Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Consumer Responsiveness to Environmental Labelling:

Assessing the Impact of Colour-Coded Labels on Purchase Decisions

Name: Leon Hommerich
Student number: [redacted]
Date: 07.06.2023
Supervisor: Dr. A.E. Urai
Second reader: Dr.ir. R.E. de Kleijn
Word count: 11428
Cognitive Psychology
Thesis Msci Applied Cognitive Psychology

Abstract

This study examines the efficacy of a climate label intervention to affect consumer buying behaviour in an Amsterdam based grocery store. Three label categories were introduced, green, yellow, and red, representing low, medium, and high climate impact, respectively. Data from one month before the label implementation of the labelling system and one month after were collected and analysed. The results show no significant differences between the three label groups when analysing a change score between pre- and post-label implementation. However, when combining red and negative labels into a negative label group to raise power, we find marginally significant differences between changes in the positive label group and the negative label group, showing that the intervention may have had a marginally statistically significant negative impact on negative label product purchases. In addition to these findings, our research expresses the importance of policy changes, including climate labels, to influence climate change and explores future avenues for research.

Consumer Responsiveness to Environmental Labelling: Assessing the Impact of Colour-Coded Labels on Purchase Decisions

The Escalating Climate Change Threat

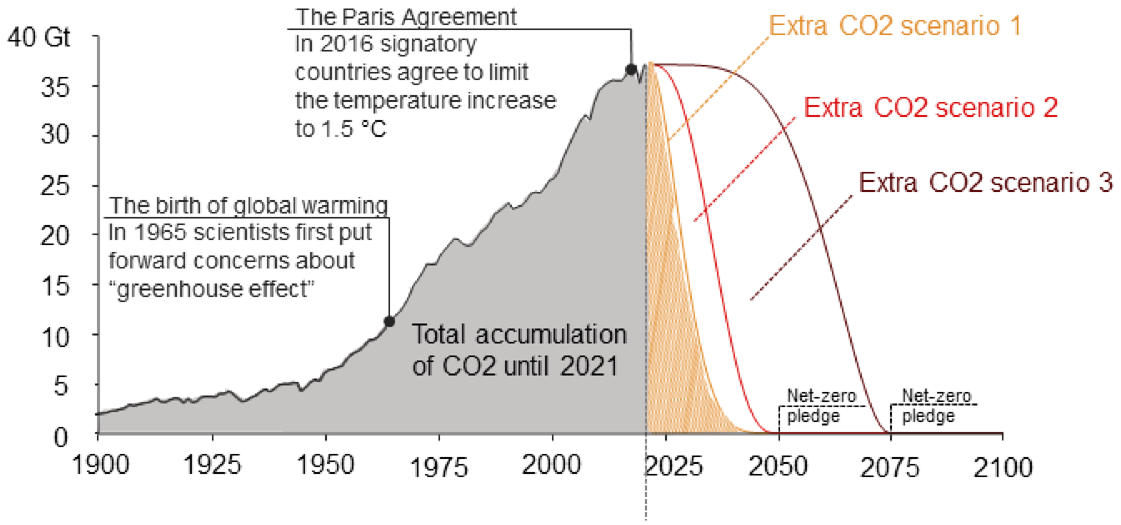
Current efforts to reduce global warming are not enough. The international community is not reaching the goals set by the Paris Agreement. Countries across the globe continue to experience growing climate change impacts (United Nations Environment Programme, 2022). One major greenhouse gas contributing sector is the production and consumption of food (Clark et al., 2020). A climate label intervention may aid emission reductions by influencing consumer choices and incentivizing corporate social responsibility (Taufique et al., 2022).

The Emissions Gap Report 2022 by the UN environment program states, that the Paris Agreement's target of keeping global warming to well below 2°C, ideally 1.5°C, is far from being met. Updated national commitments made since COP26, the 2021 United Nations Climate Change Conference, barely affect predicted emissions for 2030. Current policies point to a 2.8°C temperature rise by the end of the century (COP26, 2021; United Nations Environment Programme, 2022).

Greenhouse gas emissions must decrease as the effects of climate change become more widespread. In order to get on track for keeping global warming to 1.5°C, global annual greenhouse gas emissions must be reduced by 45% from projections under current policies in less than eight years, and they must continue to fall quickly after 2030 to prevent further anthropogenic climate warming (United Nations Environment Programme, 2022). Figure 1 shows the current net-zero scenarios that are consistent with the Paris Climate Agreement (Rogelj et al., 2019; Urai & Kruk, 2023).

Figure 1

Net-zero scenarios

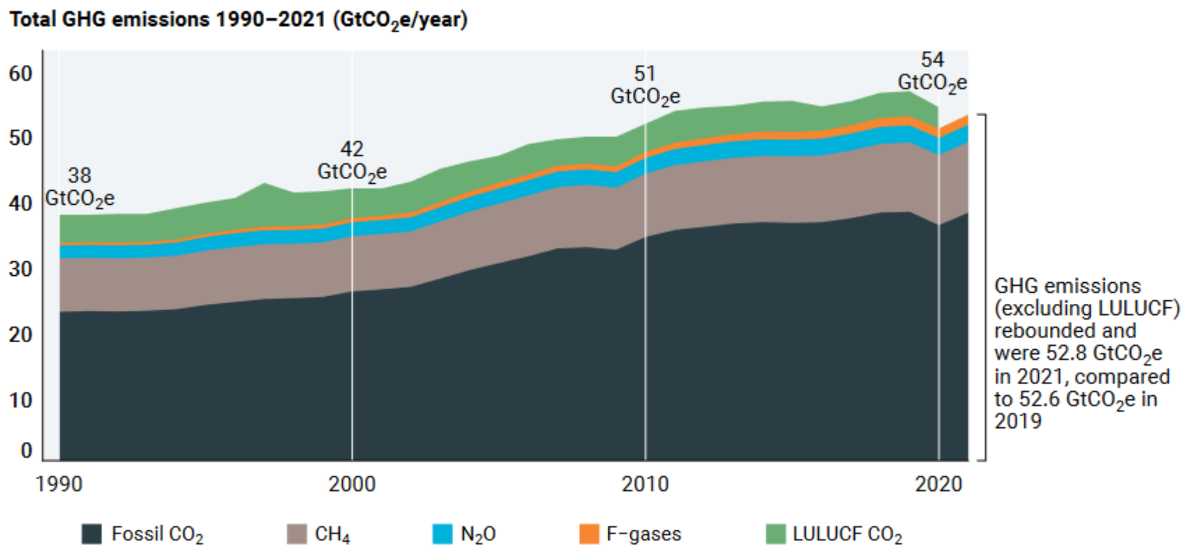


Note. From Kruk & Urai (2023). Net-zero scenarios. Zenodo, https://doi.org/10.5281/zenodo.7767084

Currently, countries appear to be off track to achieve even the highly insufficient nationally determined contributions submitted since COP26. Based on current efforts, global annual greenhouse gas emissions are estimated to be at 58 $GtCO_2e$ in 2030. The gap between the 2030 estimate and the necessary reduction to reach the 2°C warming limit is 15 $GtCO_2e$. For the 1.5°C pathway the estimated global emissions are 23 $GtCO_2e$ too high (Climate Action Tracker, 2021; den Elzen et al., 2022; United Nations Environment Programme, 2022).

Figure 2

Total GHG emissions 1990–2021 and comparison of LULUCF (Land Use, Land-Use Change, and Forestry) estimate



Note. Figure 2 illustrates the global GHG emissions trends across recent decades. It highlights a deceleration in the rate of growth in emissions from 2.6% per year (2000-2009) to 1.1% per year (2010-2019). The figure underscores a record high emission level in 2019, with total emissions averaging 54.4 GtCO2e between 2010 and 2019. Although the 2021 data is incomplete, initial estimates project emissions levels comparable to or possibly exceeding those of 2019, suggesting the potential for a new record (United Nations Environment Programme, 2022).

To achieve the significant reductions required to limit greenhouse gas emissions by 2030, a system-wide transformation must take place immediately. This transformation must occur across all sectors, including energy and power, transportation, construction, and food production and consumption (IPCC, 2021; United Nations Environment Programme, 2022).

The sector including production and consumption of food is a major cause of climate change and other environmental issues. Researchers estimate the global environmental impact of food production to be between 21% and 37%, with research papers frequently stating the impact to be about "one third" of total global greenhouse gas emissions (Crippa et al., 2021; Poore & Nemecek, 2018; Rosenzweig et al., 2020).

The environmental impact of food production goes beyond greenhouse gas pollution, perpetuating loss of biodiversity, changes in land structure and use, the depletion of freshwater supplies, and the contamination of water and land-based ecosystems due to the use of fertilizer and manure (Cordell & White, 2014; Crippa et al., 2021; Diaz & Rosenberg, 2008; Foley et al., 2005; Intergovernmental Panel on Climate Change, 2019; Robertson & Vitousek, 2009; Shiklomanov & Rodda, 2003; Wada et al., 2010; Willett et al., 2019). In this paper the we will refer to the above described sector including food production and consumption as "food systems".

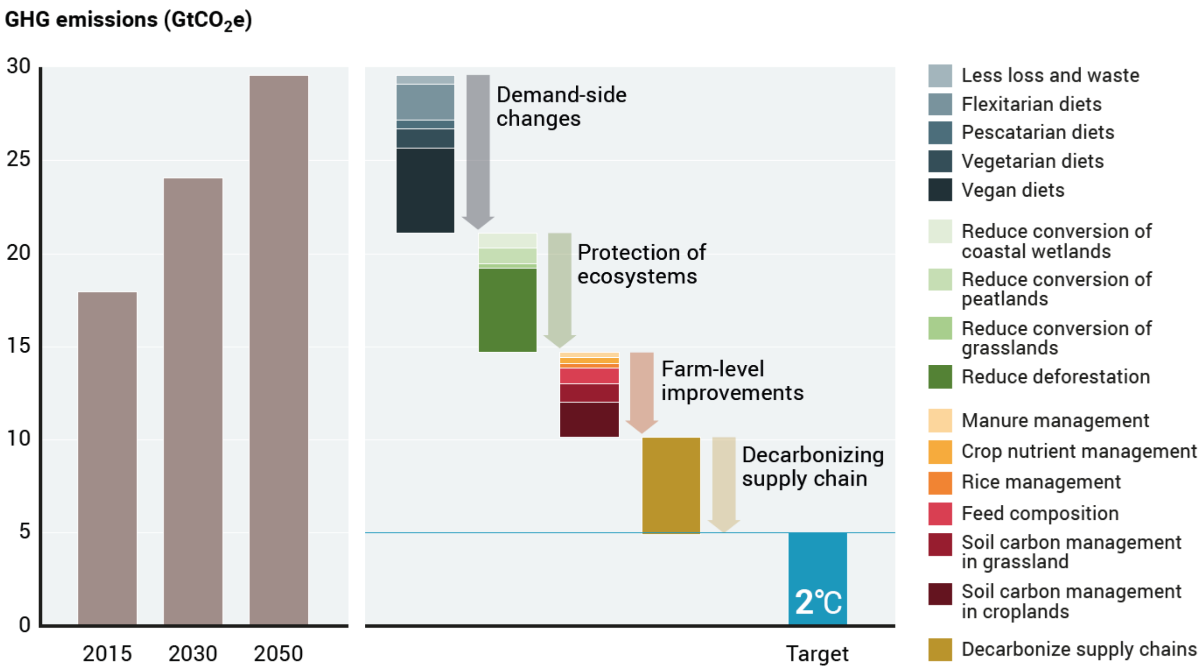
Food Systems and Food Systems Emissions

The current and even rising levels of food system emissions contradict the rapid emissions reductions required to meet the Paris Agreement goals (Clark et al., 2020; IPCC, 2022). With continuation of current dietary behaviours and agricultural practices, food consumption alone could contribute between 0.7 ± 0.2 and 0.9 ± 0.2 °C additional increase in global temperatures. Even if fossil fuel emissions were to be rapidly decreased, the emissions from the food systems alone would make it impossible to meet the target of limiting global warming to below 1.5°C and global warming will quickly approach the 2°C threshold (Clark et al., 2020; Ivanovich et al., 2023). Major changes in food systems are needed to meet the goals of the Paris Agreement (Clark et al., 2020).

As can be seen in Figure 3, 20-30% of the reduction potential to lower climate warming may come from demand-side changes that include consumer-behaviour, lifestyle modifications and other factors in end-use sectors (Creutzig et al., 2016, 2018).

Figure 3

Food systems emissions trajectory and mitigation potentials by transformation domain



Note. *Food systems emissions trajectory and mitigation potentials by transformation domain. From United Nations Environment Programme (2022). Emissions Gap Report 2022: The Closing Window — Climate crisis calls for rapid transformation of societies. Nairobi. https://www.unep.org/emissions-gap-report-2022*

The conventional IPCC category of “Agriculture, Forestry and Other Land Use” might systematically underestimate how much the agriculture and food sector really contribute to overall GHG emissions. In this paper we will use the updated categorisation of food systems, as suggested by recent research (Tubiello et al., 2021; United Nations Environment Programme, 2022).

The categorization of food systems contains: (1) Land use and land-use change, including CO_2 removals from soils, known as “Land Use and Land-Use Changes”. (2) Agricultural emissions, including emissions from animals, manure management, the burning of crop waste, agricultural soils, and indirect nitrous oxide in agriculture. (3) Food supply chain emissions, including energy usage from on-farm use, transporting, fertilizer manufacturing, packaging, and retail. Industrial processes, such as refrigeration from retail and waste, including solid food waste, waste burning and industrial and domestic water waste (Crippa et al., 2021; Intergovernmental Panel on Climate Change, 2019; Tubiello et al., 2021; United Nations Environment Programme, 2022).

According to the IPCC (2019) and a recent paper by Crippa and colleagues (2021), food systems are responsible for one third of total GHG emissions. That amounts to about 18 gigatons of CO_2 equivalent ($GtCO_2e$) per year. Agriculture production (7.1 Gt CO2e, or 39%), which includes the production of inputs like fertilizers, is the largest contributor. Land use changes (5.7 Gt CO2e, or 32%) and supply chain activities are the next two major contributors (5.2 GtCO2e, 29 per cent). The latter comprises industrial operations, packaging, waste management, fuel production, retail, transportation, and consumption.

Currently, developing countries account for 73% of emissions from food systems, yet due to their large populations, they have up to four times smaller per capita emission footprints than industrialized nations (Crippa et al. 2021)

Within agriculture production Agriculture herding of red meat, such as beef, and other animal related produces are the most problematic factors due to their high methane output and land use. In general, animal products markedly exceed the environmental impacts of vegetable substitutes and alternatives, with meat, aquaculture, eggs, and dairy using roughly 83% of the world's farmland and contributing to about 57% of food's combined emissions. That is despite animal products only providing 37% of our dietary protein and making up merely 18% of our total calories (European Commission, 2008; Poore & Nemecek, 2018; Steinfeld et al., 2006).

Without targeted interventions and if existing trends in population growth and dietary shifts toward more animal source foods continue, especially in low- and middle-income countries, food system emissions are anticipated to rise by up to 60-90 percent between 2010 and 2050 (Mbow et al., 2019; Riahi et al., 2017; Springmann et al., 2018).

Various food systems domains have been identified that need to be transformed to reduce global warming and favourably shift the climate predictions. (1) Demand-side changes are needed, including the adaptation of more sustainable diets by the world population, as well as reductions in food waste. (2) The natural ecosystems need to be protected by reducing deforestation and degradation of land for agricultural purposes. (3) Improvements in farm level food production are necessary, which should consist of modifications in the composition of animal food, as well as improved rice, manure, and crop nutrient management. (4) Finally, it is important to decarbonize the food supply chain, which can be done through CO_2 neutral retail, less polluting transportation, changing fuel use, improved industrial processes and waste management, along with more climate friendly packaging (Clark et al., 2020; IPCC, 2022; Roe et al., 2019; Springmann et al., 2018; United Nations Environment Programme, 2022).

The current study will focus on exploring demand-side changes through the application of climate labelling. Our research will add to the existing research on the efficacy of a climate label intervention on retail store consumer buying behaviour.

Climate Labelling to Induce Demand-Side Changes and Improve the Supply Chain

The introduction of climate labelling on food products could incentivize the decarbonization and efficiency of the food supply chain. Moreover, climate labels may help to influence consumer choices to encourage more climate conscious shopping (Vandenbergh et al., 2011). Both may occur at the same time, with changes in consumer buying behaviour forcing producers, retailers and intermediaries to adopt lower carbon production methods, while lower climate labels on the food products may act as positive brand advertising to the consumer (Röös & Tjärnemo, 2011; Taufique et al., 2022). Some research has shown that environmental labelling can help shift corporate behaviour even if consumer responses are moderate (Bullock, 2017; Darnall & Aragón-Correa, 2014; Kitzmueller & Shimshack, 2012; van der Ven, 2019).

Labelling interventions are currently implemented either through private organisations or through policymakers, and by extension governments, as, for example, can be seen in Japan. Japan has created their own labelling standards which offers a uniform labelling method and effectively regulates how the label data is sourced and calculated (Liu et al., 2016; Shi, 2010).

Researchers suggest that next to the current strong focus on technological advances and restructuring of existing systems and methodologies in transportation, food production, housing and other sectors, more attention should be paid to the role of demand-side approaches through policy changes as they can support the goal of limiting global warming to 1.5°C (Creutzig et al., 2016, 2018; Mundaca et al., 2019; Riahi et al., 2015).

In the food sector, interventions enforced by either governments, non-government organizations or for profit companies, may include (1) public awareness campaigns that can educate the public about the environmental consequences of their purchase decisions, (2) marketing regulations that ensure that companies accurately communicate about the environmental impact of their products and services, (3) pricing incentives, such as discounts or rebates, for low-carbon products, and (4) taxes and fees on high-emission products that may discourage consumption of such and promote sustainable alternatives (Caillavet et al., 2019; Creutzig et al., 2016, 2022; Edjabou & Smed, 2013; García-Muros et al., 2017; Harris et al., 2009; Macdiarmid et al., 2016).

