



Final Project 2

PMG 497

Midocean University

Faculty of Management

Department/ Specialization



Brand Loyalty in Retail Stores in Saudi Arabia

Adham Ehab Elkhoully

.....

Final Project - 2

Faculty of Management - Midocean University

Supervisor

Dr.

Faculty member- Faculty of management- Midocean University

1/8/2026 – Second Semester

Acknowledgment

Dedication

Abstract

Purpose:

Methodology:

Findings:

Originality & Implication:

Keywords:

Table of Contents

Acknowledgment	2
Dedication	4
Abstract	5
1. Chapter 1: Introduction	9
1.1. Research Background.....	11
1.2. Research Problem.....	12
1.2.1. Theoretical Gaps.....	13
1.2.2. Practical Gaps.....	14
1.3. Research Significance.....	15
1.3.1. At the theoretical level.....	15
1.3.2. At the practical level.....	16
1.4. Research Questions.....	17
1.5. Research Objectives.....	18
1.6. Research Hypotheses.....	18
1.7. Research Model.....	19
1.8. Research Methodology.....	20
1.9. Scope of The Study.....	21
1.10. The Suggested Project Structure.....	21
2. Chapter 2: Literature Review	23
2.1. Introduction.....	24
2.2. The independent variable.....	25
2.3. The dependent variable.....	26
2.4. The relationship between the independent and the dependent variables.....	27
2.5. Conclusion.....	28
3. Chapter 3: Research Methodology	29
3.1. Introduction.....	26
3.2. Operational definitions for research variables.....	15

3.3.	Research design (research onion).....	15
3.4.	Research population & Sampling methods and technique.....	16
3.5.	Data collection method.....	16
3.6.	Types of data.....	16
3.7.	Statistical analysis procedures.....	17
3.8.	Conclusion.....	17
4.	Chapter 4: Statistical Analysis & Findings discussion	18
4.1.	Introduction.....	18
4.2.	Sample profile (descriptive statistics for research sample).....	19
4.3.	Validity & reliability tests for the questionnaires.....	19
4.4.	Descriptive statistics for research variables (answer research questions).....	19
4.5.	Testing Hypotheses & Findings Discussion.....	20
4.6.	Conclusion.....	20
5.	Chapter five: Research limitations, Recommendations & Conclusion	21
5.1.	Research Limitations & Future Research.....	21
5.2.	Research Recommendations.....	22
	References	22
	Appendix	23

List of Figures:

List of Tables:

1. Chapter 1: Introduction

Brand Loyalty is now considered as one of the most critical factors in the long term success of the retail businesses worldwide.

In a growing competitive market place where consumers are always exposed to a vast array of choices, the ability for a retail brand to own and sustain loyalty among its consumer base, has become a key differentiator.

Saudi Arabia's retail sector has undergone huge transformation in the recent decades, presenting a very interesting scene to examine this phenomenon, given the country's cultural, economic and demographic characteristics.

Kingdom of Saudi Arabia is witnessing an unprecedented growth in the retail industry, driven by factors such as young and digitally connected population, rising disposable incomes and adding also to that the ambition Vision 2030 that aims to diversify the economy away from oil dependency. These forces joined, reshaped the consumer behaviour and expectations, making it crucial for retailers to understand what drives the brand loyalty with this distinctive market environment.

1.1. Research Background

Concept of brand has a rich history within marketing literature, originating in the early 20th century as a straightforward practice of measuring repeated purchases behaviour and has since expanded to a multidimensional construct encompassing attitudinal commitment, emotional attachment and behavioural intentions.

Scholars as Oliver (1999) have proposed that loyalty might not be only limited to habitual repurchase, but can hold commitment to rebuy or re-patronize a preferred product/service in a consistent way in the future, which leads to repetitive same brand purchasing despite of situational influences or marketing efforts that can trigger switching behaviour.

The Saudi Arabian retail business evolved significantly since the 1990s, transitioning from small, and family owned establishments to large scale multinational retail formats that includes hypermarkets, shopping malls, speciality stores and e-commerce platforms.

International brands as Zara, H&M, IKEA have established a very strong presence alongside the homegrown retailers like Panda, Al Othaim, Jarir and Extra.

This coexistence of global and local brands made the competition intensify and raise the importance of the brand loyalty as a mechanism for sustaining the competitive advantage.

Saudi Arabia's Unique social and cultural context distinguishes it as an interesting and compelling research setting.

The country's islamic values, tribal affiliations, family oriented consumer behaviour and strong affirmation on social status and prestige

all of these factors influence the purchasing decisions in a way that differs totally from the western markets, adding to that the predominance of young population and smartphone penetration rates with active social media engagement, created a generation of consumers whose loyalty is shaped by digital experience alongside in-store interactions.

1.2. Research Problem

Despite the growing strategic importance of brand loyalty in the Saudi retail context, there remains a significant lack of comprehensive empirical research that systematically examines the factors driving brand loyalty in this specific market. Much of the existing literature on brand loyalty either focuses on Western consumer markets or examines broad emerging market contexts without addressing the particularities of the Gulf Cooperation Council (GCC) or Saudi Arabian retail environment. This gap undermines the ability of retail managers and policymakers in Saudi Arabia to develop evidence-based strategies for cultivating lasting customer relationships.

Furthermore, the rapid digital transformation of the Saudi retail sector—accelerated by the COVID-19 pandemic, which prompted a dramatic shift toward online shopping—has introduced new dimensions of brand loyalty that traditional frameworks may not adequately capture. The interplay between omnichannel retail experiences, digital brand engagement, customer trust in online platforms, and physical store loyalty presents a complex research problem that requires updated theoretical and empirical frameworks. Retailers are struggling to understand how loyalty earned in physical stores translates, or fails to translate, to the digital shopping environment.

The research problem addressed by this study is therefore twofold: first, to identify and validate the key antecedents of brand loyalty in Saudi retail stores; and second, to assess whether established theoretical frameworks developed in other cultural contexts adequately explain brand loyalty behavior among Saudi consumers. By addressing these concerns, the study aims to fill a critical void in the literature and provide actionable insights for stakeholders in the Saudi retail industry.

1.2.1. Theoretical Gaps

From a theoretical standpoint, the existing literature has not sufficiently addressed how cultural dimensions specific to Saudi Arabia—such as collectivism, religious values, and national identity—interact with conventional brand loyalty antecedents like service quality, customer satisfaction, and perceived value. Most theoretical models of brand loyalty, including those proposed by Aaker (1991), Oliver (1999), and Chaudhuri and Holbrook (2001), were developed and validated primarily in Western, individualistic cultural settings, limiting their direct applicability to the Saudi context.

Additionally, there is a notable theoretical gap concerning the role of digital and social media factors in shaping brand loyalty in Saudi Arabia, which boasts one of the highest social media usage rates in the world. The theoretical integration of electronic word-of-mouth, influencer marketing, and personalized digital experiences into brand loyalty models remains underdeveloped, particularly within the GCC retail context. Existing models have yet to satisfactorily account for how these digital touchpoints moderate or mediate the relationship between traditional loyalty antecedents and actual loyalty outcomes.

A further theoretical gap exists in relation to the moderating role of demographic variables such as gender, age, and income level in the Saudi retail brand loyalty context. Saudi Arabia has recently undergone substantial social changes, including the relaxation of gender-segregation policies in public spaces and the increased economic participation of women, which may have significant implications

for brand loyalty patterns that have not yet been theoretically or empirically explored in the existing literature.

1.2.2. Practical Gaps

In terms of practical gaps, Saudi retail managers frequently lack sector-specific data and insights about the loyalty drivers most relevant to their consumer base. While multinational retailers operating in Saudi Arabia may draw on global best practices, these are often not calibrated to local market realities, leading to the implementation of loyalty programs and marketing strategies that fail to resonate with Saudi consumers. The absence of locally validated measurement tools for brand loyalty further compounds this challenge for practitioners.

There is also a practical gap related to the effectiveness of loyalty programs in the Saudi retail market. Despite the widespread adoption of loyalty cards, reward points, and membership programs by major Saudi retailers, there is limited empirical evidence assessing their actual impact on consumer loyalty behavior in this context. Practitioners are therefore operating without a clear understanding of the return on investment of these initiatives or how they compare in effectiveness with other loyalty-building strategies such as personalized service, store atmosphere, and community engagement.

Finally, small and medium-sized retailers in Saudi Arabia—which constitute a significant portion of the retail sector—face a practical gap in that they lack access to the research and analytical resources needed to understand and cultivate brand loyalty. Without targeted, context-specific guidance, these businesses are at a disadvantage compared to large multinational chains. This study seeks to provide practical insights applicable to retailers of all scales operating within the Kingdom.

1.3. Research Significance

This research significance comes from both the theoretical contribution and the academic literature on the brand loyalty and the practical implications on the Saudi retail industry.

By conducting rigorous investigation within a unique and under studied culture and market context, this study advances the understanding of how the brand loyalty is formed and sustained in a dynamic emerging retail market.

