One Small Step for Customers, One Giant Leap for Food Systems: Analyzing Purchasing Trends for Increasing Sustainability and Reducing Food Waste

by
Camilla Carpinelli

Dissertation submitted to the School of Computer Science
at Reykjavík University in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

March 3rd, 2025

Supervisors:	María Óskarsdóttir, Associate Professor, Reykjavík University, Iceland
	Anna Sigríður Islind, Associate Professor, Reykjavík University, Iceland
Examining Board:	Ewa Ryszarda Lazarczyk Carlson, Associate Professor, Reykjavík University, Iceland
	Livia Norström, Assistant Professor, University West, Sweden
	Seppe vanden Broucke, Assistant Professor, Ghent University, Belgium
Examiner:	Luis Aburto Lafourcade, Assistant Professor, Universidad Adolfo Ibañez, Chile

ISBN 978-9935-539-66-3 Print version
ISBN 978-9935-539-67-0 Electronic version
ORCID 0000-0003-0329-2689

One Small Step for Customers, One Giant Leap for Food Systems: Analyzing Purchasing Trends for Increasing Sustainability and Reducing Food Waste

Camilla Carpinelli

March 3rd, 2025

Abstract

Food waste is one of the most pressing challenges within the food system, contributing significantly to environmental degradation, resource depletion, and climate change. It has been shown that a large amount of food is wasted at the consumption stage. Therefore, addressing this issue requires an investigative approach that examines customer behavior and the effectiveness of demand forecasting in reducing waste. This thesis focuses on understanding the factors driving food purchasing and explores strategies to mitigate food waste impact through data-driven insights and sustainable practices.

The research applies a data-driven approach investigating sustainable grocery sales, with a particular focus on purchasing habits, decision-making drivers, and the role of demand forecasting in minimizing overproduction and consequent waste. Leveraging comprehensive data from a major supermarket chain covering nearly the entire Icelandic market, this study offers a rare opportunity to explore waste-reduction practices at a national scale, providing insights that can inform broader global efforts.

The thesis is structured around four different studies, adopting a pragmatism research paradigm. It analyzes the drivers of sustainable purchasing behavior, emphasizing the roles of digital media, social influences, and nudging techniques in shaping customer decisions. Special attention is given to discovering the gap between customers' purchasing intentions and actual purchasing behaviors, revealing opportunities for educational and digital interventions to bridge this divide. Furthermore, one of the studies presented in the thesis highlights the potential of advanced demand forecasting models to align production with actual customer demand, reducing food waste while enhancing supply chain efficiency.

The findings underscore the shared responsibility of customers, retailers, and policymakers in addressing food waste. While individual customer choices have significant ripple effects, the study emphasizes systemic changes, such as promoting digital literacy to combat misinformation, launching targeted campaigns to encourage sustainable purchasing, and incentivizing the adoption of advanced forecasting models like subsidies.

This thesis advances sustainability by connecting customer behavior research with operational strategies to address food waste. Its key contribution is providing actionable insights to promote sustainable consumption while emphasizing demand forecasting as an essential tool for reducing waste. By leveraging real-world data, the work provides a foundation for scaling these strategies to diverse contexts.

Keywords: Sustainability; Purchasing behavior; Data Science; Food waste

Acknowledgements

Those who have known me for a long time are aware that these are acknowledgments I never thought I would write, given my long-standing aversion to pursuing a Ph.D. Yet, if there is one thing I've learned since 2021, it's that nothing ever goes as expected. In a single moment, life can change completely, and that's exactly what happened after a "short five-day trip to Iceland." It wasn't easy for me to break free from the patterns my mind had set, nor was it simple to embrace the decision to change my life, country, career perspectives, and leave behind friends, family, and my beloved routine. However, against all odds, I have never regretted it—not for a single day. Not once have I looked back wishing for anything other than the life I was living, and I owe this to the people who supported me, even from afar, throughout this journey.

First and foremost, I want to thank my supervisors. I genuinely believe I couldn't have been luckier. Thank you, María, for always being a guide and a source of support for me. Thank you for reading and rereading my papers, indulging my need for reassurance even when the changes involved something as minor as a comma. You have been an incredible mentor, and I wish you all the best for the new chapter awaiting you. Thank you, Anna Sigga, for always having excellent ideas for me, suggesting the most interesting conferences, and for sharing Italian aperitifs with me.

Thank you to my Committee Members: Livia, Ewa, Seppe, and my examiner Luis. Your feedback and support was precious for me.

Thank you to my family: Mom, Dad, and Michela. Thank you, Mom and Dad, for the trust you have always placed in me, supporting every life choice I made, even when it seemed the most reckless. Thank you for enduring the pain of having me so far away, seeing me only every six months or less, without ever complaining. Thank you, Michela, for always believing in my abilities and encouraging me to live as many adventures as possible.

Thank you, Riki, for the beautiful life you've given me over these years and for the decisions your presence led me to make, which always resulted in something wonderful. Thank you for sharing these amazing years with me, for always having complete faith in me, and for being both my greatest challenge and my safest haven. I can't imagine my life without you. You and Gilli are my family, and I'm infinitely grateful for that. Ti amo.

Thank you, Emilia, Fabio, and Giovanna. Even though we're far apart and don't see or hear from each other often, you are my best friends. Not once have I doubted our friendship, and you are the reason I am always happy to return to Salerno.

Thank you to my colleagues: Ioana and Nidia. Sharing the office with you was a wonderful experience. You both lightened even the most stressful days, and I'll miss being a source of distraction for you and with you. Thank you, Ioana, for treating me like a friend from the very beginning, even when my words to answer you were so limited. Thank you for the great conversations and for always being there for me. Thank you, Nidia, for your valuable advice and information, always acting like a big sister and sharing your experience with us.

Thank you to the Italian friends who shared these years with me: FedeK, Ida, FedeP, Lorena, Michelangelo, Gabriele, Carmine, Damiano, and Gianmarco. With you, I've lived unforgettable moments, and I'm grateful every day to have met you. You're my little Italian crew that made me feel at home, even though I was so far away. I could write special thanks for each of you, but forgive me — you are too many. Just know that I care deeply for you all, and whether it's a night out dancing, a board game, or tea and gossip (with my fairies), every day with you is incredible. I'll keep annoying you even after leaving the university and Iceland, don't worry.

Last, but not least, I want to thank every single person who, even if only briefly, crossed paths with me here in Iceland. Each of you has contributed to this journey in your own way, giving me happy and peaceful moments. If today I look back at this experience with so much joy, it's also thanks to you.

Chi mi conosce da tanto sa che questi sono i ringraziamenti che mai avrei pensato di scrivere, data la mia profonda avversione al dottorato da sempre. Eppure, se c'è una cosa che ho imparato dal 2021 ad oggi è che niente va come ti aspetti, che un attimo può cambiare una vita intera e dopo un "viaggio breve di 5 giorni in Islanda" così è stato. Non è stato facile per me uscire fuori dagli schemi che la mia mente aveva prefissato, non è stato facile accettare di voler cambiare vita, Paese, prospettive lavorative, lasciando indietro amici, familiari, la mia tanto amata routine. Nonostante ciò, contro ogni aspettativa non mi sono pentita neanche un giorno. Neanche un singolo giorno ho guardato indietro desiderando altro che non fosse quello che stavo vivendo, e questo lo devo alle persone che mi sono state vicine, anche quando lontane, in tutto questo percorso.

Prima di tutti, devo ringraziare le mie supervisors. Penso sinceramente che non potessi essere più fortunata. Grazie María, per essere stata sempre una guida e un supporto per me. Grazie per aver letto e riletto i miei paper, assecondando il mio bisogno di conferme anche quando le modifiche riguardavano una virgola. Grazie per essere stata un mentore magnifico, ti auguro il meglio per la nuova vita che ti attende. Grazie Anna Sigga, per aver avuto sempre ottime idee per me, per aver sempre proposto le conferenze più interessanti e aver condiviso gli aperitivi italiani con me.

Grazie ai miei famigliari: Mamma, Papà e Michela. Grazie Mamma e Papà, per la fiducia che avete sempre riposto in me, fidandovi di ogni mia scelta di vita anche quando sembrava la più folle. Grazie per aver sopportato il dispiacere di avermi tanto lontana, vedendomi ogni sei mesi o meno, senza mai lamentarvi. Grazie Michela, per aver sempre creduto nelle mie capacità e avermi sempre spronato a vivere quante più avventure possibili.

Grazie Riki, per la vita bellissima che mi hai regalato in questi anni, e per le scelte che la tua esistenza mi ha portato a prendere che hanno portato sempre a qualcosa di bello. Grazie per aver condiviso con me degli anni bellissimi, per avermi sempre dato piena fiducia e per essere contemporaneamente il mio punto di scontro e il mio porto sicuro. Non voglio immaginare la mia vita senza di te. Tu e Gilli siete la mia famiglia, e ne sono infinitamente grata. Ti amo.

Grazie Emilia, Fabio e Giovanna. Nonostante siamo lontani,

nonostante ci vediamo e sentiamo raramente, siete i miei migliori amici. Non ho mai dubitato un singolo giorno della nostra amicizia, siete la mia conferma, e il motivo per cui sono e sarò sempre felice di ritornare a Salerno.

Grazie alle mie colleghe: Ioana e Nidia. Aver condiviso l'ufficio con voi è stata un'esperienza bellissima, avete entrambe reso le mie giornate stressanti sempre un po' più leggere, mi mancherà essere fonte di disturbo con voi. Grazie Ioana, per avermi trattato come un'amica fin dall'inizio, quando le parole a mia disposizione erano così limitate. Grazie per le belle chiacchierate e per esserci sempre stata per me. Grazie Nidia, per gli utili consigli e le informazioni che mi hai dato, comportandoti sempre un po' come una sorella maggiore e condividendo la tua esperienza con noi.

Grazie ai ragazzi italiani che hanno condiviso con me tutti questi anni: FedeK, Ida, FedeP, Lorena, Michelangelo, Gabriele, Carmine, Damiano e Gianmarco. Con voi ho vissuto momenti bellissimi e sono grata ogni giorno di avervi incontrati, la mia piccola crew italiana che mi ha fatta sentire a casa anche se da casa ero tanto lontana. Avrei ringraziamenti speciali per ognuno di voi, ma non me ne vogliate, siete troppi. Sappiate solo che vi voglio benissimo, e che sia una serata a ballare o un gioco da tavolo o un thè spettegolando (con le mie fatine), ogni giorno con voi è fantastico. Continuerò a rompervi le scatole anche quando sarò fuori dall'Università e dall'Islanda, non temete.

Ultimo, ma non per importanza, voglio ringraziare ogni singola persona che è stata anche solo di passaggio nella mia vita per pochi mesi qui in Islanda. Tutti voi avete contribuito a questo mio viaggio a modo vostro, donandomi momenti felici e sereni, e se ad oggi sono contentissima di quest'esperienza vissuta, è anche merito vostro.

Contents

Publications List

Publications included in the thesis

Chapter 4: Paper 1: Green Intentions vs. Behavior

C. Carpinelli, E. T. Einarsson Reynis, A. S. Islind, *et al.*, "Green intentions: Field research and data-driven analysis of customers' purchasing patterns", *Sustainability*, vol. 14, no. 16, p. 9863, 2022

Chapter 5: Paper 2: Impact of Social Media

C. Carpinelli, A. S. Islind, and M. Óskarsdóttir, "The quiet power of social media: Impact on fish-oil purchases in iceland during covid-19", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 18, 2024, pp. 203–213

Chapter 6: Paper 3: Nudging Towards Sustainability

C. Carpinelli, I. Visescu, A. S. Islind, *et al.*, "Digital nudging towards sustainability: Exploring strategies to encourage greener choices among adolescent girls", in *Proceedings of the Italian Chapter of AIS (Association for Information Systems)*, 2024

Chapter 7: Paper 4: Forecasting Against Overproduction

C. Carpinelli, B. Baesens, S. V. Broucke, *et al.*, "Novel insights on time series forecasting comparing deep learning and statistical models to avoid overproduction", *Paper Submitted to Expert Systems With Applications*, 2024

Other publications by author

C. Carpinelli, S. V. Broucke, B. Baesens, *et al.*, "Optimizing sustainability: Benchmarking time-series forecasting models on purchasing data to reduce overproduction", in *Proceedings of the Italian Chapter of AIS (Association for Information Systems)*, 2024

Abstracts presented at international scientific conferences

C. Carpinelli, A. S. Islind, C. A. Bravo, *et al.*, "Extreme climate events and credit risk: A stress testing approach for loans to smes based on multilayer network analysis", 2024. Abstract presented at International Conference of The Association for the Advancement of Artificial Intelligence (AAAI)

List of Tables

List of Figures

Chapter 1

Introduction

Food waste is a global issue with significant environmental, social, and economic consequences. Each year, approximately one-third of all food produced globally is wasted, amounting to about 1.3 billion tons [7]. This waste occurs across the entire food supply chain—from production and processing to retail and consumption. Food waste contributes to environmental pollution through the unnecessary consumption of natural resources, including water, land, and energy, as well as the release of greenhouse gases during decomposition in landfills [8]. For example, wasted food accounts for approximately 8-10% of global greenhouse gas emissions, making it a significant contributor to climate change [9]. Addressing food waste is critical not only to reduce these environmental impacts but also to ensure food security and improve resource efficiency.

Food waste is one of the primary challenges within the broader food system, which is inherently complex due to the numerous activities involved in producing, processing, distributing, and consuming food [10]. This system is a vast, interconnected network that starts with crop seeds and livestock and extends through various stages (cultivation, processing, transportation, and retail) to grocery stores, supermarkets, and restaurants where customers purchase and consume food [11]. A complex web of actors, including farmers, food processors, distributors, retailers, and customers, contribute to the operation of this multifaceted system (see Figure 1.1). Customers, grocery stores, production industries, and governments all play distinct roles, each with unique characteristics and responsibilities. Customers influence demand and purchasing patterns, while grocery stores and industries impact inventory management, production planning, and distribution efficiency [11], [12]. Governments, in turn, shape the regulatory and policy framework that governs sustainable practices.

Sustainable practices, in the context of this thesis, refer to actions, products, and behaviors that minimize environmental impact, conserve natural resources, and support long-term ecological balance. This intricate network of contributors makes addressing food waste a multifaceted challenge requiring targeted solutions tailored to the specific dynamics of each stakeholder [12].

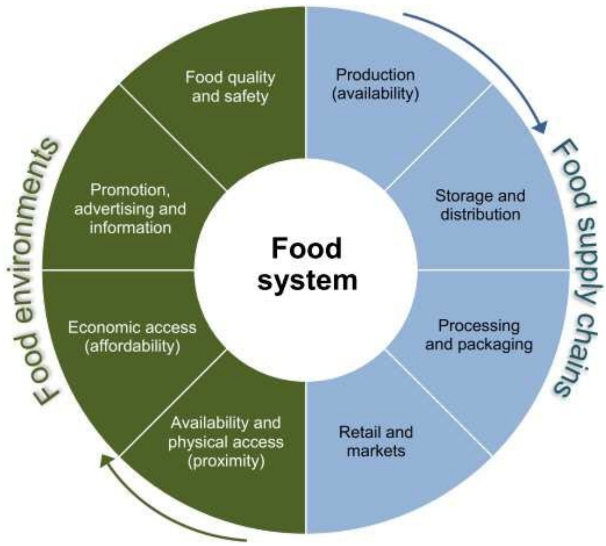


Figure 1.1: The food system. Figure extracted from [13].

The complexity of the food system generates vast and varied datasets that are essential for understanding and mitigating food waste. These datasets encompass customer purchasing behaviors, inventory and demand trends, production schedules, and environmental impact metrics [14]. In this context, sustainable practices encompass actions that aim to reduce food waste, promote responsible sourcing, and encourage environmentally conscious consumption.

For instance, groceries and supermarkets collect data on product sales, shelf-life durations, and waste rates, while production industries generate information on crop yields, processing timelines, and energy consumption. By integrating these diverse datasets, it becomes possible to uncover patterns, identify inefficiencies, and develop strategies to reduce food waste at multiple levels of the supply chain [15].

1.1 Grocery Sales for Sustainable Practices

Grocery sales are one of the actors in the food system and are pivotal in advancing sustainable food practices [16]. For instance, customers, through their purchasing decisions, affect demand patterns, which in turn influence production practices, distribution methods, and waste generation at various points in the supply chain [17]. On the other hand, producers, retailers, and policymakers can shape customer behavior through pricing, product availability, marketing strategies, and regulations, thus creating a feedback loop between actors that ultimately drives the sustainability outcomes of the food system [18].

The importance of addressing customer behavior in this context lies in its potential to both drive and transform the food system [19]. Customers represent the endpoint of the supply chain and their actions, whether consciously or unconsciously, shape much of what happens earlier in the process [18], [19]. However, despite being a crucial part of the system, customer behavior does not exist in isolation. It is influenced by external factors such as marketing, social norms, economic incentives, and policy measures, all of which contribute to the decisions customers make [20]. It is necessary to deeply understand these influences to make sustainability initiatives aimed at changing customer behavior more effective.

By recognizing the interconnectedness of all the actors in the food system, it is possible to better understand how customers' purchasing choices are not just the result of individual preferences but also shaped by broader structural forces [21], [22]. Increasing the food system's sustainability requires more than just encouraging customers to make "green" choices; it involves understanding how those choices interact with and are constrained by other parts of the system [23]. "Green products" refer to products that have a lower environmental impact compared to conventional alternatives.

Thus, the focus on customers is motivated by their critical role in driving the demand that shapes production and waste outcomes [24]. Moreover, understanding the gaps between customers' intentions and actual behaviors can inform strategies that align sustainable practices across the entire food system [25]. This focus provides insights not only into how to reduce food waste but also how to optimize the roles of all actors in creating a more sustainable food system.

Data science plays a crucial role in navigating grocery sales data. It enables researchers and decision makers to follow changes in purchasing behavior, investigate demand fluctuations, and develop predictive models to optimize resource allocation and reduce waste [14]. By applying data-driven methodologies on purchasing data, the food system can become more sustainable through interventions such as demand forecasting, inventory optimization, and customer education. These tools not only help minimize environmental impact, but also improve operational efficiency and support sustainable food systems [26].

1.2 Influencing Factors in Customer Purchasing Behavior

Customer purchasing behavior is influenced by various factors, from personal preferences and social norms to advertising and economic considerations [27]. While environmental awareness has encouraged some customers to adopt greener practices, everyday choices are still heavily driven by convenience, price, personal health, and societal influences [27]. Social media and digital platforms have added new dimensions to these influences, shaping customer attitudes and behaviors in significant ways [28].

Social Media

Social media plays a significant role in shaping customer purchasing behavior by influencing perceptions, preferences, and decisions [29]. Through targeted advertisements, influencer endorsements, and user-generated content, platforms like Instagram, Twitter, and Facebook create constant exposure to products and trends

[28]. Social media's visual and interactive nature also allows brands to engage customers more effectively, often personalizing advertisements based on user data and behavior [30].

The impact of social media on purchasing behavior is particularly strong for products associated with lifestyle, health, and wellness. For example, during the COVID-19 pandemic, social media heavily promoted supplements as immune-boosting, increasing customer interest and purchases [31].

Nudging

Nudging is a subtle intervention encouraging specific behaviors without restricting choices or changing economic incentives [32]. In sustainability, nudging can be an effective tool for influencing customer behavior in favor of eco-friendly products and practices [33]. For example, displaying eco-labels or positioning sustainable products at eye level in stores can encourage customers to make environmentally friendly choices [34]. Similarly, offering information on the environmental impact of specific products or highlighting their positive effects on health and the planet can nudge customers towards greener options [35].

Digital nudging, particularly in online shopping environments, leverages techniques such as personalized recommendations, default settings, and visual cues to guide customers' decisions [36]. For instance, e-commerce platforms might default to green delivery options or highlight sustainable products in search results. Nudging, both online and offline, thus plays a crucial role in aligning customer behavior with sustainability goals, encouraging small shifts that collectively contribute to large-scale change [37].

1.3 Encouraging Sustainable Consumption and Reducing Overproduction

Given the scale of environmental challenges, promoting sustainable choices in purchasing and consumption is essential. Sustainable consumption requires a decision-making process that considers not only individual needs and desires but also social and environmental responsibilities [38]. However, changing ingrained purchasing habits remains challenging, as convenience, cost, and habit often drive daily decisions, making customers resist change, even when they express support for sustainability and recognize its benefits [39]. The conflict between intention and behavior in customer choices underscores the complexity of promoting sustainable practices within current consumption patterns [40]. This complexity is mirrored on the supply side, where demand forecasting plays a crucial role in shaping what stores offer. Supply and demand are two sides of the same coin: customers' purchasing decisions influence what retailers stock, while the options made available by stores guide what customers buy [41]. Aligning these dynamics is essential for fostering a sustainable food system.

To encourage sustainable behavior, it is vital to motivate customers and producers toward environmentally responsible decisions, creating a cultural shift that values sustainable choices as the norm [42]. Achieving this goal requires a nuanced understanding of customer behavior, including not only what influences purchasing decisions but also the contexts in which those decisions are made. This understanding can be informed by studying trends and identifying specific motivators that encourage sustainable actions [43]. Strategies such as nudging have effectively influenced customer choices. By implementing these nudging techniques, such as highlighting sustainable products at critical points of sale, companies can make eco-friendly options more visible, accessible, and appealing [44].

Reducing overproduction and overconsumption is also critical for the food system, as it addresses the environmental impact of wasted resources and the financial losses associated with unsold products [45]. Methods like demand forecasting, powered by data analytics, offer promising solutions to minimize food waste by aligning production volumes with actual customer needs [46]. Through accurate demand forecasting, retailers and producers can better manage inventory and avoid surplus production, which not only reduces waste but also conserves the resources and energy used in production [47]. This approach contributes to a more sustainable food system by balancing supply with demand, reducing environmental impacts, and fostering more efficient and responsible resource use.

1.4 Motivations and Research Focus

This thesis adopts a comprehensive data-driven approach to analyze grocery sales data, combining diverse data sources to gain a deeper understanding of purchasing patterns and their implications for sustainability. By integrating fine-grained longitudinal purchasing data from supermarkets with survey data representing customers' purchasing intentions, the research bridges the gap between what consumers say they intend to do and their actual behavior. This dual perspective enables the identification of trends, behaviors, and discrepancies in consumer choices, offering valuable insights into the factors driving purchasing decisions.

Using data science techniques, the analysis highlights opportunities to align consumer behavior with sustainability goals. These insights can inform strategies to raise customer awareness, encourage more sustainable choices, and reduce food waste, thereby fostering a more environmentally conscious food system. Ultimately, this approach demonstrates how data analytics can play a pivotal role in promoting sustainability at both the individual and systemic levels.

The aim of this thesis is three-fold:

A1 First, it aims to examine what kind of purchasing behavior customers have towards green products and what their perceptions of food purchasing and food waste are.

A2 Second, it aims to analyze and discover customers' purchasing trends, how external factors - like social media - contribute to changing these purchasing trends, and which methods can nudge them towards more sustainable choices.

A3 Finally, the aim is to analyze and predict food demand in grocery stores to decrease overproduction and consequent food waste.

From these aims, it was possible to formulate the following research questions(RQ):

RQ1 *How do grocery customers perceive their purchasing intentions towards green products and how do they really purchase?*

RQ2 *How do social media impact purchasing behavior?*

RQ3 *How effective are certain nudging techniques at influencing customers towards greener purchasing?*

RQ4 *Which is the most accurate forecasting model to predict customers' demand in order to avoid overproduction?*

Research Question	**Aim**
RQ1	A1
RQ2	A2
RQ3	A2
RQ4	A3

Table 1.1: Aims related to RQs.

The problems addressed in this thesis represent only part of the problems in the food system that continue to contribute to climate change [48]. In this thesis, the main focus is on customers and their consumption in grocery stores and how every purchasing decision is linked to production in order to front some of the issues that arise during the production phase. For this reason, the thesis relies on the data from one of the largest supermarket chains in Iceland, and according to the different studies conducted, different types of data are collected based on the specific problem addressed.

The data used in this thesis enlighten a large part of Iceland's purchasing patterns and provide novel insights into grocery purchases. It can further shed light on ways to reduce food waste.

This thesis makes a key contribution by providing a comprehensive analysis of customer behavior and sustainability in the food system. Through the use of extensive real purchasing data, it bridges theory and practice, delivering actionable insights to encourage sustainable consumption and minimize food waste.

1.5 Outline and Contributions

This thesis focuses on analyzing customer purchasing behavior within the food system, exploring how it can inform sustainable choices and reduce food waste. This section further outlines the thesis structure, composed of three published papers and one submitted. This section summarizes the thesis objectives, primary findings, and chapter contributions. Collectively, each chapter builds a comprehensive view of how, when, and why customers make purchasing decisions and how these choices can be influenced to promote sustainability.

The contribution of this thesis also concerns one of the 17 Sustainable Development Goals (SGD) [49]: Goal 12, which aims to ensure sustainable consumption and production and purchase of products in general [50].

Presented in Figure 1.2 is an overview of all the interactions discussed in this thesis, accompanied by the corresponding chapter(s) addressing them. At the center is the customer, for whom we study purchasing behavior from various aspects as outlined below.

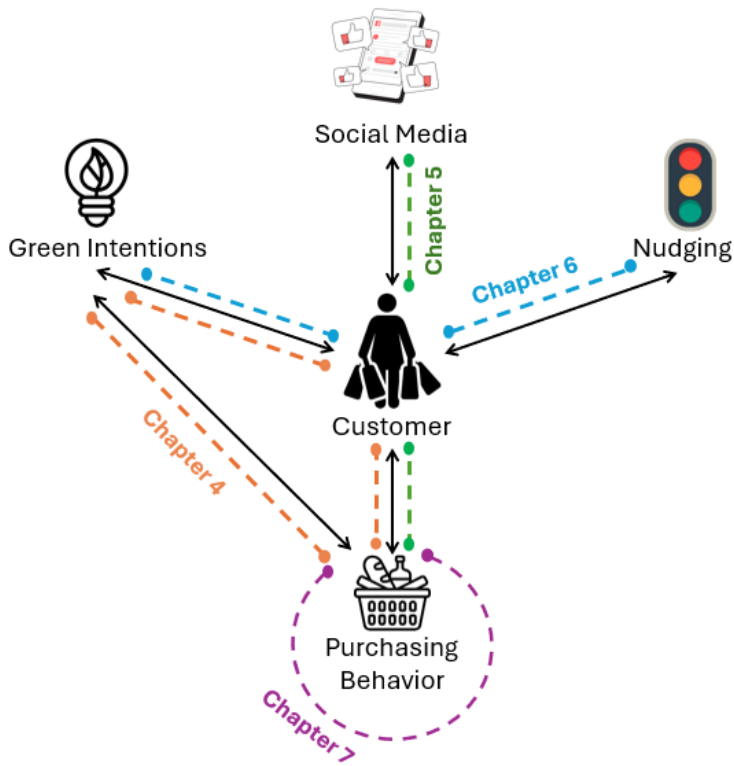


Figure 1.2: Overview of the chapters in the thesis.

Chapter 1 The current chapter provides an overview of the food system as well as its challenges. It also outlines the motivation and research focus, along with a general overview of the chapters included in the thesis, the relationships between different actors in the studies conducted, the data used, and their main findings.

Chapter 2 This chapter elaborates on the related work conducted to date on applying data science to understand customers purchasing behavior, their influence factors and food demand to reduce food waste. It also explores strategies to influence customers towards greener choices and how external factors such as social media impact purchasing behavior.

Chapter 3 This chapter outlines the research paradigm adopted in the thesis and presents a methodological overview of the diverse studies included, describing the data employed in each.

Chapter 4 In this chapter, we explore the relationships: customers $\leftrightarrow$ green intentions $\leftrightarrow$ purchasing behavior. We present a study conducted in 2021, in which we gathered subjective data asking potential customers to fill out a survey with the aim of understanding their environmental concerns and their intentions toward purchasing green products. Then, we compared the results with objective purchasing data pulled from one of the largest supermarket chains in Iceland.

We discover that women tend to have more environmental concerns and are more intent on buying green products than men. However, there is no actual correlation between customers' intentions and their actual purchasing behavior. Apparently, customers want to be greener, but in the end, they are not.

The full reference of this published paper is:

C. Carpinelli, E. T. Einarsson Reynis, A. S. Islind, *et al.*, "Green intentions: Field research and data-driven analysis of customers' purchasing patterns", *Sustainability*, vol. 14, no. 16, p. 9863, 2022

Chapter 5 This chapter examines the social media ↔ customer ↔ purchasing behavior relationship, focusing on Iceland during the COVID-19 pandemic. Specifically, we investigate how omega-3 (fish oil) supplements, widely advertised on social media as potential preventatives or remedies for COVID-19, may have impacted customer purchasing behavior.

We analyze Twitter posts before, during, and after the pandemic, comparing them with fish oil purchasing data from Iceland. Our findings reveal a statistically significant correlation between social media activity and purchasing patterns, suggesting that Icelandic customers were likely influenced by social media in buying fish oil supplements.

The full reference of this published paper is:

C. Carpinelli, A. S. Islind, and M. Óskarsdóttir, "The quiet power of social media: Impact on fish-oil purchases in iceland during covid-19", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 18, 2024, pp. 203–213-Under review.

Chapter 6 This chapter examines the nudging ↔ customer ↔ green intentions relationship, explicitly exploring how variations of the traffic light nudging technique impact different groups of teenage girls. We designed a survey based on conjoint analysis principles, where four distinct groups of girls were asked to choose between three snack options with varying levels of sustainability, ranging from least to most sustainable. Each group encountered a different version of the traffic light technique to assess which, if any, approach most effectively nudged them toward greener choices.

Our findings reveal that the variant displaying only a green traffic light on the most sustainable snack was the most effective at influencing choices, while other combinations showed less impact. Interestingly, even when a red traffic light marked the least sustainable option, participants were generally unwilling to avoid it, suggesting limitations to this nudging method in discouraging less sustainable choices.

The full reference of this published paper is:

C. Carpinelli, I. Visescu, A. S. Islind, *et al.*, "Digital nudging towards sustainability: Exploring strategies to encourage greener choices among adolescent girls", in *Proceedings of the Italian Chapter of AIS (Association for Information Systems)*, 2024

Chapter 7 This chapter analyzes purchasing behavior as a standalone factor to forecast demand across various time horizons and aggregation levels. The goal is to determine which forecasting approach yields the most accurate predictions under different scenarios, ranging from deep learning models to classical statistical methods. Accurate demand forecasting is essential to minimize overproduction and reduce food waste.

To assess model performance, we aggregate purchasing data at daily, weekly, and monthly levels and predict demand for both identical and progressive time horizons across these granularities. This analysis helps identify the optimal data granularity and prediction horizon for the most accurate results. We further introduce scenarios with missing data to evaluate model resilience under challenging conditions.

Interestingly, our findings reveal that traditional models like ARIMA [51] perform comparably well against newer methods, such as TimeGPT [52], suggesting an ongoing debate on the best forecasting approach between classical and deep-learning techniques.

The full reference of this submitted paper is:

C. Carpinelli, B. Baesens, S. V. Broucke, *et al.*, "Novel insights on time series forecasting comparing deep learning and statistical models to avoid overproduction", *Paper Submitted to Expert Systems With Applications*, 2024

Chapter 8 This chapter discusses the significance of the findings in previous chapters, how they contribute to increasing sustainability, and how they face some of the challenges of the food system.

In this chapter, the main implications of the thesis concerning data, resources, and support are presented and contextualized.

This chapter also focuses on discussing the research limitations, potential biases, and constraints of the thesis.

Chapter 9 The final chapter concludes the thesis by summarizing key insights, contributions, and closing reflections.

It also looks forward to outlining future research directions and highlighting new questions and challenges that have emerged from this work.

Chapter 2

Related Work

The previous chapter introduced the food system in general, focusing on the food waste issue, and outlined the research focus of this thesis. This chapter focuses deeper into the role of data science in analyzing purchasing behavior by examining studies on customer purchasing and exploring how this behavior can be influenced through nudging techniques and social media. It reviews prior research on how data-driven insights can guide customers toward more sustainable choices, highlighting both the benefits and potential drawbacks of behavioral interventions in the food system, particularly against food waste.

2.1 Purchasing Behavior Analysis

In the past few decades, data science has revolutionized the study and understanding of customer purchasing behavior, leveraging large datasets to uncover meaningful patterns, predict future trends, and inform decision-making [53], [54]. In recent years, the influx of data from various sources, such as point-of-sale (POS) systems, loyalty programs, online transactions, social media, and even sensor data from IoT-enabled devices, has provided a wealth of information that can be analyzed to obtain information on customer behavior [55]. The intersection of data science and purchasing behavior analysis offers important opportunities for companies to optimize their operations, adapt marketing efforts, predict demand, and better understand the factors driving customer choices [56]. Data science can analyze large amounts of transaction data from multiple sources and, in real time, provide a continuous flow of information that reveals patterns across different customer segments, periods, and geographic locations [53].

Data used in the analysis of purchasing behavior encompass a variety of types, each providing unique insights into customer habits and trends. One key source is transaction data, which is captured through point-of-sale systems. These systems record every purchase, detailing products, quantities, prices, and the time and location of transactions. This data has been widely studied to uncover purchasing trends, such as seasonal variations, popular products, and peak buying periods. For example, research leveraging transaction data has revealed significant shifts in purchasing patterns linked to holidays, weather changes, or economic conditions [57]–[59].

Customer data, encompassing demographic profiles, loyalty program participation, and personal opinions, serves as a valuable resource for gaining insights into consumer behavior [60]. Often gathered through surveys or customer interactions, this predominantly qualitative data enables effective profiling and segmentation, allowing businesses to tailor their strategies to specific customer groups [61].

A persistent challenge in fostering sustainable consumption is the gap between customers' stated environmental intentions and their actual purchasing behaviors [62]. While many consumers express a desire to make eco-friendly choices, their actions often diverge due to barriers such as convenience, cost, or insufficient information. Data science offers solutions to bridge this gap by identifying behavioral patterns and designing targeted interventions that promote sustainable decisions. For example, tailored recommendations, real-time feedback, or personalized incentives can nudge customers toward choosing environmentally friendly products. By leveraging these insights, companies can craft marketing strategies that not only resonate with individual preferences but also encourage sustainable choices [63]–[65].

Social media and web data also play a significant role in understanding purchasing behavior. Platforms like Facebook, Instagram, and e-commerce sites generate vast amounts of unstructured data, including product reviews, likes, shares, and comments. These interactions reflect customer sentiment, brand perception, and responses to marketing campaigns. Analyzing such data can help businesses assess public opinions and adapt their strategies accordingly, creating a more dynamic connection with customers [66], [67].

External environmental data, such as weather patterns, local events, and economic indicators, further enriches the analysis of purchasing behavior [62]. These external factors can have a profound influence on customer decisions, prompting demand fluctuations that are often predictable with the right data models. Incorporating this type of information allows businesses to anticipate changes in customer needs and adjust supply chains or promotional strategies proactively [68], [69].

Together, these diverse data sources provide a comprehensive view of purchasing behavior, enabling the development of sophisticated models and strategies that align with customer demand while supporting sustainability goals.

Over the years, different studies have analyzed purchasing behavior from various perspectives. Some studies provide personalized recommendations based on a customer's past purchases and browsing history [70], [71].

Other studies apply demand prediction for efficient inventory management [72]. Machine learning models can forecast demand by considering historical data, seasonal trends, and external factors. Accurate demand forecasting helps prevent stockouts and reduces waste, especially in sectors like grocery retail, where products have limited shelf life [73].

Customer segmentation is another practice that is widely used [74]–[76]. Based on purchasing behavior, companies can target marketing efforts more effectively. Additionally, predictive models can estimate the lifetime value of customers, allowing businesses to allocate resources toward retaining high-value customers [77]. Thanks to data science, it is also possible to assess the effectiveness of promotions by tracking changes in purchasing behavior [78]. By comparing purchasing data before and after a promotion, companies can determine its impact and optimize future promotions for greater effectiveness [78].

Advances in machine learning, artificial intelligence, and big data analytics have pushed data science to the forefront of analyzing purchasing behavior [79]. Today, data science enables more accurate customer segmentation, nuanced demand forecasts, and the prediction of complex, multi-factor customer behaviors [80].

