27.4%) — RR: 6.87%	60,983 (24.9%) — RR: 3.05%	5,039 (25.8%) — RR: 2.98%	31,559 (24.0%) — RR: 6.85%
Night (0–6)	4,859 (8.3%) — RR: 1.87%	12,554 (8.6%) — RR: 7.54%	19,710 (8.0%) — RR: 3.63%	1,681 (8.6%) — RR: 3.57%	11,653 (8.9%) — RR: 10.82%
Numerical features					
Initial # items	3.18 (± 1.65)	3.83 (± 1.97)	3.13 (± 1.65)	3.26 (± 1.55)	3.23 (± 1.67)
Initial order amount (€)	64.37 (± 45.87)	81.37 (± 74.54)	65.99 (± 52.54)	210.84 (± 171.89)	67.36 (± 52.95)

Table G.3: Overview clusters LCC - orders - with shares and return rates

Scatterplots returned / not returned

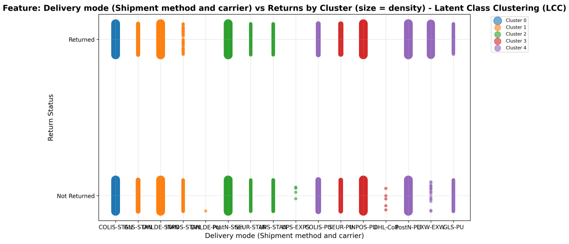


Figure G.22: Scatterplot delivery mode

Figure G.23: Scatterplot part of the day

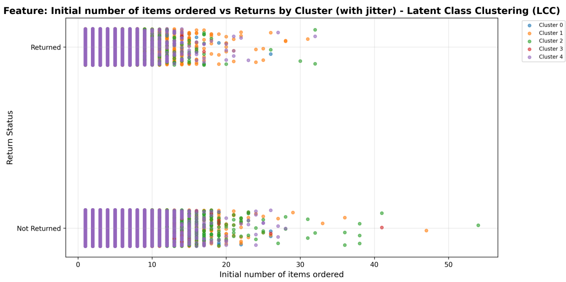


Figure G.24: Scatterplot initial number of items ordered

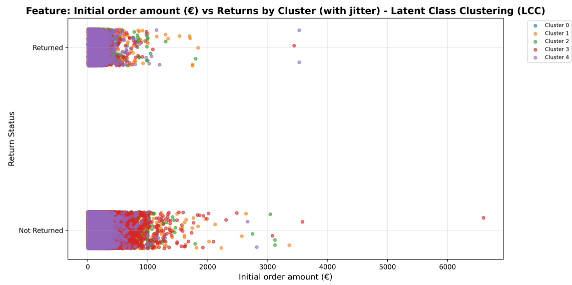


Figure G.25: Scatterplot initial order amount

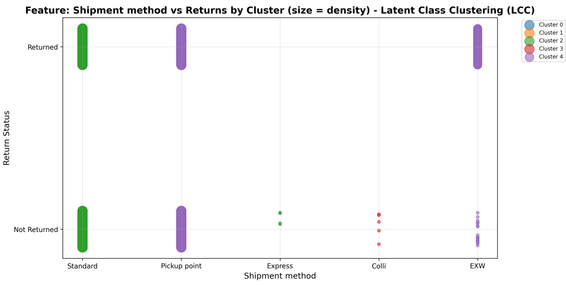


Figure G.26: Scatterplot shipment method

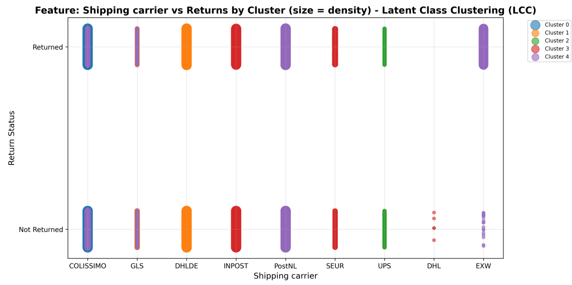


Figure G.27: Scatterplot shipping carrier

G.2. Product Level

G.2.1. K-prototyping

Overview clusters

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Products (Share)	127,258 (8.1%)	35,310 (2.3%)	1,000,487 (63.8%)	404,729 (25.8%)
Return rate	6.54%	7.22%	6.24%	8.38%
Product Category				
Cat L	68,686 (54.0%) — RR: 5.43%	25,371 (71.9%) — RR: 6.77%	122,933 (12.3%) — RR: 5.29%	139,265 (34.4%) — RR: 6.16%
Cat I	29,187 (22.9%) — RR: 6.93%	4,488 (12.7%) — RR: 9.09%	—	—
Cat H	7,053 (5.5%) — RR: 2.28%	1,353 (3.8%) — RR: 1.63%	377,729 (37.8%) — RR: 5.61%	34,681 (8.6%) — RR: 3.54%
Cat A	—	—	74,201 (7.4%) — RR: 5.81%	—
Cat F	—	—	—	66,363 (16.4%) — RR: 21.61%
Product Color				
Black	31,121 (24.5%) — RR: 8.22%	8,317 (23.6%) — RR: 7.20%	214,055 (21.4%) — RR: 7.90%	144,940 (35.8%) — RR: 12.01%
Unknown	27,651 (21.7%) — RR: 5.92%	4,977 (14.1%) — RR: 7.29%	493,875 (49.4%) — RR: 5.57%	86,218 (21.3%) — RR: 4.55%