Findings are expected to enrich the global brand loyalty literature with local culture insights that challenge and refine existing frameworks.

From a wider perspective, this study will contribute to the field of consumer behaviour by demonstrating how the macro level factors interact with micro level individual consumer variables to shape the brand loyalty outcomes.

This integration perspective is valuable at a time when researchers and practitioners are struggling with the complexity of consumer decision making in the multichannel retail across the globe.

1.3.1. At the theoretical level

At the theoretical level, this research makes several important contributions.

First, it extends established brand loyalty models by incorporating culturally specific variables that reflect the Saudi consumer context, thereby enhancing the cross-cultural generalizability of these models.

By empirically testing the applicability of frameworks such as Oliver's (1999) four-stage loyalty model and Aaker's (1991) brand equity framework in the Saudi retail environment, the study provides critical evidence for or against their universal validity.

Second, the study contributes to theory by developing and testing a culturally adapted conceptual model of brand loyalty that integrates traditional antecedents—such as service quality, perceived value, customer satisfaction, and trust—with emerging digital and social factors specific to the Saudi context. This integrated model represents a novel theoretical

contribution that future researchers can employ and build upon in other GCC and Middle Eastern market contexts.

Third, the research contributes methodologically by validating or developing measurement scales appropriate for the Saudi retail context.

The creation of culturally validated instruments is an important contribution in itself, as it provides researchers working in similar cultural settings with reliable and valid tools for future empirical investigations. This methodological contribution enhances the rigor and comparability of future brand loyalty research in the Arab world.

1.3.2. At the practical level

At the practical level, this research provides Saudi retail managers with empirically grounded insights into the most influential drivers of brand loyalty among their target consumers. Armed with this knowledge, managers can make more informed decisions about resource allocation, customer relationship management strategies, loyalty program design, and service quality improvement initiatives. The study's findings will help retailers prioritize the factors that yield the greatest return in terms of building lasting customer loyalty.

For marketing practitioners, the research offers specific guidance on how to craft loyalty-enhancing communication strategies that resonate with Saudi consumers' cultural values, aspirations, and digital behaviors.

Understanding the relative importance of emotional versus functional loyalty drivers, for instance, can inform the tone, content, and channel selection for loyalty marketing campaigns. Similarly, insights into the role of social influence and community membership in shaping loyalty can guide the design of referral programs and social media engagement strategies.

For investors and entrepreneurs entering the Saudi retail market, this research provides a valuable orientation to the competitive dynamics of brand loyalty in the Kingdom. Understanding what makes Saudi consumers loyal to particular retail brands is essential intelligence for new market entrants seeking to establish a foothold and grow a sustainable customer base in one of the Middle East's most promising retail markets.

1.4. Research Questions

This study is guided by a set of carefully formulated research questions designed to address the theoretical and practical gaps in a systematic manner. The primary research question asks: What are the key factors that influence brand loyalty among consumers in retail stores in Saudi Arabia? This question is supported by a series of subsidiary questions that explore specific dimensions of the brand loyalty phenomenon in the Saudi retail context.

The subsidiary research questions include

- 1) What is the relationship between customer satisfaction and brand loyalty in Saudi retail stores?
- 2) How does perceived service quality influence brand loyalty among Saudi consumers? 3) What role does brand trust play in mediating the relationship between satisfaction and loyalty?
- 4) To what extent do digital engagement and social media interactions contribute to brand loyalty in the Saudi retail sector?
- 5) Do demographic variables such as age, gender, and income moderate the relationships between brand loyalty antecedents and outcomes?

These research questions collectively provide a structured framework for the investigation.

By addressing each question through the analysis of primary survey data collected from Saudi retail consumers, the study will be able to construct a comprehensive picture of the brand loyalty landscape in Saudi Arabia. The answers to these questions

will form the empirical core of the research findings and will directly inform the theoretical and practical contributions described in the previous sections.

1.5. Research Objectives

The primary objective is to analyze and identify the most important determinants of brand loyalty between customers in the Saudi Arabian retail stores.

This objective will be operationalized through a set of sub objectives that will provide a comprehensive and also multi-faceted understanding from the brand loyalty phenomenon in this particular market.

The research objectives are:

- 1) Review and synthesize the theoretical and empirical literature on brand loyalty that already exists, with attention to studies conducted in emerging and culturally diverse markets.
- 2) Develop a culturally adapted model of the brand loyalty that is appropriate for the Saudi retail context
- 3) Design and validate measurement indicators for the key constructs pre identified in the conceptual model
- 4) Collect and analyze data from consumers in Saudi Arabia to test the model
- 5) Draw theoretically meaningful and useful conclusions from the empirical findings

1.6. Research Hypotheses

Based on the framework in chapter 2 and the conceptual model in section 1.7, this study proposes the following hypotheses for testing.

H1: Service quality has a significant positive effect on customer satisfaction in Saudi retail stores

H2: Customer satisfaction has a significant positive effect on brand loyalty in Saudi retail stores.

H3: Brand trust mediates the relationship between customer satisfaction and brand loyalty.

H4: Perceived value has a significant positive effect on brand loyalty.

H5: Digital engagement with a retail brand positively influences brand loyalty among Saudi consumers.

H6: Demographic variables (age, gender, income) significantly moderate the relationship between brand loyalty antecedents and brand loyalty outcomes.

These hypotheses are designed to be testable through quantitative analysis of the primary survey data.

The direction of each relationship is supported by prior research findings and theoretical reasoning.

1.7. Research Model

The conceptual model in this research integrates multiple perspectives to provide comprehensive framework for understanding brand loyalty in Saudi Arabian retail context.

The model points that the brand loyalty comprising both attitudinal and behavioural loyalty are influenced by a set of interconnected variables that operate at the functional and operational level of the consumer experience.

The independent variables in the model include service quality, perceived value, customer satisfaction, brand trust, and digital engagement.

Service quality and perceived value are hypothesized to exert direct effects on both customer satisfaction and brand loyalty, while brand trust is conceptualized as a mediator that channels the effects of satisfaction on loyalty.

Digital engagement is included as a contemporary antecedent that reflects the growing importance of online and social media interactions in shaping Saudi consumers' brand relationships. Demographic variables are incorporated as moderators that may amplify or attenuate the main effects.

The model draws on several theoretical foundations, including Oliver's (1999) satisfaction-loyalty framework, Parasuraman, Zeithaml, and Berry's (1988) SERVQUAL model of service quality, Chaudhuri and Holbrook's (2001) brand trust and brand affect model, and more recent frameworks addressing digital brand engagement. By synthesizing these perspectives into a single integrated model tailored

to the Saudi context, the research provides a theoretically robust and empirically testable representation of the brand loyalty formation process in Saudi retail stores.

1.8. Research Methodology

This study uses a quantitative research design to test the relationships in the conceptual model. Data will come from a structured, self-administered questionnaire given to adult consumers who have shopped at retail stores in Saudi Arabia. The questionnaire is based on established brand loyalty scales, adapted for the Saudi cultural and retail context, and will be available in both Arabic and English to make it accessible and encourage quality responses.

The survey will be shared online and in person. Online, it will be distributed through social media and consumer panels to reach a broad and representative group of Saudi retail consumers. In-person surveys will be conducted at selected retail stores in major cities such as Riyadh, Jeddah, and Dammam to ensure geographic and retail diversity. The goal is to collect responses from 400 participants, which is enough for structural equation modeling (SEM) analysis.

Data analysis will be Data analysis will take place in two main stages. First, descriptive statistics and confirmatory factor analysis (CFA) will be used to check the reliability and validity of the measurement tools. Next, structural equation modeling (SEM) with AMOS or SmartPLS software will test the relationships in the conceptual model and evaluate the proposed hypotheses. Moderation analyses will use multigroup SEM or hierarchical regression to examine how demographic variables affect the results. This approach aims to produce rigorous, reliable, and generalizable findings.

1.9. Scope of The Study

This research defines its scope by theme, location, time, and demographics. The main focus is brand loyalty in Saudi Arabian retail stores, including supermarkets, hypermarkets, specialty stores, fashion retailers, and electronics retailers. The study does not cover service industries like banking, hospitality, or telecommunications, but relevant insights from these sectors are included in the literature review.

The research is limited to Saudi Arabia, with data collected from consumers in Riyadh, Jeddah, Makkah, Madinah, and Dammam. These cities were chosen because they represent a wide range of Saudi consumers and most of the country's retail activity. Although the results are most relevant to Saudi Arabia, the study's theories and methods may also help researchers looking at brand loyalty in other GCC and Arab countries.

Temporally, the study's primary data collection is planned for the year 2026, capturing a snapshot of Saudi consumer behavior. The study will collect its main data in 2026, aiming to capture Saudi consumers' brand loyalty attitudes and behaviors during a time of major change in the retail sector. It will focus on Saudi adults aged 18 and over who have shopped at retail stores in the past six months. This ensures that participants have recent experience to provide useful information about what influences brand loyalty and its results.

