Beyond Who the Customer Is: Behavioral Drivers of Value in Fashion Retail

An Empirical Study of How Behavioral Patterns and Predefined Customer Segments Explain Visit-Level Value Creation

Authors

Johan Danha

Teo Paganus

Supervisor

Alexander Mafael

Stockholm School of Economics

Retail Management Program

Abstract

This study examines how value is created in physical retail by comparing customer composition and in-store behavior as explanatory factors. While retailers commonly rely on Ideal Customer Profiles (ICPs) to identify high-value customers, it remains unclear whether these predefined segments are empirically linked to value creation within the store. Using a single-case study centered on one focal fashion retail store, the analysis combines aggregated in-store tracking data from Indivd with structured observational data from comparable retail environments. Due to the anonymized and visit-based nature of the data, value is operationalized at the visit level using conversion and value per visitor as outcome measures. The findings show that ICP-based customer composition provides a useful descriptive framework for understanding store performance but has limited explanatory power in predicting value outcomes. In contrast, behavioral patterns, particularly those reflecting active evaluation such as product interaction, cognitive engagement and fitting room usage, show more direct and consistent associations with purchase and estimated value outcomes. Behaviors indicating friction are negatively associated with purchase likelihood, while general exploration shows no significant association. These results suggest that visit-level value creation in this case is more directly associated with how customers behave during the store visit than with predefined customer composition alone. The study therefore highlights the importance of complementing traditional segmentation approaches with behavioral analysis and demonstrates how behavioral data can be used to inform retail decision-making.

Keywords: In-Store Behavior, Value Creation, Ideal Customer Profile, Fashion Retail, Customer Behavior

Acknowledgements

We would like to express our sincere gratitude to everyone who supported us throughout the thesis process. We especially thank our supervisor, *Alexander Mafael*, for his guidance, insightful feedback and continuous support. His availability and willingness to engage in discussions have been invaluable. We also thank Indivd for providing the data that made this study possible and in particular *Fredrik Amréus Hammargården*, CEO of Indivd, for his valuable insights and support. Finally, we thank the stores that allowed us to conduct observations in their environments, whose openness made an important part of this study possible.

Table of Contents

1. Introduction

1.1 Background

The physical retail environment plays a central role in shaping customer behavior and influencing purchase outcomes (Bitner, 1992; Donovan et al., 1994). Traditionally, retailers have focused on increasing store traffic and optimizing overall performance metrics such as footfall and conversion. However, it is increasingly recognized that not all visitors contribute equally to value creation (Gupta et al., 2004). As a result, understanding what characterizes high-value customers, and how they behave in-store, has become an important strategic question.

A common approach to this challenge is the use of Ideal Customer Profiles (ICPs), which aim to define the retailer's most valuable customer segments. These profiles are typically used to guide both marketing efforts and in-store decision-making (Wedel & Kamakura, 2000). However, despite their widespread use, ICPs are rarely evaluated empirically in physical retail environments. In particular, it remains unclear whether these predefined customer segments are actually associated with higher value outcomes in practice.

At the same time, advances in in-store tracking technologies have made it possible to observe customer behavior at a much more detailed level (Hui et al., 2009a; Larson et al., 2005). This has shifted attention towards how customers interact with the retail environment, through movement, attention and engagement, as a potential driver of value (Puccinelli et al., 2009; Chandon et al., 2009). This raises a key question: is value in physical retail primarily determined by who the customer is or by how the customer behaves?

This distinction can be clarified by how the two perspectives are approached in this study. The "who the customer is" perspective is represented through the retailer's Ideal Customer Profile (ICP), which defines customer segments considered strategically valuable, but is based on assumptions rather than direct observation of value creation within the store. In contrast, the "how the customer behaves" perspective focuses on how customers interact with the retail environment through movement, attention and engagement. These observable actions provide insight into how customers progress through the decision process and therefore offer a more direct lens for analyzing how value is created during the visit. Building on this, the study examines how these two perspectives relate to visit-level value outcomes using a combination of aggregated tracking data and structured observations.

1.2 Purpose and Research Question

The purpose of this study is to compare customer composition and in-store behavior as explanations of value creation in a physical retail setting. Customer composition refers to the distribution of visitors across predefined customer segments, such as the retailer's Ideal Customer Profile (ICP), within the store. Specifically, the study evaluates whether the retailer's current ICP definition is associated with observed visit-level value outcomes and whether behavioral patterns provide a stronger explanation of value creation. Rather than treating ICP as a fixed representation of high-value customers, it is approached as a testable construct. This allows the study to move beyond descriptive segmentation and instead assess how well predefined customer groups align with observed value outcomes.

A key challenge in this context is that individual-level customer value cannot be directly observed. In physical retail settings, data is typically anonymized and visit-based, making it difficult to link behavior to long-term measures such as Customer Lifetime Value (CLV). This requires a shift in analytical focus from high-value customers to high-value visits, where value is assessed based on observable outcomes within each store visit. In the empirical analysis, different outcome measures are used depending on the analytical perspective. Conversion is used to evaluate whether predefined customer segments are associated with purchase outcomes, while value per visitor is used to examine how behavioral patterns relate to value creation. While these measures are related, value per visitor reflects an estimated proxy based on the retailer's internal valuation model.

The study is guided by the following research question:

- *To what extent do customer composition and in-store behavior explain visit-level value creation in a physical fashion retail setting?*

In addressing this question, the study examines not only whether ICP-aligned visitors convert into purchases, but also how behavioral patterns, such as engagement and friction, are associated with value outcomes. This allows for an empirical comparison between customer composition and behavior as explanatory factors of value.

1.3 Contribution

The primary contribution of this thesis is to demonstrate how value creation in physical retail can be analyzed using behavioral data at the level of the store visit. By shifting the analytical focus from predefined customer segments to observable in-store behavior, the study reframes value as something that emerges through how customers interact with the retail environment, rather than as a fixed characteristic of specific customer groups. Building on this, the study provides an empirical comparison between customer composition and in-store behavior as explanatory factors of value. By treating the retailer's Ideal Customer Profile (ICP) as a testable construct, rather than a predefined representation of high-value customers, the analysis evaluates how well ICP-aligned visitors are associated with observed value outcomes. The findings show that while ICP provides a useful descriptive perspective, behavioral patterns offer a stronger and more consistent explanation of value creation. In addition, the study illustrates how retailers can analyze in-store performance under their existing ICP definition through the use of a zone scorecard. By linking customer composition to zone-level outcomes, the analysis shows how different areas of the store attract and convert target customer segments, providing a structured way to identify high-performing zones and areas for improvement.

Importantly, the study does not aim to establish causal relationships or provide generalizable conclusions beyond the specific case. Instead, its contribution lies in demonstrating how behavioral data can be used to analyze and interpret value creation in a physical retail context. The applicability of this approach to other retail settings is therefore positioned as an avenue for future research.

2. Literature Review

2.1 From Segmentation to Behavioral Analysis

Customer segmentation has traditionally been based on observable and identifiable characteristics such as demographics, purchase history and long-term value measures like Customer Lifetime Value (CLV) (Gupta et al., 2004; Wedel & Kamakura, 2000). These approaches are widely used to identify high-value customer segments and guide strategic decision-making. However, while segmentation provides a basis for identifying target customers, it does not explain how value is created within the store. In particular, there is limited understanding of how predefined customer segments actually behave in the physical retail environment and how such insights can be translated into concrete in-store design decisions. Existing research has primarily focused on identifying valuable customers, rather than on linking these segments to observable in-store behavior and actionable design implications (Hosseini & Shabani, 2015; Tsiptsis & Chorianopoulos, 2010).

This limitation is further reinforced by the nature of data in physical retail. Customer data is typically anonymized and recorded at the level of the visit, making it difficult to directly connect individual characteristics to observed behavior and value outcomes. As a result, segmentation approaches such as ICP are typically based on transactional or longitudinal data, rather than on direct observation of how value is generated within individual store visits. At the same time, advances in in-store tracking technologies have made it possible to observe how customers move, interact and engage with the retail environment in real time (Hui et al., 2009a; Larson et al., 2005). These technologies enable a shift from focusing on who the customer is to analyzing how the customer behaves during the store visit. This creates an opportunity to move beyond segmentation as a purely descriptive tool and instead examine how value is generated through observable in-store behavior. This shift forms the starting point for the analysis developed in the following sections.

2.2 Store Environment and Behavior as Value Signals

Building on the shift from segmentation to behavioral analysis, the next step is to understand what shapes customer behavior in the store environment. Customer behavior in a store does not occur randomly but is shaped by how the store is designed. Elements such as layout, product placement and service points influence where customers go, what they interact with and how they move through the store. This relationship is well established in retail research.

Bitner (1992) shows that the physical environment affects customer responses, which in turn influence behavior, while Donovan et al. (1994) demonstrate that positive in-store experiences are associated with increased time spent in-store and a higher likelihood of purchase. Together, these studies support the idea that store design plays a direct role in shaping customer actions.

For the purpose of this study, however, the key point is not only that the environment influences behavior, but that behavior can be interpreted as providing insight into how customers make decisions. Observable actions, such as where customers spend time, whether they interact with products or whether they return to the same area, can therefore be understood as indicators of how they are progressing through the purchase process. This is particularly important given the nature of the available data. Since customer value cannot be directly observed at the individual level, value must be inferred from observable actions during the visit. Behavior therefore becomes the primary source of information for understanding value creation within the store. This defines the analytical focus of the study. Given this limitation, the analysis relies on observable in-store behavior at the level of individual visits as a proxy for value creation.

While this section establishes that behavior is shaped by the store environment, it does not yet explain how behavior should be interpreted in relation to value creation. The following section therefore examines how observable behaviors can be understood as signals of underlying decision processes.

2.3 In-Store Behavior as Decision Signals

Building on this, the focus now shifts to how observable in-store behavior can be interpreted as signals of customers' decision-making processes. The key question is not only what customers do but what their behavior indicates about their likelihood to purchase. For behavior to be useful in analyzing value creation, it must therefore be interpreted in relation to underlying decision processes. A central distinction in this context is between attention and evaluation. Chandon et al. (2009) show that visual attention, such as noticing a product or stopping in a zone, is a necessary first step in the decision process, but does not by itself indicate purchase intent. A customer may attend to a product without actively considering it. This means that common behavioral measures such as dwell time or stops, are inherently ambiguous. On their own, they primarily reflect exposure rather than evaluation or purchase readiness. To become meaningful indicators of value, they must be combined with behaviors that signal active evaluation, such as revisiting products, comparing alternatives or structured movement between zones.

However, behavior does not only reflect attention and evaluation. It may also indicate friction in the decision process. Puccinelli et al. (2009) show that different psychological processes such as goal activation, information processing, involvement and affect, influence behavior at different stages of the shopping journey. As a result, the same observable behavior, such as time spent in a zone, can reflect very different underlying processes. For example, extended time in a zone may indicate active evaluation but it may also signal difficulty in finding or assessing products.