As the volume of available data grows, data science will play an increasingly central role in purchasing behavior analysis, enabling companies to respond dynamically to customer needs, optimize operations, and promote sustainable consumption [81]. By leveraging data-driven insights, businesses can not only increase their competitive edge but also contribute to more efficient, ethical, and customer-centric practices across industries [82].

2.1.1 Social Media Influence on Purchasing Behavior

Social media significantly shapes customer purchasing behavior by influencing preferences, awareness, and the perceived social value of products [83]. The platform's reach, especially among younger customers, allows companies to drive engagement and build brand loyalty through targeted advertising, influencer marketing, and direct interactions [84]. Brands can tailor ads based on user interests, online behaviors, and demographics, making recommendations that resonate individually and often lead to increased purchase intent [29].

The impact of social media on purchasing behavior is amplified by influencer marketing, where influencers, often trusted by their audiences, endorse products [85]. This form of peer endorsement creates a perception of authenticity and credibility that traditional advertising may lack. Customers are more likely to trust product recommendations from people they follow, integrating these endorsements into their decision-making process [86].

By analyzing social media data, companies can track customer sentiment and assess the impact of influencer marketing [87]. Social media sentiment analysis can reveal shifts in brand perception and help companies respond proactively to customer feedback [87].

2.2 Data Science for Increasing Sustainability

Data science has become a powerful tool for promoting sustainable purchasing, enabling businesses and customers to make more environmentally conscious choices [88]. Through predictive analytics, real-time monitoring, and advanced modeling techniques, data science supports sustainable purchasing efforts by optimizing resource use, reducing waste, and promoting greener products [89], [90].

2.2.1 Nudging Toward Sustainable Choices

Nudging, a behavioral science technique, has emerged as an influential tool for subtly shaping customer behavior, particularly in the realm of sustainable purchasing [91]. Unlike mandates or strict incentives, nudging employs gentle interventions designed to guide customers toward more eco-friendly choices while preserving their autonomy [92]. By restructuring the decision-making environment to emphasize sustainable options, nudging helps customers make greener choices without feeling overwhelmed or coerced [92]. With the rise of digital platforms, this concept has evolved into digital nudging, which extends traditional techniques into the online domain, offering new opportunities for fostering sustainable consumption behaviors [93].

Traditional nudging techniques have been widely implemented in physical retail settings to encourage sustainable purchasing [94]. One prominent example is traffic light labeling, where products are marked with simple, color-coded indicators—green for sustainable options, yellow for moderate environmental impact, and red for high impact [95], [96]. This straightforward visual system allows customers to quickly assess a product's sustainability at a glance, eliminating the need for extensive research or label reading [96].

Another common approach involves shelf placement and product positioning. Sustainable products are often placed at eye level or near high-traffic items to boost visibility and likelihood of selection [97]. Research indicates that customers are more inclined to choose products that are conveniently located or prominently displayed, making strategic placement a highly effective nudging technique [98].

Eco-certification and badging also play a significant role in nudging. Products labeled as organic, fair-trade, or environmentally friendly frequently carry distinctive marks that signal their sustainable credentials. These labels appeal to customers' values of ethical consumption and environmental responsibility, subtly guiding them toward more eco-conscious choices [99].

By integrating these methods into both physical and digital retail spaces, nudging fosters an environment where sustainable options are more accessible and appealing, encouraging customers to adopt greener purchasing behaviors with minimal resistance.

With the rapid shift to digital shopping platforms, digital nudging has become a key strategy in promoting sustainable choices [93]. Digital nudging operates in the virtual environment, leveraging website layouts, recommendation algorithms, notifications, and other online features to guide customers' behavior [100]. In the online context, digital nudges can adapt to user behaviors in real-time and can be personalized, creating highly effective and responsive methods to promote sustainable choices [100]. While nudging holds great potential to encourage sustainable behaviors, it also raises important ethical considerations. Transparency is crucial; customers should be aware that nudges are in place and should be informed about how and why their choices are being influenced [101]. As data science and digital technologies continue to evolve, the future of nudging in sustainability looks promising. Advancements in AI and machine learning will enable even more personalized and responsive nudging techniques, enhancing the relevance and effectiveness of sustainability-promoting nudges [102]. Additionally, integration across platforms—such as linking e-commerce sites with social media, fitness apps, and eco-footprint calculators—could provide a seamless and holistic approach to influencing sustainable behavior [103].

2.2.2 Demand Forecasting and Sustainability

Demand prediction, or demand forecasting, involves using historical data, seasonal trends, external variables, and machine learning algorithms to predict future purchasing behavior [104]. This can help businesses optimize their stock levels and production schedules, making it possible to respond efficiently to actual demand without overproducing [105].

Demand forecasting today utilizes a wide array of models, from traditional statistical approaches to advanced machine learning techniques, each offering unique strengths for different forecasting needs [106]. Classical models, such as ARIMA (Auto-Regressive Integrated Moving Average), remain popular, particularly for scenarios with stable demand patterns [107]. These models leverage historical data to identify seasonality and trends, making them highly effective for applications where demand is predictable and consistent.

Machine learning approaches, including Random Forest, Gradient Boosting, and deep learning models like Long Short-Term Memory (LSTM) networks, have brought new accuracy to demand forecasting [104]. These advanced models excel when demand is complex and driven by multiple, often interdependent factors, capturing non-linear relationships and providing more robust forecasts in situations with high volatility or unexpected events [108]. Such models have demonstrated success in forecasting applications where demand patterns shift rapidly or are influenced by a wide range of external variables [72].

The potential of accurate demand forecasting for sustainability is significant. For example, in the grocery industry, accurate demand predictions can prevent the overstocking of perishable items, such as fresh produce and dairy, which have short shelf lives [73]. By anticipating how much stock will actually sell within a specific time frame, supermarkets can reduce the volume of unsold products and decrease food waste [109]. Beyond food, demand prediction is valuable for industries such as fashion, where overproduction often leads to significant waste when unsold items are discarded [110].

As data science continues to advance, the precision of demand forecasting will only improve. This evolution offers valuable opportunities for supporting sustainable purchasing by aligning production closely with actual customer needs, reducing overproduction, conserving resources, and, ultimately, building a more sustainable future for both businesses and the environment.

Chapter 3

Research Design

This chapter describes the comprehensive methodological approach adopted in this thesis, detailing the data sources, analytical frameworks, and techniques applied in each of the studies. It aims to provide an overarching view of the methodology, illustrating how the research stages were structured to address the central research questions and contribute to understanding sustainable customer behavior and food waste optimization.

The methodological approach of this thesis is grounded in the pragmatism approach [111]. Pragmatism is characterized by its focus on practical problem-solving and its flexibility in combining different research paradigms to address complex research questions [112]. Pragmatists believe that reality is dynamic and shaped by changing contexts, which necessitates adaptable approaches to research. This approach supports the use of both quantitative and qualitative techniques, blending positivist and interpretivist methods as needed to achieve the most effective outcomes [113].

In the context of this thesis, pragmatism allows for the integration of diverse data sources, analytical techniques, and perspectives to comprehensively explore the multifaceted issues of sustainable consumption and food waste reduction. The research design reflects this by employing a mixed-methods approach that combines quantitative data analysis with qualitative insights, ensuring that each method aligns with the specific research questions and objectives.

The research follows a seven-stage process, depicted in Figure 3.1. Each stage is designed to build incrementally toward robust insights and conclusions, incorporating diverse analytical approaches tailored to the specific requirements of each study.

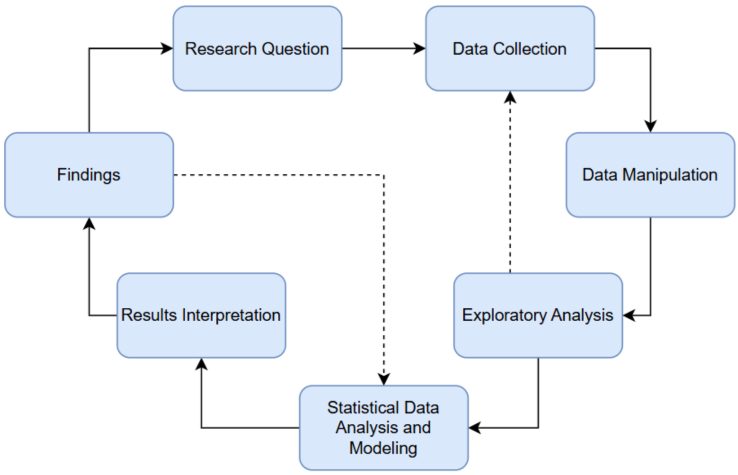


Figure 3.1: Overview of the research approach stages of this thesis.

The process begins with the definition of the research question, where the key problem is identified, and the scope of the inquiry is outlined. This foundational stage includes a critical review of the literature to contextualize the research and identify relevant gaps. The research questions were designed to address both theoretical and practical aspects of sustainable consumption. On the theoretical side, this involves contributing to the academic understanding of concepts like sustainable consumption and food waste management. Practically, it seeks to solve real-world problems or offer actionable insights in these areas. This dual focus underscores the relevance of the research for both academic audiences and industry practitioners. By doing so, it positions the thesis as both scientifically rigorous and practically applicable.

The second step is data collection, which involves multiple sources, each chosen based on the specific aims of the chapters. Table 3.1 provides a summary of the data sources utilized across the thesis. These included purchasing data from one of the largest Icelandic supermarket chains, survey responses, and social media insights.

The datasets were diverse, capturing customer behavior from multiple perspectives: transactional data for time series analysis, survey data for understanding customer attitudes, and social media content for assessing digital influence on purchasing patterns.

Chapter	Data Source
[4] Green Intentions: Field Research and Data-Driven Analysis of Customers' Purchasing Patterns	Survey (1) Purchasing Data
[5] The Quiet Power of Social Media: Impact on Fish-Oil Purchases in Iceland during COVID-19	Social Media Data Purchasing Data
[6] Nudging Towards Sustainability: Exploring Strategies to Encourage Greener Choices Among Adolescent Girls	Survey (2)
[7] Novel Insights on Time Series Forecasting Comparing Deep Learning and Statistical Models to Avoid Overproduction	Purchasing Data

Table 3.1: Data sources included in each chapter of the thesis.

Following collection, the data underwent rigorous preparation and cleaning. Transactional data was formatted for time-series analysis, addressing issues such as missing values and outliers. Survey responses were anonymized, coded, and categorized to facilitate statistical analysis. For social media data, text was preprocessed to remove noise. For survey data concerning the nudging experiment, sentiment analysis was applied to assess reliability. These steps ensured data quality and compatibility with the chosen analytical methods.

The thesis incorporates a variety of analytical techniques, reflecting the interdisciplinary nature of the research objectives. These techniques were applied depending on the specific data and research questions of each study.

In Chapters [4], traditional statistical methods were employed to analyze survey responses. Tests such as chi-square, t-tests, and correlation analysis were used to examine relationships between customer attitudes and purchasing behaviors. Structural break analysis was applied in Chapter [5] to assess significant shifts in purchasing patterns linked to social media posts during COVID-19.

Chapter [7] employed advanced time-series analysis techniques to model purchasing patterns and develop forecasting models. Statistical methods, such as ARIMA [51], were compared with deep learning approaches to determine the most effective strategies for reducing overproduction and waste.

Survey data in Chapter [6] was analyzed using sentiment analysis and choice-based conjoint analysis to understand whether the nudging techniques had an effect on the participants.

The final stages of the research process involved interpreting the results and synthesizing the findings into coherent conclusions. Each chapter's results were contextualized within the broader objectives of the thesis, connecting customer behavior insights to actionable recommendations for promoting sustainability and reducing food waste.

By integrating diverse methodologies and data sources, this thesis offers a multifaceted exploration of sustainable customer behavior, providing valuable insights for academia, industry, and policymakers. This structured methodological framework ensures that each chapter contributes meaningfully to the overarching narrative, building a cohesive argument for sustainability in food systems.

3.1 Researcher's role

Following Reykjavik University's rules, a declaration of authorship contribution must be submitted to the Research Group on Constitutional Studies in the Computer Science Department. The declaration reports the researcher's degree of involvement at different stages of the research and publication process of the papers included in this thesis.

As outlined in Appendix A, Table 3.2 details my contributions to the four papers included in this thesis. The appendix provides a template for categorizing a Ph.D. student's effort across various stages of the publication process using abbreviations (ME, EE, CE, LE). This framework focuses on standardizing authorship contributions and does not specify the student's specific role, such as project lead or coordinator.

Paper name	**Idea**	**Related work & literature**	**Data gathering**	**Research design**	**Artifact design**	**Analysis & synthesis**	**Draft**	**Administration**
Green Intentions: Field Research and Data-Driven Analysis of Customers' Purchasing Patterns	LE	LE	CE	EE	ME	ME	ME	ME
The Quiet Power of Social Media: Impact on Fish-Oil Purchases in Iceland during COVID-19	ME	ME	CE	ME	ME	ME	ME	ME
Nudging Towards Sustainability: Exploring Strategies to Encourage Greener Choices Among Adolescent Girls	ME	CE	ME	EE	ME	ME	ME	ME
Novel Insights on Time Series Forecasting Comparing Deep Learning and Statistical Models to Avoid Overproduction	EE	ME	ME	EE	ME	ME	ME	ME

Table 3.2: Declaration of authorship contribution.

Chapter 4

Green Intentions: Field Research and Data-Driven Analysis of Customers' Purchasing Patterns

This paper explores the gap between customers' intentions and their actual purchasing behavior regarding green products. We conduct a two-part study: a survey and a data analysis of purchasing data from a large Icelandic retail store. The survey, conducted online, collects data on customers' attitudes and intentions toward green purchases, while the purchasing data provides insights into real-world buying patterns. The study finds a significant difference between customers' expressed intentions and their actual purchasing behavior. Despite a high environmental concern among the surveyed participants, the purchasing data shows a lack of responsiveness to products with lower environmental impacts.

Full reference:

C. Carpinelli, E. T. Einarsson Reynis, A. S. Islind, *et al.*, "Green intentions: Field research and data-driven analysis of customers' purchasing patterns", *Sustainability*, vol. 14, no. 16, p. 9863, 2022

4.1 Introduction

In recent years, environmental issues have escalated significantly, including increased carbon emission, which further enhances global warming, as well as excessive pollution due to toxic emissions and massive amounts of garbage [114]. As one means to counteract this evolution, there has been a surge in the invention and production of so-called "green products", often made of recycled or low-polluting materials.

Despite this, one of the most important parts of managing the imminent environmental challenge is customers' awareness rooted in individual dedication and commitment towards the solution [115], [116]. The motivation for green purchasing should begin with the customer [115], [117]. The most common motivations behind the increasing demand for sustainable production and products are compliance and avoiding negative effects; these two factors are the most frequent drivers of companies to act sustainably [117].

As the environmental issues grow, it becomes vital to make plans for the future. Sustainability poses multiple challenges and opportunities for businesses [114]. Consequently, tremendous benefits lie ahead for businesses that develop sustainable products and services. It is therefore important to study how customers are responding to products' lowered environmental impact to be able to adjust green marketing strategies accordingly and for manufacturers to see value in moving towards sustainable manufacturing processes.

In this paper, we study customers' attitudes towards green products and compare them to actual purchases of such goods. Our approach is two-fold. On the one hand, we study customers' attitudes and intentions towards green products through a survey. The survey presents a direct response from the customer and allows us to study their views and intentions to choose environmentally friendly products. On the other hand, we take a data-driven approach to customers' purchasing behavior specifically focusing on products that have gone through a 'green' transition towards decreasing their environmental impact. The data was provided by a large retail store in Iceland. Analysis of the purchasing data allows us to investigate the customers' actual response to the introduction of green products and whether this introduction actually affects their purchasing behavior.

Our research is based on the following two research questions, in which the first one is oriented towards survey data analysis, while the second one is oriented towards purchasing data analysis:

i *What type of attitude do customers have when it comes to green purchases?*

ii *Does the environmental impact of products have an effect on customers purchasing behavior?*

The main contribution of this paper is conceptualizing customers' green purchases. We show, through our survey data, that customers, in general, have high maturity when it comes to environmental awareness and would like to make green purchases, but through the purchasing data, we show that there is a low follow-through. The purchases are generally not green. This calls for new types of initiatives that can support customers in their in-store decision-making processes since the customers *want* to act more sustainably but are, to a large extent, unable to do so.

The rest of this paper is organized as follows. In the next section, we present literature related to our research, focusing on green products and purchasing. The third section explains the research approach and the methods used. Section four presents the results and their interpretation. Finally, section five presents the discussion of this research, and section six presents the conclusions and direction for future work.

4.2 Related Work

Nowadays, there is a growing popularity of 'green' products to increase sustainability and decrease the carbon footprint. To date, there is no unified definition of the concept 'green products.' One definition of green products is "when its environmental and societal performance, in production, use and disposal, is significantly improved and improving in comparison to conventional or competitive products offerings" [118].

More simply put, green products are products that have a lower environmental impact compared either to other products, or to prior versions of the same product. Producing a product with a low environmental impact means ensuring that, throughout the complex life cycle (from the extraction of raw materials to disposal as waste (or recycling)), the contribution to environmental change, in at least one of its matrices, e.g., air, water, and soil, is reduced. The reduction of environmental impact can be achieved by those who make the product through improvements in production processes or technologies, either directly managed or influenced upstream or downstream of their position in the production chain.

From a purchasing perspective, with increasing threats to the environment, an increasing number of people have started to pay attention to sustainable development to protect the environment and society. The concepts of green purchase and green marketing have gradually become popular [119], [120]. Theories such as the Theory of Reasoned Action and Theory of Planned behavior underline the research in customers purchasing behavior [121]–[123]. These models aim to explain volitional behaviors and have been used to predict behavioral intentions and behaviors [124], and identifying strategies for changing behaviou. These theories have been used to a large extent, to shed light on the sustainability debate. Green purchase intention (GPI) is the concept of customers willingness and probability to give preference to products who have environmentally friendly features over other products in their purchase intention. It plays customers experience with green products an influential role in their green purchase decisions [125]. In relation to customers inquisitiveness to gain knowledge on the environmental aspects and features of green products, the customers strive to gain knowledge on their ingredients, the environmental impact and their functionality [120]. This further influences customers purchasing decision and enable them to make right informed purchasing decisions and develop their willingness-to-pay higher for green products [126], [127]. Purchasing decisions are in a form of purchasing green products, supporting green companies [128], [129], and adopting sustainable consumption [130]. Green customers are those who take the environmental impact of their consumption pattern into consideration and intend to modify their purchasing and consumption pattern to reduce their environmental impact [119]. It is interesting to understand what the difference is between customers' perception of their actions towards sustainable purchases and what their behavior is, studying purchasing patterns [131], [132]. There has been already other studies that compare customers' attitudes and their actual behavior towards green products, like the study in South Africa that compared females' attitudes and purchase intentions towards green cosmetics [133]. That study demonstrates that there is a positive correlation between women's intention and their purchasing behavior. However, other research shows that customers are inclined to buy green products only when there are immediate and tangible benefits, as well as being more environmentally friendly [134].

4.3 Methods

This research study relied on a quantitative data analysis approach, as the research focuses on a numerical analysis of purchase data and analysis of survey results. The quantitative approach is structured for numerical and statistical analysis where prearranged processes, questions and design which pose certain flexibility issues[135].

To answer the research question *"Does the environmental impact of products have an effect on customers purchasing behavior?"*, data is obtained by two methods.

First, we conducted a survey that provided data directly from the customers and gave a perspective on their purchasing attitudes and intentions toward green products. secondly, we gathered customer purchasing data from a large retail store on select products. The purpose of the purchasing data analysis is to study how customers react to products lowered environmental impact. The results are then compared to research how customers attitude and intentions differ from their actual purchasing behavior.

4.3.1 Intentions

Survey Methodology

To answer our first research question, we created a survey in a web- and email-based configuration. The Web-based section was distributed to customers using social media. This proves beneficial in reaching a large number of respondents with a quick turnaround

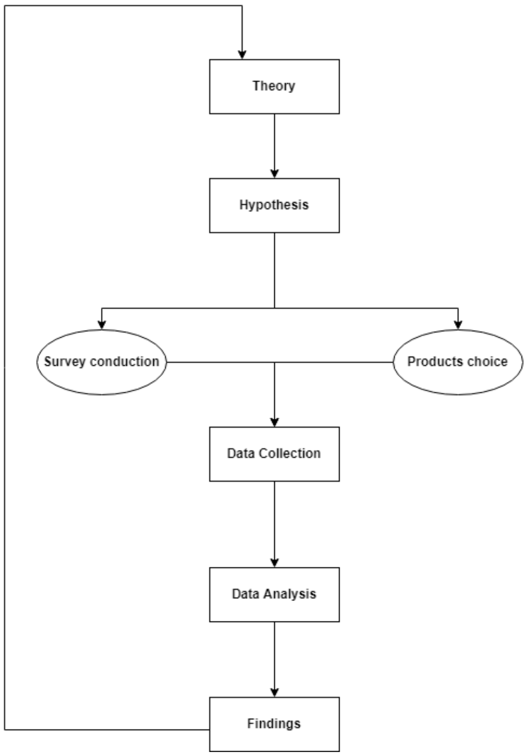


Figure 4.1: Research process.

time. The Reykjavik University (RU) email service was utilized to conduct the email-based survey. The respondents are students and faculty at the university.

For the structure of the survey, we used a 5 point Likert scale. The possible answers were: *Strongly agree*, *Agree*, *Neither agree nor disagree*, *Disagree* and *Strongly disagree*.

Each possibility was given a corresponding number for Statistical Package for the Social Sciences (SPSS) analysis ranging from 5 to 1, with 5 representing strongly agree, and 1 representing strongly disagree. These options should represent a symmetric and balanced choice selection with the position of the neutrality exactly between the two extremes of strongly disagree and strongly agree [136].

Survey Questions

The questions in the survey were derived from a number of researches on similar topics [123], [125], [137], [138], but structured in a way to fit this research more adequately. Table 4.1 shows the survey questions.

Table 4.1: Survey questions

Description Frequency (percentage)	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
1. What is your age?					
2. What is your gender?					
3. I read the products label before deciding on purchasing the product.	27 (16)	88 (53)	34 (20)	9 (5)	8 (<5)
4. I am an environmentally friendly customer.	17 (10)	81 (49)	57 (34)	11 (7)	0 (0)
5. I would change my purchasing behavior because of social pressure.	6 (4)	70 (42)	49 (30)	31 (19)	10 (6)
6. I think that the environmental impact of products is stated clearly enough on product packaging.	1 (1)	18 (11)	43 (26)	75 (45)	28 (17)
7. I would consider using an app with informations on products' environmental impact.	29 (17)	67 (40)	29 (17)	25 (15)	16 (10)
8. I know my carbon footprint.	4 (2)	22 (13)	31 (19)	57 (34)	52 (31)
9. I have taken measures to reduce my carbon footprint.	20 (12)	71 (43)	32 (19)	32 (19)	10 (6)
10. I want manufacturers to offset their carbon footprint through Kolviður or other similar platforms.	40 (24)	69 (42)	47 (28)	7 (4)	3 (2)

The survey was sent by email to the students of Reykjavik University and by social media on March 30, 2021 and ended on April 6, 2021. The survey contained ten questions, took two minutes to complete and participants were encouraged to answer all questions consciously. Anonymity and confidentiality was also guaranteed.

Survey Analysis

Survey analysis is used to answer the first research question. The survey questions are analyzed through Reliability analysis, Principal Component Analysis (PCA) and Ordinal Logistic Regression Analysis (OLRA). This procedure is necessary to study the reliability of the scale used and how questions correlate with each other for guidance on classifying the questions in components.

Reliability analysis allows to study the properties of measurement scales and the items that compose the scales [139]. As parameters for the analysis we used Corrected Item-to-total correlation (CITC), Cronbach's alpha, variance and standard deviation. The Cronbach's alpha coefficient is a commonly used indicator of internal consistency and a measure of the underlying construct [140]. A Cronbach's alpha of 0.7 or higher is considered acceptable [141]. But, since the survey contained fewer than ten questions, low Cronbach's values are commonly found [140]. The CITC value depicts how each question correlates to all other questions in the analysis.

The Principal Component Analysis (PCA) method transforms variables into a set of new composite variables (principal components) that are not directly correlated [142]. This approach is useful to reduce the number of related variables to smaller dimensions to explore the underlying structure of the set variables [140]. The questions were then classified in sections depending on the joint components between questions.

Ordinal Logistic Regression Analysis (OLRA) is a statistical analysis method used in this research to answer the first research question. PCA analysis in section 4.4.1 were used to analyse both GPI and Environmentally Convern (EC) and a mean value calculated.

4.3.2 Behavior

Purchasing data

To study the actual purchasing behavior, we analyzed time series data on the quantities of daily product purchases provided by a major Icelandic supermarket. We had information about the daily number of quantities sold of each product, spanning over two years, from January 1, 2019 until mid March, 2021.

As the goal of this research is to study the consumption of green products, we narrowed the dataset to products that met the following requirements:

1 The product underwent changes to lower its environmental impact;

2 The product has been on the market for longer than the period which the retrieved data covers;

3 The product has information on the lowered environmental impact on its exterior or packaging.

Among all the brands and products available in the dataset, the Tresemmé products satisfied all the above mentioned conditions, therefore this research will focus on the Tresemmé product line. Their products had undergone a transition in its environmental impact in the time frame that this research is focusing, with information on these changes clearly stated on the products container. Moreover, they launched a marketing campaign stating that their new packaging was more environmental friendly. That, in addition to the noticeable information clearly stated on the product container about the lower carbon footprint, rendered the product feasible. While the Tresemmé containers are comprised of plastic, they are recyclable but have been indistinguishable by the sorting machine at the recycling centre. To sort this issue, Tresemmé developed a new detectable pigment for the plastic containers which allows them to be detected by sorting machines at the recycling centres. According to Tresemmé, these changes will save up to 2.500 tonnes of black plastic going to waste each year. These changes have already been implemented to Tresemmé products and with hopes to decrease black plastic waste. At this point we would like to highlight that the products illustrated specifically (for instance Tresemmé), are used for illustrative purposes, and as such, it is not the product per se that we focus on, instead the focal point is on the impact that the product represents.

The purchasing data covers over two years, and depicts ten different Tresemmé products bought from 28 stores from, early January in 2019, until middle of March in 2021. The data is split in three sections, Tresemmé regular (REG) , Tresemmé recyclable (REC) and the comparison products (NORM). REG are the Tresemmé products which have not experienced any changes in their containers in terms of recycling, REC are the Tresemmé products that have undergone the pigment modifications. NORM are the Tresemmé products used for direct comparison with REG and REC in terms of correlation between 2019 and 2020.

Analysis of Purchasing Data

The data was analyzed by conducting a correlation analysis between Tresemmé NORM, Tresemmé REC and Tresemmé REG and comparing with the correlations of the same sections from a period before the lowered environmental impact of the Tresemmé products. This allowed us to study customers reaction towards these changes. A Pearson correlation analysis for 2019/2020 was conducted after normalising the data and ANOVA linearity testing was established.

To study the significance of correlation coefficients between 2019 and 2020, calculation of correlations between NORM20Q2 and REC20Q2, and NORM19 and REC19 will be conducted by the following equation:

$$Z_{obs} = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}} \tag{4.1}$$

N_1 and N_2 are respectively the total usable days for 2019-2020 and for 2020 Q2 - 2021 Q1.

Using Correlation Coefficient Table in [140] and considering NORM20Q2 and REG20Q2 as r1 and NORM19 and REG19 as r2, inserting r1 = .125 and gaining a corresponding Z_1 value of .126 and same for the r2 = .212 with a Z_{obs} value of .213, the equation gives with N1 = 348 and N2 = 363 a Z_{obs} value of 1.154. If the Z_{obs} value is between -1.96 and +1.96, there is not a statistically significant difference between the two correlation coefficients [140].

Same procedure is done for NORM20Q2 and REG20Q2 using r1 = .283 and NORM19 and REG19 with r2 = .270, the results give a Z_{obs} value of 0.212 which is also between the -1.96 and +1.96 range signifying that there is not a statistically significant difference between the two correlation coefficients.

4.4 Results

The results of the present research work are divided in Survey results and Data results. All analyses was conducted using SPSS.

4.4.1 Intentions

First of all, we present the main characteristics (gender, age) of the survey participants. Then, we describe the results of analysis and testing.

Participants Characteristics

The target audience for the survey were people 18 years or older, since the subject of sustainability and Environmentally Friendly Products (EFP) may be difficult to comprehend for minors [143]. A total of 166 people responded to the survey, of whom 93 were male or 56%, 71 were female or 43% and two individuals or 1% self-identified as other. The participants age range was from 20 to 69 years old, with 92 or 55% aged between 20 - 29 years old. Respondent origin was predominantly from Iceland, 161 or 97%, and the email-based survey accounted for 118 or 71% of the participants. The participation ratio from the email-based survey was around 5.5%, but the ratio is uncertain from the social media survey.

Table 4.2: Survey's participants characteristics

Variables	Categories	Frequency	Percentage
Gender	Male	93	56
	Female	71	43
	Other	2	1
Age	20 - 29	92	55
	30 - 39	38	23
	40 - 49	20	12
	50 - 59	10	6
	60 - 69	6	4
Origin	Iceland	161	97
	Other	5	3
Survey	E-mail	118	71
	Facebook	46	28
	Google	1	<1
	Other	1	<1

Reliability Analysis

The internal scale consistency is one of the main concerns, or the degree by which the questions that make up the scale "hang together".

Question 6 (Q6) showed a negative CITC value of -0.052 which may have indicated that the question was negatively worded and was moving in a different direction compared with other questions. To adjust for this, a re-coding process was conducted where each value in the 1 - 5 Likert scale was reversed to correlate with the other questions [144]. After re-coding Q6, the CITC value was positive at 0.052 but still too low [140], and was therefore excluded from further analysis. Table 4.3 depicts the reliability analysis.

Table 4.3: Reliability Analysis

Question	Corrected Item-to-total correlation	Cronbach's (alpha)	Var	Std.dev
Q3	.397	.736	19.612	4.429
Q4	.462			
Q5	.264			
Q7	.520			
Q8	.384			
Q9	.623			
Q10	.525			

All CITC values are within the acceptable range of 0.3 [140], apart from Q5 which is then revised. Additionally the Cronbach's alpha value is above 0.7 which is acceptable, with the mean Inter-item Correlation factor of .288 which is within the range of .2 - .4 [140]. These questions were used in a Principal Component Analysis for guidance on forming the Environmental Concern (EC) and Green Purchase Intention (GPI) sections.

Principal Component Analysis

Table 4.4 shows the correlation matrix of the PCA, i.e. how each question relates to one another. This should not be directly compared to the CITC value explained above as the CITC value compares each question the other questions as a total corrected value.

Table 4.4: Correlation Table about Survey's questions

Question	Q3	Q4	Q5	Q7	Q8	Q9	Q10
Q3	1	.240	.092	.262	.314	.375	.239
Q4	.240	1	.172	.287	.275	.377	.424
Q5	.092	.172	1	.290	.064	.188	.243
Q7	.262	.287	.290	1	.240	.454	.423
Q8	.314	.275.	.064	.240	1	.398	.201
Q9	.375	.377	.188	.454	.398	1	.491
Q10	.239	.424	.243	.423	.201	.491	1

As per the litterature the correlation values should preferably be 0.3 or above for the majority of the matrix [140]. Here, of the 21 values, 8 are above 0.3 and 7 are close to the 0.3 value, and are used in further analysis.

Figure 4.2 demonstrates how the variance is distributed among the components graphically.

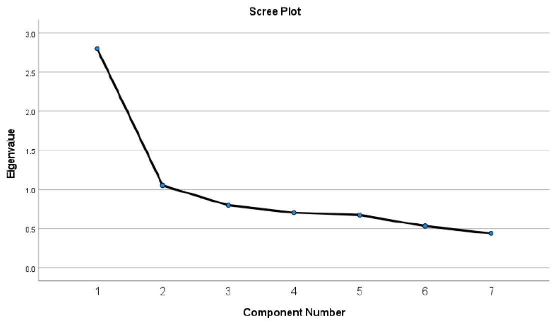


Figure 4.2: Plot of variance distribution.

For the principal components (labelled component) shown on the x-axis, the corresponding eigenvalue is plotted on the y-axis [145]. Since component one and two have eigenvalues greater than one, two components are used to divide the questions into components, with component one and two accounting for 55% of the total variance. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) value is required to be 0.6 or higher and the Bartlett's Test of Sphericity value to be significant ($p<0.05$) [140]. These values are crucial for verification before proceeding with the principal component analysis, both of these requirements are met with the KMO of 0.806 and Bartlett's Test of Sphericity significant value of ($p<0.01$).

Table 4.5: Pattern and Component Matrix for PCA with Oblimin rotation

	Communalities		**Component Matrix**		**Pattern Matrix**	
	Initial	Extraction	C1	C2	C1	C2
Q9	1	.631	.785	-117	**.597**	.378
Q10	1	.570	.717	.235	.261	**.636**
Q7	1	.534	.689	.245	.234	**.627**
Q4	1	.411	.641	.007	.402	**.395**
Q3	1	.493	.562	-.421	**.707**	-.017
Q8	1	.578	.556	-.519	**.784**	-.104
Q5	1	.635	.399	.690	-.318	**.829**

Table 4.5 shows the Communalities matrix, Component matrix and the Pattern matrix. The table depicts how questions Q3, Q8 and Q9 fit in component 1 which represents that they have commonalities or underlying traits, and will be classified together

to represent Green Purchase Intention (GPI). Same for Q5, Q7 and Q10, their values favour component 2 and will be classified together to represent the Environmental Concern (EC) section. As seen in the pattern matrix, Q4 is closely related to both sections but will be classified in the EC section since the wording of the question adheres more to that section.

Reliability Analysis on PCA

Table 4.6: Reliability analysis on GPI and EC classified questions

Section	Q	Corrected Item-to-total Correlation	Cronbach's	Var	Std.dev
GPI	Q3	.412	**.630**	5.834	2.415
	Q8	.432			
	Q9	.478			
EC	Q4	.386	**.628**	7.277	2.698
	Q5	.318			
	Q7	.465			
	Q10	.508			

Table 4.6 depicts the sectioning of questions to EC and GPI. A reliability analysis reveals the CITC values are all above the acceptable limit of 0.3 [140], and the Cronbach's values as above 0.6. As discussed in chapter 4.4.1 the Cronbach's values may be low when there are fewer than 10 questions.

Ordinal Logistic Regression Analysis

Before conducting the OLRA, a normality test of the data was carried out which depicted a Goodness-of-Fit values for Pearsons Chi-Square values of 159.479 with a significance value of 0.850, and a Deviance Chi-Square value of 148.285 with a significant value of 0.955. A test of normality under Kolmogorov-Smirnov was carried out, and it showed a significance value of ($p<0.05$), which indicates that the data is not normally distributed. Additionally, a model fitting Chi-squared value of 42.149 with a significant value of ($p<0.01$).

These factors indicate that the model fits the data well but since the data does not satisfy the normally distributed condition, the Spearman's correlation coefficient is used in answering to the first research question [140].

Table 4.7 shows the correlation between EC, GPI, age and gender. Age proves not to be statistically significant to either EC or

Table 4.7: OLRA Correlations

Correlations						
Nonparametric Correlations			**EC**	**GPI**	**Age**	**Gender**
Spearman's rho		Correlation Coefficient	1	.465**	-.152	.228**
	EC	Sig (2-tailed)	.	<.001	.050	.003
		N	166	166	166	166
		Correlation Coefficient	.465**	1	.002	.065
	GPI	Sig (2-tailed)	<.001	.	.977	.404
		N	166	166	166	166
		Correlation Coefficient	-.152	.002	1	-.101
	Age	Sig (2-tailed)	.050	.977	.	.195
		N	166	166	166	166
		Correlation Coefficient	.228**	.065	-.101	1
	Gender	Sig (2-tailed)	.003	.404	.195	.
		N	166	166	166	166
****. Correlation is significant at the 0.01 level (2-tailed)**						

GPI. However, gender depicts a positive correlation to EC, gender of participants was categorised and females given a value of 2 and males 1. This correlation is positive, which indicates that females have more correlation with EC than males, with the correlation value indicating the strength of the relationship between the two variables. According to [146] in [140], the range for a medium correlation value is rho = 0.30 to 0.49. The correlation value between EC and GPI is $\rho = 0.465$ which is then considered as a medium positive correlation value. With the significance level of ($p<0.01$) for EC and GPI, the result of the first research question is that there a positive correlation between customers EC and GPI.