1.10. The Suggested Project Structure

This research paper is divided into five chapters, with each chapter building on the last to explore brand loyalty in Saudi retail stores. Chapter 1, which you are reading now, introduces the background, research problem, significance, questions, objectives, hypotheses, model, methodology, and scope. It explains why this research is important and describes the framework for the study.

Chapter 2 reviews the main academic literature on brand loyalty, covering theories, key factors, and findings from different cultures and markets. This review highlights gaps in current knowledge and supports the study's conceptual model.

Chapter 3 explains the research methods, including design, sampling, questionnaire development, data collection, and analysis. Chapter 4 shares the study's results, including descriptive analysis, confirmatory factor analysis, and structural equation modeling. Chapter 5 discusses what the findings mean in theory and practice, summarizes the conclusions, notes the study's limitations, and suggests areas for future research.

References are provided at the end of the document, following the American Psychological Association (APA) 7th edition citation format. Appendices, if applicable, will include the survey questionnaire in both Arabic and English, along with extra statistical results that support Chapter 4. This structure aims to make the research clear, logical, and complete.

2.Chapter 2: Literature Review

2.1. Introduction

Brand loyalty is considered one of the most valuable assets for retail firms. Since the 1950s, when researchers have studied how repeat purchasing sets loyal customers apart from others. Early ideas from Brown (1953) and Cunningham (1956) focused on behavior, defining loyalty as simply buying the same brand repeatedly. However, this view was criticized for not separating true loyalty, which comes from real preference and commitment, from spurious loyalty that happens out of habit, inertia, or limited choices.

Later theories added attitudinal aspects to the idea of brand loyalty. Researchers realized that truly loyal customers not only buy the same brand repeatedly but also have positive attitudes, strong emotional connections, and are less likely to switch to competitors. Scholars like Day (1969), Jacoby and Chestnut (1978), and Dick and Basu (1994) developed a framework that identified four types of loyalty: true loyalty, latent loyalty, spurious loyalty, and no loyalty. These ideas have shaped how brand loyalty is understood, measured, and managed in both research and business.

Today, brand loyalty has changed because of new technology, more shopping options, and shifting consumer expectations. E-commerce, mobile shopping, social media, and loyalty platforms have introduced new ways for customers to connect with brands. In Saudi Arabia, the mix of traditional values and modern digital habits offers a unique setting to study how these changes affect the basic reasons people stay loyal to brands.

2.2. The independent variable

The independent variables in this study's conceptual model are service quality, perceived value, customer satisfaction, brand trust, and digital engagement. These factors are the main ways brand loyalty is built in retail. Service quality, as described in the SERVQUAL model by Parasuraman, Zeithaml, and Berry (1988), means how consumers judge the overall excellence or superiority of a service compared to what they expect. The SERVQUAL model identifies five dimensions of service quality: reliability, assurance, tangibles, empathy, and responsiveness. These have been widely studied in retail and are consistently found to predict customer loyalty.

Perceived value is another important independent variable in the model. It is defined as how consumers judge the usefulness of a product or service by comparing what they get to what they give. Zeithaml (1988) introduced this idea as a key link between quality and loyalty, noting that consumers consider quality along with the price and effort needed to get the product or service. In Saudi retail, perceived value may be especially important because major retailers use competitive pricing and some consumer groups are very price-sensitive.

Brand trust, which is the willingness of consumers to rely on a brand to do what it promises, has been widely studied as both a cause and a link to brand loyalty. Morgan and Hunt (1994) said that trust is essential for any long-term relationship between a brand and its customers, as it gives people the confidence to buy again. In Saudi Arabia, trust may be shaped by cultural values like honesty and reliability, which are important in Islamic business ethics. This makes trust an important variable to study in the model.

2.3. The dependent variable

In this study, brand loyalty is the dependent variable and includes both attitudinal and behavioral aspects. Attitudinal loyalty is when a consumer feels positively about a retail brand, is emotionally connected, and stays committed by preferring, recommending, and not switching brands. Behavioral loyalty is seen in actions such as repeat purchases, higher spending, and continued shopping at the same store.

This study uses Oliver's (1999) four-stage model to define brand loyalty: cognitive, affective, conative, and action loyalty. Cognitive loyalty is when someone prefers a brand based on its features. Affective loyalty develops from repeated positive experiences with the brand. Conative loyalty means a strong commitment to buy the brand again, driven by positive feelings. Action loyalty is the intention to repurchase, even if there are obstacles.

In the context of Saudi retail stores, measuring brand loyalty requires particular attention to the cultural specificities of how loyalty is expressed and communicated among Saudi consumers. For instance, word-of-mouth recommendations—a key indicator of attitudinal loyalty—may operate differently in the Saudi cultural context, where family and social network endorsements carry especially high weight in shaping purchasing decisions. Similarly, the willingness to pay a premium price for a preferred retail brand, another indicator of behavioral loyalty, must be assessed in relation to the specific economic circumstances and value orientations of the Saudi consumer population.

2.4. The relationship between the independent and the dependent variables

Researchers have studied how service quality, perceived value, customer satisfaction, brand trust, and digital engagement relate to brand loyalty. Service quality is often seen as one of the strongest predictors of brand loyalty in retail. Studies by Zeithaml, Berry, and Parasuraman (1996) show that better service quality leads to higher customer retention, more positive word-of-mouth, and a greater willingness to pay premium prices, all of which signal brand loyalty. This link has been found in many types of retail and across different cultures, supporting its use in this model.

Research shows that customer satisfaction is a key link between service quality, perceived value, and brand loyalty. The connection between satisfaction and loyalty is not straightforward. Usually, dissatisfaction hurts loyalty more than satisfaction helps it. This means that avoiding service failures may be more important for keeping loyal customers than providing exceptional service. In Saudi Arabia, where service expectations are influenced by both Western and local traditions, this relationship may have unique features that deserve further study.

Brand trust is important for building and maintaining loyalty because it gives customers a sense of security and confidence to keep choosing the same brand. Chaudhuri and Holbrook (2001) found that both brand trust and brand affect help explain the link between brand loyalty, purchase loyalty, and attitudinal loyalty. This means that both feelings and thoughts about a brand shape loyalty. In Saudi Arabia, where trust and reputation are especially valued, brand trust is likely to play an even bigger role in creating loyalty.

2.5. Conclusion

This chapter has provided a comprehensive review of the theoretical and empirical literature relevant to the study of brand loyalty in retail stores, with particular emphasis on the factors most pertinent to the Saudi Arabian context. The review has established that brand loyalty is a complex, multidimensional construct influenced by a constellation of interrelated antecedents, including service quality, perceived value, customer satisfaction, brand trust, and digital engagement. The relative importance of these antecedents, and the mechanisms through which they influence loyalty, may vary across cultural and market contexts, underscoring the value of conducting context-specific empirical research such as the present study.

The review has also identified several important theoretical gaps that this study seeks to address, including the limited cross-cultural validation of brand loyalty models in the Saudi and broader Arab world context, the insufficient theoretical integration of digital engagement as a loyalty antecedent, and the lack of attention to the moderating effects of cultural and demographic variables in existing loyalty frameworks. By addressing these gaps, the study aims to make meaningful contributions to the brand loyalty literature and provide theoretically grounded, practically useful insights for the Saudi retail industry.

The theoretical foundations laid out in this chapter will inform the development of the research methodology in the following chapter and provide the interpretive framework through which the empirical findings will be discussed. The next chapter will detail the research design, sampling approach, data collection procedures, and analytical techniques that will be employed to rigorously test the conceptual model and hypotheses presented in Chapter 1, ensuring that the study's findings are both internally valid and externally generalizable.

3. Chapter 3: Research Methodology

3.1. Introduction

The preceding chapters established the aims and significance of this study. Chapter 1 defined the research problem, questions, objectives, and hypotheses, and introduced the conceptual model positioning service quality, perceived value, customer satisfaction, brand trust, and digital engagement as antecedents of brand loyalty in Saudi retail stores. Chapter 2 situated this model within the existing literature, tracing the evolution of brand loyalty theory from behavioural to composite attitudinal–behavioural frameworks, and identifying the cross-cultural and digital gaps addressed by the present study. The purpose of this chapter is to translate the conceptual framework into an actionable empirical plan. It outlines how each construct will be operationalized, the philosophical and design choices guiding the inquiry, the data sources and collection methods, and the statistical procedures for testing the six hypotheses.

Methodological transparency is regarded as a substantive requirement rather than a procedural formality. To credibly assess the applicability of Western-developed loyalty frameworks in the Saudi retail context, the study must explicitly detail its measurement instruments, sampling rationale, and analytical decision rules to enable replication. Accordingly, this chapter explains not only the methodological choices made but also the rationale for rejecting alternatives, and it acknowledges the limitations each decision imposes on result interpretation. This approach aligns with the principle of methodological alignment, which requires that philosophy, approach, strategy, time horizon, and technique form a coherent and consistent methodological chain (Saunders et al., 2019).