Taken together, this suggests that in-store behavior should be interpreted in terms of three analytically distinct patterns:

- **Attention without evaluation**, such as brief stops or passive browsing, which indicates exposure but limited purchase intent
- **Active evaluation,** such as revisits, comparisons, and structured interaction, which reflects deeper consideration and higher purchase likelihood
- **Friction or confusion**, such as prolonged time without progression or erratic movement, which may indicate difficulty in navigating or evaluating products

This distinction highlights an important analytical principle: meaningful insights are derived from patterns of behavior rather than from single indicators. Individual measures are often ambiguous, but when combined, they provide a clearer picture of how customers progress through the decision process. Path data provides a way to capture these patterns. Hui et al. (2009a) show that customers' movement through a store, including which zones they visit and in what sequence, reflects how decisions unfold over time. These movement patterns are not random but can be grouped into distinct behavioral types (Larson et al., 2005), supporting the idea that different behavioral profiles are associated with different value outcomes. Similarly, Hui et al. (2009b) suggest that in-store behavior must be interpreted in context, as movement patterns may reflect more than simple exploration.

For the purpose of this study, this implies that behavior should not be analyzed as isolated actions, but as patterns that indicate how customers move through different stages of the decision process. These patterns provide the conceptual basis for linking observable behavior to value outcomes in the empirical analysis, even if the underlying processes cannot be directly observed.

2.4 Fashion Context: Diagnostic Behaviors

While the previous section establishes general principles for interpreting in-store behavior, the meaning of these behaviors depends on the retail context. Not all behaviors mean the same thing in every category and their interpretation depends on the type of product being evaluated. This section specifies how in-store behaviors should be understood in the context of fashion retail. In fashion, certain behaviors are especially informative because of how purchase decisions are made. Incorporating the fashion context therefore allows the framework to identify which behaviors are most diagnostic of evaluation and purchase readiness.

Fashion products have important material and body-related attributes such as texture, drape and fit, which cannot be fully assessed visually. This increases the importance of sensory interaction, particularly touch. Peck and Childers (2003) show that haptic information plays a key role in evaluating products with varying material properties. They also find that restricting touch can reduce consumers' confidence in their evaluations and increase frustration. This provides a clear implication: behaviors such as touching, picking up and handling products are strong signals of active evaluation, as they reflect a deeper level of information processing than visual attention alone. Touch also has an important emotional (affective) role, not just an informational one. Peck and Wiggins (2006) show that touch can enhance positive feelings and increase persuasion, especially when the sensory experience is neutral or pleasant. For a behavior-based ICP, this means that touch-intensive behavior can signal both deeper evaluation and more positive attitudes toward the product. In this sense, touch functions as a dual signal, reflecting both cognitive evaluation and emotional engagement.

Another central behavior in fashion retail is fitting room use, which represents a key stage in the decision process where customers move from browsing to active evaluation. At this stage, uncertainty about fit and style is reduced through direct trial. Research shows that fitting room usage is closely linked to purchase outcomes, with moderate use associated with higher sales (Lee et al., 2020). At the same time, excessive usage can create congestion and reduce customers' ability to evaluate products effectively. This highlights that the fitting room is not only central to evaluation, but also sensitive to operational constraints that can introduce friction. Industry data provides further insight into how this process unfolds in practice. While fitting room usage is generally associated with higher conversion, performance varies depending on store conditions. In particular, conversion tends to decrease during periods of high traffic, indicating that capacity constraints can disrupt the transition from evaluation to purchase (Indivd, 2026c). This shows that the fitting room functions as a critical evaluation stage, where value is more likely to be created, but also where operational friction can reduce that value.

A broader perspective on why physical interaction matters in fashion is provided by Zhang et al. (2022), who show that deeper product inspection in physical stores enhances learning and can strengthen future customer value. Although long-term outcomes cannot be directly observed in anonymous visit-level data, this supports treating behaviors that reflect deep inspection, such as repeated comparisons, revisits to the same zone and extended interaction with products, as meaningful indicators of high-value visits.

Taken together, the fashion context refines the behavior-to-inference logic established earlier. In a category where decisions depend on sensory and situational validation, specific behaviors such as touch, fitting room use and comparison sequences, are more informative indicators of purchase potential than general measures like total time spent in-store. This means that in fashion retail, behaviors related to product interaction and fitting room use provide particularly useful indicators when analyzing which visits are likely to result in purchase. In the context of this study, these behavioral interpretations should be understood as indicative rather than definitive, as the empirical analysis observes behavioral proxies rather than directly measuring underlying evaluation processes.

While these insights clarify how behavior can be interpreted in a fashion context, the analysis also requires a clear definition of value that is consistent with the available data. The following section therefore addresses how value is conceptualized and measured in this study.

2.5 From CLV to Visit-Based Value

A central challenge in this study is how to define and measure customer value in a context where individual customers cannot be tracked over time. In traditional customer value literature, concepts such as Customer Lifetime Value (CLV) and customer equity are central to decision-making. These approaches define value as the discounted stream of future revenue generated by a customer, based on retention and repeat purchases (Gupta et al., 2004; Drèze & Bonfrer, 2008).

A key assumption underlying these concepts is that customer behavior can be observed at the individual level across multiple interactions. This makes it possible to estimate retention, predict future purchases and calculate long-term value. However, in many physical retail environments, this assumption does not hold. Behavioral data is typically anonymized and recorded at the level of the visit, meaning that individual customers cannot be reliably linked across time. As a result, CLV is not only difficult to estimate, but also conceptually misaligned with the available data (Bradlow et al., 2017). This creates a necessary shift in how value is defined. Instead of focusing on long-term customer value, value must be operationalized at the level that is observable in the data: the store visit. In this study, value is therefore defined using visit-level proxies that capture the observable or estimated economic outcome of a visit.

A practical way to operationalize this is to distinguish between different dimensions of visit-level value. First, conversion captures whether a visit results in a purchase. Second, value per visitor reflects the expected economic value generated during a visit. In this study, this measure is based on the retailer's internal valuation model and should therefore be interpreted as an estimated proxy rather than a direct measure of transaction-level value. While both measures are related, they capture different aspects of value creation: conversion reflects the occurrence of purchase, whereas value per visitor reflects variation in the magnitude of value across visits. This distinction is important for empirical analysis. Conversion is used to evaluate whether predefined customer segments, such as the ICP, are associated with purchase outcomes. In contrast, value per visitor is used to analyze how behavioral patterns relate to differences in value generation across visits. In this sense, the two measures serve complementary roles in assessing value creation.

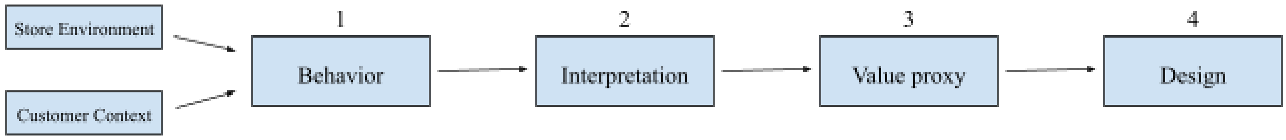
This shift in the definition of value also requires a reconsideration of how ICP is understood. In traditional segmentation approaches, ICP is often treated as a predefined group of customers assumed to generate high value. However, in a setting where value is defined at the level of the visit, this interpretation becomes less straightforward. Instead, ICP in this study is approached as an empirical construct that must be evaluated against observed outcomes. This means that rather than assuming that certain customer characteristics or profiles correspond to high value, the analysis examines whether visits associated with the ICP actually result in higher conversion or higher value.

At the same time, the use of behavioral data allows for an alternative perspective on ICP. Rather than defining ICP purely in terms of who the customer is, it becomes possible to interpret ICP in terms of how customers behave during a visit. From this perspective, ICP can be extended beyond static customer characteristics and partly understood through behavioral patterns that are associated with higher expected value outcomes. Importantly, this does not imply that ICP represents a fixed category. Instead, it should be understood as a probabilistic and continuous concept, where different behavioral patterns are associated with different likelihoods of generating value. This interpretation is consistent with segmentation theory, which allows for situational and behavior-based segmentation rather than fixed group boundaries (Dickson, 1982). Overall, these considerations establish how value and behavior can be meaningfully linked within the constraints of visit-level data. The following section integrates these insights into a conceptual framework that guides the empirical analysis.

2.6 Conceptual Framework

The literature reviewed provides the foundation for a structured framework that links the store environment, customer behavior and value outcomes. In this framework, behavior serves as the central mechanism through which value is generated and can be analyzed. Customer behavior does not occur in isolation. It is shaped by both the store environment and the customer context. The store environment, including layout, assortment, service elements and fitting room capacity, represents a set of controllable factors that influence how customers navigate and evaluate products. At the same time, customer-specific factors, such as goals, urgency and preferences, shape how the environment is experienced. Together, these factors influence how customers move through the store and interact with products, resulting in observable in-store behavior. This behavior reflects how customers progress through different stages of the decision process and therefore forms the basis for analyzing value creation. To structure this framework analytically, the study outlines four key elements that describe how observable behavior is interpreted, linked to visit-level value and translated into design implications:

Conceptual Framework



(1) Observable behavior

Customers' in-store actions, such as movement patterns, time spent in zones, product interaction and fitting room use, represent observable expressions of how they engage with the store environment.

(2) Psychological interpretation

These behaviors are interpreted as signals of underlying decision processes, such as attention, evaluation and friction. Since individual behaviors are often ambiguous, interpretation focuses on patterns of behavior rather than isolated actions.

(3) Value (visit-level)

Based on this interpretation, behavioral patterns are linked to value outcomes at the level of the visit. Value is captured using observable measures such as conversion and value per visitor, reflecting both the likelihood and magnitude of purchase.

(4) Design implications

Finally, behavioral patterns associated with higher or lower value are linked back to elements of the store environment. This makes it possible to identify how layout, product placement and service capacity can be adjusted to facilitate value-generating behavior and reduce friction in the customer journey.

The framework describes a directional process: store design and customer context shape behavior, behavior is interpreted and linked to value and these insights are then used to inform design decisions. This creates a feedback loop in which the store environment can be continuously adjusted to better support value-generating behavior. This framework provides a direct link between theory and empirical analysis, enabling a structured comparison between customer composition and behavioral patterns as explanations of value creation within the store.

2.7 Theoretical Propositions

This section translates the conceptual framework into a set of theoretically grounded expectations that guide the empirical analysis. The propositions do not represent formal hypotheses to be statistically tested, but instead specify how different types of in-store behavior are expected to relate to visit-level value outcomes. In this sense, they provide a structured basis for interpreting the empirical results rather than testing isolated theoretical constructs.

Proposition 1: Attention Versus Evaluation

Attention-related behaviors, such as dwell time and stops, are necessary but not sufficient indicators of value. Their relationship with value outcomes is expected to be weak unless combined with behaviors that signal active evaluation, such as revisits, comparisons or structured interaction with products.