4.4.2 Behavior

In this section, we present first the Trend analysis of Tresemmé purchasing products through the entire time period (2019 -2021), then we analyse the correlation between comparison products (NORM), Tresemmé REC and Tresemmé REG, for 2019 and 2020 - 2021 separately.

Trend Analysis

The purchasing of all the Tresemmé products through the time period is depicted in Fig 4.3.

A linear trend line fits the data well with a R squared value of

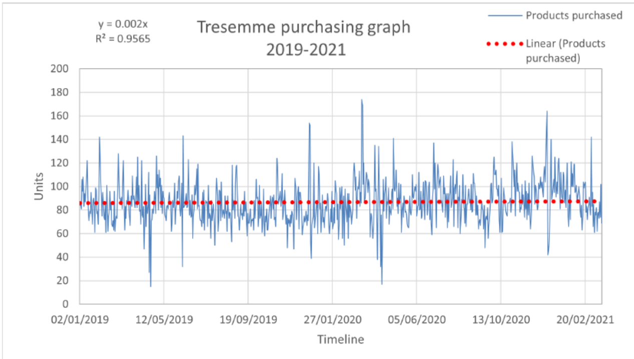


Figure 4.3: Purchasing of all Tresemmé products.

R^2 = 0.9565, indicating a slight increase in quantity bought. This purchasing graph is comprised of the cumulative purchasing of both REC and REG Tresemmé products, which is then separated and analyzed separately in tables 4.8 and 4.9.

Figure 4.4 with a linear trend line with a R squared value of R^2 = 0.9429, shows the purchasing data of the 45 different products which is used for comparison with Tresemmé REG and Tresemmé REC data shown in figure 4.3.

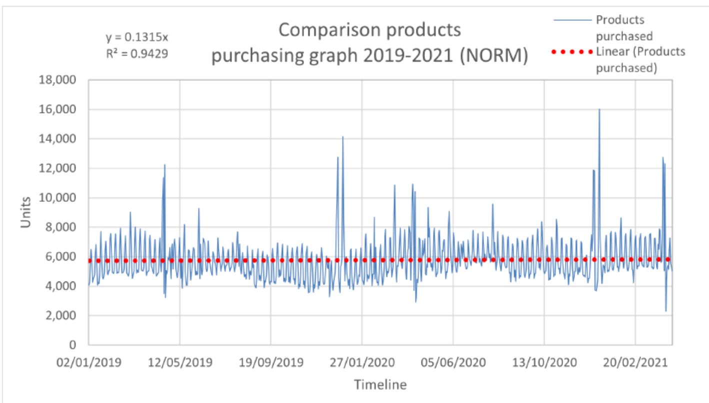


Figure 4.4: Purchasing of comparison products.

By analyzing the correlation between the comparison products

(NORM), Tresemmé recyclable products (REC) and Tresemmé regular products (REG) and comparing with the correlations of the same sections from a period before the lowered environmental impact of the Tresemmé products. It enables for researching the customers reaction towards these changes. A Pearson correlation analysis for 2019 is conducted after normalising the data and ANOVA linearity testing to check differences between different groups, respectively: NORM/REC and NORM/REG for 2019 and NORM/REC NORM/REG for 2020.

Table 4.8 depicts the correlation in 2019 for:

1 NORM19 and REG19 as r = 0.270;

2 NORM19 and REC19 as r = 0.212;

3 REG19 and REC19 as r = 0.215.

All this correlation have a significant value of ($p<0.01$).

Table 4.8: Correlation analysis for 2019

	Correlations		
	NORM19	**REC19**	**REG19**
NORM19 **Pearson Correlation**	1	.212**	.270**
Sig. (2-tailed)		<.001	<.001
N	363	363	363
REC19 **Pearson Correlation**	.212**	1	.215**
Sig. (2-tailed)	<.001		<.001
N	363	363	363
REG19 **Pearson Correlation**	.270**	.215**	1
Sig. (2-tailed)	<.001	<.001	
N	363	363	363
****. Correlation is significant at the 0.01 level (2-tailed)**			

According to Tresemmé in [147] the containers pigment changes and recyclable labelling entered the customers market in 2019. However, Nathan & Olsen the Tresemmé products distributor in Iceland, states that the recyclable labelled containers would have entered the Icelandic market in Q2 of 2020 [148]. In table 4.9 accounts for this and depicts the correlation between:

1 REC20Q2 and NORM20Q2 as r = 0.125 with a significance level of ($p<0.05$);

2 REG20Q2 and NORM20Q2 is r = 0.283 with a significance level of ($p<0.01$);

3 REG20Q2 and REC20Q2 as r = 0.199 with a significance level of ($p<0.01$).

Table 4.9: Correlation analysis Q2 2020 to Q1 2021

	Correlations NORM20Q2	REC20Q2	REG20Q2
NORM20Q2			
Pearson Correlation	1	.125*	.283**
Sig. (2-tailed)		.019	<.001
N	348	348	348
REC20Q2			
Pearson Correlation	.125*	1	.199**
Sig. (2-tailed)	.019		<.001
N	348	348	348
REG20Q2			
Pearson Correlation	.283**	.199**	1
Sig. (2-tailed)	<.001	<.001	
N	348	348	348

***. Correlation is significant at the 0.05 level (2-tailed)**
****. Correlation is significant at the 0.01 level (2-tailed)**

4.5 Discussions

The current food system neither works for all habitats of the planet, nor does it work for the environment. Changing our food system to a system based on the principles of circular economy to a larger extent, would be one of the utmost important steps towards building biodiversity on the one hand, and tackling climate change on the other hand [149]. One small step towards building a more sustainable food system, one that could over time lead to a regenerative food production and in turn lead to more stable soil, improved biodiversity, better air and water quality with both local and global benefits, would be to change the way each individual makes their individual choices [150]. Although eliminating food waste and shifting towards the principles of circular economy are the ultimate target, the individual choice, and follow-through, is where we are contributing to the change. Reaching towards the sustainability goals and analyzing aspects of managing the grand challenge related to the environmental impact of our actions, is a vital way forward in studying how customers react to products attributes and environmental impact [150].

The surveys results in Section 4.4.1 show how customers Environmental Concern and their Green Purchase Intention are positively correlated by a value of 0.465, a medium correlation value. So, for the first research question, the results show that although customers may be concerned about the environment and intend to purchase green products, their actions do not reflect these concerns and intentions. Additionally, table 4.7 shows how female customers correlate more with Environmental Concern than males. This indicates that women have a higher tendency towards environmental concern. However, when analyzing the purchasing behavior of customers, through actual purchasing data, the findings do not indicate high environmental concern in action, as customers do not seem to have responded with Tresemmé's products lowered environmental impact.

We would like to highlight that the products highlighted specifically in the purchasing data, are used for illustrative purposes, and as such, it is not the product per se that we focus on, instead the focal point is on the impact that the product represents. As depicted in table 4.9, the correlation is lower for Tresemmé recyclable and the comparison data, than for Tresemmé regular and the comparison data. With Tresemmé recyclable products in this period now displaying the recyclable label, it can be assumed that with lower correlation than Tresemmé regular and the comparison data, customers may be responding to Tresemmé's recyclable labelling. But as paragraph 4.3.2 depicts, the difference between correlations in 2019 and 2020 does not reach statistical significance. Consequently, for the second research question, the results indicate that even though customers are aware of their purchasing behavior, they do not react to products lowered environmental impact.

The results are in accordance with the findings by [151], where the results indicated that the perception of green products, labels, packaging and ingredients did not influence customers' perception. Other product attribute played an important role in the customer purchasing process, such as price and quality [152].

The results of this research can be deepened and used for further psychological and economic studies to understand how to influence customers towards the purchase of green products. In addition, the use of data-driven analysis was useful to understand the difference between customers' perceptions of their actions and their actual behavior. As pointed out earlier, our aim was not to link specific individuals perceptions of their actions and their behavior, but instead we used the two-fold analysis to shed light on the views of the general public, and the purchasing behavior of a wide-range of shoppers within the general public too. We do that to illustrate the intention of the general public, and actions of the general public but not to pinpoint the intentions and actions of certain individuals.

4.5.1 Limitations and Future Work

The present research has certain limitations. The sample data is limited in two ways, firstly, the survey had a limited sample of 166 participants, and secondly, the purchasing data was collected from a single retail store with a limited selection of suitable products for customers purchasing data analysis. Additionally, part of the data is related to COVID-19 period, so the purchasing data results may be skewed because of the restrictions.

Then, the two parts of our analysis, i.e. on the purchase intention and the actual purchase, did not include the same participants, and we do not know if they are from the same social groups. Also, the survey more broadly examined environmental impact, while our purchasing data shows purchasing patters for specific product categories. This could potentially affect the results in some way. Since our intention was to shed light on the gap between attitude and intent versus customers' green purchasing behavior and not to draw specific conclusions, the data suffices for that purpose.

Also, before conducting a survey of similar nature as this research, it would be wise to use a control group for testing the preliminary results in conducting a quantitative analysis to ensure the suitability for its purpose in the research. Furthermore, there are multiple purchase considerations in the customers decision process. Bias towards the author of this study may be present in the surveys participants from Reykjavik University or through social media. Additionally, the author of this study had previously encountered the products used in the study which may have influenced the decision process in search of suitable products for data analysis. Measures should be taken by future researchers to minimise such bias if possible, or take the bias into consideration when interpreting the results.

For further research in this area, it could be interesting to conduct a more complete analysis and include purchasing patterns of a larger assortment of products that meet the requirements discussed in 4.3.2. Additionally, future research could emphasise other relevant factors, such as environmental-, health-, personal- and social benefits of better aligning environmental intentions with purchasing actions. This could furthermore be extended to situations where more emphasis is on customers behavior towards the environment, whereby conducting in-depth interviews with customers could be explored to hopefully gain enriched understanding of green customer purchasing behavior.

4.6 Conclusions

This study presents a two-fold analysis of survey data and purchasing data in Iceland. In terms of a theoretical contribution our analysis shows that there is a discrepancy between customers concern for the environment and their intention to purchase green products and actual purchasing of green products. Based on the findings from the survey and the purchasing data analysis, the answer to the research question is that despite high environmental concern detected within the general population, there is low concern detected from purchasing data from customers when grocery shopping. Ergo, customers have difficulty translating their intentions into purchasing behavior based on the environmental impact of the products available to them. Based on that we conclude that there is a discrepancy between customers intention and their action, this discrepancy can be defined as a gap related to how customers perceive their purchasing behavior versus how they actually conduct their purchasing. They see themselves as environmentally responsible, but when analyzing purchasing patterns, there is less impact to be detected. In terms of a practical contribution we suggest that there is a need for change in the way grocery stores line up their products, in order to nudge customers to align intentions with actions. This could be done by more visibly portraying products with a lower environmental impact, in eyesight. It is clear that customers want to make an impact, so allowing them to more easily grab green products compared and less easily products with higher carbon footprint, would be beneficial for the environment.

Chapter 5

The Quiet Power of Social Media: Impact on Fish-Oil Purchases in Iceland during COVID-19

This paper examines the impact of social media posts on fish oil purchases in Iceland during the COVID-19 pandemic. Using a large dataset from a supermarket chain, Google search trends, and Twitter data, we analyze the influence of online information on offline purchasing behavior. The findings indicate a strong correlation between the spread of social media posts promoting fish oil as a remedy for COVID-19 and a significant increase in fish oil sales in Iceland during the pandemic. This suggests that customers were influenced by social media posts, potentially leading to increased fish oil consumption. The analysis reveals a notable lack of online research about the effectiveness of fish oil, suggesting that customers may have relied heavily on social media information without seeking further verification. The study highlights the potential for social media to impact purchasing behavior and emphasizes the need to critically evaluate online information, particularly during periods of uncertainty and heightened anxiety.

Full reference:

C. Carpinelli, A. S. Islind, and M. Óskarsdóttir, “The quiet power of social media: Impact on fish-oil purchases in iceland during covid-19”, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 18, 2024, pp. 203–213

5.1 Introduction

In the last decade, social media have become an integral part of our daily lives and have radically changed the way we communicate, gather information and make purchasing decisions [153], [154]. With the advent of social media platforms, customers now have access to a vast amount of information about products and services that were once not easily accessible [155].

Social media have a huge impact on purchasing behavior, as customers increasingly rely on reviews and recommendations from their peers and social media influencers to make informed decisions [84], [156]. Platforms such as Instagram, Facebook and Twitter, have become a powerful marketing tool for companies to showcase their products and services and build brand awareness [157].

At the beginning of the COVID-19 pandemic, fear and uncertainty dominated people’s perception and reaction to this new and unknown disease, which was not yet fully studied, and a cure was not well known. Consequently, in this atmosphere of uncertainty, people were impressionable in advice on how to avoid and treat the disease, and both true and false news spread quickly using the internet, and in particular social media platforms [158], [159].

Furthermore, there were several such news and advice regarding the relationship between certain diets or supplements that reduced the risk of contracting COVID-19. Some of them even went as far as stating that certain diets or supplements would cure COVID-19. As an example, before COVID-19, there was evidence supporting several micronutrients, like zinc and vitamins C and D, as key components in aiding the immune system when fighting infections [160]–[162]. Numerous other nutritional supplements exist, including omega-3 fatty acids – also known as fish oil –, probiotics, and plant isolates such as garlic [163].

In Iceland, fish oil stands out as a remarkably popular product. Renowned for its health benefits [32], [164], fish oil has been embraced by the Icelandic population for decades. Because of all the benefits it brings, fish oil was widely recommended worldwide to prevent and treat COVID-19.

Iceland's geographic isolation, heightened during the COVID-19 pandemic, underscores a unique context in which fish oil emerges as a popular and enduring supplement, with established purchasing patterns, mainly recording increased sales during the fall [165]. In light of this, establishing fish oil as beneficial in the fight against COVID-19 introduces a compelling scenario.

Although its effectiveness, in general, has been proven, research shows that opinion on the effect of fish oil on COVID-19, in particular, spread at an incredible speed during the pandemic [166].

The purpose of this paper is not to prove or debunk fish oil and its connection to COVID-19. Instead, this article aims to quantitatively assess the potential susceptibility of Icelandic individuals to fish oil promotion in the context of the ongoing pandemic. Through this investigation, we analyze the spread of social media posts regarding the benefits of fish oil versus COVID-19 and its impact on customer purchasing behavior.

We are guided by the following research question: "*How did social media posts impact the Icelandic population's purchasing behavior during the COVID-19 pandemic?*". To answer the research question, we collected and analyzed different data sources. We had the unique opportunity to collaborate with the second largest supermarket chain in Iceland, which provided us with purchasing data, so it was possible to compare the effects of the online world on the offline one. Besides, we conducted a netnography by extracting the time series of posts published advertising fish oil.

Then, we analyzed the data using different techniques, to show that the increase of fish oil purchasing was not casual but apparently influenced by social media. More information will be detailed in the methods section.

The main contribution of this paper is our conceptualization of social media influence in information systems through a data-driven approach to social media activity and its cultivation in an offline world, crystalized through changes in purchasing behavior.

5.2 Related Work

5.2.1 The Impact of Social Media on Purchasing Behavior

Social media have significantly influenced customer behavior over time, revolutionizing the way they interact with brands, make purchasing decisions, and engage with products and services [167], [168].

Social media exert a significant influence on purchasing behavior through its provision of abundant product information and reviews. Users can effortlessly access comprehensive details about products, peruse reviews from other customers, and compare various options within these platforms. This profusion of information empowers individuals to make more informed decisions when it comes to their purchases [169]. Furthermore, social media platforms have become the domain of influential individuals known as influencers who often endorse and recommend products or services through their substantial followings[170], [171]. Their opinions carry weight and hold the power to influence the decisions of their followers, resulting in increased sales for businesses [172].

Moreover, people are naturally inclined to be influenced by the choices and preferences of others [173], [174]. When they witness their peers or admired influencers endorsing specific products, it fosters a sense of validation and trust. As a result, individuals are more likely to make similar purchases, influenced by the actions of those they admire. Regardless of influencers or just regular people sharing content, social media have become a primary source of information, where customers research products, read reviews, compare prices, and gather insights before buying [170].

During sudden events like COVID-19, social media became a very important source of information [175]. The pandemic caused confusion and concern about health and wellness among the public [176]. In particular, lots of dietary news - both true and unproven - circulated on the internet [176]. Among those, there were news items stating that certain foods or supplements can protect against or even cure viruses in general and COVID-19 in particular, and as mentioned earlier in this paper, certain supplements would prevent or reduce symptoms. Although some of this news was unproven, the information was still spread through social media platforms and shared widely by netizens citizens of the internet, and they drastically changed customers' purchasing behavior [177].

5.2.2 The Relationship Between Online Behavior and Offline Behavior During COVID-19

One of the large-scale changes during the pandemic was customers' purchasing behavior and purchasing trends [178]. A study conducted by [179] shows evidence of customers changing their purchasing behavior drastically due to COVID-19, mostly moving to a larger extent towards online purchases and e-commerce. More specifically, various studies indicate that many customers opted for online shopping in order to avoid going out for an enhanced feeling of safety and to cultivate an increased feeling of protection from the spread of the virus [180], [181]. Moreover, some people were forced toward e-commerce and buying their groceries online because in some countries many shops were closed due to drastic social distancing measures taken.

In this context, data analysis can play a crucial role in understanding the changes in purchasing behavior as a result of online information [182], [183]. One way that data analysis has been used is to track changes in sales data over time [1]. For instance, by analyzing sales data before and during the pandemic, it was possible to demonstrate that we can identify changes in customer behavior directly related to the onset of the COVID-19 pandemic [184]. The power of data analysis can be utilized for a wide variety of product categories or for multiple geographical regions to identify specific areas where certain types of information may have had a greater impact on customer behavior [185] and [186].

5.2.3 Research Contribution

Inspired by this type of research, we explore the impact of advertisements online, spread by netizens and their effect on the offline world, where customer behavior plays a crucial role. Another way that data analysis can be used, is to track online searches and social media platform activity related to COVID-19 and customer products through a netnography approach, which is coupled with the data analysis herein. By analyzing these types of data sources, for instance internet data and purchasing behavior data, we show that it is possible to identify trends in the types of social media posts that circulated and their impact on customer behavior [83].

In this study, we had the unique opportunity to analyze the purchasing behavior of an entire nation. Iceland is isolated and was even more so during the pandemic [187].

In contrast to many other nations where platforms such as Amazon played a pivotal role, Iceland lacked a robust online shopping infrastructure [188]. The absence of major e-commerce platforms heightened the complexity of transitioning to online purchasing, a transition that became particularly critical during periods of restricted mobility [189].

One of the very few ways available in Iceland to do shopping online is by using the website of the supermarket you want to refer to.

The limited availability of comprehensive online marketplaces posed challenges for residents accustomed to the convenience offered by platforms like Amazon. Access to a diverse array of products online became a notable challenge, leading to an increased reliance on local businesses and traditional retail channels [190].

The pandemic was a sudden event that caused both fear and uncertainty and as such, it changed people's purchasing behavior. customers started to buy different products in different quantities, which is why some studies speak of a structural break [191]. A structural break might occur when there is a war, a major change in government policy, or some equally sudden events like for instance, a sudden onset of a worldwide pandemic. Certainly, COVID-19 represented a breaking event for the purchasing trends that had existed up to that point, and it is interesting to study how much social media contributed to this change.

5.3 Methods

As stated earlier, we set out to investigate the impact that social media posts about fish oil as a remedy against COVID-19 had on customer purchasing behavior in the offline world, through their in-store and online purchasing behavior. For that particular purpose, several data sources were utilized: we extracted Twitter data through academic Twitter API to understand the trends of certain types of posts via social media platforms, Google search trend data to see how people reacted to the two pieces of information mentioned above and if they researched further, and the purchasing data, focusing specifically on changes in fish oil purchase trends. The research process was three-fold. First, we conducted our data collection. At this stage, we collected all the data, explained in the section below. Secondly, we conducted data inspection and data analysis. In this phase, Twitter data, Google Trends and purchasing data were inspected as time series, and structural break analysis and correlation analysis were applied to all three data sources. Then, we applied the Granger Causality Test to determine whether there was evidence that social media posts time series had caused a change in the purchasing one. Finally, the results obtained were compared and discussed. Figure 5.1 shows the research process.

5.3.1 Data Collection

We divided the data collection into four phases: i) Twitter data collection, ii) Google Trends data collection, and iii) purchasing data collection.

Twitter Data Collection

In the first phase, we looked for how information spread on social media platforms, focusing on Twitter data. We used the academic Twitter API to extract Twitter posts both in English and Icelandic from the period of 01.01.2019 to 31.12.2020 where we followed the terms and tags with both capital and lower-case initial letters: Fish oil covid; Fishoil covid; Lýsi; #fishoil; #fishoil #covid; #Lýsi. For clarification, lýsi is the Icelandic word for fish oil.

We extracted the data in weekly granularity for two reasons. Firstly, because there were several zeros in the daily granularity; secondly, because the data provided by Google Trends also had weekly granularity and because of that, a comparison between the graphs would yield a more accurate view.

We found 6768 Twitter posts in total from users all around the world. Before analyzing the data, we added up the number of all the Twitter posts extracted and analyzed them as a unique time series. Figure 5.2 shows the number of Twitter posts over time.

Figure 5.3 shows a few examples of the posts found. It is interesting to notice that there are some verified accounts (the

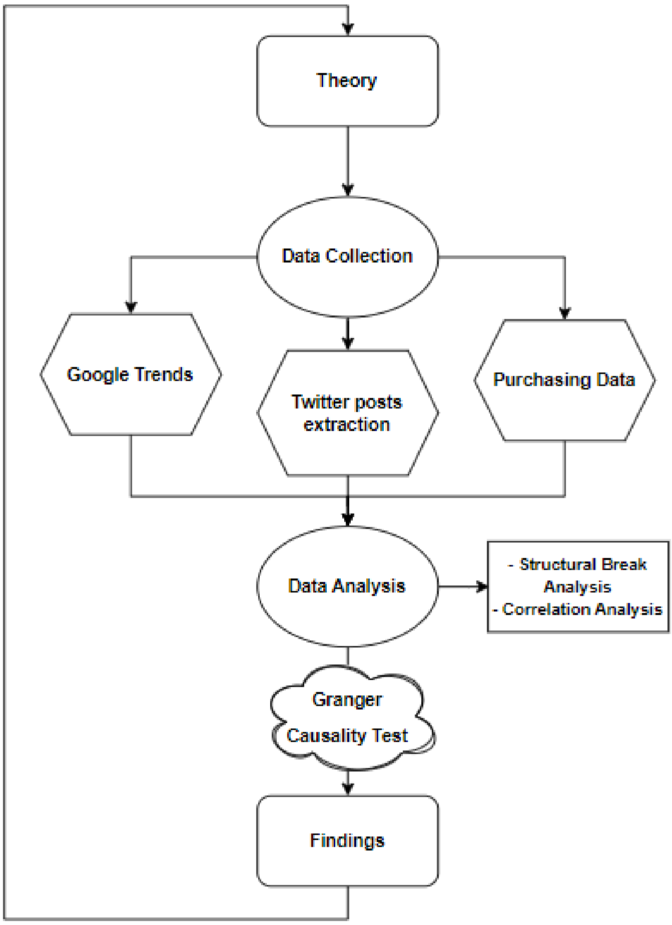


Figure 5.1: Research Process.

one with the blue badge). This means that the account that published that post is an account of public interest, with a lot of followers and hypothetically a lot of influence.

Google Trends Data Collection

In the second phase, we looked for the potential impact that news and posts on fish oil benefits had on internet searches, by ana-

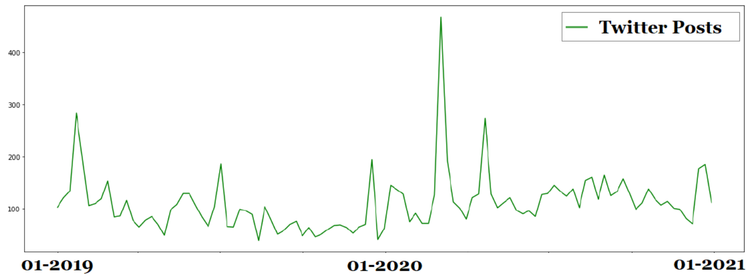


Figure 5.2: Twitter posts over time.

lyzing Google searches in Iceland between 2019 and 2021 for key terminology related to fish oil. During this phase we analyzed the following words with both capital and lower-case initial letters, in various combinations: Lysi; fish oil; Omega-3 and Cod liver oil.

Then, we combined all the searches and illustrated that in Figure 5.4 where we show the Google Trends day by day, both with initial capital letters and lower-case letters. We found 4589 Google searches in total for the selected period.

Purchasing Data Collection

The purchasing data analyzed in this paper is numerical and represents daily sales of fish oil of all brands sold in Iceland for two years: 2019 and 2020. The data was obtained thanks to our collaboration with the Icelandic second largest supermarket chain and it is from on-site and online purchases. The data included the features of date (in format dd-mm-yyyy) and quantity sold, for a total of 33707 purchases in 732 entries. Figure 5.5 shows the time series data obtained.

Before conducting data analysis and comparison between the time series, we aggregated the purchasing data in weekly granularity.

Table 5.1 illustrates the total data points identified for each data source. Following the adaptation of all three datasets to a weekly granularity, the data points are distributed across a total of 54 entries per dataset.

Figure 5.3: Some of the Twitter posts related to fish oil and COVID-19 during 2020.

Data Source	#data points
Twitter posts	6768
Google Trends	4589
Purchasing data	33707

Table 5.1: Overall data points for each data source.

5.3.2 Structural Break Analysis

The first part of our data analysis focuses on investigating whether there are significant changes in the time series data during the time when fish oil recommendations on social media were promi-

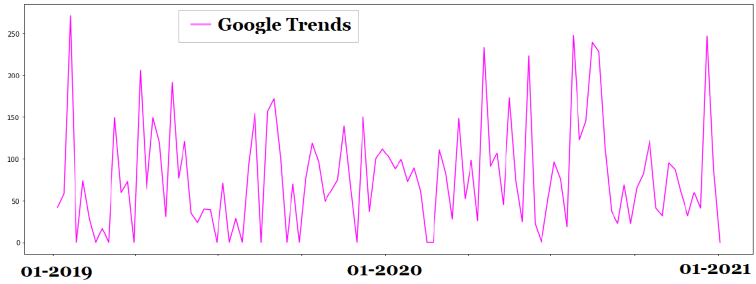


Figure 5.4: Google searches over time.

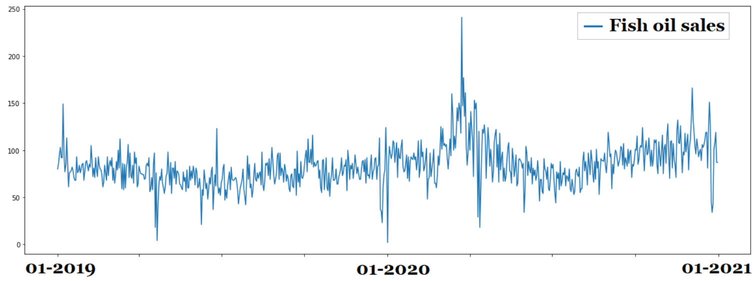


Figure 5.5: Items of fish oil sold daily over time.

nent. To embark on that analysis, we first analyzed the stationarity of the time series. We performed the Augmented Dickey-Fuller test, a statistical significance test widely recognized and frequently used. There is a hypothesis test involved with a null hypothesis and alternate hypothesis, and as a result, a test statistic is computed, and p-values reported. From the test statistic and the p-value, an inference can be made as to whether a given series is stationary or not, as we can read from [192].

A structural break is an analysis type that is used to illustrate an abrupt change in a time series. This change could involve changes in mean or a change in the other parameters of the process that produces the series [193]. Structural break analysis is used to determine when and whether there is a significant change in the data, which is what our research question calls for in this particular paper. In the present study, we looked for structural changes by performing a linear regression model on the following formula:

$$model = LinearRegression(Data \sim \frac{1}{length(Data)}) \qquad (5.1)$$

Then, we evaluated the coefficient significance to ensure that the linear model created was consistent. We calculated the breakpoints and took the number of breakpoints with the minimum Bayesian Information Criterion (BIC), which is a metric commonly used to compare the goodness of fit of different regression models, as assumed by [194].

5.3.3 Correlation Analysis

Before conducting the correlation analysis, a normality test of the data was carried out. In order to put together the three data sources, we transformed the purchase data to a weekly granularity, like the Google Trends and the Twitter data. Then, we conducted a Shapiro-Wilk normality test to see whether we should use parametric or non-parametric tests to calculate the correlation between our data. The Shapiro-Wilk test turns out to be much more accurate than other tests for small datasets [195], so we chose this one because we have a time series with weekly granularity and only 54 entries for each dataset.

The correlation analysis was conducted between the three time series. We used Spearman's correlation, a non-parametric method used when the data do not meet the assumptions of parametric tests, such as when the data are not normally distributed or when there is a non-linear relationship between variables [196].

5.3.4 Granger Causality Test

The Granger causality test is a statistical hypothesis test used to determine whether one time series can be used to predict another time series [197]. It is named after Clive Granger, who was awarded the Nobel Prize in Economic Sciences in 2003 for his work on this concept [198].

The Granger causality test helps to understand whether changes in one variable can be used to predict future changes in another variable. However, it is important to note that the term "causality" here does not imply a true cause-and-effect relationship in the way it is commonly understood in philosophy or physics [199].

Instead, it suggests that one variable provides some information about the future behavior of another variable.

Given two time series X and Y, the Granger test is based on the F-test to decide between two hypotheses:

- Null Hypothesis (H0): The past values of variable X do not Granger cause (predict) variable Y;
- Alternative Hypothesis (H1): The past values of variable X do Granger cause (predict) variable Y.

If the p-value < 0.05 it is possible to reject the null hypothesis (suggesting Granger causality), otherwise or confirm it (suggesting no Granger causality) [200].

5.4 Results

5.4.1 Structural Break Analysis

After the Augmented Dickey-Fuller test, all the time series turned out to be stationary, so no changes were necessary, and we could proceed with the analysis. We computed the linear regression model with Equation 5.1. Then we proceeded with the search for structural changes.

Twitter Posts

The number of structural breaks with the minimum BIC for the Twitter posts time series was equal to 2, so we took two break-points in our model. Each break-point has both the date component and the quantity sold because it is a specific point in the graph. The two breakpoints corresponded to the following dates: 26.02.2020 and 31.03.2020. Figure 5.6 shows the time series of Twitter posts and purchases (aggregated at a weekly level) with the structural breaks of the Twitter posts' time series.

As we can see from this image, the two graphs display the same trend, especially at the beginning of 2020. The structural breaks in the Twitter posts' time series correspond almost entirely to the points where the sales graph also changes. The highest peak in the Twitter time series occurs in the week starting from 08.03.2020, while the purchasing time series' peak is in the 22.03.2020 week.

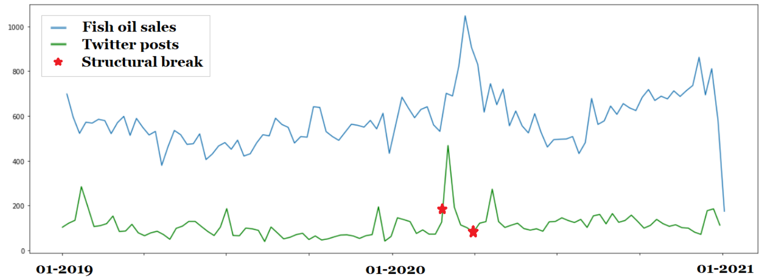


Figure 5.6: Twitter posts time series (green) with the structural breaks (red stars) compared to the purchasing time series (blue).

We notice how the peaks and troughs in the Twitter posts graph are slightly earlier than in the sales graph.

The discrete consistency observed suggests a potential association between individuals' actions and consumption of posted news. In particular, the peak of substantial purchases aligns closely with the peak of Twitter posts. This correlation leads one to explore the plausible influence exerted by social media posts and news published online on individuals' behavior, a phenomenon also recognized by previous research [201].

Google Trends

The Google search yielded no relevant results, as there was no change in the number of searches by Icelandic Google users for the terms indicated in the Data Collection section. Figure 5.4 shows minimal changes between 2019 and 2020 but the trend graph does not display a consistent number of searches over time. To determine whether this increased search volume is correlated with actual customer purchases, we attempted to calculate the structural breaks in the time series. However, the number of structural breaks with the smaller BIC was equal to zero. This result indicates that there are no structural breaks in the Google Trends time series, meaning that there is no significant change in the graph.

Figure 5.7 depicts two graphs that exhibit dissimilar trends. Google searches for fish oil remained low. The disparity between the two graphs may suggest that at the beginning of COVID-19, people increased their purchases of fish oil without conducting any

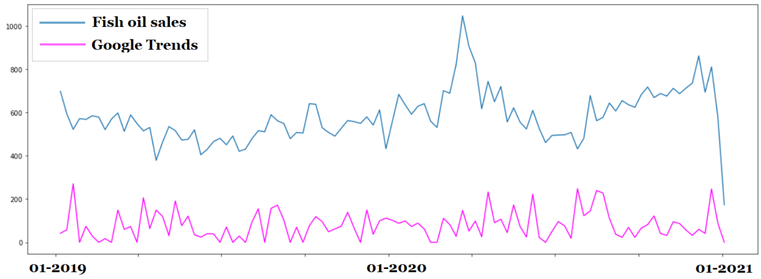


Figure 5.7: fish oil purchasing and Google searches time series.

additional research.

Purchasing Data

For the purchasing data, minimum BIC was obtained with 3 breakpoints, that corresponded to the following dates: 18.04.2019, 11.03.2020, 21.06.2020. The resulting model with the corresponding breakpoints is shown in Figure 5.8 below.

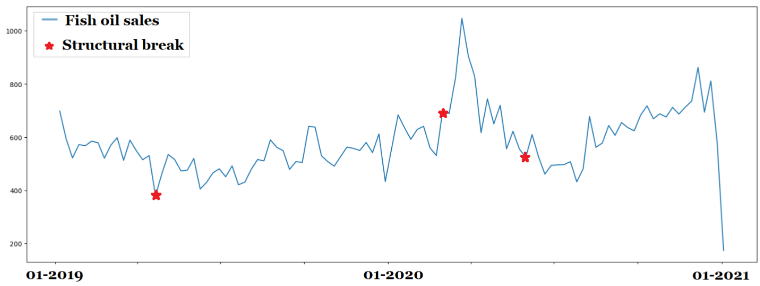


Figure 5.8: Structural breaks of the purchasing data.

In Iceland, fish oil is usually widely used during the winter to protect the immune system, so it is common for the purchasing trend to increase during the autumn and decrease during the summer. However, Figure 5.8 shows that there is an abrupt increase in fish oil purchasing that starts in March 2020. Our analysis shows that this can be supposedly attributed to the spread of posts suggesting the use of fish oil online. We then see that around the summer of 2020, there is a decrease in purchases, which brings us back to the usual pattern of purchases decreasing during the warmer seasons.

5.4.2 Correlation Analysis

The normality test we carried out depicted that our data was not normally distributed. Table 5.2 shows the correlations between the time series.

The Spaerman's correlation value between fish oil purchases and Twitter posts resulted in r = 0.301, which is considered as a medium positive correlation value according to [202]. With the significance level of ($p < 0.01$) for fish oil purchases and the number of Twitter posts published over time, the result is that there is a medium positive correlation between them.

	Purchasing data	Google Trends	Twitter posts
Purchasing data	1	0.199	**0.301***
Google Trends	0.199	1	0.108
Twitter posts	**0.301***	0.108	1

*. Correlation is significant at the 0.01 level (2-tailed).

Table 5.2: Spearman's correlations.

Google searches proved not to be statistically significantly correlated with either of the other time series. This result demonstrates again the commonality between trends on Twitter and in purchasing, and the difference with the trend in searches on Google. This might indicate that people did not research any further information about the benefits of fish oil.

5.4.3 Granger Causality Test

We applied Granger's test by first evaluating the pair (X,Y) = (Twitter posts, Purchasing data) and then (X,Y) = (Google Trends, Purchasing data). Table 5.3 shows that there is a causality between the publication of certain Twitter posts and the purchasing behavior of Icelandic customers. Indeed, the p-value for that pair is less than 0.05, and therefore we can say that there is predictive information: significant Granger causality suggests that past values of the variable X (Twitter posts) contain information that helps predict future values of a variable (Purchasing behavior).