This chapter is organized into seven substantive sections. Section 3.2 provides operational definitions for each construct in the conceptual model and identifies the source scales for measurement items. Section 3.3 presents the research design, following the layered structure of the research onion from philosophical stance to data collection technique. Section 3.4 defines the target population and justifies the sampling method, technique, and size. Section 3.5 outlines the data collection instrument and procedure, including the translation protocol and adopted ethical safeguards. Section 3.6 categorizes the types of data generated and the levels of measurement, as these determine the appropriate statistical analyses. Section 3.7 specifies the sequence of statistical analysis procedures and the decision criteria for

evaluating each hypothesis. Section 3.8 concludes by summarizing the methodological approach and linking it to the results presented in Chapter 4.

3.2. Operational definitions for research variables

Operationalisation refers to the process by which an abstract construct is translated into a set of observable indicators that can be recorded and analysed. As none of the variables in this study are directly observable, each is inferred from respondents' reported perceptions. The credibility of the empirical analysis depends on whether these indicators accurately represent the intended constructs. This study follows the established paradigm for developing marketing measures, wherein constructs are first specified conceptually, then measured using multiple items from previously validated scales, and subsequently refined through reliability and validity assessments prior to hypothesis testing (Churchill, 1979). The decision to adapt existing scales, rather than develop new ones, preserves comparability with the international literature under examination and ensures that any divergence in the Saudi results can be attributed to contextual factors rather than instrument novelty.

All six focal constructs are treated as reflective latent variables. All six focal constructs are treated as reflective latent variables, indicating that the indicators are considered manifestations of an underlying perception rather than components that collectively define the construct. Each construct is measured using a five-point Likert scale, anchored at 1 = strongly disagree and 5 = strongly agree. The five-point format was selected due to the bilingual administration of the questionnaire and the use of mall-intercept conditions, where minimizing cognitive load is essential. Additionally, five categories align with conventions in Arabic-language survey practice, reducing the likelihood of respondents defaulting to the scale midpoint due to uncertainty. Service quality is conceptualized as a multidimensional construct with five subdimensions, consistent with the SERVQUAL framework reviewed in Chapter 2, while the remaining constructs are unidimensional. Demographic variables are included as moderators and are measured using categorical items rather than Likert scales. All are set out below. Service quality is conceptually defined as the consumer's comparative judgement of the overall excellence or superiority of a retail store's service against expectation, and is operationally defined as the respondent's mean score across items measuring reliability, assurance, tangibles, empathy and responsiveness, adapted from

Parasuraman, Zeithaml and Berry (1988) and from the retail-specific refinement offered by Dabholkar, Thorpe and Rentz (1996). Perceived value is conceptually defined as the consumer's overall assessment of the utility of the retail offering based on what is received relative to what is given, and is operationally defined as the mean score across items adapted from Sweeney and Soutar (2001), covering monetary, functional and emotional value. Customer satisfaction is conceptually defined as the evaluative judgement following consumption in which perceived performance is compared with prior expectation, and is operationally defined using items adapted from Oliver (1980). Brand trust is conceptually defined as the consumer's willingness to rely on the retail brand to perform its stated function, and is operationally defined using items adapted from Chaudhuri and Holbrook (2001). Digital engagement is conceptually defined as the consumer's cognitive, emotional and behavioural investment in interactions with the retail brand through online and social media channels, and is operationally defined using items adapted from Hollebeek, Glynn and Brodie (2014). Brand loyalty, the dependent variable, is conceptually defined following Oliver (1999) as a deeply held commitment to repatronise a preferred retail brand consistently in the future, and is operationally defined as a composite of attitudinal indicators (preference, recommendation intention, resistance to switching) and behavioural indicators (repeat patronage, share of category spend, willingness to pay a premium).

Table 3.1: Operational definitions and measurement of research variables

Variable	Role in model	Conceptual definition	Operationalisation	Source of adapted scale
Service quality	Independent	Consumer judgement of the overall excellence of the store's service relative to expectation	15 items across 5 dimensions (reliability, assurance, tangibles, empathy, responsiveness); 5-point Likert	Parasuraman, Zeithaml & Berry (1988); Dabholkar, Thorpe & Rentz (1996)
Perceived value	Independent	Overall assessment of utility based on what is received relative to what is given	5 items (monetary, functional, emotional value); 5-point Likert	Zeithaml (1988); Sweeney & Soutar (2001)
Customer satisfaction	Independent / mediating	Evaluative judgement comparing perceived performance with prior expectation	4 items (overall, expectation-based, experience-based, comparative); 5-point Likert	Oliver (1980, 1999)
Brand trust	Mediating	Willingness to rely on the retail brand to deliver what it promises	4 items (dependability, honesty,	Morgan & Hunt (1994); Chaudhuri & Holbrook (2001)

Variable	Role in model	Conceptual definition	Operationalisation	Source of adapted scale
			confidence, safety); 5-point Likert	
Digital engagement	Independent	Cognitive, emotional and behavioural investment in the brand's online and social media touchpoints	6 items (cognitive processing, affection, activation); 5-point Likert	Hollebeek, Glynn & Brodie (2014)
Brand loyalty	Dependent	Deeply held commitment to repatronise a preferred retail brand consistently in future	6 items: 3 attitudinal (preference, recommendation, non-switching) and 3 behavioural (repeat purchase, share of spend, price premium)	Oliver (1999); Chaudhuri & Holbrook (2001); Zeithaml, Berry & Parasuraman (1996)
Age, gender, income	Moderating	Demographic characteristics of the respondent	Gender: nominal (2 categories); age: ordinal (5 bands); monthly household income: ordinal (5 bands)	Researcher-constructed, bands aligned to national statistical reporting conventions

Source: prepared by the researcher based on the cited studies.

Two operational decisions in Table 3.1 require explicit justification due to their implications for interpreting the results. First, customer satisfaction serves a dual function: Chapter 1 identifies it as an independent variable, while hypothesis H1 treats it as an outcome of service quality and H2 as a predictor of loyalty. Accordingly, customer satisfaction is more accurately classified as an endogenous mediating variable and will be modeled as such. Second, brand loyalty is operationalized as a composite construct, with attitudinal and behavioral indicators loading onto a single latent factor rather than being treated as separate constructs. This approach aligns with the composite-loyalty tradition established by Dick and Basu (1994) and Chaudhuri and Holbrook (2001).

However, an empirical risk exists: if attitudinal and behavioral items do not converge in the Saudi sample, the measurement model will indicate this through low loadings or unacceptable average variance extracted.

In such a case, the construct will be specified as two dimensions. This contingency is stated in advance to ensure that any respecification is recognized as a documented response to empirical evidence rather than a post hoc adjustment.

3.3. Research design

The research design follows the layered framework of Saunders et al. (2019), which organizes methodological decisions into six sequential layers, from philosophical commitments to procedural techniques. This framework requires each methodological choice to be justified in relation to the preceding layer, ensuring that the selection of a questionnaire and a structural equation model is grounded in a defensible epistemological stance regarding knowledge about consumer loyalty. As the study aims to test hypotheses concerning relationships between measurable perceptions, each layer consistently supports a quantitative approach. The subsequent paragraphs explicitly detail this progression and identify the methodological trade-offs at each stage.

At the outermost layer, the study adopts a positivist research philosophy. Ontologically, consumer perceptions of service quality, value, satisfaction, trust, engagement, and loyalty are treated as phenomena existing independently of the researcher and are represented, within acceptable error, by numerical indicators. Epistemologically, observable and measurable regularities are considered the legitimate basis for knowledge claims, requiring the researcher to remain external to the phenomenon. Positivism is appropriate because the research problem concerns whether specified relationships exist and their strength, rather than the subjective meanings consumers attach to loyalty. However, this approach limits access to deeper insights; an interpretivist design using depth interviews would likely yield richer understanding of how Islamic values, family influence, and social prestige are experienced in shopping decisions. While a pragmatist mixed-methods design could address this limitation, constraints related to time, resources, and the need for a sufficiently large sample to estimate a model with six constructs and three moderators render a mono-method quantitative approach the most feasible option.

At the second layer, the study employs a deductive approach. It draws on established theory reviewed in Chapter 2, formulates six falsifiable hypotheses in Section 1.6, operationalizes the constructs, collects data, and evaluates each hypothesis using statistical evidence. This approach is appropriate because the constructs and their expected relationships are well established in the literature; the study's contribution is to test the applicability of existing theory in an under-researched cultural context. At the third layer, the methodological choice is mono-method quantitative, utilizing a single data collection technique—a structured questionnaire—and a corresponding set of analytical procedures. At the fourth layer, the

strategy is a survey. A survey is preferred over an experiment, as variables such as perceived service quality or cultural trust cannot be ethically or practically manipulated, and over a case study, as the research aims to identify generalizable patterns across the Saudi consumer population rather than focus on a single retail organization.