While the above policy interventions may all have their own merits and drawbacks, this paper will focus on the exploration of climate labels as a low-cost intervention that can be part of a cluster of political changes that may help society to adapt more sustainable consumption practices. We focus on climate labelling specifically as it empowers the consumer to make informed decisions based on their personal values and sustainability goals. Labelling systems may serve an educational purpose, raising awareness on the climate impact of foods and food choices. Additionally, climate labelling is flexible and scalable. The labels can easily be updated in accordance with the newest findings on how to best design labels and what consumers are after when looking for climate information.

With the above in mind, it must be said that climate labelling also comes with its own challenges. The success of a labelling interventions depends on factors, such as the accuracy of the labelling, the visibility and ease of understanding of the labels, in addition to the potential for information overload, as well as the consumer trust in the labelling system. There is also the question of who should pay for the label, which depends on whether the label is driven by producers, individual food retailers, or the government, all of which may have different goals for the label initiative. In general, food choices are often deeply rooted in cultural and geographic traditions and eating habits. Changing such behaviour will always be a difficult task (Kumar & Ghodeswar, 2015; Röös & Tjärnemo, 2011).

Global changes across all environmentally relevant systems are necessary to change our current direction of global warming. However, climate labelling systems represent one tool that can help to influence consumer choices, hold companies accountable for their environmental footprint, and with that, may bridge some of the time needed to make government driven policy changes (*Advancing the Science of Climate Change*, 2010; *Limiting the Magnitude of Future Climate Change*, 2010; Vandenbergh et al., 2011).

Several studies demonstrate that customers take the environment into consideration when making purchases. The sustainability of products is a key issue for some consumers, for others it is one point of importance among others, and for some it does not matter at all (Kumar & Ghodeswar, 2015; Mainieri et al., 1997; Stern, 1999; Taufique et al., 2017). However, a sizable portion of the populace in many nations is driven by environmental concerns, and in the absence of national and international legislative action, we may capitalize on their preference for low-carbon products (Leiserowitz, 2007).

Despite the environmental consciousness of a significant proportion of the population, and potential advantages of carbon labelling, the real-world application presents various challenges. In the subsequent section, we discuss the difference between climate labelling and carbon labelling and explore the progress made in implementing carbon labels worldwide.

Difference Between Climate Labelling and Carbon Labelling

Carbon labels focus mostly on greenhouse gas (GHG) emissions associated with a product or service. They highlight the total GHG emissions throughout a product or service's life cycle - from raw material extraction to production, distribution, use, and disposal. They often include a specific CO2 equivalent value, providing a precise measurement of the product's carbon footprint. This direct representation of carbon emissions helps consumers make informed decisions based on the carbon impact of their purchases.

On the other hand, eco-labels and climate labels are products of the broader ecological footprint concept, which encompasses multiple aspects of environmental impact. They provide a comprehensive view of a product's environmental footprint, including factors such as water usage, air and water pollution, habitat destruction, and waste generation. Eco-labels aim to inform and motivate more environmentally friendly purchasing behaviours by giving consumers a wide perspective on the environmental impacts associated with their choices.

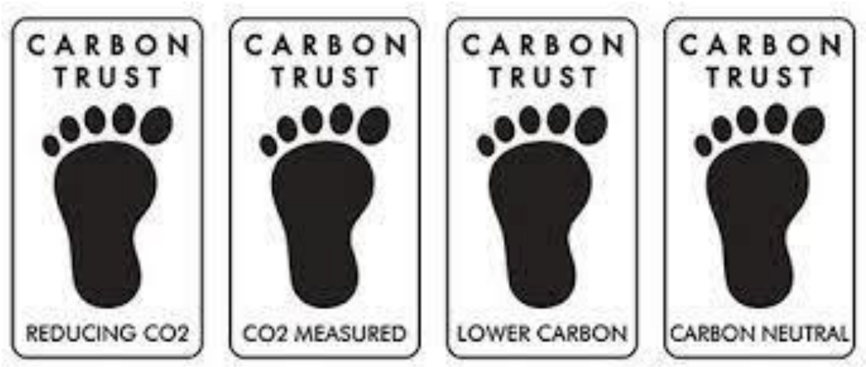
While both types of labels aim to promote more sustainable consumption patterns, they differ in their focus of environmental coverage. Carbon labels provide a specific focus on carbon emissions, while eco-labels offer a wider view of various environmental impacts (Liu et al., 2016).

Current Labelling Efforts

Climate labelling techniques have been developed for various products, with 31 carbon footprint labels being part of the 455 eco-labels listed by the Ecolabel Index across 25 sectors and 199 countries(Liu et al., 2016). The Carbon Trust has labelled tens of thousands of items, including all types of products and services. Some of the first carbon labelling initiatives were taken by large European retailers like Tesco, Casino, E.Leclerc, and RAISIO, labelling thousands of products (Boardman, 2008; Liu et al., 2016; Schaefer & Blanke, 2014). However, not all of these efforts remain active. Tesco, for example, had to abandon their plan to label all 70,000 products due to high costs (Taufique et al., 2022; Vaughan, 2012).

Figure 4

Carbon Trust Labels

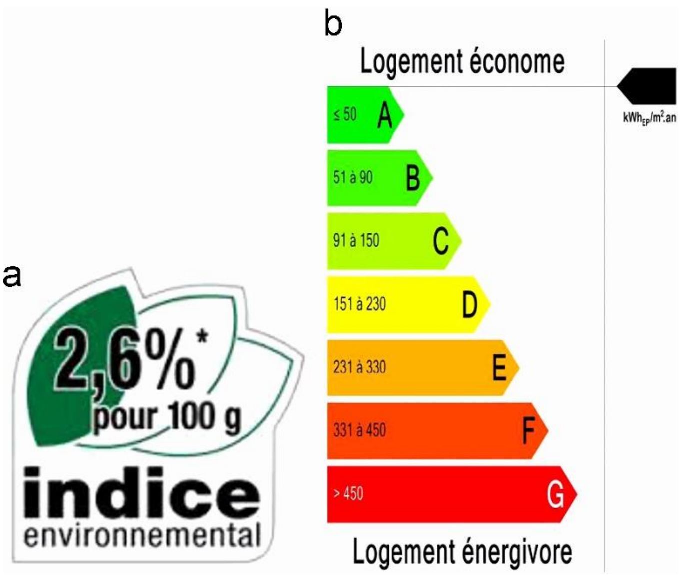


Note. There are various labels and certifications issued from Carbon Trust. More under: https://www.carbontrust.com/what-we-do/assurance-and-labelling/product-carbon-footprint-label (Carbon Trust, 2020)

Several developed and some less-developed countries have implemented carbon labelling on a nationwide scale. In 2006, the UK-based non-governmental organization, the Carbon Trust, introduced two types of carbon footprint labels - a CO2 Measured Label and a Reducing CO2 Label - positioning the country as a leader in the implementation of such labelling in the market (Liu et al., 2016).

Figure 5

Carbon labels in France. (a) Indice carbone. (b) Bilan carbone.



Note. As seen in Liu, T., Wang, Q., & Su, B. (2016). A review of carbon labeling: Standards, implementation, and impact. Renewable and Sustainable Energy Reviews, 53, 68–79. https://doi.org/10.1016/j.rser.2015.08.050

In 2008, the introduction of numeric carbon dioxide values on product prices or receipts within supermarkets marked the initiation of carbon labelling in France, with both Casino and Leclerc leading the way. In 2010, the "Grenelle 2" law established legal requirements for carbon labels, making France the pioneer in codifying carbon footprint labelling. Initially implemented as a voluntary environmental labelling scheme for all consumer goods sold in, it officially became compulsory for specific product categories in July 2012 (Djama, 2011; Liu et al., 2016).

Figure 6

Carbon Label in Japan

Note. Retrieved from Ministry of Economy, Trade and Industry. (n.d.). Carbon Footprint of Products in Japan. Retrieved 8 July 2023, from https://www.meti.go.jp/english/policy/energy_environment/cfp/pdf/cfp_products.pdf

In Japan, the government plays a crucial role in the implementation of carbon labelling schemes, serving as a regulator, while third parties are responsible for certification and labelling of goods. The country has established its own carbon labelling standards, called TSQ001 that provide a uniform national method for labelling carbon emissions, effectively regulating carbon label accounting methods within the country (Liu et al., 2016; Rugani et al., 2013; Shi, 2010). During our research we were unable to find data on the effectiveness of Japan's carbon label implementation.

While the worldwide implementation of carbon labels has been met with varying degrees of success, it is important to consider the potential opportunities that carbon labelling can offer. Opportunities may encompass supporting a demand-side change towards more green buying behaviours and business opportunities for sellers.

Carbon Labels as a Business Opportunity?

Businesses can benefit from carbon labelling. Some of the current research shows that consumers across Europe are willing to spend 20% more money on a food product that displays an explicitly climate friendly carbon label (Feucht & Zander, 2018a).

A recent study on Belgian consumers has shown that people's willingness to pay (WTP) for free range claims and animal welfare on meat products may be higher than for organic and carbon labels. However, consumers were still willing to pay a smaller price premium on carbon labelled products (Van Loo et al., 2014). Additionally, as stated by a Eurobarometer poll on sustainable consumption and production, 72% of a sample of EU citizens think that in the future, it should be required for products to carry a label that displays its carbon footprint (Eurobarometer, 2009).

Trust in Carbon Labels and Climate Education

While consumer demand for carbon labels is strong, interpretation of the stated carbon statistics on the label may be confusing. An increasing number of studies show that consumers want more information on the climate impact of a product even though their willingness to change purchasing behaviours remains low (Birkenberg et al., 2021; Gadema & Oglethorpe, 2011; Hartikainen et al., 2014; Upham et al., 2011).

Additionally, consumers seem to distrust the labels if they are unknown or not enough information is given on how the climate statistics are obtained, further limiting the potential impact of the provided information (Feucht & Zander, 2018a; Taufique et al., 2017).

Combining enough information on a label while not overwhelming the consumer at the point of sales (POS) remains a challenge for private labelling organisations and policy makers (Hornibrook et al., 2015; Meyerding et al., 2019).

Does Label Design Matter?

The aim of a carbon footprint label is to mitigate the information gap between producers and consumers by providing a quick and easily understandable assessment of a product's climate impact. This helps consumers make informed purchasing decisions, as they cannot independently verify the environmental attributes of the product at the point of sale (Meyerding et al., 2019; Weinrich & Spiller, 2016).

Studies on optimizing the design of carbon footprint labels have concluded that making key information more easily understandable is crucial in influencing consumer decisions. The findings indicate that for some label designs the key climate information may not be easily comprehended by the average consumer and highlights the challenge of balancing the need for detailed information with the risk of overwhelming consumers at the point of sale (Hartikainen et al., 2014; Thøgersen & Nielsen, 2016).

Several studies have implied that working with colours to indicate relative carbon footprint significantly increases the label effectiveness. It was found that a three-colour traffic light ranking system most effectively increased the impact of carbon labelling on consumer choices. Future labelling efforts should not only rely on simple labels with logos but include colour supported categorical information (Meyerding et al., 2019; Thøgersen & Nielsen, 2016).

The use of color-coding enhances the comparability of products within the same category and makes the information of the carbon footprint label more easily comprehended. Findings additionally suggest that negative labelling is more likely to influence consumer behaviour than positive labelling, and should be taken into consideration when designing a new carbon footprint label (Meyerding et al., 2019).

Our Carbon Labelling Approach

Our labelling approach used in this study builds on previous research on label design. It includes a three-tier traffic light system, as well as a relative measure of carbon footprint that we deemed more easily understandable than pure numerical GHG emissions. A new design has already been informed and realized during the design of this study that iterated on the current design. The current study however used the design as can be seen in Figure 7.

Research Gap

In their 2022 review of the current state of carbon climate labels, Taufique and colleagues note that most studies were either conducted in artificial lab settings using hypothetical choice experiments or in small scale field experiments, for example in canteens and restaurants. Existing studies typically focus on one product category, such as meats or vegetables (Taufique et al., 2022).

Additionally, most current studies evaluated labelling effects with a self-reported willingness to pay or purchase intentions. While self-reported willingness to pay or purchase intentions may provide useful data, they do not necessarily translate into actual purchasing behaviour. The so called “intention-behaviour” gap may be especially prominent in contexts such as environmental intentions. People often express pro-environmental intentions but fail to act on them. Therefore, relying on self-reported measures might overestimate the effect of labels on actual consumer behaviour in those studies (Carrington et al., 2010; Hassan et al., 2016).

Artificial lab settings and small scale-scale field experiments are not always generalizable to a real-world setting. Only evaluating one category prevents researchers from assessing substitution and spillover effects between other categories.

Study Aim

The main aim of this study is to investigate whether carbon labels on food products can be a successful intervention to elicit behavioural change in consumers. We will compare a store's sales before and after the implementation of carbon labels. This study will compare actual transactions instead of only measuring the consumers' willingness to pay. For this we collaborate with an Amsterdam based store: Little Plant Pantry (LPP).

Collaborations

Little Plant Pantry

Little Plant Pantry is a wholefood store focusing on ethical consumerism and minimising package waste. The store offers a wide range of healthy natural products, ranging from raw beans, lentils, nuts and seeds to herbs and spices and fresh homemade dishes (LittlePlantPantry, 2021).

They sell over 200 products and will incrementally apply carbon labels to the shelfs and packages within the store. The labels are provided by GreenSwapp.

GreenSwapp

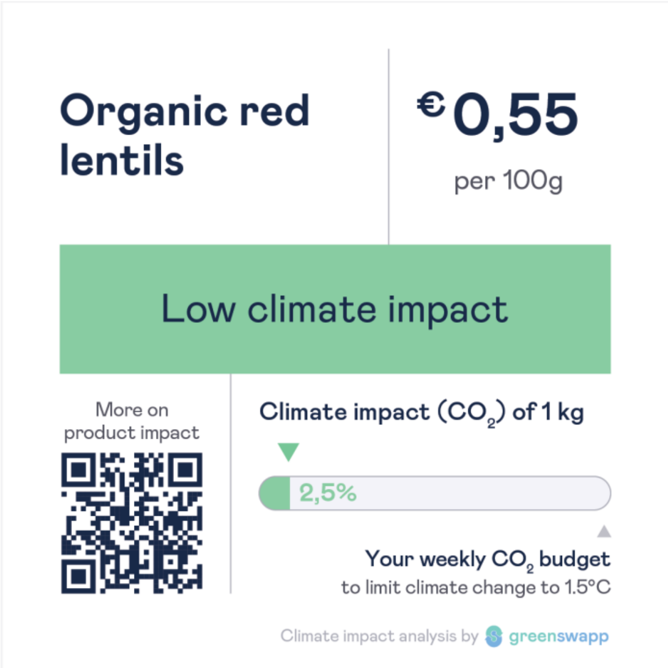
We partner with GreenSwapp to create and provide the carbon labels LPP applies. GreenSwapp, based in Amsterdam, provides an impact tracking platform where companies can track, reduce and offset the carbon footprint of their food products (GreenSwapp, 2022).

They offer carbon labels for stores and companies to communicate their products' carbon footprint to their customers. See Figure 7.

Label Design

Figure 7

Example of Current GreenSwapp Climate Impact Label



The colour on the label indicates one of three categories: Low, medium, or high impact. GreenSwapp uses life-cycle-analysis (LCA) to calculate the GHG emissions of food products. GreenSwapp follows the ISO 14040 standard for its LCA (International Organization for Standardization, 2014).

Depending on the amount of GHG emissions a food product generates it will be put in one of the categories with green indicating a low carbon footprint and red a high one.

Life-Cycle-Analysis (LCA)

LCA is a process to evaluate the environmental impact of a product, process, or activity. It tries to assess all direct and indirect factors that contribute to a potential environmental impact across the full life cycle of a product, including material acquisition, manufacturing, use and final disposal or reuse (Brusseau, 2019).

Products are categorized with a color-coded system - green, yellow, or red - based on the proportion of a consumer's weekly carbon budget they use. Green-labelled products account for up to 25% of the weekly carbon budget, indicating they have a lower environmental impact. Yellow-labelled products consume more than 25% but less than 35% of the weekly carbon budget, suggesting a moderate environmental impact. In contrast, red-labelled products utilize 35% or more of the weekly carbon budget, indicating they have a high environmental impact (GreenSwapp, 2022).

With the help of GreenSwapp and LPP we will be able to analyse how carbon labels will influence overall sales and GHG emissions.

Hypotheses

H1: Products with a green carbon label will have higher sales after the label is introduced

This expectation is based on the findings by (Feucht & Zander, 2018b) who postulate that consumers are willing to pay a 20% price premium for food products with a carbon label that explicitly states that the product is an eco-friendly choice.

H2: Products with a red or orange label will have lower sales after the label is introduced

This expectation can be substantiated with the findings by (Brunner et al., 2018), who applied carbon labels to the products in an University restaurant and found that the sales of red labelled dishes were reduced by 4.8%.

Methods

Design

The current study was designed to combine elements of experimental and longitudinal research in a real-world setting. We measured the shopping behaviour of Little Plant Pantry grocery store customers over time before and after the intervention of climate labels. The outcome on our dependent sales variable may not fully be a result of the intervention, due to the non-laboratory setting. The design of this study can be described as semi-experimental longitudinal field research.

We deliberately chose this design as it may provide valuable insights into complex social behaviours outside of a lab setting. We changed the independent variable of climate labels to observe changes in the dependent variable of sales. The independent variable was either having climate labels or not having climate labels. The labels themselves were split into three categories based on GHG emissions per product: (1) low impact with a green label, (2) medium impact with a yellow label, and (3) high impact with a red label. With this the labels of the products represent the GHG emissions a consumer would encourage by buying the labelled product, ranging from low emissions to high emissions.