Proposition 2: Diagnostic Behaviors in Fashion

In fashion retail, behaviors involving product interaction, such as touch and fitting room use, are expected to be stronger indicators of value than general exposure measures, as they reflect active evaluation of product attributes that cannot be assessed visually.

Proposition 3: Behavioral Patterns Versus Single Indicators

Value outcomes are expected to be better explained by patterns of behavior rather than by individual metrics. Different movement and interaction patterns, such as goal-directed versus exploratory behavior, are therefore expected to have different relationships with value outcomes.

Proposition 4: Friction and Disruption
Behaviors that indicate friction in the decision process, such as prolonged time without progression, repeated searching, or erratic movement, are expected to be negatively associated with value outcomes.

Proposition 5: Visit-Level Value and ICP
In the absence of individual-level tracking, value is defined at the level of the visit. ICP is therefore expected to function as an approximate or partial indicator of value, rather than a precise representation of high-value customers.

These propositions guide the interpretation of the empirical analysis by linking observable behavioral patterns to visit-level value outcomes and by providing a structured basis for comparing the explanatory roles of customer composition and in-store behavior.

3. Methodology

3.1 Research Design

This study adopts a single-case, exploratory research design to examine how customer composition and in-store behavior relate to value creation in a physical retail setting. The empirical context is a fashion retail environment, where behavioral data is used to analyze how visitors interact with the store and how these interactions are associated with value outcomes. The exploratory design reflects both the novelty of the data and the aim of the study. Rather than testing predefined causal relationships, the analysis seeks to identify patterns and assess how well different approaches explain value creation. In particular, the study is structured around a comparison between two analytical perspectives: one based on predefined customer segments (ICP) and one based on observed behavioral patterns. This design is consistent with the overall objective of evaluating whether value in physical retail is better explained by customer composition or by how the customers behave during the store visit.

A single-case approach is appropriate for this study, as the objective is to develop an in-depth understanding of how behavioral data can be used to analyze value creation within a specific retail context. Case study research is particularly suitable for exploratory research where the aim is to examine complex phenomena in real-world settings and where contextual factors play an important role (Yin, 2018). In this study, the store environment, customer behavior and value outcomes are closely intertwined, making a case-based approach well suited for capturing these relationships.

3.2 Data Sources

The empirical analysis is based on two complementary data sources: large-scale in-store tracking data and structured in-store observations. The primary dataset is provided by Indivd and is based on behavioral tracking data from one focal fashion retail store. The data captures in-store behavior using camera-based tracking systems and converts this into anonymized, aggregated visitor-flow statistics (Indivd, 2026a). The system records aggregated information on how visitors move through the store, how long they spend in different zones and how they interact with different parts of the retail environment. The data is anonymized and recorded at an aggregated level, ensuring that individual visitors cannot be identified or tracked across time (Indivd, 2026a, 2026b).

To complement this, structured in-store observations were conducted in three fashion stores in central Stockholm. 124 observations were carried out using a predefined observation instrument (see Appendix A1). These observations capture behavior at a more detailed level, focusing on indicators such as product interaction, comparison, revisits and signs of evaluation or friction. Observation is a suitable method in this context because it allows customer behavior to be captured as it unfolds in a natural retail setting, rather than relying on self-reported data (Baker, 2006). This is particularly important given that in-store behavior is often context-dependent, meaning that similar observable actions may reflect different underlying decision processes, such as evaluation or friction. Structured observation enables these nuances to be captured by recording both what customers do and how they interact with the store environment. In exploratory research settings, observation is also well suited for identifying and interpreting behavioral patterns without relying on predefined models (Tenny et al., 2022).

Observations were conducted during the afternoon of a Friday to capture a period of relatively high customer activity. One of the observed stores corresponds to the focal store used in the Indivd dataset, while the two additional stores were included to provide complementary observations beyond a single-store setting. These stores were chosen since they share a similar fashion retail orientation, central urban location and customer journey characteristics. The observations were carried out by the two authors using a systematic convenience approach. Customers entering the store environment were selected when their customer journey could be clearly observed, and each observed journey was recorded only once to reduce the risk of double-counting.

The two data sources serve different but complementary roles in the study. The Indivd tracking data provides a large-scale, aggregated view of customer behaviour and value-related outcomes across zones and time in the focal store. This makes it possible to analyse broader relationships between customer composition, engagement, conversion and value per visitor. However, because the data is anonymized and aggregated, it cannot show how individual customer journeys unfold or what specific behaviours mean in practice. The structured observations address this limitation by capturing visit-level behaviours such as product interaction, comparison, revisits, fitting room use and signs of friction. The observational data therefore adds a more detailed behavioural layer that helps interpret the mechanisms behind the aggregate patterns identified in the tracking data. In this way, the two datasets are not used as separate or competing sources of evidence, but as complementary levels of analysis. The tracking data shows where and when value-related patterns appear at scale, while the observations help explain how those patterns may emerge through specific customer behaviours during the store visit.

3.3 Data Structure and Variables

The Indivd dataset is structured at the level of zone–time observations, where each observation represents activity within a specific store zone during a defined time interval (hourly). This means that the unit of analysis is not the individual customer, but aggregated visitor activity within a given zone and time period. This reflects the system's focus on anonymous group-level visitor-flow statistics rather than individual tracking (Indivd, 2026a). As a result, all variables in the dataset reflect group-level patterns rather than individual behavior.

3.3.1 Customer composition (ICP variables)

Customer composition is captured through retailer-defined customer segments (ICP groups). In this study, ICP groups are based on average order value and demographic characteristics. However, it is important to note that such classifications are not inherently fixed or objective. In practice, retailers can define ICPs based on a wide range of attributes, depending on what they consider strategically relevant, including transactional metrics, demographic characteristics, behavioral indicators or brand-related considerations. As such, ICP should be understood as a managerial construct that reflects underlying assumptions about which customers are most valuable.

For each zone–time observation, the dataset includes the number of visitors belonging to each ICP group. Based on this, ICP share is calculated as the proportion of visitors in a given zone–time observation that belongs to the ICP group(s), relative to the total number of qualified visitors in that observation. Since ICP classifications are defined by the retailer, they are treated as an external classification that is evaluated in the analysis, rather than assumed to reflect actual value.

3.3.2 Behavioral Variables

Behavioral variables capture how visitors interact with the store environment at an aggregated level (Indivd, 2026a). These include:

- movement between zones
- time spent in specific areas (dwell time)
- engagement indicators, such as whether visitors remain in a zone or quickly exit

These variables reflect collective patterns of behavior rather than individual actions. As such, they provide an indication of how visitors engage with different parts of the store, but require interpretation to be meaningfully linked to decision processes. To support this interpretation, behavioral patterns are further analyzed using the observational data. This allows aggregated behavioral measures to be interpreted in terms of underlying processes such as attention, active evaluation and friction.

3.3.3 Outcome variables (value measures)

Value is operationalized at the level of the visit using two complementary outcome measures:

- **Conversion rate**, defined as the proportion of visitors in a given zone–time observation who make a purchase
- **Value per visitor**, defined as the estimated economic value generated per visitor within a given observation

These measures capture different dimensions of value creation. Conversion reflects whether a purchase occurs, while value per visitor captures variation in the magnitude of value across visits.

3.3.4 Implications of data structure

Due to the aggregated and anonymized nature of the data: individual customers cannot be tracked across zones or over time, behavioral patterns cannot be directly linked to individual purchase decisions and the analysis focuses on relationships between variables at an aggregated level. This implies that the results should be interpreted as associations between variables at the aggregated level, reflecting patterns of visitor behavior rather than behavior of individual customers. The analysis therefore identifies how different compositions and behavioral patterns are associated with value outcomes across observations, rather than how individual customers behave or make decisions.

3.4 Analytical Approach

Building on the aggregated and visit-level structure of the data, the empirical analysis is structured around two complementary analytical approaches, reflecting the study's objective of comparing customer composition and observable in-store behavior as explanations of visit-level value creation.

3.4.1 Approach 1: ICP-Based Analysis (customer composition)

The first part of the analysis examines store performance based on the retailer's existing ICP definition. This includes a descriptive zone-level analysis using a zone scorecard, where ICP share and conversion rates are compared across different store zones. This provides a benchmark of store performance under the retailer's current segmentation logic. In addition, regression analysis is used to test whether ICP share is associated with conversion outcomes. This allows the study to evaluate whether the retailer's predefined customer segments function as meaningful indicators of value.

3.4.2 Approach 2: Behavior-Based Analysis

The second part of the analysis focuses on behavioral patterns as an alternative explanation of value creation. Behavioral variables are analyzed in relation to value per visitor, examining how different forms of engagement are associated with variation in value outcomes across observations. In practice, this involves constructing aggregated behavioral measures and analyzing their relationship with value. To further interpret these results, the observational data is used both to test how specific visit-level behaviors are associated with purchase probability and to unpack the behavioral mechanisms underlying the aggregated variables. In particular, the observations are used to distinguish between different types of behavior, such as cognitive engagement (e.g., focused attention and revisits) and friction (e.g., prolonged inactivity or repeated searching). This approach allows the analysis to move beyond aggregate measures and provide a more detailed understanding of how specific behavioral patterns relate to value creation.

Together, these two approaches enable a direct comparison between customer composition and behavior as explanatory factors of value creation. While the ICP-based analysis evaluates the retailer's predefined segmentation, the behavior-based analysis examines whether value outcomes can be more effectively explained through observed patterns of in-store behavior.

3.5 Limitations

Several limitations should be noted. First, the data is aggregated and anonymized, which prevents analysis at the individual level and limits the ability to establish causal relationships. Second, the analysis does not control for external factors such as promotions, seasonality or staffing conditions, which may influence observed outcomes. In addition, the observational component is limited in scale and is intended to provide interpretive support by helping clarify the behavioral meaning of the aggregated variables, rather than statistically generalizable evidence on its own. Finally, the primary study is based on a single retail case, which limits the generalizability of the findings. The results should therefore be interpreted as context-specific and illustrative rather than universally applicable.

4. Empirical Results

4.1 Introduction to the Analysis of Indivd's Data

Based on the raw datasets provided by Indivd, several analyses have been conducted and will be presented in this section. While parts of the analytical approach, particularly the zone scorecard, are inspired by Indivd's methodology, all calculations have been independently performed using Python and Excel.

For this dataset, the fashion retailer has defined a set of Ideal Customer Profiles (ICPs). These segments are based on average order value and further differentiated by demographic attributes such as gender and age. The figure below presents the four most important ICP groups for this retailer.

ICP 1 (Women 30-40)	2.0x AOV = 1,500 SEK
ICP 2 (Women 20-30)	1.5x AOV = 1,125 SEK
ICP 3 (Women 40-50)	1.2x AOV = 900 SEK
ICP 4 (Men 30-40)	1.2x AOV = 900 SEK

While the validity of these ICP groups as the most value-generating segments will be evaluated in a later analysis, they represent the retailer's current strategic focus. Therefore, the purpose of the first analysis is to assess how well the store performs in attracting and converting these target customers across different zones.