At the fifth layer, the time horizon is cross-sectional. Data will be collected at a single point in 2026, providing a snapshot of the relationships among the constructs. This approach aligns with the study's objective of characterizing the current loyalty landscape during a period of rapid retail transformation and is the only option compatible with the project timeline.

However, this design imposes a key interpretive limitation, discussed further in Section 5.1: cross-sectional data can demonstrate covariation among variables as predicted by the hypotheses but cannot establish temporal precedence. Therefore, any causal language in the hypotheses should be interpreted as theoretically motivated rather than empirically confirmed. A longitudinal design, tracking the same consumers over multiple waves, would be necessary to observe the formation and decay of loyalty. At the innermost layer, the study employs a bilingual self-administered questionnaire distributed both online and in person, with data analyzed using partial least squares structural equation modeling, supported by preliminary descriptive and screening analyses in SPSS. Further details are provided in Sections 3.5 and 3.7.

Table 3.2: Summary of research design choices by layer

Design layer	Choice adopted	Principal justification and trade-off
Philosophy	Positivism	Constructs are treated as measurable and researcher-independent; forgoes the interpretive depth an interpretivist design would provide
Approach to theory	Deductive	Theory is well established, so the contribution lies in testing transferability rather than generating new constructs
Methodological choice	Mono-method quantitative	Feasible within project constraints; a mixed-methods design would strengthen cultural interpretation
Strategy	Survey	Antecedents cannot be manipulated experimentally; generalisable patterns are sought rather than single-organisation depth
Time horizon	Cross-sectional	Captures the current transformation snapshot; cannot establish temporal precedence or causality
Techniques and procedures	Bilingual structured questionnaire; PLS-SEM analysis	Suits a complex model with mediation and moderation and does not assume multivariate normality

Source: prepared by the researcher, structured according to Saunders et al. (2019).

3.4. Research population & Sampling methods and technique

The target population for this study comprises adult consumers resident in Saudi Arabia who are aged eighteen years or above and who have made at least one purchase at a physical or online retail store in the Kingdom within the six months preceding data collection. The unit of analysis is the individual consumer, and the referent for every perceptual item is the retail brand at which the respondent shops most frequently, which the respondent names at the beginning of the questionnaire. Anchoring all items to a single self-nominated brand is methodologically important: without it, respondents would evaluate service quality with one retailer in mind and loyalty with another, and the resulting correlations would be uninterpretable. The recency screen serves a comparable function, ensuring that judgements are grounded in retrievable experience rather than in general impressions of the retail sector. Consistent with the scope defined in Section 1.9, the population is limited to consumers of supermarkets, hypermarkets, specialty stores, fashion retailers and electronics retailers, and excludes consumers of banking, hospitality and telecommunications services.

No accessible sampling frame exists that enumerates this population. There is no publicly available register of Saudi retail consumers, and although national statistical authorities publish population estimates by region, age and nationality, these cannot be converted into a list from which individuals could be randomly selected. Probability sampling is therefore not attainable, and the study uses non-probability sampling. Specifically, it combines convenience sampling with quota controls: respondents are recruited through accessible online and in-store channels, but recruitment is monitored against predetermined quotas for city of residence and gender so that no single subgroup dominates the sample. Quota controls are also imposed on retail format so that grocery-oriented and non-grocery formats are both adequately represented. This design does not permit calculation of sampling error and does not license strict statistical generalisation to the national population; that limitation is acknowledged rather than minimised, and it is revisited in Section 5.1. What the quota structure does achieve is protection against the most damaging forms of composition bias, particularly the over-representation of young, urban, digitally active respondents that purely social-media-based recruitment tends to produce and which would systematically inflate the apparent influence of digital engagement on loyalty.

The target sample size is 400 usable responses, a figure supported by three independent lines of reasoning. First, for a population of the order of several million, standard sample size determination tables indicate that approximately 384 responses are required to estimate a proportion within a five per cent margin of error at the ninety-five per cent confidence level, so 400 satisfies that conventional threshold (Krejcie & Morgan, 1970). Second, the instrument contains approximately forty reflective indicators, and the widely applied rule of at least ten observations per indicator implies a minimum of four hundred cases for stable parameter estimation in structural equation modelling (Hair et al., 2010). Third, and most binding, hypothesis H6 requires multigroup comparison across demographic categories; because meaningful group comparison requires roughly one hundred cases per group, a total of four hundred is the minimum that permits the sample to be partitioned into four demographic cells without rendering the subgroup estimates unstable. To protect against attrition from incomplete responses, straight-line answering and failed screening questions, distribution will target approximately 480 initiated responses on the expectation that eighty to eighty-five per cent will prove usable.

Table 3.3: Planned quota allocation of the target sample

City	Target n	Share of sample	Rationale for allocation
Riyadh	120	30%	Largest metropolitan retail market and the widest range of retail formats in the Kingdom
Jeddah	100	25%	Second commercial centre; distinct Hijazi consumer culture and strong international brand presence
Dammam / Eastern Province	80	20%	Industrial and higher-income consumer base with mature mall and hypermarket provision
Makkah	60	15%	Distinctive retail dynamic shaped by pilgrimage-related demand alongside resident consumption
Madinah	40	10%	Smaller retail market included to avoid an exclusively large-city sample
Total	400	100%	Gender quota applied within each city at approximately 50:50; retail format quota applied across the full sample

Source: prepared by the researcher. Allocation is proportionate to the relative concentration of retail activity in the five cities specified in Section 1.9, not to resident population.

3.5. Data collection method

Primary data are collected using a structured, self-administered questionnaire organized into five sequential sections. Section one includes screening questions to establish residence, age, and retail purchase recency, and requests that respondents identify the retail brand they shop at most frequently; this brand serves as the referent for all subsequent perceptual items. Section two measures perceptions of the independent variables, with the service quality, perceived value, and digital engagement blocks presented in rotated order to mitigate systematic order effects. Section three assesses customer satisfaction and brand trust. Section four measures attitudinal and behavioral indicators of brand loyalty. Section five gathers demographic variables necessary for moderation analysis. Demographic questions are intentionally placed last to minimize stereotype-consistent responding on substantive items, although this placement may slightly increase item non-response on the final page. The complete instrument is provided in Appendix A (English) and Appendix B (Arabic).

Given that the target population is predominantly Arabic-speaking and the source scales are published in English, translation quality represents a substantive threat to validity. To address this, a back-translation protocol is implemented (Brislin, 1970). Initially, a bilingual translator translates the English instrument into Arabic. Subsequently, a second bilingual translator, working independently and without access to the original, translates the Arabic version back into English. The two English versions are compared item by item, and any discrepancies are resolved by revising the Arabic wording. Special attention is devoted to items where English marketing terminology lacks an exact Arabic equivalent, such as the distinction between "trust" and "confidence" in the brand trust block, as well as to items whose literal translation may be socially insensitive. The reconciled instrument is then reviewed by two academic staff members in marketing to assess content validity, ensuring that each item is comprehensible, unambiguous, and relevant to its intended construct. Following expert review, a pilot study is conducted with thirty consumers from the target population, who are excluded from the main sample. The pilot study verifies item comprehension, estimates completion time, and calculates Cronbach's alpha coefficients. Any construct with an alpha below the 0.70 threshold at the pilot stage prompts item rewording or replacement prior to launching the main survey.

The reconciled instrument is distributed through two parallel channels to maximize reach and minimize channel-specific bias. Online, the questionnaire is hosted on a web survey platform

and disseminated via social media, messaging applications, and consumer panels. While the online channel efficiently accesses respondents across all five cities, it tends to favor digitally active individuals. To address this limitation, in-person administration is also employed, utilizing mall-intercept and in-store approaches at selected retail locations in the five cities. Fieldwork is intentionally scheduled across weekdays and weekends, as well as morning and evening trading periods, to ensure the sample is not limited to consumers with specific shopping patterns. Respondents approached in person may complete the questionnaire on a tablet or on paper, according to their preference. Ethical safeguards are consistently applied across both channels: participation is voluntary and may be discontinued at any time without consequence; a cover statement communicates the academic purpose of the research, the estimated ten-minute completion time, and the assurance that no personally identifying information is collected. No financial incentives are provided, as such inducements may encourage satisficing behavior. Completed responses are stored securely in password-protected form and reported only in aggregate. Since all variables are measured from a single respondent at a single point in time, common method variance is a recognized risk. Accordingly, procedural remedies recommended by Podsakoff et al. (2003) are incorporated: predictor and criterion items are separated by intervening blocks, anonymity is explicitly assured to reduce socially desirable responding, scale anchors are varied between sections, and a small number of reverse-coded items are included to detect inattentive responding. A statistical test for method bias is detailed in Section 3.7.