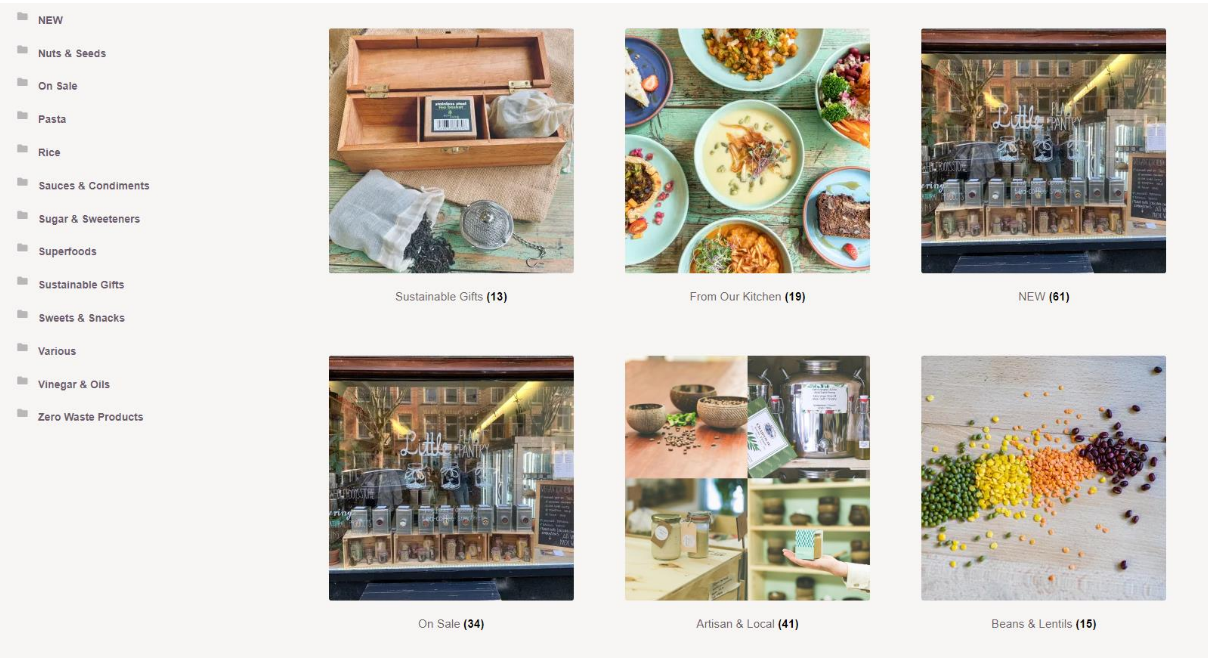
The dependent variable we were observing is the sales of the product. We expected the sale of high impact products to slightly decrease when red labels were present as compared to no labels. We also expected the sales of low impact products to slightly increase after the implementation of green labels.

We would like to reiterate that as a semi-experimental field research, it will be hard to control for confounders that arise from various sources, such as the level of climate change awareness per subject or the economic changes of the market. For that reason, only limited causality can be inferred from the impact of the climate label intervention on the product sales.

For the data analysis we obtained full access to the anonymous sales data of the LittlePlantPantry grocery store. With this we were able to gather one year of historical data on sales pre-label and measured one year of intervention data post-label.

Figure 8

Screenshot of the Little Plant Pantry Online Store



Note. A screenshot showing the overview provided on the Little Plant Pantry webstore. Vegan Organic Plastic-Free Shop. (2023). Little Plant Pantry. Retrieved 3 July 2023, from https://littleplantpantry.com/shop/

The Little Plant Pantry store offered over 600 different products in 2021 and 2022. This includes a wide range of products, ranging from raw lentils, through lower emission chocolate, over to in house self-made vegan dishes like tiramisus and brownies. Their website provides an overview of all the products they currently have on offer (LittlePlantPantry, 2021).

182 of these products were assigned a label. The assignation to certain products did not have any influence from this study but was based on decisions of the shop owner. 166 labels were low impact and green, 9 labels were medium impact and yellow, and 7 labels were high impact and red. This label distribution is heavily skewed towards green labels. That is because the store is already aiming to provide sustainable and low-carbon products as the business premise. Hence, a majority of the products are low-impact and as such the label distribution represents the impact distribution across all products of the store.

As a demo, some labels were applied on 29.01.2022, others followed on 19.03.2022 and the majority and rest of the labels were applied on 10.04.2022. These different times of label implementation were statistically corrected for in the analysis of this data.

Participants

The data we were able to access was fully anonymous, not including any direct personal information, such as the name of the customer. However, the Zettle payment system the store used for recording transaction data includes a repeat customer statistic based on the encrypted debit and credit card numbers used at the point of sale. This allows us to differentiate the number of repeat customers from the number of unique customers. Assuming that repeat customers visit the shop every week, 4 times a month, we can estimate the number of unique customers. In the year 2021 we estimate 5954 unique customers and in the year 2022 we estimate 5830 unique customers with a total of 11784 unique customers in our data not accounting for repeat customer overlap between the years.

The participants of this study were not recruited by us. We analysed the full dataset of the Little Plant Pantry store including every transaction without exclusion criteria. We were not able to assess any specific characteristics of the customers. As such, our sample consists of every customer in the Little Plant Pantry store without exclusions. The data was fully anonymous.

Materials

Figure 9

GreenSwapp climate label applied to products in the Little Plant Pantry store

This study made use of climate labels consisting of three different colours, green, yellow and red. The labels included the name of the product, the pricing, the colour based on the product's emissions, and an estimate of how much of the customers weekly CO^2 budget this product would take up. The latter being a calculation from GreenSwapp.

Figure 10

High Climate Impact Label (Red) As Implemented in the Little Plant Pantry Store

Figure 11

GreenSwapp Information Flyer That Was Hung at the Entrance of the Little Plant Pantry Store

Measures

We accessed historical sales data before the label implementation and measured the number of sales on labelled products after the label implementation via the Zettle payment system (Zettle, 2023).

Procedure

The participants were not guided in any way in this study. The Little Plant Pantry customers simply went about their normal shopping. After the inclusion of the labels, the customers were able to find an information flyer about the labels at the beginning of the store and they encountered the climate labels on the products themselves.

Data Analysis

The data for this study was gathered with the help of the Zettle by PayPal payment system at the point of sale. Zettle was the system used by the LittlePlantPantry store to manage their sales and transaction data of the customers. We were able to download the data from the web interface of the payment system for the years 2021 and 2022 in .xlsx Excel format.

The initial data from the payment system was not usable for further data analysis. Due to LittlePlantPantry being a small store being run by two people, their limited time was understandably not focused on keeping the data clean and to make it available in an orderly fashion for other people to access. In fact, they never expected anyone to analyse their data before our research. For that reason, the data needed a substantial amount of data pre-processing. The data processing, restructuring, plotting and the analyses were done using Python with the libraries Pandas, Numpy, Matplotlib, Seaborn, Scipy and Pingouin (*Matplotlib/Matplotlib*, 2011/2023; *Numpy/Numpy*, 2010/2023; *Scipy/Scipy*, 2011/2023; The pandas development team, 2010/2023; Vallat, 2018/2023; Waskom, 2012/2023).

In the following paragraphs we will outline the data pre-processing process. To allow for the exact reproduction of our findings and to clearly see what has been done to the data and how, we published the Python based Jupyter Notebooks on GitHub: https://github.com/Carragos/EnvironmentProject

The code can be used under the GNU General Public License v3.0. We utilized three distinct notebooks for various stages of our data processing, which included data importing, data cleaning, and data analysis. In the data analysis, we generated plots and calculated both descriptive and inferential statistics.

Data Import and Pre-processing

The raw transaction data, as obtained from the Zettle payment system for 2021 and 2022, were processed using Python's Pandas library. We created two-dimensional tabular data structures, or dataframes, from the raw data. These dataframes were concatenated to form one comprehensive dataset.

We continued to standardise the "unit" column, removing spelling mistakes from the initial data. Additionally, we numerically enumerated the transactions.

A pivotal part of the pre-processing stage involved the integration of the CO2 product data from GreenSwapp. To reconcile the differences between the product names in the GreenSwapp data and the LittlePlantPantry store data, a manual matching process was undertaken to create unique identifiers for each product. This process included refining the product name column by eliminating unnecessary characters and white spaces, and correcting spelling errors. For instance, similar products with minor naming variations, like "Sea Salt Coarse" and "Sea Salt Fine", were unified under a common name – "Sea Salt".

However, during this process, it was identified that some products had multiple names in the database, despite representing the same unique store product. When these anomalies could not be resolved, such products, including “Sauercrowd Jar”, “Sauercrowd Golden turmeric”, and “Sauercrowd Purple rain kraut”, were excluded from the dataset.

Notably, considerable discrepancies emerged between the number of unique products in the 2021 data and the 2022 data. Collaboration with the Little Plant Pantry owners revealed significant changes in product naming during the study period, introducing multiple duplicates and slightly altered product names. Despite attempts to solve these duplicate naming issues using a 'fuzzy matching' technique, satisfactory results were not achieved due to the ambiguity in name matches. Consequently, the scope of the analysis was narrowed to encompass only one month prior to, and following, the label implementation. A timeframe where names were not changed yet.

Critical to the analysis was the standardisation of the label implementation time. This was achieved by creating a “TimeDeltaDays” column, which represented the time difference between the date of transaction and the label implementation date.

Considering these pre-processing steps, our final analysis focused on products which remained consistent during the data collection period. This approach resulted in a dataset comprising 184 products, each marked with green, orange, or red labels based on their CO2 impact. Future research may improve the name matching procedure, potentially increasing the studied timeframe and number of products for analysis.

Data Analyses and Plotting

We conducted a thorough examination of the data, producing descriptive and inferential statistics, as well as numerous visualizations. For simplicity, this section concentrates solely on analyses crucial to our study's outcomes. Complete details and additional analyses can be found with our code and data on our GitHub page: https://github.com/Carragos/EnvironmentProject

Our analysis aimed to quantify the impact of a climate label intervention on transaction numbers and product sales volume over one month pre- and post-intervention. By examining changes in these measures, we aimed to infer the efficacy of the intervention and its potential to reduce food products' CO2 impact.

For our assessment, we plotted the sum of the number of transactions for each label colour (green, yellow, red) for the month before the labels were implemented and for the month after the labels have been implemented (Figure 14). As the different label categories included a greatly varying number of labels, we plotted the same graph again but this time with adjusting for the number of labels in each category by dividing the sum of product amount sold by the number of labels in the respective category.

- T_i represents the total amount of product sold for a particular label category (like green, yellow, or red).
- N_i represents how many labels there are in that category
- P_i is the average product amount sold per label in the category, which we get by dividing T_i by N_i

$$P_i = \frac{T_i}{N_i}$$

We repeated the same plots for the sum of transactions: one plot simply showing the sum of transactions before and after the label intervention, and the other one showing the sum but adjusted by the number of labels in the respective label category.

Again, the store implemented 166 green labels, 9 yellow labels, and 7 red labels. However, within the timeframe of our analysis, 114 unique green labelled products were sold in the month before, and 118 unique green labelled products were sold in the month after. Yellow labels were recorded with 6 unique products in both months and red labels were recorded with 5 unique products in both months.

To better understand the impact of the label intervention on our data, we computed a 'change score' for each product in our dataset. This change score represents the absolute variation in transaction numbers for each product category from pre to post-intervention. It is calculated by subtracting the number of transactions prior to the intervention from those following it.

However, considering that some products have a naturally higher transaction frequency than others (e.g., tomatoes are purchased more often than shampoo), it was crucial to account for this inherent variability. Therefore, we focused on changes that could be specifically attributed to the intervention, rather than these baseline differences in transaction frequency.

To achieve this, we added a 'percentage change' column to our dataset. This was computed by dividing the absolute change score by the number of transactions preceding the intervention and multiplying the result by 100. This provided a relative measure of change, reflecting how sales were affected by the intervention, independent of baseline transaction rates.

$$Percentage\ Change = \left(\frac{Post - Pre}{Pre}\right) * 100$$

During this process, we encountered 'infinity' values, which arose when the pre-intervention transaction count was zero, resulting in division by zero. To handle this, we replaced infinity values with NaN (Not a Number) and subsequently dropped these rows from our data, thereby removing all transactions that led to division by zero-values from the data, effectively removing all transactions that had infinity values.

The same normalisation and standardisation of our data was applied to the data specifying the amount of product sold per product.

Assessing Normality.

To assess normality, we used graphical and numerical methods. For both the data about transactions and the data describing the amount of product sold we created histograms and QQ-plots to assess the distribution of the data. Additionally, we used the Shapiro–Wilk on our data to statistically confirm or disconfirm the assumption of normality. You can find the graphs and tests for normality in the appendix. The statistical tests used in our analysis, such as ANOVA, t-test, Kruskal-Wallis test and Mann-Whitney-U test are based on the different assumption of normality that we assessed.

Reducing Variance and Increasing Power.

The reduce the high variance between the label colour groups in hopes to increase statistical power, we reran our initial tests on groups that combined the yellow and red label groups into a “Negative Labels” group and renamed the green label group into a “Positive Labels” group. With these groups we were able to increase the number of labels that were compared to the green or positive labels group, and we were able to use a different array of statistical tests as further outlined below.

Statistical Tests.

The data on the label colour and pre- and -post intervention groups that was normally distributed was further assessed by comparing the variances across means through an analysis of variance (ANOVA). Data that did not approach a normal distribution was analysed with the Kruskal–Wallis test by ranks.

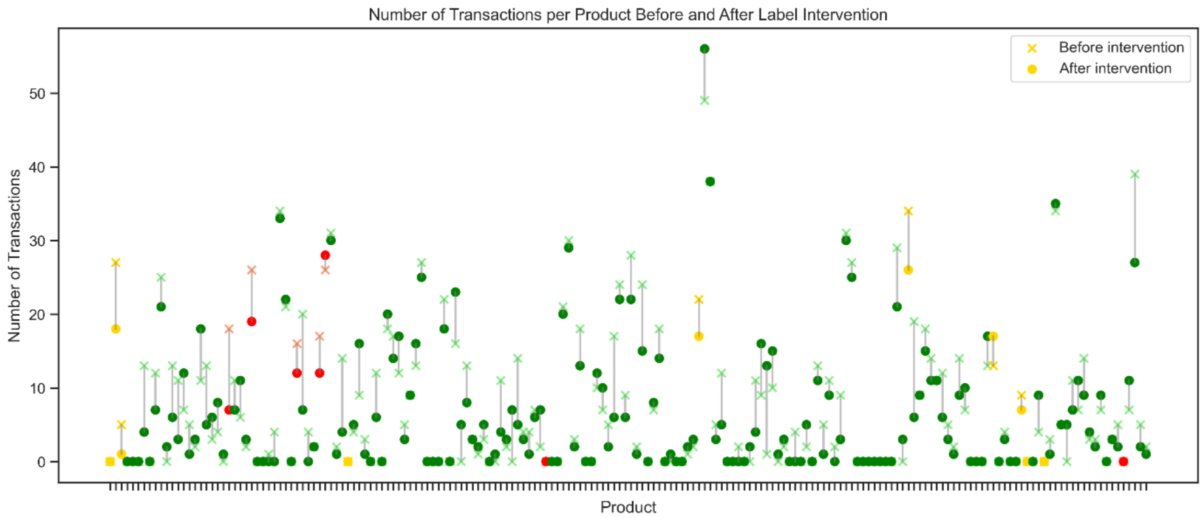
The data on the combined groups of positive and negative labels was analysed with an unpaired t-test in case of our normally distributed data, and with a Mann-Whitney-U test in case of the not normally distributed data.

Results

Removing outliers and filtering of data

The following section will first present the data that includes all products after our pre-processing described above. This unfiltered dataset consists of 184 unique products. 167 of which are labelled as green products, 10 as yellow and 7 as red products.

Figure 12

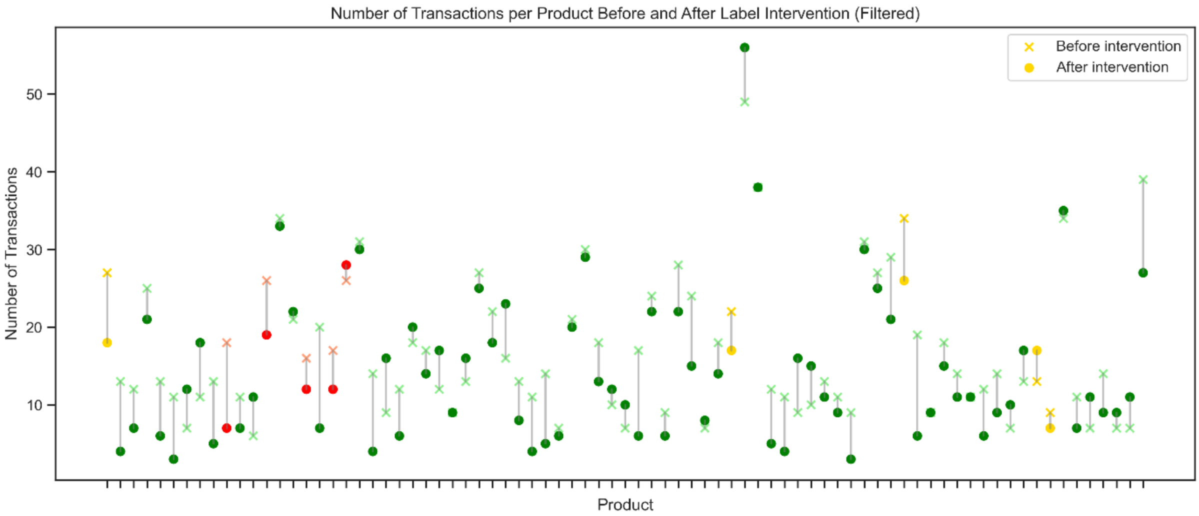


The filtered dataset will follow the section about the unfiltered data and consists of 79 products. 69 of which have the green label, 5 the yellow label and 5 the red label.