4.2 Zone Scorecard under Retailer-Defined ICPs

The first analysis is a zone scorecard, designed to provide a descriptive benchmark of store performance under the retailer's current ICP definition. Rather than evaluating whether the ICP groups are optimal, this analysis examines how the store performs in attracting and converting these predefined target customers across different zones. The purpose is therefore to answer the following question: given the retailer's existing segmentation, where in the store do these target groups appear and how effectively are they converted? This analysis is based on two datasets:

- *"12. Raw - zone performance"*
- *"13. Raw - cross-conversion"*

The first dataset contains hourly zone-level metrics such as total visitors, bounce rate, dwell time, qualified visitors and ICP visitors. This dataset is used to calculate the weighted ICP share for each zone (see Appendix B1). ICP share is defined as the proportion of ICP visitors relative to qualified visitors within a zone. To ensure an accurate representation across time, the metric is aggregated as a weighted average based on qualified visitors. According to Indivd's definition, qualified visitors are themselves weighted by dwell time engagement, meaning that visitors who spend enough time in a zone have a greater influence on the metric. This measure indicates the extent to which each zone attracts the retailer's predefined target customer segments. A high ICP share suggests that a large proportion of visitors in that zone belong to strategically important customer groups.

The second dataset, *"13. Raw - cross-conversion"*, provides hourly conversion rates to any tills for each zone. This dataset is used to calculate the weighted conversion rate per zone (see Appendix B2). Similar to ICP share, conversion is aggregated using a weighted approach to account for differences in traffic across time. This ensures that periods with higher visitor volumes have a proportionally larger impact on the final metric. The conversion rate reflects the extent to which visitors in each zone complete a purchase, providing a measure of commercial performance.

To interpret the results, a threshold-based framework is applied to classify each zone according to its performance and strategic role. Median thresholds were used because they provide a simple relative benchmark within the dataset rather than imposing external performance standards. The thresholds are based on the median values of the dataset:

- Median ICP share: **61.6% (rounded to 62%)**
- Median conversion rate: **5.7%**

Using these thresholds, each zone is categorized into one of four groups:

Framework quadrants (thresholds)

Core	ICP share >= 62% AND conversion rate >= 5.7%.
Broad	ICP share < 62% AND conversion rate >= 5.7%.
Potential	ICP share >= 62% AND conversion rate < 5.7%.
Weak	ICP share < 62% AND conversion rate < 5.7%.

- **Core zones**: Zones with both high ICP share and high conversion. These represent the most strategically valuable areas of the store.
- **Broad zones**: Zones with high conversion but lower ICP share. These generate strong commercial performance but attract a broader customer base beyond the target segments.
- **Potential zones**: Zones with high ICP share but lower conversion. These indicate strong presence of target customers but underperformance in converting them, representing key opportunities for improvement.
- **Weak zones**: Zones with both low ICP share and low conversion. These areas currently contribute limited strategic or commercial value.

The resulting classification is summarized in the zone scorecard below. Only zones with available data for both ICP share and conversion are included. Certain zones, such as structural areas, are excluded due to their limited relevance, while others (e.g. fitting rooms and café) are retained despite differing functional roles.

Zone scorecard based on ICP share and conversion rate

Zone	ICP %	Conv %	Qual.	Conv
CORE (W11-W13)				
Home 2	78.5%	7.7%	*19,888*	*1,531*
Women 4	74.8%	7.4%	*32,942*	*2,438*
Women 3	70.7%	5.7%	*34,910*	*1,990*
Men 1	64.9%	6.3%	*14,083*	*887*
BROAD (W11-W13)				
Fitting room	59.6%	12.0%	*3,597*	*432*
Men 2	59.0%	9.6%	*1,884*	*181*
Home 1	55.8%	6.8%	*5,772*	*392*
Men 4	52.5%	5.7%	*2,713*	*155*
Women 8	50.5%	6.4%	*2,713*	*174*
Women 6	30.7%	6.9%	*2,663*	*184*
POTENTIAL (W11-W13)				
Women 7	66.6%	5.6%	*19,465*	*1,090*
Men 3	71.6%	3.1%	*18,875*	*585*
Women 1	65.1%	4.0%	*37,128*	*1,485*
Fitting room queue	64.8%	4.6%	*16,963*	*780*
WEAK (W11-W13)				
Women 2	61.6%	5.0%	*3,677*	*184*
Café	49.1%	2.7%	*3,520*	*95*
Children 1	33.9%	5.5%	*2,050*	*113*

As shown in the scorecard, both ICP share and conversion rate are color-graded based on quartile distributions to improve interpretability (see Appendix B3 for the quartile-based color classification).

The analysis identifies several zones as **core areas**, including *Home 2, Women 4, Women 3*, and *Men 1*, where both ICP share and conversion are above the defined thresholds. These zones combine a high presence of the retailer's target customers under the current ICP definition with strong conversion performance. The women's sections in particular stand out, not only due to strong ICP alignment but also because of relatively high visitor volumes, indicating a substantial contribution to overall store performance.

A number of zones fall into the **potential category**, including *Men 3, Women 1, Women 7*, and *Fitting room queue*. These zones attract a high proportion of the retailer's target customers but show comparatively lower conversion rates, suggesting that the presence of the target customers does not automatically translate into purchase outcomes.

The **broad category** includes zones such as *Fitting room, Men 2, Home 1*, and *Women 8*, where conversion rates are strong but the customer mix is less aligned with the defined ICPs. These zones deliver strong conversion performance, but with a lower share of the retailer's target segments, suggesting that value is also created outside the predefined ICP groups.

Finally, **weak zones** such as *Café, Children 1*, and *Women 2* show lower performance across both dimensions. However, it is important to note that certain zones, such as the café, serve alternative purposes and therefore should not be evaluated solely on conversion metrics.
A more detailed interpretation of these findings and their strategic implications will be discussed in the discussion section.

Importantly, this analysis does not assess whether the ICP definition itself captures the most value-generating customers. Instead, it serves as a descriptive benchmark of store performance under the retailer's current segmentation logic. The explanatory strength of the ICP definition is therefore evaluated in the following analysis.

4.3 Does the Current ICP Definition Capture Value?

While the zone scorecard provides a clear picture of store performance under the retailer's current ICP definition, it does not assess whether these predefined customer segments are actually associated with higher value creation. To address this, the following analysis examines whether ICP share is empirically related to conversion performance. This represents a shift from a descriptive perspective to a more analytical evaluation of the retailer's segmentation logic.

Due to the lack of individual-level tracking data and the fact that order value is estimated based on retailer assumptions, there is no direct way to fully validate the ICP definition. However, it is possible to approximate this by analyzing the relationship between ICP presence and conversion behavior. Using the *"13. Raw - cross-conversion"* dataset, conversion rates are available per zone and time, while the *"12. Raw - zone performance"* dataset provides ICP share per zone and time. By merging these datasets, it becomes possible to test whether higher ICP presence within a given zone and time is associated with higher conversion rates.

During the merging process, some observations do not align across the datasets or contain missing values. These rows were excluded to ensure that the analysis only included observations with complete information on both ICP share and conversion rate, resulting in a final dataset of 4,280 observations containing both ICP share and conversion rate (see Appendix C1 for more information).

The regression results are presented in Appendix C2. The estimated coefficient for ICP share is 0.0044, indicating a small positive effect on conversion. This effect is statistically significant. In practical terms, assuming both variables are measured as proportions, this means that a 10 percentage point increase in ICP share is associated with an increase in conversion rate of approximately 0.044 percentage points. However, given the large number of observations, statistical significance should be interpreted cautiously and evaluated together with effect size and explanatory power. While the relationship is positive, the magnitude of the effect is economically small. Furthermore, the model has very low explanatory power ($R^2 \approx 0.003$), meaning that ICP share explains only a negligible proportion of the variation in conversion rates. This suggests that although ICP presence is statistically related to conversion, it is not a strong standalone indicator of conversion performance.

To complement the row-level analysis, a second regression is performed using aggregated zone-level data (full regression results are presented in Appendix C3). At this level, the explanatory power of the model increases slightly ($R^2 \approx 0.106$), indicating that ICP share explains a somewhat larger, but still limited, portion of the variation in conversion rates across zones. However, the coefficient for ICP share is not statistically significant, meaning that zones with higher ICP share do not consistently exhibit higher conversion rates.

Taken together, the results indicate that ICP share has a limited explanatory power for conversion. While there is a small and statistically significant relationship at the row level, the effect is economically weak and does not translate into a statistically significant relationship across zones. This suggests that customer composition alone is insufficient for explaining conversion performance. Although the analysis has limitations, it provides an important indication that the retailer's current ICP definition may not fully capture the most value-generating customers. This creates a clear contrast: while the ICP definition is strategically meaningful from the retailer's perspective, it appears to be empirically weak as a signal of value creation. This raises the question of whether other factors, particularly behavioral patterns, provide a more accurate explanation of value, which is explored in the following section.

4.4 Which Broad Behaviors Are Associated with Value Creation?

Building on the previous analysis, which showed that ICP share has limited explanatory power for conversion, this section examines whether behavioral patterns provide a stronger explanation of value creation.

Importantly, this analysis differs from the previous section in terms of the outcome variable. While section 4.3 focuses on conversion rate, this section uses value per visitor as a proxy for economic value. This measure is constructed based on the retailer's internal valuation model and reflects estimated value rather than actual transaction-level spending. Rather than aiming to measure true customer value, the purpose of this analysis is to examine how broad behavioral engagement relates to the retailer's own valuation logic. This means that the comparison between customer composition and behavioral engagement should be understood as complementary evidence across different available value proxies, rather than as a strict model-to-model comparison using the same dependent variable.

A multiple linear regression model is used to estimate how behavioral and contextual factors relate to value per visitor (see Appendix D for full regression results). The main variable of interest is engagement, measured as the share of visitors who stay in the store rather than leaving quickly. The model also controls for visitor volume, time segment, weekday versus weekend and differences across store zones.

The results show that engagement has a strong and highly statistically significant positive relationship with value per visitor ($\beta = 440.73$, $p < 0.01$). Observations with higher engagement are associated with substantially higher value, indicating that customers who interact more actively with the store environment contribute more to value creation. This effect is not only statistically robust, but also meaningful in magnitude, suggesting that changes in customer behavior can have a noticeable impact on value outcomes. At the same time, visitor volume is negatively associated with value per visitor ($\beta = -9.88$, $p < 0.01$). This implies that higher traffic does not necessarily lead to higher value on a per-customer basis. One possible interpretation is that periods with more visitors may include a larger share of less purchase-oriented customers, or that increased congestion reduces the quality of in-store interactions.

The results also show that time plays an important role. In particular, evening periods are associated with significantly higher value per visitor ($\beta = 77.00$, $p < 0.05$) compared to other times of day, suggesting that customers visiting during these periods may be more focused or more likely to engage in value-generating behavior. In addition, there are substantial differences across store zones. Several zones are associated with lower value relative to the baseline, indicating that value creation is not evenly distributed across the store. Instead, it depends on where customers spend their time and how they interact with specific areas.