3.6. Types of data

This study utilizes both primary and secondary data, with a distinct separation of their respective roles. Primary data consist of questionnaire responses from four hundred targeted consumers and serve as the exclusive evidentiary basis for testing the six hypotheses. The collection of primary data is essential, as the constructs of interest are subjective perceptions not captured by existing datasets. While national statistical agencies report retail sales values, outlet counts, and household expenditure, they do not provide data linking perceived service quality to brand trust at the individual consumer level. Secondary data include peer-reviewed journal articles, scholarly books, industry reports, and official publications reviewed in Chapters 1 and 2, such as national policy documents like the Vision 2030 programme and official statistical releases on population and household spending. Secondary sources fulfill

three functions: they justify the research problem, supply the theoretical constructs and validated measurement items adapted by the primary instrument, and provide external benchmarks for assessing the achieved sample profile in Section 4.2. Secondary sources are not used to test hypotheses, and no hypothesis in this study is evaluated solely on the basis of secondary evidence.

The primary data encompass three levels of measurement, which determine the appropriate statistical operations. Nominal data represent unordered categories, such as gender, city of residence, retail format, and preferred shopping channel. These variables are analyzed using frequencies, modes, and cross-tabulation, and serve as grouping variables for multigroup comparisons. Ordinal data consist of ranked categories with unequal or unknown intervals, including age bands, monthly household income bands, education categories, and reported shopping frequency. These variables are analyzed using medians and rank-based procedures, and when used as moderators, they are treated categorically rather than as continuous variables. Interval-treated data refer to the Likert-scored perceptual items. Although a five-point agreement scale technically yields ordinal responses, as the psychological distance between response options is not guaranteed to be equal, the established convention in consumer research is to treat multi-item Likert composites as approximately interval. This approach permits the calculation of means, standard deviations, correlations, and the estimation of a structural equation model. This convention is explicitly stated, as it underpins the validity of the parametric analyses conducted.

Two additional characteristics of the data influence the analytical strategy. First, the data are cross-sectional and self-reported, resulting in all constructs sharing a common measurement source. This introduces the potential for common method variance, which is addressed procedurally in Section 3.5 and statistically in Section 3.7. Second, the perceptual data are unlikely to exhibit multivariate normality. Consumer evaluations of retailers they have chosen to patronize often cluster toward the agreement end of the scale, resulting in negative skew. This pattern is anticipated across the satisfaction, trust, and loyalty constructs. Instead of transforming the data or discarding cases to achieve normality, the study employs an estimation method that does not require distributional normality. This consideration is a principal reason for selecting partial least squares over covariance-based structural equation modeling, as discussed in Section 3.7.

Table 3.4: Types of data and levels of measurement in the study

Data type	Level of measurement	Variables concerned	Permissible treatment in this study
Primary	Nominal	Gender, city, retail format, preferred channel	Frequencies, percentages, mode, cross-tabulation; grouping variable for multigroup analysis
Primary	Ordinal	Age band, income band, education, shopping frequency	Frequencies, median, rank-based tests; treated categorically when used as moderators
Primary	Interval (by convention)	All Likert-scored items for the six focal constructs	Mean, standard deviation, correlation, reliability coefficients, structural equation modelling
Secondary	Documentary and aggregate statistical	Prior empirical findings, published scales, national statistics and policy documents	Problem justification, scale sourcing, benchmarking of the achieved sample; not used for hypothesis testing

3.7. Statistical analysis procedures

The analysis follows five sequential stages, with the sequence being critical: structural relationships are not interpreted until the underlying measurement model is demonstrated to be reliable and valid. The first stage involves data screening and preparation using SPSS. Responses that fail any screening question are excluded. Cases with substantial missing data are deleted, while isolated missing values on Likert items are replaced with the item mean when they affect fewer than 5% of responses for that item. Straight-line responses, defined as identical answers across an entire block including reverse-coded items, are removed as evidence of inattentive completion. Univariate outliers are identified by standardised scores exceeding an absolute value of 3.29, and multivariate outliers are detected using Mahalanobis distances at $p < 0.001$; retention or deletion is determined on a case-by-case basis. Skewness and kurtosis are calculated for each item to document distributional shape, with thresholds of approximately ± 2 for skewness and ± 7 for kurtosis indicating severe departures. Finally, early and late respondents are compared on substantive constructs to assess non-response bias (Armstrong & Overton, 1977).

The second stage generates the descriptive results presented in Sections 4.2 and 4.4. Frequencies and percentages are used to describe the demographic profile of the sample and to compare it against the quota targets in Table 3.3. Means, standard deviations, and rank

orders are calculated for each construct and individual item. Construct means are interpreted using a three-band convention for a five-point scale: values from 1.00 to 2.33 indicate a low level, 2.34 to 3.67 a moderate level, and 3.68 to 5.00 a high level. This stage addresses the descriptive components of the research questions in Section 1.4 by establishing the extent to which Saudi consumers perceive each antecedent and report their loyalty, independent of relationships among constructs.

The third stage evaluates the measurement model using SmartPLS, adhering to the reporting standards established by Hair et al. (2019). Indicator reliability is assessed via outer loadings, with a preferred threshold of 0.708. Indicators with loadings between 0.40 and 0.708 are retained only if their removal does not increase composite reliability or average variance extracted above the required level; indicators below 0.40 are deleted. Internal consistency reliability is measured using Cronbach's alpha and composite reliability, both of which must exceed 0.70. Composite reliability above 0.95 is flagged as evidence of redundant items rather than superior reliability (Cronbach, 1951; Nunnally, 1978). Convergent validity requires an average variance extracted of at least 0.50 for each construct, indicating that the construct explains the majority of variance in its indicators (Fornell & Larcker, 1981). Discriminant validity is assessed using three complementary criteria: the Fornell–Larcker criterion, where the square root of each construct's average variance extracted must exceed its correlations with other constructs; examination of cross-loadings; and the heterotrait–monotrait ratio of correlations, which must fall below 0.90, with a stricter threshold of 0.85 for conceptually adjacent constructs such as satisfaction and trust (Henseler et al., 2015). Common method bias is evaluated through a full collinearity test, requiring variance inflation factors from a model regressing all constructs on a marker variable to remain below 3.3 (Kock, 2015).

Stage four evaluates the structural model and tests H1, H2, H4 and H5. The fourth stage assesses the structural model and tests hypotheses H1, H2, H4, and H5. Lateral collinearity among predictors is examined first, with variance inflation factors required to remain below 5, and preferably below 3. Path coefficients are estimated, and their significance is determined through bootstrapping with 5,000 subsamples, using bias-corrected and accelerated confidence intervals. A hypothesis is supported when the coefficient has the predicted sign and its t-value is at least 1.96, corresponding to $p < 0.05$ on a two-tailed test. Explanatory power is reported as the coefficient of determination for each endogenous construct, interpreted using benchmarks of 0.25, 0.50, and 0.75 for weak, moderate, and

substantial explanation in consumer behaviour research. The practical importance of each predictor is reported as the effect size f^2 , interpreted using thresholds of 0.02, 0.15, and 0.35 for small, medium, and large effects, respectively. Predictive relevance is evaluated using the cross-validated Q^2 statistic, with values above zero indicating predictive relevance beyond the estimation sample. The mediation of the satisfaction–loyalty relationship is tested using the bootstrapped indirect effect procedure, in which the product of the satisfaction-to-trust and trust-to-loyalty paths is computed across 5,000 subsamples, and mediation is concluded when the bias-corrected 95 per cent confidence interval for that indirect effect excludes zero (Preacher & Hayes, 2008). The nature of the mediation is then classified: if the direct effect of satisfaction on loyalty remains significant alongside a significant indirect effect, the mediation is partial and complementary; if the direct effect becomes non-significant, the mediation is full. This bootstrapping approach is preferred to the older causal-steps logic of Baron and Kenny (1986) because it does not assume a normal sampling distribution for the indirect effect and reports greater statistical power. H6, which posits demographic moderation, is tested by multigroup analysis. Measurement invariance across groups is established first through the measurement invariance of composite models procedure, since comparing path coefficients across groups with differing measurement properties would confound moderation with measurement artefact (Henseler et al., 2016). Where invariance holds, group-specific path coefficients are compared using the parametric approach and the permutation test, and moderation is concluded when the difference between coefficients is significant at $p < 0.05$ (Sarstedt et al., 2011). Analysis is conducted in SPSS version 29 for screening and description and in SmartPLS 4 for the measurement, structural, mediation and moderation stages. Partial least squares is chosen in preference to covariance-based estimation for three reasons: the model is complex, containing six latent constructs, a mediating chain and a moderation test; the data are not expected to be multivariate normal; and the study’s orientation is towards explaining and predicting a target construct rather than confirming the global fit of a theoretical structure (Hair et al., 2019).