We filtered out all products that were sold less than or equal to 5 times, as in number of transactions, in the month before our label implementation. All products that matched the following equation with “T” representing the number of transactions, were excluded from the dataset:

$$0 \leq T \leq 5$$

Figure 13



Applying this filter helps remove potential bias and noise from infrequently sold products. These products could skew results and lead to inaccurate conclusions due to their extreme relative changes with minor absolute variations. By excluding them, we can focus our analysis on regularly purchased products, providing a more reliable understanding of the labelling system's impact on general consumer behaviour.

We will continue to reference the raw data we received from the store after pre-processing as “unfiltered” and the data where the low sales products were removed as “filtered”. One may call it “outliers removed”, but the products we removed were not extreme outliers of any sort, but simply products which were either not sold at all during the time of our measurements, or very infrequently bought.

For assessments of normality, please refer to the Appendix.

Unfiltered Data

Number of transactions pre- and -post label implementations

Figure 14

Number of Transactions One Month Before and One Month After Label Implementation

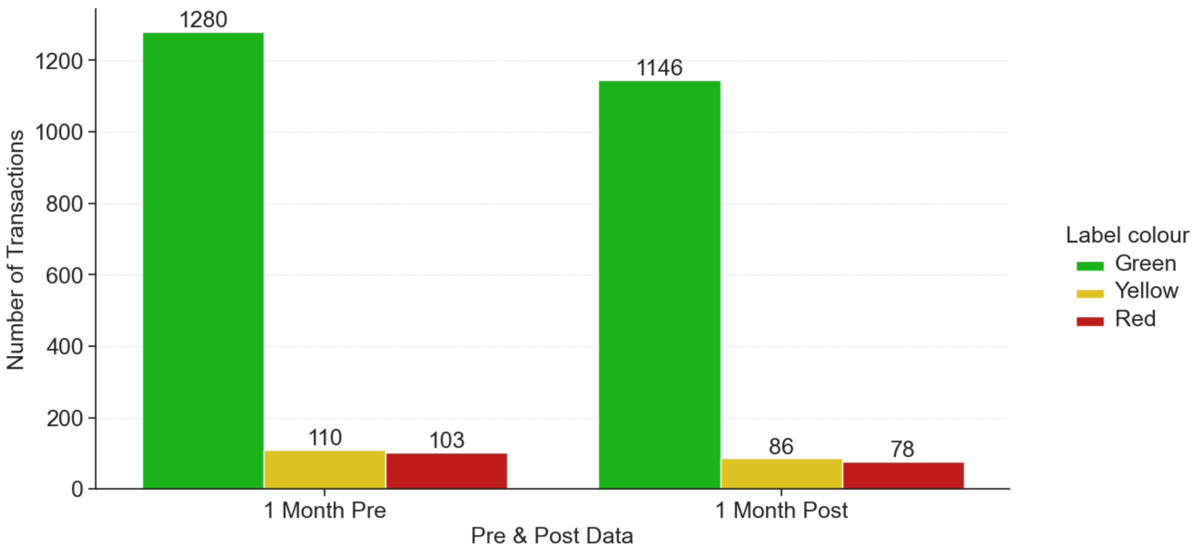


Figure 14 indicates a change in the number of transactions between the pre- and post-label sample. In the month prior to label implementation (pre), there were 1280 transactions for green products, 110 transactions for yellow products, and 103 transactions for red products. Following label implementation (post), the number of transactions decreased to 1146 for green products, 86 for yellow products, and 78 for red products.

The total sales experienced a decline of 183 transactions, corresponding to a 12.25% reduction between the pre- and post-implementation periods of the labelling system. This overall decrease in transactions was observed across the different label colours as follows: a 10.47% decline for green labelled products, a 21.82% reduction for yellow labelled products, and a 24.27% decrease for red labelled products.

Figure 15

Average Transactions per Product Category: One Month Before and One Month After Label Implementation

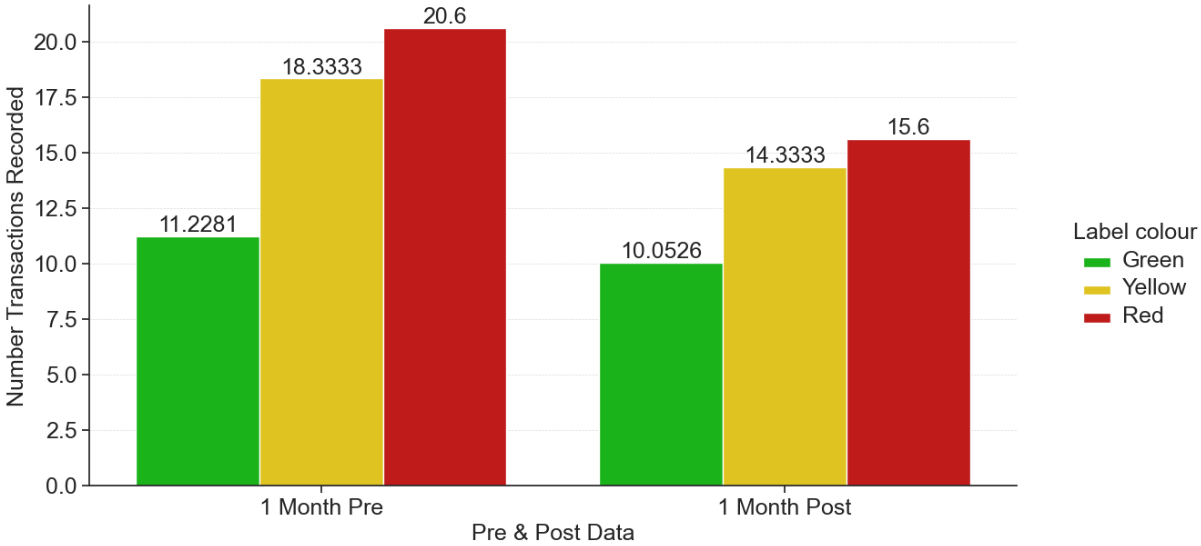


Figure 15 is showing the average number of transactions for each product per month, normalized by the number of products in each category. This allows for a more balanced comparison between categories. Green labels show a slight decrease in transaction numbers from 11.2281 to 10.0526 after the intervention. Yellow labels display a decrease in transaction numbers from 18.3333 to 14.3333, and red labels also show a decline from 20.6 to 15.6. This suggests a reduction in transaction numbers for all label categories after the implementation of the nutritional labelling system.

Figure 16

Average Percentage Change in the Number of Transactions per Label Category with Individual Products Overlayed

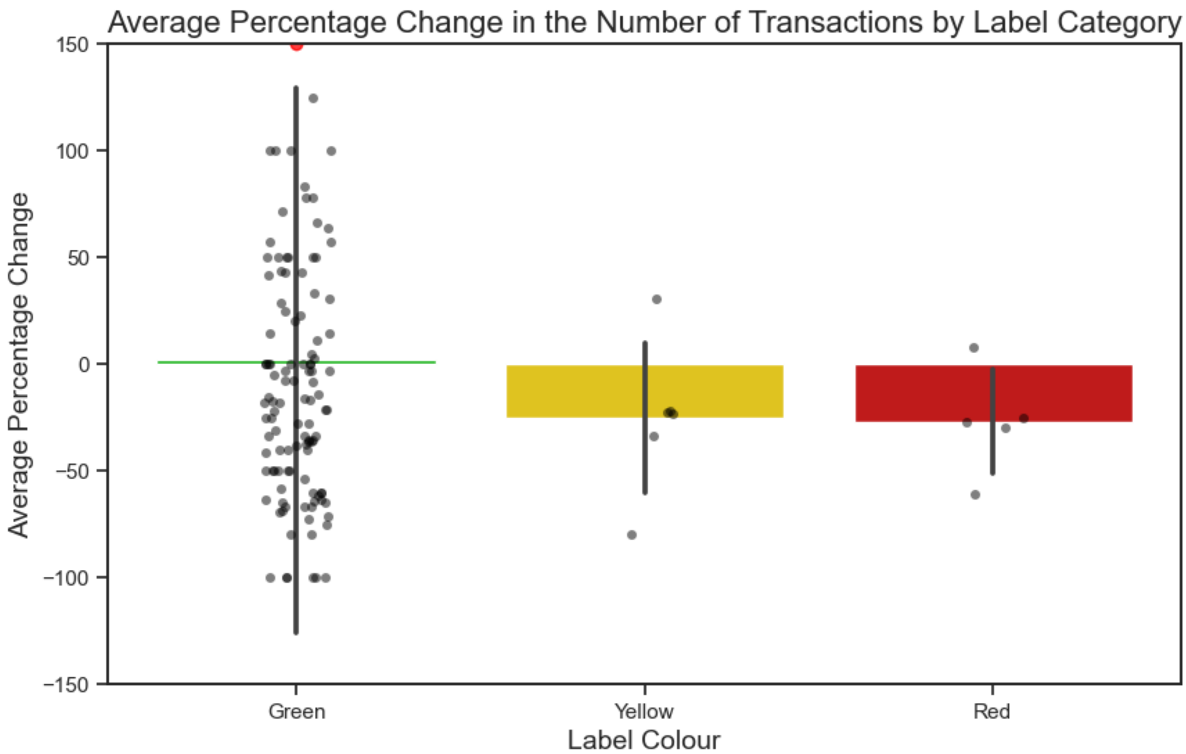


Figure 16 presents the average percentage change in the number of transactions per label category, with individual products overlaid on the graph. The green bar represents a slight decrease in transactions, with an average percentage change of -2%. In contrast, the yellow and red bars exhibit more substantial declines, with average percentage changes of -25.2% and -27%, respectively. The error bars denote the standard deviation, giving an understanding of the variability within each label category.

These findings suggest that products with yellow and red labels experienced a more pronounced decrease in transactions compared to those with green labels.

ANOVA on Transaction Change Score.

Table 1

Results of the ANOVA on the Transaction Change Score

Source	SS	DF	MS	F	p	np2
Label Colour	95.042526	2	47.521263	2.279117	0.106717	0.036017
Within	2543.789474	122	20.850733	NaN	NaN	NaN

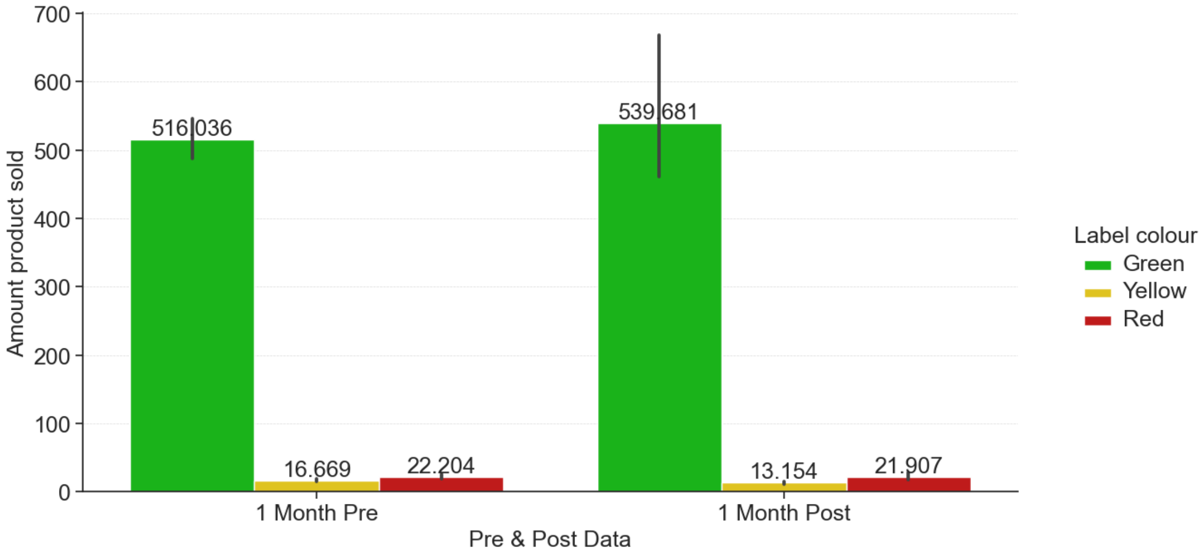
T-test on the Transaction Change Score With Combined Label Groups.

An independent samples t-test was conducted to compare the transaction change scores between the positive labels and the negative labels groups. The results indicated a marginally significant difference in scores, $t(12.06) = 2.13$, $p = .054$, with a 95% confidence interval ranging from -0.07 to 6.13. The effect size, as measured by Cohen's d, was 0.67, suggesting a medium to large effect. The Bayes Factor (BF10) was 1.872, providing anecdotal evidence in favour of the alternative hypothesis. The statistical power of the test was 0.55. In other words, the mean transaction change score for the positive label group is marginally significantly different from the mean transaction change score of the negative label group.

Amount of product sold pre- and -post label intervention

Figure 17

Number of Product Amount Sold One Month Before and One Month After Label Implementation

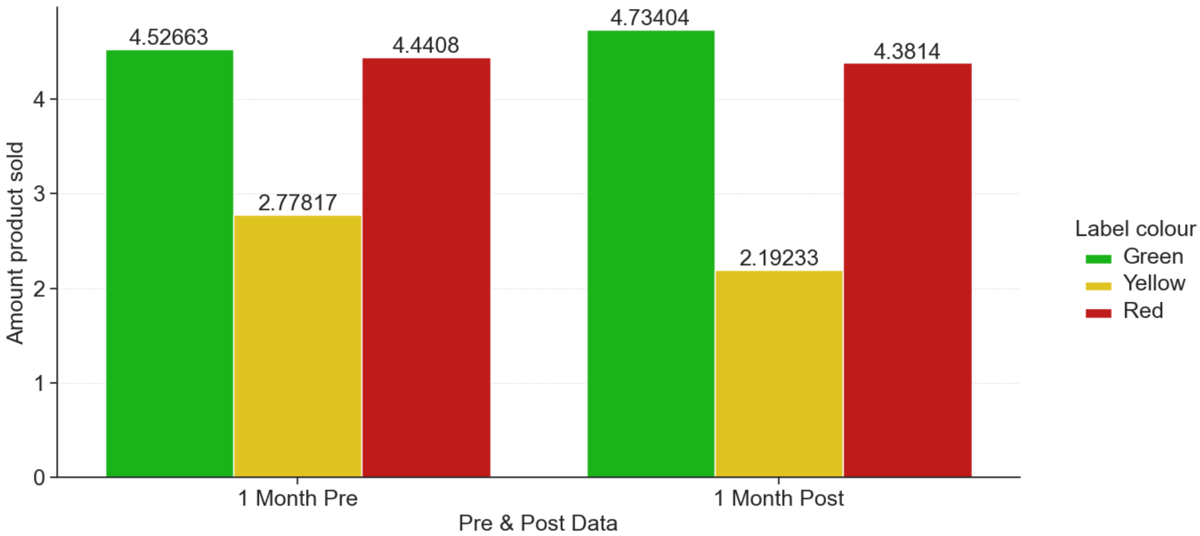


The sum of the products sold (product amount), measured in either kilograms or packaged units, was analysed for each label colour category before and after the label implementation. The results indicate that before the label implementation green labels accounted for 516.036 units, yellow labels for 16.669 units, and red labels for 22.204 units. After the label implementation, green labels showed an increase in product amount sold, totalling 539.681 units, while yellow and red labels experienced a decrease, with 13.154 and 21.907 units, respectively.

As such, green labels experienced an increase of 4.57%, while yellow labels saw a decrease of 21.09%, and red labels exhibited a slight decline of 1.34%

Figure 18

Average Product Amount Sold per Product Category: One Month Before and One Month After Label Implementation



The graph illustrates the analysis of the average amount of products sold, either in kilograms or packaged units, for each label colour category before and after the label implementation. Before the label implementation, green labels represented 4.53 units, yellow labels 2.78 units, and red labels 4.44 units. Post-implementation, green labels experienced a mild increase to 4.73 units (+4.57%), while both yellow and red labels saw a decrease to 2.19 units (-21.09%) and 4.38 units (-1.34%), respectively.

The results highlight varying impacts of the labels on the sale of products, with green labels driving an increase in sales, and yellow and red labels causing a slight decline.

Figure 19

Average Percentage Change in the Amount of Product Sold per Label Category with Individual Products Overlaid

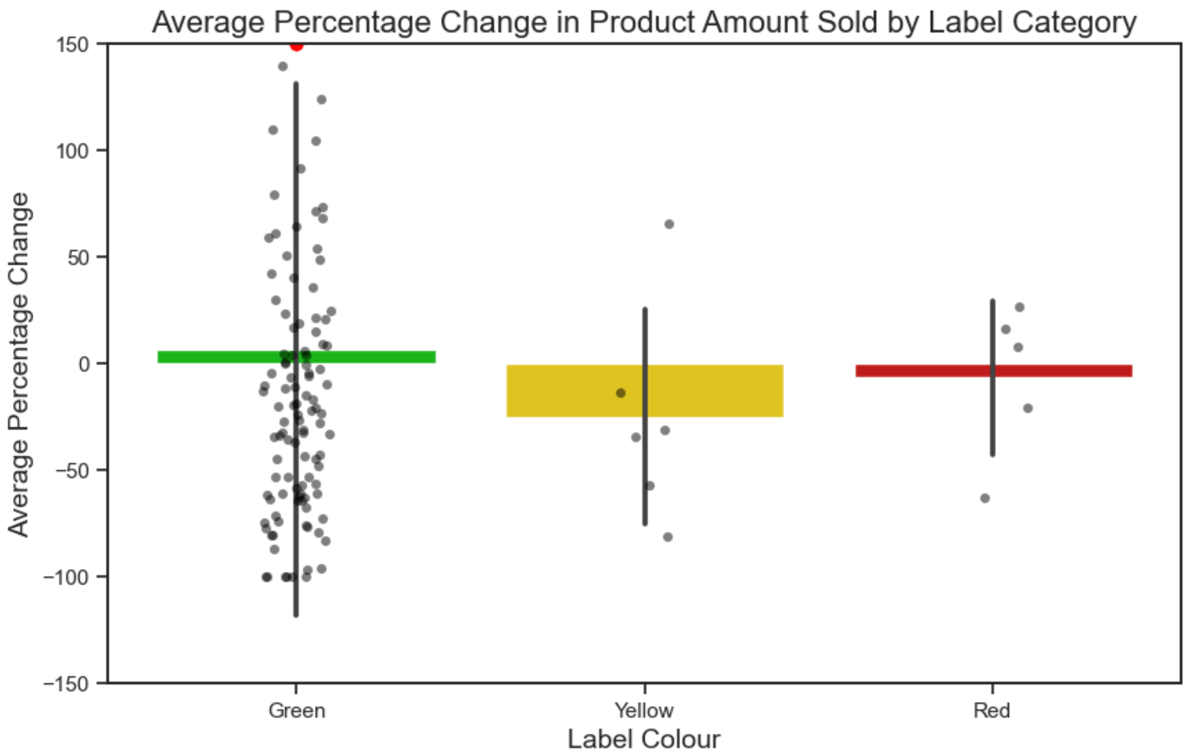


Figure 19 presents the average percentage change in the amount of product sold per label category, with individual products overlaid on the graph. The green bar represents an increase in the amount of product sold, with an average percentage change of +6.66%. Conversely, the yellow and red bars exhibit significant decreases, with average percentage changes of -25.08% and -6.49%, respectively. The error bars denote the standard deviation, giving an understanding of the variability within each label category.