Overall, the analysis suggests that broad behavioral engagement is positively associated with the retailer's estimated value-per-visitor measure. Together with the ICP analysis, this indicates that behavioral variables provide a more direct and informative explanation of visit-level outcomes than ICP share alone. However, because the ICP and behavioral analyses rely partly on different outcome measures, the comparison should be interpreted as complementary evidence rather than as a strict model-to-model comparison.

4.5 What Do Those Behaviors Appear to Mean in Practice?

4.5.1 Logistic Regression

To complement the analyses based on Indivd's aggregated behavioral data, a second analysis was conducted using manually collected observational data from individual in-store customer visits. While the previous analyses examined value creation using Indivd's aggregated data, this analysis uses manually collected observation data to examine how specific visit-level behaviors are associated with purchase probability. This analysis includes a formal regression model, but its primary role is complementary and interpretive, providing additional insight into the behavioral mechanisms underlying the aggregated results.

The observation-based approach complements the Indivd dataset by enabling a more fine-grained analysis of behavioral mechanisms during the customer journey. Whereas the Indivd data captures aggregated behavioral patterns across store zones, the observational data makes it possible to directly examine how specific customer behaviors, such as product interaction, fitting room usage, revisits, and signs of friction, are associated with conversion outcomes.

The analysis is based on 124 structured in-store observations. Each observation represents one customer visit and includes systematically recorded behavioral indicators related to movement patterns, product interaction, evaluation behavior and purchase outcome. Purchase was operationalized as a binary variable based on whether a purchase was observed during the visit. Because the purpose of this part of the analysis is to examine whether specific behaviors are associated with conversion, purchase was treated as a binary outcome rather than as a continuous value measure. This makes logistic regression appropriate, since the model estimates how different behavioral dimensions are associated with the probability that a visit ends in purchase.

To operationalize the theoretical framework into measurable analytical variables, the observed behaviors were grouped into five behavioral dimensions based on the structured observation items presented in Appendix A1. These dimensions are derived from the theoretical distinction between attention, active evaluation, and friction developed in the literature review and represent different stages of the in-store decision-making process. While grounded in this framework, the variables represent observable approximations of these concepts rather than direct measurements of the underlying decision processes. This grouping was used because single behaviors are often ambiguous when interpreted alone. For example, touching a product, revisiting a zone, or moving through several areas may have different meanings depending on the broader journey context. Grouping behaviors into theoretically grounded dimensions therefore makes the analysis more aligned with the conceptual framework and reduces the risk of overinterpreting isolated actions.

- *Product engagement* captures active interaction with products through touching (C1) and comparing items (C2).
- *Cognitive engagement* represents behaviors associated with deeper evaluation and focused consideration, measured through revisiting zones (C3) and repeated product handling (D1).
- *High intent* captures behaviors associated with later stages of the purchase process, measured through fitting room usage (C4).
- *Friction* represents behaviors interpreted as hesitation, uncertainty, or obstacles during the shopping journey, including long pauses without action (E1), repeated searching (E2) and repeatedly picking up and abandoning products (E3).
- *Exploration* captures how extensively customers moved through the store, measured through the number of zones visited and transitions between zones during the observed visit (B: Zone movement map).

These behavioral dimensions were subsequently used as independent variables in a binary logistic regression model, where the dependent variable was whether a purchase was observed during the visit. Before estimating the model, correlation analysis was conducted to assess whether the behavioral dimensions were too strongly related to be included in the same model (see Appendix E). The correlation matrix showed moderate positive correlations between several engagement-related variables, particularly product engagement, cognitive engagement, high intent, and exploration. However, none of the correlations were high enough to indicate severe multicollinearity. The variables were therefore retained in the final model.

The final logistic regression model converged successfully and showed good exploratory fit for an exploratory observation-based analysis (see Appendix F for the full regression output):

- Number of observations: 124
- Pseudo R^2: 0.49
- Log-likelihood ratio test: $p < 0.001$

The regression results are summarized below.

Logistic regression results

Variable	Coefficient	Odds ratio	p-value
Product engagement	1.64	5.17	0.003
Cognitive engagement	1.27	3.56	0.005
High intent	1.99	7.30	0.008
Friction	-0.70	0.50	0.024
Exploration	-0.04	0.96	0.781

The results indicate that conversion in physical fashion retail is strongly associated with behaviors reflecting active product evaluation and progression within the decision-making process. Since logistic regression coefficients are expressed in log-odds, odds ratios are included to make the results easier to interpret. An odds ratio above 1 indicates higher odds of purchase, while an odds ratio below 1 indicates lower odds of purchase, holding the other variables constant. Product engagement shows a positive and statistically significant relationship with purchase probability. The odds ratio of 5.17 indicates that visits involving active product interaction, such as touching and comparing products, were associated with substantially higher odds of purchase. This aligns with prior research emphasizing the importance of physical interaction in fashion retail evaluation processes.

Similarly, cognitive engagement demonstrates a significant positive relationship with conversion. Behaviors such as revisiting zones and repeated product handling appear to reflect deeper involvement and more advanced stages of evaluation during the shopping journey. The odds ratio of 3.56 indicates that higher cognitive engagement was associated with substantially higher odds of purchase.

The strongest positive relationship is observed for the high-intent dimension. Customers who used fitting rooms were significantly more likely to complete a purchase. The odds ratio of 7.30 indicates that fitting room usage was associated with much higher odds of purchase, supporting its interpretation as a later-stage evaluation behavior closely connected to purchase readiness.

In contrast, friction demonstrates a statistically significant negative relationship with purchase probability. The odds ratio of 0.50 indicates that behaviors interpreted as hesitation or disruption, such as long inactive pauses, repeated searching, or repeatedly picking up and abandoning products, were associated with lower odds of purchase. This supports the distinction developed earlier in the literature review between behaviors reflecting active evaluation and behaviors reflecting uncertainty or disruption within the shopping journey. Exploration was not found to have a statistically significant relationship with purchase probability when controlling for the other behavioral dimensions. This suggests that moving through more zone steps in the store does not necessarily indicate stronger purchase intention, and that conversion is more strongly associated with the nature of customer engagement than with movement itself.

Overall, the analysis supports the conceptual framework developed earlier in the thesis by demonstrating that observable in-store behaviors can function as meaningful indicators of underlying decision-making processes and visit-level purchase outcomes. More specifically, the findings suggest that behaviors associated with active evaluation and progression toward purchase increase conversion likelihood, while behaviors associated with friction reduce it. Importantly, this analysis differs from the earlier ICP-based zone scorecard. Rather than evaluating predefined customer segments, the logistic regression model instead identifies which behavioral mechanisms are associated with conversion itself. As a result, the analysis contributes a complementary behavioral perspective on value creation within the physical store environment and provides the primary statistical basis for interpreting how specific behaviors are associated with conversion.

4.5.2 Supplementary Linear Probability Model

To assess whether the behavioral pattern identified in the logistic regression is robust across model specifications, a supplementary linear probability model was estimated using the same behavioral dimensions and the same binary purchase outcome. Since purchase is operationalized as a binary variable, the logistic regression remains the primary model for this analysis. The purpose of the linear probability model is therefore limited to a robustness check, rather than serving as a co-equal analytical approach.

The results were consistent with the logistic regression. Product engagement, cognitive engagement, and high-intent behavior remained positively and statistically significantly associated with purchase, while friction remained negatively and statistically significantly associated with purchase. Exploration remained statistically insignificant. This suggests that the main behavioral pattern is not dependent on the specific model specification. The standardized coefficients from the supplementary model are presented below.

Supplementary linear probability model results

Variable	Standardized coefficient (β)	t-value	p-value
Product engagement	0.157	3.815	<0.001
Cognitive engagement	0.150	3.554	0.001
High intent	0.131	3.435	0.001
Friction	-0.076	-2.293	0.024
Exploration	0.003	0.077	0.939

The standardized coefficients also provide a limited additional insight by showing the relative contribution of the behavioral dimensions on the same scale. Product engagement showed the strongest relative positive association with purchase, followed by cognitive engagement and high-intent behavior. Friction had a smaller but negative association, while exploration contributed almost no additional explanatory value.

Overall, the supplementary linear probability model supports the interpretation from the logistic regression: purchase outcomes are more closely associated with active evaluation and progression toward purchase than with general movement through the store. The full linear probability model output is presented in Appendix G.

4.6 Summary of Empirical Results

The table below summarizes the empirical analyses conducted in this section and shows how each analysis contributes to the comparison between customer composition and in-store behavior.

Summary of empirical results

Analysis	**Outcome variable**	**Main result**
Zone scorecard	ICP share and conversion rate	Zones differ in how strongly they combine ICP presence and conversion. Several zones show mismatches between ICP share and conversion.
ICP regression	Conversion rate	ICP share has a statistically significant but economically small effect at row level. At zone level, the relationship is not statistically significant.
Behavior-value regression	Estimated value per visitor	Engagement is positively associated with estimated value per visitor, while visitor volume is negatively associated with value per visitor.
Logistic observation model	Purchase yes/no	Product engagement, cognitive engagement and high intent are positively associated with purchase. Friction is negatively associated, while exploration is not significant.

5. Discussion

5.1 Understanding Value Creation: ICP vs Behavior

5.1.1 Rethinking ICP as a Value Indicator

The results show that customer composition, such as the retailer's existing ICP definition, can be analyzed within a useful descriptive framework for understanding store performance, but has limited explanatory power for visit-level value creation. In particular, ICP makes it possible to describe how predefined customer segments are distributed across the store and how different zones perform in aggregate in terms of conversion. However, when moving from a descriptive to an analytical perspective, its ability to explain variation in value outcomes becomes more limited. The regression analysis shows that ICP share has only a weak association with conversion outcomes and explains little meaningful variation across zones. This suggests that while ICP may reflect the retailer's strategic priorities, it does not function as a strong empirical indicator of visit-level value creation. Importantly, this does not mean that ICP lacks relevance. ICP definitions can still serve broader strategic purposes, such as guiding brand positioning, assortment decisions, and long-term customer targeting. The limitation is rather that ICP alone provides limited insight into how value is generated during the store visit.

A key reason for this limitation is that ICP is typically based on assumptions or indirect indicators of value, rather than on observed behavior linked to actual outcomes. As a result, it does not fully capture the situational and dynamic nature of in-store decision-making. Customers who belong to the same segment may behave very differently depending on context such as their goals, time constraints or interaction with the store environment. This variation is not reflected in a static segmentation such as ICP. In contrast, visit-level value in physical retail appears to be more closely tied to what customers do during the visit.

The findings suggest that ICP should not be treated as a direct indicator of value but rather as a starting assumption that must be validated and refined through behavioral analysis. This shifts the role of ICP from a definitive representation of high-value customers to a more tentative and testable construct.