In other words, changes in X may provide insights or improve forecasts for Y [203].

Granger Causality Test	Res.Df	Df	F	Pr(>F)
Twitter Posts - FishOil Purchasing	21	-2	5.8266	**0.01063***
Google Searches - FishOil Purchasing	21	-2	2.6871	0.09383

*. Significant at the 0.01 level

Table 5.3: Granger Causality Test.

It is crucial to emphasize that Granger causality does not establish a true causal relationship in the traditional sense. It merely demonstrates statistical predictability. Other factors or variables not considered in the analysis could be responsible for the observed relationship [204].

However, the many different analyses performed in this study also reinforce the result obtained from this last test, as we can see that in every analysis performed there is always a significant result between the Twitter time series and purchasing one.

5.5 Discussion

Within the scope of shedding light on the way social media impact on people's beliefs and behavior, attention and attention-grabbing has become a crucial factor. Attention has become recognized as one of the essential critical resources for influence. As a large portion of the media today is driven by clicks, there is a wave of clicks that comes with successfully grabbing the readers' attention [205], [206]. Since clicks and attention are valued over correctness, there is an epidemic of false or misleading information ongoing through social media, so there is a need to examine further and contribute with empirical findings [176]. Parallel to the focus on clicks and attention-grabbing, there is ideological polarization. Ideological polarization outlines the strengthening of the existing beliefs of netizens which are reinforced by the repetition of similar information within a closed system, also known as filter bubbles [207], [208] or echo chambers [209]. Although these concepts represent slightly different phenomena, they all resemble each other in terms of their impact. When information enters into reinforced repetition within a closed loop system, the beliefs of the netizens are further reinforced and certain types of news are, by extension, more difficult to break [207], [208].

Furthermore, since governmental policy-making, news agendas, social movements and election campaigns are run in online settings and through social media platforms in combination with trusted news outlets, it has become increasingly vital to study the correctness of the information spread online [206], [210] and to try to limit the influence that certain information can have on the users. Moreover, most of the research out there examines the way social media news spread and its evolution, but there is a gap regarding how the impact of social media information spread in online settings can impact the offline world. In this paper, we contribute to that particular gap with our study of the online world, through a netnography of how one news spread, while also studying behavior in offline settings through our purchasing behavior data derived from the supermarket chain. Through that data, we are able to contribute with novel findings on how online information impacts our offline behavior. With that said, it is now time to examine our methodological approach, to tackle our problem area, which we did here.

Google Trends shows a lack of interest in fish oil searches. Both purchasing and Google data were only collected for Iceland, so we can see how the nation reacted to the news. The analysis of Google Trends showed that people did not bother to look for the truthfulness of news read on social media, but trusted them blindly and started to buy more fish oil. This is one of the most worrying circumstances because it is the time when false information can spread fastest and create negative consequences for people [211].

Another reason why people were highly influenceable during COVID-19 is due to the high levels of uncertainty and anxiety that people experienced [212], [213]. When people are anxious or uncertain, they are more likely to seek out information to help them make sense of what is happening around them [212].

Further cause that contributed to people's susceptibility to certain types of news read on social media was the sheer volume of information that was available during the pandemic. Figure 5.8 shows precisely how the biggest peak is in March 2020, when the pandemic started to spread around the world and there was such a large amount of information every day [212]. The structural break analysis reveals how abrupt changes in the graph of purchases and the graph of Twitter posts almost match. The fact that there is a structural break in the Twitter data and not in the Google Trends is really interesting because it shows how much people believe what they read and do not search further [214]. It is important to see that also the correlation analysis confirms what we interpreted from the structural break analysis because we can see in Table 5.2 that fish oil purchasing is related to Twitter posts but not to Google searches. Moreover, in the Granger causality test, there is a significant result between the Twitter posts and the purchasing time series, but not between the Google searches and the purchasing ones.

These results suggest a strong likelihood that social media significantly influences our lives, potentially leading to a reduced inclination to thoroughly research online news due to implicit trust [215]. In the contemporary landscape, the emergence of influencers appears to play a role in influencing our decisions or purchases, even if we may not be fully conscious of it [215].

5.5.1 Limitations

Like all research, this study contains some limitations. Although the data comes from one of the largest supermarket chains in Iceland, with a large geographical spread, we are referring only to one supermarket chain out of four that we can find in Iceland. Because of that we also do not make claims for generalizing for the world based on these findings from Iceland. Moreover, data comes from only one social media which means that the analysis and the data itself could include biases. To counteract this limitation, we could have analyzed all the data from all the supermarket chains in Iceland and all the most popular social media, but due to complexity, that was not possible. Since our data is derived from a very popular supermarket chain and one social network that is widely used in Iceland, we can assume that the present study is an acceptable approximation of the impact of social media on the Icelandic population during COVID-19.

5.5.2 Future Work

Our study paves the way for further research in multiple directions. Firstly, we intend to broaden our analysis by examining various social media platforms to enhance the scope of our work. Additionally, we aim to delve into other case studies by differentiating between ordinary users and influencers, while employing social network analysis techniques. This will allow us to investigate the dissemination of social media posts promoting specific product types across the network and compare the outcomes with the tangible influence on purchasing behavior. Finally, it would be interesting to get a worldwide perspective by analyzing other markets besides Iceland.

5.6 Conclusions

In this study, we had the unique opportunity to analyze data from one of the biggest supermarket chains in a single country concerning a major event that impacted society as a whole. We investigated the relationship between social media and its impact on behavior considering all of Iceland as a case study. We zoomed in on one particular product, *fish oil*, whose effectiveness against the virus circulated through websites and social media at the start of the pandemic. We took a data-driven approach and collected data from various sources, including purchases of fish oil through a statistical analysis combined with internet searches and visibility on social media platforms through a netnography approach. We answered our research question "*How did social media posts impact the Icelandic population's purchasing behavior during the COVID-19 pandemic?*" by showing that the publication of Twitter posts provided some information about the Icelandic customers' purchasing behavior.

The acquired results suggest a considerable likelihood that customers were influenced by social media platforms, manifesting a perceptible alteration in their offline behavior as evidenced by an increased consumption of fish oil compared to the established pre-pandemic consumption patterns.

In our analysis, we present the presence of numerous breakpoints, which show that especially in March and April 2020, when COVID-19 began to spread throughout Europe and with it the fear of contracting the virus, different claims spread on social media platforms and fish oil sales significantly increased. Usually the purchase of fish oil increases during the fall, so it is surprising to see such a large increase in sales in the spring of 2020, which leads us to assume an influence dictated by social media. Correlation analysis and Granger causality tests continued to confirm the hypothesis we had from the Structural breaks. The effects we have shown are on changing purchasing behavior, but social media platforms may affect the population in many other ways.

Chapter 6

Nudging Towards Sustainability: Exploring Strategies to Encourage Greener Choices Among Adolescent Girls

This paper investigates the impact of digital nudging techniques, specifically traffic light labeling, on the snack choices of teenage girls. Using Choice-Based Conjoint Analysis, we conduct a survey with four variants: one without traffic lights, one with all three traffic lights (red, yellow, green), and two featuring only red or green traffic lights. The study finds that a green traffic light effectively encourages the selection of eco-friendly snacks, aligning with the researchers' hypothesis, while a red traffic light has minimal influence on choices. The study also observes a surprising predisposition towards sustainable snack choices even without traffic lights, highlighting the potential for adolescents to already possess inherent eco-conscious tendencies. The results highlight the importance of simplicity and clarity in nudging mechanisms and the need for further research to explore the long-term effects

of these techniques on sustainable behaviors.

Full reference:

C. Carpinelli, B. Baesens, S. V. Broucke, *et al.*, "Novel insights on time series forecasting comparing deep learning and statistical models to avoid overproduction", *Paper Submitted to Expert Systems With Applications*, 2024

6.1 Introduction

In the face of increasing environmental challenges, the global imperative to encourage sustainable choices has become prominent [216]. While customers' intentions are often geared toward sustainable decisions, their behavior is the opposite [1]. Hence, customers show a clear gap between their commitment to sustainability and purchasing decisions. Bridging this gap is critical to move toward a more eco-conscious future.

Addressing this challenge requires innovative solutions that integrate both social and technical elements, leading to the emergence of behavioral interventions like digital nudging as pivotal tools in guiding individuals toward more sustainable decisions [162], [216]. Digital nudging, which leverages digital platforms and technologies grounded in behavioral economics, gently steers decision-making processes by exploiting the natural human tendency to rely on cognitive shortcuts in complex situations [217]. This approach represents the intersection of social behavior and technical solutions.

Our study, inspired by the urgent need to turn good intentions into actual actions for a greener planet [1], investigates how digital nudging can be strategically applied to influence sustainable snack choices among adolescents, a demographic that significantly shapes future market trends [163]. We focus on the socio-technical dynamics of decision-making, where technical interventions (nudging techniques) interact with the social context (adolescent customers) to shape behavior [218].

Recognizing the disparity between intentions and actions, we explore the application of nudging techniques in digital platforms as a critical tool for encouraging customer choices toward sustainability, and we employ data analysis to assess the efficacy of the techniques used.

To achieve this purpose, we experimented in the context of the Girls in ICT Day event at the University of Reykjavik, Iceland. During the event, over 400 fourteen- and fifteen-year-old adolescents, primarily girls, actively participated in workshops and experiments to reduce the gender gap in technology. The Girls in ICT context presented a distinctive opportunity to investigate and evaluate the potential influence of digital nudging techniques on adolescents. The deliberate choice of this demographic group has intentional and substantive value. As representatives of the future adult customer base, adolescents influence future market trends considerably [166]. Examining how they respond to digital nudging becomes critical to understanding how these techniques can be strategically leveraged to have a lasting and sustainable impact on customer choice [164]. As today's adolescents navigate the complex landscape of consumption choices, influencing their preferences toward sustainability becomes a strategic imperative [32].

The experiment involved presenting the participants with an online Choice-Based Conjoint Analysis survey featuring four variants: the first one was a regular survey, then the next three contained different combinations of traffic lights (red, yellow, green) as nudging techniques, designed to guide their choices towards more sustainable options. Then, we applied different analyses to measure the performance of the designed technique. The subsequent Methods section will explain further details and a comprehensive understanding of the methods employed.

Our guiding hypotheses were as follows:

1. (H1) *The presence of three traffic lights (red, yellow, green) would exhibit a more favorable influence on the choices for products labeled with yellow and green lights compared to the absence of traffic lights.*

2. (H2) *The presence of a red traffic light will negatively influence the choice of that product.*

3. (H3) *The presence of a green traffic light will positively influence the choice of that product.*

We derived our hypotheses from prior research, particularly the work by [219], [220] and [221]. The first study indicated that the traffic light method effectively promotes sustainable choices, with the presence of the green light typically influencing decision-making. The second and the third studies, conducted in a hospital and a university cafeteria setting, respectively, revealed that customers tended to avoid dishes labeled with the red traffic light. However, no studies have yet explored the impact of these nudging techniques on adolescents. Therefore, our study seeks to ascertain whether these findings hold within this demographic, shedding light on their effectiveness in influencing toward more sustainable choices.

Our study takes place against a growing awareness of the importance of responsible nudging. By situating our research within the context of socio-technical systems, we aim to contribute to the growing discourse on how technical innovations, such as digital nudging, can address social challenges like sustainability. Our findings offer valuable insights into the design of sustainable choice architectures for adolescents, thereby advancing the theoretical and practical understanding of responsible nudging within socio-technical frameworks [222].

6.2 Related Work

Nudging, derived from the fields of economics and behavioral psychology, is a way of guiding behavior through design choices [223]. Through nudges, it is attempted to define default options or alter the decision-making environment [101], [224], meaning to bring a desired outcome without restricting options or engaging significant economic, social, or time penalties or incentives [225].

The concept of nudging has garnered substantial attention across various disciplines, including information systems, where it challenges the traditional notion of rational decision-making. Cognitive limitations and biases often lead individuals to make suboptimal choices, which can be mitigated through well-designed nudges. Recognized globally as an effective tool for policy and behavioral change, nudging has been widely adopted to promote pro-environmental and pro-health behaviors [32], [33], [226]–[232]. Governments and organizations use nudges to guide citizen choices while preserving individual autonomy, reflecting a socio-technical approach where technical interventions shape social behaviors.

In the digital age, nudges have evolved into digital nudges, which leverage online platforms and technologies to influence behavior [233], [234]. These digital interventions harness psychological principles to guide user decisions, suggesting that the efficacy of traditional nudges can be extended to online environments, albeit with certain contextual limitations [233]–[236].

Worldwide, seen as an activity where nudging can be highly effective, food consumption can be a means to approach more sustainable choices [230]. Nudging has been deployed to address obesity concerns and environmental impacts, such as reducing meat consumption [237] and minimizing food waste [238]. However, sustainability is approached differently depending on the gender of the customer, as we see women are more likely to choose a more sustainable option. At the same time, men do not seem to be persuaded by labeling the same way [239]. When shopping for food, people have been found to focus primarily on factors such as price and health [217], but these factors can be prone to nudging depending on the context set. Nudging in the context of food consumption and environmental awareness has been approached in multiple studies using various methodologies. However, research on interventions with digital nudges is still limited and presents inconsistencies that may affect replicability [240]. These inconsistencies refer to but are not limited to describing nudging strategies used and differentiating between digital nudging and nudging in digital spaces. Several studies show potential in using digital nudging for sustainable food choices. Still, the methodology is essential, and the efficacy of the intervention varies accordingly, being affected by factors such as technostress or nudging employed [241]–[243].

Regarding methodologies, in some studies, researchers explore nudging by introducing a third option, also called the compromise effect. This option serves as a middle ground between two existing choices, often regarding attributes like calorie content or product size. People tend to prefer this middle option when making decisions, especially when the original choices seem too extreme or polarized [244], [245]. While numerically informing customers of the calorie content or other nutritional factors of their purchase did not show a significant change in their choices [246]–[248], traffic-light labeling has shown promising results in nudging people towards healthier or more sustainable choices [220], [249], and choices with lower carbon emissions [96], most likely due to the simplicity and straight-forwardness of the color-based coding, compared to numerical values only [250], [251]. However, labeling related to carbon emissions brings mixed results; some studies show the traffic-light system to work in these cases too **rw30**, while others show that although customers prefer the traffic light labeling over others, the carbon emissions of a product do not have a significant effect on their purchasing choices [252]. Hence, traffic lights' effectiveness can vary based on cultural, contextual, and individual factors [253]. Understanding how this technique influences decision-making in the specific context of sustainable choices remains a subject of ongoing exploration in behavioral science and environmental psychology [254]. As researchers and practitioners continue to refine and adapt nudging techniques, the traffic lights metaphor is a compelling tool in the collective effort to encourage more sustainable behaviors.

Despite the extensive literature on nudging and its efficacy in guiding behavior, there is a noticeable gap in exploring its impact in digital environments and adolescents, particularly with a focus on the gendered aspect. A study focused on Indian youth and employing physical nudging showed them attaching increased importance to labels, with a focus on price and health [255]. However, the study defines youth as the population between 18 and 30 years old. Furthermore, while studies highlight gender differences in sustainability choices, with women being more responsive to eco-friendliness [256], even in their younger years [166], there is a lack of research specifically addressing the nuanced dynamics influencing teenagers through labeling and digital nudging. Understanding the influence of digital nudging strategies, such as traffic-light labeling, in shaping teenagers' choices regarding food consumption and environmental awareness is crucial for developing targeted interventions and policies to foster sustainable behaviors in this demographic.

In conclusion, nudging applications span various domains, including pro-environmental and pro-health initiatives, with governments and organizations worldwide adopting nudges to guide choices without restricting freedom. While studies have demonstrated the effectiveness of nudging in promoting sustainable food consumption, mainly through methods like traffic-light labeling, there remains a notable gap in understanding its impact in digital environments, especially among adolescents, with limited research addressing gender dynamics. While existing studies offer insights into the preferences of youth regarding labeling, the focus on older age groups and the lack of nuanced exploration into gendered influences among teenagers underscore the need for further research. Investigating the socio-technical interactions that shape adolescents' responses to digital nudging strategies, such as traffic-light labeling, will contribute to developing targeted interventions to foster sustainable behaviors in this demographic.

6.3 Methods

In this section, we outline the methodology employed to investigate the effectiveness of nudging techniques in promoting sustainable choices among adolescent girls through digital platforms. We describe the use of choice-based conjoint analysis, the implementation of traffic lights as a nudging technique, the survey design, and the sensitivity analysis conducted to assess survey reliability.

6.3.1 Experimental Setup

The experiment was conducted at Reykjavik University, Iceland, in May 2023, during the annual Girls in ICT Day. This event is specifically designed for high school girls and non-binary teenagers aged 14-15, offering them the opportunity to visit the university's technology departments, including computer science and engineering. The departments organize various workshops and activities with the overarching goal to reduce the gender gap in technology fields in both academia and industry. As part of the 2023 Girls in ICT Day, we divided the participants into four groups of 100 teenagers each. Each group was asked to complete a different survey version designed following the principles of Choice-Based Conjoint Analysis. These survey variants were crafted to explore how different digital nudging techniques, such as traffic-light labeling, influence snack choices—a scenario that blends social behavior with technical intervention in a digital context.

Of the over 400 girls and non-binary teenagers who received the survey variants, 362 responded. Since the percentage of non-binary teenagers among the respondents was very small (around 1%), we decided to exclude this group from the analyses as there were too few to make statistically based conclusions. This decision was made to ensure the reliability and relevance of the socio-technical analysis, as the primary aim was to understand the intersection of digital nudging and gendered decision-making in a controlled environment that reflects real-world social dynamics.

6.3.2 Choice-Based Conjoint Analysis

Choice-Based Conjoint Analysis (CBCA) is a powerful research technique to understand how individuals choose different product or service offerings [257]. Unlike traditional conjoint analysis, where respondents rank or rate product profiles, CBCA presents respondents with choice scenarios, and they must choose their preferred option from a set of alternatives [258]. In this specific study, in which we used snacks as products, CBCA works as follows:

1. Attributes and levels definition: identify and specify attributes, such as Nutrition Content, Vegan, Local Ingredients, etc., along with their respective levels (e.g., low, medium, high environmentally friendly features).

2. Profile creation: combine the attribute levels to create diverse product profiles representing various combinations of snack features.

3. Choice sets formation: develop choice sets comprising two or more product profiles presented to participants for preference selection.

4. Data collection: gather responses from participants as they make choices, providing insights into their preferences for specific snack attributes and combinations.

5. Analysis: utilize statistical models to analyze collected data. By estimating part-worth utilities through models like Ordinal Logistic Regression, we ascertain the relative importance of each attribute level. In our experiment, data was transformed into binary values, and analysis provided valuable insights into the features most influential in shaping teenagers' snack preferences.

CBCA, in this context, enabled us to go beyond mere preferences and understand the trade-offs participants are willing to make in their snack choices. It provided a realistic and actionable framework for measuring the effectiveness of the nudging technique and uncovering the key drivers of decision-making in the specific domain of environmentally friendly snack consumption among teenage girls.

6.3.3 Traffic Lights Method

The traffic lights nudging technique is a well-known method in behavioral economics that uses simple visual cues like traffic signals to influence decision-making [221]. It borrows the red, yellow, and green colors to tap into people's associations with these signals in the context of traffic, guiding them in different situations. In eco-friendly choices, the traffic lights nudging technique has been employed to guide people towards environmentally sound options [96].

Our experiment used them on products to show how they impact the environment. Green means the product is eco-friendly, yellow suggests a moderate impact, and red indicates a less eco-friendly option. This color-coded system could enable teenagers to make informed decision-making, potentially better matching their values [96].

6.3.4 The Survey

The survey was designed following the CBCA methodology, showcasing three distinct snacks, each characterized by eight features reflecting key aspects such as Nutrition Content, Veganism, Local Ingredients, Additives, Packaging, Flavor, Price, and Taste.

The snacks were deliberately designed to represent varying levels of environmental friendliness, which is in line with CBCA. To assess the effectiveness of digital nudging, each group received a different QR code leading to a different survey variant, including those without traffic lights and those incorporating different combinations of red, yellow, and green lights to influence choices (see Figure 6.1). The participants were asked to select one product. At the end of the survey, they were asked, "*Given what you normally buy, would you buy the product you chose?*". We then applied sensitivity analysis to assess the robustness of our results by examining how variations in the responses to the reliability question affect the overall interpretation of the survey data.

Responses from the participants were meticulously collected, resulting in four distinct datasets. Integrating traffic light indicators as features in the datasets facilitated a nuanced understanding of the impact of nudging on choices. After Sensitivity Analysis and some Exploratory Data Analysis (EDA), Conjoint Analysis was employed to unravel the features that predominantly influenced the participants' snack choices, shedding light on the efficacy of digital nudging in the context of sustainable decision-making and the usefulness of Data Analysis in measuring digital nudging performances.

Considering the three snacks with the following characteristics, which product would you buy?

	SNACK 1	SNACK 2	SNACK 3
NUTRITION CONTENT	Low nutrition content	Medium nutrition content	Optimal nutrition content
VEGAN	No	No	Yes
MADE OF LOCAL INGREDIENTS	No	No	Yes
CONTAINS ADDITIVES	Yes	Yes	No
PACKAGING	Plastic	Paper	Recycled paper
FLAVOUR	Sweet	Salty	Salty
PRICE	4$	4$	6$
TASTE	Good taste	Good taste	Very good taste

Snack 1

Snack 2

Snack 3

Variant 1: No Traffic Lights

Given what you normally buy, would you really buy the product you chose?

Yes

No

The Survey

	HIGH ENVIRONMENTAL IMPACT SNACK 1	MEDIUM ENVIRONMENTAL IMPACT SNACK 2	LOW ENVIRONMENTAL IMPACT SNACK 3
NUTRITION CONTENT	Low nutrition content	Medium nutrition content	Optimal nutrition content
VEGAN	No	No	Yes
MADE OF LOCAL INGREDIENTS	No	No	Yes
CONTAINS ADDITIVES	Yes	Yes	No
PACKAGING	Plastic	Paper	Recycled paper
FLAVOUR	Sweet	Salty	Salty
PRICE	4$	4$	6$
TASTE	Good taste	Good taste	Very good taste

Variant 2: Three Traffic Lights

	HIGH ENVIRONMENTAL IMPACT SNACK 1	SNACK 2	SNACK 3
NUTRITION CONTENT	Low nutrition content	Medium nutrition content	Optimal nutrition content
VEGAN	No	No	Yes
MADE OF LOCAL INGREDIENTS	No	No	Yes
CONTAINS ADDITIVES	Yes	Yes	No
PACKAGING	Plastic	Paper	Recycled paper
FLAVOUR	Sweet	Salty	Salty
PRICE	4$	4$	6$
TASTE	Good taste	Good taste	Very good taste

Variant 3: Red Traffic Light

	SNACK 1	SNACK 2	LOW ENVIRONMENTAL IMPACT SNACK 3
NUTRITION CONTENT	Low nutrition content	Medium nutrition content	Optimal nutrition content
VEGAN	No	No	Yes
MADE OF LOCAL INGREDIENTS	No	No	Yes
CONTAINS ADDITIVES	Yes	Yes	No
PACKAGING	Plastic	Paper	Recycled paper
FLAVOUR	Sweet	Salty	Salty
PRICE	4$	4$	6$
TASTE	Good taste	Good taste	Very good taste

Variant 4: Green Traffic Light

Figure 6.1: The four different variants of the survey used for the experiment showing the setup of the CBCA and the use of traffic lights for snack presentation.

6.3.5 Sensitivity Analysis

Sensitivity analysis examines how specific parameters or criteria changes affect the results [259]. In assessing a survey's reliability, applying a sensitivity analysis to understand how varying thresholds for a "reliable" response impact certain conclusions is possible.

We started by establishing different thresholds or criteria for what constitutes a "reliable" response to the question about purchase intentions: "*Given what you normally buy, would you buy the product you chose?*" in the four different survey variants. The choice of a threshold for reliability depends on various factors, including the study context, the specific question being asked, and the desired confidence level in the survey responses [260]. In survey methodology, researchers often aim for a reliability level of at least 60% to 80% to ensure confidence in the data [261]. This range allows for a balance between the desired level of reliability and practical considerations such as sample size and survey constraints. We chose a medium threshold, set at a minimum of 60%. Our primary objective is to find out their real purchase intentions, and if the reliability of the survey variant falls below this threshold, the validity of our results and the usefulness of the traffic light as a stimulus tool decreases.

6.4 Results

In this section, we present the results of our experiment. While our hypotheses were based on the conventional concept that a red light would discourage choices and a green light would encourage ecological selections, our results unveiled a narrative that did not confirm all these hypotheses.

6.4.1 Sensitivity Analysis

We first conducted the sensitivity analysis on the survey question: "*Given what you normally buy, would you really choose the product you selected?*". Figure 6.2 shows the percentage of "yes" and "no" given for each variant of the survey, and Table 6.1 shows the result of the sensitivity analysis.

As we can see in the table, the "Threshold Level" column represents the different threshold levels ranging from 50% to 80%.

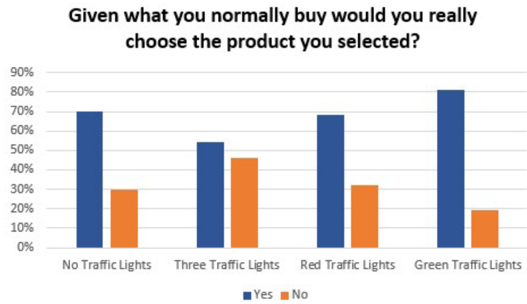


Figure 6.2: Answers to the last question of the surveys.

Table 6.1: Sensitivity Analysis.

Threshold Level	**Proportion of Reliable Survey Variants**
50%	1.0
60%	0.75
70%	0.25
80%	0.25

The "Proportion of Reliable Survey Variants" column indicates the percentage of survey variants that meet or exceed the respective threshold level. Higher values in the "Proportion of Reliable Survey Variants" column suggest a greater proportion of survey variants considered reliable based on the chosen threshold.

For instance, at a threshold of 70%, the corresponding proportion of reliable survey variants is 0.25, indicating that 25% of the variants (1 out of 4) had at least 70% "yes" responses and are thus considered reliable according to this criterion.

We observed that at the 60% threshold we set, only 75% of the variants met the criterion, equating to 3 out of 4. As illustrated in Figure 2, the survey variant that failed to meet this threshold featured the three traffic lights, prompting our decision to exclude it from further analysis. For this reason, we could not confirm hypothesis H1: "*The presence of three traffic lights (red, yellow, green) would exhibit a more favorable influence on the choices for products labeled with yellow and green lights compared to the* *absence of traffic lights.*".

This suggests that the presence of three traffic lights may not have as strong an influence on choices as initially thought.

6.4.2 Exploratory Data Analysis

The next aspect we analyzed involved examining the percentage distribution of choices for each snack within each survey variant. Figure 6.3 illustrates the three survey variants considered for analysis and their respective choice distributions.

As we can see, the variant featuring only the red traffic light indicates a preference of 35.8% for both Snack 1 and Snack 3, with a lower 28% preference for Snack 2. In contrast, the one with only the green traffic light displays a notable 54.3% preference for Snack 3, accompanied by preferences of 25.4% for Snack 2 and 20.1% for Snack 1. Surprisingly, even the variant without traffic lights yields promising results, with preferences of 52.5% for Snack 3, 27.1% for Snack 2, and 20.3% for Snack 1.

The results obtained from the survey depicting the red traffic light are surprising because it is contrary to what is expected from the literature that informed our second hypothesis [220], [221]. On the other hand, the choices made in the survey depicting the green traffic light could be in line with the study that led to the formulation of the third hypothesis [219].

The subsequent subsection will explore a detailed discussion of these results, coupled with the findings from the choice-based conjoint analysis.

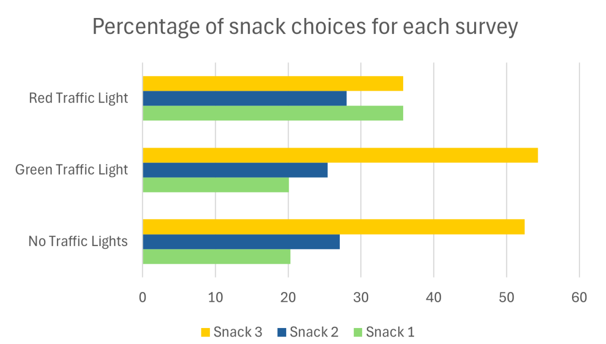


Figure 6.3: Percentage of snack choices in each survey.

6.4.3 Choice-Based Conjoint Analysis

Across all survey variants, packaging features, nutrition content, and flavor emerged as consistently influential in shaping choices. Regardless of the specific nudging technique, these universal influencers held significant sway over participants' preferences. Figure 6.4 shows the relative importance of every feature for each survey variant.

We can notice in Figure 6.4 that the conjoint analysis underscored a surprising finding in the survey variant featuring only the red traffic light. Among the nine features considered, the red traffic light held a meager 0.8% importance. This marginal influence, coupled with the unexpected choice patterns, led us to reconsider the impact of the red light. Figure 6.3 shows that participants did not avoid the less eco-friendly option marked by the red light, challenging assumptions and prompting a rejection of hypothesis H2: *"The presence of a red traffic light will negatively influence the choice of that product"* and suggesting a nuanced response to this particular nudging cue.

Conversely, the survey with only the green traffic light showcased a more positive outcome. While the traffic light's importance in the conjoint analysis was moderate at 8.90%, it could have contributed to steering choices toward the sustainable option.

The reliability of this survey, marked by an impressive 81% affirmative response to the concluding question (see Figure 6.2), confirmed the potential success of the green light in promoting eco-friendly choices. This aligns with our hypothesis H3: *"The presence of a green traffic light will positively influence the choice of that product.*

Surprisingly, the survey variant without any traffic light interventions also yielded a high proportion of sustainable choices, with a reliability score of 70%. Despite its lower reliability than the green light survey shown in Figure 6.2, this finding suggests that participants leaned towards more environmentally friendly options in the absence of explicit nudging.

The comprehensive analysis conducted in this study accentuates the complex nature of digital nudging interventions when applied to adolescents' decision-making processes. In particular, the results revealed a varied interaction between different uses of the selected nudging technique, such as the difference between

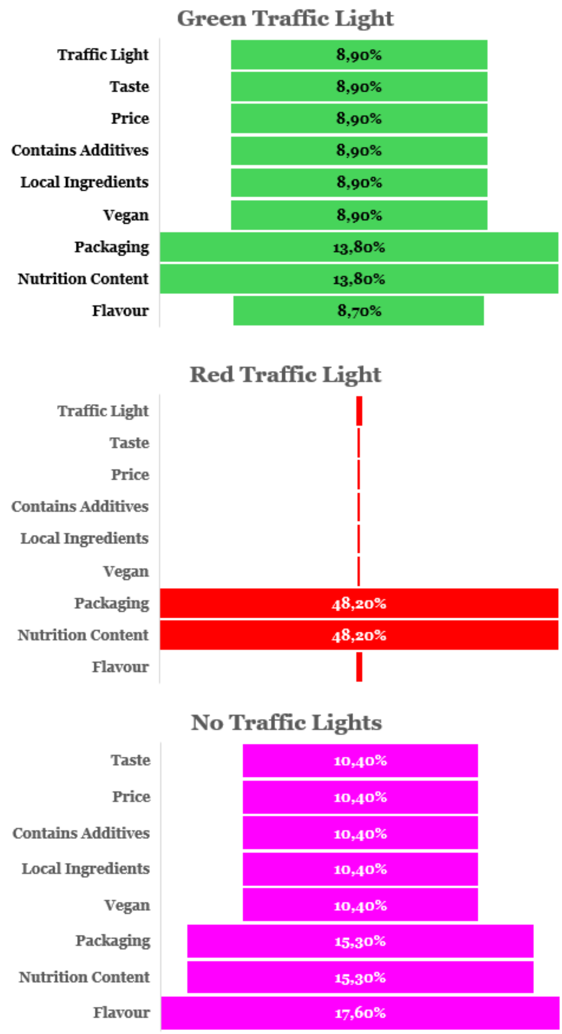


Figure 6.4: Relative importance of features in the different survey variants.

red and green traffic lights and the preferences expressed towards sustainable choices when not subjected to digital nudging. These findings suggest that while some nudges may evoke expected responses, others may produce unexpected or counterintuitive results [262]. This complexity emphasizes the importance of considering various factors, including individual characteristics and contextual influences, when designing effective nudging strategies suitable for adolescent populations.

6.5 Discussion

The findings of this study underscore the significant role adolescents play in shaping sustainable behaviors, a critical consideration for both current and future socio-technical systems [263]. As the future customer base and next generation of decision-makers, adolescents will inherit and address the environmental challenges we face today.

Encouraging sustainable choices among adolescents is thus paramount for fostering a more environmentally conscious society in the long run [264]. Using digital nudging and IT solutions holds particular promise in influencing adolescent behavior [230], [265], [266]. Given the pervasive presence of digital technology in adolescents' lives, leveraging these platforms for nudging interventions can have a profound impact. Digital nudging techniques can seamlessly integrate into adolescents' digital experiences, subtly guiding their choices towards more sustainable options without imposing undue restrictions [267].

Furthermore, applying IT solutions allows for future work on personalized and targeted nudging strategies tailored to adolescents' preferences and decision-making processes [268]. By harnessing data analytics algorithms, IT solutions can analyze vast amounts of data to discern patterns in adolescent behavior and deliver tailored nudges effectively [269]. This approach enhances the efficacy of nudging interventions and ensures that they resonate with adolescents on a personal level. Integrating digital nudging and IT solutions presents a compelling opportunity to empower adolescents to make sustainable choices [270]. By harnessing the power of technology, we can engage adolescents in environmentally friendly behaviors in a seamless, engaging, and impactful way. This not only enhances the efficacy of nudging interventions but also ensures that these interventions are aligned with the values and habits of the target demographic, reinforcing the socio-technical nature of these interactions.

In investigating the impact of digital nudging techniques on the decision-making of adolescent girls, our study has uncovered fascinating insights into the efficacy of strategies to foster environmentally friendly choices within this demographic. By exploring the use of red and green traffic lights, we gained a valuable understanding of the nuanced responses of teenagers.

Contrary to expectations, the red traffic light showed minimal influence on snack choices. The findings revealed that despite expectations, adolescent girls did not demonstrate avoidance of snacks labeled with the red traffic light, leading us to deny hypothesis H2. This unexpected behavior suggests that factors beyond the traffic light cue, such as curiosity or a lack of awareness, may have influenced their snack choices [271]. However, the green traffic light emerged as a potentially powerful driver, guiding choices towards more sustainable options, confirming our hypothesis H3. This success underscores the importance of tailoring digital nudging techniques to align with the values and preferences of the target population [272].

Interestingly, teenagers were predisposed to sustainable choices even without implementing traffic lights. The survey variant without traffic lights revealed a notable preference for environmentally friendly snacks, highlighting the inherent motivation of this demographic toward eco-conscious decisions [273]. This observation aligns with broader research suggesting that women generally tend to exhibit a higher propensity for eco-conscious purchasing [1]. The fact that such tendencies are already apparent during adolescence raises important questions about the necessity of nudging techniques in this context. It suggests that the inclination towards sustainable behaviors may be an existing trait among teenagers rather than one that needs to be artificially induced. This prompts a deeper exploration into whether digital nudging strategies enhance these pre-existing inclinations or whether they may be redundant in contexts where such eco-conscious behaviors are already prevalent. Understanding this dynamic is crucial for determining the most effective approaches to fostering sustainable choices among adolescents.

Unfortunately, the survey variant with three traffic lights yielded unexpected results, as only 54% of the participants said they would really pick the snack they chose based on what they usually like. This response rate fell below the predetermined threshold for reliability, leading us to exclude this survey from further analysis and denying hypothesis H1. Despite its exclusion, this result underscores the importance of maintaining robust standards in data analysis and highlights the potential impact of survey complexity on participant responses [260]. This outcome prompted us to consider whether simpler signals, such as a single traffic light, might be more effective for adolescents.

Drawing from the literature on behavioral economics, particularly studies by [243] and [223], established frameworks guided our exploration into digital nudging dynamics. By examining how adolescents respond to different traffic light signals in their choices, our study delves into how users interact with and are influenced by technology. This user-technology interaction is a key element of socio-technical systems [274]. Our findings align with the evolving understanding of digital nudging effectiveness in specific contexts, emphasizing the need for tailored strategies **dis15**.