Gold	12,921 (10.2%) — RR: 3.81%	—	—	—
Pink	—	5,765 (16.3%) — RR: 6.50%	81,011 (8.1%) — RR: 5.20%	—
Purple	—	—	—	44,261 (10.9%) — RR: 6.13%
Product Size				
Unknown	123,105 (96.7%) — RR: 5.90%	35,215 (99.7%) — RR: 7.19%	941,438 (94.1%) — RR: 5.69%	341,028 (84.3%) — RR: 5.82%
Small	1,246 (1.0%) — RR: 24.72%	—	15,526 (1.6%) — RR: 15.61%	19,087 (4.7%) — RR: 23.06%
Large	1,057 (0.8%) — RR: 27.81%	—	—	14,851 (3.7%) — RR: 23.28%
One Size	—	—	20,368 (2.0%) — RR: 12.64%	—
Medium	—	30 (0.1%) — RR: 23.33%	—	—
Extra Large	—	25 (0.1%) — RR: 12.00%	—	—
Time-of-day				
Afternoon (13–18)	45,384 (35.7%) — RR: 6.08%	12,479 (35.3%) — RR: 6.82%	352,069 (35.2%) — RR: 5.73%	143,420 (35.4%) — RR: 7.82%
Evening (19–24)	39,157 (30.8%) — RR: 6.80%	10,196 (28.9%) — RR: 7.27%	313,642 (31.3%) — RR: 6.70%	126,632 (31.3%) — RR: 8.94%
Morning (7–12)	31,956 (25.1%) — RR: 6.31%	9,515 (26.9%) — RR: 7.08%	250,073 (25.0%) — RR: 5.85%	101,507 (25.1%) — RR: 8.15%
Night (0–6)	10,761 (8.5%) — RR: 8.28%	3,120 (8.8%) — RR: 9.10%	84,703 (8.5%) — RR: 7.75%	33,170 (8.2%) — RR: 9.42%
Numerical features				
Product price (€)	24.68 ($\pm$ 5.54)	59.74 ($\pm$ 34.14)	2.37 ($\pm$ 1.61)	9.47 ($\pm$ 2.78)
Ordered quantity of product	1.05 ($\pm$ 0.39)	1.06 ($\pm$ 0.45)	1.02 ($\pm$ 0.22)	1.04 ($\pm$ 0.27)

Table G.4: Overview clusters K-prototyping - products - with shares and return rates

Distributions across clusters

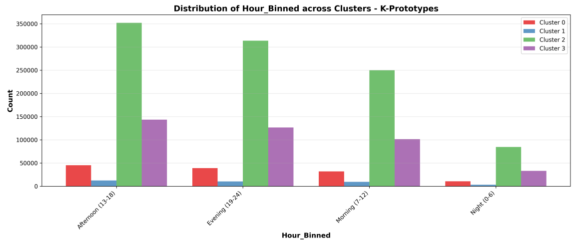


Figure G.28: Distribution of time of day in clusters

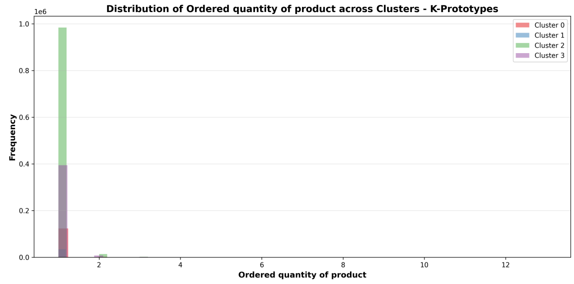


Figure G.29: Distribution of ordered quantities in clusters

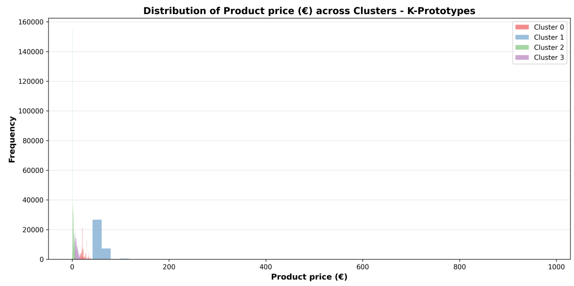


Figure G.30: Distribution of product price in clusters

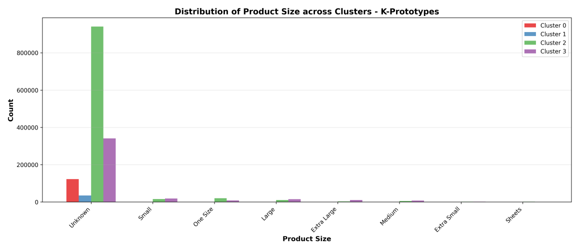
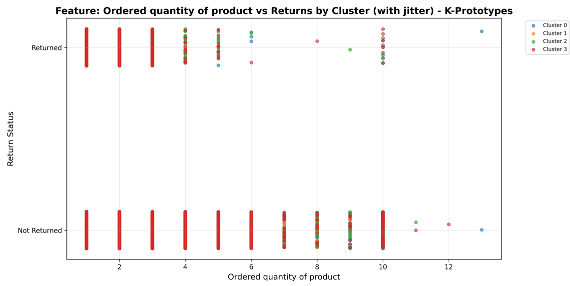
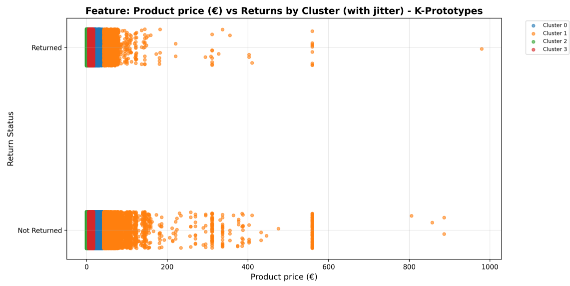