These findings suggest that products with yellow labels experienced a steep decline in the amount sold, while green-labelled products saw a slight increase, and red-labelled products underwent a moderate decrease.

Kruskal-Wallis Test by Ranks on Product Amount Change Score.

A Kruskal-Wallis test was conducted to compare the product amount change scores across the different label colour groups. The results indicated that there was no significant difference in the product amount change scores between the groups, $H(2) = 1.07$, $p = .584$. Therefore, we failed to reject the null hypothesis and concluded that label colour did not have a significant effect on the product amount change scores.

Mann-Whitney-U on Product Amount Change Score for Combined Label Groups.

A Mann-Whitney U test was conducted to compare the product amount change scores between the combined label groups (positive labels vs. negative labels). The results indicated that there was no significant difference between the two groups, $U = 632.0$, $p = .930$, two-sided. The rank-biserial correlation was -0.017, and the common language effect size (CLES) was 0.508, suggesting a negligible difference between the groups.

Filtered Data

In the following section we will present our findings with the products filtered as described above.

Number of Transactions Pre- and -Post Label Implementations (Filtered)

Figure 20

Number of transactions pre- and -post label implementations (Filtered)

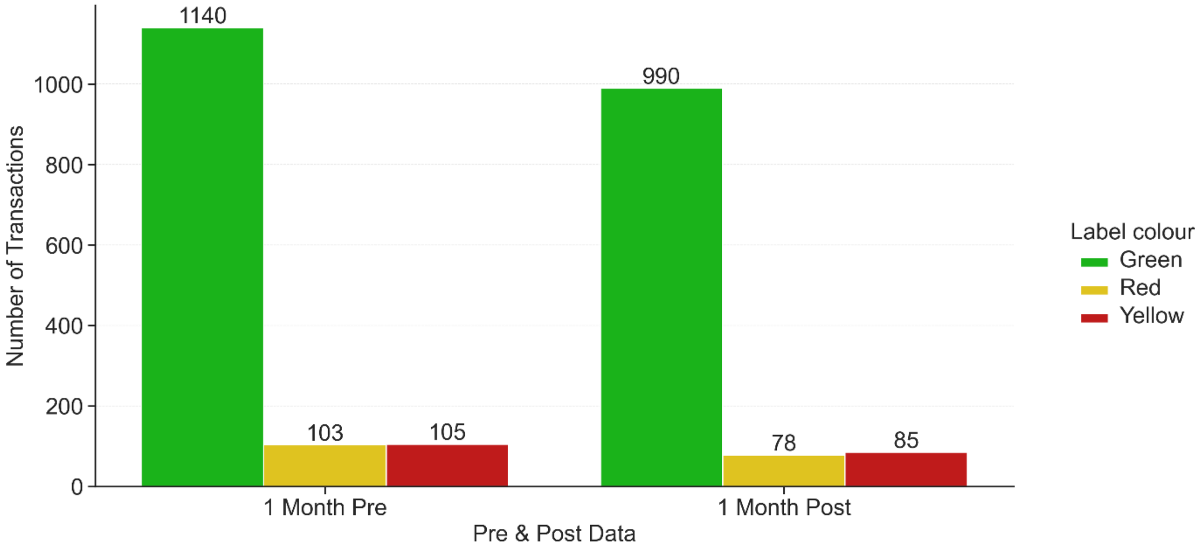


Figure 20 displays the transition in the number of transactions between the pre- and post-label implementation period, considering the filtered data. In the month leading up to the label implementation (pre), there were 1140 transactions for green labelled products, 103 transactions for yellow labelled products, and 105 transactions for red labelled products. After the label implementation (post), the number of transactions fell to 990 for green products, 78 for yellow products, and 85 for red products.

The total transactions experienced a decrease of 195, representing a 13.22% drop between the pre- and post-implementation periods of the labelling system. This overall decrease in transactions was observed across the different label colours as follows: a 13.16% decline for green labelled products, a 24.27% reduction for yellow labelled products, and a 19.05% decrease for red labelled products.

Figure 21

Average Transactions per Product Category: One Month Before and One Month After Label Implementation (Filtered)

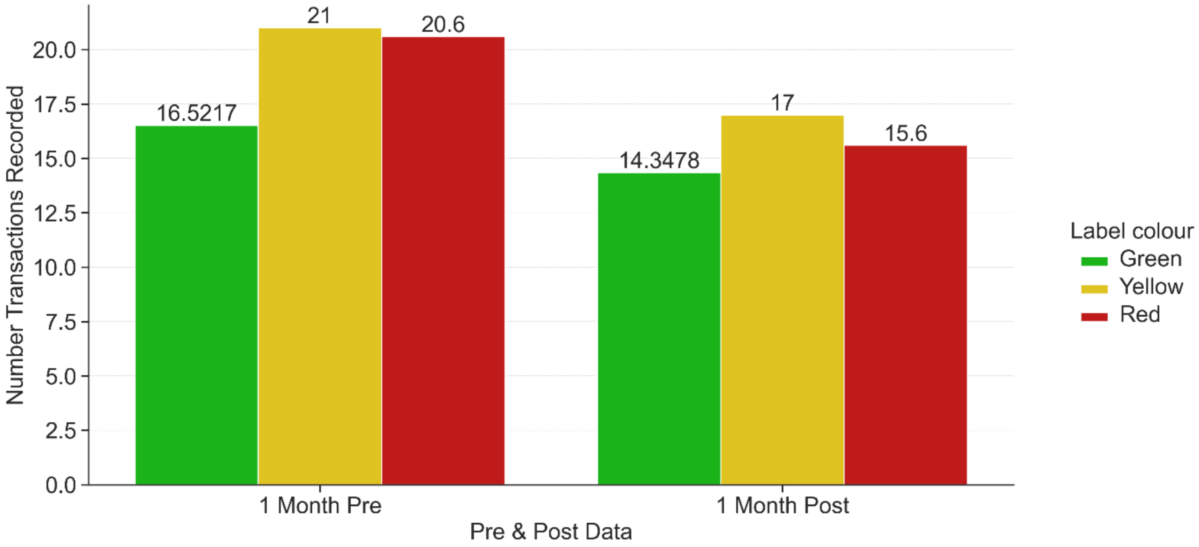


Figure 21, based on the filtered data, represents the average number of transactions for each product per month, normalized by the number of products in each category. This normalization allows for a more equitable comparison among categories.

Before the intervention, the average transaction numbers for green, yellow, and red labels were 16.5217, 21, and 20.6, respectively. Following the implementation of the labelling system, the transaction numbers for these categories changed to 14.3478, 17, and 15.6.

These figures indicate a decrease in transaction numbers across all label categories post-implementation. Green labels experienced a drop of around 13.17%, yellow labels saw a reduction of approximately 19.05%, and red labels recorded a decline of about 24.27%.

Figure 22

Average Percentage Change in the Number of Transactions per Label Category with Individual Products Overlaid (Filtered)

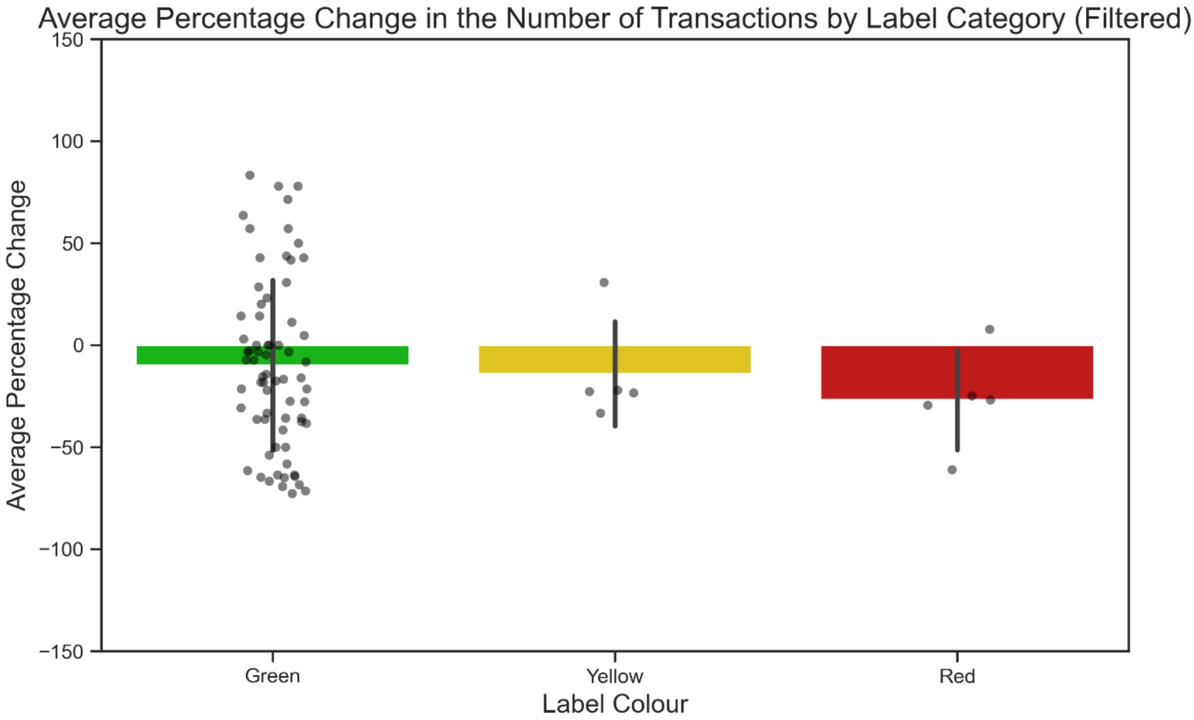


Figure 22 illustrates the average percentage change in the number of transactions per label category, with individual products as dots on the graph. The green bar indicates a moderate decrease in transactions, with an average percentage change of -10%. The yellow and red bars show larger decreases, with average percentage changes of -14.2% and -27.0%, respectively. The error bars indicate the standard deviation, providing an insight into the variability within each label category.

These findings imply that products labelled as green experienced a moderate decrease in transactions, while yellow-labelled products faced a more substantial reduction, and red-labelled products underwent the most significant decline following the implementation of the labelling system.

ANOVA on Transaction Change Score (Filtered).

Table 2

Results of the ANOVA on the Transaction Change Score

Source	SS	DF	MS	F	p	np2
Label Colour	49.757843	2	24.878921	0.920583	0.402682	0.023653
Within	2053.913043	76	27.025172	NaN	NaN	NaN

T-test on the Transaction Change Score With Combined Label Groups (Filtered).

An independent samples t-test was conducted to compare the transaction change scores between the positive labels and the negative labels groups, using the filtered data. The results indicated a marginally significant difference in scores, $t(12.39) = 1.43$, $p = .0177$, with a 95% confidence interval ranging from -1.21 to 5.86. This indicates that there is no statistically significant difference between the mean transaction change score of green labelled products and the combined group of yellow and red-labelled products.

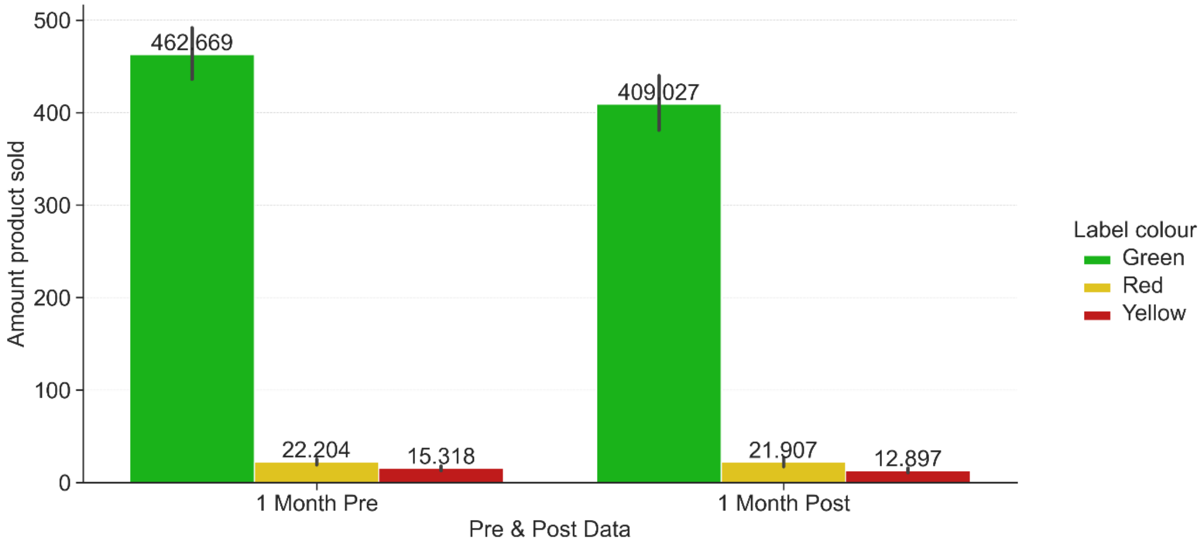
The effect size, as measured by Cohen's d, was 0.45. This value indicates a small to medium effect size according to Cohen's classification. The Bayes Factor (BF10) was 0.715, providing anecdotal evidence slightly in favour of the null hypothesis (no significant difference) over the alternative hypothesis (a significant difference).

The statistical power of the test was 0.26, which is relatively low. This indicates that the test has a relatively high risk of Type II error. The sample size might be too small to detect a significant effect if it exists.

Amount of Product Sold Pre- and -Post Label Intervention (Filtered)

Figure 23

Number of Product Amount Sold One Month Before and One Month After Label Implementation (Filtered)

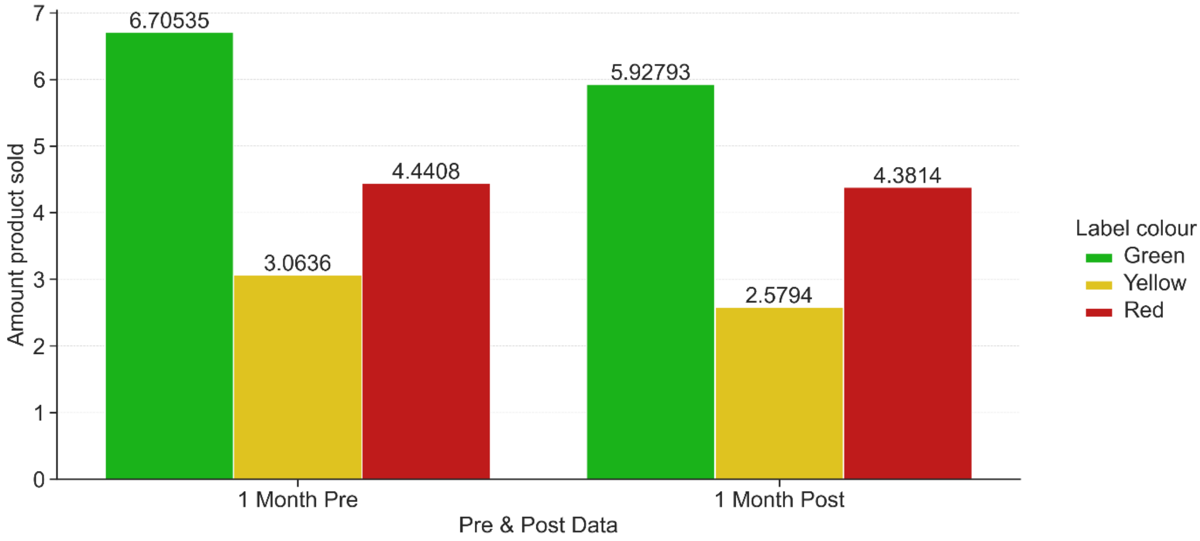


The total volume of products sold, by weight in kilograms or count of packaged units, was examined for each label colour category before and after the label introduction. The analysis reveals that before the label application, green labelled items accounted for 462.669 units, yellow labelled products for 22.204 units, and red labelled products for 15.318 units. After the labels were introduced, the volume of green-labelled products sold dropped to 409.027 units, while yellow and red labels experienced declines as well, falling to 21.907 and 12.897 units, respectively

In terms of percentage change, green labels saw a decrease of approximately 11.61%, yellow labels experienced a slight decline of 1.34%, and red labels showed a more pronounced decrease of approximately 15.80%.