Table 3.5: Mapping of hypotheses to analytical techniques and decision criteria

Hyp.	Relationship tested	Technique	Decision criterion
H1	Service quality → customer satisfaction	PLS-SEM path coefficient; bootstrapping (5,000 subsamples)	Positive β , $t \geq 1.96$, $p < 0.05$, CI excludes zero

Hyp.	Relationship tested	Technique	Decision criterion
H2	Customer satisfaction → brand loyalty	PLS-SEM path coefficient; bootstrapping	Positive β , $t \geq 1.96$, $p < 0.05$, CI excludes zero
H3	Satisfaction → trust → loyalty (mediation)	Bootstrapped indirect effect (Preacher & Hayes, 2008)	Bias-corrected 95% CI for indirect effect excludes zero
H4	Perceived value → brand loyalty	PLS-SEM path coefficient; bootstrapping	Positive β , $t \geq 1.96$, $p < 0.05$, CI excludes zero
H5	Digital engagement → brand loyalty	PLS-SEM path coefficient; bootstrapping	Positive β , $t \geq 1.96$, $p < 0.05$, CI excludes zero
H6	Age, gender, income moderate the antecedent–loyalty paths	MICOM invariance test, then PLS-MGA and permutation test	Significant between-group difference in β at $p < 0.05$ after invariance is established

3.8. Conclusion

This chapter has established the methodological framework through which the conceptual model introduced in Chapter 1 and substantiated in Chapter 2 will undergo empirical testing. The six constructs have been operationally defined and mapped to validated source scales, ensuring that each abstract perception is represented by multiple observable indicators whose adequacy can be systematically evaluated. The research design has been justified at each layer as positivist in philosophy, deductive in approach, mono-method quantitative in method, survey-based in strategy, and cross-sectional in temporal scope, with the trade-offs of each layer explicitly acknowledged. The target population has been specified, and a quota-controlled non-probability sample of four hundred consumers across five cities has been justified on three independent grounds. The data collection instrument, its bilingual translation protocol, pilot validation, and the ethical and procedural safeguards governing its administration have been described in sufficient detail to enable replication.

Three features of the research design warrant particular emphasis due to their direct implications for interpreting the findings. First, the sequencing of the analysis is substantive: measurement quality is assessed prior to examining structural relationships, ensuring that any path coefficient reported in Chapter 4 is based on constructs that have been demonstrated to be both reliable and empirically distinct. Second, the decision criteria for each hypothesis have been pre-specified in Table 3.5, thereby eliminating the possibility of selecting favorable thresholds post hoc and ensuring that unsupported hypotheses are recognized as substantive findings rather than study failures. Third, the limitations associated with non-probability sampling, cross-sectional timing, and single-source self-report are inherent structural features

of the design, not oversights, and they delineate the boundaries within which the study's conclusions are valid.

This chapter, therefore, presents a methodological plan that is internally coherent and appropriately aligned with the research questions. Chapter 4 implements this plan by reporting the achieved sample profile relative to the quota targets specified herein, presenting evidence of reliability and validity for the measurement instrument, describing the levels of each construct as perceived by Saudi retail consumers, and reporting the results of the hypothesis tests alongside a discussion that contextualizes each finding within the literature reviewed in Chapter 2. Chapter 5 subsequently addresses the limitations that shape these results, offers recommendations for practitioners in the Saudi retail sector, and provides the overall conclusion of the study.

Survey Link : <https://form.jotform.com/262186092919062>

4. Chapter 4: Statistical Analysis & Findings discussion

4.1. Introduction

This chapter presents the empirical results of the study and interprets their significance within the theoretical framework outlined in Chapter 2. The analytical sequence established in Chapter 3, which prioritizes verification of measurement quality before interpretation of structural relationships, determines the structure of this chapter. Section 4.2 details the achieved sample and compares it to the quota targets specified in Table 3.3, as the sample composition influences the reliability of subsequent findings. Section 4.3 provides evidence of the reliability and validity of the measurement instrument, assessing whether the adapted scales perform adequately in the Saudi context. Section 4.4 presents descriptive statistics for each construct, thereby addressing the descriptive aspects of the research questions. Section 4.5 outlines the results of the six hypothesis tests and evaluates each outcome in relation to previous research. Section 4.6 offers concluding remarks.

A key interpretive consideration applies throughout this chapter and is articulated at the outset. Since the study utilizes scales originally developed and validated in Western consumer markets, the findings in Section 4.3 possess theoretical significance beyond serving as a technical prerequisite. Should the adapted scales demonstrate acceptable reliability and discriminant validity within the Saudi sample, this would provide evidence supporting the cross-cultural robustness of the underlying constructs and directly address the second component of the research problem identified in Section 1.2. Conversely, if constructs that are empirically distinct in Western contexts do not separate in the Saudi sample, the measurement outcomes themselves constitute a substantive insight into the organization of loyalty perceptions in this market. Accordingly, the reporting in Section 4.3 is structured to reflect this dual significance, rather than functioning merely as a procedural step preceding the main results.

A further interpretive consideration addresses the distinction between statistical significance and practical relevance. Given a sample size of four hundred, even modest path coefficients

may achieve conventional significance thresholds, potentially resulting in statistical support for hypotheses that do not reflect relationships strong enough to inform managerial decisions. Therefore, this chapter reports effect sizes in addition to significance for each structural path, and the discussion in Section 4.5 differentiates between hypotheses that are merely supported and those that are sufficiently robust to justify the practical recommendations presented in Section 5.2. The same approach is applied to descriptive results: while a high construct mean suggests favorable consumer perception, the corresponding standard deviation reveals whether this perception is broadly shared or masks substantial subgroup differences.

4.2. Sample profile (descriptive statistics for research sample)

This section reports the demographic and behavioural composition of the achieved sample. Three quantities are reported before the profile itself. The first is the number of responses initiated, which Section 3.4 anticipated at approximately 480. The second is the number of responses removed and the reason for each removal, disaggregated into failures of the screening questions, incomplete submissions, straight-line responding detected through the reverse-coded items, and cases deleted as multivariate outliers. The third is the number of usable responses, against the target of 400, together with the usable response rate calculated as usable responses divided by responses initiated. Reporting the attrition transparently matters because a high deletion rate concentrated in one recruitment channel would indicate a channel-specific data quality problem rather than random attrition, and would need to be disclosed as a qualification on the results.

Table 4.1: Demographic and behavioural profile of the achieved sample

Characteristic	Category	Frequency	Per cent	Quota target (%)
Gender	Male			50
	Female			50
Age	18–24			—
	25–34			—
	35–44			—
	45–54			—
	55 and above			—
City of residence	Riyadh			30
	Jeddah			25
	Dammam / Eastern Province			20
	Makkah			15
	Madinah			10
Education	Secondary or below			—
	Diploma			—
	Bachelor's degree			—
	Postgraduate			—
Monthly household income (SAR)	Below 5,000			—
	5,000–9,999			—
	10,000–14,999			—
	15,000–24,999			—
	25,000 and above			—
Primary retail format	Supermarket / hypermarket			—
	Fashion retailer			—
	Electronics retailer			—
	Specialty store			—
Shopping frequency	Weekly or more often			—

Characteristic	Category	Frequency	Per cent	Quota target (%)
	2–3 times per month			—
	Monthly or less often			—
Preferred channel	Predominantly in-store			—
	Predominantly online			—
	Both approximately equally			—

The sample profile is evaluated along three dimensions. First, quota attainment is measured by comparing the achieved city and gender distributions with the targets specified in the final column. Minor deviations of a few percentage points are expected in quota-controlled non-probability sampling. However, substantial shortfalls in the smaller-city quotas would effectively render the sample a three-city Riyadh, Jeddah, and Dammam sample, necessitating qualification of the scope claims presented in Section 1.9. Second, external comparability is evaluated by comparing the achieved age, gender, and income distributions with published national statistics for the adult Saudi population. This comparison serves a diagnostic rather than corrective function. Because the sample is non-probability, close alignment with national parameters does not justify statistical generalisation. However, pronounced divergence highlights which subgroups are over- or under-represented and indicates which findings require the greatest caution.

Third, and most consequentially for the subsequent analysis, the sample profile determines whether the multigroup comparison required by H6 is feasible as specified. Section 3.4 justified the sample size on the requirement that meaningful subgroup comparisons necessitate 100 cases per group. If the achieved distribution results in a near-even gender split and age bands that can be combined into two balanced groups, the multigroup analysis proceeds as planned. However, if any comparison group contains substantially fewer than one hundred cases, the corresponding moderation test loses statistical power, rendering a non-significant between-group difference uninformative because it cannot be distinguished from insufficient power to detect a true difference. In such cases, the affected demographic bands are combined into broader categories prior to analysis, and this adjustment is reported explicitly. The reduced granularity is acknowledged as a limitation in Section 5.1.

The behavioural characteristics presented in the lower portion of Table 4.1 serve an analytical purpose beyond mere description. The distribution of primary retail format and preferred channel establishes the empirical context for interpreting the digital engagement results in

Section 4.5. If the sample is heavily weighted toward respondents who primarily shop online, a strong digital engagement effect on loyalty may reflect sample composition as much as population-level dynamics, and this qualification should be reported. Conversely, if the sample is predominantly in-store and digital engagement remains a significant predictor of loyalty, the finding is more robust than the raw coefficient suggests, as it holds among consumers whose primary relationship with the retailer is physical rather than digital.