Figure G.31: Distribution of Product Size in clusters

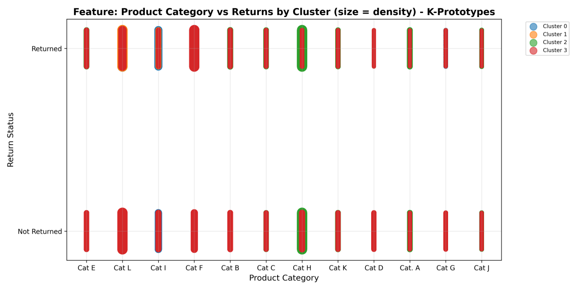
Scatterplots returned / not returned



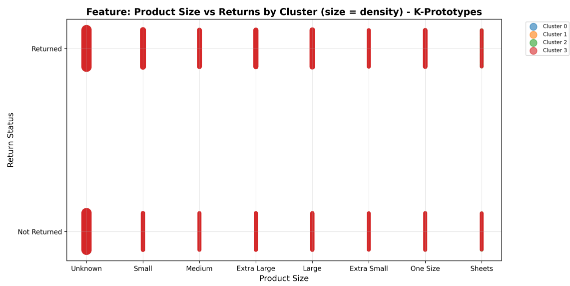
(a) Scatterplot ordered quantities



(b) Scatterplot product price



(c) Scatterplot Product categories



(d) Scatterplot Product size

G.2.2. CAVE

Overview clusters

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Products (share)	547,884 (34.9%)	749,463 (47.8%)	223,976 (14.3%)	46,461 (3.0%)
Return rate	7.45%	6.04%	7.96%	7.06%
Product Category				
Cat L	164,380 (30.0%) — RR: 5.78%	65,523 (8.7%) — RR: 5.08%	97,931 (43.7%) — RR: 5.87%	28,421 (61.2%) — RR: 6.88%
Cat H	99,868 (18.2%) — RR: 4.10%	304,572 (40.6%) — RR: 5.93%	14,189 (6.3%) — RR: 2.73%	2,187 (4.7%) — RR: 1.83%
Cat F	77,047 (14.1%) — RR: 18.72%	—	25,659 (11.5%) — RR: 23.92%	—
Cat E	46,295 (8.4%) — RR: 7.40%	—	—	—
Cat A	—	68,829 (9.2%) — RR: 5.86%	—	—
Cat K	—	59,953 (8.0%) — RR: 4.85%	—	—
Cat I	—	—	49,021 (21.9%) — RR: 6.02%	5,708 (12.3%) — RR: 8.67%
Cat G	—	—	—	5,498 (11.8%) — RR: 6.38%
Product Color				
Black	166,184 (30.3%) — RR: 10.73%	149,115 (19.9%) — RR: 7.15%	73,232 (32.7%) — RR: 11.21%	9,902 (21.3%) — RR: 7.63%
Unknown	157,418 (28.7%) — RR: 4.47%	404,503 (54.0%) — RR: 5.82%	39,813 (17.8%) — RR: 5.37%	10,987 (23.6%) — RR: 6.20%
Pink	50,707 (9.3%) — RR: 5.69%	62,534 (8.3%) — RR: 5.07%	—	5,923 (12.7%) — RR: 6.62%
Purple	48,890 (8.9%) — RR: 6.42%	36,851 (4.9%) — RR: 5.32%	19,218 (8.6%) — RR: 6.56%	—
Blue	—	36,851 (4.9%) — RR: 5.32%	—	—
Magenta	—	—	—	5,375 (11.6%) — RR: 7.57%
Product Size				
Unknown	473,004 (86.3%) — RR: 5.63%	721,324 (96.2%) — RR: 5.76%	200,210 (89.4%) — RR: 5.92%	46,248 (99.5%) — RR: 6.98%
Small	24,495 (4.5%) — RR: 19.83%	5,201 (0.7%) — RR: 13.59%	6,141 (2.7%) — RR: 25.52%	—
Large	17,216 (3.1%) — RR: 20.79%	4,650 (0.6%) — RR: 13.96%	5,037 (2.2%) — RR: 26.66%	59 (0.1%) — RR: 28.81%
One Size	16,174 (3.0%) — RR: 14.48%	11,577 (1.5%) — RR: 11.63%	—	—
Extra Large	—	—	5,247 (2.3%) — RR: 24.57%	45 (0.1%) — RR: 33.33%
Medium	—	—	—	59 (0.1%) — RR: 25.42%
Time-of-day				
Afternoon (13–18)	192,884 (35.2%) — RR: 6.95%	264,245 (35.3%) — RR: 5.54%	79,719 (35.6%) — RR: 7.37%	16,504 (35.5%) — RR: 6.59%
Evening (19–24)	172,203 (31.4%) — RR: 7.93%	234,750 (31.3%) — RR: 6.48%	69,168 (30.9%) — RR: 8.53%	13,506 (29.1%) — RR: 7.32%
Morning (7–12)	137,327 (25.1%) — RR: 7.16%	187,068 (25.0%) — RR: 5.66%	56,212 (25.1%) — RR: 7.70%	12,444 (26.8%) — RR: 6.86%
Night (0–6)	45,470 (8.3%) — RR: 8.63%	63,400 (8.5%) — RR: 7.67%	18,877 (8.4%) — RR: 9.17%	4,007 (8.6%) — RR: 8.76%
Numerical features				
Product price (€)	6.52 ($\pm$ 2.11)	1.60 ($\pm$ 0.98)	18.59 ($\pm$ 6.15)	54.13 ($\pm$ 31.40)
Ordered quantity	1.03 ($\pm$ 0.25)	1.02 ($\pm$ 0.21)	1.05 ($\pm$ 0.36)	1.06 ($\pm$ 0.47)