Figure 24

Average Product Amount Sold per Product Category: One Month Before and One Month After Label Implementation (Filtered)



The graph demonstrates the average quantity of products sold, either by weight in kilograms or count of packaged units, for each label colour category, both prior to and following the label implementation. Before the labels were introduced, green-labelled items represented 6.70535 units on average, products with yellow labels accounted for 3.0636 units, and products with red labels made up 4.4408 units.

After the label application, a drop in product quantity sold was observed. Green-labelled products saw a decline to an average of 5.92793 units, representing a decrease of approximately 11.60%. Both yellow and red-labelled products also experienced reductions, with yellow dropping to an average of 2.5794 units (-15.82%) and red decreasing to 4.3814 units (-1.34%).

These results show the varied impacts of the labels on product sales. While all label categories saw a decrease in the average quantity of products sold, the effect was most pronounced for green and yellow-labelled products.

Figure 25

Average Percentage Change in the Amount of Product Sold per Label Category with Individual Products Overlaid (Filtered)

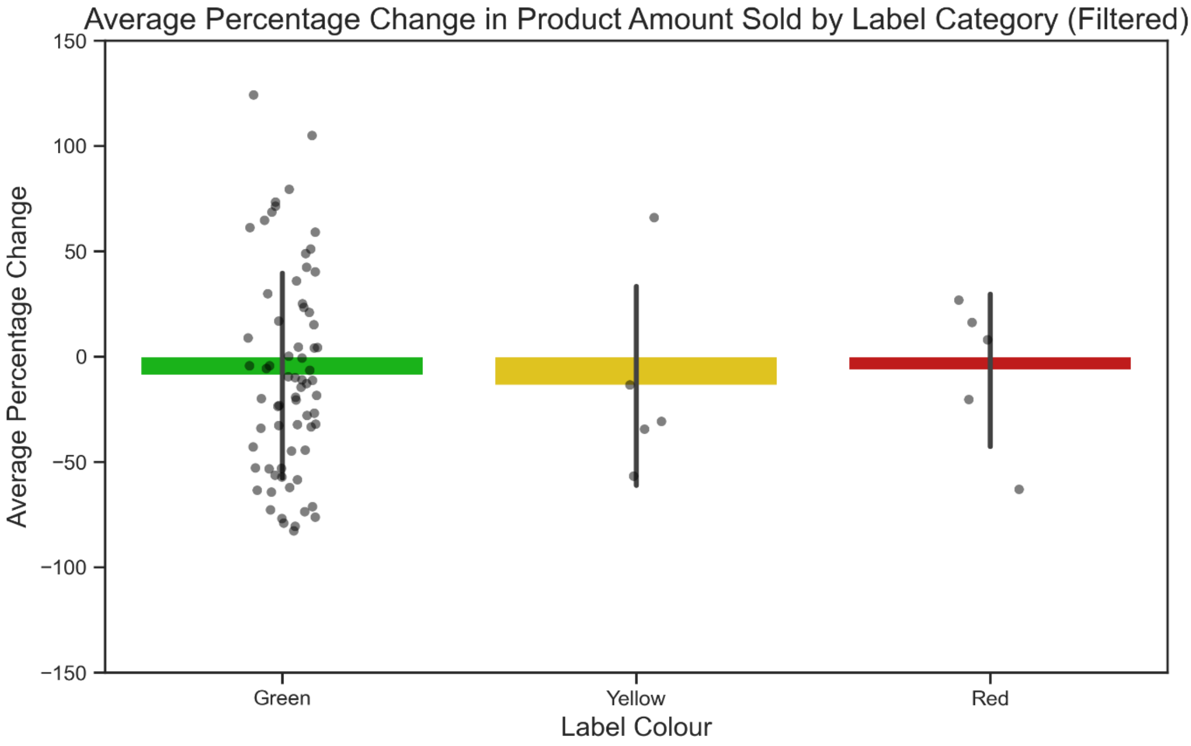


Figure 25 illustrates the average percentage change in the quantity of product sold for each label category, with individual product data points also depicted. The green bar indicates a reduction in the quantity of products sold, showing an average percentage change of -9.1%. In contrast, the yellow bar demonstrates a larger decrease and the red bar a smaller decrease, with average percentage changes of -13.9% and -6.5%, respectively. The error bars represent the standard deviation, providing an insight into the variability within each label category.

These results suggest that all products, regardless of their label colour, experienced a decline in the amount sold. The decrease was most pronounced for yellow-labelled products, followed by green-labelled ones, while red-labelled products saw a slightly smaller reduction.

Kruskal-Wallis Test by Ranks on Product Amount Change Score (Filtered).

A Kruskal-Wallis test was performed to compare the product amount change scores across the different label colour groups. The results revealed that there was no significant difference in the product amount change scores among the groups, $H(2) = 0.69$, $p = .709$. As such, we failed to reject the null hypothesis and concluded that label colour did not exert a significant effect on the product amount change scores.

Mann-Whitney-U on Product Amount Change Score for Combined Label Groups (Filtered).

A Mann-Whitney U test was conducted to compare the product amount change scores between the combined label groups (positive labels vs. negative labels). The results demonstrated that there was no significant difference between the two groups, $U = 311.0$, $p = .621$, two-sided. The rank-biserial correlation was 0.099, and the common language effect size (CLES) was 0.451, suggesting a minimal difference between the groups.

Only Marginal Statistical and No Statistical Significance

Despite the above tests being statistically only marginally significant, the analysis of transaction data reveals a distinct trend in product sales across the three label categories following the implementation of the labelling system. While the overall number of transactions decreased for all label categories, the decline was more pronounced for yellow and red products compared to green products. However, it is to note that there was a big difference between the number of products in the green label category, and the number of products in the yellow and red label category.

Interestingly, despite the reduction in transactions, the total amount of green products sold increased from 516.036 units to 539.681 units post-implementation. This suggests a potential shift in consumer preferences towards green-labelled products. In contrast, both yellow and red-labelled products experienced a decline in the total product amount sold, with yellow products exhibiting a considerable decrease from 16.669 to 13.154 units, while red products experienced a relatively smaller decline from 22.204 to 21.907 units.

Overall, the findings indicate that the introduction of the labelling system may have led to a more pronounced decrease in transactions for yellow and red-labelled products, while promoting an increase in the total amount of green products sold. However, this is only true when visually analysing the graphs we created. Statistically there are no significant changes found between pre- and post-label implementation transactions or product amount sales.

In general, with the high variability of products counts between the different label categories, we want to be cautious when generalizing our findings.

Within the filtered data, our findings suggest that the introduction of the labelling system led to a decrease in both transactions and product volume sold, with varying impacts across different label categories. However, statistical analysis does not conclusively prove a significant difference caused by the labels. The results of the filtered data underline the lack of power due to the limited number of products that we tested in this study. Additionally, all categories seemed to experience a decline, which could have simply been due to a general decline is sales in the store.

In conclusion, our statistical results point towards no effect of our climate label intervention on consumer behaviour. While in the unfiltered data there seems to be a slight trend towards more green categorized products bought and less yellow and red categorized products bought, in the filtered data this difference is not present.

Discussion

In our study we investigated the effect of a climate label intervention with green, yellow, and red labels on the number of sales and the amount of product sold in their respective colour categories.

The purpose of this study was to gain a better understanding of how climate label interventions may impact consumer buying behaviour and to add to the literature exploring whether climate labels can be an effective part of the global and multifaceted effort to combat climate change.

The study underscores the importance of policy changes, including climate labels, in influencing climate change and explores future avenues for research. However, it also highlights the challenges associated with climate labelling. The success of a labelling intervention depends on factors such as the accuracy of the labelling, the visibility and ease of understanding of the labels, the potential for information overload, and consumer trust in the labelling system. There is also the question of who should bear the cost of the label, which depends on whether the label is driven by producers, individual food retailers, or the government, all of which may have different goals for the label initiative.

The results of the present study reject the hypothesis that the implementation of three categories of climate labels leads to higher sales for products with green climate labels. Additionally, the results provide evidence against the hypothesis that products with a red or orange label will have lower sales after the label introduction.

The study yielded three pivotal observations. First, the implementation of climate labels in the grocery store did not significantly influence consumer buying behaviour. This was observed through the comparison of data collected one month before and after the label implementation, which showed no significant differences among the three label groups.

Second, when the red and negative labels were combined into a single group to increase power, a marginally significant difference was observed. This observation suggests a possible marginal decrease in purchases of negatively labelled products, though it is important to interpret this with caution given its borderline significance. Third, upon refining our analysis to exclude products that were purchased in very small quantities, we found that the statistically significant results observed in the unfiltered data could not be replicated. This led us to the conclusion that both our initial hypotheses must be rejected.

Our results are partially consistent with previous literature that suggests that consumers are slightly influenced by labels and that these labels may affect purchasing decisions (Brunner et al., 2018; Canavari & Coderoni, 2019; Feucht & Zander, 2018a; Osman & Thornton, 2019; Taufique et al., 2022). However, the results of our study are more in line with previous literature that reported either very small or no evidence for the hypothesis that climate or carbon labelling can positively influence consumer behaviour towards more eco-friendly purchasing decisions (Emberger-Klein & Menrad, 2018; Hornibrook et al., 2015; Slapø & Karevold, 2019; Taufique et al., 2022).

Taken together, our findings suggest that it is conceivable that a climate label intervention might exert a minimal impact on consumer buying behaviour and subsequently on overall sales. However, it is pivotal to underscore that the absence of statistically significant results in our study limits the generalizability and applicability of our findings. Graphical analyses tentatively indicate potential differences between the positive and the negative label groups, warranting further investigation, as the statistical analyses did not yield conclusive evidence of substantial effects. Additionally, the filtered data also suggests no statistically significant effects of climate labels on purchasing behaviour.

The Challenges of Climate Labelling

In addition to our findings above, we would like to outline multiple general challenges of climate labelling.

One of the primary challenges is the so-called 'attitude-behaviour gap'. While there is a large proportion of the population that expresses a positive attitude towards caring for the environment, this sentiment often does not translate into actual behaviour. Consumers might appreciate the value of climate-friendly choices, but when it comes to making purchase decisions, they frequently choose convenience, familiarity, or cost over environmental considerations.

Another significant challenge is habitual purchasing behaviour. Many consumers purchase products out of habit, which can be difficult to change. Without direct incentives or noticeable benefits, it can be hard to break these habits and encourage people to try alternative, more environmentally friendly options. Climate labelling can provide the necessary information but shifting strongly learned consumer habits will likely require more multifaceted approach.

Trust is also a significant factor. For climate labels to be effective, consumers must trust the information presented to them. If there is doubt about the accuracy of the carbon footprint information or if the labelling is perceived as a marketing gimmick, the effectiveness of the climate labelling initiative will be significantly undermined. Building this trust will require transparency about the methods used to calculate the carbon footprint and the oversight of the labelling process.

Moreover, consumers can only process a limited amount of information. A label packed with detailed environmental data might be overwhelming and, paradoxically, discourage its use. For climate labelling to be effective, it needs to be easily comprehensible and convey the key information in a manner that can readily inform purchase decisions.

While climate labelling has the potential to be an effective tool in the fight against global warming, these challenges need to be addressed for it to have a significant impact. Further research is needed to understand how best to design and implement climate labels and how to communicate this information effectively to influence consumer behaviour positively. At the same time, climate labelling should be part of a broader strategy, which includes policies and incentives that support and encourage more sustainable consumption.

Challenges of This Study

Additionally, there are at least three potential limitations concerning the results of this study. First, the lack of control over confounding variables in our field research study is a significant limitation. As we did not collect any participant characteristics, we were unable to control for factors such as socioeconomic status, age, gender, education level, health and environment consciousness, prior experiences with similar interventions, and others.

Second, the small sample size of the labels themselves posed a limitation. This led to a lower likelihood of statistical significance due to decreased statistical power.

Third, the lack of control over the Little Plant Pantry store and the data they provided was a limitation. Internal changes to the products or product shelves were not accounted for. Any change in the setup of the store or the tracking of sales during the intervention could have influenced the outcome of this study.

During our study, we faced additional challenges that may be avoided in future studies with better cooperation with the supermarket or store that provides the data. For instance, it was difficult to have a higher level of control over the products sold and the changes made to those products within the Little Plant Pantry store. While this is to be expected while doing field-research, there are multiple points of improvement:

All the labels should have been put at the same point in time, so that there is a clear cut-off between pre- and post-intervention.

The product naming schemes should not have been changed during the duration of the study or there should have been unique identifiers for each product from the start. The latter point made it difficult for us to keep the data clean and have a clear insight into what product exactly had which transaction count over the course of multiple months. Due to this change in naming conventions in the Zettle payment system, we were forced to only assess one month prior and one month after the label introduction, as that was the timeframe in which the naming conventions remained unchanged.

Certain limitations of this study could be addressed in future research. For example, the label distribution could be more balanced so that variances between label categories are small. Additionally, more labels in bigger markets could be observed to increase the sample size and the variety of products. Future studies may also incorporate a greater variety of products, including more meat products. A greater variety of products will give a better overview of how climate labels will influence consumer behaviour within different product categories. For example, across meat and non-meat products.

Given these substantial limitations, our results tentatively suggest some potential theoretical and practical implications. Future studies on the effect of climate labels should carefully choose the supermarket they are cooperating with and set guidelines and boundaries in cooperation with the store to ensure frictionless exchange of data. Additionally, the store may share potential changes to the store layout or products that could act as confounders to influence the results of the study in unwanted or unexpected ways.

This study can serve as a preliminary resource for other researchers, policymakers, and regulatory bodies, offering initial insights into what might potentially be expected from a label intervention on the overall product sales and thus greenhouse gas expenditure of supermarkets and the food industry.

Investigating more long-term effects of the intervention can assess the sustainability of the interventions impact on consumer behaviour and may reveal additional advantages and challenges resulting from a climate label intervention. Markets and food stores should be chosen that attract a wider variety of customers that better represent the average population.

The present study serves as an exploratory step in understanding the intricacies of climate label interventions and their implications on consumer behaviour. The findings, tempered by the study’s limitations, underscore the need for more robust and diverse research to unravel the potential and challenges of climate labels in contributing to environmental sustainability. We hope that the current research will stimulate further investigation of climate labels as a supporting intervention in the complex array of strategies that are necessary to reduce global warming.

References

Advancing the Science of Climate Change (p. 12782). (2010). National Academies Press. https://doi.org/10.17226/12782

Birkenberg, A., Narjes, M. E., Weinmann, B., & Birner, R. (2021). The potential of carbon neutral labeling to engage coffee consumers in climate change mitigation. *Journal of Cleaner Production*, *278*, 123621. https://doi.org/10.1016/j.jclepro.2020.123621

Boardman, B. (2008). Carbon labelling: Too complex or will it transform our buying? *Significance*, *5*(4), 168–171. https://doi.org/10.1111/j.1740-9713.2008.00322.x

Brunner, F., Kurz, V., Bryngelsson, D., & Hedenus, F. (2018). Carbon Label at a University Restaurant – Label Implementation and Evaluation. *Ecological Economics*, *146*(January 2017), 658–667. https://doi.org/10.1016/j.ecolecon.2017.12.012

Brusseau, M. L. (2019). Sustainable Development and Other Solutions to Pollution and Global Change. In *Environmental and Pollution Science* (pp. 585–603). Elsevier. https://doi.org/10.1016/B978-0-12-814719-1.00032-X

Bullock, G. (2017). Green Grades: Can Information Save the Earth? In *Green Grades: Can Information Save the Earth?* (p. 390). https://doi.org/10.7551/mitpress/11121.001.0001

Caillavet, F., Fadhuile, A., & Nichèle, V. (2019). Assessing the distributional effects of carbon taxes on food: Inequalities and nutritional insights in France. *Ecological Economics*, *163*, 20–31. https://doi.org/10.1016/j.ecolecon.2019.04.020

Canavari, M., & Coderoni, S. (2019). Green marketing strategies in the dairy sector: Consumer-stated preferences for carbon footprint labels. *Strategic Change*, *28*(4), 233–240. https://doi.org/10.1002/jsc.2264

Carbon Trust. (2020, January 19). *Product carbon footprint label*. Https://Www.Carbontrust.Com/What-We-Do/Assurance-and-Labelling/Product-

Carbon-Footprint-Label. https://www.carbontrust.com/what-we-do/assurance-and-labelling/product-carbon-footprint-label

Carrington, M. J., Neville, B. A., & Whitwell, G. J. (2010). Why Ethical Consumers Don't Walk Their Talk: Towards a Framework for Understanding the Gap Between the Ethical Purchase Intentions and Actual Buying Behaviour of Ethically Minded Consumers. *Journal of Business Ethics*, *97*(1), 139–158. https://doi.org/10.1007/s10551-010-0501-6

Clark, M. A., Domingo, N. G. G., Colgan, K., Thakrar, S. K., Tilman, D., Lynch, J., Azevedo, I. L., & Hill, J. D. (2020). Global food system emissions could preclude achieving the 1.5° and 2°C climate change targets. *Science*, *370*(6517), 705–708. https://doi.org/10.1126/science.aba7357

Climate Action Tracker. (2021). *Warming Projections Global Update—November 2021*. https://climateactiontracker.org/documents/997/CAT_2021-11-09_Briefing_Global-Update_Glasgow2030CredibilityGap.pdf

COP26. (2021). *COP26 Outcomes*. UN Climate Change Conference (COP26) at the SEC – Glasgow 2021. https://ukcop26.org/the-conference/cop26-outcomes/

Cordell, D., & White, S. (2014). Life's Bottleneck: Sustaining the World's Phosphorus for a Food Secure Future. *Annual Review of Environment and Resources*, *39*(1), 161–188. https://doi.org/10.1146/annurev-environ-010213-113300

Creutzig, F., Fernandez, B., Haberl, H., Khosla, R., Mulugetta, Y., & Seto, K. C. (2016). Beyond Technology: Demand-Side Solutions for Climate Change Mitigation. *Annual Review of Environment and Resources*, *41*(1), 173–198. https://doi.org/10.1146/annurev-environ-110615-085428