6.5.1 Limitations

Our study focused on a specific demographic—girls adolescents aged 14 and 15. While this homogeneity offers depth and specificity, it also poses significant limitations regarding the generalizability of our findings to a broader teenage population. The exclusion of boys and older adolescents may restrict our ability to capture diverse responses influenced by age, gender, culture, and socio-economic status [275].

Although we aimed to address a gap in research by focusing on this particular group, as existing literature highlights gender differences in sustainability preferences, the narrow demographic focus limits the broader applicability of our findings. Additionally, while our study provides insights into how digital nudging affects this demographic, it does not account for other potential variables, such as cultural background or socio-economic status, which may also influence behavior.

6.6 Conclusions

Our study bridges the gap between social behavior and technical interventions, making it a socio-technical investigation into how digital nudging can foster sustainable choices among adolescents.

Through the examination of various nudging approaches, including the use of traffic light signals, we uncovered nuanced responses within this demographic.

While the survey featuring three traffic lights proved ineffective, the success of more straightforward signals, such as the green traffic light, underscores the importance of tailored nudging strategies for adolescents [272]. However, the ineffective response to the red traffic light underscores the need for further exploration and refinement in digital nudging approaches.

Our findings emphasize the pivotal role of digital technologies in promoting sustainable behaviors among the youth [166]. By harnessing data analysis, IT solutions can deliver personalized nudges that resonate with adolescents, fostering a generation of environmentally conscious customers [268]. Moving forward, interdisciplinary collaborations and continued exploration into digital nudging dynamics are essential for leveraging technology to address sustainability challenges and shape a more eco-conscious future [276].

While our study offers valuable insights into digital nudging techniques among adolescent girls, it also highlights the need for continued exploration into the broader landscape of behavioral interventions. As we strive to promote sustainable decision-making, effectively nudging diverse populations across digital platforms becomes increasingly crucial for researchers and policymakers [277].

6.6.1 Future Work

To enhance the applicability and generalizability of our findings, future research should include a more diverse participant pool encompassing different age groups, genders, and socio-economic backgrounds. This broader approach would provide a more comprehensive understanding of how digital nudging influences snack choices across various segments of the adolescent population, thereby offering more widely applicable insights. Additionally, expanding the scope to include cultural differences could yield valuable information on how digital nudging strategies might need to be adapted for different contexts.

Further studies could also explore the long-term effects of digital nudging on sustainable behavior, particularly how these interventions might influence habits and decision-making as adolescents age. Finally, interdisciplinary collaborations incorporating insights from behavioral economics, psychology, and information systems could further refine digital nudging strategies to be more effective across diverse populations.

Chapter 7

Novel Insights on Time Series Forecasting Comparing Deep Learning and Statistical Models to Avoid Overproduction

This paper examines the performance of seven time series forecasting models in predicting purchasing demand for a supermarket chain with three different versions of data aggregations (daily, weekly, monthly). The models are tested under three scenarios: (i) the same prediction horizon, (ii) the same number of prediction steps, and (iii) data with missing values. We employ a novel weighted Mean Absolute Scaled Error (wMASE) metric to penalize overestimation errors, which are especially detrimental in the food production industry. The study finds that AutoARIMA and TimeGPT consistently outperform other models across all scenarios, demonstrating their robustness and adaptability. However, the choice of model depends on factors like computational resources and the need for interpretability. This research offers valuable insights for food supply chain management by highlight-

ing the strengths and weaknesses of different forecasting models, enabling businesses to select the best approach for their specific needs and minimizing food waste and overproduction.

Full reference:

C. Carpinelli, B. Baesens, S. V. Broucke, *et al.*, "Novel insights on time series forecasting comparing deep learning and statistical models to avoid overproduction", *Paper Submitted to Expert Systems With Applications*, 2024

7.1 Introduction

For decades, supermarkets have fostered a perception of limitless selection, implying that customers can always find what they need [278]. The endless selection and demand for availability leads to overproduction and a higher risk of waste. This phenomenon has become the norm, often leading society and production companies to neglect the hidden environmental cost of overproduction and subsequent waste [279]. Fueled by the illusion of a constant consumption need, our purchasing decisions propagate through the entire food supply chain [1], [278]. One of the drivers of food waste is overproduction, caused by two factors: (i) customers' tendency to buy in excess and (ii) supermarkets' pressure to offer ever-expanding selections. This creates a vicious circle - the more that is produced, the more is bought, and the more we produce, leading to more waste [280].

Food waste is both an ethical concern and a significant environmental issue. The resources used or consumed during the production process, packaging, and transporting of unconsumed products contribute significantly to climate change and resource depletion [281]. Overproduction perpetuates a cycle of environmental damage [5]. One way to address this challenge is to enhance decision support through demand forecasting in order to craft strategic plans to reduce excess production [282]. To reduce food waste, local implementations of decision-support for forecasting purposes need to be explored [283].

We investigate the potential of using time series forecasting models to enhance inventory management and minimize food waste for supermarkets. We contribute to the literature by applying various data aggregations using a unique data set of five years of historical sales data from Iceland's largest supermarket chain. While previous research often focuses on single-company data and rarely on food purchasing data [284]–[286], our analysis provides insights into the dynamics of the supermarket sector, enabling the development of more generalizable forecasting models. A notable gap remains in the availability of comprehensive, nationwide data sets and how to use food purchasing data for sustainability purposes, trying to provide support systems to decrease overproduction and waste. The unique data set we analyze in our paper allows us to explore the dynamics of the supermarket sector beyond the confines of individual companies.

Our approach is the following: first, we aggregate the data into three granularities: daily, weekly, and monthly. Then, we evaluate the forecasting performances of seven modern and traditional forecasting models in three case studies: (i) using the same forecasting horizons—one, two, and three months—and (ii) the same forecast steps—four, eight, and 12, (iii) simulating scenarios with varying percentages of missing data to test the robustness of the models under incomplete data sets.

We use the Mean Absolute Scaled Error (MASE) metric for model evaluation. We choose MASE because it provides a robust, scale-independent measure of forecasting accuracy [287]. Unlike traditional metrics, MASE can be applied consistently across different data sets and scales [288]. It also compares the forecast accuracy to a simple baseline model, making it particularly effective in identifying whether it improves prediction accuracy relative to historical averages [288]. MASE also handles outliers better and is less affected by data variance, making it suitable for evaluating time series models in real-world, fluctuating environments [287]. To better assess model performance, we introduce a novel MASE metric (wMASE) that penalizes overestimation errors more heavily [289], thereby aligning the evaluation with the goal of helping the food production industry and supermarkets with decision support systems for reducing overproduction and waste [290], [291]. Moreover, we performed statistical tests and post-hoc analyses to identify significant differences in model performance across data aggregations.

This study answers the following research questions (RQ): *Which forecasting models are most effective in predicting purchasing demand, helping minimize overproduction and waste in* *the supermarket sector, considering different prediction horizons, levels of granularity, and missing data?* By demonstrating the potential of forecasting models to predict purchasing demand, we contribute to developing sustainable food supply chain strategies.

In Section 7.3, we provide a detailed description of the selected forecasting models and data aggregation methods. This is followed by the subsequent Section 7.4 and 7.5, in which we examine the complex interplay between these factors, ultimately addressing the research question.

7.2 Related Work

7.2.1 The Overproduction Issue in Environmental Terms

Overproduction, a complex issue with far-reaching implications, occurs when production exceeds customer demand, leading to overstocking and waste of resources [292]. This economic inefficiency results in a significant environmental burden, depleting natural resources and exacerbating ecosystem pressures **Talia2019customer**, [293].

The environmental consequences of overproduction are profound and manifold. Increased energy consumption, amplified greenhouse gas emissions, and the generation of large amounts of waste throughout the production and distribution chain contribute to environmental degradation [1], [294]. Every stage of the food supply chain leaves an ecological footprint, from the extraction of raw materials to the disposal of finished products [294]. A key factor intensifying this issue is the bullwhip effect, where small fluctuations in customer demand grow into exaggerated production responses further up the food supply chain [295]. This effect magnifies inefficiencies, leading to increased waste at each stage of production and distribution [295].

The food production industry is particularly susceptible to overproduction, with oversupply often leading to food waste and spoilage [296]. A complex interaction between market volatility, agricultural subsidies, food supply chain inefficiencies, and customer behavior drives the issue [45], [293], [297]. The environmental impact is considerable, as excessive food production depletes vital resources such as land, water, and energy while intensifying challenges such as deforestation, soil erosion, and climate change [45], [297]. The bullwhip effect further exacerbates food waste, causing more significant inefficiencies in agricultural production, leading to excessive food waste throughout the food supply chain [298].

Reducing food waste requires a comprehensive approach, including sustainable agricultural practices, optimized distribution networks, and informed customer choices. Collaboration between farmers, policymakers, businesses, and customers is essential to develop effective strategies to combat overproduction and build a resilient and sustainable food system [299]–[301].

7.2.2 Time Series Forecasting to Decrease Overproduction

Forecasting can be essential in mitigating overproduction by enhancing inventory management and production planning across various sectors, including agriculture, manufacturing, and retail [302]–[304].

By analyzing past sales data, demographics, and time, supermarkets can more accurately predict what customers want and when [305], [306]. This allows supermarkets to stay closer to actual needs, reducing the likelihood of undesirable excesses of soon-to-expire food. Time series forecasting can also aid in price setting. Knowing how demand could ebb and flow, supermarkets can adjust prices strategically [307]. For example, labeling items nearing expiration can encourage buying and avoiding waste. Forecasting can also aid with inventory management [302], [308], [309]. In addition, forecasting can allow supermarkets to identify the strong point for stock levels, ensuring that shelves are neither bare nor overflowing [302]. Targeted promotions become another essential tool in the fight against waste [310]. The advantages of using this technology extend beyond the walls of the supermarket. Forecast data can be shared with upstream suppliers for better production planning. This reduces the production of excess goods that could be wasted at the supermarket level [311].

A crucial element of forecasting is its capacity to predict fluctuations in customer demand, seasonal trends, and market dynamics [312]. Additionally, accurate time series forecasting can help mitigate the bullwhip effect, as demonstrated by [313], [314]. By using reliable forecasts, production and retail companies and supermarkets can smooth out demand signals and make more accurate inventory and production decisions. This reduces exaggerated production responses that often result in increased waste throughout the food supply chain [315].

However, inaccurate forecasts can lead to significant loss and wasted resources through overproduction, environmental harm, or underproduction [302], [316]. To mitigate these risks, evaluating and comparing multiple time series forecasting models to identify the most accurate and reliable approaches for predicting future demand can be beneficial. This comparative analysis is crucial in selecting models that minimize the risks of underproduction or overproduction [317].

Over the years, numerous studies have investigated demand forecasting for sustainability, employing various models from basic statistical methods to advanced machine learning models and neural networks [104], [290], [318]–[326]. For instance, some research suggests that fuzzy time series models excel in forecasting electricity consumption [319], while others argue that statistical models like ARIMA are more suitable for predicting food consumption [318], [325]. Still, other studies have found artificial neural networks to be the most effective in predicting agricultural prices, such as wheat sales in China [320], [326]. Yet, some studies assess that the best model is a basic machine learning model [290], [321]. However, the effectiveness of these models can vary significantly depending on the specific context and data set, making it challenging to generalize their success across different scenarios. Table 7.1 shows an overview of the leading papers found alongside their characteristics, allowing for comparison within the existing literature in the present paper.

The diversity of approaches reflects ongoing research efforts to identify the most suitable forecasting models for different contexts and applications [327]. Benchmarking the effectiveness of various forecasting models can become essential for evaluating prediction models because it provides a standardized framework to assess their accuracy and reliability. By establishing benchmarks, researchers and practitioners can identify the strengths and weaknesses of each model, allowing for informed decisions about which model is best suited for specific contexts [328], [329]. Benchmarking also enables comparing new models with existing ones, promoting innovation and the continuous improvement of forecasting models [330].

Given the variability in data sets and study objectives, benchmarking is particularly important for selecting and/or recommending the most appropriate prediction models for specific problems. Different forecasting models may perform better in various contexts, making benchmarking a crucial tool for identifying the most effective models tailored to the unique characteristics of each data set and research question [330]. This approach supports data-driven decisions, guiding future research and helping the food supply chain choose the most suitable models for their forecasting needs [330], [331].

7.3 Methods

The following subsections describe the data sources, techniques used for cleaning and preparing the data sets, the model selection, and the criteria used for model evaluation.

7.3.1 Experimental Setup

In this paper, we have the unique opportunity to use a comprehensive purchasing data set from Iceland's leading supermarket chain, which remains anonymous and whose data cannot be shared due to the confidentiality agreement outlined in a signed non-disclosure agreement. Our data set contains all items sold every single day in the entire nation between the years of 2018 and 2023.

The data set is initially recorded at the lowest level of granularity, resulting in billions of entries, as each scanned product is logged individually. To make the data more manageable and meaningful for forecasting, we aggregate it into three different granularities, (i) daily, (ii) weekly, and (iii) monthly, and 20 different features: BakingProducts, Biscuits, Bread, Breakfast, CakesAndPastry, Cheese, Coffee, Fish, FrozenProducts, Fruit, HealthProducts, HotFood, MeatTable, Milk, Prepack, ProcessedMeatAndFish, Sauces, Soup, Toppings, Vegetables. This aggregation results in 1,818 data entries for the daily granularity, 261 for the weekly granularity, and 60 for the monthly granularity.

Our methodology is divided into three key phases. In the first phase, we forecast each data set across identical forecasting horizons: one, two, and three months. For daily granularity, we forecast 30, 60, and 90 days; for weekly granularity, we forecast

Table 7.1: Literature overview of different forecasting models chosen as best models for different data sets.

Authors (Year)	Title	Data Set	Performance Metrics	Best Forecasting Model
M. H. L. Lee et al. (2022)	A Comparative Study of Forecasting Electricity Consumption Using Machine Learning Models	Monthly electricity consumption data from seven European countries	RMSE and AFE	Fuzzy Time Series
W. Niu and Z. Feng (2021)	Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management	Historical flow data from Manwan Hydropower	RMSE, MAPE, R, Nash-Sutcliffe E	Extreme learning machine, Gaussian process regression and support vector machine
H. Abbasimehr et al. (2020)	An optimized model using LSTM network for demand forecasting	Furniture company's sales data	Unknown	Multi-layer LSTM Network
F. Rustam et al. (2020)	COVID-19 Future Forecasting Using Supervised Machine Learning Models	COVID-19 confirmed cases	R^2 Score, R^2 Adjusted, RMSE, MAE, MASE	Exponential Smoothing
M. A. Maas and C. Bisognin (2020)	Forecast of the historical series of revenues of the Brazilian food industry using forecasting techniques	Brazilian food industry's revenues	RMSE, MAPE, MAE	Combination of SARIMA and Holt-Winters Multiplicative
Y.T. Chen et al. (2020)	Machine learning with parallel neural networks for analyzing and forecasting electricity demand	Electricity consumption in Australia	RMSE	Parallel Neural Network
B. P. BV and M. Dakshayini (2020)	Computational Performance Analysis of Neural Network and Regression Models in Forecasting the Societal Demand for Agricultural Food Harvests	Food crops sales	Unspecified	Multiple Linear Regression and Artificial Neural Networks
J. Feizabadi (2020)	Machine learning demand forecasting and supply chain performance	Purchasing data from a steel manufacturing company	MAPE	Hybrid method using ARIMAX and Neural Network
Fathimath Zuha Maksood and Geetha Achuthan (2017)	Sustainability in Oman: Energy Consumption Forecating using R	Residential and industrial energy consumption data	RMSE	Trend and Seasonal component model
R. Carbonneau et al. (2008)	Application of machine learning techniques for supply chain demand forecasting	Simulated supply chain and Canadian Foundiers orders	Unknown	Recurrent Neural Network and Support Vector Machine
H. Zou et al. (2007)	An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting	Chinese food grain prices	RMSE	Artificial Neural Network
C. Carpinelli et al. (2024)	Sustainable Time Series Forecasting Comparing Deep Learning and Statistical Models: Novel Insights	Icelandic historical sales data from supermarkets in the entire country	MASE	AutoARIMA and Transformer-Based Forecasting Model (TimeGPT)

four, eight, and 12 weeks; and for monthly granularity, one, two, and three months. This multi-horizon approach allows us to test the adaptability of the models across both short- and long-term forecasts.

In the second phase, we maintain a consistent number of forecasting steps across each granularity — four, eight, and twelve steps — forecasting four, eight, and 12 days for daily granularity, four, eight, and 12 weeks for weekly granularity, and four, eight, and 12 months for monthly granularity. This approach helps us analyze the models' consistency in handling forecasts over different temporal scales while maintaining comparable step sizes.

Finally, we simulate scenarios with varying percentages of missing data to assess the robustness of the forecasting models. Using the weekly granularity data set, we forecast four weeks, randomly deleting 1%, 5%, 10%, 15%, 20%, and 30% of the data to replicate common real-world challenges such as incomplete or delayed sales records. This allows us to evaluate how well the models retain predictive accuracy under imperfect data conditions, providing insight into their ability to handle uncertainty and incomplete information while maintaining reliable performance.

The specifics of evaluating the forecasting models with the novel wMASE and how we modify it according to our goals are further discussed in subsection 7.3.3. To assess the performance of the forecasting models, we conduct robust statistical tests on the rankings of the error metrics rather than the raw error values themselves. Specifically, after computing the wMASE for each model, we rank the models based on their wMASE scores. These rankings are then analyzed using the Friedman test to determine if there are statistically significant differences between the models. We apply the Nemenyi post-hoc test to pinpoint which specific models differed significantly. This approach allows us to evaluate the models' relative performance more robustly, ensuring that our conclusions are based on consistent patterns across the various data sets and granularities. Figure 7.1 illustrates the methodological approach of the study.

7.3.2 Forecasting Models

In our selection process for forecasting models, we aim for a comprehensive exploration of models that represent diverse architectures and innovations in time series forecasting. Our primary fo-

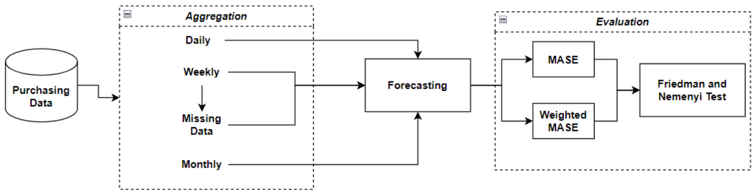


Figure 7.1: Methodological approach of the paper.

cus is on models developed and validated between 2023 and 2024, ensuring they reflect the most recent advancements in the field. A key criterion in our search is that the models be stable, publicly accessible, and not in beta to guarantee their reliability and maturity for robust analysis [332].

To ensure diversity in our chosen time series forecasting models, we include a range of architectures, from cutting-edge machine-learning models to classical statistical approaches. Among the recent models considered were Chronos [333], TimeGPT [52], Lag-Llama (in its zero-shot version) [334], Moirai [335], and Moment [336]. These models, released within the past year, represent significant advancements and were prioritized for their innovative approaches to time series forecasting.

In addition to these cutting-edge models, we include classic models such as Holt-Winters [337], and AutoARIMA [51], sourced from the Python library *statsforecast* and maintained by the same developers of TimeGPT. These classical models serve as benchmarks for comparison, allowing us to gauge the performance of newer models against well-established forecasting techniques.

Table B.1 summarizes the chosen forecasting models, providing details on their architecture, key features, and year of release. This structured comparison ensures a thorough and fair evaluation across various contexts. Through this blend of classical and modern techniques, we aim to identify the most effective models for supermarket demand forecasting.

7.3.3 Weighted MASE

The Mean Absolute Scaled Error (MASE) is a metric used to assess the accuracy of forecasting models, especially for time series data [338]. It is designed to provide a standardized way of evaluating and comparing the performance of different models [287].

The formula for MASE involves two main components: the Mean Absolute Error (MAE) of the forecasting model and the Mean Absolute Error of a naive forecasting approach. The formula for MASE is as follows:

$$MASE = \frac{MAE}{MAE_{naive}} = \frac{\frac{1}{h}\sum_{t=n+1}^{n+h}|d_t - f_t|}{\frac{1}{n-1}\sum_{t=2}^{n}|d_t - d_{t-1}|} \tag{7.1}$$

The numerator of the formula calculates the MAE, which is the average of the absolute differences between the actual observed values and the values predicted by the model over a given forecast horizon. This represents the model's average error in predicting future values. The denominator of the formula represents the MAE of a naive model, which typically assumes that the forecast for any future time period is simply the actual value from the previous period. This naive approach serves as a baseline for comparison. By dividing the MAE of the forecasting model by the MAE of the naive model, MASE provides a scale-invariant measure of accuracy. If the resulting MASE value is less than one, the forecasting model performs better than the naive approach, meaning it has a lower error. A MASE value equal to one suggests that the model performs similarly to the naive model. In contrast, a value greater than one indicates that the model underperforms compared to the naive approach. MASE aims to achieve a value as close to zero as possible, indicating superior predictive accuracy.

In evaluating time series forecasting models, the Mean Absolute Scaled Error (MASE) has emerged as a crucial and reliable performance measure [339], [340]. MASE offers a significant advantage over traditional error metrics by providing a scale-invariant and interpretable measure of forecast accuracy, particularly when compared to a naive baseline model [340].

The utility of MASE lies in its ability to standardize error across different time series, making it an ideal choice for comparing models in diverse contexts [340]. The goal is to achieve a MASE value close to zero, reflecting superior predictive accuracy. Its sensitivity to scaling and capacity to offer insights into models' relative performance has solidified its role in benchmarking and assessing time series forecasting methodologies [341].

Given its advantages, adopting MASE for consistent and comparable results in forecast accuracy evaluations is recommended, ensuring more reliable decision-making and ultimately leading to improved forecasting methodologies [341].

In our paper, we introduce a novel modified version of the MASE metric, wMASE, using a weighted approach to better align with our goal of reducing overproduction, a significant contributor to pollution and waste [289]. This wMASE is designed to penalize overestimations more heavily than underestimations, reflecting the more significant environmental and economic consequences of producing more than needed. By incorporating a weight variable into the error calculation, we assigned greater importance to overestimation errors, ensuring that models that predict higher-than-actual demand are penalized more, thus encouraging more accurate and sustainable forecasting practices.

The novel formula that supports our paper is as follows:

$$wMASE = w{*}MASE = w{*}\frac{MAE}{MAE_{naive}} = w{*}\frac{\frac{1}{h}\sum_{t=n+1}^{n+h}|d_t - f_t|}{\frac{1}{n-1}\sum_{t=2}^{n}|d_t - d_{t-1}|} \tag{7.2}$$

$w \in [1, 2]$

In our paper, we test two versions of wMASE. First, we set the weight to 1, which applied no additional penalization, treating overestimation and underestimation equally. Then, we set the weight to 2, introducing a penalty that specifically targeted overestimation errors, thereby emphasizing the reduction of overproduction.

7.4 Results

In this Results section, we present the outcomes of our forecasting experiments in a structured manner. First, we describe the forecasting performances for the same forecasting horizon and same horizon steps, without missing data, both with and without penalization ($w = 2$ and $w = 1$). Next, we present the results with different percentages of missing data and the average overestimation level of each model. Lastly, we outline the findings of the statistical tests, including the Friedman and Nemenyi post-hoc tests used to identify statistically significant differences between the forecasting models. This comprehensive results section offers a detailed comparison of model performance across varying error penalizations and statistical evaluations.

7.4.1 Forecasting Performances

Same Forecasting Horizon

Figure 7.2 shows the results comparing the performance of different forecasting models across the same forecasting horizons, both without penalization (left plot) and with penalization for overestimation (right plot).

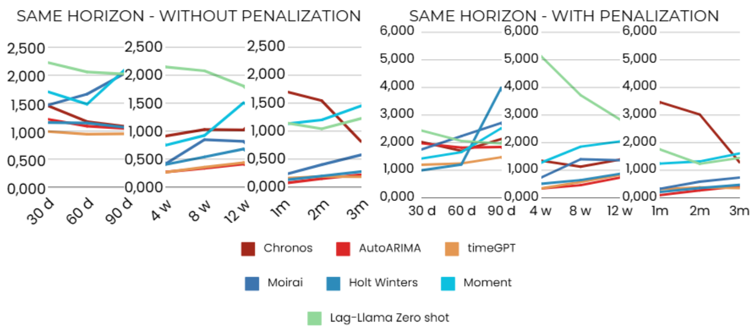


Figure 7.2: Forecasting models performances with or without penalized overestimation with same prediction horizon.

When overestimations are not penalized, TimeGPT stands out for having the best overall performance, especially in the short-term forecasts (30, 60, and 90 days). This indicates its strong adaptability to the multi-horizon approach. AutoARIMA also performs well, particularly in small forecasting steps with a monthly granularity (1m, 2m), achieving the lowest wMASE value. Holt-Winters and AutoARIMA show competitive performance for a monthly granularity, where the forecasting steps are smaller but begin to lag as the forecasting horizon extends. In contrast, Moment and Lag-Llama show much higher wMASE values, indicating poorer performance. Lag-Llama, in particular, performs the worst across most horizons, consistently showing the highest errors.

When penalizing overestimations, the performance of several models significantly declines, most notably Chronos, which sees a substantial increase in error across all horizons. TimeGPT remains resilient, although its performance deteriorates slightly, it maintains a competitive edge. Holt-Winters also shows robustness, though it encounters issues in more extended forecasts, such as a dramatic spike in error for the 90-day forecast. AutoARIMA and Moirai also show increased errors under penalization. Still, AutoARIMA remains competitive, particularly in monthly horizons, where it maintains the lowest error rates, especially for 1 month forecast. Lag-Llama, which already struggled without penalization, performs even worse under penalization, particularly in longer-term forecasts.

Moirai's performance is notably inconsistent. For the shorter-term forecasts (e.g., 30, 60, and 90 days), Moirai tends to underperform compared to other models like TimeGPT and AutoARIMA, often showing more significant errors. Its error rate increases significantly as the forecast horizon extends into longer periods (e.g., two and three months), suggesting that Moirai struggles with long-term accuracy.

In particular, the 90-day forecast shows a sharp rise in error, indicating that Moirai is less reliable as the forecasting horizon increases. Interestingly, when penalization for overestimation is applied, Moirai's performance deteriorates even further, especially for long-term forecasts, reflecting a greater vulnerability to over-forecasting. Despite some relatively stable results in mid-range forecasts like 12 weeks, Moirai lacks the adaptability and resilience to handle both short and long-term predictions as effectively as other models in the comparison.

Same Forecasting Steps

The results in Figure 7.3 illustrate the performance of various models when forecasting at the same horizon steps, without and with the penalization of overestimation, respectively.

AutoARIMA and TimeGPT perform consistently well in the first scenario, especially in shorter horizons. AutoARIMA achieves lower errors in forecasts of four to 12 days and four to eight weeks, with its best performance at four months. TimeGPT follows closely, slightly outperforming AutoARIMA in four-day and four-week predictions. These models maintain lower error rates, demonstrating robustness, particularly in short-term forecasts. Conversely, models like Chronos and Holt-Winters show slightly higher error rates, particularly at longer horizons like eight and 12 months. Moment and Lag-Llama consistently underperform across all forecasting horizons, showing more significant forecast deviations. Their errors notably increase at longer horizons, high-

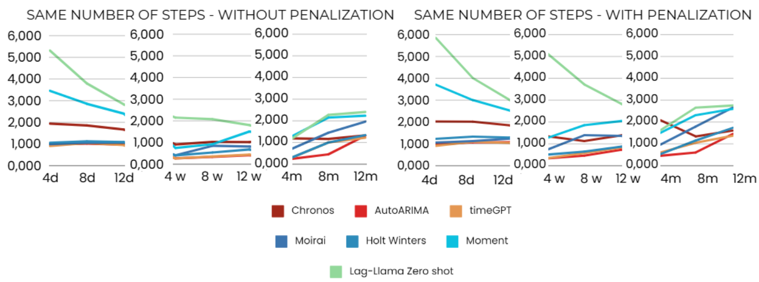


Figure 7.3: Forecasting models performances with or without penalized overestimation with same forecasting steps.

lighting their limitations when projecting further into the future.

When overestimation is penalized, the performance of all models is impacted, but AutoARIMA remains resilient, showing only a modest increase in error. It continues to perform well, particularly at the four-day and four-week horizons. TimeGPT shows a similar trend, with slightly higher errors, though it maintains competitiveness in shorter horizons. Chronos, however, suffers significantly from the penalization, particularly in longer horizons, where its error rates spike dramatically. Moment and Lag-Llama still perform poorly, with Lag-Llama especially failing to adapt to penalization, as seen in the four-week and longer-term forecasts.

Moirai's performance is quite varied across the different forecast horizons. In scenarios without penalization, it generally shows middling results, often getting outperformed by AutoARIMA and TimeGPT. Its accuracy is notably weaker in long-term forecasts, such as 8 and 12 months, where the error margins increase noticeably. However, in the mid-term ranges, like 12 weeks, Moirai demonstrates better results, placing it closer to the top-performing models.

When penalization is applied, Moirai's error margins grow significantly, especially in the longer horizons. For example, at the 12-month forecast, Moirai's error nearly doubles, highlighting its sensitivity to penalization for overestimation. While the model performs decently in shorter horizons, it struggles to maintain accuracy under conditions where overestimation is penalized, suggesting that it may be better suited for mid-term forecasts but lacks robustness for long-term, penalized predictions.

Missing Data

Figure 7.4 shows how different models perform when predicting with weekly granularity under varying levels of missing data. The data is removed randomly in percentages ranging from 1% to 30%, and the forecasting horizon is set to four weeks. This setup was chosen because the weekly granularity generally demonstrated the lowest error in prior experiments. The models are evaluated without penalizing overestimations ($w = 1$) to examine their robustness against missing data without additional complexities.

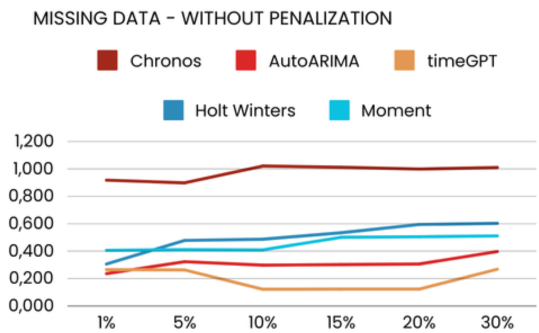


Figure 7.4: Forecasting models performances with different percentages of missing data without penalizing overestimation.

AutoARIMA consistently demonstrates strong resistance to missing data, maintaining low error rates even as the percentage of missing data increases. For instance, at 1% and 5% missing data, its wMASE is low, and even at 30% missing data, it shows a relatively low error, indicating robust performance.

TimeGPT is another model that performs exceptionally well under missing data conditions. In particular, it achieves the lowest error at 10%, 15%, and 20% missing data. However, at 30% missing data, its performance decreases slightly but remains competitive.

While performing moderately well at lower percentages of missing data, Holt-Winters starts showing significant increases in error as more data is removed. By the time 30% of the data is missing, its wMASE rises, indicating that it is less resilient to higher levels of missing data.

Chronos exhibits a consistent error trend across all levels of missing data, with wMASE values hovering around 1.0. This suggests that Chronos is less effective at handling missing data than AutoARIMA and TimeGPT, but it remains more stable than Holt-Winters when the data removal percentage increases.

Moment struggles somewhat more than the other models, with its errors staying in the 0.4-0.5 range as the missing data percentage increases. Even with 30% of data removed, Moment's error rises slightly, showing that it handles missing data reasonably well, but not as robustly as AutoARIMA or TimeGPT.

The forecasting models not present in this case study do not support missing data.

Average Overestimation

In this subsection, we explore the average overestimation across all models, referring to Figure 7.5, which summarizes the overestimation tendencies of each forecasting method based on the case studies.

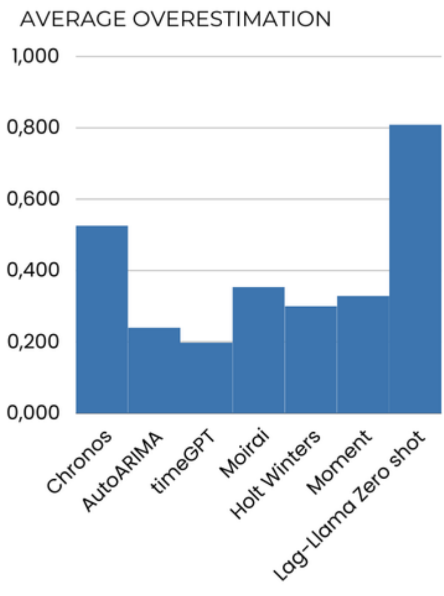


Figure 7.5: Average overestimation for each model.

Lag-Llama Zero Shot exhibits the highest average overestimation error, reflecting its tendency to predict values higher than the actual outcomes. This consistent bias toward overestimation could indicate limitations in adjusting for inevitable fluctuations in demand patterns, particularly in scenarios with volatile data.

Chronos follows closely, suggesting that this model consistently struggles with fine-tuning its predictions to more realistic values, potentially due to its complexity or generalization issues when dealing with the data set specifics.

Moirai also displays a significant upward bias in its forecasts, though it's less severe than Chronos and Lag-Llama. While it does perform better in medium-range forecasts, as discussed earlier, this metric suggests that Moirai often overshoots its predictions, which could make it less reliable in practical use cases where avoiding surplus is critical.

AutoARIMA and TimeGPT, in contrast, show much more controlled overestimations. As previously noted, these models maintain a balanced prediction strategy, leading to better performance across different forecast horizons. This makes them more dependable choices when precise predictions are required, particularly in minimizing waste or overstock scenarios.

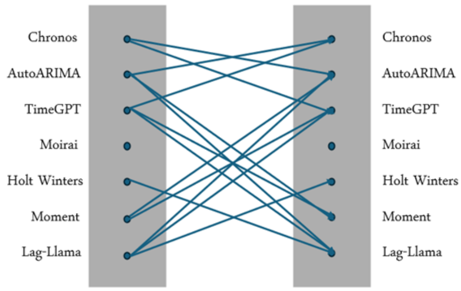
Holt-Winters and Moment also show moderate overestimations. Their tendencies to overshoot are lower than that of Moirai but still higher than TimeGPT and AutoARIMA, placing them in a middle ground between robustness and a slight upward bias.

7.4.2 Statistical Tests

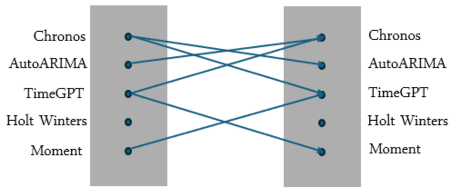
The statistical analysis of the forecasting models across three case studies—using the same prediction horizon, the same number of prediction steps, and handling missing data—highlights important differences in their performance.

Friedman's test results show a very low p-value for all case studies, indicating significant statistical differences among the forecasting models. To further investigate these distinctions, we applied the Nemenyi post-hoc test to identify which models performed significantly better than others.

In the first case study, all granularities shared the same prediction horizon (both with and without penalization for overestimation). Similarly, the second case study, which focuses on the same number of prediction steps, yields comparable statistical significance, both with and without penalizing overestimation.



(a)



(b)

Figure 7.6: Statistically significant difference between time series forecasting models, forecasting both with same prediction horizon and same number of prediction steps (a) and handling different percentages of missing data in the data set (b).

The first and second case studies demonstrate the same statistically significant differences among the forecasting models, as illustrated in Figure 7.6(a).

From the figure, we observe that AutoARIMA and TimeGPT are statistically distinct from models such as Chronos, Moment, and Lag-Llama. Furthermore, Tables B.2, B.3, B.4, and B.5 confirm that AutoARIMA and TimeGPT are statistically superior. Additionally, Holt-Winters is shown to be statistically better than Lag-Llama, though it does not differ significantly from the other forecasting models.

The Nemenyi post-hoc test for the third case study, which deals with missing data, reveals a similar pattern of statistically significant differences as in the previous two case studies. These results are presented in Table B.6 and visualized in Figure 7.6(b).

Figure 7.6(b) demonstrates that the statistical differences observed are comparable to those in the earlier case studies, albeit with slight variations. Chronos, once again, performs significantly worse than both AutoARIMA and TimeGPT. This time, only TimeGPT is statistically better than Moment, while Holt-Winters shows no significant differences from the other models.