Table G.5: Overview clusters CAVE — products — with shares and return rates

Distributions across clusters

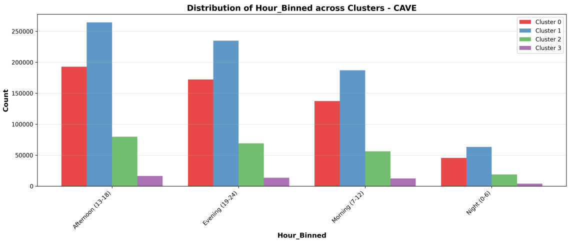


Figure G.33: Distribution of part of the days in clusters

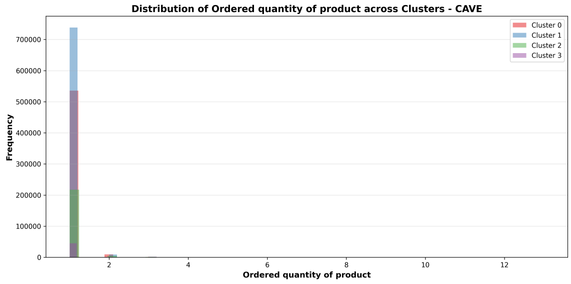


Figure G.34: Distribution of ordered quantities in clusters

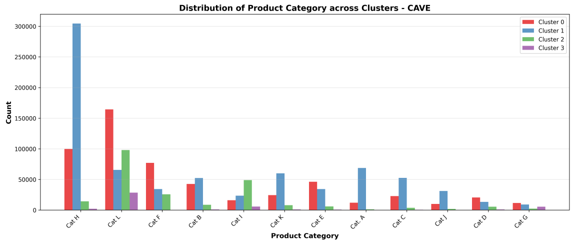


Figure G.35: Distribution of Product categories in clusters

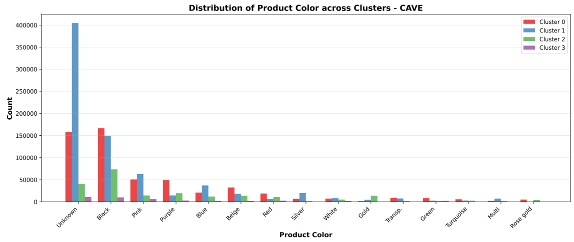


Figure G.36: Distribution of Product Colors in clusters

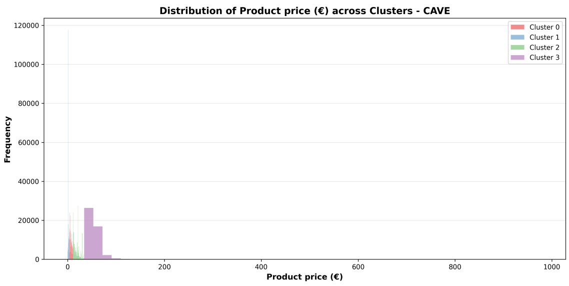


Figure G.37: Distribution of product price in clusters

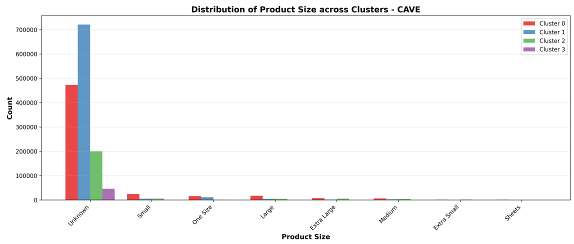
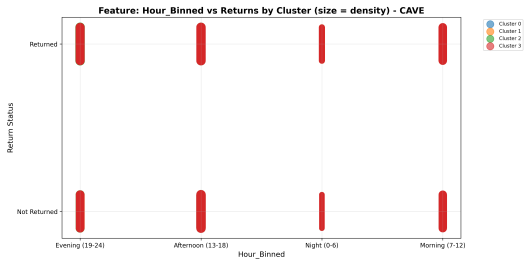
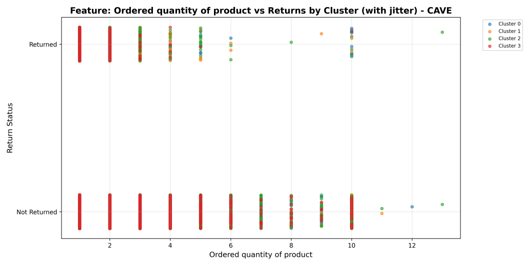