Creutzig, F., Niamir, L., Bai, X., Callaghan, M., Cullen, J., Díaz-José, J., Figueroa, M., Grubler, A., Lamb, W. F., Leip, A., Masanet, E., Mata, É., Mattauch, L., Minx, J. C., Mirasgedis, S., Mulugetta, Y., Nugroho, S. B., Pathak, M., Perkins, P., … Ürge-

Vorsatz, D. (2022). Demand-side solutions to climate change mitigation consistent with high levels of well-being. *Nature Climate Change*, *12*(1), Article 1. https://doi.org/10.1038/s41558-021-01219-y

Creutzig, F., Roy, J., Lamb, W. F., Azevedo, I. M. L., Bruine de Bruin, W., Dalkmann, H., Edelenbosch, O. Y., Geels, F. W., Grubler, A., Hepburn, C., Hertwich, E. G., Khosla, R., Mattauch, L., Minx, J. C., Ramakrishnan, A., Rao, N. D., Steinberger, J. K., Tavoni, M., Ürge-Vorsatz, D., & Weber, E. U. (2018). Towards demand-side solutions for mitigating climate change. *Nature Climate Change*, *8*(4), Article 4. https://doi.org/10.1038/s41558-018-0121-1

Crippa, M., Solazzo, E., Guizzardi, D., Monforti-Ferrario, F., Tubiello, F. N., & Leip, A. (2021). Food systems are responsible for a third of global anthropogenic GHG emissions. *Nature Food*, *2*(3), Article 3. https://doi.org/10.1038/s43016-021-00225-9

Darnall, N., & Aragón-Correa, J. A. (2014). Can Ecolabels Influence Firms' Sustainability Strategy and Stakeholder Behavior? *Organization & Environment*, *27*(4), 319–327. https://doi.org/10.1177/1086026614562963

den Elzen, M. G. J., Dafnomilis, I., Forsell, N., Fragkos, P., Fragkiadakis, K., Höhne, N., Kuramochi, T., Nascimento, L., Roelfsema, M., van Soest, H., & Sperling, F. (2022). Updated nationally determined contributions collectively raise ambition levels but need strengthening further to keep Paris goals within reach. *Mitigation and Adaptation Strategies for Global Change*, *27*(5), 33. https://doi.org/10.1007/s11027-022-10008-7

Diaz, R. J., & Rosenberg, R. (2008). Spreading Dead Zones and Consequences for Marine Ecosystems. *Science*, *321*(5891), 926–929. https://doi.org/10.1126/science.1156401

Djama, M. (2011). Regulating the globalised economy: Articulating private voluntary standards and public regulations. *Perspective - Cirad*, 1–4. https://doi.org/10.19182/agritrop/00052

Edjabou, L. D., & Smed, S. (2013). The effect of using consumption taxes on foods to promote climate friendly diets – The case of Denmark. *Food Policy*, *39*, 84–96. https://doi.org/10.1016/j.foodpol.2012.12.004

Emberger-Klein, A., & Menrad, K. (2018). The effect of information provision on supermarket consumers' use of and preferences for carbon labels in Germany. *Journal of Cleaner Production*, *172*, 253–263. https://doi.org/10.1016/j.jclepro.2017.10.105

Eurobarometer, F. (2009). Europeans' attitudes towards the issue of sustainable consumption and production. *Flash Eurobarometer*, *256*, 1–18.

European Commission. (2008). *Environmental improvement potentials of meat and dairy products*. Publications Office. https://data.europa.eu/doi/10.2791/38863

Feucht, Y., & Zander, K. (2018a). Consumers' preferences for carbon labels and the underlying reasoning. A mixed methods approach in 6 European countries. *Journal of Cleaner Production*, *178*, 740–748. https://doi.org/10.1016/j.jclepro.2017.12.236

Feucht, Y., & Zander, K. (2018b). Consumers' preferences for carbon labels and the underlying reasoning. A mixed methods approach in 6 European countries. *Journal of Cleaner Production*, *178*, 740–748. https://doi.org/10.1016/j.jclepro.2017.12.236

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., & Snyder, P. K. (2005). Global Consequences of Land Use. *Science*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Gadema, Z., & Oglethorpe, D. (2011). The use and usefulness of carbon labelling food: A policy perspective from a survey of UK supermarket shoppers. *Food Policy*, *36*(6), 815–822. https://doi.org/10.1016/j.foodpol.2011.08.001

García-Muros, X., Markandya, A., Romero-Jordán, D., & González-Eguino, M. (2017). The distributional effects of carbon-based food taxes. *Journal of Cleaner Production*, *140*, 996–1006. https://doi.org/10.1016/j.jclepro.2016.05.171

GreenSwapp. (2022). *GreenSwapp – Climate impact tracking API for food*.

Harris, J. L., Pomeranz, J. L., Lobstein, T., & Brownell, K. D. (2009). A Crisis in the Marketplace: How Food Marketing Contributes to Childhood Obesity and What Can Be Done. *Annual Review of Public Health*, *30*(1), 211–225. https://doi.org/10.1146/annurev.publhealth.031308.100304

Hartikainen, H., Roininen, T., Katajajuuri, J.-M., & Pulkkinen, H. (2014). Finnish consumer perceptions of carbon footprints and carbon labelling of food products. *Journal of Cleaner Production*, *73*, 285–293. https://doi.org/10.1016/j.jclepro.2013.09.018

Hassan, L. M., Shiu, E., & Shaw, D. (2016). Who Says There is an Intention–Behaviour Gap? Assessing the Empirical Evidence of an Intention–Behaviour Gap in Ethical Consumption. *Journal of Business Ethics*, *136*(2), 219–236. https://doi.org/10.1007/s10551-014-2440-0

Hornibrook, S., May, C., & Fearne, A. (2015). Sustainable Development and the Consumer: Exploring the Role of Carbon Labelling in Retail Supply Chains: Exploring the Role of Carbon Labelling in Retail Supply Chains. *Business Strategy and the Environment*, *24*(4), 266–276. https://doi.org/10.1002/bse.1823

Intergovernmental Panel on Climate Change. (2019). *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*. Intergovernmental Panel on Climate Change. https://www.ipcc.ch/site/assets/uploads/2019/11/SRCCL-Full-Report-Compiled-191128.pdf

International Organization for Standardization. (2014, August 12). *ISO 14040:2006*. ISO. https://www.iso.org/standard/37456.html

IPCC. (2021). Summary for Policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 3−32). Cambridge University Press. https://doi.org/10.1017/9781009157896.001

IPCC. (2022). Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In *Climate Change 2022: Mitigation of Climate Change.* Cambridge University Press. https://www.ipcc.ch/report/sixth-assessment-report-working-group-3/

Ivanovich, C. C., Sun, T., Gordon, D. R., & Ocko, I. B. (2023). Future warming from global food consumption. *Nature Climate Change*, *13*(3), Article 3. https://doi.org/10.1038/s41558-023-01605-8

Kitzmueller, M., & Shimshack, J. (2012). Economic Perspectives on Corporate Social Responsibility. *Journal of Economic Literature*, *50*(1), 51–84. https://doi.org/10.1257/jel.50.1.51

Kumar, P., & Ghodeswar, B. M. (2015). Factors affecting consumers' green product purchase decisions. *Marketing Intelligence & Planning*, *33*(3), 330–347. https://doi.org/10.1108/MIP-03-2014-0068

Leiserowitz, A. (2007). *International Public Opinion, Perception, and Understanding of Global Climate Change* (Human Development Occasional Papers (1992-2007) HDOCPA-2007-31). Human Development Report Office (HDRO), United Nations

Development Programme (UNDP). https://EconPapers.repec.org/RePEc:hdr:hdocpa:hdocpa-2007-31

Limiting the Magnitude of Future Climate Change (p. 12785). (2010). National Academies Press. https://doi.org/10.17226/12785

LittlePlantPantry. (2021). *Organic Zero-Waste Grocery Shop Amsterdam*.

Liu, T., Wang, Q., & Su, B. (2016). A review of carbon labeling: Standards, implementation, and impact. *Renewable and Sustainable Energy Reviews*, *53*, 68–79. https://doi.org/10.1016/j.rser.2015.08.050

Macdiarmid, J. I., Douglas, F., & Campbell, J. (2016). Eating like there's no tomorrow: Public awareness of the environmental impact of food and reluctance to eat less meat as part of a sustainable diet. *Appetite*, *96*, 487–493. https://doi.org/10.1016/j.appet.2015.10.011

Mainieri, T., Barnett, E. G., Valdero, T. R., Unipan, J. B., & Oskamp, S. (1997). Green Buying: The Influence of Environmental Concern on Consumer Behavior. *The Journal of Social Psychology*, *137*(2), 189–204. https://doi.org/10.1080/00224549709595430

Matplotlib/matplotlib. (2023). [Python]. Matplotlib Developers. https://github.com/matplotlib/matplotlib (Original work published 2011)

Mbow, C., Rosenzweig, C., Barioni, L., Benton, T., Herrero, M., Krishnapillai, M., Liwenga, E., Pradhan, P., Rivera-Ferre, M., Sapkota, T., & others. (2019). *Chapter 5 Food Security in Climate Change and Land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. IPCC (IPCC). https://www. ipcc. ch/srccl.

Meyerding, S., Schaffmann, A.-L., & Lehberger, M. (2019). Consumer Preferences for Different Designs of Carbon Footprint Labelling on Tomatoes in Germany—Does Design Matter? *Sustainability*, *11*(6), 1587. https://doi.org/10.3390/su11061587

Mundaca, L., Ürge-Vorsatz, D., & Wilson, C. (2019). Demand-side approaches for limiting global warming to 1.5 °C. *Energy Efficiency*, *12*(2), 343–362. https://doi.org/10.1007/s12053-018-9722-9

Numpy/numpy. (2023). [Python]. NumPy. https://github.com/numpy/numpy (Original work published 2010)

Osman, M., & Thornton, K. (2019). Traffic light labelling of meals to promote sustainable consumption and healthy eating. *Appetite*, *138*, 60–71. https://doi.org/10.1016/j.appet.2019.03.015

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science*, *360*(6392), 987–992. https://doi.org/10.1126/science.aaq0216

Riahi, K., Kriegler, E., Johnson, N., Bertram, C., den Elzen, M., Eom, J., Schaeffer, M., Edmonds, J., Isaac, M., Krey, V., Longden, T., Luderer, G., Méjean, A., McCollum, D. L., Mima, S., Turton, H., van Vuuren, D. P., Wada, K., Bosetti, V., … Edenhofer, O. (2015). Locked into Copenhagen pledges—Implications of short-term emission targets for the cost and feasibility of long-term climate goals. *Technological Forecasting and Social Change*, *90*, 8–23. https://doi.org/10.1016/j.techfore.2013.09.016

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., … Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Robertson, G. P., & Vitousek, P. M. (2009). Nitrogen in Agriculture: Balancing the Cost of an Essential Resource. *Annual Review of Environment and Resources*, *34*(1), 97–125. https://doi.org/10.1146/annurev.environ.032108.105046

Roe, S., Streck, C., Obersteiner, M., Frank, S., Griscom, B., Drouet, L., Fricko, O., Gusti, M., Harris, N., Hasegawa, T., Hausfather, Z., Havlík, P., House, J., Nabuurs, G.-J., Popp, A., Sánchez, M. J. S., Sanderman, J., Smith, P., Stehfest, E., & Lawrence, D. (2019). Contribution of the land sector to a 1.5 °C world. *Nature Climate Change*, *9*(11), Article 11. https://doi.org/10.1038/s41558-019-0591-9

Rogelj, J., Forster, P. M., Kriegler, E., Smith, C. J., & Séférian, R. (2019). Estimating and tracking the remaining carbon budget for stringent climate targets. *Nature*, *571*(7765), Article 7765. https://doi.org/10.1038/s41586-019-1368-z

Röös, E., & Tjärnemo, H. (2011). Challenges of carbon labelling of food products: A consumer research perspective. *British Food Journal*, *113*(8), 982–996. https://doi.org/10.1108/00070701111153742

Rosenzweig, C., Mbow, C., Barioni, L. G., Benton, T. G., Herrero, M., Krishnapillai, M., Liwenga, E. T., Pradhan, P., Rivera-Ferre, M. G., Sapkota, T., Tubiello, F. N., Xu, Y., Mencos Contreras, E., & Portugal-Pereira, J. (2020). Climate change responses benefit from a global food system approach. *Nature Food*, *1*(2), 94–97. https://doi.org/10.1038/s43016-020-0031-z

Rugani, B., Vázquez-Rowe, I., Benedetto, G., & Benetto, E. (2013). A comprehensive review of carbon footprint analysis as an extended environmental indicator in the wine sector. *Journal of Cleaner Production*, *54*, 61–77. https://doi.org/10.1016/j.jclepro.2013.04.036

Schaefer, F., & Blanke, M. (2014). Opportunities and Challenges of Carbon Footprint, Climate or CO2 Labelling for Horticultural Products. *Erwerbs-Obstbau*, *56*(2), 73–80. https://doi.org/10.1007/s10341-014-0206-6

Scipy/scipy. (2023). [Python]. SciPy. https://github.com/scipy/scipy (Original work published 2011)

Shi, X. (2010). Carbon Footprint Labeling Activities in the East Asia Summit Region: Spillover Effects to Less Developed Countries. *Working Papers*, Article DP-2010-06. https://ideas.repec.org//p/era/wpaper/dp-2010-06.html

Shiklomanov, I. A., & Rodda, J. C. (Eds.). (2003). *World water resources at the beginning of the twenty-first century*. Cambridge University Press.

Slapø, H. B., & Karevold, K. I. (2019). 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*, *3*. https://www.frontiersin.org/articles/10.3389/fsufs.2019.00040

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., de Vries, W., Vermeulen, S. J., Herrero, M., Carlson, K. M., Jonell, M., Troell, M., DeClerck, F., Gordon, L. J., Zurayk, R., Scarborough, P., Rayner, M., Loken, B., Fanzo, J., … Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature*, *562*(7728), 519–525. https://doi.org/10.1038/s41586-018-0594-0

Steinfeld, H., Gerber, P., Wassenaar, T. D., Castel, V., Rosales M., M., & Haan, C. de. (2006). *Livestock's long shadow: Environmental issues and options*. Food and Agriculture Organization of the United Nations.

Stern, P. C. (1999). Information, Incentives, and Proenvironmental Consumer Behavior. *Journal of Consumer Policy*, *22*(4), 461–478. https://doi.org/10.1023/A:1006211709570

Taufique, K. M. R., Nielsen, K. S., Dietz, T., Shwom, R., Stern, P. C., & Vandenbergh, M. P. (2022). Revisiting the promise of carbon labelling. *Nature Climate Change*, *12*(2), 132–140. https://doi.org/10.1038/s41558-021-01271-8

Taufique, K. M. R., Vocino, A., & Polonsky, M. J. (2017). The influence of eco-label knowledge and trust on pro-environmental consumer behaviour in an emerging market. *Journal of Strategic Marketing*, *25*(7), 511–529. https://doi.org/10.1080/0965254X.2016.1240219

The pandas development team. (2023). *pandas-dev/pandas: Pandas* [Python]. https://github.com/pandas-dev/pandas (Original work published 2010)

Thøgersen, J., & Nielsen, K. S. (2016). A better carbon footprint label. *Journal of Cleaner Production*, *125*, 86–94. https://doi.org/10.1016/j.jclepro.2016.03.098

Tubiello, F. N., Rosenzweig, C., Conchedda, G., Karl, K., Gütschow, J., Xueyao, P., Obli-Laryea, G., Wanner, N., Qiu, S. Y., Barros, J. D., Flammini, A., Mencos-Contreras, E., Souza, L., Quadrelli, R., Heiðarsdóttir, H. H., Benoit, P., Hayek, M., & Sandalow, D. (2021). Greenhouse gas emissions from food systems: Building the evidence base. *Environmental Research Letters*, *16*(6), 065007. https://doi.org/10.1088/1748-9326/ac018e

United Nations Environment Programme. (2022). *Emissions Gap Report 2022: The Closing Window—Climate crisis calls for rapid transformation of societies.* (p. 65). https://www.unep.org/emissions-gap-report-2022

Upham, P., Dendler, L., & Bleda, M. (2011). Carbon labelling of grocery products: Public perceptions and potential emissions reductions. *Journal of Cleaner Production*, *19*(4), 348–355. https://doi.org/10.1016/j.jclepro.2010.05.014

Urai, A., & Kruk, K. (2023). *Net-zero scenarios.* Zenodo. https://doi.org/10.5281/zenodo.7767084

Vallat, R. (2023). *Raphaelvallat/pingouin* [Python]. https://github.com/raphaelvallat/pingouin (Original work published 2018) van der Ven, H. (2019). *Beyond Greenwash: Explaining Credibility in Transnational Eco-LabelingExplaining Credibility in Transnational Eco-Labeling*. https://doi.org/10.1093/oso/9780190866006.001.0001

Van Loo, E. J., Caputo, V., Nayga, R. M., & Verbeke, W. (2014). Consumers' valuation of sustainability labels on meat. *Food Policy*, *49*, 137–150. https://doi.org/10.1016/j.foodpol.2014.07.002

Vandenbergh, M. P., Dietz, T., & Stern, P. C. (2011). Time to try carbon labelling. *Nature Climate Change*, *1*(1), 4–6. https://doi.org/10.1038/nclimate1071

Vaughan, A. (2012, January 30). Tesco drops carbon-label pledge. *The Guardian*. https://www.theguardian.com/environment/2012/jan/30/tesco-drops-carbon-labelling

Wada, Y., van Beek, L. P. H., van Kempen, C. M., Reckman, J. W. T. M., Vasak, S., & Bierkens, M. F. P. (2010). Global depletion of groundwater resources: GLOBAL GROUNDWATER DEPLETION. *Geophysical Research Letters*, *37*(20), n/a-n/a. https://doi.org/10.1029/2010GL044571

Waskom, M. (2023). *seaborn: Statistical data visualization* [Python]. https://github.com/mwaskom/seaborn (Original work published 2012)

Weinrich, R., & Spiller, A. (2016). Developing food labelling strategies: Multi-level labelling. *Journal of Cleaner Production*, *137*, 1138–1148. https://doi.org/10.1016/j.jclepro.2016.07.156

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L. J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J. A., De Vries, W., Majele Sibanda, L., … Murray, C. J. L. (2019). Food in the Anthropocene: The EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet*, *393*(10170), 447–492. https://doi.org/10.1016/S0140-6736(18)31788-4

Zettle. (2023). *Zettle NL (was iZettle) | POS, pinapparaten en betaaloplossingen*. https://www.zettle.com/nl

Appendix A

Figure 26

Histograms and Q-Q plots for every data sample pre-label implementation.