5.1.2 Behavior as a Stronger Signal of Value

Behavioral patterns change throughout the shopping process and reflect how customers engage with products, evaluate alternatives and progress towards a purchase decision. As such, a fixed segmentation based on customer characteristics is less able to capture these dynamics. This limitation of ICP creates the rationale for the behavioral analysis. If predefined customer composition explains only a limited part of visit-level value creation, the next question is whether observed behavior provides a more informative signal of how customers progress through the shopping process.

Given the limited explanatory power of ICP, the behavioral analysis provides a complementary way to understand visit-level value creation. While the aggregated analysis captures value through the retailer's estimated value-per-visitor proxy, the observation analysis provides a complementary perspective based on directly observed purchase outcomes. Across the empirical analyses, behavioral variables show a stronger and more consistent association with value outcomes than customer composition. This suggests that value in physical retail is more closely linked to what customers do during the visit than to who they are assumed to be under a predefined segmentation logic. This finding supports the conceptual framework developed in the literature review, where behavior is interpreted as a signal of underlying decision processes. More broadly, the results indicate that value creation is not driven by activity alone, but by the type of engagement, as different forms of engagement reflect how customers progress through the decision process.

An additional insight is that behaviors closer to the final decision stage show the strongest relationship with purchase outcomes. In the observational analysis, high-intent behaviors are associated with the highest purchase likelihood, followed by product engagement and cognitive engagement. This suggests a progression in the decision process, where early-stage behaviors contribute to value, but behaviors reflecting active evaluation and purchase readiness are most strongly associated with conversion. Together, these findings highlight that value creation is not explained by activity alone, but by the type and quality of engagement. Behavioral patterns provide a more direct and context-sensitive explanation of value, as they capture how customers interact with the store environment and progress through the decision process in real time.

This conclusion should be interpreted in relation to the data context of the study. Since value is observed at the level of the visit, behavioral variables provide a more direct link to outcomes, while customer composition remains an indirect measure. In settings where individual-level and longitudinal data are available, the relative explanatory power of segmentation-based approaches may differ. In this study, however, the findings suggest that while ICP reflects who the retailer aims to serve, visit-level value is more closely associated with how customers behave within the store.

5.2 Implications for In-Store Design

While the previous section demonstrated the limited explanatory power of the retailer-defined ICP in explaining conversion outcomes, the zone scorecard analysis highlights its strength as a practical and managerially intuitive diagnostic tool. Rather than explaining why value is created, the scorecard provides a structured way to diagnose where the retailer's predefined target customers are present, where conversion outcomes are stronger or weaker, and where mismatches between ICP presence and conversion occur. The classification of zones into core, potential, broad, and weak categories makes these mismatches visible and translates the scorecard into concrete store-level design questions.

For instance, core zones, such as Home 2 and Women 4, combine a high share of ICP visitors with strong conversion rates. These areas indicate zones where the retailer's current target segments are not only present, but where conversion outcomes are also relatively strong. As such, they provide useful reference points for identifying where the current store setup appears to be associated with stronger commercial outcomes. In contrast, potential zones, such as Women 7 and the fitting room queue, reveal a different type of insight. These zones attract a high proportion of ICP visitors but underperform in conversion. This mismatch suggests that while the retailer's predefined target customers are present, the store environment may not sufficiently support their progression from attention to evaluation and purchase. From a managerial perspective, these areas represent clear opportunities for targeted interventions such as adjustments in assortment, layout or in-store guidance. Broad zones further nuance the interpretation. These areas achieve strong conversion rates despite a lower share of ICP visitors, indicating that purchase outcomes are not limited to the retailer's predefined target groups. This reinforces the idea that the current ICP definition does not fully capture where observed conversion performance is strongest. Finally, weak zones show lower performance across both ICP presence and conversion, although some of these areas, such as the café, may serve alternative functions that are not fully captured by these metrics.

These findings demonstrate how the zone scorecard can be used as a diagnostic tool for prioritizing store-level investigation and potential action. By explicitly mapping the relationship between customer composition (ICP share) and performance (conversion), the analysis allows managers to move beyond aggregate metrics and instead focus on specific areas of the store. Importantly, the value of this approach does not depend on the ICP being perfectly defined. Even with a potentially imperfect segmentation, the scorecard enables retailers to identify where mismatches occur, such as high ICP presence combined with low conversion, and to generate hypotheses about underlying causes.

However, the findings from the behavioral analyses suggest that the zone scorecard could become more meaningful when combined with behavioral interpretation. Rather than relying solely on predefined ICP groups based on demographic and transactional characteristics, retailers may benefit from incorporating behavioral patterns associated with purchase progression and estimated value outcomes. In this context, the ICP should be viewed as a flexible construct that can be iteratively refined based on observed in-store behavior. Applying the zone scorecard through a more behavior-informed lens could improve its diagnostic precision by not only identifying where predefined target customers are present, but also where behaviors associated with active evaluation and purchase progression occur and how effectively these are supported by the store environment. This highlights the complementary roles of the two approaches: while the scorecard provides a scalable framework for linking customer composition and performance at a granular level, behavioral interpretation helps explain what may be happening within those zones. Together, this enables a shift from simply targeting predefined customer segments to shaping store environments that support evaluation, progression toward purchase and reduced friction.

5.3 Managerial Implications

The findings suggest that improving store performance requires a shift in how behavioral data is used in managerial decision-making. Rather than treating behavioral insights as a supplementary layer, they should be integrated as a central tool for understanding how visit-level value emerges within the store and for guiding ongoing optimization efforts. A key implication is that predefined target customer groups should not be assumed to be the primary indicators of in-store value but instead treated as hypotheses that need to be empirically tested against observed outcomes. This shifts the focus of store management from relying on static segmentation towards continuously evaluating how customers actually behave during the visit and how the store environment influences that behavior.

One important implication concerns the distinction between attention and evaluation. Much of traditional store optimization focuses on attracting attention through layout, displays or visual communication. However, the findings suggest that attention alone is not strongly associated with value creation. Instead, behaviors reflecting active evaluation, such as interacting with products, comparing alternatives or revisiting items, are more closely linked to purchase outcomes. From a managerial perspective, this implies that store environments should be designed not only to capture attention but to actively support customers in progressing beyond it. This includes enabling interaction, facilitating comparison and reducing barriers to continued engagement.

The results also highlight the importance of specific points in the customer journey where purchase decisions are more likely to occur. In particular, behaviors associated with later stages of the decision process, such as fitting room usage, show the strongest relationship with purchase outcomes. These areas can therefore be understood as critical decision points within the store. Managing these points effectively, for example by ensuring accessibility, reducing congestion and maintaining product availability, becomes directly linked to improving purchase progression. At the same time, the analysis highlights the importance of identifying and reducing friction within the store. Behavioral patterns such as hesitation, repeated searching or lack of progression are consistently associated with lower purchase likelihood and indicate disruptions in the decision process. Importantly, such friction may not be visible through traditional performance metrics alone. This suggests that managers should place greater emphasis on diagnosing where customers encounter difficulties and on addressing structural issues in layout, navigation or assortment that hinder progression.

These implications point towards a more adaptive approach to retail management. Rather than treating store design and segmentation as relatively fixed decisions, managers can use behavioral patterns to continuously monitor how customers move through the decision process and where improvements are needed. In this sense, the role of the retailer shifts from targeting predefined customer groups to actively managing and shaping the conditions under which customers are more likely to progress toward purchase.

5.4 Limitations

A central limitation of the study is that the Indivd data is aggregated and anonymized at the zone and time level. As a result, the analysis does not follow individual customers through the store or across visits but instead captures patterns in visitor behavior at an aggregated level. This has two key implications. First, the findings cannot be interpreted as individual-level effects. For example, a positive association between engagement and estimated value per visitor does not imply that every more engaged customer generates higher actual value, but rather that observations with higher aggregated engagement tend to be associated with higher estimated value outcomes. Second, the data structure limits causal inference. Since individual trajectories are not observed and the store environment is not experimentally manipulated, the analysis cannot establish causal relationships between behavior and outcomes. The findings should therefore be interpreted as associations between behavioral patterns, customer composition and value outcomes at the aggregated level. This defines the appropriate level of interpretation, while still demonstrating how behavioral data can be used to identify meaningful value patterns within the store.

A further limitation concerns how value is defined in the study. Due to the anonymized and visit-based nature of the data, value is operationalized at the level of the individual store visit using measures such as conversion and estimated value per visitor. While this allows for a consistent analysis of observable outcomes, it does not capture the full economic value of a customer over time. This has important implications for the interpretation of the findings. In particular, the limited explanatory power of the ICP should be understood in relation to visit-level outcomes rather than overall customer value. A customer who does not purchase during a specific visit may still be valuable in the long term through future purchases, brand engagement or repeat visits. As a result, segmentation approaches such as ICP may have stronger explanatory power when evaluated against longer-term customer value measures. The findings should therefore be interpreted as applying to visit-level value creation within the store, rather than as a general assessment of the relationship between customer segmentation and total customer value.

A third limitation concerns the observational data. The analysis is based on a relatively limited number of in-store observations and should therefore not be interpreted as statistically representative of all customer behavior in physical fashion retail. Instead, it serves as an exploratory complement to the aggregated data, providing a more detailed interpretation of behavioral patterns. In addition, the observational data involves a degree of interpretation. Behaviors such as focused attention, hesitation or friction are identified through structured observation but still require judgment from the observer, meaning that they cannot fully capture underlying customer intentions. For example, a customer who pauses for a long time may be carefully evaluating a product but the same behavior could also reflect confusion or uncertainty. To reduce this limitation, individual observations are grouped into broader, theoretically grounded behavioral dimensions. This makes the analysis less dependent on isolated actions and more aligned with the conceptual framework. However, the findings should still be interpreted as indicative patterns rather than definitive evidence of customer intention.

Another limitation is that customer behavior and purchase outcomes are influenced by external factors that cannot be fully controlled in a real-world retail environment. Variables such as promotions, staffing levels, product availability, store traffic and time of day may affect both behavior and conversion independently of the behavioral dimensions analyzed in the study. As a result, the analysis cannot determine the precise extent to which observed outcomes are driven by behavioral mechanisms versus situational retail conditions. The findings should therefore be interpreted as reflecting patterns within a dynamic store context rather than isolated behavioral effects. At the same time, this limitation reflects an important strength of the study. Since the analyses are conducted in a real-world retail setting rather than a controlled laboratory environment, the findings capture behavioral patterns under realistic commercial conditions. This increases the practical relevance of the results, even though it reduces the ability to isolate causal effects fully.

A further limitation is that the study is conducted as a single-case study within a physical fashion retail context, where the primary analysis is based on one focal store and complemented by observational data from two additional comparable stores. While this allows for a detailed and contextually grounded analysis of behavioral patterns and value creation, it limits the extent to which the findings can be generalised to other retail settings. Customer behavior and decision processes are likely to differ across retail categories, store formats and customer segments. In fashion retail, behaviors such as product interaction, comparison and fitting room use are particularly central, while in other contexts, such as grocery or convenience retail, customer journeys may be more routine and goal-directed. As a result, the relationship between behavior and value may vary across settings. The findings should therefore be interpreted as context-specific rather than universally applicable. At the same time, the case-based approach provides depth and allows for a more detailed connection between behavioral theory, in-store activity and store design decisions.