7.5 Discussions

This paper evaluates the performance of several time series forecasting models under three different experimental setups: (i) the same prediction horizon, (ii) the same number of prediction steps, and (iii) datasets with missing data. In the first two scenarios, the forecasting models were tested both with and without penalizing overestimating, thanks to our developed wMASE. Our analysis uses Friedman's test to identify statistical differences among the models, followed by the Nemenyi post-hoc test to determine which models significantly outperformed others. Additionally, we analyze the average overestimation level of each model to assess their practical applicability, especially for demand forecasting in supermarket inventory management, with the goal of providing the tools to decrease food overproduction.

Our research is driven by the question: *Which forecasting models are most effective in predicting purchasing demand, helping minimize overproduction and waste in the supermarket sector, considering different prediction horizons, levels of granularity, and missing data?*. However, there is no unique answer to our RQ. The results of our analysis do not point to a single best model; instead, the effectiveness of a forecasting model depends on the data characteristics, specific scenarios, and available computational resources. Table 7.2 shows a summary of the results obtained.

Our key contribution lies in demonstrating that different models offer different strengths depending on different factors, and our analysis provides clear recommendations based on the trade-offs observed. Besides, the development of wMASE offers valuable insights into evaluating time series forecasting from different perspectives in situations where not all the forecasts have the same impact.

The first case study compares all models using the same prediction horizon across various granularities. The results indicate

Table 7.2: Summary of the results obtained and the applicability of each forecasting model.

Method	Best Condition	Explanation
AutoARIMA	Short to medium-term horizons	AutoARIMA performs exceptionally well with both short and medium prediction horizons, with or without missing data. Minimal overestimation across all cases.
TimeGPT	Medium to long-term horizons, Handling missing data	Consistently robust across time horizons and missing data scenarios, with the lowest overestimation levels. Best suited for real-world conditions with data gaps.
Holt-Winters	Stable environments, Short-term horizons	Performs well in shorter-term, stable forecasting scenarios but struggles when data becomes incomplete or horizons grow longer.
Moirai	Mid-range performance, Stable environments	Shows moderate performance in short-term and stable settings. Does not excel in handling missing data or longer horizons.
Chronos	Stable data without missing values	Performs acceptably in stable conditions but struggles significantly with missing data or longer-term horizons. Prone to overestimation.
Moment	Short-term horizons	Performs reasonably in short-term settings but cannot compete with leading models in terms of accuracy or resilience to missing data.
Lag-LLama	Rarely performs best	Struggles across most forecasting scenarios, especially with missing data and long horizons. Exhibits the highest error and overestimation levels.

that TimeGPT and AutoARIMA are the most reliable models overall. Both models consistently perform across multiple forecasting horizons, especially in short-term forecasts (four days to four weeks). TimeGPT stands out with the lowest error rates in penalized and non-penalized setups, reinforcing its position as the top-performing model.

However, when overestimation is penalized, the performance of some models, such as Chronos and Lag-Llama Zero-Shot, drops significantly. These models, already prone to overestimating demand, show an apparent vulnerability when introducing penalties. On the other hand, AutoARIMA maintains its strong performance across all horizons, demonstrating robustness even when penalizations were applied, making it ideal for scenarios where overproduction is particularly costly.

While performing moderately well in the non-penalized scenario, Moirai shows significant degradation when penalties are introduced, especially in mid-to-long-term forecasts (12 weeks and beyond). This suggests that Moirai may be unreliable for practical applications requiring precise demand forecasts without overestimating supply.

The second case study examines the models' performance when tasked with the same number of prediction steps, again with and without penalizing overestimations. Here, the results mirror those of the first scenario, with TimeGPT and AutoARIMA once again outperforming other models. Notably, AutoARIMA excels in medium-term predictions (four to 12 weeks), while TimeGPT continues to dominate across all horizons, showing minimal error rates even in the most challenging forecasting environments.

Introducing penalties for overestimations with our novel wMASE amplifies the differences between the models. Chronos, which initially demonstrates moderate performance in the non-penalized scenario, sees its error rates increase when penalizations are applied, particularly in short-term forecasts. Lag-Llama Zero-Shot and Moment also struggle in this scenario, confirming their unsuitability for real-world demand forecasting where overproduction has significant cost implications.

Moirai once again displays inconsistent behavior. While it shows competitive performance in the non-penalized scenario, its error rates surge when penalties are introduced, making it less practical for environments that require careful inventory management to avoid waste.

The third case study focuses on how the forecasting models perform when handling datasets with varying percentages of missing data. Missing data is a common issue in real-world forecasting applications, and this scenario provided valuable insights into the robustness of each model.

TimeGPT and AutoARIMA again emerge as the most resilient models. They display minimal performance degradation even when faced with large amounts of missing data, underscoring their ability to adapt to imperfect datasets. This robustness is essential for demand forecasting in industries like supermarkets, where data collection can be incomplete or inconsistent due to logistical challenges.

Moirai and Lag-Llama Zero-Shot, in contrast, show significant performance drops in the presence of missing data, indicating that these models lack the flexibility required for real-world applications. Although slightly more resilient than Moirai, Chronos still performs poorly, especially in long-term forecasts with missing data, making it a less reliable option in such scenarios.

Examining average overestimation levels, TimeGPT and AutoARIMA stand out as the models with the lowest tendency to overestimate demand. This makes them highly suitable for environments where minimizing waste and avoiding overproduction are critical. Chronos and Lag-Llama Zero-Shot had the highest overestimation levels, making them less reliable in scenarios where precise demand forecasting is required to avoid surplus.

The statistical analysis using Friedman's test reveals significant differences in the models' performance across all three case studies. The Nemenyi post-hoc test confirms that TimeGPT and AutoARIMA are statistically better than Chronos, Moment, and Lag-Llama Zero-Shot models. Furthermore, while generally competitive, Holt-Winters does not show a statistically significant difference from the top performers in case studies, positioning it as a solid alternative but not the top choice.

It is interesting to notice how the scenarios that do not penalize overestimation reveal the greatest statistically significant results in model comparison. This indicates that even the best-performing models are prone to overestimation under certain circumstances, which can have critical implications in practical applications like inventory management, where overestimations lead to overproduction and, subsequently, waste. By identifying this trend, we provide actionable insights for practitioners in the food supply chain industry, where minimizing waste is essential for both environmental sustainability and economic efficiency [309].

In contrast, the scenarios where overestimation is penalized highlight the tendency of some models to perform poorly when they overestimate demand. The ability to penalize overestimation offers a closer reflection of real-world conditions where businesses must balance having enough stock to meet demand and avoiding a surplus that could lead to waste [45].

Another crucial factor we address is the ability of models to handle missing data. In real-world forecasting, it is common for data sets to have missing or incomplete information [342]. Our paper finds that not all models can support missing data; for example, models such as Moirai and Lag-Llama Zero-Shot cannot handle incomplete data sets, making them unsuitable for practical application in imperfect data collection environments. On the other hand, TimeGPT and AutoARIMA prove to be highly resilient even in this regard, handling missing data without significant performance degradation. This robustness is critical for food supply chain forecasting, as data gaps are common due to various logistical challenges [342].

7.5.1 Main Contribution

Through these analyses, we inspect different highly relevant scenarios for real-life applications, particularly in food supply chain management. The three case studies represent the challenges businesses face when forecasting demand under uncertain conditions, such as incomplete data or the risks of overestimation. After thoroughly evaluating the performance of each model, we recommend AutoARIMA and TimeGPT as the best overall models. Both models consistently outperformed others regarding accuracy and resilience across all scenarios. However, it is essential to note that TimeGPT requires online tokens, which may not always be feasible for all users due to computational or financial limitations. In the absence of online access or when computational power is constrained, AutoARIMA becomes the preferred model, offering high accuracy without additional online resources. It is particularly noteworthy that a classical model like AutoARIMA consistently performs at the top alongside modern, more complex models [343]. This demonstrates the enduring value of traditional statistical methods, which—despite their simplicity—can still compete with advanced machine learning approaches [344]. AutoARIMA's ability to adapt and handle varying data conditions while remaining computationally efficient suggests that well-established models still hold significant relevance. This finding emphasizes that newer, more complex solutions are not always inherently superior, especially in practical, resource-constrained applications like food supply chain forecasting [343].

In addition to assessing performance, we examine the interpretability and stability of each forecasting model, as these qualities are essential for practical applications where understanding a model's decision-making process is critical [345]. Classic models like ARIMA and Holt-Winters are particularly valued for their interpretability [346]. This transparency is valuable for domain experts, who can easily adjust parameters based on their understanding of the data, facilitating trust and ease of use in operational settings [347].

In contrast, advanced models, such as TimeGPT, often function as "black boxes," providing high accuracy but limited insight into how predictions are made. The complexity of these models means that practitioners are generally unable to trace individual feature contributions or understand the internal mechanics, making it challenging to explain or justify predictions. This opacity can be problematic in fields like supply chain management, where accountability and explainability are crucial, and decisions based on obscure logic could lead to unanticipated risks [348].

Model stability further underscores the divide between traditional and deep-learning approaches. While classic models like ARIMA remain stable once trained, many deep-learning models dynamically update with new data, adapting continuously but potentially losing previously learned patterns. While beneficial in rapidly changing environments, this characteristic can introduce inconsistencies and bias over time, complicating long-term forecasting reliability. For practitioners prioritizing interpretability, stability, and ease of control, these attributes highlight the advantage of traditional models in environments where a clear, stable forecast logic is as important as accuracy.

The primary contribution of our research lies in providing a structured comparison of forecasting models to help food supply chain managers minimize overproduction and reduce waste. Our findings are precious for industries like supermarkets, where sales forecasting directly impacts inventory management [349]. Overproduction results in financial losses and contributes to environmental harm through increased waste. By offering a clear framework for selecting the appropriate forecasting model based on specific operational conditions, we equip supermarkets and similar businesses with the tools to make informed decisions and optimize their operations.

7.6 Conclusions

Our study provides a comprehensive analysis of forecasting models in various real-world scenarios. We have shown that AutoARIMA and TimeGPT are the best models for food supply chain forecasting, with AutoARIMA being the preferred choice when online resources are unavailable, and stability and interoperability are essential requirements. By addressing overestimation with our novel wMASE, we challenge overproduction and waste reduction, and we contribute to sustainability in supply chain operations, helping businesses optimize their forecasting practices and reduce their environmental footprint. Through this research, we offer a practical solution for one of the most pressing issues facing today's food production industry.

Looking forward, future studies could explore several avenues to build upon the foundation laid in this work. First, future research could investigate the performance of these models over longer forecasting horizons or under extreme scenarios, such as sudden demand spikes or food supply chain disruptions. Understanding how models like AutoARIMA and TimeGPT behave in highly volatile environments would provide deeper insights into their robustness and applicability across different industries. Another potential direction for future research is the cost-benefit analysis of model selection, considering the computational resources required to run more complex models like TimeGPT. Understanding the trade-off between accuracy and computational cost is essential for businesses that may not have access to powerful computational resources but still require high-performing models.

Chapter 8

Discussion

This thesis's discussion chapter synthesizes the findings of the previous chapters, each addressing critical dimensions of customer behavior. By integrating these diverse yet interrelated areas, this chapter explores the overarching insights and implications that emerge, advancing our understanding of sustainable customer behavior, the digital influence, and the reliability of forecasting models for predicting demand within complex, real-world contexts. However, while these findings contribute to the literature, they also present certain limitations that must be considered when interpreting the results.

This chapter provides a comprehensive overview of the key findings and discusses the broader implications for sustainability and operational efficiency in various domains, particularly customer behavior and food supply chains.

8.1 Key Findings

Chapter 4 addresses the first RQ: *How do grocery customers perceive their purchasing intentions towards green products and how do they really purchase?*. The results reveal a nuanced picture of customer behavior. The study found a positive correlation between customers' environmental concern and their green purchase intention, indicating a genuine concern for the environment among participants. HHowever, the findings from the purchase data reveal that customer behavior does not always translate into action, as shown by the minimal response to eco-friendly product labels

[350]. Further analysis revealed that female customers correlated more strongly with environmental concerns than males, suggesting a gendered dimension in sustainability attitudes [351]. Nevertheless, these results should be interpreted with caution, as the gender-based differences might be influenced by other confounding factors, such as income levels, household responsibilities, or cultural expectations.

This finding also highlighted a broader gap between intentions and actions, a disconnect that has been echoed in previous research, which suggests that despite positive intentions, attributes such as price and product quality often outweigh environmental considerations when it comes to actual purchasing decisions [352], [353]. However, this study did not account for other potential exogenous factors that might influence purchasing decisions, such as availability of green products, in-store marketing, or habit formation, all of which could impact consumer behavior. This discrepancy between intentions and behavior indicates that while eco-labeling may be useful, it is unlikely to be a standalone solution for driving substantial changes in customer purchasing habits [354].

Chapter 5 focuses on the impact of social media, particularly during high-stakes periods such as the COVID-19 pandemic, when customers were faced with unprecedented uncertainty. This Chapter addresses the second RQ: *How do social media impact purchasing behavior?*. The study highlights that customers often trust social media information without verifying its accuracy, leading to significant shifts in purchasing behavior [355]. This was particularly evident in the correlation between Twitter posts and increased fish oil purchases, even without any corresponding increase in Google search trends. The findings from this paper underscore the power of social media as an information driver [30]. Additionally, the strong Granger causality between Twitter mentions and fish oil purchases suggests that there might be a causality between social media and purchases, prompting important ethical considerations regarding the responsibilities of online platforms [30]. However, while strong Granger causality was found between Twitter mentions and fish oil purchases, it is important to recognize that correlation does not imply direct causation. There may be additional underlying factors that explain this relationship, such as traditional media coverage, word-of-mouth effects, or concurrent health trends that were not controlled for in this study. These findings also stress the importance of critically evaluating information online, especially in periods of crisis when unproven news can spread rapidly and unchallenged [356].

Chapter 6 partially answers the third RQ: *How effective are certain nudging technique when influencing customers towards greener purchasing?* by studying a specific demographic of potential customers and a specific type of nudging technique. Given their role as future decision-makers and customers, we examine adolescents, a key demographic for promoting sustainable food purchases. The study reveals that adolescents exhibit an inherent preference for eco-friendly options, even in the absence of overt digital nudges such as traffic-light labeling systems. However, this finding is limited by the scope of the study, as it does not account for differences in socio-economic status, prior environmental awareness, or cultural influences that could shape adolescent decision-making.

This preference aligns with the results of Chapter 4 suggesting that women—and by extension, adolescent girls—are more likely to engage in eco-conscious purchasing behavior. Interestingly, the study also shows that while the red traffic light, designed to dissuade unsustainable choices, had minimal impact on adolescents' snack choices, the green traffic light effectively promoted sustainable choices. This indicates that positive reinforcement aligns more with adolescent values than negative reinforcement [357]. These findings may suggest that adolescents in certain socio-cultural-economical environments are inclined toward sustainable behaviors, but their motivations are complex and likely shaped by curiosity, awareness, and social influences. Thus, while the results suggest potential for positive nudging techniques, additional experiments using a broader range of nudges are necessary to fully answer the RQ.

The final study (Chapter 7) focuses on the fourth and last RQ: *Which is the most accurate forecasting model to predict customers' demand in order to avoid overproduction?*. This Chapter evaluates various time series forecasting models to identify the most effective methods for demand forecasting in the food supply chain, a sector where precise demand predictions are crucial for minimizing waste [358]. The study rigorously tested several forecasting models across various scenarios, including fixed prediction horizons, fixed prediction steps, and datasets with missing data. The findings reveal that model choice must be context-dependent, as different models perform better under different conditions. For instance, TimeGPT and AutoARIMA consistently performed well, particularly when penalized for overestimation, making them suitable for environments where overproduction is costly. The results suggest that while advanced models like TimeGPT offer higher accuracy, traditional models like AutoARIMA perform similarly well without relying on computationally demanding resources. This highlights that simplicity can rival complexity in practical settings, especially when computational resources are limited [359]. Furthermore, the robustness of TimeGPT and AutoARIMA in handling missing data indicates that these models are better suited to the real-world challenges faced by supply chain managers. The ability to handle incomplete data is particularly valuable in reducing food waste and ensuring inventory levels align more closely with actual demand, a key goal for promoting sustainability in the food system [360]. However, it is crucial to recognize that these models were evaluated based on past data, and their real-world performance could be affected by unforeseen market changes, economic shocks, or shifts in customer behavior. Future studies should compare model outputs with actual reductions in food waste to establish a clearer causal relationship.

The discussion also touches on the importance of model interpretability, especially in the critical area of food supply forecasting. While more complex models like TimeGPT may offer higher accuracy, simpler models such as ARIMA and Holt-Winters are valuable for their transparency, allowing domain experts to adjust and trust parameters based on their understanding of the data [361]. This balance between accuracy and interpretability is vital for decision-making in high-stakes environments where trust and clarity are essential [362].

8.2 Theoretical Implications

The findings across the four studies provide valuable insights into how customer intentions, external influences, demographics, and technology interact to shape sustainable consumption. This thesis deepens theoretical understanding by exploring the gap between people's stated environmental intentions and their actual purchasing behaviors, as discussed in Chapters 4 and 6. This gap supports existing ideas in customer psychology and behavioral economics and highlights the need for models that consider social and psychological factors to better predict eco-friendly purchasing patterns.

Chapters 5 and 7 expand theories on digital influence and operational efficiency. Chapter 5 offers key insights into how digital misinformation affects customer behavior during crises. Chapter 7 connects sustainability and operations research, showing how accurate forecasting models can align environmental goals with economic priorities, offering practical solutions for sustainable supply chains.

Together, these studies challenge and enrich theories from multiple disciplines by showing how internal factors (like intentions and concerns) interact with external ones (like social and technological influences) to shape sustainable behaviors. They highlight the need for a combined theoretical approach that links behavioral, digital, and operational insights to better understand and encourage sustainability.

8.3 Practical Implications and Recommendations

The practical contributions of this thesis lie in its actionable recommendations for fostering sustainable behaviors and improving systemic operations.

This thesis focuses on customers: their intentions, purchasing habits, and the various factors influencing their choices. While individual customer choices are significant, they operate within a larger system influenced by powerful actors, including supermarkets, supply chain managers, food producers, and governments. These entities not only shape the purchasing environment but also bear substantial responsibility for enabling and supporting sustainable practices. Figure 8.1 illustrates the dynamic interaction between customer choices and the broader supply chain, emphasizing the reciprocal influence between demand patterns and systemic responses. By analyzing comprehensive customer data, this thesis provides a unique opportunity to make informed recommendations to these influential actors, with the aim of reducing waste and fostering sustainable consumption from the top down.

First, while customers play a role in driving demand, larger actors must take the lead in creating an environment conducive

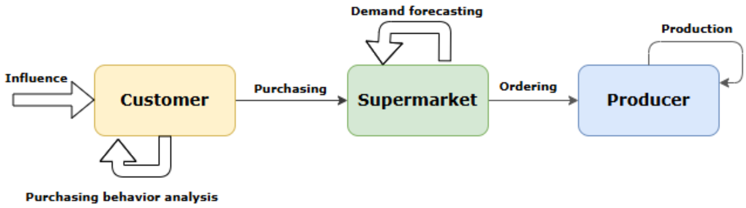


Figure 8.1: Overview of how customer choices impact supermarket operations and broader supply chain decisions.

to sustainable decision-making. Governments and supply chain stakeholders should implement policies and practices that facilitate sustainable purchasing, such as incentivizing the availability of eco-friendly products, ensuring transparent labeling, and promoting affordable sustainable alternatives.

Educational campaigns targeting customers remain critical as they bridge the gap between environmental concerns and purchasing behavior. These initiatives should focus on raising awareness of the long-term environmental impacts of purchasing decisions and provide practical guidance for sustainable choices.

Simultaneously, social media platforms must be regulated to ensure the accurate dissemination of sustainability-related information, creating a digital environment that supports informed choices. Given the demonstrated influence of social media on purchasing behavior, targeted campaigns promoting sustainable consumption could have a profound impact, reaching customers at scale.

Lastly, as Chapter 7 underscores, it is important to adopt accurate demand forecasting models to reduce food waste and align production with actual customer demand. Production companies and/or grocery stores should adopt advanced forecasting techniques to minimize overproduction and waste. To ensure their effectiveness, these models must remain transparent and accessible, fostering trust and informed decision-making.

The food system can become more resilient and environmentally sustainable by integrating customer education, digital literacy, targeted interventions, and improved forecasting models. This collaborative effort lays the foundation for a future in which sustainable practices are embedded throughout the supply chain, benefiting both the environment and society as a whole.

8.3.1 Limitations

While this thesis provides valuable insights into sustainable customer behavior, digital influences, adolescent decision-making, and forecasting models, several limitations should be considered. Firstly, the sample sizes and contexts of the studies may limit the generalizability of the findings. For instance, the research on customer behavior (Chapter 4) and social media influence (Chapter 5) focused on specific demographics and products, which may not fully represent diverse global customer groups. Similarly, the study on adolescents' preferences for eco-friendly products (Chapter 6) may not capture the full range of factors influencing sustainability decisions across different age groups, cultures, or socio-economic backgrounds. These limitations suggest the need for future studies that include a broader, more representative sample and consider socio-economic and cultural factors that may affect sustainable consumption behaviors.

Another limitation is the reliance on cross-sectional data in some studies, which prevents the analysis of long-term trends or causality. Specifically, in Chapters 4 and 5, the data collected at a single point in time may not fully reflect shifts in customer behavior over extended periods or in response to changing social and economic factors. Future work should aim to use longitudinal data to track changes in behavior over time and explore causality more robustly, especially in the context of external factors like economic shifts or public crises.

Additionally, this research did not delve deeply into other potential influencing factors, such as socioeconomic status, digital literacy, or cultural differences, which could further shape customer behavior, decision-making, and the effectiveness of digital or nudge-based interventions. A more comprehensive exploration of these factors could provide a richer understanding of the barriers and motivators that impact sustainable consumption, allowing for more targeted and effective interventions.

Finally, while the models for demand forecasting presented in Chapter 7 show promising results, they were evaluated based on historical data, which may not fully account for unpredictable external influences, such as changes in consumer trends, economic fluctuations, or unforeseen events like supply chain disruptions. Thus, the real-world application of these models must be tested in dynamic and complex settings where external factors might influence demand forecasts.

Chapter 9

Conclusion

This thesis explores sustainable customer behavior, the role of digital media, the impact of nudging techniques, and the potential of demand forecasting to enhance food supply chain efficiency. Using comprehensive purchasing data from a leading supermarket chain serving the entire Icelandic market, this research provides a nationwide perspective on customer behavior. The insights gained are not only applicable locally but also have the potential to inform sustainable practices on a global scale, offering valuable contributions to real-world decision-making.

The broader implications of this work emphasize the interconnectedness of individual purchasing decisions, supply chain dynamics, and environmental outcomes. As the food system—one of the largest contributors to global greenhouse gas emissions—faces the challenge of feeding a growing population while reducing its environmental impact, this thesis underscores the role of customer behavior in driving systemic change. From improving demand forecasting to minimizing waste, the findings highlight actionable strategies to promote sustainability across the food system. By addressing the factors that influence sustainable choices, this thesis offers insights to encourage behavioral shifts, align production with demand, and reduce the environmental footprint of food production and consumption.

The main contribution of this thesis lies in its in-depth analysis of customer behavior and sustainability in the food system. By integrating psychological, social, and technological perspectives and leveraging extensive real-world data, it bridges the gap between theory and practice, offering concrete strategies to foster sustainable consumption and combat food waste.

9.1 Future Work

Despite the limitations presented in the previous chapter, the findings point to several promising avenues for future research and policy development aimed at promoting environmental responsibility and reducing ecological impact. As mentioned, addressing environmental sustainability in its entirety is complex and goes beyond the scope of this thesis. However, this study lays the groundwork for several next steps, each expanding on the contributions of this work.

Future research could broaden the scope of nudging techniques beyond the adolescent demographic explored here, extending their application to a wider population. This would help understand how various demographic groups—across ages, cultures, and socioeconomic backgrounds—respond to different nudging strategies. Such research could help identify how these strategies can be optimized to promote eco-friendly behaviors more effectively across diverse groups, and could also examine how cultural and socioeconomic factors influence the success of certain nudges.

Additionally, while this research relied on Icelandic purchasing data, there is significant potential for comparative studies across multiple countries. Future studies could shed light on how factors like labeling, pricing, and marketing influence environmentally conscious behavior in different cultural and economic contexts by examining data from multinational retailers. Such research could identify region-specific barriers and enablers for sustainable consumption, helping to tailor interventions for specific cultural environments and providing a broader understanding of global consumer behavior.

The demand forecasting models presented in this thesis demonstrate considerable practical potential in minimizing food waste and improving operational efficiency. Future work could adapt these models for use in larger production and retail settings, where their ability to accurately forecast demand could significantly reduce environmental impact. However, real-world applications should focus on testing these models across various industries, as model performance can differ depending on the complexity and volatility of specific markets.

Furthermore, the thesis provides a foundation for exploring innovative ways to encourage responsible purchasing decisions. By combining data-driven insights with proven nudging techniques, future studies could develop new approaches that integrate personalized digital nudges, eco-labels highlighting the positive environmental impact of customer choices, or reward systems designed to foster long-term eco-conscious habits. It would be valuable to explore how these interventions align with customers' values and motivations in the long term, as well as to test their effectiveness across various sectors and consumer groups.

This thesis not only offers valuable insights into responsible consumer behavior but also paves the way for future research that can extend these findings on a broader scale. Through data-driven strategies and targeted interventions, there is a significant opportunity to foster environmentally sustainable actions not only in customer markets and supply chains but also across broader societal systems. Ultimately, these efforts could contribute to shaping a greener, more sustainable future.

Bibliography

[1] C. Carpinelli, E. T. Einarsson Reynis, A. S. Islind, H. Stefánsson, and M. Óskarsdóttir, "Green intentions: Field research and data-driven analysis of customers' purchasing patterns", *Sustainability*, vol. 14, no. 16, p. 9863, 2022.

[2] C. Carpinelli, A. S. Islind, and M. Óskarsdóttir, "The quiet power of social media: Impact on fish-oil purchases in iceland during covid-19", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 18, 2024, pp. 203–213.

[3] C. Carpinelli, I. Visescu, A. S. Islind, and M. Óskarsdóttir, "Digital nudging towards sustainability: Exploring strategies to encourage greener choices among adolescent girls", in *Proceedings of the Italian Chapter of AIS (Association for Information Systems)*, 2024.

[4] C. Carpinelli, B. Baesens, S. V. Broucke, A. S. Islind, and M. Óskarsdóttir, "Novel insights on time series forecasting comparing deep learning and statistical models to avoid overproduction", *Paper Submitted to Expert Systems With Applications*, 2024.

[5] C. Carpinelli, S. V. Broucke, B. Baesens, and A. Sigridur, "Optimizing sustainability: Benchmarking time-series forecasting models on purchasing data to reduce overproduction", in *Proceedings of the Italian Chapter of AIS (Association for Information Systems)*, 2024.

[6] C. Carpinelli, A. S. Islind, C. A. Bravo, R. Calabrese, and M. Óskarsdóttir, "Extreme climate events and credit risk: A stress testing approach for loans to smes based on multilayer network analysis", 2024.

[7] K. Flanagan, K. Robertson, and C. Hanson, "Reducing food loss and waste", *Setting the Global Action Agenda: Washington, DC, USA*, 2019.

[8] M. Zurek, A. Hebinck, and O. Selomane, "Climate change and the urgency to transform food systems", *Science*, vol. 376, no. 6600, pp. 1416–1421, 2022.

[9] Y.-T. Tang, "Accounting for the environmental impact of food waste on water resources and climate change", in *Food Industry Wastes*, Elsevier, 2020, pp. 305–329.

[10] B. Bajželj, T. E. Quested, E. Röös, and R. P. Swannell, "The role of reducing food waste for resilient food systems", *Ecosystem services*, vol. 45, p. 101 140, 2020.

[11] M. Vittuari, F. De Menna, L. García-Herrero, M. Pagani, L. Brenes-Peralta, and A. Segrè, "Food systems sustainability: The complex challenge of food loss and waste", in *Sustainable Food Supply Chains*, Elsevier, 2019, pp. 249–260.

[12] J. Tollefson, "Soot a major contributor to climate change", *Nature*, vol. 15, no. 15, pp. 10–1038, 2013.

[13] W. Tumwesigye, A. Aschalew, W. Wilber, and A. Destra, "Impact of climate change on food systems: A narrative", *Journal of Water Resources and Ocean Science*, vol. 8, no. 4, pp. 50–55, 2019.

[14] J. H. Faghmous and V. Kumar, "A big data guide to understanding climate change: The case for theory-guided data science", *Big data*, vol. 2, no. 3, pp. 155–163, 2014.

[15] Z. Irani, A. M. Sharif, H. Lee, *et al.*, "Managing food security through food waste and loss: Small data to big data", *Computers & Operations Research*, vol. 98, pp. 367–383, 2018.

[16] L. Neumayr and C. Moosauer, "How to induce sales of sustainable and organic food: The case of a traffic light eco-label in online grocery shopping", *Journal of Cleaner Production*, vol. 328, p. 129 584, 2021.

[17] S. G. Azevedo, H. Carvalho, and V. C. Machado, "The influence of green practices on supply chain performance: A case study approach", *Transportation research part E: logistics and transportation review*, vol. 47, no. 6, pp. 850–871, 2011.

[18] T. Allen and P. Prosperi, "Modeling sustainable food systems", *Environmental management*, vol. 57, no. 5, pp. 956–975, 2016.

[19] G. Zeinstra, S. van der Haar, and G. van Bergen, *Drivers, barriers and interventions for food waste behaviour change: A food system approach.* Wageningen Food & Biobased Research, 2020.

[20] P. D. de Araújo, W. M. C. Araújo, L. Patarata, and M. J. Fraqueza, "Understanding the main factors that influence consumer quality perception and attitude towards meat and processed meat products", *Meat Science*, vol. 193, p. 108 952, 2022.

[21] M. C. Campbell, J. J. Inman, A. Kirmani, and L. L. Price, *In times of trouble: A framework for understanding consumers' responses to threats*, 2020.

[22] A. S. Al-Adwan, M. M. Al-Debei, and Y. K. Dwivedi, "E-commerce in high uncertainty avoidance cultures: The driving forces of repurchase and word-of-mouth intentions", *Technology in Society*, vol. 71, p. 102 083, 2022.

[23] M. Herrero, P. K. Thornton, D. Mason-D'Croz, *et al.*, "Articulating the effect of food systems innovation on the sustainable development goals", *The Lancet Planetary Health*, vol. 5, no. 1, e50–e62, 2021.

[24] R. M. Y. Özbük and A. Coşkun, "Factors affecting food waste at the downstream entities of the supply chain: A critical review", *Journal of Cleaner Production*, vol. 244, p. 118 628, 2020.

[25] I. Vermeir, B. Weijters, J. De Houwer, *et al.*, "Environmentally sustainable food consumption: A review and research agenda from a goal-directed perspective", *Frontiers in psychology*, vol. 11, p. 1603, 2020.

[26] G.-J. Nychas, E. Sims, P. Tsakanikas, and F. Mohareb, "Data science in the food industry", *Annual Review of Biomedical Data Science*, vol. 4, no. 1, pp. 341–367, 2021.

[27] F. P. Salvatore, N. Adamashvili, and F. Conto, "Factors affecting consumer purchasing behavior of functional food: A comparative analysis for consumer management", *British Food Journal*, vol. 124, no. 5, pp. 1519–1536, 2022.

[28] S. Varghese and M. Agrawal, "Impact of social media on consumer buying behavior", *Saudi Journal of Business and Management Studies*, vol. 6, no. 3, pp. 51–55, 2021.

[29] B. Oumayma, "Social media made me buy it: The impact of social media on consumer purchase behavior", in *Proceedings of the 4th International Conference on Smart City Applications*, 2019, pp. 1–7.

[30] Y. K. Dwivedi, E. Ismagilova, D. L. Hughes, *et al.*, "Setting the future of digital and social media marketing research: Perspectives and research propositions", *International journal of information management*, vol. 59, p. 102 168, 2021.

[31] O. Djaoudene, A. Romano, Y. D. Bradai, *et al.*, "A global overview of dietary supplements: Regulation, market trends, usage during the covid-19 pandemic, and health effects", *Nutrients*, vol. 15, no. 15, p. 3320, 2023.

[32] M. Gammone, G. Riccioni, G. Parrinello, and N. D'Orazio, "Omega-3 polyunsaturated fatty acids: Benefits and endpoints in sport", *Nutrients*, vol. 11, p. 46, Dec. 2018.

[33] M. Cesareo, A. Sorgente, M. Labra, *et al.*, "The effectiveness of nudging interventions to promote healthy eating choices: A systematic review and an intervention among italian university students", *Appetite*, vol. 168, p. 105 662, 2022.

[34] M. Proi, E. C. Dudinskaya, S. Naspetti, E. Ozturk, and R. Zanoli, "The role of eco-labels in making environmentally friendly choices: An eye-tracking study on aquaculture products with italian consumers", *Sustainability*, vol. 15, no. 5, p. 4659, 2023.

[35] J. Vandenbroele, I. Vermeir, M. Geuens, H. Slabbinck, and A. Van Kerckhove, "Nudging to get our food choices on a sustainable track", *Proceedings of the Nutrition Society*, vol. 79, no. 1, pp. 133–146, 2020.

[36] M. Mirbabaie, J. Marx, and L. Erle, “Digital nudge stacking and backfiring: Understanding sustainable e-commerce purchase decisions”, *Pacific Asia Journal of the Association for Information Systems*, vol. 15, no. 3, p. 3, 2023.

[37] M. Petel, “Nudging towards sustainability? a critical perspective on behavioral economics”, *Environs: Envtl. L. & Pol'y J.*, vol. 43, p. 223, 2020.

[38] S. Lorek and J. H. Spangenberg, “Sustainable consumption within a sustainable economy–beyond green growth and green economies”, *Journal of cleaner production*, vol. 63, pp. 33–44, 2014.

[39] A. Mazar, G. Tomaino, Z. Carmon, and W. Wood, “Habits to save our habitat: Using habit psychology to promote sustainability”, *Behavioral Science & Policy*, vol. 7, no. 2, 2021.

[40] H. Dreyer, N. Sonnenberg, and D. Van der Merwe, “Transcending linearity in understanding green consumer behaviour: A social–cognitive framework for behaviour changes in an emerging economy context”, *Sustainability*, vol. 14, no. 22, p. 14 855, 2022.

[41] K. Wohlfarth, E. Worrell, and W. Eichhammer, “Energy efficiency and demand response–two sides of the same coin?”, *Energy Policy*, vol. 137, p. 111 070, 2020.

[42] A. Sajjad, Q. Zhang, F. Asmi, M. A. Anwar, and M. Bhatia, “Identifying the motivating factors to promote socially responsible consumption under circular economy: A perspective from norm activation theory”, *J. Retail. Consum. Serv.*, vol. 76, no. 103544, 2024.

[43] S. S. Kamble, A. Gunasekaran, and S. A. Gawankar, “Sustainable industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives”, *Process Saf. Environ. Prot.*, vol. 117, pp. 408–425, 2018.

[44] A. Nygaard, “Green promotion and marketing communication”, pp. 63–81, 2024.

[45] R. Messner, H. Johnson, and C. Richards, “From surplus-to-waste: A study of systemic overproduction, surplus and food waste in horticultural supply chains”, *Journal of Cleaner Production*, vol. 278, p. 123 952, 2021.

[46] M. C. Annosi, F. Brunetta, F. Bimbo, and M. Kostoula, "Digitalization within food supply chains to prevent food waste. drivers, barriers and collaboration practices", *Ind. Mark. Manag.*, vol. 93, pp. 208–220, 2021.

[47] Y. Tadayonrad and A. B. Ndiaye, "A new key performance indicator model for demand forecasting in inventory management considering supply chain reliability and seasonality", *Supply Chain Analytics*, vol. 3, no. 100026, 2023.

[48] Thin, *Still an unhappy marriage*, `https://news.thin-ink.net/p/still-an-unhappy-marriage`, Accessed: 2024-11-4, 2022.

[49] *The 17 goals*, `https://sdgs.un.org/goals`, Accessed: 2024-11-4.

[50] *Goal 12*, `https://sdgs.un.org/goals/goal12`, Accessed: 2024-11-4.