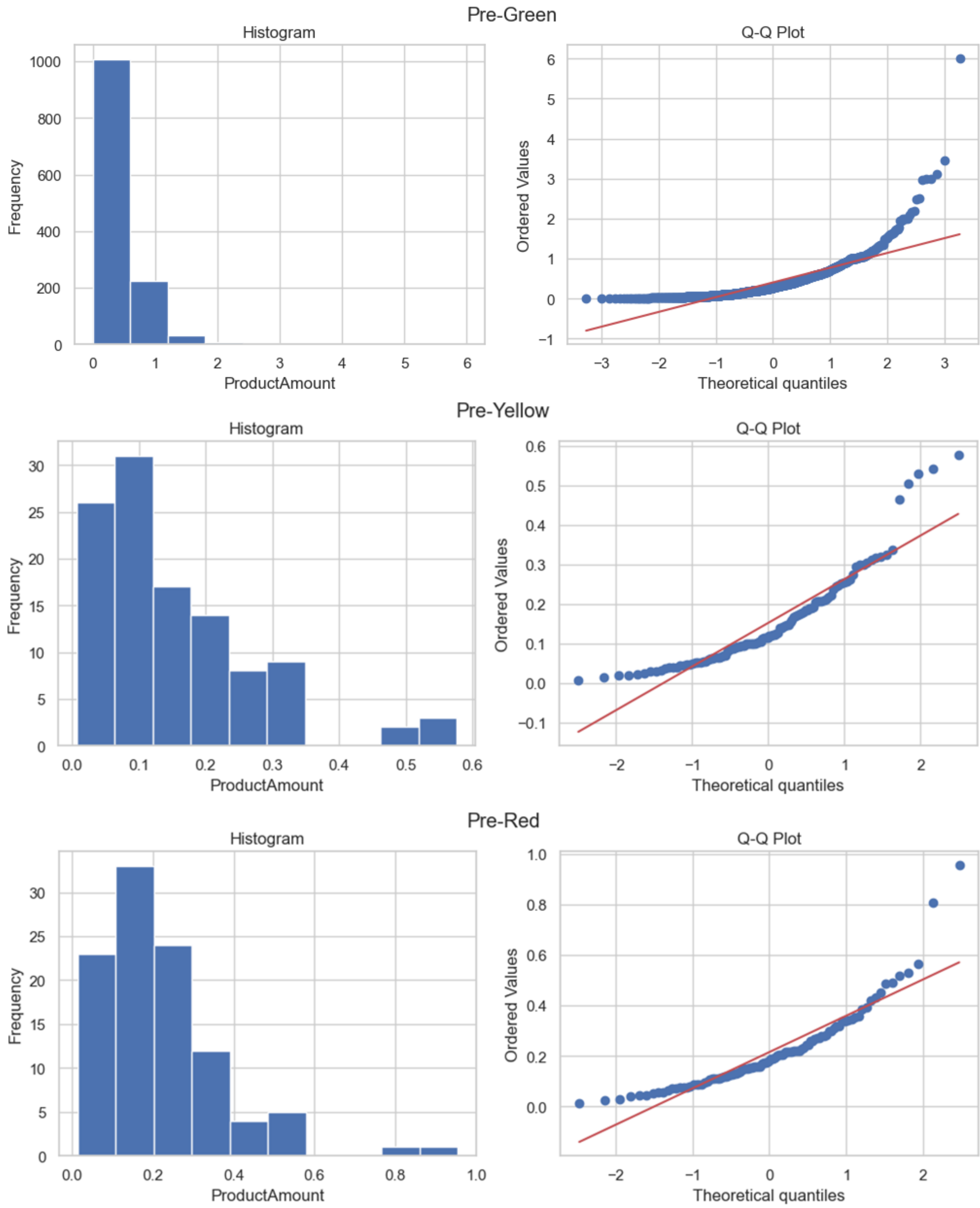
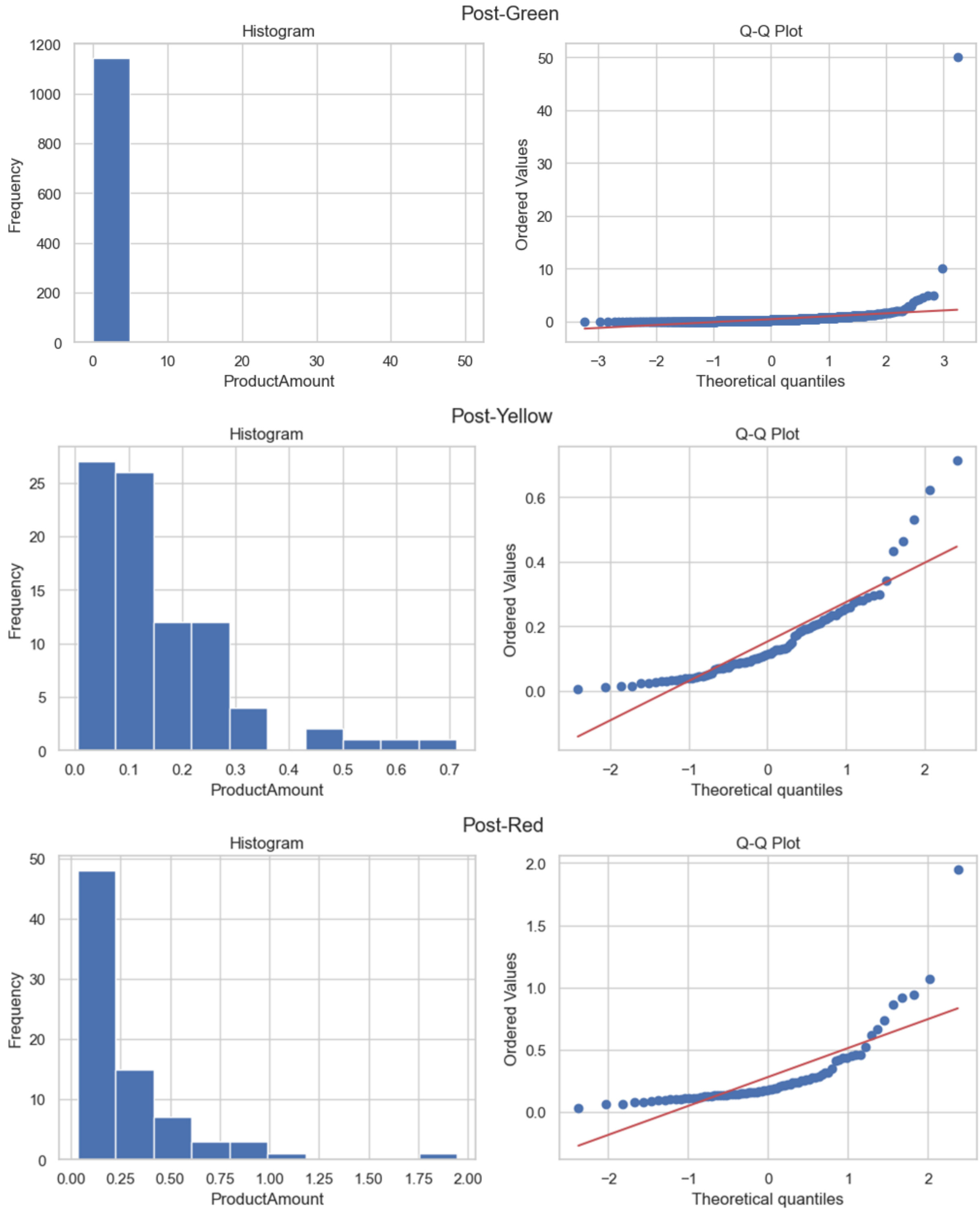


Figure 27

Histograms and Q-Q plots for every data sample post-label implementation.



Testing for Normality With Plots.

Figure 28

Histogram and QQ-plot Transaction Change Score

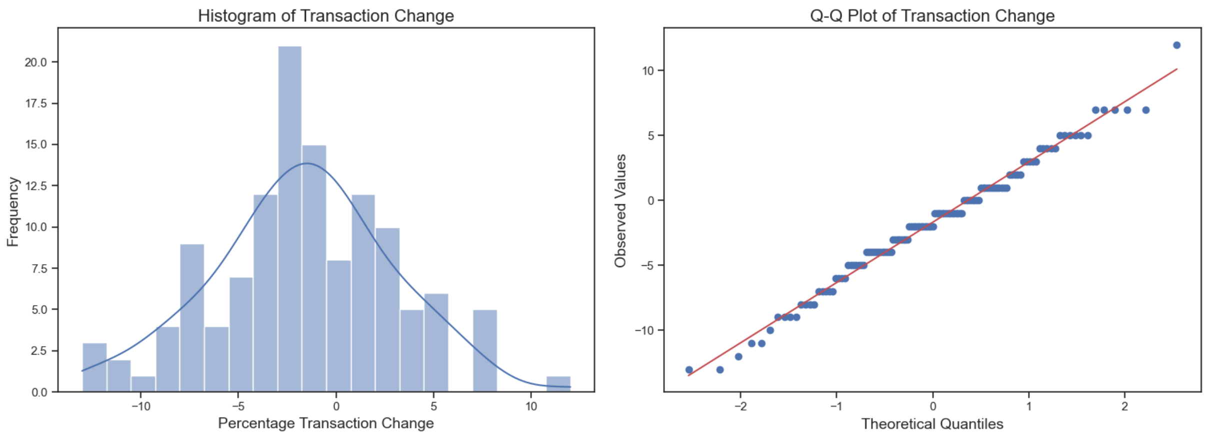
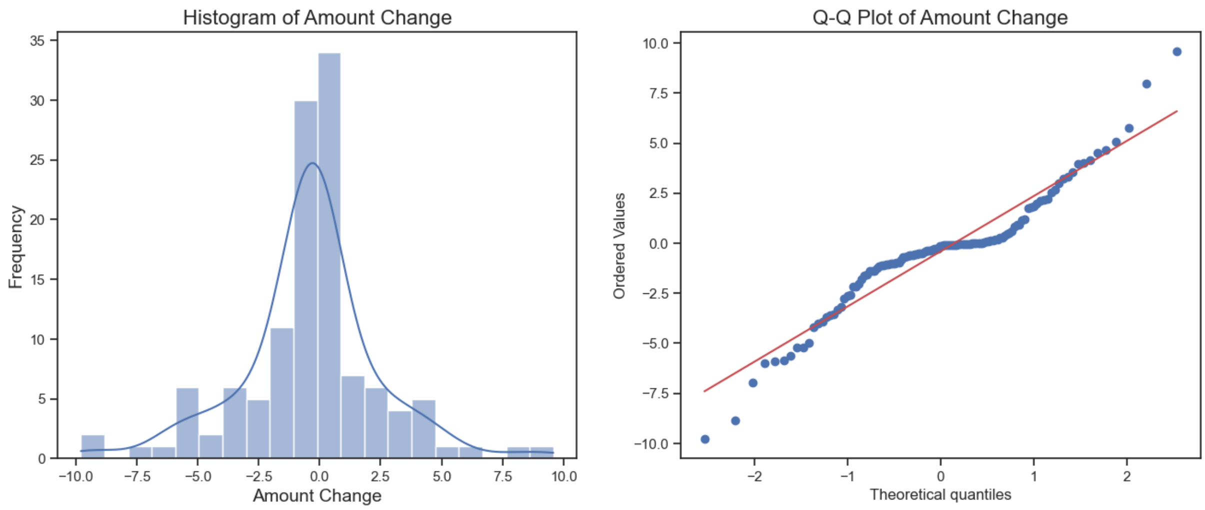


Figure 29

Histogram and QQ-plot Product Amount Change Score



Testing Normality With the Shapiro-Wilk Test.

In addition to our graphical assessment of normality we wanted to provide a statistical normality assessment to confirm or disconfirm our previous findings.

For our data on the transaction change score a Shapiro-Wilk test was conducted to assess the normality of the distribution. It was found that the data did not significantly deviate from a normal distribution, $W = 0.989$, $p = .445$. For that reason, we failed to reject the null hypothesis and can conclude that the sample approximately followed a normal distribution.

An additional Shapiro-Wilk test was conducted for our data on the product amount change score. We found that the data significantly deviated from a normal distribution, $W = 0.929$, $p < .001$. As such, we rejected the null hypothesis and conclude that the sample did not approximately follow a normal distribution.

Table 3

Shapiro-Wilk Test p-values

Group	Green	Yellow	Red
Pre	0.0000	0.0000	0.0000
Post	0.0000	0.0000	0.0000

The Shapiro-Wilk test was employed to evaluate the normality of the data within each group of product sales, divided by label colour and implementation period (pre and post). The resulting p-values for all six groups were found to be 0.0000, indicating that the null hypothesis of normality can be rejected for every group. Consequently, it can be inferred that the distribution of product sales in each group does not follow a normal distribution. This deviation from normality was taken into consideration when selecting appropriate statistical tests for further analysis, as it may impact the validity of the assumptions made by parametric tests.

Testing for Normality With Plots (Filtered).

Figure 30

Histogram and QQ-plot Transaction Change Score (Filtered)

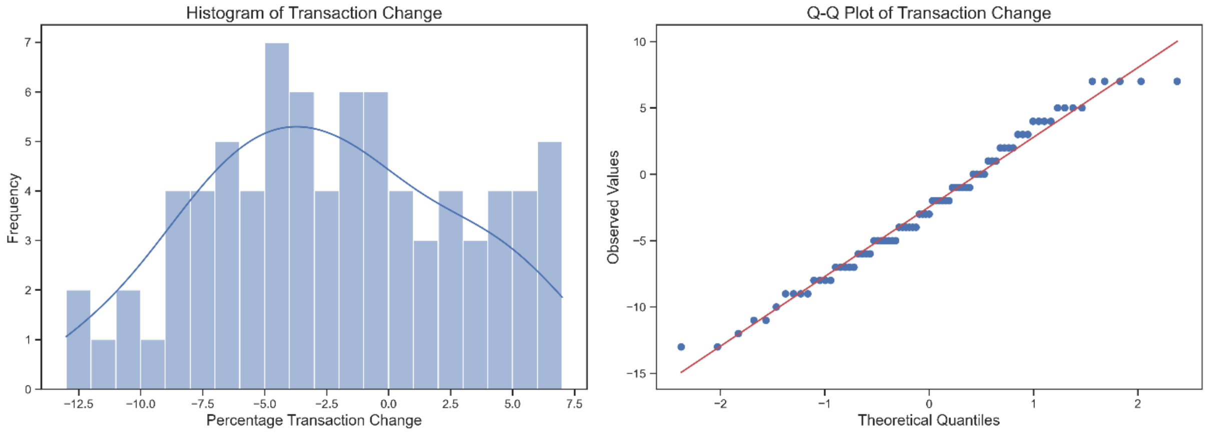
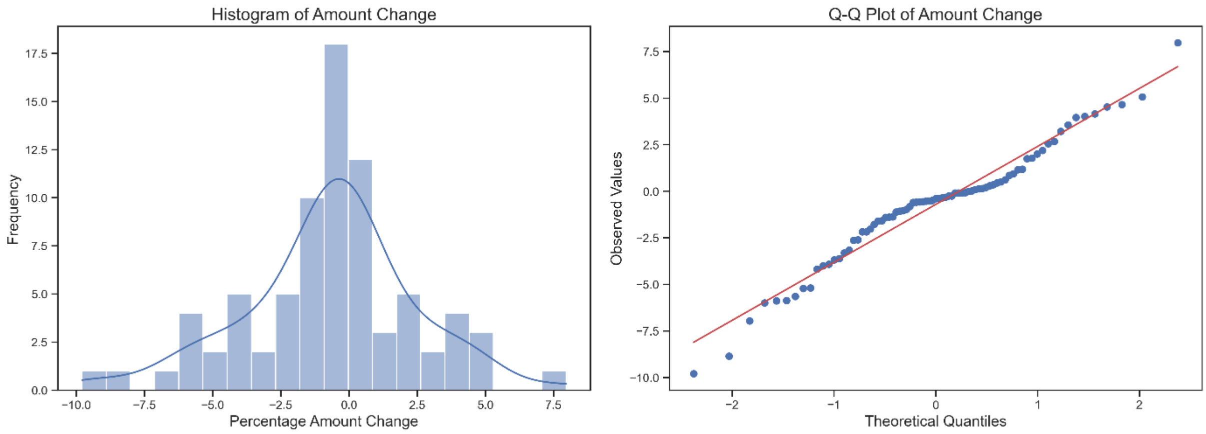


Figure 31

Histogram and QQ-plot Product Amount Change Score (Filtered)



Testing Normality With the Shapiro-Wilk Test.

In addition to our graphical assessment of normality we wanted to provide a statistical normality assessment to confirm or disconfirm our previous findings.

For our data on the transaction change score a Shapiro-Wilk test was conducted to assess the normality of the distribution. It was found that the data did not significantly deviate from a normal distribution, $W = 0.977$, $p = .159$. For that reason, we failed to reject the null hypothesis and can conclude that the sample approximately followed a normal distribution.

An additional Shapiro-Wilk test was conducted for our data on the product amount change score. We found that the data significantly deviated from a normal distribution, $W = 0.968$, $p = .0463$. As such, we rejected the null hypothesis and conclude that the sample did not approximately follow a normal distribution.