4.3. Validity & reliability tests for the questionnaires

The measurement model is checked before interpreting any structural relationships, following the steps and thresholds described in Section 3.7. Four properties are reviewed in order: indicator reliability using outer loadings, internal consistency reliability using Cronbach's alpha and composite reliability, convergent validity using average variance extracted, and discriminant validity using the Fornell–Larcker criterion and the heterotrait–monotrait ratio. Table 4.2 shows the first three properties for each construct. Any indicator loading below 0.708 is noted, and the decision to keep or remove it is explained. Indicators can be kept if their loading is above 0.40 and removing them would not improve composite reliability or average variance extracted beyond the threshold. Indicators with loadings below 0.40 are always removed. All deletions are reported, since an instrument that only meets reliability standards after removing several items is not as strong as one that meets them without changes. Readers have the right to know which items were removed.

Table 4.2: Indicator loadings, internal consistency and convergent validity

Construct	No. of items retained	Loading range	Cronbach's α	CR	AVE
Service quality					
Perceived value					
Customer satisfaction					
Brand trust					
Digital engagement					
Brand loyalty					
Threshold	—	≥ 0.708	≥ 0.70	0.70–0.95	≥ 0.50

Source: primary data, SmartPLS 4 output. Thresholds per Hair et al. (2019), Fornell and Larcker (1981) and Nunnally (1978). CR = composite reliability; AVE = average variance extracted.

Discriminant validity is assessed using the two criteria presented in Tables 4.3 and 4.4. The Fornell–Larcker criterion requires that the square root of each construct’s average variance extracted, shown on the diagonal, exceeds that construct’s correlation with all other constructs, which are reported off the diagonal. The heterotrait–monotrait ratio is considered a more sensitive test and is regarded as decisive when the two criteria yield conflicting results, particularly when constructs are conceptually adjacent. A general threshold of 0.90 is applied, with a stricter threshold of 0.85 for the satisfaction–trust and satisfaction–loyalty pairs, as these constructs are theoretically similar and a ratio above 0.85 would suggest that respondents do not distinguish between them. The satisfaction–trust pair warrants special attention because H3 posits a mediating relationship between these constructs. If empirical distinction between the two cannot be established, the mediation hypothesis cannot be tested as specified and this limitation must be reported accordingly.

Table 4.3: Discriminant validity — Fornell–Larcker criterion

Construct	SQ	PV	CS	BT	DE	BL
Service quality (SQ)						
Perceived value (PV)						
Customer satisfaction (CS)						
Brand trust (BT)						
Digital engagement (DE)						
Brand loyalty (BL)						

Source: primary data, SmartPLS 4 output. Diagonal values are the square root of AVE; off-diagonal values are construct correlations. The diagonal value must exceed all values in its row and column.

Table 4.4: Discriminant validity — heterotrait–monotrait (HTMT) ratio

Construct	SQ	PV	CS	BT	DE	BL
Service quality (SQ)	—					
Perceived value (PV)		—				
Customer satisfaction (CS)			—			
Brand trust (BT)				—		
Digital engagement (DE)					—	
Brand loyalty (BL)						—

Source: primary data, SmartPLS 4 output. Threshold < 0.90 in general; < 0.85 applied to the conceptually adjacent CS–BT and CS–BL pairs, per Henseler et al. (2015).

Measurement results are evaluated for both theoretical significance and technical adequacy, as outlined in Section 4.1. Three distinct outcomes are possible, each with specific implications. If all constructs satisfy every threshold with all items retained, this outcome supports the cross-cultural applicability of the measurement instruments and facilitates comparison with international literature. If thresholds are achieved only after item deletion, the pattern of these deletions provides valuable insight. For instance, if deletions are concentrated within a single SERVQUAL dimension, this may indicate that the dimension lacks relevance in the Saudi retail context, thereby informing the discussion on the dimensional stability of SERVQUAL across cultures. If discriminant validity is not established between satisfaction and trust or between satisfaction and loyalty, this suggests that Saudi consumers in this sample do not differentiate these constructs as distinctly as Western-based models propose. In such cases, the structural model will be revised, and all modifications will be documented transparently.

Two additional validation checks are also reported in this section. The first is the full collinearity test for common method bias, as detailed in Section 3.7. Variance inflation factors derived from regressing all constructs on a common factor must remain below 3.3. Values exceeding this threshold indicate that a substantial portion of the observed covariance may be attributable to the measurement method rather than to theoretical relationships, and such findings must be reported with this caveat (Kock, 2015). The second check involves comparing early and late respondents on the six construct means to evaluate non-response bias. Non-significant differences suggest that respondents who required follow-up do not differ systematically from those who responded promptly (Armstrong & Overton, 1977). Both validation checks are reported irrespective of their outcomes to maintain transparency and reliability.

4.4. Descriptive statistics for research variables

This section presents the levels at which Saudi retail consumers perceive each construct, independent of the interrelationships among constructs. This distinction is important because these two types of evidence address different research questions and may lead to different practical implications. For example, a construct may serve as a strong statistical predictor of loyalty while being perceived at only a moderate level, identifying it as a potential area for improvement. Conversely, a construct may be perceived very favourably yet exert minimal

influence on loyalty, suggesting possible over-investment. Construct means are evaluated using the three-band convention outlined in Section 3.7, where values from 1.00 to 2.33 indicate a low level, 2.34 to 3.67 a moderate level, and 3.68 to 5.00 a high level on the five-point scale. Standard deviations are reported alongside the means and are interpreted as substantive indicators. A standard deviation approaching or exceeding 1.00 on a five-point scale suggests that the mean reflects considerable variation among respondents. In such cases, the mean alone may be misleading, and the multigroup results in Section 4.5 provide more informative evidence.

Table 4.5: Descriptive statistics for the research variables

Construct	Mean	SD	Min	Max	Rank	Level
Service quality						
Perceived value						
Customer satisfaction						
Brand trust						
Digital engagement						
Brand loyalty						

Source: primary data, SPSS 29 output. Interpretation bands: 1.00–2.33 = low; 2.34–3.67 = moderate; 3.68–5.00 = high

The descriptive evidence addresses the research questions outlined in Section 1.4 as follows. The primary research question concerns the key factors influencing brand loyalty in Saudi retail stores. Descriptive results provide the initial insight by establishing the relative importance of the five candidate antecedents based on consumers' evaluations. Structural results in Section 4.5 then determine which antecedents significantly influence loyalty. Subsidiary questions one and two, which focus on customer satisfaction and perceived service quality, are addressed through construct means and item-level means within the service quality block. These results clarify whether consumers evaluate tangible aspects of the store environment differently from interpersonal aspects such as staff empathy and responsiveness. Subsidiary question four, regarding digital engagement, is addressed by interpreting the digital engagement mean in conjunction with the channel distribution reported in Table 4.1, as the level of digital engagement is meaningful only when considered alongside the proportion of the sample that transacts online.

Item-level descriptive statistics are presented for each construct, in addition to the construct-level summary in Table 4.5, to avoid obscuring important variation among

individual items. Within the service quality block, the five SERVQUAL dimensions are reported and ranked separately. For example, if tangibles and assurance receive higher scores than empathy and responsiveness, this would suggest that Saudi retailers have prioritized investments in store environment and product availability over staff interaction. Such findings have direct implications for the recommendations in Section 5.2. In the brand loyalty block, the three attitudinal and three behavioural items are reported separately and their means compared. A discrepancy where attitudinal loyalty is high but behavioural loyalty is lower indicates latent loyalty, as described in Chapter 2 following Dick and Basu (1994): consumers express positive attitudes but are limited in repeat patronage by factors such as price, convenience, or accessibility. Conversely, if behavioural indicators exceed attitudinal ones, this suggests spurious loyalty driven by habit or limited alternatives, which would alter the interpretation of an apparently loyal customer base.

The willingness-to-pay-a-premium item within the behavioural loyalty block is examined individually, as anticipated in Section 2.3. Willingness to pay a premium represents the most stringent behavioural indicator of loyalty, as it requires consumers to forgo tangible monetary benefits to remain with a preferred brand. Chapter 2 emphasized that interpretation of this indicator must consider respondents' economic circumstances. If this item yields a substantially lower mean than other behavioural indicators, particularly if it is the lowest-scoring item, this should be interpreted as evidence of pronounced price sensitivity coexisting with loyalty, rather than a lack of loyalty among Saudi consumers. This interpretation can be tested by cross-tabulating the item with the income bands reported in Table 4.1. The outcome directly informs whether loyalty in Saudi retail can be monetised through premium pricing or should instead be monetised through retention and increased share of wallet.

4.5. Testing Hypotheses & Findings Discussion

4.6. Conclusion

5. Chapter five: Research limitations, Recommendations & Conclusion

5.1. Research Limitations & Future Research

5.2. Research Recommendations

5.3. Conclusion

References

Appendix