[51] G. Mélard and J.-M. Pasteels, "Automatic arima modeling including interventions, using time series expert software", *International Journal of Forecasting*, vol. 16, no. 4, pp. 497–508, 2000.

[52] A. Garza and M. Mergenthaler-Canseco, "Timegpt-1", *arXiv preprint arXiv:2310.03589*, 2023.

[53] I. A. Adeniran, C. P. Efunniyi, O. S. Osundare, and A. Abhulimen, "Transforming marketing strategies with data analytics: A study on customer behavior and personalization", *International Journal of Management & Entrepreneurship Research*, vol. 6, no. 8, 2024.

[54] K. Chaudhary, M. Alam, M. S. Al-Rakhami, and A. Gumaei, "Machine learning-based mathematical modelling for prediction of social media consumer behavior using big data analytics", *Journal of Big Data*, vol. 8, no. 1, p. 73, 2021.

[55] M. R. bin Abdullah and K. Iqbal, "Machine learning's impact on customer behavior analytics: Current trends, challenges, and future directions", *Quarterly Journal of Emerging Technologies and Innovations*, vol. 9, no. 3, pp. 4–14, 2024.

[56] A. T. Rosário and J. C. Dias, "How has data-driven marketing evolved: Challenges and opportunities with emerging technologies", *International Journal of Information Management Data Insights*, vol. 3, no. 2, p. 100 203, 2023.

[57] S. Arefin, R. Parvez, T. Ahmed, *et al.*, "Retail industry analytics: Unraveling consumer behavior through rfm segmentation and machine learning", in *2024 IEEE International Conference on Electro Information Technology (eIT)*, IEEE, 2024, pp. 545–551.

[58] A. A. Mir, "Adaptive fraud detection systems: Real-time learning from credit card transaction data", *Advances in Computer Sciences*, vol. 7, no. 1, 2024.

[59] A. Dunn, K. Hood, and A. Driessen, *Measuring the effects of the COVID-19 pandemic on consumer spending using card transaction data*. US Department of Commerce, Bureau of Economic Analysis, 2020.

[60] W. N. Wassouf, R. Alkhatib, K. Salloum, and S. Balloul, "Predictive analytics using big data for increased customer loyalty: Syriatel telecom company case study", *Journal of Big Data*, vol. 7, 2020. DOI: `10.1186/s40537-020-00290-0`.

[61] M. Kasem, M. Hamada, and I. Taj-Eddin, "Customer profiling, segmentation, and sales prediction using ai in direct marketing", *Neural Computing and Applications*, pp. 1–11, 2023. DOI: `10.48550/arXiv.2302.01786`.

[62] N. Kim and K. Lee, "Environmental consciousness, purchase intention, and actual purchase behavior of eco-friendly products: The moderating impact of situational context", *International Journal of Environmental Research and Public Health*, vol. 20, 2023. DOI: `10.3390/ijerph20075312`.

[63] V. S. P. Nimmagadda, "Artificial intelligence for customer behavior analysis in insurance: Advanced models, techniques, and real-world applications", *Journal of AI in Healthcare and Medicine*, vol. 2, no. 1, pp. 227–263, 2022.

[64] X. Qi, H. Yu, and A. Ploeger, "Exploring influential factors including covid-19 on green food purchase intentions and the intention–behaviour gap: A qualitative study among consumers in a chinese context", *International journal of*

environmental research and public health, vol. 17, no. 19, p. 7106, 2020.

[65] M. Alshurideh, A. Gasaymeh, G. Ahmed, H. Alzoubi, and B. Kurd, "Loyalty program effectiveness: Theoretical reviews and practical proofs", *Uncertain Supply Chain Management*, vol. 8, no. 3, pp. 599–612, 2020.

[66] D. Vrontis, A. Makrides, M. Christofi, and A. Thrassou, "Social media influencer marketing: A systematic review, integrative framework and future research agenda", *International Journal of Consumer Studies*, vol. 45, no. 4, pp. 617–644, 2021.

[67] F. Fadillah and N. Kusumawati, "Factors affecting makeup products online impulsive buying behavior on tiktok", *Proceeding Book of The 6th ICMEM, no. August*, pp. 11–13, 2021.

[68] S. Kimiagari and N. S. A. Malafe, "The role of cognitive and affective responses in the relationship between internal and external stimuli on online impulse buying behavior", *Journal of Retailing and Consumer Services*, vol. 61, p. 102 567, 2021.

[69] X. Zhang and F. Dong, "Why do consumers make green purchase decisions? insights from a systematic review", *International journal of environmental research and public health*, vol. 17, no. 18, p. 6607, 2020.

[70] M.-C. Chiu, J.-H. Huang, S. Gupta, and G. Akman, "Developing a personalized recommendation system in a smart product service system based on unsupervised learning model", *Computers in Industry*, vol. 128, p. 103 421, 2021.

[71] R. Karthik and S. Ganapathy, "A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce", *Applied Soft Computing*, vol. 108, p. 107 396, 2021.

[72] M. K. Sahu, "Advanced ai techniques for optimizing inventory management and demand forecasting in retail supply chains", *Journal of Bioinformatics and Artificial Intelligence*, vol. 1, no. 1, pp. 190–224, 2021.

[73] V. L. Miguéis, A. Pereira, J. Pereira, and G. Figueira, "Reducing fresh fish waste while ensuring availability: Demand forecast using censored data and machine learning", *Journal of cleaner Production*, vol. 359, p. 131 852, 2022.

[74] K. Tabianan, S. Velu, and V. Ravi, "K-means clustering approach for intelligent customer segmentation using customer purchase behavior data", *Sustainability*, vol. 14, no. 12, p. 7243, 2022.

[75] V. KATRAGADDA, "Dynamic customer segmentation: Using machine learning to identify and address diverse customer needs in real-time", *IRE Journals*, vol. 5, no. 10, pp. 278–279, 2022.

[76] Y. Li, X. Chu, D. Tian, J. Feng, and W. Mu, "Customer segmentation using k-means clustering and the adaptive particle swarm optimization algorithm", *Applied Soft Computing*, vol. 113, p. 107 924, 2021.

[77] D. Egorenkov, "Ai-powered predictive customer lifetime value: Maximizing long-term profits", *Valley International Journal Digital Library*, pp. 7339–7354, 2024.

[78] S. Aiolfi, S. Bellini, and D. Pellegrini, "Data-driven digital advertising: Benefits and risks of online behavioral advertising", *International Journal of Retail & Distribution Management*, vol. 49, no. 7, pp. 1089–1110, 2021.

[79] J. P. Bharadiya, "A comparative study of business intelligence and artificial intelligence with big data analytics", *American Journal of Artificial Intelligence*, vol. 7, no. 1, p. 24, 2023.

[80] T. Davenport, A. Guha, D. Grewal, and T. Bressgott, "How artificial intelligence will change the future of marketing", *Journal of the Academy of Marketing Science*, vol. 48, pp. 24–42, 2020.

[81] N. Bora and Y. Sujana, "Evaluating the impact of artificial intelligence on customer buying decisions: A case study of electronic goods purchases through online platforms", *Journal of Marketing*, vol. 13, no. 2, pp. 1–15, 2023.

[82] A. Haque, N. Akther, I. Khan, K. Agarwal, and N. Uddin, "Artificial intelligence in retail marketing: Research agenda based on bibliometric reflection and content analysis (2000–2023)", in *Informatics*, MDPI, vol. 11, 2024, p. 74.

[83] A. M. Shah, S. Z. Zahoor, and I. H. Qureshi, “Social media and purchasing behavior: A study of the mediating effect of customer relationships”, *Journal of Global Marketing*, vol. 32, no. 2, pp. 93–115, 2019.

[84] G. Sogari, T. Pucci, B. Aquilani, and L. Zanni, “Millennial generation and environmental sustainability: The role of social media in the consumer purchasing behavior for wine”, *Sustainability*, vol. 9, no. 10, p. 1911, 2017.

[85] H. K. A. Wahab, M. Tao, A. Tandon, M. Ashfaq, and A. Dhir, “Social media celebrities and new world order. what drives purchasing behavior among social media followers?”, *Journal of Retailing and Consumer Services*, vol. 68, p. 103 076, 2022.

[86] A. Rodrigo and T. Mendis, “Impact of social media influencers’ credibility on millennial consumers’ green purchasing behavior: A concept paper on personal and social identities”, *Management Matters*, vol. 20, no. 2, pp. 134–153, 2023.

[87] A. Wiliam, W. Kosasih, Y. Indrianti, *et al.*, “Sentiment analysis of social media engagement to purchasing intention”, in *Understanding Digital Industry*, Routledge, 2020, pp. 362–365.

[88] J. R. Borre, G. C. Romero, and J. Ramírez, “Sustainable purchasing trends in new consumer generations: Ai
s role”, *Procedia Computer Science*, vol. 241, pp. 539–545, 2024.

[89] K. Lin and S. Wei, “Advancing the industrial circular economy: The integrative role of machine learning in resource optimization”, *Journal of green economy and low-carbon development*, vol. 2, no. 3, pp. 122–136, 2023.

[90] K. Sagar, A. Awasthi, S. P. Dwivedi, *et al.*, “Smart resource management: An innovative way of balancing materials, energy, and waste for a green future”, in *E3S Web of Conferences*, EDP Sciences, vol. 453, 2023, p. 01 037.

[91] J. M. Bauer, S. C. Aarestrup, P. G. Hansen, and L. A. Reisch, “Nudging more sustainable grocery purchases: Behavioural innovations in a supermarket setting”, *Technological Forecasting and Social Change*, vol. 179, p. 121 605, 2022.

[92] R. Bucea-Manea-Țoniș, O. M. Dourado Martins, D. Ilic, *et al.*, "Green and sustainable public procurement—an instrument for nudging consumer behavior. a case study on romanian green public agriculture across different sectors of activity", *Sustainability*, vol. 13, no. 1, p. 12, 2020.

[93] M. Berger, T. Lange, and B. Stahl, "A digital push with real impact–mapping effective digital nudging elements to contexts to promote environmentally sustainable behavior", *Journal of cleaner production*, vol. 380, p. 134 716, 2022.

[94] M. Gossen, S. Jäger, M. L. Hoffmann, F. Bießmann, R. Korenke, and T. Santarius, "Nudging sustainable consumption: A large-scale data analysis of sustainability labels for fashion in german online retail", *Frontiers in Sustainability*, vol. 3, p. 922 984, 2022.

[95] M. M. Jayaraj and A. Chandrakala, "Nudging for sustainability: Innovative practices in sustainable tourism and public policy", in *3rd International Conference on Reinventing Business Practices, Start-ups and Sustainability (ICRBSS 2023)*, Atlantis Press, 2024, pp. 973–983.

[96] M. Osman and K. Thornton, "Traffic light labelling of meals to promote sustainable consumption and healthy eating", *Appetite*, vol. 138, pp. 60–71, 2019.

[97] T. Bucher, C. Collins, M. E. Rollo, *et al.*, "Nudging consumers towards healthier choices: A systematic review of positional influences on food choice", *British Journal of Nutrition*, vol. 115, no. 12, pp. 2252–2263, 2016.

[98] L. Young, M. Rosin, Y. Jiang, *et al.*, "The effect of a shelf placement intervention on sales of healthier and less healthy breakfast cereals in supermarkets: A co-designed pilot study", *Social Science & Medicine*, vol. 266, p. 113 337, 2020.

[99] E. J. Kim, S. Tanford, and L. A. Book, "The effect of priming and customer reviews on sustainable travel behaviors", *Journal of Travel Research*, vol. 60, no. 1, pp. 86–101, 2021.

[100] S. Zimmermann, A. Hein, T. Schulz, H. Gewald, and H. Krcmar, "Digital nudging toward pro-environmental behavior: A literature review.", *PACIS*, p. 226, 2021.

[101] A. T. Schmidt and B. Engelen, "The ethics of nudging: An overview", *Philosophy compass*, vol. 15, no. 4, e12658, 2020.

[102] S. Joachim, A. R. M. Forkan, P. P. Jayaraman, A. Morshed, and N. Wickramasinghe, "A nudge-inspired ai-driven health platform for self-management of diabetes", *Sensors*, vol. 22, no. 12, p. 4620, 2022.

[103] A.-K. Jung, S. Stieglitz, T. Kissmer, M. Mirbabaie, and T. Kroll, "Click me...! the influence of clickbait on user engagement in social media and the role of digital nudging", *Plos one*, vol. 17, no. 6, e0266743, 2022.

[104] J. Feizabadi, "Machine learning demand forecasting and supply chain performance", *International Journal of Logistics Research and Applications*, vol. 25, pp. 119–142, 2020.

[105] A. Aamer, L. Eka Yani, and I. Alan Priyatna, "Data analytics in the supply chain management: Review of machine learning applications in demand forecasting", *Operations and Supply Chain Management: An International Journal*, vol. 14, no. 1, pp. 1–13, 2020.

[106] E. Spiliotis, S. Makridakis, A.-A. Semenoglou, and V. Assimakopoulos, "Comparison of statistical and machine learning methods for daily sku demand forecasting", *Operational Research*, vol. 22, no. 3, pp. 3037–3061, 2022.

[107] H. Bousqaoui, I. Slimani, and S. Achchab, "Comparative analysis of short-term demand predicting models using arima and deep learning", *International Journal of Electrical and Computer Engineering*, vol. 11, no. 4, p. 3319, 2021.

[108] M. Abolghasemi, E. Beh, G. Tarr, and R. Gerlach, "Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion", *Computers & Industrial Engineering*, vol. 142, p. 106 380, 2020.

[109] M. Mahmoudyan and A. Zeqiri, "Time series forecasting using neural networks minimizing food waste by forecasting demand in retail sales", 2021.

[110] C. Giri and Y. Chen, "Deep learning for demand forecasting in the fashion and apparel retail industry", *Forecasting*, vol. 4, no. 2, pp. 565–581, 2022.

[111] H. Maarouf, "Pragmatism as a supportive paradigm for the mixed research approach: Conceptualizing the ontological, epistemological, and axiological stances of pragmatism", *International Business Research*, vol. 12, no. 9, pp. 1–12, 2019.

[112] G. S. Parvez, O. Mufti, and M. Wahab, "Pragmatism for mixed method research at higher education level", *Business & Economic Review*, vol. 8, no. 2, pp. 67–79, 2016.

[113] V. Kaushik and C. A. Walsh, "Pragmatism as a research paradigm and its implications for social work research", *Social sciences*, vol. 8, no. 9, p. 255, 2019.

[114] M. Charter and U. Tischner, Eds., *Sustainable Solutions: Developing Products and Services for the Future.* Sheffield: Taylor and Francis Group, 2001.

[115] N. Divyapriyadharshini, "Consumer awareness towards green products and its impact", *2454-6186*, 2019.

[116] H. Vazifehdoust, M. Taleghani, F. Esmaeilpour, K. Nazari, and M. Khadang, "Purchasing green to become greener: Factors influence consumers green purchasing behavior", *Management Science Letters*, vol. 3, Sep. 2013.

[117] G. Vörösmarty, I. Dobos, and T. Tátrai, "Motivations behind sustainable purchasing", in *Environmental Management Accounting and Supply Chain Management*, R. Burritt, S. Schaltegger, M. Bennett, T. Pohjola, and M. Csutora, Eds., May 2011, pp. 41–54.

[118] H. Baumann, F. Boons, and A. Bragd, "Mapping the green product development field: Engineering, policy and business perspectives", *Journal of Cleaner Production*, vol. 10, no. 5, pp. 409–425, 2002.

[119] X. Zhang and F. Dong, "Why do consumers make green purchase decisions? insights from a systematic review", 2011.

[120] M. Laroche, J. Bergeron, and G. Barbaro-Forleo, "Targeting consumers who are willing to pay more for environmentally friendly products.", *Journal of Consumer Marketing*, vol. 18, pp. 503–520, 2001.

[121] R. Vallerand, P. Deshaies, J.-P. Cuerrier, L. Pelletier, and C. Mongeau, "Ajzen and fishbeins theory of reasoned action as applied to moral behavior: A confirmatory analysis", *Journal of Personality and Social Psychology*, vol. 62, pp. 98–109, Jan. 1992.

[122] I. Ajzen, "The theory of planned behavior", *Organizational behavior and human decision processes*, vol. 50, no. 2, pp. 179–211, 1991.

[123] J. Paul, A. Modi, and J. Patel, "Predicting green product consumption using theory of planned behavior and reasoned action", *Journal of retailing and consumer services*, vol. 29, pp. 123–134, 2016.

[124] T. Madden, P. Ellen, and I. Ajzen, "A comparison of the theory of planned behavior and the theory of reasoned action", *Personality and Social Psychology Bulletin*, vol. 18, pp. 3–9, Feb. 1992.

[125] P. Kumar and B. Ghodeswar, "Factors affecting consumers green product purchase decisions", *Marketing Intelligence & Planning*, vol. 33, pp. 330–347, Apr. 2015.

[126] H.-h, "Zhao", *Q. Gao, Y.-p. Wu, Y. Wang, and X.-d. Zhu*, vol. 63, pp. 143–151, Jan. 2014.

[127] B. Nelson, D. Taylor, and S. Strick, "Wine consumers environmental knowledge and attitudes: Influence on willingness to purchase", *International Journal of Wine Research*, vol. 1, Feb. 2009.

[128] T. Albayrak, S. Aksoy, and M. Caber, "The effect of environmental concern and scepticism on green purchase behavior", *Marketing Intelligence & Planning*, vol. 31, Jan. 2013.

[129] B. Schlegelmilch, G. M. Bohlen, and A. Diamantopoulos, "The link between green purchasing decisions and measures of environmental consciousness", *European Journal of Marketing*, vol. 30, pp. 35–55, 1996.

[130] D. Gadenne, B. Sharma, D. Kerr, and T. Smith, "The influence of consumers environmental beliefs and attitudes on energy saving behaviors", *Energy Policy*, vol. 39, pp. 7684–7694, Dec. 2011.

[131] H. Lisnawati and A. Sinaga, "Data mining with associated methods to predict consumer purchasing patterns", *International Journal of Modern Education and Computer Science (IJMECS)*, vol. 12, no. 5, pp. 16–28, 2020.

[132] S. Fedushko and T. Ustyianovych, "E-commerce customers behavior research using cohort analysis: A case study of covid-19", *J. Open Innov. Technol. Mark. Complex*, vol. 8, p. 12, 2022.

[133] A. S. Shimul, I. Cheah, and B. B. Khan, "Investigating female shoppers attitude and purchase intention toward green cosmetics in south africa", *Journal of Global Marketing*, 2022.

[134] K. Chomova, "Evolution of sustainability in marketing.", 2022.

[135] M. Charter and U. Tischner, Eds., *A step-by-step guide for beginners*, 5th. Sage Publications, 2019.

[136] A. Joshi, S. Kale, S. Chandel, and D. Pal, "Likert scale: Explored and explained", *British Journal of Applied Science and Technology*, vol. 7, pp. 396–403, Jan. 2015.

[137] J. A. Roberts, "Green consumers in the 1990s: Profile and implications for advertising", *Journal of Business Research*, vol. 36, no. 3, pp. 217–231, 1996.

[138] W. Kilbourne and G. Pickett, "How materialism affects environmental beliefs, concern, and environmentally responsible behavior", *Journal of Business Research*, vol. 61, no. 9, pp. 885–893, 2008.

[139] R. A. Statistics. "Ibm". June 27, 2022, (n.d.), [Online]. Available: `https://www.ibm.com/docs/en/spss-statistics/25.0.0?topic=analysis-reliability-statistics`.

[140] J. Pallant, "Spss survival manual: Step by step guide to data analysis using spss", 2010.

[141] L. Tay and A. Jebb, "Scale development", *The SAGE Encyclopedia of Industrial and Organizational Psychology*, vol. 2, 2016.

[142] C. Gan, H. Y. Wee, L. Ozanne, and T. Kao, "Consumers purchasing behavior towards green products in new zealand", *Innovative Marketing*, vol. 4, no. 1, pp. 93–102, 2008.

[143] R. Chan, "Determinants of chinese consumers green purchase behavior", *Psychology and Marketing*, vol. 18, pp. 389–413, Apr. 2001.

[144] T. Kyriazos and A. Stalikas, "Applied psychometrics: The steps of scale development and standardization process", *Psychology*, vol. 09, pp. 2531–2560, Jan. 2018.

[145] S. Landau and B. S. Everitt, "A handbook of statistical analyses using spss", *Chapman & Hall/CRC Press*, vol. LLC, 2003.

[146] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd. Group: Taylor & Francis, 1988.

[147] Unilever, "Private communication (email)", Feb. 2021.

[148] Nathan and Olsen, "Private communication (email)", Apr. 2021.

[149] L. Norström, S. Ganesh, M. Preeti, J. Lindman, and A. S. Islind, "Beyond social auditing: Towards self-governance and empowerment of textile workers", *45th Information Systems Research Seminar in Scandinavia, Helsinge, Denmark*, 2022.

[150] A. S. Islind and L. Norström, "Learning sustainable work through critical design: A case study of a hackathon to prepare the future workforce", *Journal of Workplace Learning*, vol. 32, no. 8, pp. 641–651, Jan. 2020.

[151] C. DSouza, M. Taghian, P. Lamb, and R. Peretiatkos, "Green products and corporate strategy: An empirical investigation", *Society and Business Review*, vol. 1, pp. 144–157, May 2006.

[152] L. M. Johri and K. Sahasakmontri, "Green marketing of cosmetics and toiletries in thailand", *Journal of Consumer Marketing*, vol. 15, pp. 265–281, 1998.

[153] P. Cao, S. Meister, and O. Klante, "How social media influence apparel purchasing behavior", *Marketing Review St. Gallen*, vol. 31, pp. 77–86, 2014.

[154] H. M. Wandabwa, M. A. Naeem, F. Mirza, and R. Pears, "Follow-back recommendations for sports bettors: A twitter-based approach.", in *HICSS*, 2020, pp. 1–10.

[155] A. Young, L. Selander, and E. Vaast, "Digital organizing for social impact: Current insights and future research avenues on collective action, social movements, and digital technologies", *Information and Organization*, vol. 29, no. 3, p. 100 257, 2019.

[156] B. Bynum Boley, V. P. Magnini, and T. L. Tuten, "Social media picture posting and souvenir purchasing behavior: Some initial findings", *Tourism Management*, vol. 37, pp. 27–30, 2013, ISSN: 0261-5177.

[157] S. Bhanot, "Use of social media by companies to reach their customers.", *SIES Journal of Management*, vol. 8, no. 1, pp. 47–55, 2012, ISSN: 09742956.

[158] T. M. Awan, M. Aziz, A. Sharif, T. R. Ch, T. Jasam, and Y. Alvi, "Fake news during the pandemic times: A systematic literature review using prisma", *Open Information Science*, vol. 6, no. 1, pp. 49–60, 2022.

[159] K. Miyazaki, T. Uchiba, K. Tanaka, J. An, H. Kwak, and K. Sasahara, "" this is fake news": Characterizing the spontaneous debunking from twitter users to covid-19 false information", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 17, 2023, pp. 650–661.

[160] P. C. Calder, "Nutrition, immunity and covid-19", *BMJ nutrition, prevention & health*, vol. 3, no. 1, pp. 74–92, 2020.

[161] A. F. Gombart, A. Pierre, and S. Maggini, "A review of micronutrients and the immune system–working in harmony to reduce the risk of infection", *Nutrients*, vol. 12, no. 1, 2020, ISSN: 2072-6643.

[162] A. Martineau, D. Jolliffe, R. Hooper, *et al.*, "Vitamin d supplementation to prevent acute respiratory tract infections: Systematic review and meta-analysis of individual participant data", *BMJ*, vol. 356, p. i6583, Feb. 2017.

[163] M. Lentjes, "The balance between food and dietary supplements in the general population", *Proceedings of the Nutrition Society*, vol. 78, pp. 1–13, Oct. 2018.

[164] J. P. de Magalhaes, M. Muller, G. Rainger, and W. Steegenga, "Fish oil supplements, longevity and aging", *Aging*, vol. 8, Aug. 2016.

[165] A. Haraldsdottir, J. E. Torfadottir, U. A. Valdimarsdottir, *et al.*, "Fish and fish-liver oil consumption in adolescence and midlife and risk of chd in older women", *Public health nutrition*, vol. 19, no. 2, pp. 318–325, 2016.

[166] D. Hathaway III, K. Pandav, M. Patel, *et al.*, "Omega 3 fatty acids and covid-19: A comprehensive review", *Infection and Chemotherapy*, vol. 52, pp. 478–495, Dec. 2020.

[167] C. Shirky, "The political power of social media: Technology, the public sphere, and political change", *Foreign affairs*, pp. 28–41, 2011.

[168] L. Lee and M. G. Ocepek, "From virtual to physical: An exploratory study on how online social networks and communities influence decision-making in everyday crafting", 2023.

[169] P. M. Leonardi, "The social media revolution: Sharing and learning in the age of leaky knowledge", *Information and Organization*, vol. 27, no. 1, pp. 47–59, 2017.

[170] H. Masuda, S. H. Han, and J. Lee, "Impacts of influencer attributes on purchase intentions in social media influencer marketing: Mediating roles of characterizations", *Technological Forecasting and Social Change*, vol. 174, p. 121 246, 2022.

[171] B. Kurdi, M. Alshurideh, I. Akour, E. Tariq, A. AlHamad, and H. Alzoubi, "The effect of social media influencers' characteristics on consumer intention and attitude toward keto products purchase intention", *International Journal of Data and Network Science*, vol. 6, no. 4, pp. 1135–1146, 2022.

[172] A. Jansom and S. Pongsakornrungsilp, "How instagram influencers affect the value perception of thai millennial followers and purchasing intention of luxury fashion for sustainable marketing", *Sustainability*, vol. 13, no. 15, p. 8572, 2021.

[173] C. Bauer and B. Ferwerda, "The effect of ingroup identification on conformity behavior in group decision-making: The flipping direction matters", in *Proceedings of the 56th Hawaii International Conference on System Sciences*, University of Hawaii, 2023, pp. 2242–2251.

[174] C. Ling, K. P. Gummadi, and S. Zannettou, “" learn the facts about covid-19": Analyzing the use of warning labels on tiktok videos”, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 17, 2023, pp. 554–565.

[175] X. Han, J. Wang, M. Zhang, and X. Wang, “Using social media to mine and analyze public opinion related to covid-19 in china”, *International journal of environmental research and public health*, vol. 17, no. 8, p. 2788, 2020.

[176] D. Baines, R. J. Elliott, *et al.*, “Defining misinformation, disinformation and malinformation: An urgent need for clarity during the covid-19 infodemic”, *Discussion papers*, vol. 20, no. 06, pp. 20–06, 2020.

[177] J. Roozenbeek, C. R. Schneider, S. Dryhurst, *et al.*, “Susceptibility to misinformation about covid-19 around the world”, *Royal Society open science*, vol. 7, no. 10, p. 201 199, 2020.

[178] S. Mehta, T. Saxena, and N. Purohit, “The new consumer behaviour paradigm amid covid-19: Permanent or transient?”, *Journal of health management*, vol. 22, no. 2, pp. 291–301, 2020.

[179] P. Jílková and P. Králová, “Digital consumer behaviour and ecommerce trends during the covid-19 crisis”, *International Advances in Economic Research*, vol. 27, no. 1, pp. 83–85, 2021.

[180] S. Sayyida, S. Hartini, S. Gunawan, and S. N. Husin, “The impact of the covid-19 pandemic on retail consumer behavior”, *Aptisi Transactions on Management (ATM)*, vol. 5, no. 1, pp. 79–88, 2021.

[181] S. Gomes and J. M. Lopes, “Evolution of the online grocery shopping experience during the covid-19 pandemic: Empiric study from portugal”, *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 17, no. 3, pp. 909–923, 2022.

[182] S. Akter, T. Ashrafi, and V. Waligo, “Changes in consumer purchasing behavior due to covid-19 pandemic”, *Changes*, vol. 77, pp. 33–46, 2021.

[183] S. Laato, A. N. Islam, A. Farooq, and A. Dhir, “Unusual purchasing behavior during the early stages of the covid-19 pandemic: The stimulus-organism-response approach”, *Journal of Retailing and Consumer Services*, vol. 57, p. 102 224, 2020.

[184] O. Tyrväinen and H. Karjaluoto, “Online grocery shopping before and during the covid-19 pandemic: A meta-analytical review”, *Telematics and Informatics*, p. 101 839, 2022.

[185] J. Fong, T. Guo, and A. Rao, “Debunking misinformation about consumer products: Effects on beliefs and purchase behavior”, *Available at SSRN 3875665*, 2021.

[186] P. Hu, M. A. Bhuiyan, M. K. Rahman, M. M. Hossain, and S. Akter, “Impact of covid-19 pandemic on consumer behavioural intention to purchase green products”, *Plos one*, vol. 17, no. 10, e0275541, 2022.

[187] D. Cook and L. Jóhannsdóttir, “Impacts, systemic risk and national response measures concerning covid-19—the island case studies of iceland and greenland”, *Sustainability*, vol. 13, no. 15, p. 8470, 2021.

[188] S. Kassem, “Amazon in the time of coronavirus”, *HesaMag*, vol. 22, pp. 14–17, 2020.

[189] S. Gudmundsson, “Online marketing of tourism companies: Cases from borgarfjörður, iceland”, Ph.D. dissertation, 2010.

[190] M. Ghersetti, J. G. Ólafsson, and S. Ólafsdóttir, “Watchdogs and government megaphones: The dual democratic roles of the news media during the covid-19 pandemic in iceland and sweden”, 2023.

[191] Y. Karavias, P. K. Narayan, and J. Westerlund, “Structural breaks in interactive effects panels and the stock market reaction to covid-19”, *Journal of Business and Economic Statistics*, 2022.

[192] S. Prabhakaran. “Augmented dickey-fuller (adf) test – must read guide - ml+. machine learning plus”. (2022), [Online]. Available: https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/.

[193] P. Muthuramu and T. Maheswari, “Tests for structural breaks in time series analysis: A review of recent development”, *Shanlax International Journal of Economics*, vol. 7, pp. 66–79, Aug. 2019.

[194] Zach. “How to calculate bic. statology”. (2021), [Online]. Available: `https://www.statology.org/bic-in-r/`.

[195] P. Mishra, C. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, “Descriptive statistics and normality tests for statistical data”, *Annals of Cardiac Anaesthesia*, vol. 22, pp. 67–72, Jan. 2019.

[196] H. Akoglu, “User’s guide to correlation coefficients”, *Turkish Journal of Emergency Medicine*, vol. 18, no. 3, pp. 91–93, 2018, ISSN: 2452-2473.

[197] A. Seth, “Granger causality”, *Scholarpedia*, vol. 2, no. 7, p. 1667, 2007.

[198] K. J. Blinowska, R. Kuś, and M. Kamiński, “Granger causality and information flow in multivariate processes”, *Physical Review E*, vol. 70, no. 5, p. 050 902, 2004.

[199] M. Maziarz, “A review of the granger-causality fallacy”, *The journal of philosophical economics: Reflections on economic and social issues*, vol. 8, no. 2, pp. 86–105, 2015.

[200] A. Shojaie and E. B. Fox, “Granger causality: A review and recent advances”, *Annual Review of Statistics and Its Application*, vol. 9, pp. 289–319, 2022.

[201] Y. Lin, Z. Hu, H. Alias, and L. P. Wong, “Influence of mass and social media on psychobehavioral responses among medical students during the downward trend of covid-19 in fujian, china: Cross-sectional study”, *J Med Internet Res*, vol. 22, no. 7, e19982, Jul. 2020, ISSN: 1438-8871.

[202] B. Keskin, “Statistical power analysis”, Sep. 2013.

[203] D. TjØstheim, “Granger-causality in multiple time series”, *Journal of Econometrics*, vol. 17, no. 2, pp. 157–176, 1981, ISSN: 0304-4076.

[204] D. I. Stern, “From correlation to granger causality”, *Crawford School Research Paper*, no. 13, 2011.

[205] B. A. Conway, K. Kenski, and D. Wang, "The rise of twitter in the political campaign: Searching for intermedia agenda-setting effects in the presidential primary", *Journal of Computer-Mediated Communication*, vol. 20, no. 4, pp. 363–380, 2015.

[206] C. Wells, D. Shah, J. Lukito, A. Pelled, J. C. Pevehouse, and J. Yang, "Trump, twitter, and news media responsiveness: A media systems approach", *New media & society*, vol. 22, no. 4, pp. 659–682, 2020.

[207] C. H. Au, K. K. Ho, and D. K. Chiu, "The role of online misinformation and fake news in ideological polarization: Barriers, catalysts, and implications", *Information Systems Frontiers*, pp. 1–24, 2021.

[208] M. Tomlein, B. Pecher, J. Simko, *et al.*, "An audit of misinformation filter bubbles on youtube: Bubble bursting and recent behavior changes", in *Proceedings of the 15th ACM Conference on Recommender Systems*, 2021, pp. 1–11.

[209] M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media", *Proceedings of the National Academy of Sciences*, vol. 118, no. 9, e2023301118, 2021.

[210] D. Freelon, M. Bossetta, C. Wells, J. Lukito, Y. Xia, and K. Adams, "Black trolls matter: Racial and ideological asymmetries in social media disinformation", *Social Science Computer Review*, vol. 40, no. 3, pp. 560–578, 2022.

[211] J. Pomerance, N. Light, and L. E. Williams, "In these uncertain times: Fake news amplifies the desires to save and spend in response to covid-19", *Journal of the Association for Consumer Research*, vol. 7, no. 1, pp. 45–53, 2022.

[212] P. Millroth and R. Frey, "Fear and anxiety in the face of covid-19: Negative dispositions towards risk and uncertainty as vulnerability factors", *Journal of Anxiety Disorders*, vol. 83, p. 102 454, 2021.

[213] L. L. Chen, S. R. Wilson, S. Lohmann, and D. V. Negraia, "What are you anxious about? examining subjects of anxiety during the covid-19 pandemic", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 17, 2023, pp. 137–148.

[214] G. Pennycook and D. G. Rand, "The psychology of fake news", *Trends in cognitive sciences*, vol. 25, no. 5, pp. 388–402, 2021.

[215] W. Akram and R. Kumar, "A study on positive and negative effects of social media on society", *International Journal of Computer Sciences and Engineering*, vol. 5, no. 10, pp. 351–354, 2017.

[216] E. Holden, K. Linnerud, and D. Banister, "The imperatives of sustainable development", *Sustainable development*, vol. 25, no. 3, pp. 213–226, 2017.

[217] V. Kalnikaitė, J. Bird, and Y. Rogers, "Decision-making in the aisles: Informing, overwhelming or nudging supermarket shoppers?", *Personal and Ubiquitous Computing*, vol. 17, no. 6, pp. 1247–1259, 2012.

[218] E. Coiera, "Putting the technical back into socio-technical systems research", *International journal of medical informatics*, vol. 76 Suppl 1, S98–103, 2007.

[219] C. Tørris and H. Mobekk, "Improving cardiovascular health through nudging healthier food choices: A systematic review", *Nutrients*, vol. 11, 2019.

[220] H. B. Slapø and K. I. Karevold, "Simple eco-labels to nudge customers toward the most environmentally friendly warm dishes: An empirical study in a cafeteria setting", *Frontiers in Sustainable Food Systems*, vol. 3, 2019.

[221] M. C. Mazza, L. Dynan, R. M. Siegel, and A. L. Tucker, "Nudging healthier choices in a hospital cafeteria: Results from a field study", *Health Promotion Practice*, vol. 19, no. 6, pp. 925–934, 2017.

[222] B. Latour, P. Mauguin, and G. Teil, "A note on socio-technical graphs", *Social Studies of Science*, vol. 22, pp. 33–57, 1992.

[223] R. H. Thaler and C. R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Penguin Books, 2008.

[224] G. Weßel, E. Altendorf, M. Schwalm, Y. Canpolat, C. Burghardt, and F. Flemisch, "Self-determined nudging: A system concept for human–machine interaction", *Cognition, Technology & Work*, vol. 21, no. 4, pp. 621–630, 2019.

[225] D. M. Hausman and B. Welch, “Debate: To nudge or not to nudge”, *Journal of Political Philosophy*, vol. 18, no. 1, pp. 123–136, 2010.

[226] V. J. V. Broers, C. De Breucker, S. Van den Broucke, and O. Luminet, “A systematic review and meta-analysis of the effectiveness of nudging to increase fruit and vegetable choice”, *European Journal of Public Health*, vol. 27, no. 5, pp. 912–920, 2017.