Despite these limitations, the study demonstrates how aggregated behavioral data, when combined with structured in-store observations, can be used to interpret patterns of visit-level purchase and estimated value within physical fashion retail. The limitations therefore define the appropriate scope and level of interpretation of the findings, rather than undermining the analytical contribution of the study.

5.5 Future Research

The limitations of this study also highlight several opportunities for future research. In particular, future studies could strengthen behavioral retail analytics by improving value measurement, testing causal mechanisms more directly, examining different retail contexts, and developing more dynamic behavioral segmentation approaches.

Future research could strengthen the measurement of value by linking behavioral data with actual transaction data. In this study, value is captured through visit-level proxies such as conversion and estimated value per visitor, as individual customers cannot be followed across the full purchase journey. A stronger approach would connect observed behaviors such as product interaction, revisits and fitting room use, to outcomes such as basket size, transaction value and repeat purchases. This would allow researchers to examine not only whether certain behaviors are associated with purchase, but also whether they are linked to higher economic value. It would also enable a more direct evaluation of ICP definitions, by identifying which behavioral patterns are associated with higher spending and longer-term customer value, providing a stronger basis for behavior-informed segmentation and store design decisions.

Future research could also apply experimental designs to test whether specific store changes influence customer behavior and value outcomes. While this study identifies associations between in-store behaviors and purchase outcomes, it does not establish whether changes in the store environment cause these behaviors to change. Future studies could therefore test targeted interventions such as changes in layout, product placement, fitting room capacity or merchandising displays. For example, if friction is associated with lower purchase likelihood, experiments could examine whether clearer navigation or reduced fitting room queues increase conversion. Similarly, if product engagement is linked to purchase, redesigned displays could be tested to increase evaluation behavior. Such designs would allow researchers to move beyond behavioral associations and directly test how store interventions shape customer behavior and value outcomes.

Future research could apply the same analytical approach in other retail contexts to assess whether the findings are specific to fashion retail or more broadly applicable. This study focuses on fashion, where behaviors such as product interaction, comparison and fitting room use are central to purchase decisions. In other retail categories, these behaviors may have different meanings. For example, grocery shopping is often more routine and goal-directed, while categories such as electronics or furniture may involve more extended comparison and evaluation. Applying the framework across different contexts would help identify which behavioral patterns are category-specific and which reflect more general mechanisms of value creation.

Future research could explore real-time behavioral segmentation as an alternative to static ICP models. In many retail settings, ICPs are predefined based on demographic characteristics, historical purchasing patterns, or broad customer categories. However, the findings of this study suggest that value creation may be more closely connected to observable behavioral progression during the store visit than to fixed customer profiles alone. A promising direction would therefore be to develop dynamic ICP models that adapt based on real-time customer behavior. Rather than classifying customers according to who they are, such models would classify customers according to what they are doing. Behaviors such as product interaction, revisits, fitting room use or hesitation could be used to identify customers progressing toward evaluation and purchase. This could enable retailers to adjust store operations in real time, for example through staff allocation, product recommendations or queue management based on observed behavioral states. Future research could examine whether such behavior-based segmentation provides stronger predictive power and more actionable insights than traditional static ICP approaches.

5.6 Conclusion

This study set out to examine whether value creation in physical retail is better explained by customer composition or by in-store behavior. The findings show that while customer composition, such as the retailer's ICP, can be analyzed within a useful descriptive framework, it has limited explanatory power as a standalone predictor of visit-level conversion. In contrast, behavioral patterns, particularly those reflecting product engagement, cognitive engagement and fitting room use, are more directly and consistently associated with purchase outcomes, while behaviors indicating friction are negatively related to purchase likelihood. These results suggest that value in physical fashion retail is more closely linked to how customers behave during the store visit than to predefined customer segments. Rather than providing causal proof, the study contributes a framework for using behavioral data to interpret visit-level value creation. The findings should be understood as context-specific, reflecting physical fashion retail environments and visit-based data conditions.

6. References

Baker, L. M. (2006). Observation: A complex research method. *Library Trends,* 55(1), 171–189. https://www.ideals.illinois.edu/items/3831

Bitner, M. J. (1992). Servicescapes: The impact of physical surroundings on customers and employees. *Journal of Marketing,* 56(2), 57–71. https://doi.org/10.1177/002224299205600205

Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The role of big data and predictive analytics in retailing. *Journal of Retailing, 93*(1), 79–95. https://doi.org/10.1016/j.jretai.2016.12.004

Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing, 73*(6), 1–17. https://doi.org/10.1509/jmkg.73.6.1

Dickson, P. R. (1982). Person–situation: Segmentation's missing link. *Journal of Marketing, 46*(4), 56–64. https://doi.org/10.1177/002224298204600407

Donovan, R. J., Rossiter, J. R., Marcoolyn, G., & Nesdale, A. (1994). Store atmosphere and purchasing behavior. *Journal of Retailing, 70*(3), 283–294. https://doi.org/10.1016/0022-4359(94)90037-X

Drèze, X., & Bonfrer, A. (2008). *An empirical investigation of the impact of communication timing on customer equity*. Journal of Interactive Marketing, 22(1), 36–50. https://doi.org/10.1002/dir.20103

Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing customers. Journal of Marketing Research, 41(1), 7–18. https://doi.org/10.1509/jmkr.41.1.7.25084

Hosseini, M., & Shabani, M. (2015). New approach to customer segmentation based on changes in customer value. *Journal of Marketing Analytics, 3*(3), 110–121. https://doi.org/10.1057/jma.2015.10

Hui, S. K., Fader, P. S., & Bradlow, E. T. (2009a). Path data in marketing: An integrative framework and prospectus for model building. *Marketing Science, 28*(2), 320–335. https://doi.org/10.1287/mksc.1080.0400

Hui, S. K., Fader, P. S., & Bradlow, E. T. (2009b). Research note: The traveling salesman goes shopping: The systematic deviations of grocery paths from TSP optimality. Marketing Science, 28(3), 566–572. https://doi.org/10.1287/mksc.1080.0402

Indivd. (2026a). DPIA guide: People counting. https://support.indivd.com/hc/en-gb/articles/19486223140370-DPIA-Guide-People-Counting

Indivd. (2026b). Indivd's approach to compliance: Development and validation history. https://support.indivd.com/hc/en-gb/articles/19919829002514-Indivd-s-approach-to-compliance-development-and-validation-history

Indivd. (2026c). What fitting room data reveals about store performance. https://www.indivd.com/stories/what-fitting-room-data-reveals-about-store-performance

Larson, J. S., Bradlow, E. T., & Fader, P. S. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing, 22*(4), 395–414. https://doi.org/10.1016/j.ijresmar.2005.09.005

Lee, H. S., Kesavan, S., & Deshpande, V. (2020). Managing the impact of fitting-room traffic on retail sales: Using labor to reduce phantom stockouts. *Manufacturing & Service Operations Management, 23*(6), 1580–1596. https://doi.org/10.1287/msom.2020.0884

Peck, J., & Childers, T. L. (2003). To have and to hold: The influence of haptic information on product judgments. *Journal of Marketing, 67*(2), 35–48. https://doi.org/10.1509/jmkg.67.2.35.18612

Peck, J., & Wiggins, J. (2006). It just feels good: Customers' affective response to touch and its influence on persuasion. Journal of Marketing, 70(4), 56–69. https://doi.org/10.1509/jmkg.70.4.056

Puccinelli, N. M., Goodstein, R. C., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer experience management in retailing: Understanding the buying process. *Journal of Retailing, 85*(1), 15–30. https://doi.org/10.1016/j.jretai.2008.11.003

Tenny, S., Brannan, G. D., & Brannan, J. M. (2022). Qualitative study. In *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK470395/

Tsiptsis, K., & Chorianopoulos, A. (2010). *Data mining techniques in CRM: Inside customer segmentation*. Wiley. https://doi.org/10.1002/9780470685815

Wedel, M., & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations* (2nd ed.). Springer. https://doi.org/10.1007/978-1-4615-4651-1

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

Zhang, J. Z., Chang, C.-W., & Neslin, S. A. (2022). How physical stores enhance customer value: The importance of product inspection depth. *Journal of Marketing, 86*(2), 166–185. https://doi.org/10.1177/00222429211012106

7. Appendix

AI Usage Disclosure

During the writing process of this thesis, generative AI tools, primarily ChatGPT developed by OpenAI, were used as a supportive aid in selected parts of the work. The tool was mainly used to assist with language-related tasks such as improving grammar, sentence structure, clarity and flow in already written text. In some cases, ChatGPT was also used to discuss alternative formulations, improve transitions between sections and support the structuring of arguments and explanations.

Generative AI was not used to independently generate the theoretical framework, empirical findings, statistical analyses or conclusions of the study. All analytical decisions, interpretations, coding procedures, statistical models and final evaluations were conducted and reviewed by the authors. Literature selection, theoretical positioning, methodological choices and interpretation of results were likewise carried out independently by the authors.

In addition, AI-generated suggestions were critically reviewed, modified and adapted before inclusion in the thesis. The authors remain fully responsible for the final content, academic quality, interpretations and conclusions presented in the study.

Appendix A1. Structured Observation Instrument

This appendix presents the structured observation sheet used to record in-store customer behavior during the manual observations. The items were used to construct the behavioral dimensions included in the observation-based analysis.

A. Basic Information

1. Observation ID: ____________
2. Date: ____________
3. Time: ____________
4. Customer type: ☐ Alone ☐ Pair ☐ Group
5. Shopping purpose (observed):
   ☐ Mainly for self
   ☐ Mainly for others
   ☐ Mixed / unclear

6. Gender
7. Age

B. Movement (Path Data)

1. Zone sequence (in order): ________________________

   (Example: Z1 → Z3 → Z5 → Z3 → Z6)

C. Behavioral Indicators

Mark based on observed behavior during the visit:

1. Touched products: ☐ Yes ☐ No
2. Compared products: ☐ Yes ☐ No
3. Revisited zone(s): ☐ Yes ☐ No
4. Used fitting room: ☐ Yes ☐ No
5. Interacted with staff: ☐ Yes ☐ No

D. Signs of Evaluation

1. Repeated product handling: ☐ Yes ☐ No

E. Signs of Friction

1. Long pause without action: ☐ Yes ☐ No
2. Repeated searching (looking around, scanning): ☐ Yes ☐ No
3. Picks up and abandons items multiple times: ☐ Yes ☐ No

F. Visit length (Estimate)

1. ☐ Short (<5 min)
2. ☐ Medium (5–15 min)
3. ☐ Long (>15 min)

G. Outcome (Observed)

1. Approached checkout: ☐ Yes ☐ No
2. Purchase observed: ☐ Yes ☐ No ☐ Unclear
3. Left with visible bag: ☐ Yes ☐ No ☐ Unclear

Appendix A2: Coding Notes for the Observation Instrument

General information

Each observation is for one customer/party at a time. The observation starts when they enter the store and ends when they leave the store.