Figure G.38: Distribution of Product Size in clusters

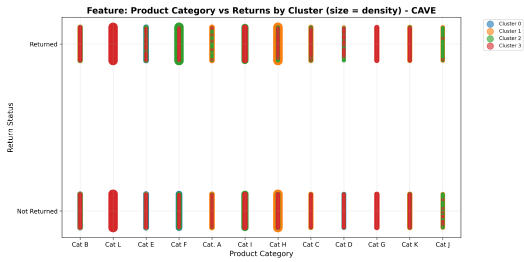
Scatterplots returned / not returned



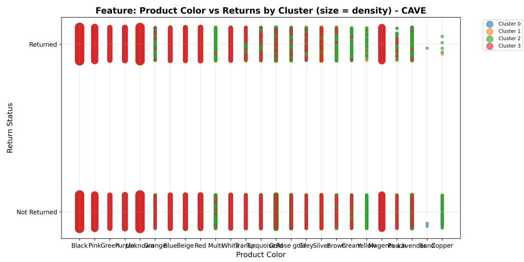
(a) Scatterplot part of the days



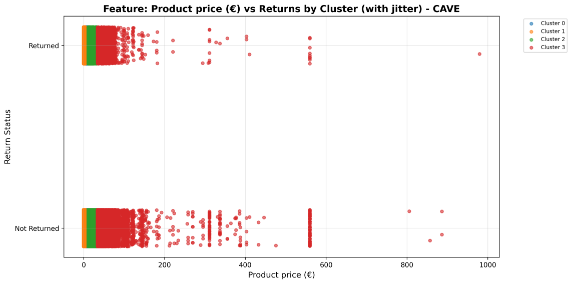
(b) Scatterplot ordered quantities



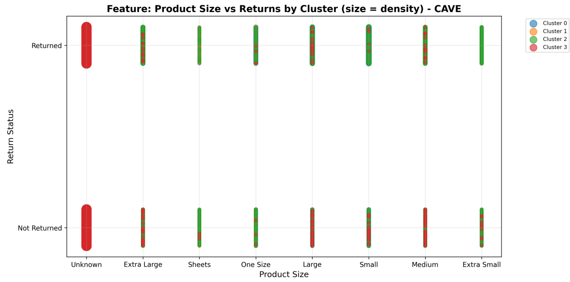
(c) Scatterplot Product categories



(d) Scatterplot Product Colors



(a) Scatterplot product price



(b) Scatterplot Product Size

G.2.3. LCC

Distributions across clusters

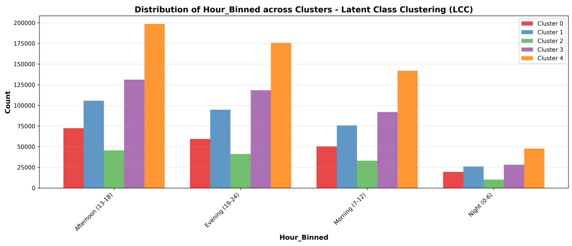


Figure G.41: Distribution of part of the days in clusters

Figure G.42: Distribution of ordered quantities in clusters

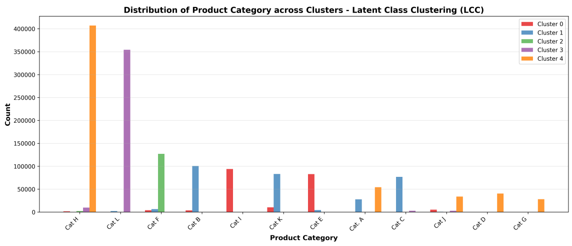


Figure G.43: Distribution of Product categories in clusters

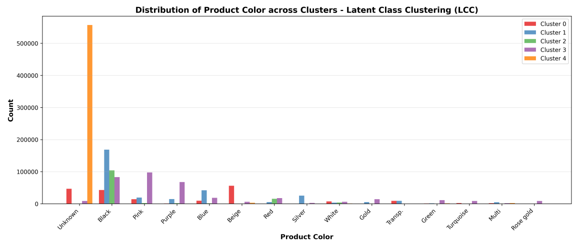


Figure G.44: Distribution of Product Colors in clusters

Figure G.45: Distribution of product price in clusters

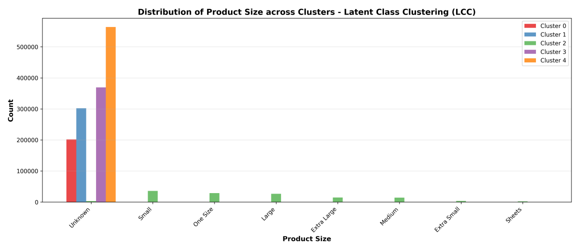


Figure G.46: Distribution of Product Size in clusters

Overview clusters

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Products (share)	201,866 (12.9%)	302,270 (19.3%)	130,048 (8.3%)	369,593 (23.6%)	564,007 (36.0%)
Return rate	6.40%	6.24%	18.63%	5.68%	5.36%
Product Category					
Cat I	94,080 (46.6%) — RR: 5.80%	—	—	—	—
Cat E	82,755 (41.0%) — RR: 7.16%	—	—	—	—
Cat K	10,415 (5.2%) — RR: 5.58%	83,002 (27.5%) — RR: 5.03%	—	—	—
Cat B	—	100,626 (33.3%) — RR: 6.91%	—	—	—
Cat C	—	76,754 (25.4%) — RR: 6.56%	—	2,759 (0.7%) — RR: 7.39%	—
Cat A	—	—	—	—	54,317 (9.6%) — RR: 5.42%
Cat F	—	—	127,169 (97.8%) — RR: 18.91%	—	—
Cat L	—	—	—	354,056 (95.8%) — RR: 5.76%	—
Cat H	—	—	2,089 (1.6%) — RR: 5.31%	9,710 (2.6%) — RR: 2.14%	407,058 (72.2%) — RR: 5.46%
Cat J	—	—	495 (0.4%) — RR: 5.25%	—	—
Cat D	—	—	—	—	40,551 (7.2%) — RR: 3.89%