[227] P. DeCosta, P. Møller, M. B. Frøst, and A. Olsen, “Changing children’s eating behaviour - a review of experimental research”, *Appetite*, vol. 113, pp. 327–357, 2017.

[228] M. C. Harbers, J. W. J. Beulens, F. Rutters, *et al.*, “The effects of nudges on purchases, food choice, and energy intake or content of purchases in real-life food purchasing environments: A systematic review and evidence synthesis”, *Nutrition Journal*, vol. 19, no. 1, 2020.

[229] R. Karlsen and A. Andersen, “Recommendations with a nudge”, *Technologies*, vol. 7, no. 2, p. 45, 2019.

[230] M. Lehner, O. Mont, and E. Heiskanen, “Nudging – a promising tool for sustainable consumption behaviour?”, *Journal of Cleaner Production*, vol. 134, no. 1, pp. 166–177, 2016.

[231] E. Laiou, I. Rapti, R. Schwarzer, *et al.*, “Review: Nudge interventions to promote healthy diets and physical activity”, *Food Policy*, vol. 102, p. 102 103, 2021.

[232] M. I. Marcano-Olivier, P. J. Horne, S. Viktor, and M. Erjavec, “Using nudges to promote healthy food choices in the school dining room: A systematic review of previous investigations”, *Journal of School Health*, vol. 90, no. 2, pp. 143–157, 2019.

[233] S. Bammert, U. M. König, M. Roeglinger, and T. Wruck, *Exploring potentials of digital nudging for business processes. Business Process Management Journal.* ahead-of-print(ahead-of-print, 2020.

[234] L. Schmitz, E. Richert, M. K. Lárusdóttir, E. S. Arnardóttir, and A. S. Islind, “Nudging with dignity: A critical examination of when and how to use digital nudging”, in *The 15th Scandinavian Conference on Information Systems*, 2024.

[235] A. R. Dennis, L. (V. Y.) Yuan, X. Feng, E. Webb, and C. J. Hsieh, “Digital nudging: Numeric and semantic priming in e-commerce”, *Journal of Management Information Systems*, vol. 37, no. 1, pp. 39–65, 2020.

[236] T. Mirsch, C. Lehrer, and R. Jung, “Digital nudging: Altering user behavior in digital environments”, *Wirtschaftsinformatik Und Angewandte Informatik*, pp. 634–648, 2017.

[237] V. Kurz, “Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant”, *Journal of Environmental Economics and Management*, vol. 90, pp. 317–341, 2018.

[238] C. von Kameke and D. Fischer, “Preventing household food waste via nudging: An exploration of consumer perceptions”, *Journal of Cleaner Production*, vol. 184, pp. 32–40, 2018.

[239] H. E. Piester, C. M. DeRieux, J. Tucker, N. R. Buttrick, J. N. Galloway, and T. D. Wilson, *I’ll try the veggie burger.* Increasing purchases of sustainable foods with information about sustainability and taste, 2020.

[240] E. Valenčič, E. Beckett, C. E. Collins, B. K. Seljak, and T. Bucher, *Digital nudging in online grocery stores: A scoping review on current practices and gaps.* Trends in Food Science & Technology, 2022.

[241] M. Berger, C. Müller, and N. Nüske, “Digital nudging in online grocery stores – towards ecologically sustainable nutrition”, *ICIS*, vol. 2020 Proceedings, 2020.

[242] M. Jesse and D. Jannach, “Digital nudging with recommender systems: Survey and future directions”, *Computers in Human Behavior Reports*, vol. 3, p. 100 052, 2021.

[243] L. Peters and D. Wieland, “I tell you what you want, what you really really want: Digital information nudging as a bridge towards a sustainable purchase decision”, *AMCIS*, vol. 2023 Proceedings, 2023.

[244] R. Carroll and B. Vallen, “Compromise and attraction effects in food choice”, *International Journal of Consumer Studies*, vol. 38, no. 6, pp. 636–641, 2014.

[245] K. M. Sharpe, R. Staelin, and J. Huber, "Using extremeness aversion to fight obesity: Policy implications of context dependent demand", *Journal of Consumer Research*, vol. 35, no. 3, pp. 406–422, 2008.

[246] K. M. Kiszko, O. D. Martinez, C. Abrams, and B. Elbel, "The influence of calorie labeling on food orders and consumption: A review of the literature", *Journal of Community Health*, vol. 39, no. 6, pp. 1248–1269, 2014.

[247] M. W. Long, D. K. Tobias, A. L. Cradock, H. Batchelder, and S. L. Gortmaker, "Systematic review and meta-analysis of the impact of restaurant menu calorie labeling", *American Journal of Public Health*, vol. 105, no. 5, e11–e24, 2015.

[248] S. E. Sinclair, M. Cooper, and E. D. Mansfield, "The influence of menu labeling on calories selected or consumed: A systematic review and meta-analysis", *Journal of the Academy of Nutrition and Dietetics*, vol. 114, no. 9, pp. 1375–1388, 2014.

[249] A. C. Fernandes, R. C. Oliveira, R. P. C. Proença, C. C. Curioni, V. M. Rodrigues, and G. M. R. Fiates, "Influence of menu labeling on food choices in real-life settings: A systematic review", *Nutrition Reviews*, vol. 74, no. 8, pp. 534–548, 2016.

[250] B. Cook, J. Costa Leite, M. Rayner, S. Stoffel, E. van Rijn, and W. Jan, "Consumer interaction with sustainability labelling on food products: A narrative literature review", *Nutrients*, vol. 15, no. 17, pp. 3837–3837, 2023.

[251] J. C. Hersey, K. C. Wohlgenant, J. E. Arsenault, K. M. Kosa, and M. K. Muth, "Effects of front-ofpackage and shelf nutrition labeling systems on consumers", *Nutrition Reviews*, vol. 71, no. 1, pp. 1–14, 2013.

[252] A. Emberger-Klein and K. Menrad, "The effect of information provision on supermarket consumers' use of and preferences for carbon labels in germany", *Journal of Cleaner Production*, vol. 172, pp. 253–263, 2018.

[253] S. Ranchordás, "Nudging citizens through technology in smart cities. international review of law", *Computers & Technology*, vol. 34, no. 3, pp. 254–276, 2019.

[254] L. Rikkelman, *Nudging and Manipulation of Choice: how Perception of action and outcome changes in Time when Autonomy is restricted*. Lex Rikkelman, 2022.

[255] N. Kumar and S. Kapoor, "Do labels influence purchase decisions of food products? study of young consumers of an emerging market", *British Food Journal*, vol. 119, no. 2, pp. 218–229, 2017.

[256] A. R. Brough, J. E. B. Wilkie, J. Ma, M. S. Isaac, and D. Gal, "Is eco-friendly unmanly? the green-feminine stereotype and its effect on sustainable consumption", *Journal of Consumer Research*, vol. 43, no. 4, pp. 567–582, 2016.

[257] K. Chrzan and B. Orme, "An overview and comparison of design strategies for choice-based conjoint analysis", Sawtooth software, research paper series, 2000.

[258] F. Eggers, H. Sattler, T. Teichert, and F. Völckner, *Choice-Based Conjoint Analysis*. Springer, 2022, pp. 781–819.

[259] E. Hartman and M. Huang, "Sensitivity analysis for survey weights", *Political Analysis*, vol. 32, no. 1, pp. 1–16, 2024.

[260] P. G. Curran, "Methods for the detection of carelessly invalid responses in survey data", *Journal of Experimental Social Psychology*, vol. 66, pp. 4–19, 2016.

[261] W. D. Perreault and L. E. Leigh, "Reliability of nominal data based on qualitative judgments", *Journal of Marketing Research*, vol. 26, no. 2, pp. 135–148, 1989.

[262] A. Mallas, M. Xenos, and M. Karavasili, "Evaluating digital nudging effectiveness using alternative questionnaires design. human interface and the management of information", *Information Presentation and Visualization*, pp. 49–60, 2021.

[263] K. Lee, "Gender differences in hong kong adolescent consumers' green purchasing behavior", *Journal of Consumer Marketing*, vol. 26, no. 2, pp. 87–96, 2009.

[264] D. Neumark-Sztainer, M. Story, C. Perry, and M. A. Casey, "Factors influencing food choices of adolescents", *Journal of the American Dietetic Association*, vol. 99, no. 8, pp. 929–937, 1999.

[265] S. Thomaes, S. Grapsas, J. van de Wetering, J. Spitzer, and A. Poorthuis, "Green teens: Understanding and promoting adolescents' sustainable engagement", *One Earth*, vol. 6, no. 4, 2023.

[266] S. M. Kamaruddin, A. Misni, P. Ahmad, and R. Yaman, "Application of principal component analysis: Deriving influencing factors of teenagers' recycling practice", in *IOP Conference Series: Earth and Environmental Science, 385, 012008*, 2019.

[267] C. Alves De Castro, I. Oreilly, A. Carthy, and I. O'reilly, "The role of influencers in adolescents' consumer decision-the role of influencers in adolescents' consumer decision-making process: A sustainability approach making process: A sustainability approach the role of influencers in adolescents' consumer decision-making process: A sustainability approach", *Critical Letters in Economics and Finance*, vol. 1, no. 1, 2022.

[268] S. Dalecke and R. Karlsen, "Designing dynamic and personalized nudges", in *Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics*, 2020.

[269] H. Masaki, K. Shibata, S. Hoshino, T. Ishihama, N. Saito, and K. Yatani, "Exploring nudge designs to help adolescent sns users avoid privacy and safety threats", in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[270] E. Prits, *From Intentions to Actions: Nudging Generation Z towards Sustainable Fashion Choices through Service Design*. Www.theseus.fi, 2023.

[271] C. Kawa, W. H. Gijselaers, J. F. H. Nijhuis, and P. M. Ianiro-Dahm, "Effects of a thin body shape nudge and other determinants of adolescents' healthy and unhealthy food consumption in a school setting", *Food Quality and Preference*, vol. 96, p. 104 388, 2022.

[272] J. Guthrie, L. Mancino, and C.-T. J. Lin, "Nudging consumers toward better food choices: Policy approaches to changing food consumption behaviors", *Psychology & Marketing*, vol. 32, no. 5, pp. 501–511, 2015.

[273] G. L. Nelles and M. B. Ressler, "Youth eco-consciousness and environmentalist identity development at a summer camp", *Interdisciplinary Journal of Environmental and Science Education*, vol. 19, no. 2, e2308, 2023.

[274] H. Ahlborg, I. Ruiz-Mercado, S. Molander, and O. Masera, "Bringing technology into social-ecological systems research—motivations for a socio-technical-ecological systems approach. sustainability", *Sustainability*, 2019.

[275] O. C. Robinson, "Sampling in interview-based qualitative research: A theoretical and practical guide", *Qualitative Research in Psychology*, vol. 11, no. 1, pp. 25–41, 2014.

[276] C. Mele, T. Russo Spena, V. Kaartemo, and M. L. Marzullo, "Smart nudging: How cognitive technologies enable choice architectures for value co-creation", *Journal of Business Research*, vol. 129, 2020.

[277] P. John, S. Cotterill, L. Richardson, *et al.*, *Nudge, Nudge, Think, Think: Experimenting with Ways to Change Civic Behaviour.* A&C Black, 2013, Edited by Google Books.

[278] L. Young, M. Rosin, Y. Jiang, *et al.*, "The effect of a shelf placement intervention on sales of healthier and less healthy breakfast cereals in supermarkets: A co-designed pilot study.", *Social science medicine*, p. 113 337, 2020.

[279] R. Gale, "Environmental management accounting as a reflexive modernization strategy in cleaner production", *Journal of Cleaner Production*, vol. 14, pp. 1228–1236, 2006.

[280] E. D. Talia, M. Simeone, and D. Scarpato, "Consumer behaviour types in household food waste", *Journal of Cleaner Production*, 2019.

[281] D. Tonini, P. F. Albizzati, and T. Astrup, "Environmental impacts of food waste: Learnings and challenges from a case study on uk.", *Waste management*, vol. 76, pp. 744–766, 2018.

[282] L. Aburto and R. Weber, "Improved supply chain management based on hybrid demand forecasts", *Appl. Soft Comput.*, vol. 7, pp. 136–144, 2007.

[283] Z. Zhai, J. F. Martínez, V. Beltran, and N. L. Martínez, "Decision support systems for agriculture 4.0: Survey and challenges", *Computers and Electronics in Agriculture*, vol. 170, p. 105 256, 2020.

[284] Z. Wu, X. Chen, and Z. Gao, "Bayesian non-parametric method for decision support: Forecasting online product sales", *Decision support systems*, vol. 174, p. 114 019, 2023.

[285] H. Guo, D. Zhang, S. Liu, L. Wang, and Y. Ding, "Bitcoin price forecasting: A perspective of underlying blockchain transactions", *Decision Support Systems*, vol. 151, p. 113 650, 2021.

[286] T.-M. Choi, C.-L. Hui, N. Liu, S.-F. Ng, and Y. Yu, "Fast fashion sales forecasting with limited data and time", *Decision Support Systems*, vol. 59, pp. 84–92, 2014.

[287] N. G. Reich, J. Lessler, K. Sakrejda, S. A. Lauer, S. Iamsirithaworn, and D. A. Cummings, "Case study in evaluating time series prediction models using the relative mean absolute error", *The American Statistician*, vol. 70, no. 3, pp. 285–292, 2016.

[288] L. Frías-Paredes, F. Mallor, M. Gastón-Romeo, and T. León, "Dynamic mean absolute error as new measure for assessing forecasting errors", *Energy conversion and management*, vol. 162, pp. 176–188, 2018.

[289] S. Hao and S. Li, "A weighted mean absolute error metric for image quality assessment", in *2020 IEEE international conference on visual communications and image processing (VCIP)*, IEEE, 2020, pp. 330–333.

[290] W.-j. Niu and Z.-k. Feng, "Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management", *Sustainable Cities and Society*, vol. 64, p. 102 562, 2021.

[291] Y. Alharbi, D. Arribas-Bel, and F. Coenen, "Sustainable development goals monitoring and forecasting using time series analysis.", in *DeLTA*, 2021, pp. 123–131.

[292] K. Rodbertus, *Overproduction and crises.* S. Sonnenschein, 1898, vol. 688.

[293] F. Taghikhah, A. Voinov, and N. Shukla, "Extending the supply chain to address sustainability", *Journal of Cleaner Production*, 2019.

[294] X. Long and L. Gui, "Waste not want not? the environmental implications of quick response and upcycling", *SSRN Electronic Journal*, 2023.

[295] H. L. Lee, V. Padmanabhan, and S. Whang, "The bullwhip effect in supply chains", 1997.

[296] T. W. Chien, C. Lin, B. Tan, and W. C. Lee, "A neural networks-based approach for strategic planning", *Information & Management*, vol. 35, no. 6, pp. 357–364, 1999.

[297] P. Alexander, C. Brown, A. Arneth, J. Finnigan, D. Moran, and M. Rounsevell, "Losses, inefficiencies and waste in the global food system", *Agricultural Systems*, vol. 153, pp. 190–200, 2017.

[298] E. Rafati, "The bullwhip effect in supply chains: Review of recent development", *Journal of Future Sustainability*, vol. 2, no. 3, pp. 81–84, 2022.

[299] J. Aschemann-Witzel, I. D. Hooge, H. Rohm, *et al.*, "Key characteristics and success factors of supply chain initiatives tackling consumer-related food waste – a multiple case study", *Journal of Cleaner Production*, vol. 155, pp. 33–45, 2017.

[300] C. C. de Moraes, F. H. D. O. Costa, C. R. Pereira, A. L. da Silva, and I. Delai, "Retail food waste: Mapping causes and reduction practices", *Journal of Cleaner Production*, vol. 256, p. 120 124, 2020.

[301] Q. D. Read, S. Brown, A. Cuéllar, *et al.*, "Assessing the environmental impacts of halving food loss and waste along the food supply chain.", *The Science of the total environment*, vol. 712, p. 136 255, 2020.

[302] N. Kourentzes, J. R. Trapero, and D. K. Barrow, "Optimising forecasting models for inventory planning", *International Journal of Production Economics*, 2020.

[303] X. Bi, G. Adomavicius, W. Li, and A. Qu, "Improving sales forecasting accuracy: A tensor factorization approach with demand awareness", *ArXiv*, vol. abs/2011.03452, 2020. DOI: 10.1287/ijoc.2021.1147.

[304] S. Cisse, “Inventory management and demand forecasting improvement of a forecasting model based on artificial neural networks”, *Journal of Management Science Engineering research*, 2021.

[305] T. Nguyen, L. Zhou, A. Chong, B. Li, and X. Pu, “Predicting customer demand for remanufactured products: A data-mining approach”, *Eur. J. Oper. Res.*, vol. 281, pp. 543–558, 2020.

[306] T. E. Goltsos, A. Syntetos, C. Glock, and G. Ioannou, “Inventory - forecasting: Mind the gap”, *Eur. J. Oper. Res.*, vol. 299, pp. 397–419, 2021.

[307] G. Raghuwanshi, G. Zamare, R. Apar, R. Dabhade, and V. Wankhade, “Advanced forecasting of demandable products prices using machine learning algorithm”, *International Journal of Advanced Research in Science, Communication and Technology*, 2023.

[308] C. Chairy, C. Raharja, J. Syahrivar, and M. Ekananda, “Waste not: Selling near-expired bread in indonesia”, *International Review on Public and Nonprofit Marketing*, pp. 1–17, 2020.

[309] M. Z. Babai, J. Boylan, and B. R. Tabar, “Demand forecasting in supply chains: A review of aggregation and hierarchical approaches”, *International Journal of Production Research*, vol. 60, pp. 324–348, 2021.

[310] J. Gustavo, L. Trento, M. de Souza, *et al.*, “Green marketing in supermarkets: Conventional and digitized marketing alternatives to reduce waste”, *Journal of Cleaner Production*, vol. 296, p. 126 531, 2021.

[311] J. V. Belle, T. Guns, and W. Verbeke, “Using shared sell-through data to forecast wholesaler demand in multi-echelon supply chains”, *Eur. J. Oper. Res.*, vol. 288, pp. 466–479, 2021.

[312] R. Fildes, S. Ma, and S. Kolassa, “Retail forecasting: Research and practice”, *International Journal of Forecasting*, vol. 38, no. 4, pp. 1283–1318, 2022.

[313] Z. Michna, S. M. Disney, and P. Nielsen, “The impact of stochastic lead times on the bullwhip effect under correlated demand and moving average forecasts”, *Omega*, vol. 93, p. 102 033, 2020.

[314] M. Udenio, E. Vatamidou, and J. C. Fransoo, "Exponential smoothing forecasts: Taming the bullwhip effect when demand is seasonal", *International Journal of Production Research*, vol. 61, no. 6, pp. 1796–1813, 2023.

[315] X. Yuan, X. Zhang, and D. Zhang, "Analysis of the impact of different forecasting techniques on the inventory bullwhip effect in two parallel supply chains with a competition effect", *Journal of Engineering*, vol. 2020, no. 1, p. 2 987 218, 2020.

[316] G. Pradeep, T. D. V. Rayen, A. Pushpalatha, and P. K. Rani, "Effective crop yield prediction using gradient boosting to improve agricultural outcomes", *2023 International Conference on Networking and Communications (ICNWC)*, pp. 1–6, 2023.

[317] Q. Paletta, G. Arbod, and J. Lasenby, "Benchmarking of deep learning irradiance forecasting models from sky images–an in-depth analysis", *Solar Energy*, vol. 224, pp. 855–867, 2021.

[318] S. E. Abdel-Gadir and M. G. Mohammed, "Oman's green horizon: Steering towards sustainability through decarbonization and energy transition", 2024.

[319] M. H. L. Lee, Y. C. Ser, G. Selvachandran, *et al.*, "A comparative study of forecasting electricity consumption using machine learning models", *Mathematics*, vol. 10, no. 8, p. 1329, 2022.

[320] H. Abbasimehr, M. Shabani, and M. Yousefi, "An optimized model using lstm network for demand forecasting", *Computers & industrial engineering*, vol. 143, p. 106 435, 2020.

[321] F. Rustam, A. A. Reshi, A. Mehmood, *et al.*, "Covid-19 future forecasting using supervised machine learning models", *IEEE access*, vol. 8, pp. 101 489–101 499, 2020.

[322] R. Carbonneau, K. Laframboise, and R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting", *European journal of operational research*, vol. 184, no. 3, pp. 1140–1154, 2008.

[323] B. P. BV and M. Dakshayini, "Computational performance analysis of neural network and regression models in forecasting the societal demand for agricultural food harvests", *International Journal of Grid and High Performance Computing (IJGHPC)*, vol. 12, no. 4, pp. 35–47, 2020.

[324] Y.-T. Chen, E. W. Sun, and Y.-B. Lin, "Machine learning with parallel neural networks for analyzing and forecasting electricity demand", *Computational Economics*, vol. 56, no. 2, pp. 569–597, 2020.

[325] M. A. Maas and C. Bisognin, "Forecast of the historical series of revenues of the brazilian food industry using forecasting techniques.", 2020.

[326] H. Zou, G. Xia, F. Yang, and H. Wang, "An investigation and comparison of artificial neural network and time series models for chinese food grain price forecasting", *Neurocomputing*, vol. 70, no. 16-18, pp. 2913–2923, 2007.

[327] C. Kahraman, İ. Kaya, H. Z. Ulukan, and İ. Uçal, "Economic forecasting techniques and their applications", in *Business Intelligence in Economic Forecasting: Technologies and Techniques*, IGI Global, 2010, pp. 16–44.

[328] K. Hubrich and K. D. West, "Forecast evaluation of small nested model sets", *Journal of Applied Econometrics*, vol. 25, no. 4, pp. 574–594, 2010.

[329] M. Óskarsdóttir, C. Bravo, W. Verbeke, C. Sarraute, B. Baesens, and J. Vanthienen, "Social network analytics for churn prediction in telco: Model building, evaluation and network architecture", *Expert Systems with Applications*, vol. 85, pp. 204–220, 2017.

[330] Y. Luo, J. T. Randerson, G. Abramowitz, *et al.*, "A framework for benchmarking land models", *Biogeosciences*, vol. 9, no. 10, pp. 3857–3874, 2012.

[331] K. Frasier-Sleyman, "Benchmark your way to forecasting excellence", *The Journal of Business Forecasting*, vol. 11, no. 1, p. 6, 1992.

[332] Y. Liang, H. Wen, Y. Nie, *et al.*, "Foundation models for time series analysis: A tutorial and survey", in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 6555–6565.

[333] A. F. Ansari, L. Stella, C. Turkmen, *et al.*, "Chronos: Learning the language of time series", *arXiv preprint arXiv:2403.07815*, 2024.

[334] K. Rasul, A. Ashok, A. R. Williams, *et al.*, "Lag-llama: Towards foundation models for time series forecasting", *arXiv preprint arXiv:2310.08278*, 2023.

[335] G. Woo, C. Liu, A. Kumar, C. Xiong, S. Savarese, and D. Sahoo, "Unified training of universal time series forecasting transformers", *arXiv preprint arXiv:2402.02592*, 2024.

[336] M. Goswami, K. Szafer, A. Choudhry, Y. Cai, S. Li, and A. Dubrawski, "Moment: A family of open time-series foundation models", *arXiv preprint arXiv:2402.03885*, 2024.

[337] C. Chatfield, "The holt-winters forecasting procedure", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 27, no. 3, pp. 264–279, 1978.

[338] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance", *Climate research*, vol. 30, no. 1, pp. 79–82, 2005.

[339] T. A. Gerds, T. Cai, and M. Schumacher, "The performance of risk prediction models", *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, vol. 50, no. 4, pp. 457–479, 2008.

[340] M. Estrada, M. E. G. Camarillo, M. E. S. Parraguirre, *et al.*, "Evaluation of several error measures applied to the sales forecast system of chemicals supply enterprises", *International Journal of Business Administration*, vol. 11, p. 39, 2020.

[341] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy", *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[342] J. W. Graham, "Missing data analysis: Making it work in the real world", *Annual review of psychology*, vol. 60, no. 1, pp. 549–576, 2009.

[343] C. Faloutsos, J. Gasthaus, T. Januschowski, and Y. Wang, "Classical and contemporary approaches to big time series forecasting", in *Proceedings of the 2019 international conference on management of data*, 2019, pp. 2042–2047.

[344] M. E. Nor, H. M. Safuan, N. F. M. Shab, *et al.*, “Neural network versus classical time series forecasting models”, in *AIP Conference Proceedings*, AIP Publishing, vol. 1842, 2017.

[345] D. Barić, P. Fumić, D. Horvatić, and T. Lipic, “Benchmarking attention-based interpretability of deep learning in multivariate time series predictions”, *Entropy*, vol. 23, no. 2, p. 143, 2021.

[346] X. Shi, J. Wang, and B. Zhang, “A fuzzy time series forecasting model with both accuracy and interpretability is used to forecast wind power”, *Applied Energy*, vol. 353, p. 122 015, 2024.

[347] O. Ozyegen, I. Ilic, and M. Cevik, “Evaluation of interpretability methods for multivariate time series forecasting”, *Applied Intelligence*, pp. 1–17, 2022.

[348] E. E. Kosasih, E. Papadakis, G. Baryannis, and A. Brintrup, “A review of explainable artificial intelligence in supply chain management using neurosymbolic approaches”, *International Journal of Production Research*, vol. 62, no. 4, pp. 1510–1540, 2024.

[349] T. W. Liao and P.-C. Chang, “Impacts of forecast, inventory policy, and lead time on supply chain inventory—a numerical study”, *International Journal of Production Economics*, vol. 128, no. 2, pp. 527–537, 2010.

[350] P. Sörqvist, J. E. Marsh, M. Holmgren, R. Hulme, A. Haga, and P. B. Seager, “Effects of labeling a product eco-friendly and genetically modified: A cross-cultural comparison for estimates of taste, willingness to pay and health consequences”, *Food Quality and Preference*, vol. 50, pp. 65–70, 2016.

[351] R. Meinzen-Dick, C. Kovarik, and A. R. Quisumbing, “Gender and sustainability”, *Annual Review of Environment and Resources*, vol. 39, no. 1, pp. 29–55, 2014.

[352] G. ElHaffar, F. Durif, and L. Dubé, “Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions”, *Journal of cleaner production*, vol. 275, p. 122 556, 2020.

[353] H. V. Nguyen, C. H. Nguyen, and T. T. B. Hoang, "Green consumption: Closing the intention-behavior gap", *Sustainable Development*, vol. 27, no. 1, pp. 118–129, 2019.

[354] O. Tiboni-Oschilewski, M. Abarca, F. Santa Rosa Pierre, *et al.*, "Strengths and weaknesses of food eco-labeling: A review", *Frontiers in Nutrition*, vol. 11, p. 1 381 135, 2024.

[355] D. D. Gunawan and K.-H. Huarng, "Viral effects of social network and media on consumers' purchase intention", *Journal of business research*, vol. 68, no. 11, pp. 2237–2241, 2015.

[356] E. Aımeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: A review", *Social Network Analysis and Mining*, vol. 13, no. 1, p. 30, 2023.

[357] R. M. Jones, L. H. Somerville, J. Li, *et al.*, "Adolescent-specific patterns of behavior and neural activity during social reinforcement learning", *Cognitive, Affective, & Behavioral Neuroscience*, vol. 14, pp. 683–697, 2014.

[358] A. Y. Adewuyi, K. B. Adebayo, D. Adebayo, *et al.*, "Application of big data analytics to forecast future waste trends and inform sustainable planning", *World Journal of Advanced Research and Reviews*, vol. 23, no. 1, pp. 2469–2479, 2024.

[359] M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine learning, deep learning and statistical analysis for forecasting building energy consumption—a systematic review", *Engineering Applications of Artificial Intelligence*, vol. 115, p. 105 287, 2022.

[360] M. Rodrigues, V. Miguéis, S. Freitas, and T. Machado, "Machine learning models for short-term demand forecasting in food catering services: A solution to reduce food waste", *Journal of Cleaner Production*, vol. 435, p. 140 265, 2024.

[361] L. Feddersen, "Algorithmic transparency in forecasting support systems", *arXiv preprint arXiv:2411.00699*, 2024.

[362] K. Heinisch, C. Schult, and C. Stapper, "Transparency and forecasting: The impact of conditioning assumptions on forecast accuracy", *Applied Economics Letters*, pp. 1–5, 2024.

Appendix A

Declaration of Authorship Contribution

The table below is intended to serve as a template for how much effort was involved by the Ph.D. student in the various stages of the publication process of a research article. What is excluded in the table is what role the Ph.D. student had, i.e., whether the Ph.D. student took the lead in the project, coordinated it, acted as the driving force, handled all administration, etc., or not. The idea is, therefore, that one of the following abbreviations (ME, EE, CE, or LE) should be entered in each box. Below the tables, a brief explanation is given for each column in the table. This declaration of authorship contribution is to be submitted to the RGCS in the Computer Science department.

- ME = Main effort, includes the main effort in the indicated column.
- EE = Equal efforts, includes that there was a shared equal effort between at least one other author of the paper (this can, for instance, be the case when the work behind the paper was divided or when co-authorship has been equally divided between at least two authors).
- CE = Contributing effort, entails important effort, but there is someone else in the author list that delivered the main effort.

- LE = Learning effort, includes an effort of a learning character, for instance, by assisting with the data collection or the analysis. At least a LE is needed in all columns to fulfill the Vancouver rules for authorship.

Paper name	**Idea**	**Related work & literature**	**Data gathering**	**Research design**	**Artifact design**	**Analysis & synthesis**	**Draft**	**Administration**
Paper 1								
Paper 2								

Idea = Crystallising and formulating a clear and novel research idea alongside research question(s) or hypothesis.

Related work and literature = Reading up on the relevant literature and related work, finding the relevant references as well as putting them together in a coherent manner, alongside building up the research gap.

Data gathering = The gathering of data for the paper.

Research design = Decide on how the data gathering should be conducted (randomized clinical trial, qualitative data gathering, mixed methods, devices used for data gathering or quantitative data gathering, for instance).

Artifact design = In case there is a theoretical model, a method, a digital artifact of some sort (or any software), requirements to be tested, or an algorithm (or machine learning model) that was developed in this category would cover it.

Analysis and synthesis = The analysis of the data alongside the discussion and main contributions are drawn from the analysis.

Draft = The first finished draft of the paper.

Administration = Includes all work with the administration of the publication, such as the submissions of the multiple revisions alongside communication with editors, a major effort in writing the revision comments for the journal papers and all communication and inclusion of all authors in the various revision rounds.

Appendix B

Additional Material for Chapter 7

B.1 Forecasting Models

B.2 Nemenyi Test Results

Table B.1: Comparison between the chosen forecasting models.

	Producer	**Architecture**	**Missing Data**	**Multiple Time Series**	**GPU**	**Online Token**	**Pre-Trained Model**	**Year of Release**
Chronos [333]	Amazon	Transformers (Encoder/Decoder or Decoder-only)	✓	✗	✓	✗	✓	2024
TimeGPT garza2023TimeGPT	Nixtla	Transformers (Encoder/Decoder)	✓	✓	✗	✓	✓	2023
Lag-Llama (zero-shot) [334]	Mila-Québec AI Institute	Transformers (Decoder)	✗	✓	✗	✗	✓	2023
Moirai [335]	Salesforce AI	Transformers (Encoder)	✗	✓	✓	✗	✓	2024
Moment [336]	Independent/ Academic	Transformers (Encoder)	✓	✓	✓	✗	✓	2024
Holt-Winters [337]	Nixtla	Triple exponential smoothing	✓	✓	✗	✗	✗	1957
AutoARIMA [51]	Nixtla	ARIMA with automatic selection of parameters	✓	✓	✗	✗	✗	2010

Table B.2: Nemenyi results of forecasting performances for same forecasting horizon, without penalizing overestimation.

	Chronos	AutoARIMA	TimeGPT	Moirai	Holt-Winters	Moment	Lag-Llama Zero shot
Chronos	1	**0,005873****	**0,003905**	0,9	0,155707	0,9	0,9
AutoARIMA	**0,005873****	1	0,9	0,068583	0,9	**0,001064****	**0,001****
TimeGPT	**0,003905****	0,9	1	0,05034	0,893765	**0,001****	**0,001****
Moirai	0,9	0,068583	0,05034	1	0,572978	0,893765	0,572978
Holt-Winters	0,155707	0,9	0,893765	0,572978	1	0,05034	**0,008708****
Moment	0,9	**0,001064****	**0,001****	0,893765	0,05034	1	0,9
Lag-Llama Zero shot	0,9	**0,001****	**0,001****	0,572978	**0,008708****	0,9	1

** Significant at 0.01 level

Table B.3: Nemenyi results of forecasting performances for same forecasting horizon, penalizing overestimation.

	Chronos	AutoARIMA	TimeGPT	Moirai	Holt-Winters	Moment	Lag-Llama Zero shot
Chronos	1	**0,02607***	**0,018359***	0,9	0,155707	0,9	0,9
AutoARIMA	**0,02607***	1	0,9	0,155707	0,9	**0,04034***	**0,002565****
TimeGPT	**0,018359***	0,9	1	0,120596	0,9	**0,036541***	**0,001664****
Moirai	0,9	0,155707	0,120596	1	0,508818	0,9	0,82961
Holt-Winters	0,155707	0,9	0,9	0,508818	1	0,247891	**0,02607***
Moment	0,9	**0,04034***	**0,036541***	0,9	0,247891	1	0,9
Lag-Llama Zero shot	0,9	**0,002565****	**0,001664****	0,82961	**0,02607**	0,9	1

*Significant at 0.05 level

** Significant at 0.01 level

Table B.4: Nemenyi results of forecasting performances for same forecasting steps, without penalizing overestimation.

	Chronos	AutoARIMA	TimeGPT	Moirai	Holt-Winters	Moment	Lag-Llama Zero shot
Chronos	1	**0,008708****	**0,015317***	0,9	0,669215	0,9	0,572978
AutoARIMA	**0,008708****	1	0,9	0,198082	0,475223	**0,001****	**0,001****
TimeGPT	**0,015317***	0,9	1	0,275761	0,572978	**0,001****	**0,001****
Moirai	0,9	0,198082	0,275761	1	0,9	0,440184	0,068583
Holt-Winters	0,669215	0,475223	0,572978	0,9	1	0,175793	**0,015317***
Moment	0,9	**0,001****	**0,001****	0,440184	0,175793	1	0,9
Lag-Llama Zero shot	0,572978	**0,001****	**0,001****	0,068583	**0,015317***	0,9	1

*Significant at 0.05 level

** Significant at 0.01 level

Table B.5: Nemenyi results of forecasting performances for same forecasting steps, penalizing overestimation.

	Chronos	AutoARIMA	TimeGPT	Moirai	Holt-Winters	Moment	Lag-Llama Zero shot
Chronos	1	**0,008708****	**0,036541***	0,9	0,701291	0,9	0,440184
AutoARIMA	**0,008708****	1	0,9	0,091192	0,440184	**0,001****	**0,001****
TimeGPT	**0,036541***	0,9	1	0,247891	0,701291	**0,002565****	**0,001****
Moirai	0,9	0,091192	0,247891	1	0,9	0,701291	0,091192
Holt-Winters	0,701291	0,440184	0,701291	0,9	1	0,247891	**0,008708****
Moment	0,9	**0,001****	**0,002565***	0,701291	0,247891	1	0,893765
Lag-Llama Zero shot	0,440184	**0,001****	**0,001****	0,091192	**0,008708****	0,893765	1

*Significant at 0.05 level

** Significant at 0.01 level

Table B.6: Nemenyi results of forecasting performances for data set with missing data, without penalizing overestimation.

	Chronos	AutoARIMA	TimeGPT	Holt-Winters	Moment
Chronos	1	**0,004751****	**0,001****	0,182923	0,783655
AutoARIMA	**0,004751****	1	0,9	0,680037	0,122823
TimeGPT	**0,001****	0,9	1	0,261866	**0,016387***
Holt-Winters	0,182923	0,680037	0,261866	1	0,783655
Moment	0,783655	0,122823	**0,016387***	0,783655	1
Moment	0,9	**0,001****	**0,002565****	0,701291	0,247891
Lag-Llama Zero shot	0,440184	**0,001****	**0,001****	0,091192	**0,008708****

*Significant at 0.05 level

** Significant at 0.01 level