Explanations of how questions should be interpreted

B. The store zone map follows the following classification:

- Z1 = Entrance / front display
- Z2 = Transition / walkway area
- Z3 = Key display / campaign zone
- Z4 = Core product section
    - a) Women
    - b) Men
    - c) Kids
    - d) Unisex
- Z5 = Fitting room
- Z6 = Checkout

A7 - Shopping purpose: explanation of the different alternatives:

- Mainly for self if customer looks at items likely for themselves (e.g. own size/style and gender)
- Mainly for others if customer looks at items likely for someone else (e.g. different size, gender or discussing another person)
- Mixed / unclear if cannot clearly tell or shows signs of both

C1 - Touched products: counted as YES if customer lifts, holds, unfolds, or inspects an item closely

C2 - Compared Products: counted as YES if customer holds or looks at 2+ items side-by-side or sequentially and clearly evaluates alternatives. Counted as NO if only browsing randomly.

C3 - Revisited Zone: counted as YES if customer returns to a previously visited zone

C4 - Used Fitting Room: counted as YES if customer enters fitting room area with intent to try

C5 - Interacted with Staff: counted as YES if customer asks for help OR staff engages in meaningful interaction

D1 - Repeated product handling: counted as YES if the customer picks up and interacts with the same product multiple times. Example: picks it up → puts it back → picks it up again. Counts as NO if the customer only handles the item once or if the customer picks up an item once and then keeps it and goes to checkout. This measures hesitation or evaluation not a decisive action.

E1 - Long pause without action: counted as YES if Standing still > ~10 seconds without clear action

E2 - Repeated searching (looking around, scanning): counted as YES if looking around repeatedly. Scanning shelves/racks without selecting.

E3 - Picks up and abandons items multiple times: counted as YES if customer picks up an item, inspects, puts back and repeats more than two times

Appendix B. Zone Scorecard Calculations

This appendix presents the input measures used to construct the zone scorecard. Table B1 shows weighted ICP share by zone, Table B2 shows weighted conversion rate by zone, and Table B3 presents the color-classification thresholds used in the scorecard.

Table B1. Weighted ICP Share by Zone

Zone	Total qualified visitors	Weighted ICP share
Home 2	19888	78.5%
Women 4	32942	74.8%
Men 3	18875	71.6%
Tills (0)	7649	71.3%
Women 3	34910	70.7%
Floor 0	16441	68.3%
Women 7	19465	66.6%
Store	10299	66.0%
Women 1	37128	65.1%
Men 1	14083	64.9%
Fitting room queue	16963	64.8%
Women 2	3677	61.6%
Fitting room	3597	59.6%
Men 2	1884	59.0%
Floor -1	6126	57.4%
Tills (-1)	5369	57.1%
Home 1	5772	55.8%
Men 4	8050	52.5%
Women 8	2713	50.5%
Café	3520	49.1%
Children 1	2050	33.9%
Women 6	2663	30.7%

Table B2. Weighted Conversion Rate by Zone

Zone	Weighted conversion rate
Fitting room	12.0%
Men 2	9.6%
Home 2	7.7%
Women 4	7.4%
Women 6	6.9%
Home 1	6.8%
Women 8	6.4%
Men 1	6.3%
Men 4	5.7%
Women 3	5.7%
Women 7	5.6%
Children 1	5.5%
Women 2	5.0%
Fitting room queue	4.6%
Women 1	4.0%
Men 3	3.1%
Café	2.7%

Table B3. Zone Scorecard Color Classification

ICP share >= 68.7%
ICP share 62-68.7%
ICP share 51.5-62%
ICP share < 51.5%
Conv rate >= 7.15%
Conv rate 5.7-7.15%
Conv rate 4.8-5.7%
Conv rate < 4.8%

Appendix C. ICP Regression Output

This appendix shows the regression results used to evaluate whether ICP share is associated with conversion outcomes. Table C1 summarizes the merged dataset, Table C2 shows the row-level regression using total ICP share, and Table C3 shows the zone-level regression using weighted ICP share. Total ICP share is statistically significant at the row level ($p = 0.001$) but explains very little variation ($R^2 = 0.003$), while weighted ICP share is not significant at the zone level ($p = 0.203$).

Table C1. Merged Dataset Used for ICP Regression

Measure	Value
Total merged rows	5459
Rows with both ICP and conversion	4280
Percentage usable	78.4%

Table C2. Row-Level Regression of ICP Share on Conversion

Variable	Coefficient	Standard error	t-value	p-value	95% confidence interval
Constant	0.069	0.003	25.533	<0.001	[0.064, 0.075]
Total ICP share	0.004	0.001	3.470	0.001	[0.002, 0.007]

Model information	Value
Model	Ordinary Least Squares regression
Dependent variable	Conversion rate to any till
Independent variable	Total ICP share
Number of observations	4280
R-squared	0.003
Adjusted R-squared	0.003
F-statistic	12.04
Prob. F-statistic	0.0005
Degrees of freedom, model	1
Degrees of freedom, residuals	4278
Covariance type	Nonrobust

Table C3. Zone-level OLS regression of weighted ICP share on weighted conversion rate

Table C3a. Regression coefficients

Variable	Coefficient	Standard error	t-value	p-value	95% confidence interval
Constant	0.056	0.014	3.986	0.001	[0.026, 0.086]
Weighted ICP share	0.020	0.015	1.331	0.203	[-0.012, 0.052]

Table C3b. Model information

Model information	Value
Model	Ordinary Least Squares regression
Dependent variable	Weighted conversion rate
Independent variable	Weighted ICP share
Number of observations	17
R-squared	0.146
Adjusted R-squared	0.046
F-statistic	1.771
Prob. F-statistic	0.203
Degrees of freedom, model	1
Degrees of freedom, residuals	15
Covariance type	Nonrobust

Appendix D. Behavior-Based Value Regression

This appendix presents the full regression output for the aggregated individual analysis examining the relationship between broad behavioral engagement and the retailer's estimated value per visitor. Value per visitor is based on the retailer's internal valuation model and should therefore be interpreted as an estimated value measure rather than actual transaction-level spending.

Table D1. OLS regression of engagement on estimated value per visitor

(The table shows the OLS regression output including coefficients, standard errors, and z-values. Store-zone controls, weekend, and time segment are included in the model but not reported individually.)

Table D1a. Regression coefficients

Variable	**Coefficient**	**Standard error**	**z-value**	**p-value**	**95% confidence interval**
Intercept	246.098	23.031	10.685	<0.001	[200.96, 291.24]
Engagement rate	440.725	10.460	42.101	<0.001	[420.26, 461.24]
Log visitors	−9.886	3.843	−2.572	0.010	[-17.42, -2.35]
Weekend	−5.738	6.686	−0.858	0.391	[-18.84, 7.37]
Time segment: Evening	76.997	5.987	12.862	<0.001	[65.26, 88.73]
Time segment: Lunch	11.176	7.211	1.550	0.121	[-2.956, 25.309]
Store-zone controls	Included				

Table D1b. Model information

Model information	Value
Model	Ordinary Least Squares regression
Dependent variable	Estimated value per visitor
Main independent variable	Engagement rate
Controls	Weekend, time segment, log visitors, and store-zone controls
Number of observations	6993
R-squared	0.282
Adjusted R-squared	0.279
F-statistic	99.93
Prob. F-statistic	<0.001
Degrees of freedom, model	26
Degrees of freedom, residuals	6966
Covariance type	HC3 robust standard errors

Appendix E. Correlation Matrix for Observation Variables

This appendix presents the correlation matrix used to assess whether the behavioral dimensions in the observation-based analysis were too strongly correlated to be included in the same model. The variables were retained because none of the correlations indicated severe multicollinearity.

Table E1. Correlation Matrix of Behavioral Dimensions

Variable	Product engagement	Cognitive engagement	High intent	Friction	Exploration
Product engagement	1.000	0.574	0.503	0.106	0.522
Cognitive engagement	0.574	1.000	0.555	0.175	0.540
High intent	0.503	0.555	1.000	−0.174	0.485
Friction	0.106	0.175	−0.174	1.000	0.117
Exploration	0.522	0.540	0.485	0.117	0.100

Appendix F. Logistic Regression Output

This appendix presents the full logistic regression output for the observation-based analysis. The dependent variable is observed purchase, and the independent variables are the five behavioral dimensions: product engagement, cognitive engagement, high intent, friction and exploration.

Table F1. Logistic Regression Predicting Observed Purchase

Variable	Coefficient	Standard error	z-value	p-value	Odds ratio	95% confidence interval
Intercept	−3.321	1.000	−3.321	0.001	-	[-5.280, -1.361]
Product engagement	1.643	0.550	2.985	0.003	5.17	[0.564, 2.722]
Cognitive engagement	1.270	0.452	2.813	0.005	3.56	[0.385, 2.155]
High intent	1.987	0.751	2.645	0.008	7.30	[0.515, 3.459]
Friction	−0.698	0.309	−2.259	0.024	0.50	[-1.303, -0.092]
Exploration	−0.043	0.155	−0.279	0.781	0.96	[-0.347, 0.261]

Model information	Value
Model	Logistic regression
Dependent variable	Observed purchase
Independent variables	Product engagement, cognitive engagement, high intent, friction and exploration
Number of observations	124
Pseudo R-squared	0.4898
Log-likelihood	−43.445
LLR p-value	<0.001
Degrees of freedom, model	5
Degrees of freedom, residuals	118
Converged	Yes
Covariance type	Nonrobust

Appendix G. Supplementary Linear Probability Model

This appendix presents the supplementary linear probability model used as a robustness check for the observation-based logistic regression. The model uses the same dependent variable and behavioral dimensions as the logistic regression.

Table G1. Supplementary Linear Probability Model Predicting Observed Purchase

Variable	Coefficient	Standard error	t-value	p-value	95% confidence interval
Constant	0.444	0.031	14.148	<0.001	[0.381, 0.506]
Product engagement	0.157	0.041	3.815	<0.001	[0.076, 0.239]
Cognitive engagement	0.150	0.042	3.554	0.001	[0.066, 0.233]
High intent	0.131	0.038	3.435	0.001	[0.056, 0.207]
Friction	-0.076	0.033	-2.293	0.024	[-0.142, -0.010]
Exploration	0.003	0.040	0.077	0.939	[-0.076, 0.082]

Model information	**Value**
Model	Linear probability model
Dependent variable	Observed purchase
Independent variables	Product engagement, cognitive engagement, high intent, friction and exploration
Number of observations	124
R-squared	0.530
Adjusted R-squared	0.510
F-statistic	26.62
Prob. F-statistic	<0.001
Degrees of freedom, model	5
Degrees of freedom, residuals	118
Covariance type	Nonrobust