Product Color					
Beige	56,285 (27.9%) — RR: 7.15%	—	—	—	2,938 (0.5%) — RR: 3.40%
Unknown	46,883 (23.2%) — RR: 6.31%	—	—	—	556,926 (98.7%) — RR: 5.37%
Black	42,664 (21.1%) — RR: 5.94%	168,735 (55.8%) — RR: 6.52%	103,862 (79.9%) — RR: 18.88%	83,172 (22.5%) — RR: 5.20%	—
Blue	—	42,036 (13.9%) — RR: 5.40%	—	18,540 (5.0%) — RR: 5.69%	—
Pink	—	—	2,199 (1.7%) — RR: 15.60%	97,712 (26.4%) — RR: 5.19%	—
Red	—	—	15,313 (11.8%) — RR: 18.79%	—	—
White	—	—	4,056 (3.1%) — RR: 15.29%	—	—
Beige	—	—	—	—	2,938 (0.5%) — RR: 0.00%
Multi	—	—	—	—	2,034(0.4%) — RR: 0.00%
Product Size					
Unknown	201,834 (100.0%) — RR: 6.40%	302,270 (100.0%) — RR: 6.24%	—	369,470 (100.0%) — RR: 5.68%	563,998 (100.0%) — RR: 5.36%
Small	—	—	35,863 (27.6%) — RR: 19.90%	—	4 (0.0%) — RR: 0.00%
One Size	—	—	28,882 (22.2%) — RR: 13.46%	123 (0.0%) — RR: 4.07%	—
Large	32 (0.0%) — RR: 25.00%	—	26,930 (20.7%) — RR: 20.72%	—	—
Extra Large	—	—	—	—	4 (0.0%) — RR: 0.00%
Time-of-day					
Afternoon (13–18)	72,494 (35.9%) — RR: 5.70%	105,702 (35.0%) — RR: 5.68%	45,556 (35.0%) — RR: 17.85%	131,096 (35.5%) — RR: 5.29%	198,504 (35.2%) — RR: 4.93%
Evening (19–24)	59,408 (29.4%) — RR: 7.07%	94,824 (31.4%) — RR: 6.57%	41,258 (31.7%) — RR: 19.57%	118,358 (32.0%) — RR: 5.96%	175,779 (31.2%) — RR: 5.81%
Morning (7–12)	50,420 (25.0%) — RR: 5.82%	75,668 (25.0%) — RR: 5.87%	33,082 (25.4%) — RR: 18.56%	91,884 (24.9%) — RR: 5.45%	141,997 (25.2%) — RR: 4.98%
Night (0–6)	19,544 (9.7%) — RR: 8.49%	26,076 (8.6%) — RR: 8.36%	10,152 (7.8%) — RR: 18.52%	28,255 (7.6%) — RR: 7.10%	47,727 (8.5%) — RR: 6.58%
Numerical features					
Product price (€)	11.23 ($\pm$ 14.40)	4.23 ($\pm$ 11.40)	7.35 ($\pm$ 4.96)	13.48 ($\pm$ 14.75)	3.49 ($\pm$ 5.88)
Ordered quantity	1.01 ($\pm$ 0.16)	1.01 ($\pm$ 0.13)	1.02 ($\pm$ 0.22)	1.01 ($\pm$ 0.13)	1.06 ($\pm$ 0.38)

Table G.6: Overview clusters LCC - products - with shares and return rates

Scatterplots returned / not returned

Figure G.47: Scatterplot part of the days

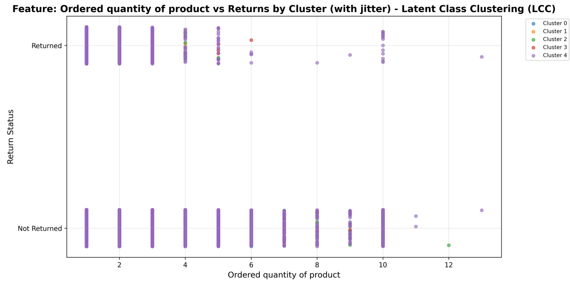


Figure G.48: Scatterplot ordered quantities

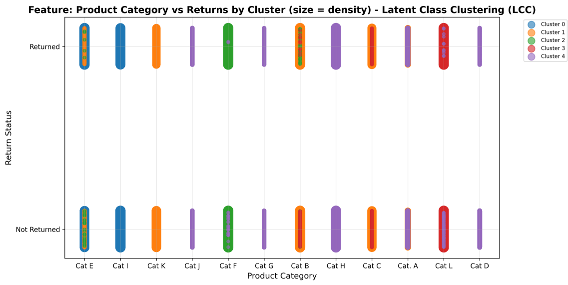


Figure G.49: Scatterplot Product categories

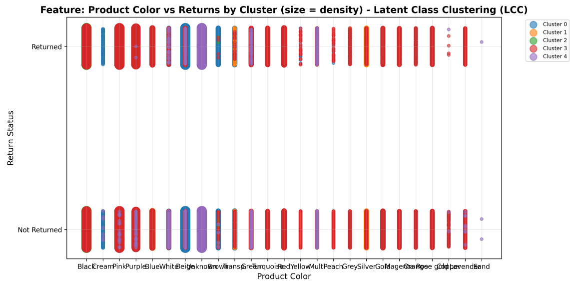


Figure G.50: Scatterplot Product Colors

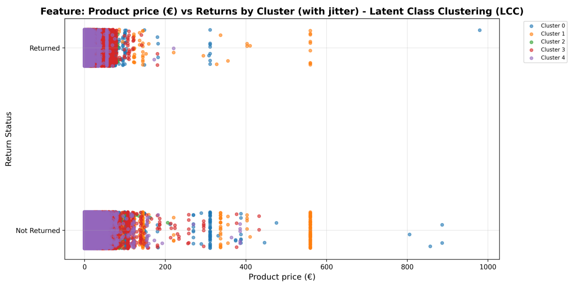


Figure G.51: Scatterplot product price

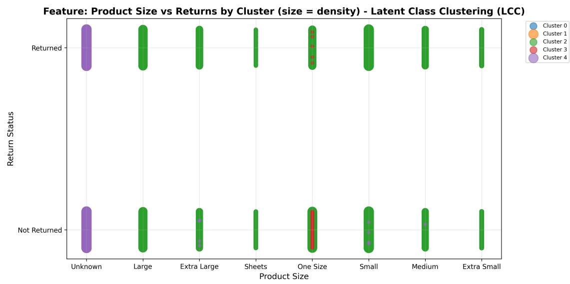


Figure G.52: Scatterplot Product Size

G.3. Extra Cluster tables

Cluster	Ass. data-points	Percentage of total	Return rate	Cluster description	Risk label
F.0	30977	22.5%	16.68%	Low priced, 1 item, One Size/Unknown	Low
F.1	26570	19.3%	15.14%	Low priced, 1 item, Small & One Size	Low
F.2	7002	5.1%	8.18%	High priced, 1,5 item, Beige, Size unknown	Low
F.3	51934	37.8%	23.34%	High priced, 1 item, Black & Red, S & L	High
F.4	20951	15.2%	19.78%	Average priced, 1 item, Black & Red, S & L	High

Note: Overall return rate of Cluster 2: 18.6%

Table G.7: Summary of clusters of Cat. F - products

Feature	Cluster 2.0	Cluster 2.1	Cluster 2.2	Cluster 2.3	Cluster 2.4
Products (share)	24,054 (18.5%) — RR: 13.84%	21,202 (16.3%) — RR: 21.31%	20,202 (15.5%) — RR: 15.51%	20,178 (15.5%) — RR: 18.04%	44,412 (34.2%) — RR: 21.63%
Product Category					
Cat F	23,884 (99.3%) — RR: 13.89%	19,294 (91.0%) — RR: 22.94%	20,137 (99.7%) — RR: 15.51%	20,084 (99.5%) — RR: 18.09%	43,770 (98.6%) — RR: 21.81%
Cat H	140 (0.6%) — RR: 4.29%	1,413 (6.7%) — RR: 4.74%	9 (0.0%) — RR: 33.33%	63 (0.3%) — RR: 4.76%	464 (1.0%) — RR: 6.90%
Cat E / Cat B / Cat J	30 (0.1%) — RR: 16.67%	495 (2.3%) — RR: 5.25%	56 (0.3%) — RR: 12.50%	31 (0.2%) — RR: 9.68%	177 (0.4%) — RR: 14.69%
Product Color					
Black	23,114 (96.1%) — RR: 13.93%	16,497 (77.8%) — RR: 22.78%	18,273 (90.5%) — RR: 15.57%	11,980 (59.4%) — RR: 19.32%	33,998 (76.6%) — RR: 21.98%
Red	428 (1.8%) — RR: 12.85%	2,230 (10.5%) — RR: 15.74%	923 (4.6%) — RR: 16.03%	4,452 (22.1%) — RR: 15.79%	10,157 (22.9%) — RR: 20.63%
White / Pink / Gold / Purple / Multi	246 (1.0%) — RR: 12.60%	1,729 (8.2%) — RR: 15.73%	288 (1.4%) — RR: 19.44%	1,806 (9.0%) — RR: 16.67%	193 (0.4%) — RR: 14.51%
Product Size					
One Size	20,142 (83.7%) — RR: 13.52%	—	7,417 (36.7%) — RR: 12.79%	—	—
Small	—	—	7,431 (36.8%) — RR: 16.61%	8,829 (43.8%) — RR: 18.79%	18,025 (40.6%) — RR: 22.07%
Medium	2,291 (9.5%) — RR: 15.58%	8,612 (40.6%) — RR: 22.86%	—	—	—
Large	—	—	4,615 (22.8%) — RR: 18.40%	6,503 (32.2%) — RR: 18.91%	14,563 (32.8%) — RR: 22.99%
Extra Large	—	6,069 (28.6%) — RR: 24.24%	—	1,971 (9.8%) — RR: 21.56%	6,711 (15.1%) — RR: 19.36%
Sheets	1,621 (6.7%) — RR: 15.24%	—	—	—	—
Extra Small	—	1,745 (8.2%) — RR: 26.59%	—	—	—
Time-of-day					
Afternoon (13–18)	8,203 (34.1%) — RR: 12.36%	7,310 (34.5%) — RR: 20.29%	7,016 (34.7%) — RR: 14.50%	7,162 (35.5%) — RR: 18.26%	15,865 (35.7%) — RR: 20.86%
Evening (19–24)	7,725 (32.1%) — RR: 15.57%	6,476 (30.5%) — RR: 22.24%	6,366 (31.5%) — RR: 15.87%	6,381 (31.6%) — RR: 18.43%	14,310 (32.2%) — RR: 22.68%
Morning (7–12)	6,044 (25.1%) — RR: 13.27%	6,076 (28.7%) — RR: 21.36%	5,193 (25.7%) — RR: 16.56%	5,078 (25.2%) — RR: 17.29%	10,691 (24.1%) — RR: 21.54%
Night (0–6)	2,082 (8.7%) — RR: 14.84%	1,340 (6.3%) — RR: 22.24%	1,627 (8.1%) — RR: 15.18%	1,557 (7.7%) — RR: 17.85%	3,546 (8.0%) — RR: 21.09%
Numerical features					
Product price (€)	4.77 (±3.12)	12.26 (±6.10)	4.17 (±0.80)	6.20 (±0.95)	8.37 (±5.18)
Ordered quantity	1.01 (±0.14)	1.04 (±0.43)	1.01 (±0.11)	1.01 (±0.11)	1.01 (±0.19)

Table G.8: Overview clusters of Cluster 2 - LCC– products

Table G.9: Overview of clusters and order features in clusters, clustered on product level

Feature	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Number of products	201,866 (12.9%)	302,270 (19.3%)	130,048 (8.3%)	369,593 (23.6%)	564,007 (36.0%)
Unique orders	165,156	193,156	68,923	279,886	337,474
Return rate	6.4%	6.2%	18.6%	5.7%	5.4%
Items Ordered (binned)	Medium (43%) RR: 5.75% Large (30.9%) RR: 7.84% Small (25.9%) RR: 5.46%	Large (47%) RR: 7.56% Medium (40.7%) RR: 5.05% Small (12.1%) RR: 4.60%	Large (48.1%) RR: 23.96% Medium (38.7%) RR: 13.46% Small (12.2%) RR: 11.67%	Medium (41%) RR: 5.25% Small (30.1%) RR: 4.61% Large (28.7%) RR: 7.08%	Medium (44.6%) RR: 4.55% Large (44.2%) RR: 6.52% Small (11.1%) RR: 3.42%
Amount Ordered (binned)	Medium+ (25.8%) RR: 5.98% High (25%) RR: 8.69% Low (25%) RR: 5.00% Medium- (24.1%) RR: 5.59%	High (26.1%) RR: 8.92% Medium+ (25.1%) RR: 6.43% Medium- (24.8%) RR: 5.08% Low (23.8%) RR: 4.06%	High (32.7%) RR: 29.84% Medium+ (27.4%) RR: 16.87% Medium- (23.1%) RR: 11.27% Low (15.9%) RR: 7.44%	High (27.4%) RR: 7.42% Medium+ (24.7%) RR: 5.83% Medium- (24.4%) RR: 4.68% Low (23.3%) RR: 4.09%	Low (28.8%) RR: 3.78% Medium- (26.3%) RR: 4.60% Medium+ (23.9%) RR: 5.75% High (20.9%) RR: 7.75%
Delivery mode (method + carrier)	PostN-Stan: 33.9% RR: 3.91% DHLDE-STAN: 19.5% RR: 8.01% PostN-PU: 18.7% RR: 5.81% COLIS-STAN: 9.9% RR: 2.84%	PostN-Stan: 38.1% RR: 4.05% DHLDE-STAN: 19.7% RR: 9.48% PostN-PU: 18.3% RR: 6.36% COLIS-STAN: 8.5% RR: 2.76%	PostN-Stan: 37.1% RR: 13.05% DHLDE-STAN: 26.8% RR: 33.13% PostN-PU: 14.0% RR: 14.45% COLIS-STAN: 9.2% RR: 5.93%	PostN-Stan: 37.7% RR: 3.86% DHLDE-STAN: 21.1% RR: 8.48% PostN-PU: 14.0% RR: 5.86% COLIS-STAN: 11.1% RR: 2.70%	PostN-Stan: 36.0% RR: 3.05% DHLDE-STAN: 25.4% RR: 7.68% PostN-PU: 15.0% RR: 5.15% COLIS-STAN: 8.5% RR: 2.14%
Shipping carrier	PostNL: 52.6% RR: 4.58% DHLDE: 19.5% RR: 8.01% COLISSIMO: 16.6% RR: 3.38% GLS: 3.6% RR: 4.60%	PostNL: 56.4% RR: 4.80% DHLDE: 19.7% RR: 9.48% COLISSIMO: 13.2% RR: 3.24% GLS: 3.8% RR: 5.84%	PostNL: 51.2% RR: 13.43% DHLDE: 26.8% RR: 33.13% COLISSIMO: 13.1% RR: 6.29% GLS: 4.1% RR: 17.75%	PostNL: 51.7% RR: 4.40% DHLDE: 21.1% RR: 8.48% COLISSIMO: 15.9% RR: 3.08% GLS: 4.1% RR: 4.96%	PostNL: 51.0% RR: 3.67% DHLDE: 25.4% RR: 7.68% COLISSIMO: 12.7% RR: 2.63% GLS: 4.5% RR: 4.83%
Shipment method	Standard: 69.1% RR: 4.94% Pickup point: 29.4% RR: 5.24%	Standard: 72.5% RR: 5.45% Pickup point: 26.6% RR: 5.66%	Standard: 79.0% RR: 19.19% Pickup point: 19.5% RR: 12.41%	Standard: 76.4% RR: 4.99% Pickup point: 22.9% RR: 5.08%	Standard: 77.0% RR: 4.58% Pickup point: 22.1% RR: 4.66%

Order Size Ranges: 1–2 (Small); 3–4 (Medium); 5+ (Large)

Order Amount bins: Low €0–49.94; Medium– €49.94–72.96; Medium+ €72.96–111.35; High €111.35–6600.56